

This document is meant purely as a documentation tool and the institutions do not assume any liability for its contents

► **B**

► **C1** COMMISSION REGULATION (EC) No 440/2008

of 30 May 2008

laying down test methods pursuant to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH)

(Text with EEA relevance) ◀

(OJ L 142, 31.5.2008, p. 1)

Amended by:

		Official Journal		
		No	page	date
► <b>M1</b>	Commission Regulation (EC) No 761/2009 of 23 July 2009	L 220	1	24.8.2009
► <b>M2</b>	Commission Regulation (EU) No 1152/2010 of 8 December 2010	L 324	13	9.12.2010
► <b>M3</b>	Commission Regulation (EU) No 640/2012 of 6 July 2012	L 193	1	20.7.2012
► <b>M4</b>	Commission Regulation (EU) No 260/2014 of 24 January 2014	L 81	1	19.3.2014
► <b>M5</b>	Commission Regulation (EU) No 900/2014 of 15 July 2014	L 247	1	21.8.2014

Corrected by:

► **C1** Corrigendum, OJ L 143, 3.6.2008, p. 55 (440/2008)

▼B

▼C1

**COMMISSION REGULATION (EC) No 440/2008  
of 30 May 2008**

**laying down test methods pursuant to Regulation (EC)  
No 1907/2006 of the European Parliament and of the Council on  
the Registration, Evaluation, Authorisation and Restriction of  
Chemicals (REACH)**

(Text with EEA relevance)

▼B

THE COMMISSION OF THE EUROPEAN COMMUNITIES,

Having regard to the Treaty establishing the European Community,

Having regard to Regulation (EC) No 1907/2006 of 18 December 2006 of the European Parliament and of the Council concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC <sup>(1)</sup>, and in particular Article 13(3) thereof,

Whereas:

- (1) Pursuant to Regulation (EC) No 1907/2006, test methods are to be adopted at Community level for the purposes of tests on substances where such tests are required to generate information on intrinsic properties of substances.
- (2) Council Directive 67/548/EEC of 27 June 1967 on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances <sup>(2)</sup>, laid down, in Annex V, methods for the determination of the physico-chemical properties, toxicity and ecotoxicity of substances and preparations. Annex V to Directive 67/548/EEC has been deleted by Directive 2006/121/EC of the European Parliament and of the Council with effect from 1 June 2008.
- (3) The test methods contained in Annex V to Directive 67/548/EEC should be incorporated into this Regulation.
- (4) This Regulation does not exclude the use of other test methods, provided that their use is in accordance with Article 13(3) of Regulation 1907/2006.

<sup>(1)</sup> OJ L 396, 30.12.2006, p. 1, as corrected by OJ L 136, 29.5.2007, p. 3.

<sup>(2)</sup> OJ 196, 16.8.1967, p. 1. Directive as last amended by Directive 2006/121/CE of the European Parliament and of the Council (OJ L 396, 30.12.2006, p. 850, as corrected by OJ L 136, 29.5.2007, p. 281).

**▼B**

- (5) The principles of replacement, reduction and refinement of the use of animals in procedures should be fully taken into account in the design of the test methods, in particular when appropriate validated methods become available to replace, reduce or refine animal testing.
- (6) The provisions of this Regulation are in accordance with the opinion of the Committee established under Article 133 of Regulation (EC) No 1907/2006,

HAS ADOPTED THIS REGULATION:

*Article 1*

The test methods to be applied for the purposes of Regulation 1907/2006/EC are set out in the Annex to this Regulation.

*Article 2*

The Commission shall review, where appropriate, the test methods contained in this Regulation with a view to replacing, reducing or refining testing on vertebrate animals.

*Article 3*

All references to Annex V to Directive 67/548/EEC shall be construed as references to this Regulation.

*Article 4*

This Regulation shall enter into force on the day following its publication in the *Official Journal of the European Union*.

It shall apply from 1 June 2008.

**▼B***ANNEX***PART A: METHODS FOR THE DETERMINATION OF PHYSICO-CHEMICAL PROPERTIES**

## TABLE OF CONTENTS

- A.1. MELTING/FREEZING TEMPERATURE
  - A.2. BOILING TEMPERATURE
  - A.3. RELATIVE DENSITY
  - A.4. VAPOUR PRESSURE
  - A.5. SURFACE TENSION
  - A.6. WATER SOLUBILITY
  - A.8. PARTITION COEFFICIENT
  - A.9. FLASH-POINT
  - A.10. FLAMMABILITY (SOLIDS)
  - A.11. FLAMMABILITY (GASES)
  - A.12. FLAMMABILITY (CONTACT WITH WATER)
  - A.13. PYROPHORIC PROPERTIES OF SOLIDS AND LIQUIDS
  - A.14. EXPLOSIVE PROPERTIES
  - A.15. AUTO-IGNITION TEMPERATURE (LIQUIDS AND GASES)
  - A.16. RELATIVE SELF-IGNITION TEMPERATURE FOR SOLIDS
  - A.17. OXIDISING PROPERTIES (SOLIDS)
  - A.18. NUMBER — AVERAGE MOLECULAR WEIGHT AND MOLECULAR WEIGHT DISTRIBUTION OF POLYMERS
  - A.19. LOW MOLECULAR WEIGHT CONTENT OF POLYMERS
  - A.20. SOLUTION/EXTRACTION BEHAVIOUR OF POLYMERS IN WATER
  - A.21. OXIDISING PROPERTIES (LIQUIDS)
  - A.22. LENGTH WEIGHTED GEOMETRIC MEAN DIAMETER OF FIBRES
- ▼M4**
- A.23. PARTITION COEFFICIENT (1-OCTANOL/WATER): SLOW-STIRRING METHOD



**▼B****A.1. MELTING/FREEZING TEMPERATURE****1. METHOD**

The majority of the methods described are based on the OECD Test Guideline (1). The fundamental principles are given in references (2) and (3).

**1.1. INTRODUCTION**

The methods and devices described are to be applied for the determination of the melting temperature of substances, without any restriction with respect to their degree of purity.

The selection of the method is dependent on the nature of the substance to be tested. In consequence the limiting factor will be according to, whether or not the substance can be pulverised easily, with difficulty, or not at all.

For some substances, the determination of the freezing or solidification temperature is more appropriate and the standards for these determinations have also been included in this method.

Where, due to the particular properties of the substance, none of the above parameters can be conveniently measured, a pour point may be appropriate.

**1.2. DEFINITIONS AND UNITS**

The melting temperature is defined as the temperature at which the phase transition from solid to liquid state occurs at atmospheric pressure and this temperature ideally corresponds to the freezing temperature.

As the phase transition of many substances takes place over a temperature range, it is often described as the melting range.

Conversion of units (K to °C)

$$t = T - 273,15$$

t: Celsius temperature, degree Celsius (°C)

T: thermodynamic temperature, kelvin (K)

**1.3. REFERENCE SUBSTANCES**

Reference substances do not need to be employed in all cases when investigating a new substance. They should primarily serve to check the performance of the method from time to time and to allow comparison with results from other methods.

Some calibration substances are listed in the references (4).

**▼ B**

## 1.4. PRINCIPLE OF THE TEST METHOD

The temperature (temperature range) of the phase transition from the solid to the liquid state or from the liquid to the solid state is determined. In practice while heating/cooling a sample of the test substance at atmospheric pressure the temperatures of the initial melting/freezing and the final stage of melting/freezing are determined. Five types of methods are described, namely capillary method, hot stage methods, freezing temperature determinations, methods of thermal analysis, and determination of the pour point (as developed for petroleum oils).

In certain cases, it may be convenient to measure the freezing temperature in place of the melting temperature.

1.4.1. **Capillary method**1.4.1.1. *Melting temperature devices with liquid bath*

A small amount of the finely ground substance is placed in a capillary tube and packed tightly. The tube is heated, together with a thermometer, and the temperature rise is adjusted to less than about 1 K/min during the actual melting. The initial and final melting temperatures are determined.

1.4.1.2. *Melting temperature devices with metal block*

As described under 1.4.1.1, except that the capillary tube and the thermometer are situated in a heated metal block, and can be observed through holes in the block.

1.4.1.3. *Photocell detection*

The sample in the capillary tube is heated automatically in a metal cylinder. A beam of light is directed through the substance, by way of a hole in the cylinder, to a precisely calibrated photocell. The optical properties of most substances change from opaque to transparent when they are melting. The intensity of light reaching the photocell increases and sends a stop signal to the digital indicator reading out the temperature of a platinum resistance thermometer located in the heating chamber. This method is not suitable for some highly coloured substances.

1.4.2. **Hot stages**1.4.2.1. *Kofler hot bar*

The Kofler hot bar uses two pieces of metal of different thermal conductivity, heated electrically, with the bar designed so that the temperature gradient is almost linear along its length. The temperature of the hot bar can range from 283 to 573 K with a special temperature-reading device including a runner with a pointer and tab designed for the specific bar. In order to determine a melting temperature, the substance is laid, in a thin layer, directly on the surface of the hot bar. In a few seconds a sharp dividing line between the fluid and solid phase develops. The temperature at the dividing line is read by adjusting the pointer to rest at the line.

**▼ B**1.4.2.2. *Melt microscope*

Several microscope hot stages are in use for the determination of melting temperatures with very small quantities of material. In most of the hot stages the temperature is measured with a sensitive thermocouple but sometimes mercury thermometers are used. A typical microscope hot stage melting temperature apparatus has a heating chamber which contains a metal plate upon which the sample is placed on a slide. The centre of the metal plate contains a hole permitting the entrance of light from the illuminating mirror of the microscope. When in use, the chamber is closed by a glass plate to exclude air from the sample area.

The heating of the sample is regulated by a rheostat. For very precise measurements on optically anisotropic substances, polarised light may be used.

1.4.2.3. *Meniscus method*

This method is specifically used for polyamides.

The temperature at which the displacement of a meniscus of silicone oil, enclosed between a hot stage and a cover-glass supported by the polyamide test specimen, is determined visually.

1.4.3. **Method to determine the freezing temperature**

The sample is placed in a special test tube and placed in an apparatus for the determination of the freezing temperature. The sample is stirred gently and continuously during cooling and the temperature is measured at suitable intervals. As soon as the temperature remains constant for a few readings this temperature (corrected for thermometer error) is recorded as the freezing temperature.

Supercooling must be avoided by maintaining equilibrium between the solid and the liquid phases.

1.4.4. **Thermal analysis**1.4.4.1 *Differential thermal analysis (DTA)*

This technique records the difference in temperatures between the substance and a reference material as a function of temperature, while the substance and reference material are subjected to the same controlled temperature programme. When the sample undergoes a transition involving a change of enthalpy, that change is indicated by an endothermic (melting) or exothermic (freezing) departure from the base line of the temperature record.

**▼ B**1.4.4.2 *Differential scanning calorimetry (DSC)*

This technique records the difference in energy inputs into a substance and a reference material, as a function of temperature, while the substance and reference material are subjected to the same controlled temperature programme. This energy is the energy necessary to establish zero temperature difference between the substance and the reference material. When the sample undergoes a transition involving a change of enthalpy, that change is indicated by an endothermic (melting) or exothermic (freezing) departure from the base line of the heat flow record.

1.4.5. **Pour point**

This method was developed for use with petroleum oils and is suitable for use with oily substances with low melting temperatures.

After preliminary heating, the sample is cooled at a specific rate and examined at intervals of 3 K for flow characteristics. The lowest temperature at which movement of the substance is observed is recorded as the pour point.

## 1.5. QUALITY CRITERIA

The applicability and accuracy of the different methods used for the determination of the melting temperature/melting range are listed in the following table:

TABLE: APPLICABILITY OF THE METHODS

A. **Capillary methods**

Method of measurement	Substances which can be pulverised	Substances which are not readily pulverised	Temperature range	Estimated accuracy <sup>(1)</sup>	Existing standards
Melting temperature devices with liquid bath	yes	only to a few	273 to 573 K	± 0,3 K	JIS K 0064
Melting temperature with metal block	yes	only to a few	293 to > 573 K	± 0,5 K	ISO 1218 (E)
Photocell detection	yes	several with appliance devices	253 to 573 K	± 0,5 K	

<sup>(1)</sup> Dependent on type of instrument and on degree of purity of the substance.

▼**B****B. Hot stages and freezing methods**

Method of measurement	Substances which can be pulverised	Substances which are not readily pulverised	Temperature range	Estimated accuracy <sup>(1)</sup>	Existing standards
Kofler hot bar	yes	no	283 to > 573 K	± 1K	ANSI/ASTM D 3451-76
Melt microscope	yes	only to a few	273 to > 573 K	± 0,5 K	DIN 53736
Meniscus method	no	specifically for polyamides	293 to > 573 K	± 0,5 K	ISO 1218 (E)
Freezing temperature	yes	yes	223 to 573 K	± 0,5 K	e.g. BS 4695

<sup>(1)</sup> Dependent on type of instrument and on degree of purity of the substance

**C. Thermal analysis**

Method of measurement	Substances which can be pulverised	Substances which are not readily pulverised	Temperature range	Estimated accuracy <sup>(1)</sup>	Existing standards
Differential thermal analysis	yes	yes	173 to 1 273 K	up to 600 K ± 0,5 K up to 1 273 K ± 2,0 K	ASTM E 537-76
Differential scanning calorimetry	yes	yes	173 to 1 273 K	up to 600 K ± 0,5 K up to 1 273 K ± 2,0 K	ASTM E 537-76

<sup>(1)</sup> Dependent on type of instrument and on degree of purity of the substance

**D. Pour point**

Method of measurement	Substances which can be pulverised	Substances which are not readily pulverised	Temperature range	Estimated accuracy <sup>(1)</sup>	Existing standards
Pour point	for petroleum oils and oily substances	for petroleum oils and oily substances	223 to 323 K	± 0,3 K	ASTM D 97-66

<sup>(1)</sup> Dependent on type of instrument and on degree of purity of the substance

## 1.6. DESCRIPTION OF THE METHODS

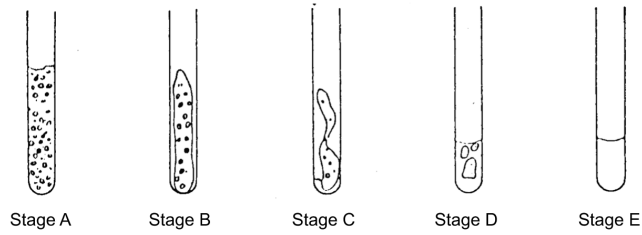
The procedures of nearly all the test methods have been described in international and national standards (see Appendix 1).

1.6.1. **Methods with capillary tube**

When subjected to a slow temperature rise, finely pulverised substances usually show the stages of melting shown in figure 1.

▼ B

Figure 1



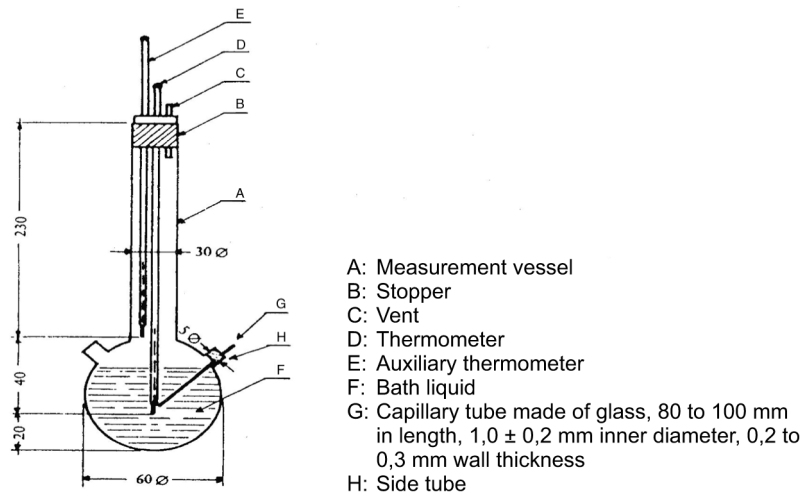
- Stage A (beginning of melting): fine droplets adhere uniformly to the inside wall of the capillary tube
- Stage B a clearance appears between the sample and the inside wall due to shrinkage of the melt
- Stage C the shrunken sample begins to collapse downwards and liquefies
- Stage D a complete meniscus is formed at the surface but an appreciable amount of the sample remains solid
- Stage E (final stage melting): there are no solid particles

During the determination of the melting temperature, the temperatures are recorded at the beginning of the melting and at the final stage.

#### 1.6.1.1. Melting temperature devices with liquid bath apparatus

Figure 2 shows a type of standardised melting temperature apparatus made of glass (JIS K 0064); all specifications are in millimeters.

Figure 2



**▼ B***Bath liquid:*

A suitable liquid should be chosen. The choice of the liquid depends upon the melting temperature to be determined, e.g. liquid paraffin for melting temperatures no higher than 473 K, silicone oil for melting temperatures no higher than 573 K.

For melting temperatures above 523 K, a mixture consisting of three parts sulphuric acid and two parts potassium sulphate (in mass ratio) can be used. Suitable precautions should be taken if a mixture such as this is used.

*Thermometer:*

Only those thermometers should be used which fulfil the requirements of the following or equivalent standards:

ASTM E 1-71, DIN 12770, JIS K 8001.

*Procedure:*

The dry substance is finely pulverised in a mortar and is put into the capillary tube, fused at one end, so that the filling level is approximately 3 mm after being tightly packed. To obtain a uniform packed sample, the capillary tube should be dropped from a height of approximately 700 mm through a glass tube vertically onto a watch glass.

The filled capillary tube is placed in the bath so that the middle part of the mercury bulb of the thermometer touches the capillary tube at the part where the sample is located. Usually the capillary tube is introduced into the apparatus about 10 K below the melting temperature.

The bath liquid is heated so that the temperature rise is approximately 3 K/min. The liquid should be stirred. At about 10 K below the expected melting temperature the rate of temperature rise is adjusted to a maximum of 1 K/min.

*Calculation:*

The calculation of the melting temperature is as follows:

$$T = T_D + 0,00016 (T_D - T_E) n$$

where:

T = corrected melting temperature in K

T<sub>D</sub> = temperature reading of thermometer D in K

T<sub>E</sub> = temperature reading of thermometer E in K

n = number of graduations of mercury thread on thermometer D at emergent stem.

1.6.1.2. *Melting temperature devices with metal block**Apparatus:*

*This consists of:*

- a cylindrical metal block, the upper part of which is hollow and forms a chamber (see figure 3),
- a metal plug, with two or more holes, allowing tubes to be mounted into the metal block,

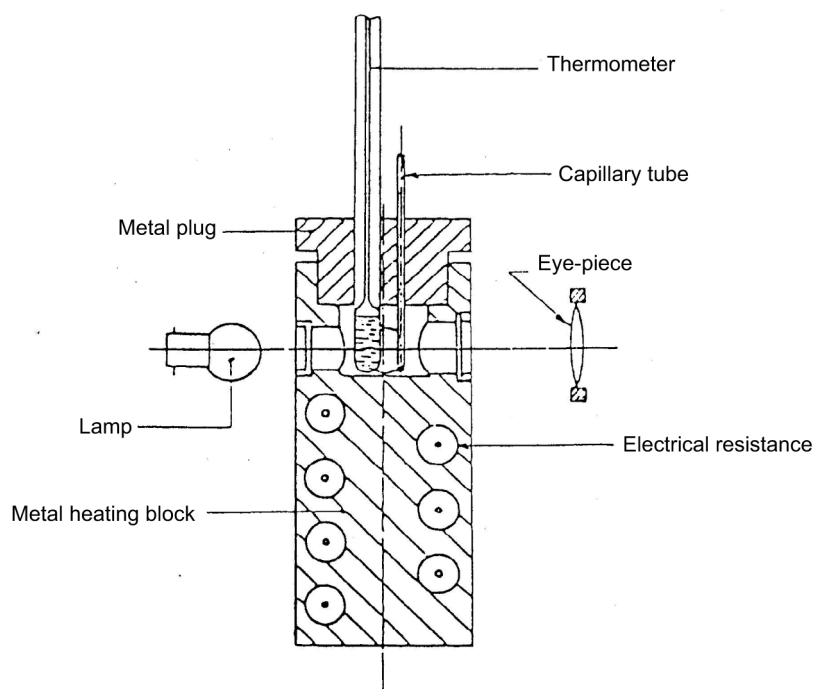
**▼ B**

- a heating system, for the metal block, provided for example by an electrical resistance enclosed in the block,
- a rheostat for regulation of power input, if electric heating is used,
- four windows of heat-resistant glass on the lateral walls of the chamber, diametrically disposed at right-angles to each other. In front of one of these windows is mounted an eye-piece for observing the capillary tube. The other three windows are used for illuminating the inside of the enclosure by means of lamps,
- a capillary tube of heat-resistant glass closed at one end (see 1.6.1.1).

*Thermometer:*

See standards mentioned in 1.6.1.1. Thermoelectrical measuring devices with comparable accuracy are also applicable.

Figure 3

1.6.1.3. *Photocell detection**Apparatus and procedure:*

The apparatus consists of a metal chamber with automated heating system. Three capillary are filled accordingly to 1.6.1.1 and placed in the oven.



**▼ B**

Several linear increases of temperature are available for calibrating the apparatus and the suitable temperature rise is electrically adjusted at a pre-selected constant and linear rate. recorders show the actual oven temperature and the temperature of the substance in the capillary tubes.

**1.6.2. Hot stages****1.6.2.1. Kofler hot bar**

See Appendix.

**1.6.2.2. Melt microscope**

See Appendix.

**1.6.2.3. Meniscus method (polyamides)**

See Appendix.

The heating rate through the melting temperature should be less than 1 K/min.

**1.6.3. Methods for the determination of the freezing temperature**

See Appendix.

**1.6.4. Thermal analysis****1.6.4.1. Differential thermal analysis**

See Appendix.

**1.6.4.2. Differential scanning calorimetry**

See Appendix.

**1.6.5. Determination of the pour point**

See Appendix.

**2. DATA**

A thermometer correction is necessary in some cases.

**3. REPORTING**

The test report shall, if possible, include the following information:

- method used,
- precise specification of the substance (identity and impurities) and preliminary purification step, if any,
- an estimate of the accuracy.

The mean of at least two measurements which are in the range of the estimated accuracy (see tables) is reported as the melting temperature.

**▼B**

If the difference between the temperature at the beginning and at the final stage of melting is within the limits of the accuracy of the method, the temperature at the final stage of melting is taken as the melting temperature; otherwise the two temperatures are reported.

If the substance decomposes or sublimes before the melting temperature is reached, the temperature at which the effect is observed shall be reported.

All information and remarks relevant for the interpretation of results have to be reported, especially with regard to impurities and physical state of the substance.

**4. REFERENCES**

- (1) OECD, Paris, 1981, Test Guideline 102, Decision of the Council C(81) 30 final.
- (2) IUPAC, B. Le Neindre, B. Vodar, eds. Experimental thermodynamics, Butterworths, London 1975, vol. II, p. 803-834.
- (3) R. Weissberger ed.: Technique of organic Chemistry, Physical Methods of Organic Chemistry, 3rd ed., Interscience Publ., New York, 1959, vol. I, Part I, Chapter VII.
- (4) IUPAC, Physicochemical measurements: Catalogue of reference materials from national laboratories, Pure and applied chemistry, 1976, vol. 48, p. 505-515.

**▼ B***Appendix*

*For additional technical details, the following standards may be consulted for example.*

1. **Capillary methods**

1.1. Melting temperature devices with liquid bath

ASTM E 324-69	Standard test method for relative initial and final melting points and the melting range of organic chemicals
BS 4634	Method for the determination of melting point and/or melting range
DIN 53181	Bestimmung des Schmelzintervalles von Harzen nach Kapillarverfahren
JIS K 00-64	Testing methods for melting point of chemical products

1.2. Melting temperature devices with metal block

DIN 53736	Visuelle Bestimmung der Schmelztemperatur von teilkristallinen Kunststoffen
ISO 1218 (E)	Plastics — polyamides — determination of ‘melting point’

2. **Hot stages**

2.1. Kofler hot bar

ANSI/ASTM D 3451-76	Standard recommended practices for testing polymeric powder coatings
---------------------	--

2.2. Melt microscope

DIN 53736	Visuelle Bestimmung der Schmelztemperatur von teilkristallinen Kunststoffen
-----------	---

2.3. Meniscus method (polyamides)

ISO 1218 (E)	Plastics — polyamides — determination of ‘melting point’
--------------	--

**▼B**

ANSI/ASTM D 2133-66	Standard specification for acetal resin injection moulding and extrusion materials
NF T 51-050	Résines de polyamides. Détermination du 'point de fusion' méthode du ménisque

3. **Methods for the determination of the freezing temperature**

BS 4633	Method for the determination of crystallising point
BS 4695	Method for Determination of Melting Point of petroleum wax (Cooling Curve)
DIN 51421	Bestimmung des Gefrierpunktes von Flugkraftstoffen, Ottokraftstoffen und Motorenbenzolen
ISO 2207	Cires de pétrole: détermination de la température de figeage
DIN 53175	Bestimmung des Erstarrungspunktes von Fettsäuren
NF T 60-114	Point de fusion des paraffines
NF T 20-051	Méthode de détermination du point de cristallisation (point de congélation)
ISO 1392	Method for the determination of the freezing point

4. **Thermal analysis**

4.1. Differential thermal analysis

ASTM E 537-76	Standard method for assessing the thermal stability of chemicals by methods of differential thermal analysis
ASTM E 473-85	Standard definitions of terms relating to thermal analysis

**▼ B**

ASTM E 472-86	Standard practice for reporting thermo-analytical data
DIN 51005	Thermische Analyse, Begriffe

## 4.2. Differential scanning calorimetry

ASTM E 537-76	Standard method for assessing the thermal stability of chemicals by methods of differential thermal analysis
ASTM E 473-85	Standard definitions of terms relating to thermal analysis
ASTM E 472-86	Standard practice for reporting thermo-analytical data
DIN 51005	Thermische Analyse, Begriffe

5. **Determination of the pour point**

NBN 52014	Echantillonnage et analyse des produits du pétrole: Point de trouble et point d'écoulement limite — Monsterneming en ontleding van aardolieproducten: Troebelingspunt en vloeipunt
ASTM D 97-66	Standard test method for pour point of petroleum oils
ISO 3016	Petroleum oils — Determination of pour point

**▼B****A.2. BOILING TEMPERATURE****1. METHOD**

The majority of the methods described are based on the OECD Test Guideline (1). The fundamental principles are given in references (2) and (3).

**1.1. INTRODUCTION**

The methods and devices described here can be applied to liquid and low melting substances, provided that these do not undergo chemical reaction below the boiling temperature (for example: auto-oxidation, rearrangement, degradation, etc.). The methods can be applied to pure and to impure liquid substances.

Emphasis is put on the methods using photocell detection and thermal analysis, because these methods allow the determination of melting as well as boiling temperatures. Moreover, measurements can be performed automatically.

The 'dynamic method' has the advantage that it can also be applied to the determination of the vapour pressure and it is not necessary to correct the boiling temperature to the normal pressure (101,325 kPa) because the normal pressure can be adjusted during the measurement by a manostat.

*Remarks:*

The influence of impurities on the determination of the boiling temperature depends greatly upon the nature of the impurity. When there are volatile impurities in the sample, which could affect the results, the substance may be purified.

**1.2. DEFINITIONS AND UNITS**

The normal boiling temperature is defined as the temperature at which the vapour pressure of a liquid is 101,325 kPa.

If the boiling temperature is not measured at normal atmospheric pressure, the temperature dependence of the vapour pressure can be described by the Clausius-Clapeyron equation:

$$\log p = \frac{\Delta H_v}{2,3RT} + \text{const.}$$

where:

$p$  = the vapour pressure of the substance in pascals

$\Delta H_v$  = its heat of vaporisation in  $\text{J mol}^{-1}$

$R$  = the universal molar gas constant =  $8,314 \text{ J mol}^{-1} \text{ K}^{-1}$

$T$  = thermodynamic temperature in K

The boiling temperature is stated with regard to the ambient pressure during the measurement.

**▼B***Conversions*

Pressure (units: kPa)

100 kPa = 1 bar = 0,1 MPa  
(‘bar’ is still permissible but not recommended)

133 Pa = 1 mm Hg = 1 Torr  
(the units ‘mm Hg’ and ‘Torr’ are not permitted)

1 atm = standard atmosphere = 101 325 Pa  
(the unit ‘atm’ is not permitted)

Temperature (units: K)

$t = T - 273,15$

t: Celsius temperature, degree Celsius (°C)

T: thermodynamic temperature, kelvin (K)

### 1.3. REFERENCE SUBSTANCES

Reference substances do not need to be employed in all cases when investigating a new substance. They should primarily serve to check the performance of the method from time to time and to allow comparison with results from other methods.

Some calibration substances can be found in the methods listed in the Appendix.

### 1.4. PRINCIPLE OF THE TEST METHOD

Five methods for the determination of the boiling temperature (boiling range) are based on the measurement of the boiling temperature, two others are based on thermal analysis.

#### 1.4.1. **Determination by use of the ebulliometer**

Ebullimeters were originally developed for the determination of the molecular weight by boiling temperature elevation, but they are also suited for exact boiling temperature measurements. A very simple apparatus is described in ASTM D 1120-72 (see Appendix). The liquid is heated in this apparatus under equilibrium conditions at atmospheric pressure until it is boiling.

#### 1.4.2. **Dynamic method**

This method involves the measurement of the vapour recondensation temperature by means of an appropriate thermometer in the reflux while boiling. The pressure can be varied in this method.

#### 1.4.3. **Distillation method for boiling temperature**

This method involves distillation of the liquid and measurement of the vapour recondensation temperature and determination of the amount of distillate.

**▼B****1.4.4. Method according to Siwoloboff**

A sample is heated in a sample tube, which is immersed in a liquid in a heat-bath. A fused capillary, containing an air bubble in the lower part, is dipped in the sample tube.

**1.4.5. Photocell detection**

Following the principle according to Siwoloboff, automatic photo-electrical measurement is made using rising bubbles.

**1.4.6. Differential thermal analysis**

This technique records the difference in temperatures between the substance and a reference material as a function of temperature, while the substance and reference material are subjected to the same controlled temperature programme. When the sample undergoes a transition involving a change of enthalpy, that change is indicated by an endothermic departure (boiling) from the base line of the temperature record.

**1.4.7. Differential scanning calorimetry**

This technique records the difference in energy inputs into a substance and a reference material as a function of temperature, while the substance and reference material are subjected to the same controlled temperature programme. This energy is the energy necessary to establish zero temperature difference between the substance and the reference material. When the sample undergoes a transition involving a change of enthalpy, that change is indicated by an endothermic departure (boiling) from the base line of the heat flow record.

**1.5. QUALITY CRITERIA**

The applicability and accuracy of the different methods used for the determination of the boiling temperature/boiling range are listed in table 1.

*Table 1:*

**Comparison of the methods**

Method of measurement	Estimated accuracy	Existing standard
Ebulliometer	± 1,4 K (up to 373 K) <sup>(1)</sup> <sup>(2)</sup> ± 2,5 K (up to 600 K) <sup>(1)</sup> <sup>(2)</sup>	ASTM D 1120-72 <sup>(1)</sup>
Dynamic method	± 0,5 K (up to 600 K) <sup>(2)</sup>	
Distillation process (boiling range)	± 0,5 K (up to 600 K)	ISO/R 918, DIN 53171, BS 4591/71
According to Siwoloboff	± 2 K (up to 600 K) <sup>(2)</sup>	
Photocell detection	± 0,3 K (up to 373 K) <sup>(2)</sup>	
Differential thermal calorimetry	± 0,5 K (up to 600 K) ± 2,0 K (up to 1 273 K)	ASTM E 537-76
Differential scanning calorimetry	± 0,5 K (up to 600 K) ± 2,0 K (up to 1 273 K)	ASTM E 537-76

<sup>(1)</sup> This accuracy is only valid for the simple device as for example described in ASTM D 1120-72; it can be improved with more sophisticated ebulliometer devices.

<sup>(2)</sup> Only valid for pure substances. The use in other circumstances should be justified.



**▼B**

## 1.6. DESCRIPTION OF THE METHODS

The procedures of some test methods have been described in international and national standards (see Appendix).

1.6.1. **Ebulliometer**

See Appendix.

1.6.2. **Dynamic method**

See test method A.4 for the determination of the vapour pressure.

The boiling temperature observed with an applied pressure of 101,325 kPa is recorded.

1.6.3. **Distillation process (boiling range)**

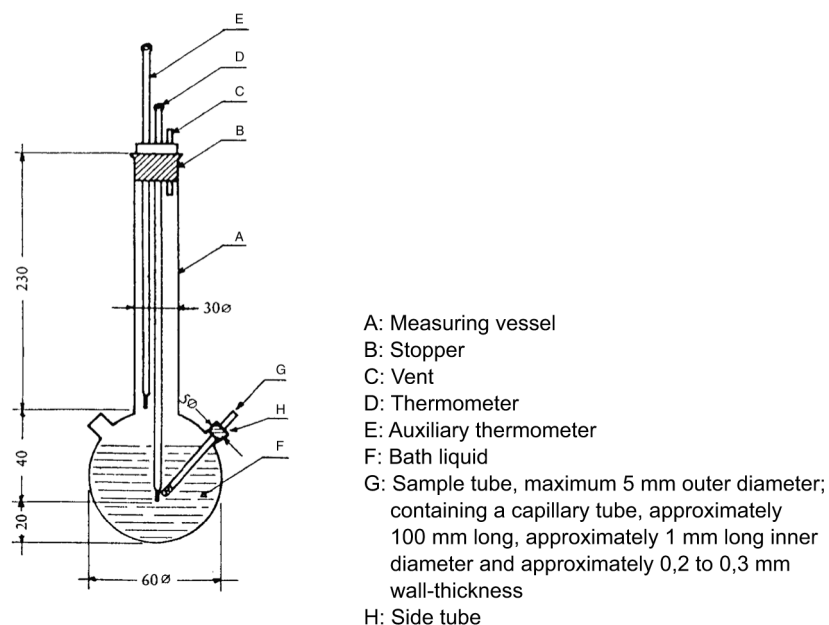
See Appendix.

1.6.4. **Method according to Siwoloboff**

The sample is heated in a melting temperature apparatus in a sample tube, with a diameter of approximately 5 mm (figure 1).

Figure 1 shows a type of standardised melting and boiling temperature apparatus (JIS K 0064) (made of glass, all specifications in millimetres).

Figure 1



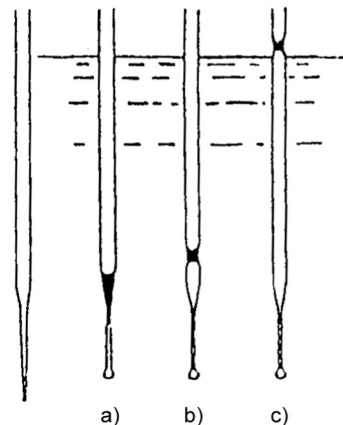
A capillary tube (boiling capillary) which is fused about 1 cm above the lower end is placed in the sample tube. The level to which the test substance is added is such that the fused section of the capillary is below the surface of the liquid. The sample tube containing the boiling capillary is fastened either to the thermometer with a rubber band or is fixed with a support from the side (see figure 2).

▼B

Figure 2

**Principle according to Siwoloboff**

Figure 3

**Modified principle**

The bath liquid is chosen according to boiling temperature. At temperatures up to 573 K, silicone oil can be used. Liquid paraffin may only be used up to 473 K. The heating of the bath liquid should be adjusted to a temperature rise of 3 K/min at first. The bath liquid must be stirred. At about 10 K below the expected boiling temperature, the heating is reduced so that the rate of temperature rise is less than 1 K/min. Upon approach of the boiling temperature, bubbles begin to emerge rapidly from the boiling capillary.

The boiling temperature is that temperature when, on momentary cooling, the string of bubbles stops and fluid suddenly starts rising in the capillary. The corresponding thermometer reading is the boiling temperature of the substance.

In the modified principle (figure 3) the boiling temperature is determined in a melting temperature capillary. It is stretched to a fine point about 2 cm in length (a) and a small amount of the sample is sucked up. The open end of the fine capillary is closed by melting, so that a small air bubble is located at the end. While heating in the melting temperature apparatus (b), the air bubble expands. The boiling temperature corresponds to the temperature at which the substance plug reaches the level of the surface of the bath liquid (c).

**1.6.5. Photocell detection**

The sample is heated in a capillary tube inside a heated metal block.

A light beam is directed, via suitable holes in the block, through the substance onto a precisely calibrated photocell.

During the increase of the sample temperature, single air bubbles emerge from the boiling capillary. When the boiling temperature is reached the number of bubbles increases greatly. This causes a change in the intensity of light, recorded by a photocell, and gives a stop signal to the indicator reading out the temperature of a platinum resistance thermometer located in the block.

This method is especially useful because it allows determinations below room temperature down to 253,15 K (– 20 °C) without any changes in the apparatus. The instrument merely has to be placed in a cooling bath.

**▼ B**1.6.6. **Thermal analysis**1.6.6.1. *Differential thermal analysis*

See Appendix.

1.6.6.2. *Differential scanning calorimetry*

See Appendix.

2. **DATA**

At small deviations from the normal pressure (max.  $\pm 5$  kPa) the boiling temperatures are normalised to  $T_n$  by means of the following number-value equation by Sidney Young:

$$T_n = T + (f_T \times \Delta p)$$

where:

$$\Delta p = (101,325 - p) \text{ [note sign]}$$

P = pressure measurement in kPa

 $f_T$  = rate of change of boiling temperature with pressure in K/kPa

T = measured boiling temperature in K

 $T_n$  = boiling temperature corrected to normal pressure in K

The temperature-correction factors,  $f_T$ , and equations for their approximation are included in the international and national standards mentioned above for many substances.

For example, the DIN 53171 method mentions the following rough corrections for solvents included in paints:

Table 2:

**Temperature — corrections factors  $f_T$** 

Temperature T (K)	Correction factor $f_T$ (K/kPa)
323,15	0,26
348,15	0,28
373,15	0,31
398,15	0,33
423,15	0,35
448,15	0,37
473,15	0,39
498,15	0,41
523,15	0,4
548,15	0,45
573,15	0,47

**▼B****3. REPORTING**

The test report shall, if possible, include the following information:

- method used,
- precise specification of the substance (identity and impurities) and preliminary purification step, if any,
- an estimate of the accuracy.

The mean of at least two measurements which are in the range of the estimated accuracy (see table 1) is reported as the boiling temperature.

The measured boiling temperatures and their mean shall be stated and the pressure(s) at which the measurements were made shall be reported in kPa. The pressure should preferably be close to normal atmospheric pressure.

All information and remarks relevant for the interpretation of results have to be reported, especially with regard to impurities and physical state of the substance.

**4. REFERENCES**

- (1) OECD, Paris, 1981, Test Guideline 103, Decision of the Council C (81) 30 final.
- (2) IUPAC, B. Le Neindre, B. Vodar, editions. Experimental thermodynamics, Butterworths, London, 1975, vol. II.
- (3) R. Weissberger edition: Technique of organic chemistry, Physical methods of organic chemistry, Third Edition, Interscience Publications, New York, 1959, vol. I, Part I, Chapter VIII.

**▼ B***Appendix*

*For additional technical details, the following standards may be consulted for example.*

1.       **Ebulliometer**
  - 1.1.      Melting temperature devices with liquid bath

ASTM D 1120-72	Standard test method for boiling point of engine anti-freezes
----------------	---
  
2.       **Distillation process (boiling range)**

ISO/R 918	Test Method for Distillation (Distillation Yield and Distillation Range)
BS 4349/68	Method for determination of distillation of petroleum products
BS 4591/71	Method for the determination of distillation characteristics
DIN 53171	Lösungsmittel für Anstrichstoffe, Bestimmung des Siedeverlaufes
NF T 20-608	Distillation: détermination du rendement et de l'intervalle de distillation
  
3.       **Differential thermal analysis and differential scanning calorimetry**

ASTM E 537-76	Standard method for assessing the thermal stability of chemicals by methods of differential thermal analysis
ASTM E 473-85	Standard definitions of terms relating to thermal analysis
ASTM E 472-86	Standard practice for reporting thermoanalytical data
DIN 51005	Thermische Analyse, Begriffe

**▼B****A.3. RELATIVE DENSITY****1. METHOD**

The methods described are based on the OECD Test Guideline (1).  
The fundamental principles are given in reference (2).

**1.1. INTRODUCTION**

The methods for determining relative density described are applicable to solid and to liquid substances, without any restriction in respect to their degree of purity. The various methods to be used are listed in table 1.

**1.2. DEFINITIONS AND UNITS**

The relative density  $D_{4}^{20}$  of solids or liquids is the ratio between the mass of a volume of substance to be examined, determined at 20 °C, and the mass of the same volume of water, determined at 4 °C. The relative density has no dimension.

The density,  $\rho$ , of a substance is the quotient of the mass,  $m$ , and its volume,  $v$ .

The density,  $\rho$ , is given, in SI units, in  $\text{kg/m}^3$ .

**1.3. REFERENCE SUBSTANCES (1) (3)**

Reference substances do not need to be employed in all cases when investigating a new substance. They should primarily serve to check the performance of the method from time to time and to allow comparison with results from other methods.

**1.4. PRINCIPLE OF THE METHODS**

Four classes of methods are used.

**1.4.1. Buoyancy methods****1.4.1.1. *Hydrometer* (for liquid substances)**

Sufficiently accurate and quick determinations of density may be obtained by the floating hydrometers, which allow the density of a liquid to be deduced from the depth of immersion by reading a graduated scale.

**1.4.1.2. *Hydrostatic balance* (for liquid and solid substances)**

The difference between the weight of a test sample measured in air and in a suitable liquid (e.g. water) can be employed to determine its density.

For solids, the measured density is only representative of the particular sample employed. For the determination of density of liquids, a body of known volume,  $v$ , is weighed first in air and then in the liquid.

**1.4.1.3. *Immersed body method* (for liquid substances) (4)**

In this method, the density of a liquid is determined from the difference between the results of weighing the liquid before and after immersing a body of known volume in the test liquid.

**▼ B****1.4.2. Pycnometer methods**

For solids or liquids, pycnometers of various shapes and with known volumes may be employed. The density is calculated from the difference in weight between the full and empty pycnometer and its known volume.

**1.4.3. Air comparison pycnometer (for solids)**

The density of a solid in any form can be measured at room temperature with the gas comparison pycnometer. The volume of a substance is measured in air or in an inert gas in a cylinder of variable calibrated volume. For the calculation of density one mass measurement is taken after concluding the volume measurement.

**1.4.4. Oscillating densitometer (5) (6) (7)**

The density of a liquid can be measured by an oscillating densitometer. A mechanical oscillator constructed in the form of a U-tube is vibrated at the resonance frequency of the oscillator which depends on its mass. Introducing a sample changes the resonance frequency of the oscillator. The apparatus has to be calibrated by two liquid substances of known densities. These substances should preferably be chosen such that their densities span the range to be measured.

**1.5. QUALITY CRITERIA**

The applicability of the different methods used for the determination of the relative density is listed in the table.

**1.6. DESCRIPTION OF THE METHODS**

The standards given as examples, which are to be consulted for additional technical details, are attached in the Appendix.

The tests have to be run at 20 °C, and at least two measurements performed.

**2. DATA**

See standards.

**3. REPORTING**

The test report shall, if possible, include the following information:

— method used,

— precise specification of the substance (identity and impurities) and preliminary purification step, if any.

The relative density,  $D_4^{20}$ , shall be reported as defined in 1.2, along with the physical state of the measured substance.

All information and remarks relevant for the interpretation of results have to be reported, especially with regard to impurities and physical state of the substance.



Table:

## Applicability of methods

Method of measurement	Density		Maximum possible dynamic viscosity	Existing Standards
	solid	liquid		
1.4.1.1. Hydrometer		yes	5 Pa s	ISO 387, ISO 649-2, NF T 20-050
1.4.1.2. Hydrostatic balance				
(a) solids	yes			ISO 1183 (A)
(b) liquids		yes	5 Pa s	ISO 901 and 758
1.4.1.3. Immersed body method		yes	20 Pa s	DIN 53217
1.4.2. Pycnometer				ISO 3507
(a) solids	yes			ISO 1183(B), NF T 20-053
(b) liquids		yes	500 Pa s	ISO 758
1.4.3. Air comparison pycnometer	yes			DIN 55990 Teil 3, DIN 53243
1.4.4. Oscillating densitimer		yes	5 Pa s	

## 4.

## REFERENCES

- (1) OECD, Paris, 1981, Test Guideline 109, Decision of the Council C(81) 30 final.
- (2) R. Weissberger ed., Technique of Organic Chemistry, Physical Methods of Organic Chemistry, 3rd ed., Chapter IV, Interscience Publ., New York, 1959, vol. I, Part 1.
- (3) IUPAC, Recommended reference materials for realization of physico-chemical properties, Pure and applied chemistry, 1976, vol. 48, p. 508.
- (4) Wagenbreth, H., Die Tauchkugel zur Bestimmung der Dichte von Flüssigkeiten, Technisches Messen tm, 1979, vol. II, p. 427-430.
- (5) Leopold, H., Die digitale Messung von Flüssigkeiten, Elektronik, 1970, vol. 19, p. 297-302.
- (6) Baumgarten, D., Füllmengenkontrolle bei vorgepackten Erzeugnissen -Verfahren zur Dichtebestimmung bei flüssigen Produkten und ihre praktische Anwendung, Die Pharmazeutische Industrie, 1975, vol. 37, p. 717-726.
- (7) Riemann, J., Der Einsatz der digitalen Dichtemessung im Brauereilaboratorium, Brauwissenschaft, 1976, vol. 9, p. 253-255.



**▼B***Appendix*

*For additional technical details, the following standards may be consulted for example.*

1. **Buoyancy methods**

## 1.1. Hydrometer

DIN 12790, ISO 387	Hydrometer; general instructions
DIN 12791	Part I: Density hydrometers; construction, adjustment and use Part II: Density hydrometers; standardised sizes, designation Part III: Use and test
ISO 649-2	Laboratory glassware: Density hydrometers for general purpose
NF T 20-050	Chemical products for industrial use — Determination of density of liquids — Areometric method
DIN 12793	Laboratory glassware: range find hydrometers

## 1.2. Hydrostatic balance

*For solid substances*

ISO 1183	Method A: Methods for determining the density and relative density of plastics excluding cellular plastics
NF T 20-049	Chemical products for industrial use — Determination of the density of solids other than powders and cellular products — Hydrostatic balance method
ASTM-D-792	Specific gravity and density of plastics by displacement
DIN 53479	Testing of plastics and elastomers; determination of density

*For liquid substances*

ISO 901	ISO 758
---------	---------

**▼ B**

DIN 51757 Testing of mineral oils and related materials; determination of density

ASTM D 941-55, ASTM D 1296-67 and ASTM D 1481-62

ASTM D 1298 Density, specific gravity or API gravity of crude petroleum and liquid petroleum products by hydrometer method

BS 4714 Density, specific gravity or API gravity of crude petroleum and liquid petroleum products by hydrometer method

1.3. Immersed body method

DIN 53217 Testing of paints, varnishes and similar coating materials; determination of density; immersed body method

2. **Pycnometer methods**

2.1. For liquid substances

ISO 3507 Pycnometers

ISO 758 Liquid chemical products; determination of density at 20 °C

DIN 12797 Gay-Lussac pycnometer (for non-volatile liquids which are not too viscous)

DIN 12798 Lipkin pycnometer (for liquids with a kinematic viscosity of less than  $100 \cdot 10^{-6} \text{ m}^2 \text{ s}^{-1}$  at 15 °C)

DIN 12800 Sprengel pycnometer (for liquids as DIN 12798)

DIN 12801 Reischauer pycnometer (for liquids with a kinematic viscosity of less than  $100 \cdot 10^{-6} \text{ m}^2 \text{ s}^{-1}$  at 20 °C, applicable in particular also to hydrocarbons and aqueous solutions as well as to liquids with higher vapour pressure, approximately 1 bar at 90 °C)

**▼ B**

DIN 12806	Hubbard pycnometer (for viscous liquids of all types which do not have too high a vapour pressure, in particular also for paints, varnishes and bitumen)
DIN 12807	Bingham pycnometer (for liquids, as in DIN 12801)
DIN 12808	Jaulmes pycnometer (in particular for ethanol — water mixture)
DIN 12809	Pycnometer with ground-in thermometer and capillary side tube (for liquids which are not too viscous)
DIN 53217	Testing of paints, varnishes and similar products; determination of density by pycnometer
DIN 51757	Point 7: Testing of mineral oils and related materials; determination of density
ASTM D 297	Section 15: Rubber products — chemical analysis
ASTM D 2111	Method C: Halogenated organic compounds
BS 4699	Method for determination of specific gravity and density of petroleum products (graduated bicapillary pycnometer method)
BS 5903	Method for determination of relative density and density of petroleum products by the capillary — stoppered pycnometer method
NF T 20-053	Chemical products for industrial use — Determination of density of solids in powder and liquids — Pycnometric method

**▼B**

## 2.2. For solid substances

ISO 1183	Method B: Methods for determining the density and relative density of plastics excluding cellular plastics
NF T 20-053	Chemical products for industrial use — Determination of density of solids in powder and liquids — Pycnometric method
DIN 19683	Determination of the density of soils

3. **Air comparison pycnometer**

DIN 55990	Part 3: Prüfung von Anstrichstoffen und ähnlichen Beschichtungsstoffen; Pulverlack; Bestimmung der Dichte
DIN 53243	Anstrichstoffe; chlorhaltige Polymere; Prüfung

**▼ M1****A.4. VAPOUR PRESSURE****1. METHOD**

This method is equivalent to OECD TG 104 (2004).

**1.1. INTRODUCTION**

This revised version of method A.4(1) includes one additional method; Effusion method: isothermal thermogravimetry, designed for substances with very low pressures (down to  $10^{-10}$  Pa). In the light of needs for procedures, especially in relation to obtaining vapour pressure for substances with low vapour pressure, other procedures of this method are re-evaluated with respect to other applicability ranges.

At the thermodynamic equilibrium the vapour pressure of a pure substance is a function of temperature only. The fundamental principles are described elsewhere (2)(3).

No single measurement procedure is applicable to the entire range of vapour pressures from less than  $10^{-10}$  to  $10^5$  Pa. Eight methods for measuring vapour pressure are included in this method which can be applied in different vapour pressure ranges. The various methods are compared as to application and measuring range in Table 1. The methods can only be applied for compounds that do not decompose under the conditions of the test. In cases where the experimental methods cannot be applied due to technical reasons, the vapour pressure can also be estimated, and a recommended estimation method is set out in the Appendix.

**1.2. DEFINITIONS AND UNITS**

The vapour pressure of a substance is defined as the saturation pressure above a solid or liquid substance.

The SI unit of pressure, which is the pascal (Pa), should be used. Other units which have been employed historically are given hereafter, together with their conversion factors:

1 Torr	= 1 mm Hg	= $1,333 \times 10^2$ Pa
1 atmosphere	= $1,013 \times 10^5$ Pa	
1 bar	= $10^5$ Pa	

The SI unit of temperature is the kelvin (K). The conversion of degrees Celsius to kelvin is according to the formula:

$$T = t + 273,15$$

where, T is the kelvin or thermodynamic temperature and t is the Celsius temperature.

▼ M1

Table 1

Measuring method	Substances		Estimated repeatability	Estimated reproducibility	Recommended range
	Solid	Liquid			
Dynamic method	Low melting	Yes	up to 25 % 1 to 5 %	up to 25 % 1 to 5 %	10 <sup>3</sup> Pa to 2 × 10 <sup>3</sup> Pa 2 × 10 <sup>3</sup> Pa to 10 <sup>5</sup> Pa
Static method	Yes	Yes	5 to 10 %	5 to 10 %	10 Pa to 10 <sup>5</sup> Pa 10 <sup>-2</sup> Pa to 10 <sup>5</sup> Pa <sup>(1)</sup>
Isoteniscope method	Yes	Yes	5 to 10 %	5 to 10 %	10 <sup>2</sup> Pa to 10 <sup>5</sup> Pa
Effusion method: vapour pressure balance	Yes	Yes	5 to 20 %	up to 50 %	10 <sup>-3</sup> to 1 Pa
Effusion method: Knudsen cell	Yes	Yes	10 to 30 %	—	10 <sup>-10</sup> to 1 Pa
Effusion method: isothermal thermo- gravimetry	Yes	Yes	5 to 30 %	up to 50 %	10 <sup>-10</sup> to 1 Pa
Gas saturation method	Yes	Yes	10 to 30 %	up to 50 %	10 <sup>-10</sup> to 10 <sup>3</sup> Pa
Spinning rotor method	Yes	Yes	10 to 20 %	—	10 <sup>-4</sup> to 0,5 Pa

<sup>(1)</sup> When using a capacitance manometer

## 1.3. PRINCIPLE OF THE TEST

In general, the vapour pressure is determined at various temperatures. In a limited temperature range, the logarithm of the vapour pressure of a pure substance is a linear function of the inverse of the thermodynamic temperature according to the simplified Clapeyron-Clausius equation:

$$\log p = \frac{\Delta H_v}{2,3RT} + \text{constant}$$

where:

p = the vapour pressure in pascals

$\Delta H_v$  = the heat of vaporisation in J mol<sup>-1</sup>

R = the universal gas constant, 8,314 J mol<sup>-1</sup> K<sup>-1</sup>

T = the temperature in K

▼ M1

## 1.4. REFERENCE SUBSTANCES

Reference substances do not need to be employed. They serve primarily to check the performance of a method from time to time as well as to allow comparison between results of different methods.

## 1.5. DESCRIPTION OF THE METHOD

1.5.1. **Dynamic method (Cottrell's method)**1.5.1.1. *Principle*

The vapour pressure is determined by measuring the boiling temperature of the substance at various specified pressures between roughly  $10^3$  and  $10^5$  Pa. This method is also recommended for the determination of the boiling temperature. For that purpose it is useful up to 600 K. The boiling temperatures of liquids are approximately  $0,1$  °C higher at a depth of 3 to 4 cm than at the surface because of the hydrostatic pressure of the column of liquid. In Cottrell's method (4) the thermometer is placed in the vapour above the surface of the liquid and the boiling liquid is made to pump itself continuously over the bulb of the thermometer. A thin layer of liquid which is in equilibrium with vapour at atmospheric pressure covers the bulb. The thermometer thus reads the true boiling point, without errors due to superheating or hydrostatic pressure. The pump originally employed by Cottrell is shown in figure 1. Tube A contains the boiling liquid. A platinum wire B sealed into the bottom facilitates uniform boiling. The side tube C leads to a condenser, and the sheath D prevents the cold condensate from reaching the thermometer E. When the liquid in A is boiling, bubbles and liquid trapped by the funnel are poured via the two arms of the pump F over the bulb of the thermometer.

Figure 1

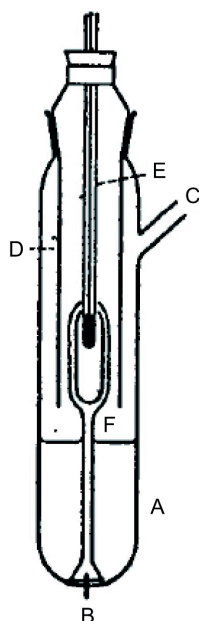
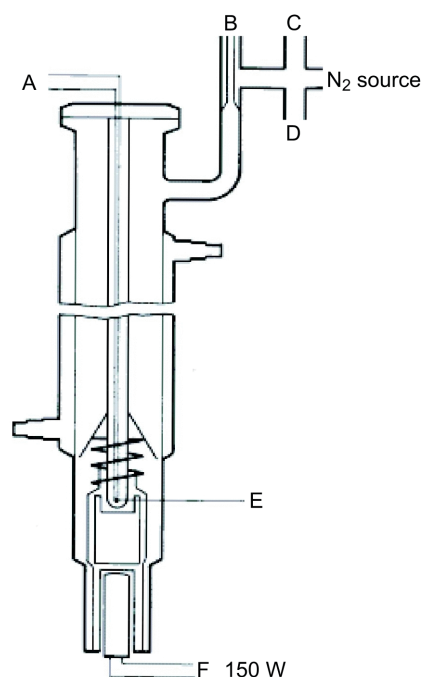


Figure 2



**▼ M1**

Cottrell pump (4)

A: Thermocouple

B: Vacuum buffer volume

C: Pressure gauge

D: Vacuum

E: Measuring point

F: Heating element c.a. 150 W

1.5.1.2. *Apparatus*

A very accurate apparatus, employing the Cottrell principle, is shown in figure 2. It consists of a tube with a boiling section in the lower part, a cooler in the middle part, and an outlet and flange in the upper part. The Cottrell pump is placed in the boiling section which is heated by means of an electrical cartridge. The temperature is measured by a jacketed thermocouple, or resistance thermometer inserting through the flange at the top. The outlet is connected to the pressure regulation system. The latter consists of a vacuum pump, a buffer volume, a manostat for admitting nitrogen for pressure regulation and manometer.

1.5.1.3. *Procedure*

The substance is placed in the boiling section. Problems may be encountered with non-powder solids but these can sometimes be solved by heating the cooling jacket. The apparatus is sealed at the flange and the substance degassed. Frothing substances cannot be measured using this method.

The lowest desired pressure is then set and the heating is switched on. At the same time, the temperature sensor is connected to a recorder.

Equilibrium is reached when a constant boiling temperature is recorded at constant pressure. Particular care must be taken to avoid bumping during boiling. In addition, complete condensation must occur on the cooler. When determining the vapour pressure of low melting solids, care should be taken to prevent the condenser from blocking.

After recording this equilibrium point, a higher pressure is set. The process is continued in this manner until  $10^5$  Pa has been reached (approximately 5 to 10 measuring points in all). As a check, equilibrium points must be repeated at decreasing pressures.

1.5.2. **Static method**

1.5.2.1. *Principle*

In the static method (5), the vapour pressure at thermodynamic equilibrium is determined at a specified temperature. This method is suitable for substances and multicomponent liquids and solids in the range from  $10^{-1}$  to  $10^5$  Pa and, provided care is taken, also in the range 1 to 10 Pa.



▼ **M1**1.5.2.2. *Apparatus*

The equipment consists of a constant-temperature bath (precision of  $\pm 0,2$  K), a container for the sample connected to a vacuum line, a manometer and a system to regulate the pressure. The sample chamber (figure 3a) is connected to the vacuum line via a valve and a differential manometer (U-tube containing a suitable manometer fluid) which serves as zero indicator. Mercury, silicones and phthalates are suitable for use in the differential manometer, depending on the pressure range and the chemical behaviour of the test substance. However, based on environmental concerns, the use of mercury should be avoided, if possible. The test substance must not dissolve noticeably in, or react with, the U-tube fluid. A pressure gauge can be used instead of a U-tube (figure 3b). For the manometer, mercury can be used in the range from normal pressure down to  $10^2$  Pa, while silicone fluids and phthalates are suitable for use below  $10^2$  Pa down to 10 Pa. There are other pressure gauges which can be used below  $10^2$  Pa and heatable membrane capacity manometers can even be used at below  $10^{-1}$  Pa. The temperature is measured on the outside wall of the vessel containing the sample or in the vessel itself.

1.5.2.3. *Procedure*

Using the apparatus as described in figure 3a, fill the U-tube with the chosen liquid, which must be degassed at an elevated temperature before readings are taken. The test substance is placed in the apparatus and degassed at reduced temperature. In the case of a multiple-component sample, the temperature should be low enough to ensure that the composition of the material is not altered. Equilibrium can be established more quickly by stirring. The sample can be cooled with liquid nitrogen or dry ice, but care should be taken to avoid condensation of air or pump-fluid. With the valve over the sample vessel open, suction is applied for several minutes to remove the air. If necessary, the degassing operation is repeated several times.

Figure 3a

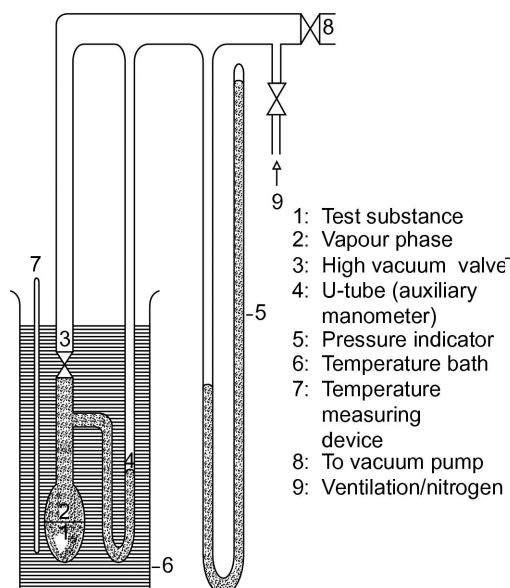
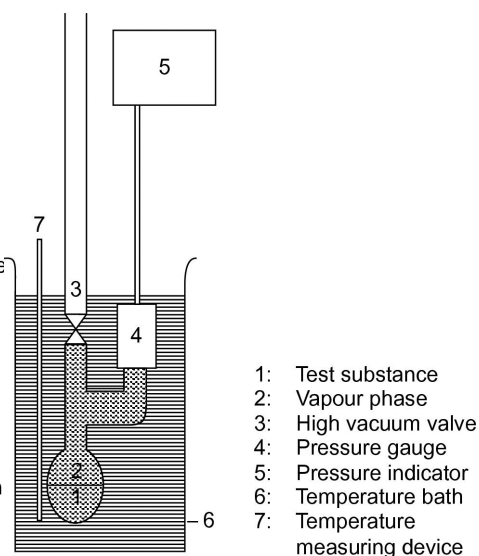


Figure 3b



**▼ M1**

When the sample is heated with the valve closed, the vapour pressure increases. This alters the equilibrium of the fluid in the U-tube. To compensate for this, nitrogen or air is admitted to the apparatus until the differential pressure indicator is at zero again. The pressure required for this can be read off the manometer or off an instrument of higher precision. This pressure corresponds to the vapour pressure of the substance at the temperature of the measurement. Using the apparatus described in figure 3b, the vapour pressure is read off directly.

The vapour pressure is determined at suitably small temperature intervals (approximately 5 to 10 measuring points in all) up to the desired temperature maximum.

Low-temperature readings must be repeated as a check. If the values obtained from the repeated readings do not coincide with the curve obtained for increasing temperature, this may be due to one of the following situations:

- (i) the sample still contains air (e.g. in the case of highly viscous materials) or low-boiling substances which is or are released during heating;
- (ii) the substance undergoes a chemical reaction in the temperature range investigated (e.g. decomposition, polymerisation).

**1.5.3. Isoteniscope Method****1.5.3.1. Principle**

The isoteniscope (6) is based on the principle of the static method. The method involves placing a sample in a bulb maintained at constant temperature and connected to a manometer and a vacuum pump. Impurities more volatile than the substance are removed by degassing at reduced pressure. The vapour pressure of the sample at selected temperatures is balanced by a known pressure of inert gas. The isoteniscope was developed to measure the vapour pressure of certain liquid hydrocarbons but it is appropriate for the investigation of solids as well. The method is usually not suitable for multicomponent systems. Results are subject to only slight errors for samples containing non-volatile impurities. The recommended range is  $10^2$  to  $10^5$  Pa.

**1.5.3.2. Apparatus**

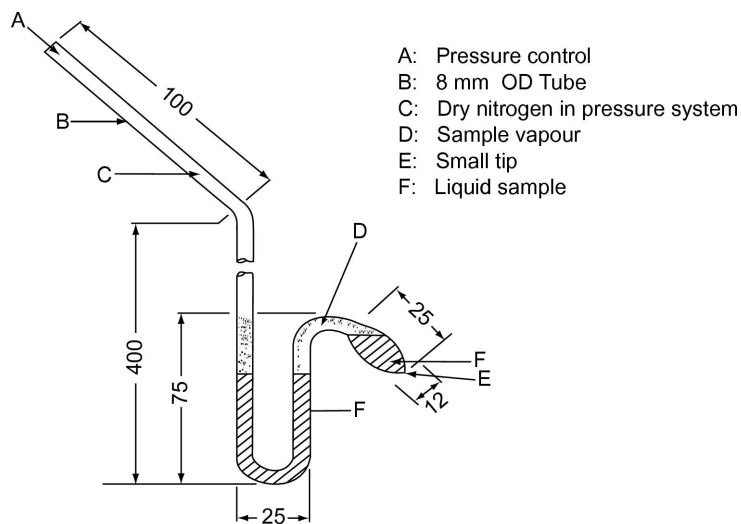
An example of a measuring device is shown in figure 4. A complete description can be found in ASTM D 2879-86 (6).

▼ **M1**

## 1.5.3.3. Procedure

In the case of liquids, the substance itself serves as the fluid in the differential manometer. A quantity of the liquid, sufficient to fill the bulb and the short leg of the manometer, is put in the isoteniscope. The isoteniscope is attached to a vacuum system and evacuated, then filled by nitrogen. The evacuation and purge of the system is repeated twice to remove residual oxygen. The filled isoteniscope is placed in a horizontal position so that the sample spreads out into a thin layer in the sample bulb and manometer. The pressure of the system is reduced to 133 Pa and the sample is gently warmed until it just boils (removal of dissolved gases). The isoteniscope is then placed so that the sample returns to the bulb and fills the short leg of the manometer. The pressure is maintained at 133 Pa. The drawn-out tip of the sample bulb is heated with a small flame until the sample vapour expands sufficiently to displace part of the sample from the upper part of the bulb and manometer arm into the manometer, creating a vapour-filled, nitrogen-free space. The isoteniscope is then placed in a constant temperature bath, and the pressure of the nitrogen is adjusted until it equals that of the sample. At the equilibrium, the pressure of the nitrogen equals the vapour pressure of the substance.

Figure 4



(Dimension in mm)

In the case of solids, and depending on the pressure and temperature ranges, manometer liquids such as silicon fluids or phthalates are used. The degassed manometer liquid is put in a bulge provided on the long arm of the isoteniscope. Then the solid to be investigated is placed in the sample bulb and is degassed at an elevated temperature. After that, the isoteniscope is inclined so that the manometer liquid can flow into the U-tube.

▼ **M1**1.5.4. **Effusion method: vapour pressure balance (7)**1.5.4.1. *Principle*

A sample of the test substance is heated in a small furnace and placed in an evacuated bell jar. The furnace is covered by a lid which carries small holes of known diameters. The vapour of the substance, escaping through one of the holes, is directed onto a balance pan of a highly sensitive balance which is also enclosed in the evacuated bell jar. In some designs the balance pan is surrounded by a refrigeration box, providing heat dissipation to the outside by thermal conduction, and is cooled by radiation so that the escaping vapour condenses on it. The momentum of the vapour jet acts as a force on the balance. The vapour pressure can be derived in two ways: directly from the force on the balance pan and also from the evaporation rate using the Hertz-Knudsen equation (2):

$$p = G \sqrt{\frac{2\pi RT \times 10^3}{M}}$$

where:

G = evaporation rate ( $\text{kg s}^{-1} \text{m}^{-2}$ )

M = molar mass ( $\text{g mol}^{-1}$ )

T = temperature (K)

R = universal gas constant ( $\text{J mol}^{-1} \text{K}^{-1}$ )

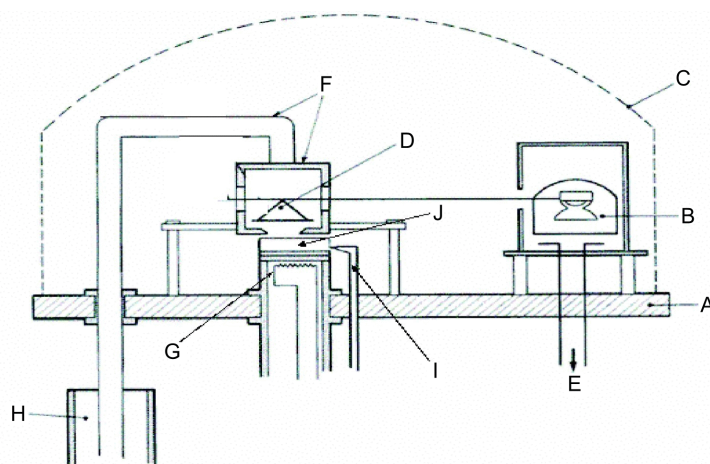
P = vapour pressure (Pa)

The recommended range is  $10^{-3}$  to 1 Pa.

1.5.4.2. *Apparatus*

The general principle of the apparatus is illustrated in figure 5.

Figure 5



- |                            |   |
|----------------------------|---|
| A: Base plate              | F: Refrigeration box and cooling bar    |
| B: Moving coil instrument  | G: Evaporator furnace                   |
| C: Bell jar                | H: Dewar flask with liquid nitrogen     |
| D: Balance with scale pan  | I: Measurement of temperature of sample |
| E: Vacuum measuring device | J: Test Substance                       |

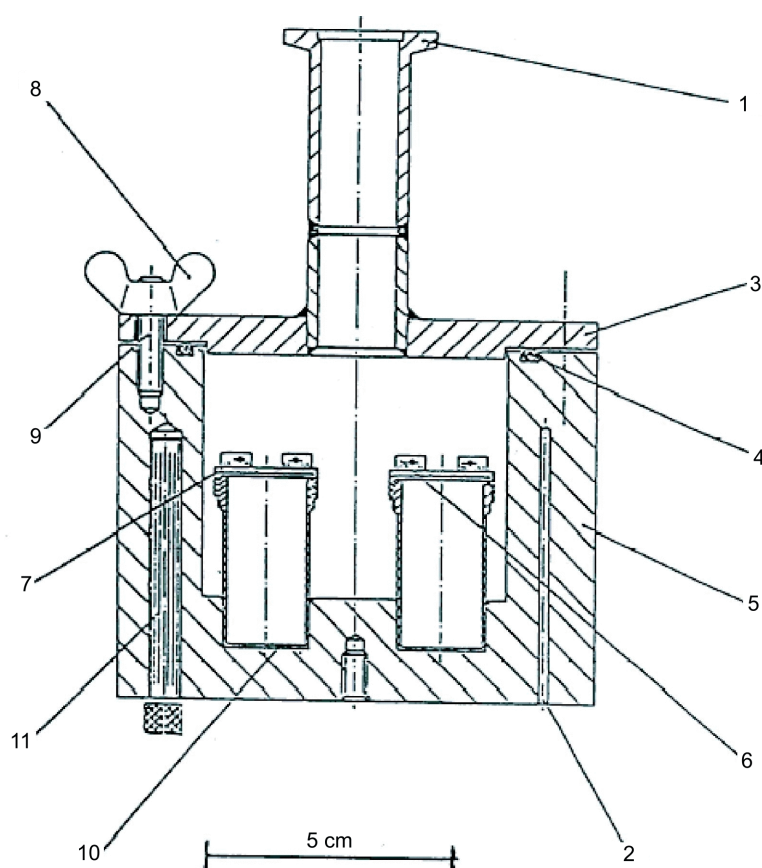
▼ **M1**1.5.5. **Effusion method: Knudsen cell**1.5.5.1. *Principle*

The method is based on the estimation of the mass of test substance flowing out per unit of time of a Knudsen cell (8) in the form of vapour, through a micro-orifice under ultra-vacuum conditions. The mass of effused vapour can be obtained either by determining the loss of mass of the cell or by condensing the vapour at low temperature and determining the amount of volatilised substance using chromatography. The vapour pressure is calculated by applying the Hertz-Knudsen relation (see section 1.5.4.1) with correction factors that depend on parameters of the apparatus (9). The recommended range is  $10^{-10}$  to 1 Pa (10)(11)(12)(13)(14).

1.5.5.2. *Apparatus*

The general principle of the apparatus is illustrated in figure 6.

Figure 6



- |  |                                    |
|--|------------------------------------|
| 1: Connection to vacuum  | 7: Threaded lid                    |
| 2: Wells from platinum resistance thermometer or temperature measurement and control | 8: Butterfly nuts                  |
| 3: Lid for vacuum tank   | 9: Bolts                           |
| 4: O-ring  | 10: Stainless steel effusion cells |
| 5: Aluminum vacuum tank  | 11: Heater cartridge               |
| 6: Device for installing and removing the effusion cells                             |                                    |

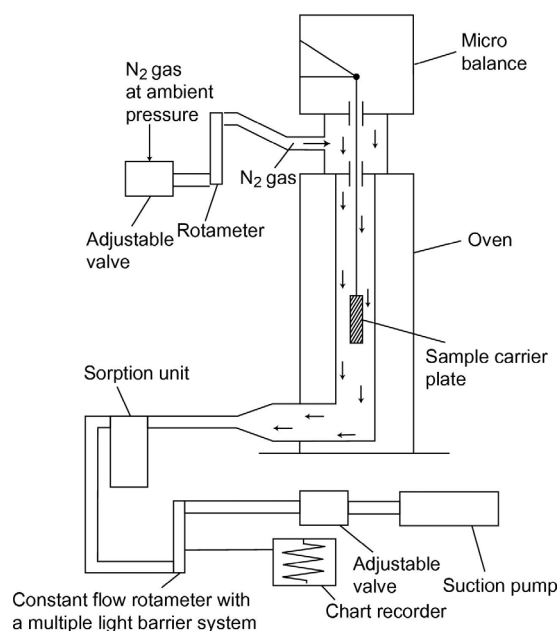
▼ **M1**1.5.6. **Effusion method: isothermal thermogravimetry**1.5.6.1. *Principle*

The method is based on the determination of accelerated evaporation rates for the test substance at elevated temperatures and ambient pressure using thermogravimetry (10)(15)(16)(17)(18)(19)(20). The evaporation rates  $v_T$  result from exposing the selected compound to a slowly flowing inert gas atmosphere, and monitoring the weight loss at defined isothermal temperatures  $T$  in Kelvin over appropriate periods of time. The vapour pressures  $p_T$  are calculated from the  $v_T$  values by using the linear relationship between the logarithm of the vapour pressure and the logarithm of the evaporation rate. If necessary, an extrapolation to temperatures of 20 and 25 °C can be made by regression analysis of  $\log p_T$  vs.  $1/T$ . This method is suitable for substances with vapour pressures as low as  $10^{-10}$  Pa ( $10^{-12}$  mbar) and with purity as close as possible to 100 % to avoid the misinterpretation of measured weight losses.

1.5.6.2. *Apparatus*

The general principle of the experimental set-up is shown in figure 7.

Figure 7



The sample carrier plate, hanging on a microbalance in a temperature controlled chamber, is swept by a stream of dry nitrogen gas which carries the vaporised molecules of the test substance away. After leaving the chamber, the gas stream is purified by a sorption unit.

1.5.6.3. *Procedure*

The test substance is applied to the surface of a roughened glass plate as a homogeneous layer. In the case of solids, the plate is wetted uniformly by a solution of the substance in a suitable solvent and dried in an inert atmosphere. For the measurement, the coated plate is hung into the thermogravimetric analyser and subsequently its weight loss is measured continuously as a function of time.

**▼ M1**

The evaporation rate  $v_T$  at a definite temperature is calculated from the weight loss  $\Delta m$  of the sample plate by

$$v_T = \frac{\Delta m}{F \cdot t} \left( \text{gcm}^{-2}\text{h}^{-1} \right)$$

where  $F$  is the surface area of the coated test substances, normally the surface area of the sample plate, and  $t$  is the time for weight loss  $\Delta m$ .

The vapour pressure  $p_T$  is calculated on the basis of its function of evaporation rate  $v_T$ :

$$\text{Log } p_T = C + D \cdot \log v_T$$

where  $C$  and  $D$  are constants specific for the experimental arrangement used, depending on the diameter of the measurement chamber and on the gas flow rate. These constants must be determined once by measuring a set of compounds with known vapour pressure and regressing  $\log p_T$  vs.  $\log v_T$  (11)(21)(22).

The relationship between the vapour pressure  $p_T$  and the temperature  $T$  in Kelvin is given by

$$\text{Log } p_T = A + B \cdot 1/T$$

where  $A$  and  $B$  are constants obtained by regressing  $\log p_T$  vs.  $1/T$ . With this equation, the vapour pressure can be calculated for any other temperature by extrapolation.

### 1.5.7. Gas saturation method (23)

#### 1.5.7.1. Principle

Inert gas is passed, at room temperature and at a known flow rate, through or over a sample of the test substance, slowly enough to ensure saturation. Achieving saturation in the gas phase is of critical importance. The transported substance is trapped, generally using a sorbent, and its amount is determined. As an alternative to vapour trapping and subsequent analysis, in-train analytical techniques, like gas chromatography, may be used to determine quantitatively the amount of material transported. The vapour pressure is calculated on the assumption that the ideal gas law is obeyed and that the total pressure of a mixture of gases is equal to the sum of the pressures of the component gases. The partial pressure of the test substance, i.e. the vapour pressure, is calculated from the known total gas volume and from the weight of the material transported.

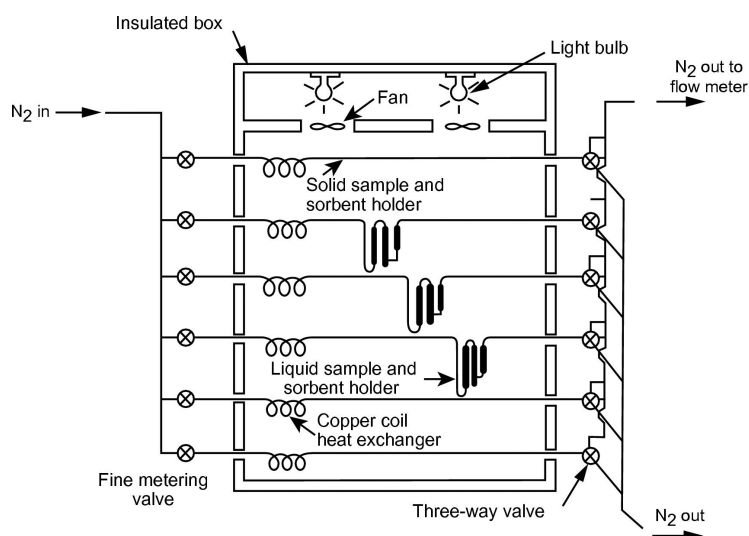
The gas saturation procedure is applicable to solid or liquid substances. It can be used for vapour pressures down to  $10^{-10}$  Pa (10)(11)(12)(13)(14). The method is most reliable for vapour pressures below  $10^3$  Pa. Above  $10^3$  Pa, the vapour pressures are generally overestimated, probably due to aerosol formation. Since the vapour pressure measurements are made at room temperature, the need to extrapolate data from high temperatures is not necessary and high temperature extrapolation, which can often cause serious errors, is avoided.

#### 1.5.7.2. Apparatus

The procedure requires the use of a constant-temperature box. The sketch in figure 8 shows a box containing three solid and three liquid sample holders, which allow for the triplicate analysis of either a solid or a liquid sample. The temperature is controlled to  $\pm 0,5$  °C or better.

▼ M1

Figure 8



In general, nitrogen is used as an inert carrier gas but, occasionally, another gas may be required (24). The carrier gas must be dry. The gas stream is split into 6 streams, controlled by needle valves (approximately 0,79 mm orifice), and flows into the box via 3,8 mm i.d. copper tubing. After temperature equilibration, the gas flows through the sample and the sorbent trap and exits from the box.

Solid samples are loaded into 5 mm i.d. glass tubing between glass wool plugs (see Figure 9). Figure 10 shows a liquid sample holder and sorbent system. The most reproducible method for measuring the vapour pressure of liquids is to coat the liquid on glass beads or on an inert sorbent such as silica, and to pack the holder with these beads. As an alternative, the carrier gas may be made to pass a coarse frit and bubble through a column of the liquid test substance.

Figure 9

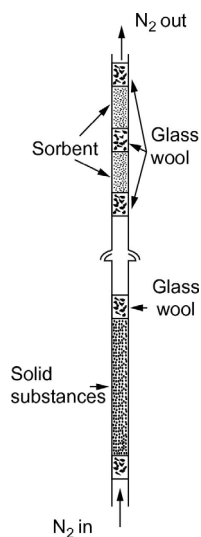
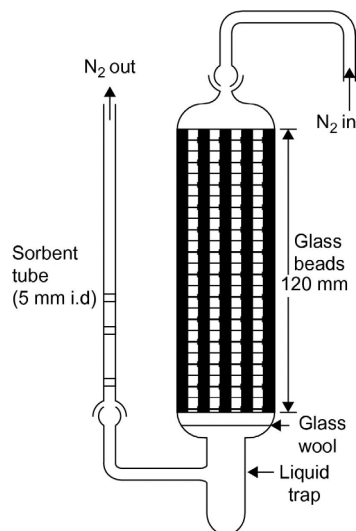


Figure 10





**▼ M1**

The sorbent system contains a front and a backup sorbent section. At very low vapour pressures, only small amounts are retained by the sorbent and the adsorption on the glass wool and the glass tubing between the sample and the sorbent may be a serious problem.

Traps cooled with solid CO<sub>2</sub> are another efficient way for collecting the vaporised material. They do not cause any back pressure on the saturator column and it is also easy to quantitatively remove the trapped material.

**1.5.7.3. Procedure**

The flow rate of the effluent carrier gas is measured at room temperature. The flow rate is checked frequently during the experiment to assure that there is an accurate value for the total volume of carrier gas. Continuous monitoring with a mass flow-meter is preferred. Saturation of the gas phase may require considerable contact time and hence quite low gas flow rates (25).

At the end of the experiment, both the front and backup sorbent sections are analysed separately. The compound on each section is desorbed by adding a solvent. The resulting solutions are analysed quantitatively to determine the weight desorbed from each section. The choice of the analytical method (also the choice of sorbent and desorbing solvent) is dictated by the nature of the test material. The desorption efficiency is determined by injecting a known amount of sample onto the sorbent, desorbing it and analysing the amount recovered. It is important to check the desorption efficiency at or near the concentration of the sample under the test conditions.

To assure that the carrier gas is saturated with the test substance, three different gas flow rates are used. If the calculated vapour pressure shows no dependence on flow rate, the gas is assumed to be saturated.

The vapour pressure is calculated through the equation:

$$p = \frac{W}{V} \times \frac{RT}{M}$$

where:

p = vapour pressure (Pa)

W = mass of evaporated test substance (g)

V = volume of saturated gas (m<sup>3</sup>)

R = universal gas constant 8,314 (J mol<sup>-1</sup> K<sup>-1</sup>)

T = temperature (K)

M = molar mass of test substance (g mol<sup>-1</sup>)

Measured volumes must be corrected for pressure and temperature differences between the flow meter and the saturator.

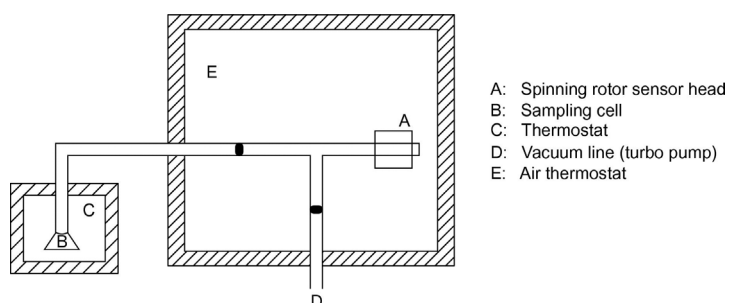
**▼ M1****1.5.8. Spinning rotor****1.5.8.1. Principle**

This method uses a spinning rotor viscosity gauge, in which the measuring element is a small steel ball which, suspended in a magnetic field, is made to spin by rotating fields (26)(27)(28). Pick-up coils allow its spinning rate to be measured. When the ball has reached a given rotational speed, usually about 400 revolutions per second, energising is stopped and deceleration, due to gas friction, takes place. The drop of rotational speed is measured as a function of time. The vapour pressure is deduced from the pressure-dependent slow-down of the steel ball. The recommended range is  $10^{-4}$  to 0,5 Pa.

**1.5.8.2. Apparatus**

A schematic drawing of the experimental set-up is shown in figure 11. The measuring head is placed in a constant-temperature enclosure, regulated within 0,1 °C. The sample container is placed in a separate enclosure, also regulated within 0,1 °C. All other parts of the set-up are kept at a higher temperature to prevent condensation. The whole apparatus is connected to a high-vacuum system.

Figure 11



- A: Spinning rotor sensor head
- B: Sampling cell
- C: Thermostat
- D: Vacuum line (turbo pump)
- E: Air thermostat

**2. DATA AND REPORTING****2.1. DATA**

The vapour pressure from any of the preceding methods should be determined for at least two temperatures. Three or more are preferred in the range from 0 to 50 °C, in order to check the linearity of the vapour pressure curve. In case of Effusion method (Knudsen cell and isothermal thermogravimetry) and Gas saturation method, 120 to 150 °C is recommended for the measuring temperature range instead of 0 to 50 °C.

**2.2. TEST REPORT**

The test report must include the following information:

- method used,

**▼ M1**

- precise specification of the substance (identity and impurities) and preliminary purification step, if any,
- at least two vapour pressure and temperature values — and preferably three or more — required in the range from 0 to 50 °C (or 120 to 150 °C),
- at least one of the temperatures should be at or below 25 °C, if technically possible according to the chosen method,
- all original data,
- a log p versus 1/T curve,
- an estimate of the vapour pressure at 20 or 25 °C.

If a transition (change of state, decomposition) is observed, the following information should be noted:

- nature of the change,
- temperature at which the change occurs at atmospheric pressure,
- vapour pressure at 10 and 20 °C below the transition temperature and 10 and 20 °C above this temperature (unless the transition is from solid to gas).

All information and remarks relevant for the interpretation of results have to be reported, especially with regard to impurities and physical state of the substance.

### 3. LITERATURE

- (1) *Official Journal of the European Communities* L 383 A, 26-47 (1992).
- (2) Ambrose, D. (1975). *Experimental Thermodynamics*, Vol. II, Le Neindre, B., and Vodar, B., Eds., Butterworths, London.
- (3) Weissberger R., ed. (1959). *Technique of Organic Chemistry*, Physical Methods of Organic Chemistry, 3rd ed., Vol. I, Part I. Chapter IX, Interscience Publ., New York.
- (4) Glasstone, S. (1946). *Textbook of Physical Chemistry*, 2nd ed., Van Nostrand Company, New York.
- (5) NF T 20-048 AFNOR (September 1985). Chemical products for industrial use — Determination of vapour pressure of solids and liquids within a range from  $10^{-1}$  to  $10^5$  Pa — Static method.
- (6) ASTM D 2879-86, Standard test method for vapour pressure — temperature relationship and initial decomposition temperature of liquids by isoteniscope.
- (7) NF T 20-047 AFNOR (September 1985). Chemical products for industrial use — Determination of vapour pressure of solids and liquids within range from  $10^{-3}$  to 1 Pa — Vapour pressure balance method.

▼ M1

- (8) Knudsen, M. (1909). *Ann. Phys. Lpz.*, 29, 1979; (1911), 34, 593.
- (9) Ambrose, D., Lawrenson, I.J., Sprake, C.H.S. (1975). *J. Chem. Thermodynamics* 7, 1173.
- (10) Schmuckler, M.E., Barefoot, A.C., Kleier, D.A., Cobranchi, D.P. (2000), Vapor pressures of sulfonylurea herbicides; *Pest Management Science* 56, 521-532.
- (11) Tomlin, C.D.S. (ed.), *The Pesticide Manual*, Twelfth Edition (2000).
- (12) Friedrich, K., Stammbach, K., Gas chromatographic determination of small vapour pressures determination of the vapour pressures of some triazine herbicides. *J. Chromatog.* 16 (1964), 22-28.
- (13) Grayson, B.T., Fosbraey, L.A., *Pesticide Science* 16 (1982), 269-278.
- (14) Rordorf, B.F., Prediction of vapor pressures, boiling points and enthalpies of fusion for twenty-nine halogenated dibenzo-p-dioxins, *Thermochimia Acta* 112 Issue 1 (1987), 117-122.
- (15) Gückel, W., Synnatschke, G., Ritttig, R., A Method for Determining the Volatility of Active Ingredients Used in Plant Protection; *Pesticide Science* 4 (1973) 137-147.
- (16) Gückel, W., Synnatschke, G., Ritttig, R., A Method for Determining the Volatility of Active Ingredients Used in Plant Protection II. Application to Formulated Products; *Pesticide Science* 5 (1974) 393-400.
- (17) Gückel, W., Kaestel, R., Lewerenz, J., Synnatschke, G., A Method for Determining the Volatility of Active Ingredients Used in Plant Protection. Part III: The Temperature Relationship between Vapour Pressure and Evaporation Rate; *Pesticide Science* 13 (1982) 161-168.
- (18) Gückel, W., Kaestel, R., Kroehl, T., Parg, A., Methods for Determining the Vapour Pressure of Active Ingredients Used in Crop Protection. Part IV: An Improved Thermogravimetric Determination Based on Evaporation Rate; *Pesticide Science* 45 (1995) 27-31.
- (19) Kroehl, T., Kaestel, R., Koenig, W., Ziegler, H., Koehle, H., Parg, A., Methods for Determining the Vapour Pressure of Active Ingredients Used in Crop Protection. Part V: Thermogravimetry Combined with Solid Phase MicroExtraction (SPME); *Pesticide Science*, 53 (1998) 300-310.
- (20) Tesconi, M., Yalkowsky, S.H., A Novel Thermogravimetric Method for Estimating the Saturated Vapor Pressure of Low-Volatility Compounds; *Journal of Pharmaceutical Science* 87-(12) (1998) 1512-20.
- (21) Lide, D.R. (ed.), *CRC Handbook of Chemistry and Physics*, 81st ed. (2000), Vapour Pressure in the Range — 25 °C to 150 °C.
- (22) Meister, R.T. (ed.), *Farm Chemicals Handbook*, Vol. 88 (2002).
- (23) 40 CFR, 796. (1993). pp 148-153, Office of the Federal Register, Washington DC.

**▼ M1**

- (24) Rordorf B.F. (1985). *Thermochimica Acta* 85, 435.
- (25) Westcott et al. (1981). *Environ. Sci. Technol.* 15, 1375.
- (26) Messer G., Röhl, P., Grosse G., and Jitschin W. (1987). *J. Vac. Sci. Technol. (A)*, 5(4), 2440.
- (27) Comsa G., Fremerey J.K., and Lindenau, B. (1980). *J. Vac. Sci. Technol.* 17(2), 642.
- (28) Fremerey, J.K. (1985). *J. Vac. Sci. Technol. (A)*, 3(3), 1715.

▼ M1*Appendix***Estimation method**

## INTRODUCTION

Estimated values of the vapour pressure can be used:

- for deciding which of the experimental methods is appropriate,
- for providing an estimate or limit value in cases where the experimental method cannot be applied due to technical reasons.

## ESTIMATION METHOD

The vapour pressure of liquids and solids can be estimated by use of the modified Watson correlation (a). The only experimental data required is the normal boiling point. The method is applicable over the pressure range from  $10^5$  Pa to  $10^{-5}$  Pa.

Detailed information on the method is given in 'Handbook of Chemical Property Estimation Methods' (b). See also OECD Environmental Monograph No.67 (c).

## CALCULATION PROCEDURE

The vapour pressure is calculated as follows:

$$\ln P_{vp} \approx \frac{\Delta H_{vb}}{\Delta Z_b R T_b} \left[ 1 - \frac{\left(3 - 2 \frac{T}{T_b}\right)^m}{\frac{T}{T_b}} - 2m \left(3 - 2 \frac{T}{T_b}\right)^{m-1} \ln \frac{T}{T_b} \right]$$

where:

T = temperature of interest

$T_b$  = normal boiling point

$P_{vp}$  = vapour pressure at temperature T

$\Delta H_{vb}$  = heat of vaporisation

$\Delta Z_b$  = compressibility factor (estimated at 0,97)

m = empirical factor depending on the physical state at the temperature of interest

Further,

$$\frac{\Delta H_{vb}}{T_b} = K_F (8,75 + R \ln T_b)$$

where,  $K_F$  is an empirical factor considering the polarity of the substance. For several compound types,  $K_F$  factors are listed in reference (b).

**▼ M1**

Quite often, data are available in which a boiling point at reduced pressure is given. In such a case, the vapour pressure is calculated as follows:

$$\ln P_{vp} \approx \ln P_1 + \frac{\Delta H_{vl}}{\Delta Z_b R T_1} \left[ 1 - \left( 3 - 2 \frac{T}{T_1} \right)^m \frac{T_1}{T} - 2m \left( 3 - 2 \frac{T}{T_1} \right)^{m-1} \ln \frac{T}{T_1} \right]$$

where,  $T_1$  is the boiling point at the reduced pressure  $P_1$ .

**REPORT**

When using the estimation method, the report shall include a comprehensive documentation of the calculation.

**LITERATURE**

- (a) Watson, K.M. (1943). *Ind. Eng. Chem*, 35, 398.
- (b) Lyman, W.J., Reehl, W.F., Rosenblatt, D.H. (1982). *Handbook of Chemical Property Estimation Methods*, McGraw-Hill.
- (c) OECD Environmental Monograph No.67. *Application of Structure-Activity Relationships to the Estimation of Properties Important in Exposure Assessment* (1993).

**▼B****A.5. SURFACE TENSION****1. METHOD**

The methods described are based on the OECD Test Guideline (1).  
The fundamental principles are given in reference (2).

**1.1. INTRODUCTION**

The described methods are to be applied to the measurement of the surface tension of aqueous solutions.

It is useful to have preliminary information on the water solubility, the structure, the hydrolysis properties and the critical concentration for micelles formation of the substance before performing these tests.

The following methods are applicable to most chemical substances, without any restriction in respect to their degree of purity.

The measurement of the surface tension by the ring tensiometer method is restricted to aqueous solutions with a dynamic viscosity of less than approximately 200 mPa s.

**1.2. DEFINITIONS AND UNITS**

The free surface enthalpy per unit of surface area is referred to as surface tension.

The surface tension is given as:

N/m (SI unit) or

mN/m (SI sub-unit)

1 N/m =  $10^3$  dynes/cm

1 mN/m = 1 dyne/cm in the obsolete cgs system

**1.3. REFERENCE SUBSTANCES**

Reference substances do not need to be employed in all cases when investigating a new substance. They should primarily serve to check the performance of the method from time to time and to allow comparison with results from other methods.

Reference substances which cover a wide range of surface tensions are given in references 1 and 3.

**1.4. PRINCIPLE OF THE METHODS**

The methods are based on the measurement of the maximum force which is necessary to exert vertically, on a stirrup or a ring in contact with the surface of the liquid being examined placed in a measuring cup, in order to separate it from this surface, or on a plate, with an edge in contact with the surface, in order to draw up the film that has formed.

Substances which are soluble in water at least at a concentration of 1 mg/l are tested in aqueous solution at a single concentration.

**1.5. QUALITY CRITERIA**

These methods are capable of greater precision than is likely to be required for environmental assessment.



**▼B**

## 1.6. DESCRIPTION OF THE METHODS

A solution of the substance is prepared in distilled water. The concentration of this solution should be 90 % of the saturation solubility of the substance in water; when this concentration exceeds 1 g/l, a concentration of 1 g/l is used for testing. Substances with water solubility lower than 1 mg/l need not be tested.

1.6.1. **Plate method**

See ISO 304 and NF T 73-060 (Surface active agents — determination of surface tension by drawing up liquid films).

1.6.2. **Stirrup method**

See ISO 304 and NF T 73-060 (Surface active agents — determination of surface tension by drawing up liquid films).

1.6.3. **Ring method**

See ISO 304 and NF T 73-060 (Surface active agents — determination of surface tension by drawing up liquid films).

1.6.4. **OECD harmonised ring method**1.6.4.1. *Apparatus*

Commercially available tensiometers are adequate for this measurement. They consist of the following elements:

- mobile sample table,
- force measuring system,
- measuring body (ring),
- measurement vessel.

1.6.4.1.1. *Mobile sample table*

The mobile sample table is used as a support for the temperature-controlled measurement vessel holding the liquid to be tested. Together with the force measuring system, it is mounted on a stand.

1.6.4.1.2. *Force measuring system*

The force measuring system (see figure) is located above the sample table. The error of the force measurement shall not exceed  $\pm 10^{-6}$  N, corresponding to an error limit of  $\pm 0,1$  mg in a mass measurement. In most cases, the measuring scale of commercially available tensiometers is calibrated in mN/m so that the surface tension can be read directly in mN/m with an accuracy of 0,1 mN/m.

**▼B**

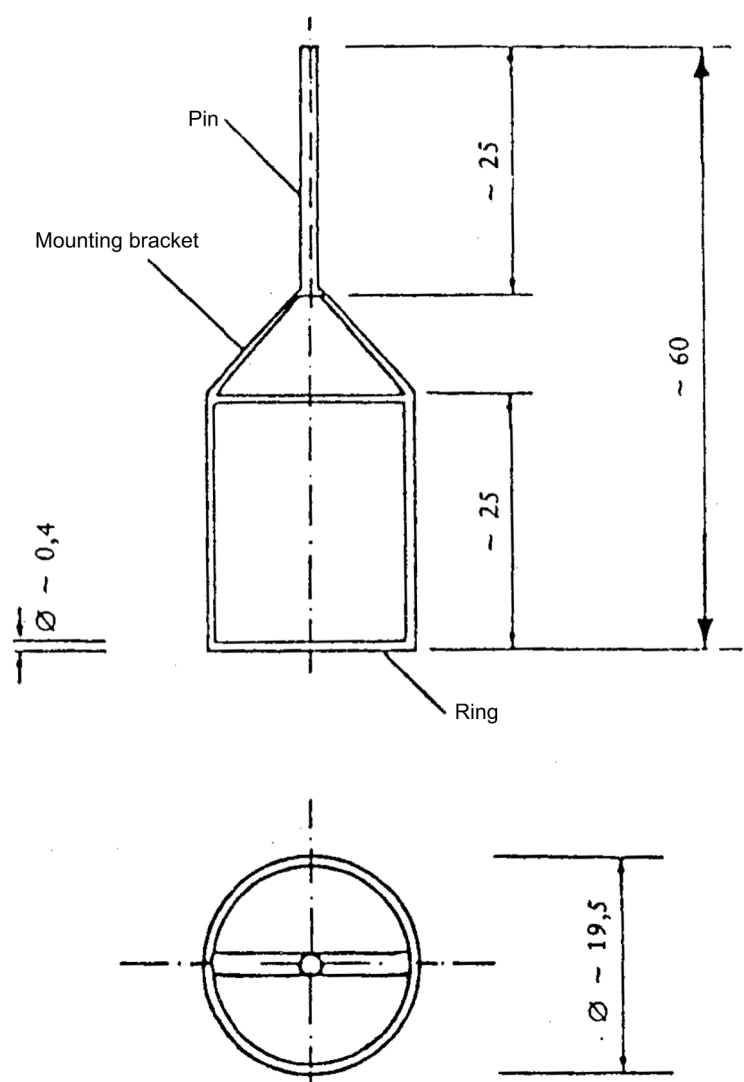
## 1.6.4.1.3. Measuring body (ring)

The ring is usually made of a platinum-iridium wire of about 0,4 mm thickness and a mean circumference of 60 mm. The wire ring is suspended horizontally from a metal pin and a wire mounting bracket to establish the connection to the force measuring system (see figure).

Figure

**Measuring body**

(All dimensions expressed in millimetres)



## 1.6.4.1.4. Measurement vessel

The measurement vessel holding the test solution to be measured shall be a temperature-controlled glass vessel. It shall be designed so that during the measurement the temperature of the test solution liquid and the gas phase above its surface remains constant and that the sample cannot evaporate. Cylindrical glass vessels having an inside diameter of not less than 45 mm are acceptable.

**▼ B**1.6.4.2. *Preparation of the apparatus*1.6.4.2.1. *Cleaning*

Glass vessels shall be cleaned carefully. If necessary they shall be washed with hot chromo-sulphuric acid and subsequently with syrupy phosphoric acid (83 to 98 % by weight of H<sub>3</sub>PO<sub>4</sub>), thoroughly rinsed in tap water and finally washed with double-distilled water until a neutral reaction is obtained and subsequently dried or rinsed with part of the sample liquid to be measured.

The ring shall first be rinsed thoroughly in water to remove any substances which are soluble in water, briefly immersed in chromo-sulphuric acid, washed in double-distilled water until a neutral reaction is obtained and finally heated briefly above a methanol flame.

*Note:*

Contamination by substances which are not dissolved or destroyed by chromo-sulphuric acid or phosphoric acid, such as silicones, shall be removed by means of a suitable organic solvent.

1.6.4.2.2. *Calibration of the apparatus*

The validation of the apparatus consists of verifying the zero point and adjusting it so that the indication of the instrument allows reliable determination in mN/m.

*Mounting:*

The apparatus shall be levelled, for instance by means of a spirit level on the tensiometer base, by adjusting the levelling screws in the base.

*Zero point adjustment:*

After mounting the ring on the apparatus and prior to immersion in the liquid, the tensiometer indication shall be adjusted to zero and the ring checked for parallelism to the liquid surface. For this purpose, the liquid surface can be used as a mirror.

*Calibrations:*

The actual test calibration can be accomplished by means of either of two procedures:

- (a) Using a mass: procedure using riders of known mass between 0,1 and 1,0 g placed on the ring. The calibration factor,  $\Phi_a$  by which all the instrument readings must be multiplied, shall be determined according to equation (1).

$$\Phi_a = \frac{\sigma_r}{\sigma_a}$$

where:

$$\sigma_r = \frac{mg}{2b} \text{ (mN/m)}$$

m = mass of the rider (g)

g = gravity acceleration (981 cm s<sup>-2</sup> at sea level)

b = mean circumference of the ring (cm)

$\sigma_a$  = reading of the tensiometer after placing the rider on the ring (mN/m).

**▼ B**

- (b) Using water: procedure using pure water whose surface tension at, for instance, 23 °C is equal to 72,3 mN/m. This procedure is accomplished faster than the weight calibration but there is always the danger that the surface tension of the water is falsified by traces of contamination by surfactants.

The calibration factor,  $\Phi_b$  by which all the instrument readings shall be multiplied, shall be determined in accordance with the equation (2):

$$\Phi_b = \frac{\sigma_o}{\sigma_g}$$

where:

$\sigma_o$  = value cited in the literature for the surface tension of water (mN/m)

$\sigma_g$  = measured value of the surface tension of the water (mN/m) both at the same temperature.

#### 1.6.4.3. *Preparation of samples*

Aqueous solutions shall be prepared of the substances to be tested, using the required concentrations in water, and shall not contain any non-dissolved substances.

The solution must be maintained at a constant temperature ( $\pm 0,5$  °C). Since the surface tension of a solution in the measurement vessel alters over a period of time, several measurements shall be made at various times and a curve plotted showing surface tension as a function of time. When no further change occurs, a state of equilibrium has been reached.

Dust and gaseous contamination by other substances interfere with the measurement. The work shall therefore be carried out under a protective cover.

#### 1.6.5. **Test conditions**

The measurement shall be made at approximately 20 °C and shall be controlled to within  $\pm 0,5$  °C.

#### 1.6.6. **Performance of test**

The solutions to be measured shall be transferred to the carefully cleaned measurement vessel, taking care to avoid foaming, and subsequently the measurement vessel shall be placed onto the table of the test apparatus. The table-top with measurement vessel shall be raised until the ring is immersed below the surface of the solution to be measured. Subsequently, the table-top shall be lowered gradually and evenly (at a rate of approximately 0,5 cm/min) to detach the ring from the surface until the maximum force has been reached. The liquid layer attached to the ring must not separate from the ring. After completing the measurements, the ring shall be immersed below the surface again and the measurements repeated until a constant surface tension value is reached. The time from transferring the solution to the measurement vessel shall be recorded for each determination. Readings shall be taken at the maximum force required to detach the ring from the liquid surface.

**▼ B****2. DATA**

In order to calculate the surface tension, the value read in mN/m on the apparatus shall be first multiplied by the calibration factor  $\Phi_a$  or  $\Phi_b$  (depending on the calibration procedure used). This will yield a value which applies only approximately and therefore requires correction.

Harkins and Jordan (4) have empirically determined correction factors for surface-tension values measured by the ring method which are dependent on ring dimensions, the density of the liquid and its surface tension.

Since it is laborious to determine the correction factor for each individual measurement from the Harkins and Jordan tables, in order to calculate the surface tension for aqueous solutions the simplified procedure of reading the corrected surface-tension values directly from the table may be used. (Interpolation shall be used for readings ranging between the tabular values.)

*Table:*

**Correction of the measured surface tension**

Only for aqueous solutions,  $\rho = 1 \text{ g/cm}^3$

r	= 9,55 mm (average ring radius)
r	= 0,185 mm (ring wire radius)

Experimental Value (mN/m)	Corrected Value (mN/m)	
	Weight calibration (see 1.6.4.2.2(a))	Water calibration (see 1.6.4.2.2(b))
20	16,9	18,1
22	18,7	20,1
24	20,6	22,1
26	22,4	24,1
28	24,3	26,1
30	26,2	28,1
32	28,1	30,1
34	29,9	32,1
36	31,8	34,1
38	33,7	36,1
40	35,6	38,2
42	37,6	40,3
44	39,5	42,3
46	41,4	44,4
48	43,4	46,5
50	45,3	48,6
52	47,3	50,7
54	49,3	52,8
56	51,2	54,9

**▼ B**

Experimental Value (mN/m)	Corrected Value (mN/m)	
	Weight calibration (see 1.6.4.2.2(a))	Water calibration (see 1.6.4.2.2(b))
58	53,2	57,0
60	55,2	59,1
62	57,2	61,3
64	59,2	63,4
66	61,2	65,5
68	63,2	67,7
70	65,2	69,9
72	67,2	72,0
74	69,2	—
76	71,2	—
78	73,2	—

This table has been compiled on the basis of the Harkins-Jordan correction. It is similar to that in the DIN Standard (DIN 53914) for water and aqueous solutions (density  $\rho = 1 \text{ g/cm}^3$  and is for a commercially available ring having the dimensions  $R = 9,55 \text{ mm}$  (mean ring radius) and  $r = 0,185 \text{ mm}$  (ring wire radius). The table provides corrected values for surface-tension measurements taken after calibration with weights or calibration with water.

Alternatively, without the preceding calibration, the surface tension call can be calculated according to the following formula:

$$\sigma = \frac{f \times F}{4\pi R}$$

where:

F = the force measured on the dynamometer at the breakpoint of the film

R = the radius of the ring

f = the correction factor (1)

### 3. REPORTING

#### 3.1. TEST REPORT

The test report shall, if possible, include the following information:

- method used,
- type of water or solution used,
- precise specification of the substance (identity and impurities),
- measurement results: surface tension (reading) stating both the individual readings and their arithmetic mean as well as the corrected mean (taking into consideration the equipment factor and the correction table),

**▼ B**

- concentration of the solution,
- test temperature,
- age of solution used; in particular the time between preparation and measurement of the solution,
- description of time dependence of surface tension after transferring the solution to the measurement vessel,
- all information and remarks relevant for the interpretation of results have to be reported, especially with regard to impurities and physical state of the substance.

**3.2. INTERPRETATION OF RESULTS**

Considering that distilled water has a surface tension of 72,75 mN/m at 20 °C, substances showing a surface tension lower than 60 mN/m under the conditions of this method should be regarded as being surface-active materials.

**4. REFERENCES**

- (1) OECD, Paris, 1981, Test Guideline 115, Decision of the Council C(81) 30 final.
- (2) R. Weissberger ed.: *Technique of Organic Chemistry, Physical Methods of Organic Chemistry*, 3rd ed., Interscience Publ., New York, 1959, vol. I, Part I, Chapter XIV.
- (3) *Pure Appl. Chem.*, 1976, vol. 48, p. 511.
- (4) Harkins, W.D., Jordan, H.F., *J. Amer. Chem. Soc.*, 1930, vol. 52, p. 1751.

**▼M4****A.6. WATER SOLUBILITY****INTRODUCTION**

1. This Test Method is equivalent to OECD Test Guideline (TG) 105 (1995). This Test Method is a revised version of the original TG 105 which was adopted in 1981. There is no difference of substance between the current version and that from 1981. Mainly the format has been changed. The revision was based on the EU Test Method 'Water Solubility' <sup>(1)</sup>.

**INITIAL CONSIDERATIONS**

2. The water solubility of a substance can be considerably affected by the presence of impurities. This Test Method addresses the determination of the solubility in water of essentially pure substances which are stable in water and not volatile. Before determining water solubility, it is useful to have some preliminary information on the test substance, like structural formula, vapour pressure, dissociation constant and hydrolysis as a function of pH.
3. Two methods, the column elution method and the flask method which cover respectively solubilities below and above  $10^{-2}$  g/l are described in this Test Method. A simple preliminary test is also described. It allows the determination of approximately the appropriate amount of sample to be used in the final test, as well as the time necessary to achieve saturation.

**DEFINITIONS AND UNITS**

4. The water solubility of a substance is the saturation mass concentration of the substance in water at a given temperature.
5. Water solubility is expressed in mass of solute per volume of solution. The SI unit is  $\text{kg/m}^3$  but g/l may also be used.

**REFERENCE CHEMICALS**

6. Reference chemicals do not need to be employed when investigating a test substance.

**DESCRIPTION OF THE METHODS****Test conditions**

7. The test is preferably run at  $20 \pm 0,5$  ° C. The chosen temperature should be kept constant in all relevant parts of the equipment.

**Preliminary test**

8. In a stepwise procedure, increasing volumes of water are added at room temperature to approximately 0,1 g of the sample (solid test substances must be pulverized) in a 10 ml glass-stoppered measuring cylinder. After each addition of an amount of water, the mixture is shaken for 10 minutes and is visually checked for any undissolved parts of the sample. If, after addition of 10 ml of water, the sample or parts of it remain undissolved, the experiment is continued in a 100 ml measuring cylinder. The approximate solubility is given in Table 1 below under that volume of water in which



**▼M4**

complete dissolution of the sample occurs. When the solubility is low, a long time may be required to dissolve a test substance and at least 24 hours should be allowed. If, after 24 hours, the test substance is still not dissolved, more time (up to 96 hours) should be allowed or further dilution should be attempted to ascertain whether the column elution method or flask method should be used.

*Table 1*

ml of water for 0,1 g soluble	0,1	0,5	1	2	10	100	> 100
approximate solubility in g/l	> 1 000	1 000 to 200	200 to 100	100 to 50	50 to 10	10 to 1	< 1

**Column elution method***Principle*

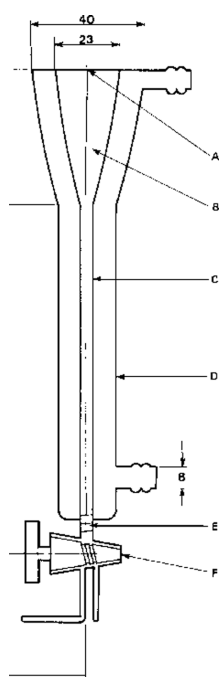
- This method is based on the elution of a test substance with water from a micro-column which is charged with an inert support material, previously coated with an excess of the test substance (2). The water solubility is given by the mass concentration of the eluate when this has reached a plateau as a function of time.

*Apparatus*

- The apparatus consists of a microcolumn (Figure 1), maintained at constant temperature. It is connected either to a recirculating pump (Figure 2) or to a levelling vessel (Figure 3). The microcolumn contains an inert support held in place by a small plug of glasswool which also serves to filter out particles. Possible materials which can be employed for the support are glass beads, diatomaceous earth, or other inert materials.
- The microcolumn shown in Figure 1 is suitable for the set-up with recirculating pump. It has a head space providing for five bed volumes (discarded at the start of the experiment) and the volume of five samples (withdrawn for analysis during the experiment). Alternatively, the size can be reduced if water can be added to the system during the experiment to replace the initial five bed volumes removed with impurities. The column is connected with tubing made of an inert material to the recirculating pump, capable of delivering approximately 25 ml/h. The recirculating pump can be, for example, a peristaltic or membrane pump. Care must be taken that no contamination and/or adsorption occur with the tube material.
- A schematic arrangement using a levelling vessel is shown in Figure 3. In this arrangement the microcolumn is fitted with a one way stopcock. The connection to the levelling vessel consists of a ground glass joint and tubing made of an inert material. The flow rate from the levelling vessel should be approximately 25 ml/h.

▼ M4

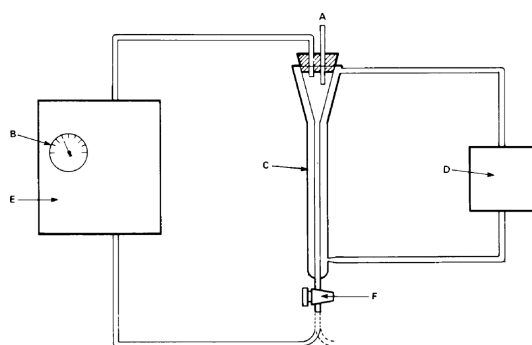
Figure 1



Dimensions in mm

- A. Connection for ground glass joint
- B. Headspace
- C. Interior 5
- D. Exterior 19
- E. Plug of glass wool
- F. Stopcock

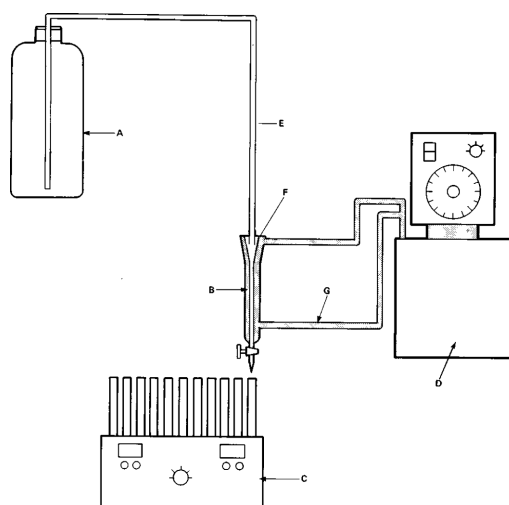
Figure 2



- A. Atmospheric equilibration
- B. Flowmeter
- C. Microcolumn
- D. Thermostatically controlled circulating pump
- E. Recirculating pump
- F. Two-way valve for sampling

▼ **M4**

Figure 3



- A. Levelling vessel (e.g. 2,5 litres chemical flask)
  - B. Column
  - C. Fraction accumulator
  - D. Thermostat
  - E. Teflon tubing
  - F. Ground glass joint
  - G. Water line (between thermostat and column, inner diameter approximately 8 mm)
13. Approximately 600 mg of support material is transferred to a 50 ml round-bottom flask. A suitable amount of test substance is dissolved in a volatile solvent of analytical reagent quality and an appropriate amount of this solution is added to the support material. The solvent is completely evaporated, e.g. using a rotary evaporator, as otherwise water saturation of the support will not be achieved during the elution step because of partitioning on the surface. The loaded support material is soaked for two hours in approximately 5 ml of water and the suspension is poured into the micro-column. Alternatively, dry loaded support material may be poured into the water-filled microcolumn and two hours are allowed for equilibrating.
14. The loading of the support material may cause problems, leading to erroneous results, e.g. when the test substance is deposited as an oil. These problems should be examined and the details reported.

*Procedure using a recirculating pump*

15. The flow through the column is started. It is recommended that a flow rate of approximately 25 ml/h, corresponding to 10 bed volumes per hour for the column described, be used. At least the first five bed volumes are discarded to remove water soluble impurities. Following this, the pump is allowed to run until equilibrium is established, as defined by five successive samples whose concentrations do not differ by more than  $\pm 30\%$  in a random fashion. These samples should be separated from each other by time intervals corresponding to the passage of at least ten bed volumes. Depending on the analytical method used, it may be preferable to establish a concentration/time curve to show that equilibrium is reached.

**▼ M4***Procedure using a levelling vessel*

16. Successive eluate fractions should be collected and analysed by the chosen method. Fractions from the middle eluate range, where the concentrations are constant within  $\pm 30\%$  in at least five consecutive fractions, are used to determine the solubility.
17. Double distilled water is the preferred eluent. Deionized water with a resistivity above 10 megohms/cm and total organic carbon content below 0,01 % can also be used.
18. Under both procedures, a second run is performed at half the flow rate of the first. If the results of the two runs are in agreement, the test is satisfactory. If the measured solubility is higher with the lower flow rate, then the halving of the flow rate must continue until two successive runs give the same solubility.
19. Under both procedures, the fractions should be checked for the presence of colloidal matter by examination of the Tyndall effect. The presence of particles invalidates the test and the test should be repeated after improvement of the filtering action of the column.
20. The pH of each sample should be measured, preferably by using special indicator strips.

**Flask method***Principle*

21. The test substance (solids must be pulverized) is dissolved in water at a temperature somewhat above the test temperature. When saturation is achieved, the mixture is cooled and kept at the test temperature. Alternatively, and if it is assured by appropriate sampling that the saturation equilibrium is reached, the measurement can be performed directly at the test temperature. Subsequently, the mass concentration of the test substance in the aqueous solution, which must not contain any undissolved particles, is determined by a suitable analytical method (3).

*Apparatus*

22. The following materials are needed:
  - normal laboratory glassware and instrumentation;
  - a device for the agitation of solutions under controlled constant temperature;
  - if required for emulsions, a centrifuge (preferably thermostated); and
  - analytical equipment.

*Procedure*

23. The quantity of test substance necessary to saturate the desired volume of water is estimated from the preliminary test. About five times that quantity is weighed into each of three glass vessels fitted with glass stoppers (e.g. centrifuge tubes, flasks). A volume of water, chosen in function of the analytical method and solubility range, is added to each vessel. The vessels are tightly stoppered and then agitated at 30 °C. A shaking or stirring device capable of operating at constant temperature should be used, e.g. magnetic stirring in a thermostated water bath. After one day, one of the vessels is equilibrated for 24 hours at the test temperature with occasional shaking. The contents of the vessel are then centrifuged at the test temperature and the concentration of the test substance in the clear aqueous phase is determined by a suitable analytical method. The other

**▼ M4**

two flasks are treated similarly after initial equilibration at 30 °C for two and three days respectively. If the concentrations measured in at least the two last vessels do not differ by more than 15 %, the test is satisfactory. If the results from vessels 1, 2 and 3 show a tendency of increasing values, the whole test should be repeated using longer equilibration times.

24. The test can also be performed without pre-incubation at 30 °C. In order to estimate the rate of establishment of the saturation equilibrium, samples are taken until the stirring time no longer influences the concentrations measured.
25. The pH of each sample should be measured, preferably by using special indicator strips.

**Analytical determinations**

26. A substance-specific method is preferred since small amounts of soluble impurities can cause large errors in the measured solubility. Examples of such methods are: gas or liquid chromatography, titration, photometry, voltammetry.

**DATA AND REPORTING****Data***Column elution method*

27. For each run, the mean value and standard deviation from at least five consecutive samples taken from the saturation plateau should be calculated. The mean values obtained from two tests with different flows should not differ by more than 30 %.

*Flask method*

28. The individual results from each of the three flasks, which should not differ by more than 15 %, are averaged.

**Test Report***Column elution method*

29. The test report must include the following information:
  - the results of the preliminary test
  - chemical identity and impurities (preliminary purification step, if any)
  - the concentrations, flow rates and pH for each sample
  - the means and standard deviations from at least five samples from the saturation plateau of each run
  - the average of at least two successive runs
  - the temperature of the water during the saturation process
  - the method of analysis
  - the nature of the support material
  - loading of the support material
  - solvent used
  - evidence of any chemical instability of the substance during the test
  - all information relevant for the interpretation of the results, in particular with regard to impurities and physical state of the test substance.

**▼ M4***Flask method*

30. The test report must include the following information:
- the results of the preliminary test
  - chemical identity and impurities (preliminary purification step, if any)
  - the individual analytical determinations and the average where more than one value was determined for each flask
  - the pH of each sample
  - the average of the values for different flasks which were in agreement
  - the test temperature
  - the analytical method
  - evidence of any chemical instability of the substance during the test
  - all information relevant for the interpretation of the results, in particular with regard to impurities and physical state of the test substance.

*LITERATURE:*

- (1) Commission Directive 92/69/EEC of 31 July 1992 adapting to technical progress for the seventeenth time Council Directive 67/548/EEC on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances (OJ L 383, 29.12.1992, p. 113).
- (2) NF T 20-045 (AFNOR) (September 1985). Chemical products for industrial use — Determination of water solubility of solids and liquids with low solubility — Column elution method.
- (3) NF T 20-046 (AFNOR) (September 1985). Chemical products for industrial use — Determination of water solubility of solids and liquids with high solubility — Flask method.

**▼B****A.8. PARTITION COEFFICIENT****1. METHOD**

The 'shake flask' method described is based on the OECD Test Guideline (1).

**1.1. INTRODUCTION**

It is useful to have preliminary information on structural formula, dissociation constant, water solubility, hydrolysis, n-octanol solubility and surface tension of the substance to perform this test.

Measurements should be made on ionisable substances only in their non-ionised form (free acid or free base) produced by the use of an appropriate buffer with a pH of at least one pH unit below (free acid) or above (free base) the pK.

This test method includes two separate procedures: the shake flask method and high performance liquid chromatography (HPLC). The former is applicable when the log  $P_{ow}$  value (see below for definitions) falls within the range - 2 to 4 and the latter within the range 0 to 6. Before carrying out either of the experimental procedures a preliminary estimate of the partition coefficient should first be obtained.

The shake-flask method applies only to essentially pure substances soluble in water and n-octanol. It is not applicable to surface active materials (for which a calculated value or an estimate based on the individual n-octanol and water solubilities should be provided).

The HPLC method is not applicable to strong acids and bases, metal complexes, surface-active materials or substances which react with the eluent. For these materials, a calculated value or an estimate based on individual n-octanol and water solubilities should be provided.

The HPLC method is less sensitive to the presence of impurities in the test compound than is the shake-flask method. Nevertheless, in some cases impurities can make the interpretation of the results difficult because peak assignment becomes uncertain. For mixtures which give an unresolved band, upper and lower limits of log P should be stated.

**1.2. DEFINITION AND UNITS**

The partition coefficient (P) is defined as the ratio of the equilibrium concentrations ( $c_i$ ) of a dissolved substance in a two-phase system consisting of two largely immiscible solvents. In the case n-octanol and water:

$$P_{ow} = \frac{c_{n\text{-octanol}}}{c_{\text{water}}}$$

The partition coefficient (P) therefore is the quotient of two concentrations and is usually given in the form of its logarithm to base 10 (log P).

**▼ B**

## 1.3. REFERENCE SUBSTANCES

*Shake-flask method*

Reference substances do not need to be employed in all cases when investigating a new substance. They should primarily serve to check the performance of the method from time to time and to allow comparison with results from other methods.

*HPLC method*

In order to correlate the measured HPLC data of a compound with its P value, a calibration graph of log P versus chromatographic data using at least six reference points has to be established. It is for the user to select the appropriate reference substances. Whenever possible, at least one reference compound should have a  $P_{ow}$  above that of the test substance, and another a  $P_{ow}$  below that of the test substance. For log P values less than 4, the calibration can be based on data obtained by the shake-flask method. For log P values greater than 4, the calibration can be based on validated literature values if these are in agreement with calculated values. For better accuracy, it is preferable to choose reference compounds which are structurally related to the test substance.

Extensive lists of values of log  $P_{ow}$  for many groups of chemicals are available (2)(3). If data on the partition coefficients of structurally related compounds are not available, then a more general calibration, established with other reference compounds, may be used.

A list of recommended reference substances and their  $P_{ow}$  values is given in Appendix 2.

## 1.4. PRINCIPLE OF THE METHOD

1.4.1. **Shake-flask method**

In order to determine a partition coefficient, equilibrium between all interacting components of the system must be achieved, and the concentrations of the substances dissolved in the two phases must be determined. A study of the literature on this subject indicates that several different techniques can be used to solve this problem, i.e. the thorough mixing of the two phases followed by their separation in order to determine the equilibrium concentration for the substance being examined.

1.4.2. **HPLC method**

HPLC is performed on analytical columns packed with a commercially available solid phase containing long hydrocarbon chains (e.g.  $C_8$ ,  $C_{18}$ ) chemically bound onto silica. Chemicals injected onto such a column move along it at different rates because of the different degrees of partitioning between the mobile phase and the hydrocarbon stationary phase. Mixtures of chemicals are eluted in order of their hydrophobicity, with water-soluble chemicals eluted first and oil-soluble chemicals last, in proportion to their hydrocarbon-water partition coefficient. This enables the relationship between the retention time on such a (reverse phase) column and the n-octanol/water partition coefficient to be established. The partition coefficient is deduced from the *capacity factor* k, given by the expression:



**▼ B**

$$k = \frac{t_r - t_0}{t_0}$$

in which,  $t_r$  = retention time of the test substance, and  $t_0$  = average time a solvent molecule needs to pass through the column (dead-time).

Quantitative analytical methods are not required and only the determination of elution times is necessary.

## 1.5. QUALITY CRITERIA

### 1.5.1. Repeatability

#### *Shake-flask method*

In order to assure the accuracy of the partition coefficient, duplicate determinations are to be made under three different test conditions, whereby the quantity of substance specified as well as the ratio of the solvent volumes may be varied. The determined values of the partition coefficient expressed as their common logarithms should fall within a range of  $\pm 0,3$  log units.

#### *HPLC method*

In order to increase the confidence in the measurement, duplicate determinations must be made. The values of log P derived from individual measurements should fall within a range of  $\pm 0,1$  log units.

### 1.5.2. Sensitivity

#### *Shake-flask method*

The measuring range of the method is determined by the limit of detection of the analytical procedure. This should permit the assessment of values of log  $P_{ow}$  in the range of - 2 to 4 (occasionally when conditions apply, this range may be extended to log  $P_{ow}$  up to 5) when the concentration of the solute in either phase is not more than 0,01 mol per litre.

#### *HPLC method*

The HPLC method enables partition coefficients to be estimated in the log  $P_{ow}$  range 0 to 6.

Normally, the partition coefficient of a compound can be estimated to within  $\pm 1$  log unit of the shake-flask value. Typical correlations can be found in the literature (4)(5)(6)(7)(8). Higher accuracy can usually be achieved when correlation plots are based on structurally-related reference compounds (9).

**▼ B**1.5.3. **Specificity***Shake-flask method*

The Nernst Partition Law applies only at constant temperature, pressure and pH for dilute solutions. It strictly applies to a pure substance dispersed between two pure solvents. If several different solutes occur in one or both phases at the same time, this may affect the results.

Dissociation or association of the dissolved molecules result in deviations from the Nernst Partition Law. Such deviations are indicated by the fact that the partition coefficient becomes dependent upon the concentration of the solution.

Because of the multiple equilibria involved, this test method should not be applied to ionisable compounds without applying a correction. The use of buffer solutions in place of water should be considered for such compounds; the pH of the buffer should be at least 1 pH unit from the pKa of the substance and bearing in mind the relevance of this pH for the environment.

## 1.6. DESCRIPTION OF THE METHOD

1.6.1. **Preliminary estimate of the partition coefficient**

The partition coefficient is estimated preferably by using a calculation method (see Appendix 1), or where appropriate, from the ratio of the solubilities of the test substance in the pure solvents (10).

1.6.2. **Shake-flask method**1.6.2.1. *Preparation*

n-Octanol: the determination of the partition coefficient should be carried out with high purity analytical grade reagent.

Water: water distilled or double distilled in glass or quartz apparatus should be employed. For ionisable compounds, buffer solutions in place of water should be used if justified.

*Note:*

Water taken directly from an ion exchanger should not be used.

1.6.2.1.1. **Pre-saturation of the solvents**

Before a partition coefficient is determined, the phases of the solvent system are mutually saturated by shaking at the temperature of the experiment. To do this, it is practical to shake two large stock bottles of high purity analytical grade n-octanol or water each with a sufficient quantity of the other solvent for 24 hours on a mechanical shaker and then to let them stand long enough to allow the phases to separate and to achieve a saturation state.

**▼B**

## 1.6.2.1.2. Preparation for the test

The entire volume of the two-phase system should nearly fill the test vessel. This will help prevent loss of material due to volatilisation. The volume ratio and quantities of substance to be used are fixed by the following:

- the preliminary assessment of the partition coefficient (see above),
- the minimum quantity of test substance required for the analytical procedure, and
- the limitation of a maximum concentration in either phase of 0,01 mol per litre.

Three tests are carried out. In the first, the calculated volume ratio of n-octanol to water is used; in the second, this ratio is divided by two; and in the third, this ratio is multiplied by two (e.g. 1:1, 1:2, 2:1).

## 1.6.2.1.3. Test substance

A stock solution is prepared in n-octanol pre-saturated with water. The concentration of this stock solution should be precisely determined before it is employed in the determination of the partition coefficient. This solution should be stored under conditions which ensure its stability.

## 1.6.2.2. Test conditions

The test temperature should be kept constant ( $\pm 1$  °C) and lie in the range of 20 to 25 °C.

## 1.6.2.3. Measurement procedure

## 1.6.2.3.1. Establishment of the partition equilibrium

Duplicate test vessels containing the required, accurately measured amounts of the two solvents together with the necessary quantity of the stock solution should be prepared for each of the test conditions.

The n-octanol phases should be measured by volume. The test vessels should either be placed in a suitable shaker or shaken by hand. When using a centrifuge tube, a recommended method is to rotate the tube quickly through 180° about its transverse axis so that any trapped air rises through the two phases. Experience has shown that 50 such rotations are usually sufficient for the establishment of the partition equilibrium. To be certain, 100 rotations in five minutes are recommended.

## 1.6.2.3.2. Phase separation

When necessary, in order to separate the phases, centrifugation of the mixture should be carried out. This should be done in a laboratory centrifuge maintained at room temperature, or, if a non-temperature controlled centrifuge is used, the centrifuge tubes should be kept for equilibration at the test temperature for at least one hour before analysis.

**▼ B**1.6.2.4. *Analysis*

For the determination of the partition coefficient, it is necessary to determine the concentrations of the test substance in both phases. This may be done by taking an aliquot of each of the two phases from each tube for each test condition and analyzing them by the chosen procedure. The total quantity of substance present in both phases should be calculated and compared with the quantity of the substance originally introduced.

The aqueous phase should be sampled by a procedure that minimises the risk of including traces of n-octanol: a glass syringe with a removable needle can be used to sample the water phase. The syringe should initially be partially filled with air. Air should be gently expelled while inserting the needle through the n-octanol layer. An adequate volume of aqueous phase is withdrawn into the syringe. The syringe is quickly removed from the solution and the needle detached. The contents of the syringe may then be used as the aqueous sample. The concentration in the two separated phases should preferably be determined by a substance-specific method. Examples of analytical methods which may be appropriate are:

- photometric methods,
- gas chromatography,
- high-performance liquid chromatography.

1.6.3. **HPLC method**1.6.3.1. *Preparation**Apparatus*

A liquid chromatograph, fitted with a pulse-free pump and a suitable detection device, is required. The use of an injection valve with injection loops is recommended. The presence of polar groups in the stationary phase may seriously impair the performance of the HPLC column. Therefore, stationary phases should have the minimal percentage of polar groups (11). Commercial microparticulate reverse-phase packings or ready-packed columns can be used. A guard column may be positioned between the injection system and the analytical column.

*Mobile phase*

HPLC grade methanol and HPLC grade water are used to prepare the eluting solvent, which is degassed before use. Isocratic elution should be employed. Methanol/water ratios with a minimum water content of 25 % should be used. Typically a 3:1 (v/v) methanol-water mixture is satisfactory for eluting compounds of log P 6 within an hour, at a flow rate of 1 ml/min. For compounds of high log P it may be necessary to shorten the elution time (and those of the reference compounds) by decreasing the polarity of the mobile phase or the column length.

Substances with very low solubility in n-octanol tend to give abnormally low log  $P_{ow}$  values with the HPLC method; the peaks of such compounds sometimes accompany the solvent front. This is probably due to the fact that the partitioning process is too slow to reach the equilibrium in the time normally taken by an HPLC separation. Decreasing the flow rate and/or lowering the methanol/water ratio may then be effective to arrive at a reliable value.

**▼ B**

Test and reference compounds should be soluble in the mobile phase in sufficient concentrations to allow their detection. Only in exceptional cases may additives be used with the methanol-water mixture, since additives will change the properties of the column. For chromatograms with additives it is mandatory to use a separate column of the same type. If methanol-water is not appropriate, other organic solvent-water mixtures can be used, e.g. ethanol-water or acetonitrile-water.

The pH of the eluent is critical for ionisable compounds. It should be within the operating pH range of the column, which is usually between 2 and 8. Buffering is recommended. Care must be taken to avoid salt precipitation and column deterioration which occur with some organic phase/buffer mixtures. HPLC measurements with silica-based stationary phases above pH 8 are not advisable since the use of an alkaline, mobile phase may cause rapid deterioration in the performance of the column.

*Solutes*

The reference compounds should be the purest available. Compounds to be used for test or calibration purposes are dissolved in the mobile phase if possible.

*Test conditions*

The temperature during the measurements should not vary by more than  $\pm 2$  K.

1.6.3.2. **Measurement***Calculation of dead time  $t_0$* 

The dead time can be determined by using either a homologous series (e.g. n-alkyl methyl ketones) or unretained organic compounds (e.g. thiourea or formamide). For calculating the dead time by using a homologous series, a set of at least seven members of a homologous series is injected and the respective retention times are determined. The raw retention times  $t_{r(n_c + 1)}$  are plotted as a function of  $t_{r(n_c)}$  and the intercept  $a$  and slope  $b$  of the regression equation:

$$t_{r(n_c + 1)} = a + b t_{r(n_c)}$$

are determined ( $n_c$  = number of carbon atoms). The dead time is then given by:

$$t_0 = a/(1 - b)$$

**▼ B***Calibration graph*

The next step is to construct a correlation plot of log k values versus log p for appropriate reference compounds. In practice, a set of between 5 and 10 standard reference compounds whose log p is around the expected range are injected simultaneously and the retention times are determined, preferably on a recording integrator linked to the detection system. The corresponding logarithms of the capacity factors, log k, are calculated and plotted as a function of the log p determined by the shake-flask method. The calibration is performed at regular intervals, at least once daily, so that possible changes in column performance can be allowed for.

*Determination of the capacity factor of the test substance*

The test substance is injected in as small a quantity of mobile phase as possible. The retention time is determined (in duplicate), permitting the calculation of the capacity factor k. From the correlation graph of the reference compounds, the partition coefficient of the test substance can be interpolated. For very low and very high partition coefficients, extrapolation is necessary. In those cases particular care has to be taken of the confidence limits of the regression line.

**2. DATA***Shake-flask method*

The reliability of the determined values of P can be tested by comparison of the means of the duplicate determinations with the overall mean.

**3. REPORTING**

The test report shall, if possible, include the following information:

- precise specification of the substance (identity and impurities),
- when the methods are not applicable (e.g. surface active material), a calculated value or an estimate based on the individual n-octanol and water solubilities should be provided,
- all information and remarks relevant for the interpretation of results, especially with regard to impurities and physical state of the substance.

*For shake-flask method:*

- the result of the preliminary estimation, if any,
- temperature of the determination,
- data on the analytical procedures used in determining concentrations,
- time and speed of centrifugation, if used,

**▼ B**

- the measured concentrations in both phases for each determination (this means that a total of 12 concentrations will be reported),
- the weight of the test substance, the volume of each phase employed in each test vessel and the total calculated amount of test substance present in each phase after equilibration,
- the calculated values of the partition coefficient (P) and the mean should be reported for each set of test conditions as should the mean for all determinations. If there is a suggestion of concentration dependency of the partition coefficient, this should be noted in the report,
- the standard deviation of individual P values about their mean should be reported,
- the mean P from all determinations should also be expressed as its logarithm (base 10),
- the calculated theoretical  $P_{ow}$  when this value has been determined or when the measured value is  $> 10^4$ ,
- pH of water used and of the aqueous phase during the experiment,
- if buffers are used, justification for the use of buffers in place of water, composition, concentration and pH of the buffers, pH of the aqueous phase before and after the experiment.

*For HPLC method:*

- the result of the preliminary estimation, if any,
- test and reference substances, and their purity,
- temperature range of the determinations,
- pH at which the determinations are made,
- details of the analytical and guard column, mobile phase and means of detection,
- retention data and literature log P values for reference compounds used in calibration,
- details of fitted regression line (log k versus log P),
- average retention data and interpolated log P value for the test compound,
- description of equipment and operating conditions,
- elution profiles,
- quantities of test and reference substances introduced in the column,
- dead-time and how it was measured.

**▼B**

4.

**REFERENCES**

- (1) OECD, Paris, 1981, Test Guideline 107, Decision of the Council C(81) 30 final.
- (2) C. Hansch and A.J. Leo, Substituent Constants for Correlation Analysis in Chemistry and Biology, John Wiley, New York, 1979.
- (3) Log P and Parameter Database, A tool for the quantitative prediction of bioactivity (C. Hansch, chairman, A.J. Leo, dir.) — Available from Pomona College Medical Chemistry Project 1982, Pomona College, Claremont, California 91711.
- (4) L. Renberg, G. Sundström and K. Sundh-Nygård, Chemosphere, 1980, vol. 80, p. 683.
- (5) H. Ellgehausen, C. D'Hondt and R. Fuerer, Pestic. Sci., 1981, vol. 12, p. 219.
- (6) B. McDuffie, Chemosphere, 1981, vol. 10, p. 73.
- (7) W.E. Hammers et al., J. Chromatogr., 1982, vol. 247, p. 1.
- (8) J.E. Haky and A.M. Young, J. Liq. Chromat., 1984, vol. 7, p. 675.
- (9) S. Fujisawa and E. Masuhara, J. Biomed. Mat. Res., 1981, vol. 15, p. 787.
- (10) O. Jubermann, Verteilen und Extrahieren, in Methoden der Organischen Chemie (Houben Weyl), Allgemeine Laboratoriumspraxis (edited by E. Muller), Georg Thieme Verlag, Stuttgart, 1958, Band I/1, p. 223-339.
- (11) R.F. Rekker and H.M. de Kort, Euro. J. Med. Chem., 1979, vol. 14, p. 479.
- (12) A. Leo, C. Hansch and D. Elkins, Partition coefficients and their uses. Chem. Rev., 1971, vol. 71, p. 525.
- (13) R.F. Rekker, The Hydrophobic Fragmental Constant, Elsevier, Amsterdam, 1977.
- (14) NF T 20-043 AFNOR (1985). Chemical products for industrial use — Determination of partition coefficient — Flask shaking method.
- (15) C.V. Eadsforth and P. Moser, Chemosphere, 1983, vol. 12, p. 1459.
- (16) A. Leo, C. Hansch and D. Elkins, Chem. Rev., 1971, vol. 71, p. 525.
- (17) C. Hansch, A. Leo, S.H. Unger, K.H. Kim, D. Nikaitani and E.J. Lien, J. Med. Chem., 1973, vol. 16, p. 1207.
- (18) W.B. Neely, D.R. Branson and G.E. Blau, Environ. Sci. Technol., 1974, vol. 8, p. 1113.
- (19) D.S. Brown and E.W. Flagg, J. Environ. Qual., 1981, vol. 10, p. 382.
- (20) J.K. Seydel and K.J. Schaper, Chemische Struktur und biologische Aktivität von Wirkstoffen, Verlag Chemie, Weinheim, New York, 1979.



**▼B**

- (21) R. Franke, *Theoretical Drug Design Methods*, Elsevier, Amsterdam, 1984.
- (22) Y.C. Martin, *Quantitative Drug Design*, Marcel Dekker, New York, Basel, 1978.
- (23) N.S. Nirrlees, S.J. Noulton, C.T. Murphy, P.J. Taylor; *J. Med. Chem.*, 1976, vol. 19, p. 615.

**▼ B***Appendix 1***Calculation/estimation methods****INTRODUCTION**

A general introduction to calculation methods, data and examples are provided in the Handbook of Chemical Property Estimation Methods (a).

Calculated values of  $P_{ow}$  can be used:

- for deciding which of the experimental methods is appropriate (shake-flask range:  $\log P_{ow}$ : - 2 to 4, HPLC range:  $\log P_{ow}$ : 0 to 6),
- for selecting the appropriate test conditions (e.g. reference substances for HPLC procedures, volume ratio n-octanol/water for shake flask method),
- as a laboratory internal check on possible experimental errors,
- for providing a  $P_{ow}$ -estimate in cases where the experimental methods cannot be applied for technical reasons.

**ESTIMATION METHOD***Preliminary estimate of the partition coefficient*

The value of the partition coefficient can be estimated by the use of the solubilities of the test substance in the pure solvents: For this:

$$P_{\text{estimate}} = \frac{\text{saturation } c_{\text{n-octanol}}}{\text{saturation } c_{\text{water}}}$$

**CALCULATION METHODS***Principle of the calculation methods*

All calculation methods are based on the formal fragmentation of the molecule into suitable substructures for which reliable  $\log P_{ow}$ -increments are known. The  $\log P_{ow}$  of the whole molecule is then calculated as the sum of its corresponding fragment values plus the sum of correction terms for intramolecular interactions.

Lists of fragment constants and correction terms are available (b)(c)(d)(e); Some are regularly updated (b).

**Quality criteria**

In general, the reliability of the calculation method decreases with increasing complexity of the compound under study. In the case of simple molecules with low molecular weight and one or two functional groups, a deviation of 0,1 to 0,3  $\log P_{ow}$  units between the results of the different fragmentation methods and the measured value can be expected. In the case of more complex molecules the margin of error can be greater. This will depend on the reliability and availability of fragment constants, as well as on the ability to recognise intramolecular interactions (e.g. hydrogen bonds) and the correct use of the correction terms (less of a problem with the computer software CLOGP-3) (b). In the case of ionising compounds the correct consideration of the charge or degree of ionisation is important.

**▼ B****Calculation procedures***Hansch  $\pi$ -method*

The original hydrophobic substituent constant,  $\pi$ , introduced by Fujira et al. (f) is defined as:

$$\pi_x = \log P_{ow}(\text{PhX}) - \log P_{ow}(\text{PhH})$$

where  $P_{ow}(\text{PhX})$  is the partition coefficient of an aromatic derivative and  $P_{ow}(\text{PhH})$  that of the parent compound

$$\text{(e.g. } \pi_{\text{Cl}} = \log P_{ow}(\text{C}_6\text{H}_5\text{Cl}) - \log P_{ow}(\text{C}_6\text{H}_6) = 2,84 - 2,13 = 0,71\text{)}.$$

According to its definition the  $\pi$ -method is applicable predominantly for aromatic substitution.  $\pi$ -values for a large number of substituents have been tabulated (b)(c)(d). They are used for the calculation of  $\log P_{ow}$  for aromatic molecules or substructures.

*Rekker method*

According to Rekker (g) the  $\log P_{ow}$  value is calculated as follows:

$$\log P_{ow} = \sum_i a_i f_i + \sum_j (\text{interaction terms})$$

where  $f_i$  represents the different molecular fragment constants and  $a_i$  the frequency of their occurrence in the molecule under investigation. The correction terms can be expressed as an integral multiple of one single constant  $C_m$  (so-called magic constant). The fragment constants  $f_i$  and  $C_m$  were determined from a list of 1 054 experimental  $P_{ow}$  values (825 compounds) using multiple regression analysis (c)(h). The determination of the interaction terms is carried out according to set rules described in the literature (e)(h)(i).

*Hansch-Leo method*

According to Hansch and Leo (c), the  $\log P_{ow}$  value is calculated from:

$$\log P_{ow} = \sum_i a_i f_i + \sum_j b_j F_j$$

where  $f_i$  represents the different molecular fragment constants,  $F_j$  the correction terms and  $a_i$ ,  $b_j$  the corresponding frequencies of occurrence. Derived from experimental  $P_{ow}$  values, a list of atomic and group fragmental values and a list of correction terms  $F_j$  (so-called factors) were determined by trial and error. The correction terms have been ordered into several different classes (a)(c). It is relatively complicated and time consuming to take into account all the rules and correction terms. Software packages have been developed (b).

*Combined method*

The calculation of  $\log P_{ow}$  of complex molecules can be considerably improved, if the molecule is dissected into larger substructures for which reliable  $\log P_{ow}$  values are available, either from tables (b)(c) or from one's own measurements. Such fragments (e.g. heterocycles, anthraquinone, azobenzene) can then be combined with the Hansch  $\pi$ -values or with Rekker or Leo fragment constants.

**Remarks**

- (i) The calculation methods can only be applied to partly or fully ionised compounds when it is possible to take the necessary correction factors into account;

**▼B**

- (ii) if intramolecular hydrogen bonds can be assumed, the corresponding correction terms (approx. + 0,6 to + 1,0 log  $P_{ow}$  units) have to be added (a). Indications for the presence of such bonds can be obtained from stereo models or spectroscopic data of the molecule;
- (iii) If several tautomeric forms are possible, the most likely form should be used as the basis of the calculation;
- (iv) the revisions of lists of fragment constants should be followed carefully.

**Report**

When using calculation/estimation methods, the test report shall, if possible, include the following information:

- description of the substance (mixture, impurities, etc.),
- indication of any possible intramolecular hydrogen bonding, dissociation, charge and any other unusual effects (e.g. tautomerism),
- description of the calculation method,
- identification or supply of database,
- peculiarities in the choice of fragments,
- comprehensive documentation of the calculation.

**LITERATURE**

- (a) W.J. Lyman, W.F. Reehl and D.H. Rosenblatt (ed.), Handbook of Chemical Property Estimation Methods, McGraw-Hill, New York, 1983.
- (b) Pomona College, Medicinal Chemistry Project, Claremont, California 91711, USA, Log P Database and Med. Chem. Software (Program CLOGP-3).
- (c) C. Hansch, A.J. Leo, Substituent Constants for Correlation Analysis in Chemistry and Biology, John Wiley, New York, 1979.
- (d) A. Leo, C. Hansch, D. Elkins, Chem. Rev., 1971, vol. 71, p. 525.
- (e) R.F. Rekker, H.M. de Kort, Eur. J. Med. Chem. -Chill. Ther. 1979, vol. 14, p. 479.
- (f) T. Fujita, J. Iwasa and C. Hansch, J. Amer. Chem. Soc., 1964, vol. 86, p. 5175.
- (g) R.F. Rekker, The Hydrophobic Fragmental Constant, Pharmacology Library, Elsevier, New York, 1977, vol. 1.
- (h) C.V. Eadsforth, P. Moser, Chemosphere, 1983, vol. 12, p. 1459.
- (i) R.A. Scherrer, ACS, American Chemical Society, Washington D.C., 1984, Symposium Series 255, p. 225.

▼B

## Appendix 2

## Recommended Reference Substances for the HPLC Method

No	Reference Substance	log P <sub>ow</sub>	pKa
1	2-Butanone	0,3	
2	4-Acetylpyridine	0,5	
3	Aniline	0,9	
4	Acetanilide	1,0	
5	Benzylalcohol	1,1	
6	p-Methoxyphenol	1,3	pKa = 10,26
7	Phenoxy acetic acid	1,4	pKa = 3,12
8	Phenol	1,5	pKa = 9,92
9	2,4-Dinitrophenol	1,5	pKa = 3,96
10	Benzonitrile	1,6	
11	Phenylacetoneitrile	1,6	
12	4-Methylbenzyl alcohol	1,6	
13	Acetophenone	1,7	
14	2-Nitrophenol	1,8	pKa = 7,17
15	3-Nitrobenzoic acid	1,8	pKa = 3,47
16	4-Chloraniline	1,8	pKa = 4,15
17	Nitrobenzene	1,9	
18	Cinnamic alcohol	1,9	
19	Benzoic acid	1,9	pKa = 4,19
20	p-Cresol	1,9	pKa = 10,17
21	Cinnamic acid	2,1	pKa = 3,89 cis 4,44 trans
22	Anisole	2,1	
23	Methylbenzoate	2,1	
24	Benzene	2,1	
25	3-Methylbenzoic acid	2,4	pKa = 4,27
26	4-Chlorophenol	2,4	pKa = 9,1
27	Trichloroethylene	2,4	
28	Atrazine	2,6	
29	Ethylbenzoate	2,6	
30	2,6-Dichlorobenzonitrile	2,6	
31	3-Chlorobenzoic acid	2,7	pKa = 3,82
32	Toluene	2,7	
33	1-Naphthol	2,7	pKa = 9,34
34	2,3-Dichloroaniline	2,8	
35	Chlorobenzene	2,8	
36	Allyl-phenylether	2,9	
37	Bromobenzene	3,0	

**▼ B**

No	Reference Substance	log P <sub>ow</sub>	pKa
38	Ethylbenzene	3,2	
39	Benzophenone	3,2	
40	4-Phenylphenol	3,2	pKa = 9,54
41	Thymol	3,3	
42	1,4-Dichlorobenzene	3,4	
43	Diphenylamine	3,4	pKa = 0,79
44	Naphthalene	3,6	
45	Phenylbenzoate	3,6	
46	Isopropylbenzene	3,7	
47	2,4,6-Trichlorophenol	3,7	pKa = 6
48	Biphenyl	4,0	
49	Benzylbenzoate	4,0	
50	2,4-Dinitro-6 sec. butyophenol	4,1	
51	1,2,4-Trichlorobenzene	4,2	
52	Dodecanoic acid	4,2	
53	Diphenylether	4,2	
54	n-Butylbenzene	4,5	
55	Phenanthrene	4,5	
56	Fluoranthene	4,7	
57	Dibenzyl	4,8	
58	2,6-Diphenylpyridine	4,9	
59	Triphenylamine	5,7	
60	DDT	6,2	
Other reference substances of low log P <sub>ow</sub>			
1	Nicotinic acid	- 0,07	

**▼B****A.9. FLASH-POINT****1. METHOD****1.1. INTRODUCTION**

It is useful to have preliminary information on the flammability of the substance before performing this test. The test procedure is applicable to liquid substances whose vapours can be ignited by ignition sources. The test methods listed in this text are only reliable for flash-point ranges which are specified in the individual methods.

The possibility of chemical reactions between the substance and the sample holder should be considered when selecting the method to be used.

**1.2. DEFINITIONS AND UNITS**

The flash-point is the lowest temperature, corrected to a pressure of 101,325 kPa, at which a liquid evolves vapours, under the conditions defined in the test method, in such an amount that a flammable vapour/air mixture is produced in the test vessel.

Units: °C

$$t = T - 273,15$$

(t in °C and T in K)

**1.3. REFERENCE SUBSTANCES**

Reference substances do not need to be employed in all cases when investigating a new substance. They should primarily serve to check the performance of the method from time to time and to allow comparison with results from other methods.

**1.4. PRINCIPLE OF THE METHOD**

The substance is placed in a test vessel and heated or cooled to the test temperature according to the procedure described in the individual test method. Ignition trials are carried out in order to ascertain whether or not the sample flashed at the test temperature.

**1.5. QUALITY CRITERIA****1.5.1. Repeatability**

The repeatability varies according to flash-point range and the test method used; maximum 2 °C.

**1.5.2. Sensitivity**

The sensitivity depends on the test method used.

**1.5.3. Specificity**

The specificity of some test methods is limited to certain flash-point ranges and subject to substance-related data (e.g. high viscosity).

**▼ B**

## 1.6. DESCRIPTION OF THE METHOD

1.6.1. **Preparations**

A sample of the test substance is placed in a test apparatus according to 1.6.3.1 and/or 1.6.3.2.

For safety, it is recommended that a method utilising a small sample size, circa 2 cm<sup>3</sup>, be used for energetic or toxic substances.

1.6.2. **Test conditions**

The apparatus should, as far as is consistent with safety, be placed in a draught-free position.

1.6.3. **Performance of the test**1.6.3.1. *Equilibrium method*

See ISO 1516, ISO 3680, ISO 1523, ISO 3679.

1.6.3.2. *Non-equilibrium method*

*Abel apparatus:*

See BS 2000 part 170, NF M07-011, NF T66-009.

*Abel-Pensky apparatus:*

See EN 57, DIN 51755 part 1 (for temperatures from 5 to 65 °C), DIN 51755 part 2 (for temperatures below 5 °C), NF M07-036.

*Tag apparatus:*

See ASTM D 56.

*Pensky-Martens apparatus:*

See ISO 2719, EN 11, DIN 51758, ASTM D 93, BS 2000-34, NF M07-019.

*Remarks:*

When the flash-point, determined by a non-equilibrium method in 1.6.3.2, is found to be  $0 \pm 2$  °C,  $21 \pm 2$  °C or  $55 \pm 2$  °C, it should be confirmed by an equilibrium method using the same apparatus.

Only the methods which can give the temperature of the flash-point may be used for a notification.

To determine the flash-point of viscous liquids (paints, gums and similar) containing solvents, only apparatus and test methods suitable for determining the flash-point of viscous liquids may be used.

See ISO 3679, ISO 3680, ISO 1523, DIN 53213 part 1.



**▼B**2. **DATA**3. **REPORTING**

The test report shall, if possible, include the following information:

- the precise specification of the substance (identification and impurities),
- the method used should be stated as well as any possible deviations,
- the results and any additional remarks relevant for the interpretation of results.

4. **REFERENCES**

None.

**▼B****A.10. FLAMMABILITY (SOLIDS)****1. METHOD****1.1. INTRODUCTION**

It is useful to have preliminary information on potentially explosive properties of the substance before performing this test.

This test should only be applied to powdery, granular or paste-like substances.

In order not to include all substances which can be ignited but only those which burn rapidly or those whose burning behaviour is in any way especially dangerous, only substances whose burning velocity exceeds a certain limiting value are considered to be highly flammable.

It can be especially dangerous if incandescence propagates through a metal powder because of the difficulties in extinguishing a fire. Metal powders should be considered highly flammable if they support spread of incandescence throughout the mass within a specified time.

**1.2. DEFINITION AND UNITS**

Burning time expressed in seconds.

**1.3. REFERENCE SUBSTANCES**

Not specified.

**1.4. PRINCIPLE OF THE METHOD**

The substance is formed into an unbroken strip or powder train about 250 mm long and a preliminary screening test performed to determine if, on ignition by a gas flame, propagation by burning with flame or smouldering occurs. If propagation over 200 mm of the train occurs within a specified time then a full test programme to determine the burning rate is carried out.

**1.5. QUALITY CRITERIA**

Not stated.

**▼B**

## 1.6. DESCRIPTION OF METHOD

1.6.1. **Preliminary screening test**

The substance is formed into an unbroken strip or powder train about 250 mm long by 20 mm wide by 10 mm high on a non-combustible, non-porous and low heat-conducting base plate. A hot flame from a gas burner (minimum diameter 5 mm) is applied to one end of the powder train until the powder ignites or for a maximum of two minutes (five minutes for powders of metals or metal-alloys). It should be noted whether combustion propagates along 200 mm of the train within the 4 minutes test period (or 40 minutes for metal powders). If the substance does not ignite and propagate combustion either by burning with flame or smouldering along 200 mm of the powder train within the four minutes (or 40 minutes) test period, then the substance should not be considered as highly flammable and no further testing is required. If the substance propagates burning of a 200 mm length of the powder train in less than four minutes, or less than 40 minutes for metal powders, the procedure described below (point 1.6.2. and following) should be carried out.

1.6.2. **Burning rate test**1.6.2.1. *Preparation*

Powdery or granular substances are loosely filled into a mould 250 mm long with a triangular cross-section of inner height 10 mm and width 20 mm. On both sides of the mould in a longitudinal direction two metal plates are mounted as lateral limitations which project 2 mm beyond the upper edge of the triangular cross section (figure). The mould is then dropped three times from a height of 2 cm onto a solid surface. If necessary the mould is then filled up again. The lateral limitations are then removed and the excess substance scraped off. A non-combustible, non-porous and low heat-conducting base plate is placed on top of the mould, the apparatus inverted and the mould removed.

Paste-like substances are spread on a non-combustible, non-porous and low heat-conducting base plate in the form of a rope 250 mm in length with a cross section of about 1 cm<sup>2</sup>.

1.6.2.2. *Test conditions*

In the case a moisture-sensitive substance, the test should be carried out as quickly as possible after its removal from the container.

1.6.2.3. *Performance of the test*

Arrange the pile across the draught in a fume cupboard.

The air-speed should be sufficient to prevent fumes escaping into the laboratory and should not be varied during the test. A draught screen should be erected around the apparatus.

A hot flame from a gas burner (minimum diameter of 5 mm) is used to ignite the pile at one end. When the pile has burned a distance of 80 mm, the rate of burning over the next 100 mm is measured.

**▼B**

The test is performed six times, using a clean cool plate each time, unless a positive result is observed earlier.

**2. DATA**

The burning time from the preliminary screening test (1.6.1) and the shortest burning time in up to six tests (1.6.2.3) are relevant for evaluation.

**3. REPORTING****3.1. TEST REPORT**

The test report shall, if possible, include the following information:

- the precise specification of the substance (identification and impurities),
- a description of the substance to be tested, its physical state including moisture content,
- results from the preliminary screening test and from the burning rate test if performed,
- all additional remarks relevant to the interpretation of results.

**3.2. INTERPRETATION OF THE RESULT**

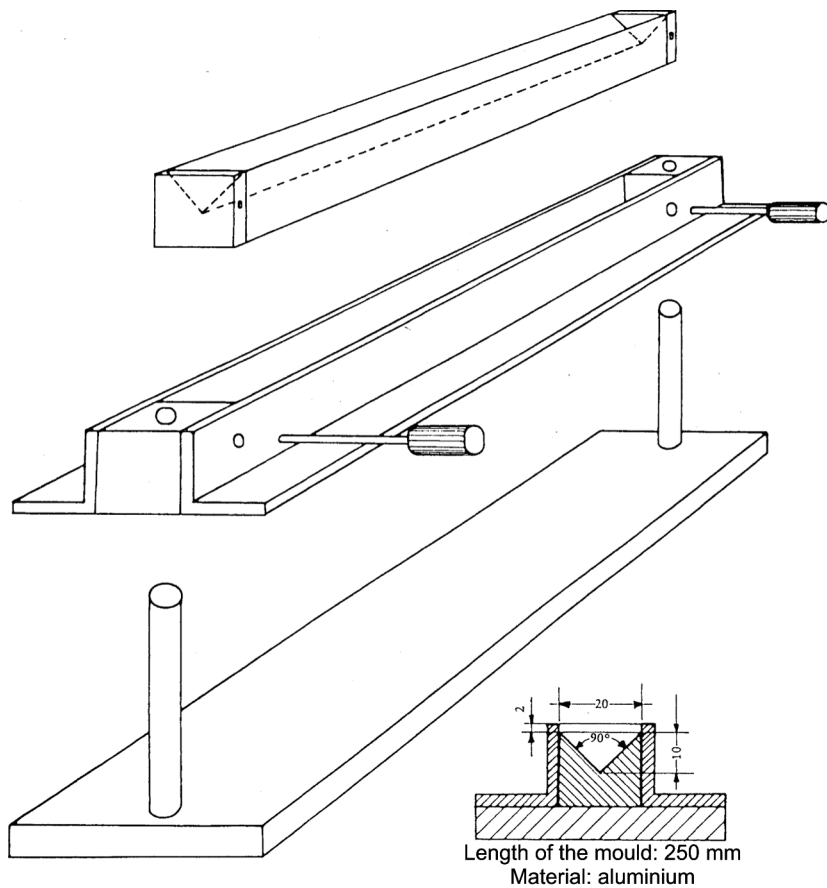
Powdery, granular or paste-like substances are to be considered as highly flammable when the time of burning in any tests carried out according to the test procedure described in 1.6.2 is less than 45 seconds. Powders of metals or metal-alloys are considered to be highly flammable when they can be ignited and the flame or the zone of reaction spreads over the whole sample in 10 minutes or less.

**4. REFERENCES**

NF T 20-042 (September 85) Chemical products for industrial use. Determination of the flammability of solids.

**▼B***Appendix**Figure***Mould and accessories for the preparation of the pile**

(All dimensions in millimetres)



**▼B****A.11. FLAMMABILITY (GASES)****1. METHOD****1.1. INTRODUCTION**

This method allows a determination of whether gases mixed with air at room temperature (circa 20 °C) and atmospheric pressure are flammable and, if so, over what range of concentrations. Mixtures of increasing concentrations of the test gas with air are exposed to an electrical spark and it is observed whether ignition occurs.

**1.2. DEFINITION AND UNITS**

The range of flammability is the range of concentration between the lower and the upper explosion limits. The lower and the upper explosion limits are those limits of concentration of the flammable gas in admixture with air at which propagation of a flame does not occur.

**1.3. REFERENCE SUBSTANCES**

Not specified.

**1.4. PRINCIPLE OF THE METHOD**

The concentration of gas in air is increased step by step and the mixture is exposed at each stage to an electrical spark.

**1.5. QUALITY CRITERIA**

Not stated.

**1.6. DESCRIPTION OF THE METHOD****1.6.1. Apparatus**

The test vessel is an upright glass cylinder having a minimum inner diameter of 50 mm and a minimum height of 300 mm. The ignition electrodes are separated by a distance of 3 to 5 mm and are placed 60 mm above the bottom of the cylinder. The cylinder is fitted with a pressure-release opening. The apparatus has to be shielded to restrict any explosion damage.

A standing induction spark of 0,5 sec. duration, which is generated from a high voltage transformer with an output voltage of 10 to 15 kV (maximum of power input 300 W), is used as the ignition source. An example of a suitable apparatus is described in reference (2).

**1.6.2. Test conditions**

The test must be performed at room temperature (circa 20 °C).

**▼B**1.6.3. **Performance of the test**

Using proportioning pumps, a known concentration of gas in air is introduced into the glass cylinder. A spark is passed through the mixture and it is observed whether or not a flame detaches itself from the ignition source and propagates independently. The gas concentration is varied in steps of 1 % vol. until ignition occurs as described above.

If the chemical structure of the gas indicates that it would be non-flammable and the composition of the stoichiometric mixture with air can be calculated, then only mixtures in the range from 10 % less than the stoichiometric composition to 10 % greater than this composition need be tested in 1 % steps.

2. **DATA**

The occurrence of flame propagation is the only relevant information data for the determination of this property.

3. **REPORTING**

The test report shall, if possible, include the following information:

- the precise specification of the substance (identification and impurities),
- a description, with dimensions, of the apparatus used,
- the temperature at which the test was performed,
- the tested concentrations and the results obtained,
- the result of the test: non-flammable gas or highly flammable gas,
- if it is concluded that the gas is non-flammable then the concentration range over which it was tested in 1 % steps should be stated,
- all information and remarks relevant to the interpretation of results have to be reported.

4. **REFERENCES**

- (1) NF T 20-041 (September 85) Chemical products for industrial use. Determination of the flammability of gases.
- (2) W. Berthold, D. Conrad, T. Grewer, H. Grosse-Wortmann 'Entwicklung einer Standard-Apparatur zur Messung von Explosionsgrenzen'. Chem.-Ing.- Tech. 1984, vol. 56, 2, 126-127., T. Redeker und H. Schacke, p. 126-127.

**▼B****A.12. FLAMMABILITY (CONTACT WITH WATER)****1. METHOD****1.1. INTRODUCTION**

This test method can be used to determine whether the reaction of a substance with water or damp air leads to the development of dangerous amounts of gas or gases which may be highly flammable.

The test method can be applied to both solid and liquid substances. This method is not applicable to substances which spontaneously ignite when in contact with air.

**1.2. DEFINITIONS AND UNITS**

Highly flammable: substances which, in contact with water or damp air, evolve highly flammable gases in dangerous quantities at a minimum rate of 1 litre/kg per hour.

**1.3. PRINCIPLE OF THE METHOD**

The substance is tested according to the step by step sequence described below; if ignition occurs at any step, no further testing is necessary. If it is known that the substance does not react violently with water then proceed to step 4 (1.3.4).

**1.3.1. Step 1**

The test substance is placed in a trough containing distilled water at 20 °C and it is noted whether or not the evolved gas ignites.

**1.3.2. Step 2**

The test substance is placed on a filter paper floating on the surface of a dish containing distilled water at 20 °C and it is noted whether or not the evolved gas ignites. The filter paper is merely to keep the substance in one place to increase the chances of ignition.

**1.3.3. Step 3**

The test substance is made into a pile approximately 2 cm high and 3 cm diameter. A few drops of water are added to the pile and it is noted whether or not the evolved gas ignites.

**1.3.4. Step 4**

The test substance is mixed with distilled water at 20 °C and the rate of evolution of gas is measured over a period of seven hours, at one-hour intervals. If the rate of evolution is erratic, or is increasing, after seven hours, the measuring time should be extended to a maximum time of five days. The test may be stopped if the rate at any time exceeds 1 litre/kg per hour.



**▼B**

## 1.4. REFERENCE SUBSTANCES

Not specified.

## 1.5. QUALITY CRITERIA

Not stated.

## 1.6. DESCRIPTION OF METHODS

1.6.1. **Step 1**1.6.1.1. *Test conditions*

The test is performed at room temperature (circa 20 °C).

1.6.1.2. *Performance of the test*

A small quantity (approximately 2 mm diameter) of the test substance should be placed in a trough containing distilled water. A note should be made of whether (i) any gas is evolved and (ii) if ignition of the gas occurs. If ignition of the gas occurs then no further testing of the substance is needed because the substance is regarded as hazardous.

1.6.2. **Step 2**1.6.2.1. *Apparatus*

A filter-paper is floated flat on the surface of distilled water in any suitable vessel, e.g. a 100 mm diameter evaporating dish.

1.6.2.2. *Test conditions*

The test is performed at room temperature (circa 20 °C).

1.6.2.3. *Performance of the test*

A small quantity of the test substance (approximately 2 mm diameter) is placed onto the centre of the filter-paper. A note should be made of whether (i) any gas is evolved and (ii) if ignition of the gas occurs. If ignition of the gas occurs then no further testing of the substance is needed because the substance is regarded as hazardous.

1.6.3. **Step 3**1.6.3.1. *Test conditions*

The test is performed at room temperature (circa 20 °C).

1.6.3.2. *Performance of the test*

The test substance is made into a pile approximately 2 cm high and 3 cm diameter with an indentation in the top. A few drops of water are added to the hollow and a note is made of whether (i) any gas is evolved and (ii) if ignition of the gas occurs. If ignition of the gas occurs then no further testing of the substance is needed because the substance is regarded as hazardous.

**▼ B**1.6.4. **Step 4**1.6.4.1. *Apparatus*

The apparatus is set up as shown in the figure.

1.6.4.2. *Test conditions*

Inspect the container of the test substance for any powder < 500 µm (particle size). If the powder constitutes more than 1 % w/w of the total, or if the sample is friable, then the whole of the substance should be ground to a powder before testing to allow for a reduction in particle size during storage and handling; otherwise the substance is to be tested as received. The test should be performed at room temperature (circa 20 °C) and atmospheric pressure.

1.6.4.3. *Performance of the test*

10 to 20 ml of water are put into the dropping funnel of the apparatus and 10 g of substance are put in the conical flask. The volume of gas evolved can be measured by any suitable means. The tap of the dropping funnel is opened to let the water into the conical flask and a stop watch is started. The gas evolution is measured each hour during a seven hour period. If, during this period, the gas evolution is erratic, or if, at the end of this period, the rate of gas evolution is increasing, then measurements should be continued for up to five days. If, at any time of measurement, the rate of gas evolution exceeds 1 litre/kg per hour, the test can be discontinued. This test should be performed in triplicate.

If the chemical identity of the gas is unknown, the gas should be analysed. When the gas contains highly flammable components and it is unknown whether the whole mixture is highly flammable, a mixture of the same composition has to be prepared and tested according to the method A.11.

2. **DATA**

The substance is considered hazardous if:

— spontaneous ignition takes place in any step of the test procedure,

or

— there is evolution of flammable gas at a rate greater than 1 litre/kg of the substance per hour.

3. **REPORTING**

The test report shall, if possible, include the following information:

— the precise specification of the substance (identification and impurities),

— details of any initial preparation of the test substance,

**▼B**

- the results of the tests (steps 1, 2, 3 and 4),
- the chemical identity of gas evolved,
- the rate of evolution of gas if step 4 (1.6.4) is performed,
- any additional remarks relevant to the interpretation of the results.

**4. REFERENCES**

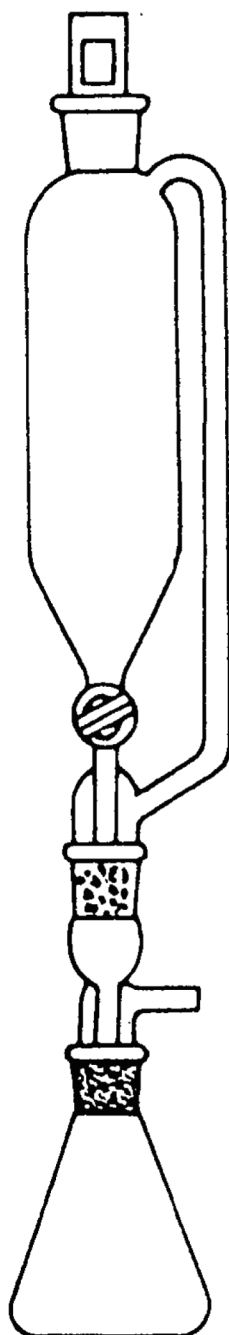
- (1) Recommendations on the transport of dangerous goods, test and criteria, 1990, United Nations, New York.
- (2) NF T 20-040 (September 85) Chemical products for industrial use. Determination of the flammability of gases formed by the hydrolysis of solid and liquid products.

▼ B

*Appendix*

*Figure*

**Apparatus**



**▼B****A.13. PYROPHORIC PROPERTIES OF SOLIDS AND LIQUIDS****1. METHOD****1.1. INTRODUCTION**

The test procedure is applicable to solid or liquid substances, which, in small amounts, will ignite spontaneously a short time after coming into contact with air at room temperature (circa 20 °C).

Substances which need to be exposed to air for hours or days at room temperature or at elevated temperatures before ignition occurs are not covered by this test method.

**1.2. DEFINITIONS AND UNITS**

Substances are considered to have pyrophoric properties if they ignite or cause charring under the conditions described in 1.6.

The auto-flammability of liquids may also need to be tested using method A.15. Auto-ignition temperature (liquids and gases).

**1.3. REFERENCE SUBSTANCES**

Not specified.

**1.4. PRINCIPLE OF THE METHOD**

The substance, whether solid or liquid, is added to an inert carrier and brought into contact with air at ambient temperature for a period of five minutes. If liquid substances do not ignite then they are absorbed onto filter paper and exposed to air at ambient temperature (circa 20 °C) for five minutes. If a solid or liquid ignites, or a liquid ignites or chars a filter paper, then the substance is considered to be pyrophoric.

**1.5. QUALITY CRITERIA**

Repeatability: because of the importance in relation to safety, a single positive result is sufficient for the substance to be considered pyrophoric.

**1.6. DESCRIPTION OF THE TEST METHOD****1.6.1. Apparatus**

A porcelain cup of circa 10 cm diameter is filled with diatomaceous earth to a height of about 5 mm at room temperature (circa 20 °C).

Note:

Diatomaceous earth or any other comparable inert substance which is generally obtainable shall be taken as representative of soil onto which the test substance might be spilled in the event of an accident.

Dry filter paper is required for testing liquids which do not ignite on contact with air when in contact with an inert carrier.

**▼B**1.6.2. **Performance of the test**(a) *Powdery solids*

1 to 2 cm<sup>3</sup> of the substance to be tested is poured from circa 1 m height onto a non-combustible surface and it is observed whether the substance ignites during dropping or within five minutes of settling.

The test is performed six times unless ignition occurs;

(b) *liquids*

Circa 5 cm<sup>3</sup> of the liquid to be tested is poured into the prepared porcelain cup and it is observed whether the substance ignites within five minutes.

If no ignition occurs in the six tests, perform the following tests:

A 0,5 ml test sample is delivered from a syringe to an indented filter paper and it is observed whether ignition or charring of the filter paper occurs within five minutes of the liquid being added. The test is performed three times unless ignition or charring occurs.

2. **DATA**

## 2.1. TREATMENT OF RESULTS

Testing can be discontinued as soon as a positive result occurs in any of the tests.

## 2.2. EVALUATION

If the substance ignites within five minutes when added to an inert carrier and exposed to air, or a liquid substance chars or ignites a filter paper within five minutes when added and exposed to air, it is considered to be pyrophoric.

3. **REPORTING**

The test report shall, if possible, include the following information:

- the precise specification of the substance (identification and impurities),
- the results of the tests,
- any additional remark relevant to the interpretation of the results.

4. **REFERENCES**

- (1) NF T 20-039 (September 85) Chemical products for industrial use. Determination of the spontaneous flammability of solids and liquids.
- (2) Recommendations on the Transport of Dangerous Goods, Test and criteria, 1990, United Nations, New York.

**▼B****A.14. EXPLOSIVE PROPERTIES****1. METHOD****1.1. INTRODUCTION**

The method provides a scheme of testing to determine whether a solid or a pasty substance presents a danger of explosion when submitted to the effect of a flame (thermal sensitivity), or to shock or friction (sensitivity to mechanical stimuli), and whether a liquid substance presents a danger of explosion when submitted to the effect of a flame or shock.

The method comprises three parts:

- (a) a test of thermal sensitivity (1);
- (b) a test of mechanical sensitivity with respect to shock (1);
- (c) a test of mechanical sensitivity with respect to friction (1).

The method yields data to assess the likelihood of initiating an explosion by means of certain common stimuli. The method is not intended to ascertain whether a substance is capable of exploding under any conditions.

The method is appropriate for determining whether a substance will present a danger of explosion (thermal and mechanical sensitivity) under the particular conditions specified in the directive. It is based on a number of types of apparatus which are widely used internationally (1) and which usually give meaningful results. It is recognised that the method is not definitive. Alternative apparatus to that specified may be used provided that it is internationally recognised and the results can be adequately correlated with those from the specified apparatus.

The tests need not be performed when available thermodynamic information (e.g. heat of formation, heat of decomposition) and/or absence of certain reactive groups (2) in the structural formula establishes beyond reasonable doubt that the substance is incapable of rapid decomposition with evolution of gases or release of heat (i.e. the material does not present any risk of explosion). A test of mechanical sensitivity with respect to friction is not required for liquids.

**1.2. DEFINITIONS AND UNITS**

Explosive:

Substances which may explode under the effect of flame or which are sensitive to shock or friction in the specified apparatus (or are more mechanically sensitive than 1,3-dinitrobenzene in alternative apparatus).

**1.3. REFERENCE SUBSTANCES**

1,3-dinitrobenzene, technical crystalline product sieved to pass 0,5 mm, for the friction and shock methods.

Perhydro-1,3,5-trinitro-1,3,5-triazine (RDX, hexogen, cyclonite — CAS 121-82-4), recrystallised from aqueous cyclohexanone, wet-sieved through a 250 µm and retained on a 150 µm sieve and dried at 103 ± 2 °C (for four hours) for the second series of friction and shock tests.

**▼ B**

## 1.4. PRINCIPLE OF THE METHOD

Preliminary tests are necessary to establish safe conditions for the performance of the three tests of sensitivity.

1.4.1. **Safety-in-handling tests (3)**

For safety reasons, before performing the main tests, very small samples (circa 10 mg) of the substance are subjected to heating without confinement in a gas flame, to shock in any convenient form of apparatus and to friction by the use of a mallet against an anvil or any form of friction machine. The objective is to ascertain if the substance is so sensitive and explosive that the prescribed sensitivity tests, particularly that of thermal sensitivity, should be performed with special precautions so as to avoid injury to the operator.

1.4.2. **Thermal sensitivity**

The method involves heating the substance in a steel tube, closed by orifice plates with differing diameters of hole, to determine whether the substance is liable to explode under conditions of intense heat and defined confinement.

1.4.3. **Mechanical sensitivity (shock)**

The method involves subjecting the substance to the shock from a specified mass dropped from a specified height.

1.4.4. **Mechanical sensitivity (friction)**

The method involves subjecting solid or pasty substances to friction between standard surfaces under specified conditions of load and relative motion.

## 1.5. QUALITY CRITERIA

Not stated.

## 1.6. DESCRIPTION OF METHOD

1.6.1. **Thermal sensitivity (effect of a flame)**1.6.1.1. *Apparatus*

The apparatus consists of a non-reusable steel tube with its re-usable closing device (figure 1), installed in a heating and protective device. Each tube is deep-drawn from sheet steel (see Appendix) and has an internal diameter of 24 mm, a length of 75 mm and wall thickness of 0,5 mm. The tubes are flanged at the open end to enable them to be closed by the orifice plate assembly. This consists of a pressure-resistant orifice plate, with a central hole, secured firmly to a tube using a two-part screw joint (nut and threaded collar). The nut and threaded collar are made from chromium-manganese steel (see Appendix) which is spark-free up to 800 °C. The orifice plates are 6 mm thick, made from heat-resistant steel (see Appendix), and are available with a range of diameters of opening.



**▼ B**1.6.1.2. *Test conditions*

Normally the substance is tested as received although in certain cases, e.g. if pressed, cast or otherwise condensed, it may be necessary to test the substance after crushing.

For solids, the mass of material to be used in each test is determined using a two-stage dry run procedure. A tared tube is filled with 9 cm<sup>3</sup> of substance and the substance tamped with 80 N force applied to the total cross-section of the tube. For reasons of safety or in cases where the physical form of the sample can be changed by compression other filling procedures may be used; e.g. if the substance is very friction sensitive then tamping is not appropriate. If the material is compressible then more is added and tamped until the tube is filled to 55 mm from the top. The total mass used to fill the tube to the 55 mm level is determined and two further increments, each tamped with 80 N force, are added. Material is then either added with tamping, or taken out, as required, to leave the tube filled to a level 15 mm from the top. A second dry run is performed, starting with a tamped quantity of a third of the total mass found in the first dry run. Two more of these increments are added with 80 N tamping and the level of the substance in the tube adjusted to 15 mm from the top by addition or subtraction of material as required. The amount of solid determined in the second dry run is used for each trial; filling being performed in three equal amounts, each compressed to 9 cm<sup>3</sup> by whatever force is necessary. (This may be facilitated by the use of spacing rings).

Liquids and gels are loaded into the tube to a height of 60 mm taking particular care with gels to prevent the formation of voids. The threaded collar is slipped onto the tube from below, the appropriate orifice plate is inserted and the nut tightened after applying some molybdenum disulphide based lubricant. It is essential to check that none of the substance is trapped between the flange and the plate, or in the threads.

Heating is provided by propane taken from an industrial cylinder, fitted with a pressure regulator (60 to 70 mbar), through a meter and evenly distributed (as indicated by visual observation of the flames from the burners) by a manifold to four burners. The burners are located around the test chamber as shown in figure 1. The four burners have a combined consumption of about 3,2 litres of propane per minute. Alternative fuel gases and burners may be used but the heating rate must be as specified in figure 3. For all apparatus, the heating rate must be checked periodically using tubes filled with dibutyl phthalate as indicated in figure 3.

1.6.1.3. *Performance of the tests*

Each test is performed until either the tube is fragmented or the tube has been heated for five minutes. A test resulting in the fragmentation of the tube into three or more pieces, which in some cases may be connected to each other by narrow strips of metal as illustrated in figure 2, is evaluated as giving an explosion. A test resulting in fewer fragments or no fragmentation is regarded as not giving an explosion.

**▼B**

A series of three tests with a 6,0 mm diameter orifice plate is first performed and, if no explosions are obtained, a second series of three tests is performed with a 2,0 mm diameter orifice plate. If an explosion occurs during either test series then no further tests are required.

1.6.1.4. *Evaluation*

The test result is considered positive if an explosion occurs in either of the above series of tests.

1.6.2. **Mechanical sensitivity (shock)**1.6.2.1. *Apparatus (figure 4)*

The essential parts of a typical fall hammer apparatus are a cast steel block with base, anvil, column, guides, drop weights, release device and a sample holder. The steel anvil 100 mm (diameter) × 70 mm (height) is screwed to the top of a steel block 230 mm (length) × 250 mm (width) × 200 mm (height) with a cast base 450 mm (length) × 450 mm (width) × 60 mm (height). A column, made from seamless drawn steel tube, is secured in a holder screwed on to the back of the steel block. Four screws anchor the apparatus to a solid concrete block 60 × 60 × 60 cm such that the guide rails are absolutely vertical and the drop weight falls freely. 5 and 10 kg weights, made from solid steel, are available for use. The striking head of each weight is of hardened steel, HRC 60 to 63, and has a minimum diameter of 25 mm.

The sample under test is enclosed in a shock device consisting of two coaxial solid steel cylinders, one above the other, in a hollow cylindrical steel guide ring. The solid steel cylinders should be of 10 (- 0,003, - 0,005) mm diameter and 10 mm height and have polished surfaces, rounded edges (radius of curvature 0,5 mm) and a hardness of HRC 58 to 65. The hollow cylinder must have an external diameter of 16 mm, a polished bore of 10 (+ 0,005, + 0,010) mm and a height of 13 mm. The shock device is assembled on an intermediate anvil (26 mm diameter and 26 mm height) made of steel and centred by a ring with perforations to allow escape of fumes.

1.6.2.2. *Test conditions*

The sample volume should be 40 mm<sup>3</sup>, or a volume to suit any alternative apparatus. Solid substances should be tested in the dry state and prepared as follows:

- (a) powdered substances are sieved (sieve size 0,5 mm); all that has passed through the sieve is used for testing;
- (b) pressed, cast or otherwise condensed substances are broken into small pieces and sieved; the sieve fraction from 0,5 to 1 mm diameter is used for testing and should be representative of the original substance.

Substances normally supplied as pastes should be tested in the dry state where possible or, in any case, following removal of the maximum possible amount of diluent. Liquid substances are tested with a 1 mm gap between the upper and lower steel cylinders.

**▼B**1.6.2.3. *Performance of the tests*

A series of six tests are performed dropping the 10 kg mass from 0,40 m (40 J). If an explosion is obtained during the six tests at 40 J, a further series of six tests, dropping a 5 kg mass from 0,15 m (7,5 J), must be performed. In other apparatus, the sample is compared with the chosen reference substance using an established procedure (e.g. up-and-down technique etc.).

1.6.2.4. *Evaluation*

The test result is considered positive if an explosion (bursting into flame and/or a report is equivalent to explosion) occurs at least once in any of the tests with the specified shock apparatus or the sample is more sensitive than 1,3-dinitrobenzene or RDX in an alternative shock test.

1.6.3. **Mechanical sensitivity (friction)**1.6.3.1. *Apparatus (figure 5)*

The friction apparatus consists of a cast steel base plate on which is mounted the friction device. This consists of a fixed porcelain peg and moving porcelain plate. The porcelain plate is held in a carriage which runs in two guides. The carriage is connected to an electric motor via a connecting rod, an eccentric cam and suitable gearing such that the porcelain plate is moved, once only, back and forth beneath the porcelain peg for a distance of 10 mm. The porcelain peg may be loaded with, for example, 120 or 360 newtons.

The flat porcelain plates are made from white technical porcelain (roughness 9 to 32  $\mu\text{m}$ ) and have the dimensions 25 mm (length)  $\times$  25 mm (width)  $\times$  5 mm (height). The cylindrical porcelain peg is also made of white technical porcelain and is 15 mm long, has a diameter of 10 mm and roughened spherical end surfaces with a radius of curvature of 10 mm.

1.6.3.2. *Test conditions*

The sample volume should be 10 mm<sup>3</sup> or a volume to suit any alternative apparatus.

Solid substances are tested in the dry state and prepared as follows:

- (a) powdered substances are sieved (sieve size 0,5 mm); all that has passed through the sieve is used for testing;
- (b) pressed, cast or otherwise condensed substances are broken into small pieces and sieved; the sieve fraction  $<$  0,5 mm diameter is used for testing.

Substances normally supplied as pastes should be tested in the dry state where possible. If the substance cannot be prepared in the dry state, the paste (following removal of the maximum possible amount of diluent) is tested as a 0,5 mm thick, 2 mm wide, 10 mm long film, prepared with a former.

**▼B**1.6.3.3. *Performance of the tests*

The porcelain peg is brought onto the sample under test and the load applied. When carrying out the test, the sponge marks of the porcelain plate must lie transversely to the direction of the movement. Care must be taken that the peg rests on the sample, that sufficient test material lies under the peg and also that the plate moves correctly under the peg. For pasty substances, a 0,5 mm thick gauge with a  $2 \times 10$  mm slot is used to apply the substance to the plate. The porcelain plate has to move 10 mm forwards and backwards under the porcelain peg in a time of 0,44 seconds. Each part of the surface of the plate and peg must only be used once; the two ends of each peg will serve for two trials and the two surfaces of a plate will each serve for three trials.

A series of six tests are performed with a 360 N loading. If a positive event is obtained during these six tests, a further series of six tests must be performed with a 120 N loading. In other apparatus, the sample is compared with the chosen reference substance using an established procedure (e.g. up-and-down technique, etc.).

1.6.3.4. *Evaluation*

The test result is considered positive if an explosion (crepitation and/or a report or bursting into flame are equivalent to explosion) occurs at least once in any of the tests with the specified friction apparatus or satisfies the equivalent criteria in an alternative friction test.

2. **DATA**

In principle, a substance is considered to present a danger of explosion in the sense of the directive if a positive result is obtained in the thermal, shock or friction sensitivity test.

3. **REPORTING**

## 3.1. TEST REPORT

The test report shall, if possible, include the following information:

- identity, composition, purity, moisture content, etc. of the substance tested,
- the physical form of the sample and whether or not it has been crushed, broken and/or sieved,
- observations during the thermal sensitivity tests (e.g. sample mass, number of fragments, etc.),
- observations during the mechanical sensitivity tests (e.g. formation of considerable amounts of smoke or complete decomposition without a report, flames, sparks, report, crepitation, etc.),
- results of each type of test,
- if alternative apparatus has been used, scientific justification as well as evidence of correlation between results obtained with specified apparatus and those obtained with equivalent apparatus must be given,

**▼B**

- any useful comments such as reference to tests with similar products which might be relevant to a proper interpretation of the results,
- all additional remarks relevant for the interpretation of the results.

**3.2. INTERPRETATION AND EVALUATION OF RESULTS**

The test report should mention any results which are considered false, anomalous or unrepresentative. If any of the results should be discounted, an explanation and the results of any alternative or supplementary testing should be given. Unless an anomalous result can be explained, it must be accepted at face value and used to classify the substance accordingly.

**4. REFERENCES**

- (1) Recommendations on the Transport of Dangerous Goods: Tests and criteria, 1990, United Nations, New York.
- (2) Bretherick, L., Handbook of Reactive Chemical Hazards, 4th edition, Butterworths, London, ISBN 0-750-60103-5, 1990.
- (3) Koenen, H., Ide, K.H. and Swart, K.H., Explosivstoffe, 1961, vol. 3, 6-13 and 30-42.
- (4) NF T 20-038 (September 85) Chemical products for industrial use — Determination of explosion risk.

▼B

## Appendix

## Example of material specification for thermal sensitivity test (see DIN 1623)

- (1) Tube: Material specification No 1.0336.505 g
- (2) Orifice plate: Material specification No 1.4873
- (3) Threaded collar and nut: Material specification No 1.3817

Figure 1

## Thermal sensitivity test apparatus

(all dimensions in millimetres)

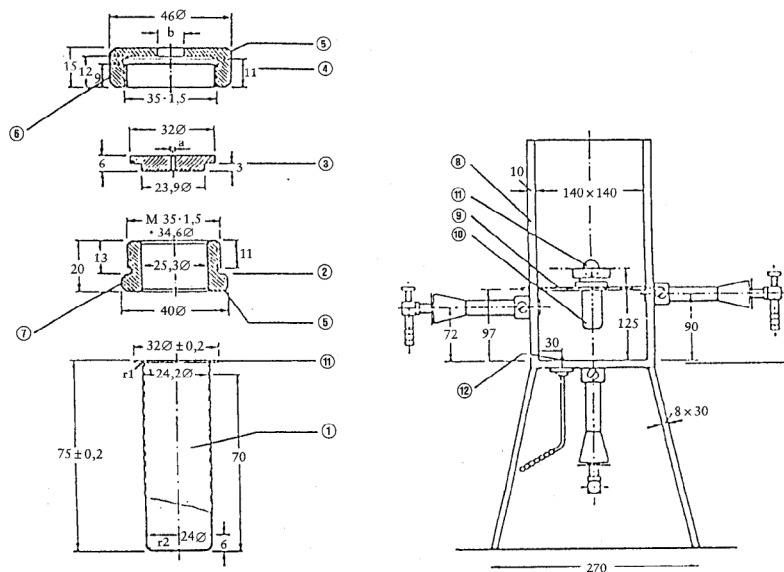


Fig. 1a Steel tube and accessories

Fig. 1b Heating and protective device

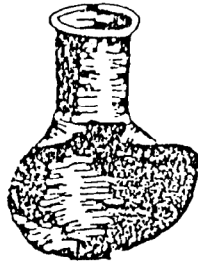
- |  |  |
|--|--|
| (1) tube   | (7) 2 flat for spanner size 36                               |
| (1a) outer flange                                | (8) splinter-proof box                                       |
| (2) threaded collar; low-friction thread         | (9) 2 supporting rods for tube                               |
| (3) orifice plate $a = 2,0$ or $6,0$ mm diameter | (10) assembled tube  |
| (4) nut $b = 10$ mm diameter                     | (11) position for rear burner; the other burners are visible |
| (5) chamfered surface                            | (12) pilot jet   |
| (6) 2 flat for spanner size 41                   |  |

▼B

Figure 2

**Thermal sensitivity test**

(example of fragmentation)



No explosion



No explosion



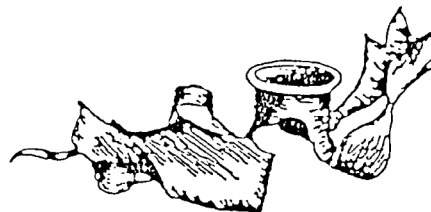
Explosion



Explosion



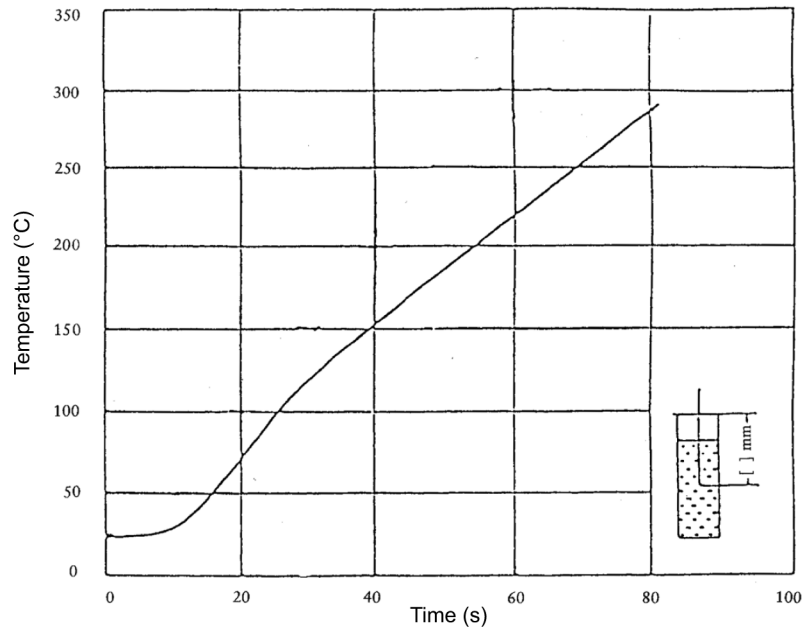
Explosion



Explosion

▼B

Figure 3

**Heating rate calibration for thermal sensitivity test**

Temperature/time curve obtained on heating dibutyl phthalate ( $27 \text{ cm}^3$ ) in a closed (1,5 mm orifice plate) tube using a propane flow rate of 3,2 litre/minute. The temperature is measured with a 1 mm diameter stainless steel sheathed chromel/alumel thermocouple, placed centrally 43 mm below the rim of the tube. The heating rate between 135 °C and 285 °C should be between 185 and 215 K/minute.



▼B

Figure 4

## Shock test apparatus

(all dimensions in millimetres)

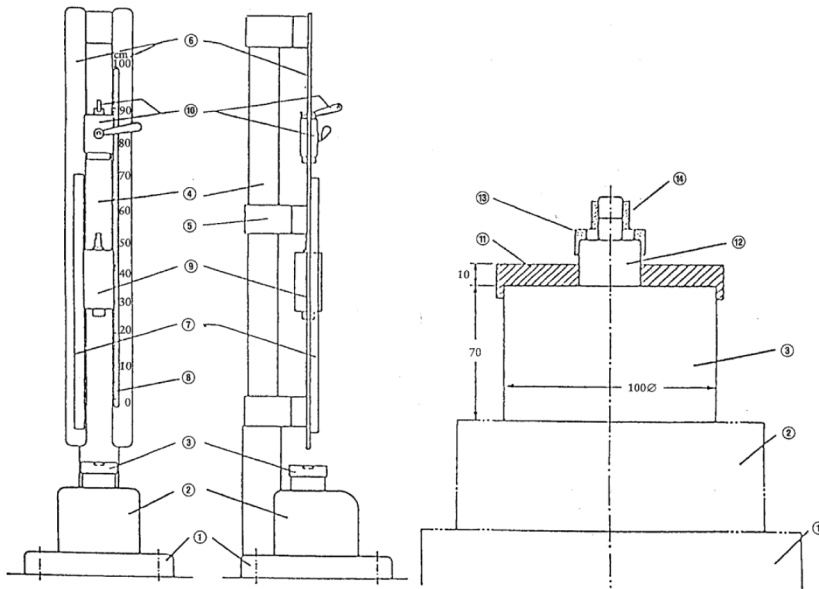


Fig. 4a Fall-hammer, front and side, general view

- (1) base, 450 x 450 x 60
- (2) steel block, 230 x 250 x 200
- (3) anvil, 100 diameter x 70
- (4) column
- (5) median cross-member
- (6) 2 guides
- (7) toothed rack

Fig. 4b Fall-hammer, lower part

- (8) graduated scale
- (9) fall-hammer (drop mass)
- (10) holding and releasing device
- (11) locating plate
- (12) intermediate anvil (interchangeable),  
26 diameter x 26
- (13) locating ring with orifices
- (14) impact device

▼B

Figure 4

Continued

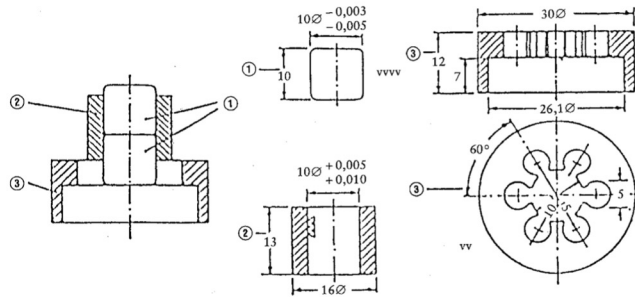


Fig. 4c Shock device for substances in powdered or paste-like form

Fig. 4d Shock device for liquid substances

- (1) steel cylinders
- (2) guide ring for steel cylinders
- (3) locating ring with orifices
  - (a) vertical section
  - (b) plan
- (4) rubber ring
- (5) liquid substance (40 mm<sup>3</sup>)
- (6) space free from liquid

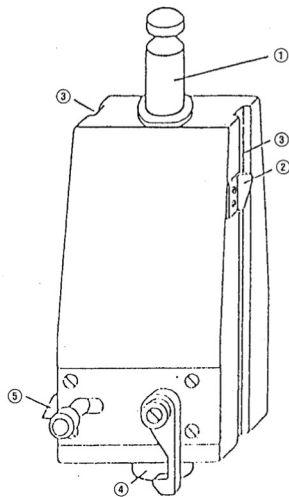
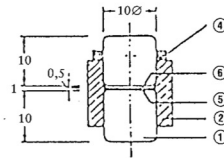


Fig. 4e Hammer (drop mass of 5 kg)

- (1) suspension spigot
- (2) height marker
- (3) positioning groove
- (4) cylindrical striking head
- (5) rebound catch

▼ B

Figure 5

## Friction sensitivity apparatus

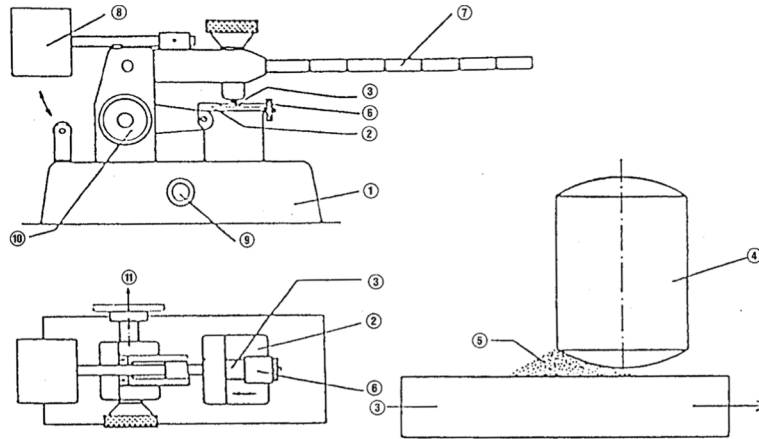


Fig. 5a Friction apparatus; elevation and plan view Fig. 5b Starting position of peg on sample

- |   |  |
|---|--|
| (1) steel base  | (6) peg-holder                                       |
| (2) movable carriage                                    | (7) loading arm                                      |
| (3) porcelain plate, 25 x 25 x 5 mm, held on carriage   | (8) counterweight                                    |
| (4) fixed porcelain peg, 10 diameter x 15 mm            | (9) switch   |
| (5) sample under test, approximately 10 mm <sup>3</sup> | (10) wheel for setting carriage at starting position |
|   | (11) direction to electric drive motor               |

**▼B**A.15. **AUTO-IGNITION TEMPERATURE (LIQUIDS AND GASES)**1. **METHOD**

## 1.1. INTRODUCTION

Explosive substances and substances which ignite spontaneously in contact with air at ambient temperature should not be submitted to this test. The test procedure is applicable to gases, liquids and vapours which, in the presence of air, can be ignited by a hot surface.

The auto-ignition temperature can be considerably reduced by the presence of catalytic impurities, by the surface material or by a higher volume of the test vessel.

## 1.2. DEFINITIONS AND UNITS

The degree of auto-ignitability is expressed in terms of the auto-ignition temperature. The auto-ignition temperature is the lowest temperature at which the test substance will ignite when mixed with air under the conditions defined in the test method.

## 1.3. REFERENCE SUBSTANCES

Reference substances are cited in the standards (see 1.6.3). They should primarily serve to check the performance of the method from time to time and to allow comparison with results from other methods.

## 1.4. PRINCIPLE OF THE METHOD

The method determines the minimum temperature of the inner surface of an enclosure that will result in ignition of a gas, vapour or liquid injected into the enclosure.

## 1.5. QUALITY CRITERIA

The repeatability varies according to the range of auto-ignition temperatures and the test method used.

The sensitivity and specificity depend on the test method used.

## 1.6. DESCRIPTION OF THE METHOD

1.6.1. **Apparatus**

The apparatus is described in the method referred to in 1.6.3.

1.6.2. **Test conditions**

A sample of the test substance is tested according to the method referred to in 1.6.3.

1.6.3. **Performance of the test**

See IEC 79-4, DIN 51794, ASTM-E 659-78, BS 4056, NF T 20-037.

**▼B****2. DATA**

Record the test-temperature, atmospheric pressure, quantity of sample used and time-lag until ignition occurs.

**3. REPORTING**

The test report shall, if possible, include the following information:

- the precise specification of the substance (identification and impurities),
- the quantity of sample used, atmospheric pressure,
- the apparatus used,
- the results of measurements (test temperatures, results concerning ignition, corresponding time-lags),
- all additional remarks relevant to the interpretation of results.

**4. REFERENCES**

None.

**▼ B****A.16. RELATIVE SELF-IGNITION TEMPERATURE FOR SOLIDS****1. METHOD****1.1. INTRODUCTION**

Explosive substances and substances which ignite spontaneously in contact with air at ambient temperature should not be submitted to this test.

The purpose of this test is to provide preliminary information on the auto-flammability of solid substances at elevated temperatures.

If the heat developed either by a reaction of the substance with oxygen or by exothermic decomposition is not lost rapidly enough to the surroundings, self-heating leading to self-ignition occurs. Self-ignition therefore occurs when the rate of heat-production exceeds the rate of heat loss.

The test procedure is useful as a preliminary screening test for solid substances. In view of the complex nature of the ignition and combustion of solids, the self-ignition temperature determined according to this test method should be used for comparison purposes only.

**1.2. DEFINITIONS AND UNITS**

The self-ignition temperature as obtained by this method is the minimum ambient temperature expressed in °C at which a certain volume of a substance will ignite under defined conditions.

**1.3. REFERENCE SUBSTANCE**

None.

**1.4. PRINCIPLE OF THE METHOD**

A certain volume of the substance under test is placed in an oven at room temperature; the temperature/time curve relating to conditions in the centre of the sample is recorded while the temperature of the oven is increased to 400 °C, or to the melting point if lower, at a rate of 0,5 °C/min. For the purpose of this test, the temperature of the oven at which the sample temperature reaches 400 °C by self-heating is called the self-ignition temperature.

**1.5. QUALITY CRITERIA**

None.

**1.6. DESCRIPTION OF THE METHOD****1.6.1. Apparatus****1.6.1.1. Oven**

A temperature-programmed laboratory oven (volume about 2 litres) fitted with natural air circulation and explosion relief. In order to avoid a potential explosion risk, any decomposition gases must not be allowed to come into contact with the electric heating elements.

**▼ B**1.6.1.2. *Wire mesh cube*

A piece of stainless steel wire mesh with 0,045 mm openings should be cut according to the pattern in figure 1. The mesh should be folded and secured with wire into an open-topped cube.

1.6.1.3. *Thermocouples*

Suitable thermocouples.

1.6.1.4. *Recorder*

Any two-channel recorder calibrated from 0 to 600 °C or corresponding voltage.

1.6.2. **Test conditions**

Substances are tested as received.

1.6.3. **Performance of the test**

The cube is filled with the substance to be tested and is tapped gently, adding more of the substance until the cube is completely full. The cube is then suspended in the centre of the oven at room temperature. One thermocouple is placed at the centre of the cube and the other between the cube and the oven wall to record the oven temperature.

The temperatures of the oven and sample are continuously recorded while the temperature of the oven is increased to 400 °C, or to the melting point if lower, at a rate of 0,5 °C/min.

When the substance ignites the sample thermocouple will show a very sharp temperature rise above the oven temperature.

2. **DATA**

The temperature of the oven at which the sample temperature reaches 400 °C by self-heating is relevant for evaluation (see figure 2).

3. **REPORTING**

The test report shall, if possible, include the following information:

- a description of the substance to be tested,
- the results of measurement including the temperature/time curve,
- all additional remarks relevant for the interpretation of the results.

4. **REFERENCES**

NF T 20-036 (September 85) Chemical products for industrial use. Determination of the relative temperature of the spontaneous flammability of solids.

▼ B

Figure 1

Pattern of 20 mm test cube

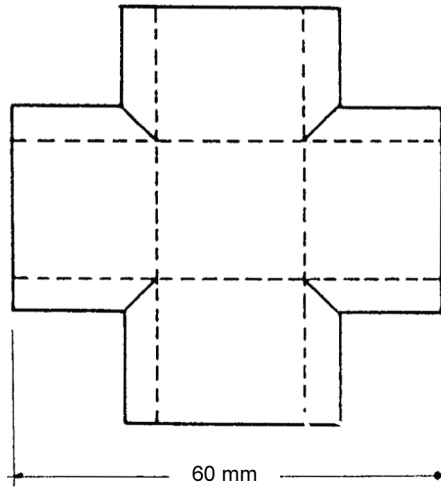
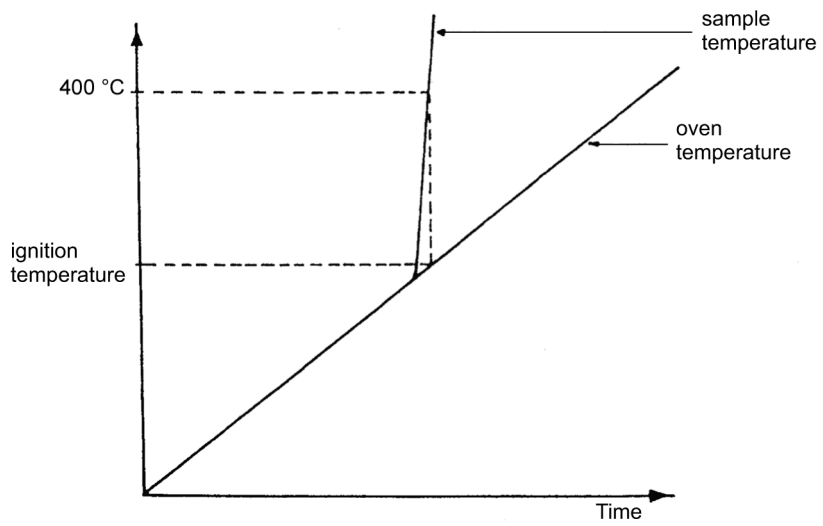


Figure 2

Typical temperature/time curve





**▼B****A.17. OXIDISING PROPERTIES (SOLIDS)****1. METHOD****1.1. INTRODUCTION**

It is useful to have preliminary information on any potentially explosive properties of the substance before performing this test.

This test is not applicable to liquids, gases, explosive or highly flammable substances, or organic peroxides.

This test need not be performed when examination of the structural formula establishes beyond reasonable doubt that the substance is incapable of reacting exothermically with a combustible material.

In order to ascertain if the test should be performed with special precautions, a preliminary test should be performed.

**1.2. DEFINITION AND UNITS**

Burning time: reaction time, in seconds, taken for the reaction zone to travel along a pile, following the procedure described in 1.6.

Burning rate: expressed in millimetres per second.

Maximum burning rate: the highest value of the burning rates obtained with mixtures containing 10 to 90 % by weight of oxidiser.

**1.3. REFERENCE SUBSTANCE**

Barium nitrate (analytical grade) is used as reference substance for the test and the preliminary test.

The reference mixture is that mixture of barium nitrate with powdered cellulose, prepared according to 1.6, which has the maximum burning rate (usually a mixture with 60 % barium nitrate by weight).

**1.4. PRINCIPLE OF THE METHOD**

A preliminary test is carried out in the interests of safety. No further testing is required when the preliminary test clearly indicates that the test substance has oxidising properties. When this is not the case, the substance should then be subject to the full test.

In the full test, the substance to be tested and a defined combustible substance will be mixed in various ratios. Each mixture is then formed into a pile and the pile is ignited at one end. The maximum burning rate determined is compared with the maximum burning rate of the reference mixture.

**1.5. QUALITY CRITERIA**

If required, any method of grinding and mixing is valid provided that the difference in the maximum rate of burning in the six separate tests differs from the arithmetic mean value by no more than 10 %.

**▼B**

## 1.6. DESCRIPTION OF THE METHOD

1.6.1. **Preparation**1.6.1.1. *Test substance*

Reduce the test sample to a particle size  $< 0,125$  mm using the following procedure: sieve the test substance, grind the remaining fraction, repeat the procedure until the whole test portion has passed the sieve.

Any grinding and sieving method satisfying the quality criteria may be used.

Before preparing the mixture the substance is dried at  $105$  °C, until constant weight is obtained. If the decomposition temperature of the substance to be tested is below  $105$  °C, the substance has to be dried at a suitable lower temperature.

1.6.1.2. *Combustible substance*

Powdered cellulose is used as a combustible substance. The cellulose should be a type used for thin-layer chromatography or column chromatography. A type with fibre-lengths of more than 85 % between  $0,020$  and  $0,075$  mm has proved to be suitable. The cellulose powder is passed through a sieve with a mesh-size of  $0,125$  mm. The same batch of cellulose is to be used throughout the test.

Before preparing the mixture, the powdered cellulose is dried at  $105$  °C until constant weight is obtained.

If wood-meal is used in the preliminary test, then prepare a soft-wood wood-meal by collecting the portion which passes through a sieve mesh of  $1,6$  mm, mix thoroughly, then dry at  $105$  °C for four hours in a layer not more than  $25$  mm thick. Cool and store in an air-tight container filled as full as practicable until required, preferably within 24 hours of drying.

1.6.1.3. *Ignition source*

A hot flame from a gas burner (minimum diameter  $5$  mm) should be used as the ignition source. If another ignition source is used (e.g. when testing in an inert atmosphere), the description and the justification should be reported.

1.6.2. **Performance of the test**

*Note:*

Mixtures of oxidisers with cellulose or wood-meal must be treated as potentially explosive and handled with due care.

1.6.2.1. *Preliminary test*

The dried substance is thoroughly mixed with the dried cellulose or wood-meal in the proportions 2 of test substance to 1 of cellulose or wood-meal by weight and the mixture is formed into a small cone-shaped pile of dimensions  $3,5$  cm (diameter of base)  $\times$   $2,5$  cm (height) by filling, without tamping, a cone-shaped former (e.g. a laboratory glass funnel with the stem plugged).

**▼B**

The pile is placed on a cool, non-combustible, non-porous and low heat-conducting base plate. The test should be carried out in a fume cupboard as in 1.6.2.2.

The ignition source is put in contact with the cone. The vigour and duration of the resultant reaction are observed and recorded.

The substance is to be considered as oxidising if the reaction is vigorous.

In any case where the result is open to doubt, it is then necessary to complete the full train test described below.

**1.6.2.2. Train test**

Prepare oxidiser cellulose-mixtures containing 10 to 90 % weight of oxidiser in 10 % increments. For borderline cases, intermediate oxidiser cellulose mixtures should be used to obtain the maximum burning rate more precisely.

The pile is formed by means of a mould. The mould is made of metal, has a length of 250 mm and a triangular cross-section with an inner height of 10 mm and an inner width of 20 mm. On both sides of the mould, in the longitudinal direction, two metal plates are mounted as lateral limitations which project 2 mm beyond the upper edge of the triangular cross-section (figure). This arrangement is loosely filled with a slight excess of mixture. After dropping the mould once from a height of 2 cm onto a solid surface, the remaining excess substance is scraped off with an obliquely positioned sheet. The lateral limitations are removed and the remaining powder is smoothed, using a roller. A non-combustible, non-porous and low heat-conducting base plate is then placed on the top of the mould, the apparatus inverted and the mould removed.

Arrange the pile across the draught in a fume cupboard.

The air-speed should be sufficient to prevent fumes escaping into the laboratory and should not be varied during the test. A draught screen should be erected around the apparatus.

Due to hygroscopicity of cellulose and of some substances to be tested, the test should be carried out as quickly as possible.

Ignite one end of the pile by touching with the flame.

Measure the time of reaction over a distance of 200 mm after the reaction zone has propagated an initial distance of 30 mm.

The test is performed with the reference substance and at least once with each one of the range of mixtures of the test substance with cellulose.

If the maximum burning rate is found to be significantly greater than that from the reference mixture, the test can be stopped; otherwise the test should be repeated five times for each of the three mixtures giving the fastest burning rate.

**▼B**

If the result is suspected of being a false positive, then the test should be repeated using an inert substance with a similar particle size, such as kieselguhr, in place of cellulose. Alternatively, the test substance cellulose mixture, having the fastest burning rate, should be retested in an inert atmosphere (< 2 % v/v oxygen content).

**2. DATA**

For safety reasons the maximum burning rate — not the mean value — shall be considered to be the characteristic oxidising property of the substance under test.

The highest value of burning rate within a run of six tests of a given mixture is relevant for evaluation.

Plot a graph of the highest value of burning rate for each mixture versus the oxidiser concentration. From the graph take the maximum burning rate.

The six measured values of burning rate within a run obtained from the mixture with the maximum burning rate must not differ from the arithmetic mean value by more than 10 %; otherwise the methods of grinding and mixing must be improved.

Compare the maximum burning rate obtained with the maximum burning rate of the reference mixture (see 1.3).

If tests are conducted in an inert atmosphere, the maximum reaction rate is compared with that from the reference mixture in an inert atmosphere.

**3. REPORT****3.1. TEST REPORT**

The test report shall, if possible, include the following information:

- the identity, composition, purity, moisture content etc. of the substance tested,
- any treatment of the test sample (e.g. grinding, drying),
- the ignition source used in the tests,
- the results of measurements,
- the mode of reaction (e.g. flash burning at the surface, burning through the whole mass, any information concerning the combustion products, etc.),
- all additional remarks relevant for the interpretation of results, including a description of the vigour (flaming, sparking, fuming, slow smouldering, etc.) and approximate duration produced in the preliminary safety/screening test for both test and reference substance,
- the results from tests with an inert substance, if any,
- the results from tests in an inert atmosphere, if any.

**▼B**

## 3.2. INTERPRETATION OF THE RESULT

A substance is to be considered as an oxidising substance when:

- (a) in the preliminary test, there is a vigorous reaction;
- (b) in the full test, the maximum burning rate of the mixtures tested is higher than or equal to the maximum burning rate of the reference mixture of cellulose and barium nitrate.

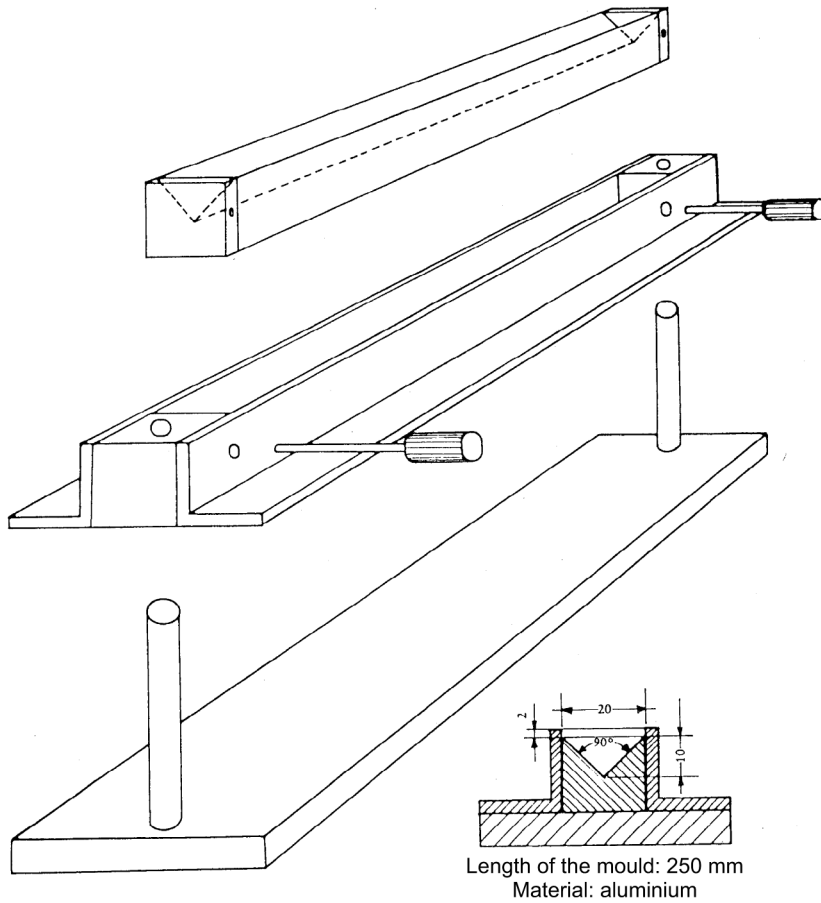
In order to avoid a false positive, the results obtained when testing the substance mixed with an inert material and/or when testing under an inert atmosphere should also be considered when interpreting the results.

## 4. REFERENCES

NF T 20-035 (September 85) Chemical products for industrial use. Determination of the oxidising properties of solids.

▼B*Appendix**Figure***Mould and accessories for the preparations of the pile**

(All dimensions in millimetres)



**▼ B****A.18. NUMBER-AVERAGE MOLECULAR WEIGHT AND MOLECULAR WEIGHT DISTRIBUTION OF POLYMERS****1. METHOD**

This Gel Permeation Chromatographic method is a replicate of the OECD TG 118 (1996). The fundamental principles and further technical information are given in reference (1).

**1.1. INTRODUCTION**

Since the properties of polymers are so varied, it is impossible to describe one single method setting out precisely the conditions for separation and evaluation which cover all eventualities and specificities occurring in the separation of polymers. In particular, complex polymer systems are often not amenable to gel permeation chromatography (GPC). When GPC is not practicable, the molecular weight may be determined by means of other methods (see Appendix). In such cases, full details and justification should be given for the method used.

The method described is based on DIN Standard 55672 (1). Detailed information about how to carry out the experiments and how to evaluate the data can be found in this DIN Standard. In case modifications of the experimental conditions are necessary, these changes must be justified. Other standards may be used, if fully referenced. The method described uses polystyrene samples of known polydispersity for calibration and it may have to be modified to be suitable for certain polymers, e.g. water soluble and long-chain branched polymers.

**1.2. DEFINITIONS AND UNITS**

The number-average molecular weight  $M_n$  and the weight average molecular weight  $M_w$  are determined using the following equations:

$$M_n = \frac{\sum_{i=1}^n H_i}{\sum_{i=1}^n H_i/M_i} \qquad M_w = \frac{\sum_{i=1}^n H_i \times M_i}{\sum_{i=1}^n H_i}$$

where,

$H_i$  is the level of the detector signal from the baseline for the retention volume  $V_i$ ,

$M_i$  is the molecular weight of the polymer fraction at the retention volume  $V_i$ , and

$n$  is the number of data points.

The breadth of the molecular weight distribution, which is a measure of the dispersity of the system, is given by the ratio  $M_w/M_n$ .

**▼ B**

## 1.3. REFERENCE SUBSTANCES

Since GPC is a relative method, calibration must be undertaken. Narrowly distributed, linearly constructed polystyrene standards with known average molecular weights  $M_n$  and  $M_w$  and a known molecular weight distribution are normally used for this. The calibration curve can only be used in the determination of the molecular weight of the unknown sample if the conditions for the separation of the sample and the standards have been selected in an identical manner.

A determined relationship between the molecular weight and elution volume is only valid under the specific conditions of the particular experiment. The conditions include, above all, the temperature, the solvent (or solvent mixture), the chromatography conditions and the separation column or system of columns.

The molecular weights of the sample determined in this way are relative values and are described as 'polystyrene equivalent molecular weights'. This means that dependent on the structural and chemical differences between the sample and the standards, the molecular weights can deviate from the absolute values to a greater or a lesser degree. If other standards are used, e.g. polyethylene glycol, polyethylene oxide, polymethyl methacrylate, polyacrylic acid, the reason should be stated.

## 1.4. PRINCIPLE OF THE TEST METHOD

Both the molecular weight distribution of the sample and the average molecular weights ( $M_n$ ,  $M_w$ ) can be determined using GPC. GPC is a special type of liquid chromatography in which the sample is separated according to the hydrodynamic volumes of the individual constituents (2).

Separation is effected as the sample passes through a column which is filled with a porous material, typically an organic gel. Small molecules can penetrate the pores whereas large molecules are excluded. The path of the large molecules is thereby shorter and these are eluted first. The medium-sized molecules penetrate some of the pores and are eluted later. The smallest molecules, with a mean hydrodynamic radius smaller than the pores of the gel, can penetrate all of the pores. These are eluted last.

In an ideal situation, the separation is governed entirely by the size of the molecular species, but in practice it is difficult to avoid at least some absorption effects interfering. Uneven column packing and dead volumes can worsen the situation (2).

Detection is effected by, e.g. refractive index or UV-absorption, and yields a simple distribution curve. However, to attribute actual molecular weight values to the curve, it is necessary to calibrate the column by passing down polymers of known molecular weight and, ideally, of broadly similar structure e.g. various polystyrene standards. Typically a Gaussian curve results, sometimes distorted by a small tail to the low molecular weight side, the vertical axis indicating the quantity, by weight, of the various molecular weight species eluted, and the horizontal axis the log molecular weight.



**▼ B**

## 1.5. QUALITY CRITERIA

The repeatability (Relative Standard Deviation: RSD) of the elution volume should be better than 0,3 %. The required repeatability of the analysis has to be ensured by correction via an internal standard if a chromatogram is evaluated time-dependently and does not correspond to the above mentioned criterion (1). The polydispersities are dependent on the molecular weights of the standards. In the case of polystyrene standards typical values are:

$M_p < 2\,000$	$M_w/M_n < 1,20$
$2\,000 \leq M_p \leq 10^6$	$M_w/M_n < 1,05$
$M_p > 10^6$	$M_w/M_n < 1,20$

( $M_p$  is the molecular weight of the standard at the peak maximum)

## 1.6. DESCRIPTION OF THE TEST METHOD

## 1.6.1. Preparation of the standard polystyrene solutions

The polystyrene standards are dissolved by careful mixing in the chosen eluent. The recommendations of the manufacturer must be taken into account in the preparation of the solutions.

The concentrations of the standards chosen are dependent on various factors, e.g. injection volume, viscosity of the solution and sensitivity of the analytical detector. The maximum injection volume must be adapted to the length of the column, in order to avoid overloading. Typical injection volumes for analytical separations using GPC with a column of 30 cm × 7,8 mm are normally between 40 and 100 µl. Higher volumes are possible, but they should not exceed 250 µl. The optimal ratio between the injection volume and the concentration must be determined prior to the actual calibration of the column.

## 1.6.2. Preparation of the sample solution

In principle, the same requirements apply to the preparation of the sample solutions. The sample is dissolved in a suitable solvent, e.g. tetrahydrofuran (THF), by shaking carefully. Under no circumstances should it be dissolved using an ultrasonic bath. When necessary, the sample solution is purified via a membrane filter with a pore size of between 0,2 and 2 µm.

The presence of undissolved particles must be recorded in the final report as these may be due to high molecular weight species. An appropriate method should be used to determine the percentage by weight of the undissolved particles. The solutions should be used within 24 hours.

## 1.6.3. Apparatus

- solvent reservoir,
- degasser (where appropriate),
- pump,

**▼B**

- pulse dampener (where appropriate),
- injection system,
- chromatography columns,
- detector,
- flowmeter (where appropriate),
- data recorder-processor,
- waste vessel.

It must be ensured that the GPC system is inert with regard to the utilised solvents (e.g. by the use of steel capillaries for THF solvent).

#### 1.6.4. **Injection and solvent delivery system**

A defined volume of the sample solution is loaded onto the column either using an auto-sampler or manually in a sharply defined zone. Withdrawing or depressing the plunger of the syringe too quickly, if done manually, can cause changes in the observed molecular weight distribution. The solvent-delivery system should, as far as possible, be pulsation-free ideally incorporating a pulse dampener. The flow rate is of the order of 1 ml/min.

#### 1.6.5. **Column**

Depending on the sample, the polymer is characterised using either a simple column or several columns connected in sequence. A number of porous column materials with defined properties (e.g. pore size, exclusion limits) are commercially available. Selection of the separation gel or the length of the column is dependent on both the properties of the sample (hydrodynamic volumes, molecular weight distribution) and the specific conditions for separation such as solvent, temperature and flow rate (1)(2)(3).

#### 1.6.6. **Theoretical plates**

The column or the combination of columns used for separation must be characterised by the number of theoretical plates. This involves, in the case of THF as elution solvent, loading a solution of ethyl benzene or other suitable non-polar solute onto a column of known length. The number of theoretical plates is given by the following equation:

$$N = 5,54 \left( \frac{V_e}{W_{1/2}} \right)^2 \quad \text{or} \quad N = 16 \left( \frac{V_e}{W} \right)^2$$

where,

N = the number of theoretical plates

$V_e$  = the elution volume at the peak maximum

**▼ B**

$W$  = the baseline peak width

$W_{1/2}$  = the peak width at half height

**1.6.7. Separation efficiency**

In addition to the number of theoretical plates, which is a quantity determining the bandwidth, a part is also played by the separation efficiency, this being determined by the steepness of the calibration curve. The separation efficiency of a column is obtained from the following relationship:

$$\frac{V_{e,M_x} - V_{e,(10M_x)}}{\text{cross sectional area of the column}} \geq 6,0 \left[ \frac{\text{cm}^3}{\text{cm}^2} \right]$$

where,

$V_{e, M_x}$  = the elution volume for polystyrene with the molecular weight  $M_x$

$V_{e,(10.M_x)}$  = the elution volume for polystyrene with a ten times greater molecular weight

The resolution of the system is commonly defined as follows:

$$R_{1,2} = 2 \times \frac{V_{e1} - V_{e2}}{W_1 + W_2} \times \frac{1}{\log_{10}(M_2/M_1)}$$

where,

$V_{e1}, V_{e2}$  = the elution volumes of the two polystyrene standards at the peak maximum

$W_1, W_2$  = the peak widths at the base-line

$M_1, M_2$  = the molecular weights at the peak maximum (should differ by a factor of 10)

The R-value for the column system should be greater than 1.7 (4).

**1.6.8. Solvents**

All solvents must be of high purity (for THF purity of 99,5 % is used). The solvent reservoir (if necessary in an inert gas atmosphere) must be sufficiently large for the calibration of the column and several sample analyses. The solvent must be degassed before it is transported to the column via the pump.

**1.6.9. Temperature control**

The temperature of the critical internal components (injection loop, columns, detector and tubing) should be constant and consistent with the choice of solvent.

**▼ B**1.6.10. **Detector**

The purpose of the detector is to record quantitatively the concentration of sample eluted from the column. In order to avoid unnecessary broadening of peaks the cuvette volume of the detector cell must be kept as small as possible. It should not be larger than 10 µl except for light scattering and viscosity detectors. Differential refractometry is usually used for detection. However, if required by the specific properties of the sample or the elution solvent, other types of detectors can be used, e.g. UV/VIS, IR, viscosity detectors, etc.

2. **DATA AND REPORTING**2.1. **DATA**

The DIN Standard (1) should be referred to for the detailed evaluation criteria as well as for the requirements relating to the collecting and processing of data.

For each sample, two independent experiments must be carried out. They have to be analysed individually.

$M_n$ ,  $M_w$ ,  $M_w/M_n$  and  $M_p$  must be provided for every measurement. It is necessary to indicate explicitly that the measured values are relative values equivalent to the molecular weights of the standard used.

After determination of the retention volumes or the retention times (possibly corrected using an internal standard),  $\log M_p$  values ( $M_p$  being the peak maxima of the calibration standard) are plotted against one of those quantities. At least two calibration points are necessary per molecular weight decade, and at least five measurement points are required for the total curve, which should cover the estimated molecular weight of the sample. The low molecular weight end-point of the calibration curve is defined by n-hexyl benzene or another suitable non-polar solute. The number average and the weight-average molecular weights are generally determined by means of electronic data processing, based on the formulas of section 1.2. In case manual digitisation is used, ASTM D 3536-91 can be consulted (3).

The distribution curve must be provided in the form of a table or as figure (differential frequency or sum percentages against  $\log M$ ). In the graphic representation, one molecular weight decade should be normally about 4 cm in width and the peak maximum should be about 8 cm in height. In the case of integral distribution curves the difference in the ordinate between 0 and 100 % should be about 10 cm.

2.2. **TEST REPORT**

The test report must include the following information:

2.2.1. **Test substance:**

— available information about test substance (identity, additives, impurities),

**▼ B**

- description of the treatment of the sample, observations, problems.

**2.2.2. Instrumentation:**

- reservoir of eluent, inert gas, degassing of the eluent, composition of the eluent, impurities,
- pump, pulse dampener, injection system,
- separation columns (manufacturer, all information about the characteristics of the columns, such as pore size, kind of separation material, etc., number, length and order of the columns used),
- number of the theoretical plates of the column (or combination), separation efficiency (resolution of the system),
- information on symmetry of the peaks,
- column temperature, kind of temperature control,
- detector (measurement principle, type, cuvette volume),
- flowmeter if used (manufacturer, measurement principle),
- system to record and process data (hardware and software).

**2.2.3. Calibration of the system:**

- detailed description of the method used to construct the calibration curve,
- information about quality criteria for this method (e.g. correlation coefficient, error sum of squares, etc.),
- information about all extrapolations, assumptions and approximations made during the experimental procedure and the evaluation and processing of data,
- all measurements used for constructing the calibration curve have to be documented in a table which includes the following information for each calibration point:
  - name of the sample,
  - manufacturer of the sample,
  - characteristic values of the standards  $M_p$ ,  $M_n$ ,  $M_w$ ,  $M_w/M_n$ , as provided by the manufacturer or derived by subsequent measurements, together with details about the method of determination,
  - injection volume and injection concentration,

**▼ B**

- $M_p$  value used for calibration,
- elution volume or corrected retention time measured at the peak maxima,
- $M_p$  calculated at the peak maximum,
- percentage error of the calculated  $M_p$  and the calibration value.

**2.2.4. Evaluation:**

- evaluation on a time basis: methods used to ensure the required reproducibility (method of correction, internal standard, etc.),
- information about whether the evaluation was effected on the basis of the elution volume or the retention time,
- information about the limits of the evaluation if a peak is not completely analysed,
- description of smoothing methods, if used,
- preparation and pre-treatment procedures of the sample,
- the presence of undissolved particles, if any,
- injection volume ( $\mu\text{l}$ ) and injection concentration ( $\text{mg/ml}$ ),
- observations indicating effects which lead to deviations from the ideal GPC profile,
- detailed description of all modifications in the testing procedures,
- details of the error ranges,
- any other information and observations relevant for the interpretation of the results.

**3. REFERENCES**

- (1) DIN 55672(1995) Gelpermeationschromatographie (GPC) mit Tetrahydrofuran (THF) als Elutionsmittel, Teil 1.
- (2) Yau, W.W., Kirkland, J.J., and Bly, D.D. eds., (1979) Modern Size Exclusion Liquid Chromatography, J. Wiley and Sons.
- (3) ASTM D 3536-91, (1991). Standard Test Method for Molecular Weight Averages and Molecular Weight Distribution by Liquid Exclusion Chromatography (Gel Permeation Chromatography-GPC) American Society for Testing and Materials, Philadelphia, Pennsylvania.
- (4) ASTM D 5296-92, (1992) Standard Test Method for Molecular Weight Averages and Molecular Weight Distribution of Polystyrene by High Performance Size-Exclusion Chromatography. American Society for Testing and Materials, Philadelphia, Pennsylvania.

**▼ B***Appendix***Examples of other methods for determination of number average molecular weight ( $M_n$ ) for polymers**

Gel permeation chromatography (GPC) is the preferred method for determination of  $M_n$ , especially when a set of standards are available, whose structure are comparable with the polymer structure. However, where there are practical difficulties in using GPC or there is already an expectation that the substance will fail a regulatory  $M_n$  criterion (and which needs confirming), alternative methods are available, such as:

**1. Use of colligative properties****1.1. Ebullioscopy/Cryoscopy**

involves measurement of boiling point elevation (ebullioscopy) or freezing point depression (cryoscopy) of a solvent, when the polymer is added. The method relies on the fact that the effect of the dissolved polymer on the boiling/freezing point of the liquid is dependent on the molecular weight of the polymer (1) (2).

Applicability,  $M_n < 20\ 000$ .

**1.2. Lowering of vapour pressure**

involves the measurement of the vapour pressure of a chosen reference liquid before and after the addition of known quantities of polymer (1) (2).

Applicability,  $M_n < 20\ 000$  (theoretically; in practice however of limited value).

**1.3 Membrane osmometry**

relies on the principle of osmosis, i.e. the natural tendency of solvent molecules to pass through a semi-permeable membrane from a dilute to a concentrated solution to achieve equilibrium. In the test, the dilute solution is at zero concentration, whereas the concentrated solution contains the polymer. The effect of drawing solvent through the membrane causes a pressure differential that is dependent on the concentration and the molecular weight of the polymer (1) (3) (4).

Applicability,  $M_n$  between 20 000 - 200 000.

**1.4 Vapour phase osmometry**

involves comparison of the rate of evaporation of a pure solvent aerosol to at least three aerosols containing the polymer at different concentrations (1)(2)(4).

Applicability,  $M_n < 20\ 000$ .

**▼ B****2. End-group analysis**

To use this method, knowledge of both the overall structure of the polymer and the nature of the chain terminating end groups is needed (which must be distinguishable from the main skeleton by, e.g. NMR or titration/derivatisation). The determination of the molecular concentration of the end groups present on the polymer can lead to a value for the molecular weight (7) (8) (9).

Applicability,  $M_n$  up to 50 000 (with decreasing reliability).

**3. References**

- (1) Billmeyer, F.W. Jr., (1984) Textbook of Polymer Science, 3rd Edn., John Wiley, New York.
- (2) Glover, C.A., (1975) Absolute Colligative Property Methods. Chapter 4. In: Polymer Molecular Weights, Part I P.E. Slade, Jr. ed., Marcel Dekker, New York.
- (3) ASTM D 3750-79, (1979) Standard Practice for Determination of Number-Average Molecular Weight of Polymers by Membrane Osmometry. American Society for Testing and Materials, Philadelphia, Pennsylvania.
- (4) Coll, H. (1989) Membrane Osmometry. In: Determination of Molecular Weight, A.R. Cooper ed., J. Wiley and Sons, pp. 25-52.
- (5) ASTM 3592-77, (1977) Standard Recommended Practice for Determination of Molecular Weight by Vapour Pressure, American Society for Testing and Materials, Philadelphia, Pennsylvania.
- (6) Morris, C.E.M., (1989) Vapour Pressure Osmometry. In: Determination of Molecular Weight, A.R. Cooper ed., John Wiley and Sons.
- (7) Schröder, E., Müller, G., and Arndt, K-F., (1989) Polymer Characterisation, Carl Hanser Verlag, Munich.
- (8) Garmon, R.G., (1975) End-Group Determinations, Chapter 3 In: Polymer Molecular Weights, Part I, P.E. Slade, Jr. ed., Marcel Dekker, New York.
- (9) Amiya, S., et al. (1990) Pure and Applied Chemistry, 62, 2139-2146.



**▼ B****A.19. LOW MOLECULAR WEIGHT CONTENT OF POLYMERS****1. METHOD**

This Gel Permeation Chromatographic method is a replicate of the OECD TG 119 (1996). The fundamental principles and further technical information are given in the references.

**1.1. INTRODUCTION**

Since the properties of polymers are so varied, it is impossible to describe one single method setting out precisely the conditions for separation and evaluation which cover all eventualities and specificities occurring in the separation of polymers. In particular, complex polymer systems are often not amenable to gel permeation chromatography (GPC). When GPC is not practicable, the molecular weight may be determined by means of other methods (see Appendix). In such cases, full details and justification should be given for the method used.

The method described is based on DIN Standard 55672 (1). Detailed information about how to carry out the experiments and how to evaluate the data can be found in this DIN Standard. In case modifications of the experimental conditions are necessary, these changes must be justified. Other standards may be used, if fully referenced. The method described uses polystyrene samples of known polydispersity for calibration and it may have to be modified to be suitable for certain polymers, e.g. water soluble and long-chain branched polymers.

**1.2. DEFINITIONS AND UNITS**

Low molecular weight is arbitrarily defined as a molecular weight below 1 000 dalton.

The number-average molecular weight  $M_n$  and the weight average molecular weight  $M_w$  are determined using the following equations:

$$M_n = \frac{\sum_{i=1}^n H_i}{\sum_{i=1}^n H_i/M_i} \quad M_w = \frac{\sum_{i=1}^n H_i \times M_i}{\sum_{i=1}^n H_i}$$

where,

$H_i$  = the level of the detector signal from the baseline for the retention volume  $V_i$ ,

$M_i$  = the molecular weight of the polymer fraction at the retention volume  $V_i$ , and  $n$  is the number of data points

The breadth of the molecular weight distribution, which is a measure of the dispersity of the system, is given by the ratio  $M_w/M_n$ .

**▼ B**

## 1.3. REFERENCE SUBSTANCES

Since GPC is a relative method, calibration must be undertaken. Narrowly distributed, linearly constructed polystyrene standards with known average molecular weights  $M_n$  and  $M_w$  and a known molecular weight distribution are normally used for this. The calibration curve can only be used in the determination of the molecular weight of the unknown sample if the conditions for the separation of the sample and the standards have been selected in an identical manner.

A determined relationship between the molecular weight and elution volume is only valid under the specific conditions of the particular experiment. The conditions include, above all, the temperature, the solvent (or solvent mixture), the chromatography conditions and the separation column or system of columns.

The molecular weights of the sample determined in this way are relative values and are described as 'polystyrene equivalent molecular weights'. This means that dependent on the structural and chemical differences between the sample and the standards, the molecular weights can deviate from the absolute values to a greater or a lesser degree. If other standards are used, e.g. polyethylene glycol, polyethylene oxide, polymethyl methacrylate, polyacrylic acid, the reason should be stated.

## 1.4. PRINCIPLE OF THE TEST METHOD

Both the molecular weight distribution of the sample and the average molecular weights ( $M_n$ ,  $M_w$ ) can be determined using GPC. GPC is a special type of liquid chromatography in which the sample is separated according to the hydrodynamic volumes of the individual constituents (2).

Separation is effected as the sample passes through a column which is filled with a porous material, typically an organic gel. Small molecules can penetrate the pores whereas large molecules are excluded. The path of the large molecules is thereby shorter and these are eluted first. The medium-sized molecules penetrate some of the pores and are eluted later. The smallest molecules, with a mean hydrodynamic radius smaller than the pores of the gel, can penetrate all of the pores. These are eluted last.

In an ideal situation, the separation is governed entirely by the size of the molecular species, but in practice it is difficult to avoid at least some absorption effects interfering. Uneven column packing and dead volumes can worsen the situation (2).

Detection is effected by e.g. refractive index or UV-absorption and yields a simple distribution curve. However, to attribute actual molecular weight values to the curve, it is necessary to calibrate the column by passing down polymers of known molecular weight and, ideally, of broadly similar structure, e.g. various polystyrene standards. Typically a Gaussian curve results, sometimes distorted by a small tail to the low molecular weight side, the vertical axis indicating the quantity, by weight, of the various molecular weight species eluted, and the horizontal axis the log molecular weight.

**▼ B**

The low molecular weight content is derived from this curve. The calculation can only be accurate if the low molecular weight species respond equivalently on a per mass basis to the polymer as a whole.

1.5. **QUALITY CRITERIA**

The repeatability (Relative Standard Deviation: RSD) of the elution volume should be better than 0,3 %. The required repeatability of the analysis has to be ensured by correction via an internal standard if a chromatogram is evaluated time-dependently and does not correspond to the above mentioned criterion (1). The polydispersities are dependent on the molecular weights of the standards. In the case of polystyrene standards typical values are:

$M_p < 2\,000$	$M_w/M_n < 1,20$
$2\,000 \leq M_p \leq 10^6$	$M_w/M_n < 1,05$
$M_p > 10^6$	$M_w/M_n < 1,20$

( $M_p$  is the molecular weight of the standard at the peak maximum)

1.6. **DESCRIPTION OF THE TEST METHOD**1.6.1. **Preparation of the standard polystyrene solutions**

The polystyrene standards are dissolved by careful mixing in the chosen eluent. The recommendations of the manufacturer must be taken into account in the preparation of the solutions.

The concentrations of the standards chosen are dependent on various factors, e.g. injection volume, viscosity of the solution and sensitivity of the analytical detector. The maximum injection volume must be adapted to the length of the column, in order to avoid overloading. Typical injection volumes for analytical separations using GPC with a column of 30 cm × 7,8 mm are normally between 40 and 100 µl. Higher volumes are possible, but they should not exceed 250 µl. The optimal ratio between the injection volume and the concentration must be determined prior to the actual calibration of the column.

1.6.2. **Preparation of the sample solution**

In principle, the same requirements apply to the preparation of the sample solutions. The sample is dissolved in a suitable solvent, e.g. tetrahydrofuran (THF), by shaking carefully. Under no circumstances should it be dissolved using an ultrasonic bath. When necessary, the sample solution is purified via a membrane filter with a pore size of between 0,2 and 2 µm.

The presence of undissolved particles must be recorded in the final report as these may be due to high molecular weight species. An appropriate method should be used to determine the percentage by weight of the undissolved particles. The solutions should be used within 24 hours.

**▼B****1.6.3. Correction for content of impurities and additives**

Correction of the content of species of  $M < 1\,000$  for the contribution from non-polymer specific components present (e.g. impurities and/or additives) is usually necessary, unless the measured content is already  $< 1\%$ . This is achieved by direct analysis of the polymer solution or the GPC eluate.

In cases where the eluate, after passage through the column, is too dilute for a further analysis it must be concentrated. It may be necessary to evaporate the eluate to dryness and dissolve it again. Concentration of the eluate must be effected under conditions which ensure that no changes occur in the eluate. The treatment of the eluate after the GPC step is dependent on the analytical method used for the quantitative determination.

**1.6.4. Apparatus**

GPC apparatus comprises the following components:

- solvent reservoir,
- degasser (where appropriate),
- pump,
- pulse dampener (where appropriate),
- injection system,
- chromatography columns,
- detector,
- flowmeter (where appropriate),
- data recorder-processor,
- waste vessel.

It must be ensured that the GPC system is inert with regard to the utilised solvents (e.g. by the use of steel capillaries for THF solvent).

**1.6.5. Injection and solvent delivery system**

A defined volume of the sample solution is loaded onto the column either using an auto-sampler or manually in a sharply defined zone. Withdrawing or depressing the plunger of the syringe too quickly, if done manually, can cause changes in the observed molecular weight distribution. The solvent-delivery system should, as far as possible, be pulsation-free ideally incorporating a pulse dampener. The flow rate is of the order of 1 ml/min.

**1.6.6. Column**

Depending on the sample, the polymer is characterised using either a simple column or several columns connected in sequence. A number of porous column materials with defined properties (e.g. pore size, exclusion limits) are commercially available. Selection of the separation gel or the length of the column is dependent on both the properties of the sample (hydrodynamic volumes, molecular weight distribution) and the specific conditions for separation such as solvent, temperature and flow rate (1) (2) (3).

**▼ B****1.6.7. Theoretical plates**

The column or the combination of columns used for separation must be characterised by the number of theoretical plates. This involves, in the case of THF as elution solvent, loading a solution of ethyl benzene or other suitable non-polar solute onto a column of known length. The number of theoretical plates is given by the following equation:

$$N = 5,54 \left( \frac{V_e}{W_{1/2}} \right)^2 \quad \text{or} \quad N = 16 \left( \frac{V_e}{W} \right)^2$$

where,

$N$  = the number of theoretical plates

$V_e$  = the elution volume at the peak maximum

$W$  = the baseline peak width

$W_{1/2}$  = the peak width at half height

**1.6.8. Separation efficiency**

In addition to the number of theoretical plates, which is a quantity determining the bandwidth, a part is also played by the separation efficiency, this being determined by the steepness of the calibration curve. The separation efficiency of a column is obtained from the following relationship:

$$\frac{V_{e,M_x} - V_{e,(10M_x)}}{\text{cross sectional area of the column}} \geq 6,0 \left[ \frac{\text{cm}^3}{\text{cm}^2} \right]$$

where,

$V_{e, M_x}$  = the elution volume for polystyrene with the molecular weight  $M_x$

$V_{e,(10.M_x)}$  = the elution volume for polystyrene with a ten times greater molecular weight

The resolution of the system is commonly defined as follows:

$$R_{1,2} = 2 \times \frac{V_{e1} - V_{e2}}{W_1 + W_2} \times \frac{1}{\log_{10}(M_2/M_1)}$$

where,

$V_{e1}, V_{e2}$  = the elution volumes of the two polystyrene standards at the peak maximum

$W_1, W_2$  = the peak widths at the base-line

$M_1, M_2$  = the molecular weights at the peak maximum (should differ by a factor of 10).

The R-value for the column system should be greater than 1,7 (4).

**▼ B****1.6.9. Solvents**

All solvents must be of high purity (for THF purity of 99,5 % is used). The solvent reservoir (if necessary in an inert gas atmosphere) must be sufficiently large for the calibration of the column and several sample analyses. The solvent must be degassed before it is transported to the column via the pump.

**1.6.10. Temperature control**

The temperature of the critical internal components (injection loop, columns, detector and tubing) should be constant and consistent with the choice of solvent.

**1.6.11. Detector**

The purpose of the detector is to record quantitatively the concentration of sample eluted from the column. In order to avoid unnecessary broadening of peaks the cuvette volume of the detector cell must be kept as small as possible. It should not be larger than 10  $\mu\text{l}$  except for light scattering and viscosity detectors. Differential refractometry is usually used for detection. However, if required by the specific properties of the sample or the elution solvent, other types of detectors can be used, e.g. UV/VIS, IR, viscosity detectors, etc.

**2. DATA AND REPORTING****2.1. DATA**

The DIN Standard (1) should be referred to for the detailed evaluation criteria as well as for the requirements relating to the collecting and processing of data.

For each sample, two independent experiments must be carried out. They have to be analysed individually. In all cases it is essential to determine also data from blanks, treated under the same conditions as the sample.

It is necessary to indicate explicitly that the measured values are relative values equivalent to the molecular weights of the standard used.

After determination of the retention volumes or the retention times (possibly corrected using an internal standard),  $\log M_p$  values ( $M_p$  being the peak maxima of the calibration standard) are plotted against one of those quantities. At least two calibration points are necessary per molecular weight decade, and at least five measurement points are required for the total curve, which should cover the estimated molecular weight of the sample. The low molecular weight end-point of the calibration curve is defined by n-hexyl benzene or another suitable non-polar solute. The portion of the curve corresponding to molecular weights below 1 000 is determined and corrected as necessary for impurities and additives. The elution curves are generally evaluated by means of electronic data processing. In case manual digitisation is used, ASTM D 3536-91 can be consulted (3).

**▼ B**

If any insoluble polymer is retained on the column, its molecular weight is likely to be higher than that of the soluble fraction, and if not considered would result in an overestimation of the low molecular weight content. Guidance for correcting the low molecular weight content for insoluble polymer is provided in the Appendix.

The distribution curve must be provided in the form of a table or as figure (differential frequency or sum percentages against log M). In the graphic representation, one molecular weight decade should be normally about 4 cm in width and the peak maximum should be about 8 cm in height. In the case of integral distribution curves the difference in the ordinate between 0 and 100 % should be about 10 cm.

**2.2. TEST REPORT**

The test report must include the following information:

**2.2.1. Test substance:**

- available information about test substance (identity, additives, impurities),
- description of the treatment of the sample, observations, problems.

**2.2.2. Instrumentation:**

- reservoir of eluent, inert gas, degassing of the eluent, composition of the eluent, impurities,
- pump, pulse dampener, injection system,
- separation columns (manufacturer, all information about the characteristics of the columns, such as pore size, kind of separation material, etc., number, length and order of the columns used),
- number of the theoretical plates of the column (or combination), separation efficiency (resolution of the system),
- information on symmetry of the peaks,
- column temperature, kind of temperature control,
- detector (measurement principle, type, cuvette volume),
- flowmeter if used (manufacturer, measurement principle),
- system to record and process data (hardware and software).

**2.2.3. Calibration of the system:**

- detailed description of the method used to construct the calibration curve,

**▼ B**

- information about quality criteria for this method (e.g. correlation coefficient, error sum of squares, etc.),
- information about all extrapolations, assumptions and approximations made during the experimental procedure and the evaluation and processing of data,
- all measurements used for constructing the calibration curve have to be documented in a table which includes the following information for each calibration point:
  - name of the sample,
  - manufacturer of the sample,
  - characteristic values of the standards  $M_p$ ,  $M_n$ ,  $M_w$ ,  $M_w/M_n$ , as provided by the manufacturer or derived by subsequent measurements, together with details about the method of determination,
  - injection volume and injection concentration,
  - $M_p$  value used for calibration,
  - elution volume or corrected retention time measured at the peak maxima,
  - $M_p$  calculated at the peak maximum,
  - percentage error of the calculated  $M_p$  and the calibration value.

**2.2.4. Information on the low molecular weight polymer content:**

- description of the methods used in the analysis and the way in which the experiments were conducted,
- information about the percentage of the low molecular weight species content (w/w) related to the total sample,
- information about impurities, additives and other non-polymer species in percentage by weight related to the total sample.

**2.2.5. Evaluation:**

- evaluation on a time basis: all methods to ensure the required reproducibility (method of correction, internal standard etc.),
- information about whether the evaluation was effected on the basis of the elution volume or the retention time,
- information about the limits of the evaluation if a peak is not completely analysed,
- description of smoothing methods, if used,
- preparation and pre-treatment procedures of the sample,
- the presence of undissolved particles, if any,



**▼B**

- injection volume ( $\mu\text{l}$ ) and injection concentration ( $\text{mg/ml}$ ),
- observations indicating effects which lead to deviations from the ideal GPC profile,
- detailed description of all modifications in the testing procedures,
- details of the error ranges,
- any other information and observations relevant for the interpretation of the results.

**3. REFERENCES**

- (1) DIN 55672 (1995) Gelpermeationschromatographie (GPC) mit Tetrahydrofuran (THF) als Elutionsmittel, Teil 1.
- (2) Yau, W.W., Kirkland, J.J., and Bly, D.D. eds. (1979) Modern Size Exclusion Liquid Chromatography, J. Wiley and Sons.
- (3) ASTM D 3536-91, (1991) Standard Test method for Molecular Weight Averages and Molecular Weight Distribution by Liquid Exclusion Chromatography (Gel Permeation Chromatography-GPC). American Society for Testing and Materials, Philadelphia, Pennsylvania.
- (4) ASTM D 5296-92, (1992) Standard Test method for Molecular Weight Averages and Molecular Weight Distribution of Polystyrene by High Performance Size-Exclusion Chromatography. American Society for Testing and Materials, Philadelphia, Pennsylvania.

**▼ B***Appendix***Guidance for correcting low molecular content for the presence of insoluble polymer**

When insoluble polymer is present in a sample, it results in mass loss during the GPC analysis. The insoluble polymer is irreversibly retained on the column or sample filter while the soluble portion of the sample passes through the column. In the case where the refractive index increment ( $dn/dc$ ) of the polymer can be estimated or measured, one can estimate the sample mass lost on the column. In that case, one makes a correction using an external calibration with standard materials of known concentration and  $dn/dc$  to calibrate the response of the refractometer. In the example hereafter a poly(methyl methacrylate) (pMMA) standard is used.

In the external calibration for analysis of acrylic polymers, a pMMA standard of known concentration in tetrahydrofuran, is analysed by GPC and the resulting data are used to find the refractometer constant according to the equation:

$$K = R/(C \times V \times dn/dc)$$

where:

K = the refractometer constant (in microvolt second/ml),

R = the response of the pMMA standard (in microvolt/second),

C = the concentration of the pMMA standard (in mg/ml),

V = the injection volume (in ml), and

$dn/dc$  = the refractive index increment for pMMA in tetrahydrofuran (in ml/mg).

The following data are typical for a pMMA standard:

R = 2 937 891

C = 1,07 mg/ml

V = 0,1 ml

$dn/dc$  =  $9 \times 10^{-5}$  ml/mg

The resulting K value,  $3,05 \times 10^{11}$  is then used to calculate the theoretical detector response if 100 % of the polymer injected had eluted through the detector.

**▼B****A.20. SOLUTION/EXTRACTION BEHAVIOUR OF POLYMERS IN WATER****1. METHOD**

The method described is a replicate of the revised version of OECD TG 120 (1997). Further technical information is given in reference (1).

**1.1. INTRODUCTION**

For certain polymers, such as emulsion polymers, initial preparatory work may be necessary before the method set out hereafter can be used. The method is not applicable to liquid polymers and to polymers that react with water under the test conditions.

When the method is not practical or not possible, the solution/extraction behaviour may be investigated by means of other methods. In such cases, full details and justification should be given for the method used.

**1.2. REFERENCE SUBSTANCES**

None.

**1.3. PRINCIPLE OF THE TEST METHOD**

The solution/extraction behaviour of polymers in an aqueous medium is determined using the flask method (see A.6 Water Solubility, Flask method) with the modifications described below.

**1.4. QUALITY CRITERIA**

None.

**1.5. DESCRIPTION OF THE TEST METHOD****1.5.1. Equipment**

The following equipment is required for the method:

- crushing device, e.g. grinder for the production of particles of known size,
- apparatus for shaking with possibility of temperature control,
- membrane filter system,
- appropriate analytical equipment,
- standardised sieves.

**1.5.2. Sample preparation**

A representative sample has first to be reduced to a particle size between 0,125 and 0,25 mm using appropriate sieves. Cooling may be required for the stability of the sample or for the grinding process. Materials of a rubbery nature can be crushed at liquid nitrogen temperature (1).

If the required particle size fraction is not attainable, action should be taken to reduce the particle size as much as possible, and the result reported. In the report, it is necessary to indicate the way in which the crushed sample was stored prior to the test

**▼B****1.5.3. Procedure**

Three samples of 10 g of the test substance are weighed into each of three vessels fitted with glass stoppers and 1 000 ml of water is added to each vessel. If handling an amount of 10 g polymer proves impracticable, the next highest amount which can be handled should be used and the volume of water adjusted accordingly.

The vessels are tightly stoppered and then agitated at 20 °C. A shaking or stirring device capable of operating at constant temperature should be used. After a period of 24 hours, the content of each vessel is centrifuged or filtered and the concentration of polymer in the clear aqueous phase is determined by a suitable analytical method. If suitable analytical methods for the aqueous phase are not available, the total solubility/extractivity can be estimated from the dry weight of the filter residue or centrifuged precipitate.

It is usually necessary to differentiate quantitatively between the impurities and additives on the one hand and the low molecular weight species on the other hand. In the case of gravimetric determination, it is also important to perform a blank run using no test substance in order to account for residues arising from the experimental procedure.

The solution/extraction behaviour of polymers in water at 37 °C at pH 2 and pH 9 may be determined in the same way as described for the conduct of the experiment at 20 °C. The pH values can be achieved by the addition of either suitable buffers or appropriate acids or bases such as hydrochloric acid, acetic acid, analytical grade sodium or potassium hydroxide or NH<sub>3</sub>.

Depending on the method of analysis used, one or two tests should be performed. When sufficiently specific methods are available for direct analysis of the aqueous phase for the polymer component, one test as described above should suffice. However, when such methods are not available and determination of the solution/extraction behaviour of the polymer is limited to indirect analysis by determining only the total organic carbon content (TOC) of the aqueous extract, an additional test should be conducted. This additional test should also be done in triplicate, using ten times smaller polymer samples and the same amounts of water as those used in the first test.

**1.5.4. Analysis****1.5.4.1. Test conducted with one sample size**

Methods may be available for direct analysis of polymer components in the aqueous phase. Alternatively, indirect analysis of dissolved/-extracted polymer components, by determining the total content of soluble parts and correcting for non polymer-specific components, could also be considered.

Analysis of the aqueous phase for the total polymeric species is possible:

either by a sufficiently sensitive method, e.g.:

— TOC using persulphate or dichromate digestion to yield CO<sub>2</sub> followed by estimation by IR or chemical analysis,

**▼ B**

- Atomic Absorption Spectrometry (AAS) or its Inductively Coupled Plasma (ICP) emission equivalent for silicon or metal containing polymers,
- UV absorption or spectrofluorimetry for aryl polymers,
- LC-MS for low molecular weight samples,

or by vacuum evaporation to dryness of the aqueous extract and spectroscopic (IR, UV, etc.) or AAS/ICP analysis of the residue.

If analysis of the aqueous phase as such is not practicable, the aqueous extract should be extracted with a water-immiscible organic solvent e.g. a chlorinated hydrocarbon. The solvent is then evaporated and the residue analysed as above for the notified polymer content. Any components in this residue which are identified as being impurities or additives are to be subtracted for the purpose of determining the degree of solution/extraction of the polymer itself.

When relatively large quantities of such materials are present, it may be necessary to subject the residue to e.g. HPLC or GC analysis to differentiate the impurities from the monomer and monomer-derived species present so that the true content of the latter can be determined.

In some cases, simple evaporation of the organic solvent to dryness and weighing the dry residue may be sufficient.

1.5.4.2. *Test conducted with two different sample sizes*

All aqueous extracts are analysed for TOC.

A gravimetric determination is performed on the undissolved/not extracted part of the sample. If, after centrifugation or filtering of the content of each vessel, polymer residues remain attached to the wall of the vessel, the vessel should be rinsed with the filtrate until the vessel is cleared from all visible residues. Following which, the filtrate is again centrifuged or filtered. The residues remaining on the filter or in the centrifuge tube are dried at 40 °C under vacuum and weighed. Drying is continued until a constant weight is reached.

## 2. DATA

### 2.1. TEST CONDUCTED WITH ONE SAMPLE SIZE

The individual results for each of the three flasks and the average values should be given and expressed in units of mass per volume of the solution (typically mg/l) or mass per mass of polymer sample (typically mg/g). Additionally, the weight loss of the sample (calculated as the weight of the solute divided by the weight of the initial sample) should also be given. The relative standard deviations (RSD) should be calculated. Individual figures should be given for the total substance (polymer + essential additives, etc.) and for the polymer only (i.e. after subtracting the contribution from such additives).

**▼ B****2.2. TEST CONDUCTED WITH TWO DIFFERENT SAMPLE SIZES**

The individual TOC values of the aqueous extracts of the two triplicate experiments and the average value for each experiment should be given expressed as units of mass per volume of solution (typically mgC/l), as well as in units of mass per weight of the initial sample (typically mgC/g).

If there is no difference between the results at the high and the low sample/water ratios, this may indicate that all extractable components were indeed extracted. In such a case, direct analysis would normally not be necessary.

The individual weights of the residues should be given and expressed in percentage of the initial weights of the samples. Averages should be calculated per experiment. The differences between 100 and the percentages found represent the percentages of soluble and extractable material in the original sample.

**3. REPORTING****3.1. TEST REPORT**

The test report must include the following information:

**3.1.1. Test substance:**

— available information about test substance (identity, additives, impurities, content of low molecular weight species).

**3.1.2. Experimental conditions:**

— description of the procedures used and experimental conditions,  
— description of the analytical and detection methods.

**3.1.3. Results:**

— results of solubility/extractivity in mg/l; individual and mean values for the extraction tests in the various solutions, broken down in polymer content and impurities, additives, etc.,  
— results of solubility/extractivity in mg/g of polymer,  
— TOC values of aqueous extracts, weight of the solute and calculated percentages, if measured,  
— the pH of each sample,  
— information about the blank values,  
— where necessary, references to the chemical instability of the test substance, during both the testing process and the analytical process,  
— all information which is important for the interpretation of the results.

**4. REFERENCES**

(1) DIN 53733 (1976) Zerkleinerung von Kunststoffzeugnissen für Prüfzwecke.

**▼B****A.21. OXIDISING PROPERTIES (LIQUIDS)****1. METHOD****1.1. INTRODUCTION**

This test method is designed to measure the potential for a liquid substance to increase the burning rate or burning intensity of a combustible substance, or to form a mixture with a combustible substance which spontaneously ignites, when the two are thoroughly mixed. It is based on the UN test for oxidising liquids (1) and is equivalent to it. However, as this method A.21 is primarily designed to satisfy the requirements of Regulation (EC) No 1907/2006, comparison with only one reference substance is required. Testing and comparison to additional reference substances may be necessary when the results of the test are expected to be used for other purposes. (1)

This test need not be performed when examination of the structural formula establishes beyond reasonable doubt that the substance is incapable of reacting exothermically with a combustible material.

It is useful to have preliminary information on any potential explosive properties of the substance before performing this test.

This test is not applicable to solids, gases, explosive or highly flammable substances, or organic peroxides.

This test may not need to be performed when results for the test substance in the UN test for oxidising liquids (1) are already available.

**1.2. DEFINITIONS AND UNITS**

Mean pressure rise time is the mean of the measured times for a mixture under test to produce a pressure rise from 690 kPa to 2 070 kPa above atmospheric.

**1.3. REFERENCE SUBSTANCE**

65 % (w/w) aqueous nitric acid (analytical grade) is required as a reference substance. (2)

(1) As, for example, in the framework of UN transport regulations.

(2) The acid should be titrated before testing to confirm its concentration.

**▼B**

Optionally, if the experimenter foresees that the results of this test may eventually be used for other purposes <sup>(1)</sup>, testing of additional reference substances may also be appropriate. <sup>(2)</sup>

**1.4. PRINCIPLE OF THE TEST METHOD**

The liquid to be tested is mixed in a 1 to 1 ratio, by mass, with fibrous cellulose and introduced into a pressure vessel. If during mixing or filling spontaneous ignition occurs, no further testing is necessary.

If spontaneous ignition does not occur the full test is carried out. The mixture is heated in a pressure vessel and the mean time taken for the pressure to rise from 690 kPa to 2 070 kPa above atmospheric is determined. This is compared with the mean pressure rise time for the 1:1 mixture of the reference substance(s) and cellulose.

**1.5. QUALITY CRITERIA**

In a series of five trials on a single substance no results should differ by more than 30 % from the arithmetic mean. Results that differ by more than 30 % from the mean should be discarded, the mixing and filling procedure improved and the testing repeated.

**1.6. DESCRIPTION OF THE METHOD****1.6.1. Preparation****1.6.1.1. Combustible substance**

Dried, fibrous cellulose with a fibre length between 50 and 250 µm and a mean diameter of 25 µm <sup>(3)</sup>, is used as the combustible material. It is dried to constant weight in a layer not more than 25 mm thick at 105 °C for four hours and kept in a desiccator, with desiccant, until cool and required for use. The water content of the dried cellulose should be less than 0,5 % by dry mass <sup>(4)</sup>. If necessary, the drying time should be prolonged to achieve this. <sup>(5)</sup> The same batch of cellulose is to be used throughout the test.

<sup>(1)</sup> As, for example, in the framework of UN transport regulations.

<sup>(2)</sup> E.g.: 50 % (w/w) perchloric acid and 40 % (w/w) sodium chlorate are used in reference 1.

<sup>(3)</sup> E.g. Whatman Column Chromatographic Cellulose Powder CF 11, catalogue No 4021 050.

<sup>(4)</sup> Confirmed by, e.g. Karl-Fisher titration.

<sup>(5)</sup> Alternatively, this water content can also be achieved by, e.g. heating at 105 °C under vacuum for 24 h.



**▼B**1.6.1.2. *Apparatus*1.6.1.2.1. *Pressure vessel*

A pressure vessel is required. The vessel consists of a cylindrical steel pressure vessel 89 mm in length and 60 mm in external diameter (see figure 1). Two flats are machined on opposite sides (reducing the cross-section of the vessel to 50 mm) to facilitate holding whilst fitting up the firing plug and vent plug. The vessel, which has a bore of 20 mm diameter is internally rebated at either end to a depth of 19 mm and threaded to accept 1" British Standard Pipe (BSP) or metric equivalent. A pressure take-off, in the form of a side arm, is screwed into the curved face of the pressure vessel 35 mm from one end and at 90° to the machined flats. The socket for this is bored to a depth of 12 mm and threaded to accept the 1/2" BSP (or metric equivalent) thread on the end of the side-arm. If necessary, an inert seal is fitted to ensure a gas-tight seal. The side-arm extends 55 mm beyond the pressure vessel body and has a bore of 6 mm. The end of the side-arm is rebated and threaded to accept a diaphragm type pressure transducer. Any pressure-measuring device may be used provided that it is not affected by the hot gases or the decomposition products and is capable of responding to rates of pressure rise of 690-2 070 kPa in not more than 5 ms.

The end of the pressure vessel farthest from the side-arm is closed with a firing plug which is fitted with two electrodes, one insulated from, and the other earthed to, the plug body. The other end of the pressure vessel is closed by a bursting disk (bursting pressure approximately 2 200 kPa) held in place with a retaining plug which has a 20 mm bore. If necessary, an inert seal is used with the firing plug to ensure a gas-tight fit. A support stand (figure 2) holds the assembly in the correct attitude during use. This usually comprises a mild steel base plate measuring 235 mm × 184 mm × 6 mm and a 185 mm length of square hollow section (S.H.S.) 70 mm × 70 mm × 4 mm.

A section is cut from each of two opposite sides at one end of the length of S.H.S. so that a structure having two flat sided legs surmounted by 86 mm length of intact box section results. The ends of these flat sides are cut to an angle of 60° to the horizontal and welded to the base plate. A slot measuring 22 mm wide × 46 mm deep is machined in one side of the upper end of the base section such that when the pressure vessel assembly is lowered, firing plug end first, into the box section support, the side-arm is accommodated in the slot. A piece of steel 30 mm wide and 6 mm thick is welded to the lower internal face of the box section to act as a spacer. Two 7 mm thumb screws, tapped into the opposite face, serve to hold the pressure vessel firmly in place. Two 12 mm wide strips of 6 mm thick steel, welded to the side pieces abutting the base of the box section, support the pressure vessel from beneath.

**▼B**

## 1.6.1.2.2. Ignition system

The ignition system consists of a 25 cm long Ni/Cr wire with a diameter 0,6 mm and a resistance of 3,85 ohm/m. The wire is wound, using a 5 mm diameter rod, in the shape of a coil and is attached to the firing plug electrodes. The coil should have one of the configurations shown in figure 3. The distance between the bottom of the vessel and the underside of the ignition coil should be 20 mm. If the electrodes are not adjustable, the ends of the ignition wire between the coil and the bottom of the vessel should be insulated by a ceramic sheath. The wire is heated by a constant current power supply able to deliver at least 10 A.

1.6.2. Performance of the test <sup>(1)</sup>

The apparatus, assembled complete with pressure transducer and heating system but without the bursting disk in position, is supported firing plug end down. 2,5 g of the liquid to be tested is mixed with 2,5 g of dried cellulose in a glass beaker using a glass stirring rod <sup>(2)</sup>. For safety, the mixing should be performed with a safety shield between the operator and mixture. If the mixture ignites during mixing or filling, no further testing is necessary. The mixture is added, in small portions with tapping, to the pressure vessel making sure that the mixture is packed around the ignition coil and is in good contact with it. It is important that the coil is not distorted during the packing process as this may lead to erroneous results <sup>(3)</sup>. The bursting disk is placed in position and the retaining plug is screwed in tightly. The charged vessel is transferred to the firing support stand, bursting disk uppermost, which should be located in a suitable, armoured fume cupboard or firing cell. The power supply is connected to the external terminals of the firing plug and 10 A applied. The time between the start of mixing and switching on the power should not exceed 10 minutes.

The signal produced by the pressure transducer is recorded on a suitable system which allows both evaluation and the generation of a permanent record of the time pressure profile obtained (e.g. a transient recorder coupled to a chart recorder). The mixture is heated until the bursting disk ruptures or until at least 60 s have elapsed. If the bursting disk does not rupture, the mixture should be allowed to cool before carefully dismantling the apparatus, taking precautions to allow for any pressurisation which may occur. Five trials are performed with the test substance and the reference substance(s). The time taken for the pressure to rise from 690 kPa to 2 070 kPa above atmospheric is noted. The mean pressure rise time is calculated.

In some cases, substances may generate a pressure rise (too high or too low), caused by chemical reactions not characterising the oxidising properties of the substance. In these cases, it may be necessary to repeat the test with an inert substance, e.g. diatomite (kieselguhr), in place of the cellulose in order to clarify the nature of the reaction.

<sup>(1)</sup> Mixtures of oxidisers with cellulose must be treated as potentially explosive and handled with due care.

<sup>(2)</sup> In practice this can be achieved by preparing a 1:1 mixture of the liquid to be tested and cellulose in a greater amount than needed for the trial and transferring 5 ± 0,1 g to the pressure vessel. The mixture is to be freshly prepared for each trial.

<sup>(3)</sup> In particular, contact between the adjacent turns of the coil must be avoided.

**▼B****2. DATA**

Pressure rise times for both the test substance and the reference substance(s). Pressure rise times for the tests with an inert substance, if performed.

**2.1. TREATMENT OF RESULTS**

The mean pressure rise times for both the test substance and the reference substances(s) are calculated.

The mean pressure rise time for the tests with an inert substance (if performed) is calculated.

Some examples of results are shown in Table 1.

*Table 1*

**Examples of results <sup>(a)</sup>**

Substance <sup>(b)</sup>	Mean pressure rise time for a 1:1 mixture with cellulose (ms)
Ammonium dichromate, saturated aqueous solution	20 800
Calcium nitrate, saturated aqueous solution	6 700
Ferric nitrate, saturated aqueous solution	4 133
Lithium perchlorate, saturated aqueous solution	1 686
Magnesium perchlorate, saturated aqueous solution	777
Nickel nitrate, saturated aqueous solution	6 250
Nitric acid, 65 %	4 767 <sup>(c)</sup>
Perchloric acid, 50 %	121 <sup>(c)</sup>
Perchloric acid, 55 %	59
Potassium nitrate, 30 % aqueous solution	26 690
Silver nitrate, saturated aqueous solution	<sup>(d)</sup>
Sodium chlorate, 40 % aqueous solution	2 555 <sup>(c)</sup>
Sodium nitrate, 45 % aqueous solution	4 133
<i>Inert substance</i>	
Water: cellulose	<sup>(d)</sup>

<sup>(a)</sup> See reference (1) for classification under the UN transport scheme.

<sup>(b)</sup> Saturated solutions should be prepared at 20 °C.

<sup>(c)</sup> Mean value from interlaboratory comparative trials.

<sup>(d)</sup> Maximum pressure of 2 070 kPa not reached.

**▼B****3. REPORT****3.1. TEST REPORT**

The test report should include the following information:

- the identity, composition, purity, etc. of the substance tested,
- the concentration of the test substance,
- the drying procedure of the cellulose used,
- the water content of the cellulose used,
- the results of the measurements,
- the results from tests with an inert substance, if any,
- the calculated mean pressure rise times,
- any deviations from this method and the reasons for them,
- all additional information or remarks relevant to the interpretation of the results.

**3.2. INTERPRETATION OF THE RESULTS <sup>(1)</sup>**

The test results are assessed on the basis of:

- (a) whether the mixture of test substance and cellulose spontaneously ignites; and
- (b) the comparison of the mean time taken for the pressure to rise from 690 kPa to 2 070 kPa with that of the reference substance(s).

A liquid substance is to be considered as an oxidiser when:

- (a) a 1:1 mixture, by mass, of the substance and cellulose spontaneously ignites; or
- (b) a 1:1 mixture, by mass, of the substance and cellulose exhibits a mean pressure rise time less than or equal to the mean pressure rise time of a 1:1 mixture, by mass, of 65 % (w/w) aqueous nitric acid and cellulose.

In order to avoid a false positive result, if necessary, the results obtained when testing the substance with an inert material should also be considered when interpreting the results.

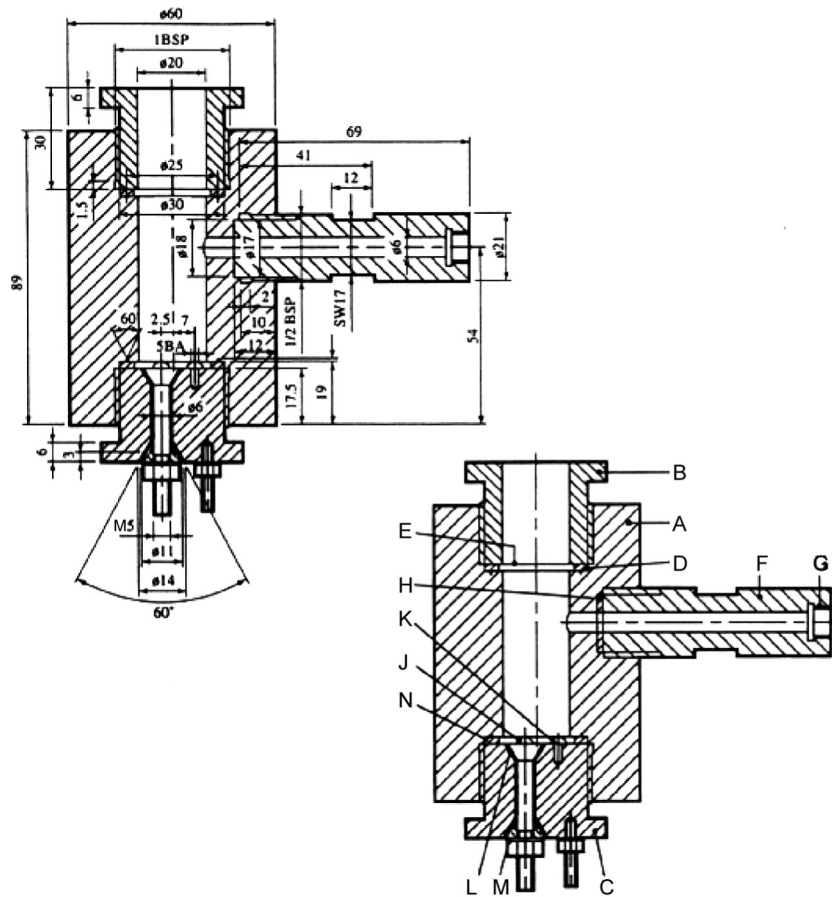
<sup>(1)</sup> See reference 1 for interpretation of the results under the UN transport regulations using several reference substances.

**▼B**

4. REFERENCES

- (1) Recommendations on the Transport of Dangerous Goods, Manual of Tests and Criteria. 3<sup>rd</sup> revised edition. UN Publication No: ST/SG/AC.10/11/Rev. 3, 1999, page 342. Test O.2: Test for oxidising liquids.

Figure 1  
Pressure vessel



- |                              |                                  |                         |
|------------------------------|----------------------------------|-------------------------|
| (A) Pressure vessel body     | (B) Bursting disk retaining plug | (C) Firing plug         |
| (D) Soft lead washer         | (E) Bursting disc                | (F) Side arm            |
| (G) Pressure transducer head | (H) Washer                       | (J) Insulated electrode |
| (K) Earthed electrode        | (L) Insulation                   | (M) Steel cone          |
| (N) Washer distorting groove |                                  |                         |

▼B

Figure 2

Support stand

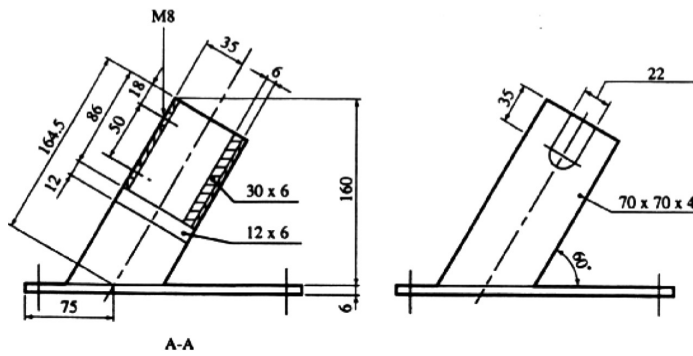
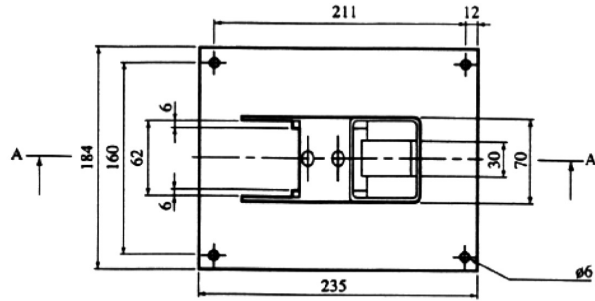
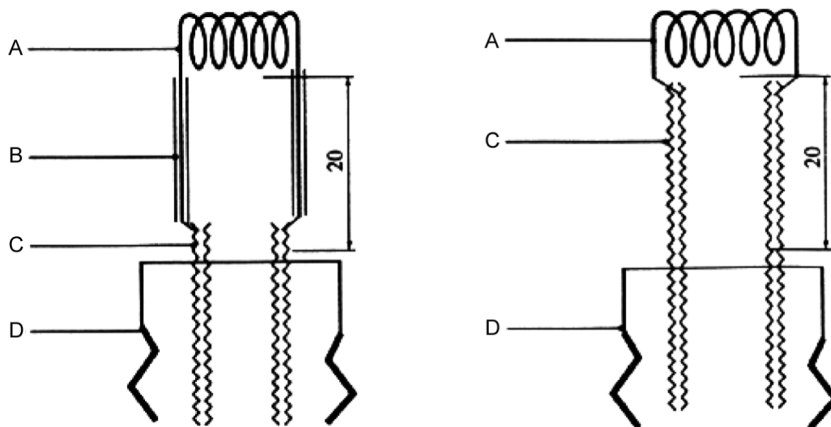


Figure 3

Ignition system

- (A) Ignition coil    (B) Insulation    (C) Electrodes    (D) Firing plug



Note: either of these configurations may be used.

**▼ M1****A.22. LENGTH WEIGHTED GEOMETRIC MEAN DIAMETER OF FIBRES****1. METHOD****1.1. INTRODUCTION**

This method describes a procedure to measure the Length Weighted Geometric Mean Diameter (LWGMD) of bulk Man Made Mineral Fibres (MMMMF). As the LWGMD of the population will have a 95 % probability of being between the 95 % confidence levels (LWGMD  $\pm$  two standard errors) of the sample, the value reported (the test value) will be the lower 95 % confidence limit of the sample (i.e. LWGMD  $-$  2 standard errors). The method is based on an update (June 1994) of a draft HSE industry procedure agreed at a meeting between ECFIA and HSE at Chester on 26/9/93 and developed for and from a second inter-laboratory trial (1, 2). This measurement method can be used to characterise the fibre diameter of bulk substances or products containing MMMFs including refractory ceramic fibres (RCF), man-made vitreous fibres (MMVF), crystalline and polycrystalline fibres.

Length weighting is a means of compensating for the effect on the diameter distribution caused by the breakage of long fibres when sampling or handling the material. Geometric statistics (geometric mean) are used to measure the size distribution of MMMF diameters because these diameters usually have size distributions that approximate to log normal.

Measuring length as well as diameter is both tedious and time consuming but, if only those fibres that touch an infinitely thin line on a SEM field of view are measured, then the probability of selecting a given fibre is proportional to its length. As this takes care of the length in the length weighting calculations, the only measurement required is the diameter and the LWGMD-2SE can be calculated as described.

**1.2. DEFINITIONS**

**Particle:** An object with a length to width ratio of less than 3:1.

**Fibre:** An object with a length to width ratio (aspect ratio) of at least 3:1.

**1.3. SCOPE AND LIMITATIONS**

The method is designed to look at diameter distributions which have median diameters from 0,5  $\mu\text{m}$  to 6  $\mu\text{m}$ . Larger diameters can be measured by using lower SEM magnifications but the method will be increasingly limited for finer fibre distributions and a TEM (transmission electron microscope) measurement is recommended if the median diameter is below 0,5  $\mu\text{m}$ .

**▼ M1**

## 1.4. PRINCIPLE OF THE TEST METHOD

A number of representative core samples are taken from the fibre blanket or from loose bulk fibre. The bulk fibres are reduced in length using a crushing procedure and a representative sub-sample dispersed in water. Aliquots are extracted and filtered through a 0,2 µm pore size, polycarbonate filter and prepared for examination using scanning electron microscope (SEM) techniques. The fibre diameters are measured at a screen magnification of  $\times 10\,000$  or greater <sup>(1)</sup> using a line intercept method to give an unbiased estimate of the median diameter. The lower 95 % confidence interval (based on a one sided test) is calculated to give an estimate of the lowest value of the geometric mean fibre diameter of the material.

## 1.5. DESCRIPTION OF THE TEST METHOD

1.5.1. **Safety/precautions**

Personal exposure to airborne fibres should be minimised and a fume cupboard or glove box should be used for handling the dry fibres. Periodic personal exposure monitoring should be carried out to determine the effectiveness of the control methods. When handling MMMF's disposable gloves should be worn to reduce skin irritation and to prevent cross-contamination.

1.5.2. **Apparatus/equipment**

- Press and dyes (capable of producing 10 MPa).
- 0,2 µm pore size polycarbonate capillary pore filters (25 mm diameter).
- 5 µm pore size cellulose ester membrane filter for use as a backing filter.
- Glass filtration apparatus (or disposable filtration systems) to take 25 mm diameter filters (e.g. Millipore glass microanalysis kit, type No XX10 025 00).
- Freshly distilled water that has been filtered through a 0,2 µm pore size filter to remove micro-organisms.
- Sputter coater with a gold or gold/palladium target.
- Scanning electron microscope capable of resolving down to 10 nm and operating at  $\times 10\,000$  magnification.
- Miscellaneous: spatulas, type 24 scalpel blade, tweezers, SEM tubes, carbon glue or carbon adhesive tape, silver dag.
- Ultrasonic probe or bench top ultrasonic bath.
- Core sampler or cork borer, for taking core samples from MMMF blanket.

<sup>(1)</sup> This magnification value is indicated for 3 µm fibres, for 6 µm fibres a magnification of  $\times 5\,000$  may be more suitable.



**▼ M1****1.5.3. Test Procedure****1.5.3.1. Sampling**

For blankets and bats a 25 mm core sampler or cork borer is used to take samples of the cross-section. These should be equally spaced across the width of a small length of the blanket or taken from random areas if long lengths of the blanket are available. The same equipment can be used to extract random samples from loose fibre. Six samples should be taken when possible, to reflect spatial variations in the bulk material.

The six core samples should be crushed in a 50 mm diameter dye at 10 MPa. The material is mixed with spatula and re-pressed at 10 MPa. The material is then removed from the dye and stored in a sealed glass bottle.

**1.5.3.2. Sample Preparation**

If necessary, organic binder can be removed by placing the fibre inside a furnace at 450 °C for about one hour.

Cone and quarter to subdivide the sample (this should be done inside a dust cupboard).

Using a spatula, add a small amount (< 0,5 g) of sample to 100 ml of freshly distilled water that has been filtered through a 0,2 µm membrane filter (alternative sources of ultra pure water may be used if they are shown to be satisfactory). Disperse thoroughly by the use of an ultrasonic probe operated at 100 W power and tuned so that cavitation occurs. (If a probe is not available use the following method: repeatedly shake and invert for 30 seconds; ultrasonic in a bench top ultrasonic bath for five minutes; then repeatedly shake and invert for a further 30 seconds.)

Immediately after dispersion of the fibre, remove a number of aliquots (e.g. three aliquots of 3, 6 and 10 ml) using a wide-mouthed pipette (2-5 ml capacity).

Vacuum filter each aliquot through a 0,2 µm polycarbonate filter supported by a 5 µm pore MEC backing filter, using a 25 mm glass filter funnel with a cylindrical reservoir. Approximately 5 ml of filtered distilled water should be placed into the funnel and the aliquot slowly pipetted into the water holding the pipette tip below the meniscus. The pipette and the reservoir must be flushed thoroughly after pipetting, as thin fibres have a tendency to be located more on the surface.

Carefully remove the filter and separate it from the backing filter before placing it in a container to dry.

**▼ M1**

Cut a quarter or half filter section of the filtered deposit with a type 24 scalpel blade using a rocking action. Carefully attach the cut section to a SEM stub using a sticky carbon tab or carbon glue. Silver dag should be applied in at least three positions to improve the electrical contact at the edges of the filter and the stub. When the glue/silver dag is dry, sputter coat approximately 50 nm of gold or gold/palladium onto the surface of the deposit.

1.5.3.3. *SEM calibration and operation*

## 1.5.3.3.1. Calibration

The SEM calibration should be checked at least once a week (ideally once a day) using a certified calibration grid. The calibration should be checked against a certified standard and if the measured value (SEM) is not within  $\pm 2\%$  of the certified value, then the SEM calibration must be adjusted and re-checked.

The SEM should be capable of resolving at least a minimum visible diameter of 0,2  $\mu\text{m}$ , using a real sample matrix, at a magnification of  $\times 2\,000$ .

## 1.5.3.3.2. Operation

The SEM should be operated at 10 000 magnification<sup>(1)</sup> using conditions that give good resolution with an acceptable image at slow scan rates of, for example, 5 seconds per frame. Although the operational requirements of different SEMs may vary, generally to obtain the best visibility and resolution, with relatively low atomic weight materials, accelerating voltages of 5-10 keV should be used with a small spot size setting and short working distance. As a linear traverse is being conducted, then a 0° tilt should be used to minimise re-focussing or, if the SEM has a eucentric stage, the eucentric working distance should be used. Lower magnification may be used if the material does not contain small (diameter) fibres and the fibre diameters are large ( $> 5\ \mu\text{m}$ ).

1.5.3.4. *Sizing*

## 1.5.3.4.1. Low magnification examination to assess the sample

Initially the sample should be examined at low magnification to look for evidence of clumping of large fibres and to assess the fibre density. In the event of excessive clumping it is recommended that a new sample is prepared.

For statistical accuracy it is necessary to measure a minimum number of fibres and high fibre density may seem desirable as examining empty fields is time consuming and does not contribute to the analysis. However, if the filter is overloaded, it becomes difficult to measure all the measurable fibres and, because small fibres may be obscured by larger ones, they may be missed.

<sup>(1)</sup> For 3  $\mu\text{m}$  fibres, see previous note.

**▼ M1**

Bias towards over estimating the LWGMD may result from fibre densities in excess of 150 fibres per millimetre of linear traverse. On the other hand, low fibre concentrations will increase the time of analysis and it is often cost effective to prepare a sample with a fibre density closer to the optimum than to persist with counts on low concentration filters. The optimum fibre density should give an average of about one or two countable fibre per fields of view at 5 000 magnification. Nevertheless the optimum density will depend on the size (diameter) of the fibres, so it is necessary that the operator uses some expert judgement in order to decide whether the fibre density is close to optimal or not.

#### 1.5.3.4.2. Length weighting of the fibre diameters

Only those fibres that touch (or cross) an (infinitely) thin line drawn on the screen of the SEM are counted. For this reason a horizontal (or vertical) line is drawn across the centre of the screen.

Alternatively a single point is placed at the centre of the screen and a continuous scan in one direction across the filter is initiated. Each fibre of aspect ratio greater than 3:1 touching or crossing this point has its diameter measured and recorded.

#### 1.5.3.4.3. Fibre sizing

It is recommended that a minimum of 300 fibres are measured. Each fibre is measured only once at the point of intersection with the line or point drawn on the image (or close to the point of intersection if the fibre edges are obscured). If fibres with non-uniform cross sections are encountered, a measurement representing the average diameter of the fibre should be used. Care should be taken in defining the edge and measuring the shortest distance between the fibre edges. Sizing may be done on line, or off-line on stored images or photographs. Semi-automated image measurement systems that download data directly into a spreadsheet are recommended, as they save time, eliminate transcription errors and calculations can be automated.

The ends of long fibres should be checked at low magnification to ensure that they do not curl back into the measurement field of view and are only measured once.

## 2. DATA

### 2.1. TREATMENT OF RESULTS

Fibre diameters do not usually have a normal distribution. However, by performing a log transformation it is possible to obtain a distribution that approximates to normal.

Calculate the arithmetic mean (mean lnD) and the standard deviation ( $SD_{\ln D}$ ) of the log to base e values (lnD) of the n fibre diameters (D).

$$\text{mean lnD} = \frac{\sum \ln D}{n} \quad (1)$$

**▼ M1**

$$SD_{\ln D} = \sqrt{\frac{\sum (\ln D - \text{mean } \ln D)^2}{n - 1}} \quad (2)$$

The standard deviation is divided by the square root of the number of measurements (n) to obtain the standard error ( $SE_{\ln D}$ ).

$$SE_{\ln D} = \frac{SD}{\sqrt{n}} \quad (3)$$

Subtract two times the standard error from the mean and calculate the exponential of this value (mean minus two standard errors) to give the geometric mean minus two geometric standard errors.

$$LWGMD - 2SE = e^{(\text{mean } \ln D - 2SE_{\ln D})} \quad (4)$$

### 3. REPORTING

#### TEST REPORT

The test report should include at least the following information:

- The value of LWGMD-2SE.
- Any deviations and particularly those which may have an effect on the precision or accuracy of the results with appropriate justifications.

### 4. REFERENCES

1. B. Tylee SOP MF 240. Health and Safety Executive, February 1999.
2. G. Burdett and G. Revell. Development of a standard method to measure the length-weighted geometric mean fibre diameter: Results of the Second inter-laboratory exchange. IR/L/MF/94/-07. Project R42.75 HPD. Health and Safety Executive, Research and Laboratory Services Division, 1994.

**▼M4****A.23. PARTITION COEFFICIENT (1-OCTANOL/WATER): SLOW-STIRRING METHOD**

## INTRODUCTION

1. This Test Method is equivalent to OECD Test Guideline (TG) 123 (2006). 1-octanol/water partition coefficient ( $P_{OW}$ ) values up to a  $\log P_{OW}$  of 8,2 have been accurately determined by the slow-stirring method (1). Therefore it is a suitable experimental approach for the direct determination of  $P_{OW}$  of highly hydrophobic substances.
2. Other methods for the determination of the 1-octanol/water partition coefficient ( $P_{OW}$ ) are the 'shake-flask' method (2), and the determination of the  $P_{OW}$  from reversed phase HPLC-retention behaviour (3). The 'shake-flask' method is prone to artifacts due to transfer of octanol micro-droplets into the aqueous phase. With increasing values of  $P_{OW}$  the presence of these droplets in the aqueous phase leads to an increasing overestimation of the concentration of the test substance in the water. Therefore, its use is limited to substances with  $\log P_{OW} < 4$ . The second method relies on solid data of directly determined  $P_{OW}$  values to calibrate the relationship between HPLC-retention behaviour and measured values of  $P_{OW}$ . A draft OECD guideline was available for determining 1-octanol/water partition coefficients of ionisable substances (4) but shall no longer be used.
3. This Test Method has been developed in The Netherlands. The precision of the methods described here has been validated and optimized in a ring-test validation study in which 15 laboratories participated (5).

## INITIAL CONSIDERATIONS

**Significance and use**

4. For inert organic substances highly significant relationships have been found between 1-octanol/water partition coefficients ( $P_{OW}$ ) and their bioaccumulation in fish. Moreover,  $P_{OW}$  has been demonstrated to be correlated to fish toxicity as well as to sorption of chemicals to solids such as soils and sediments. An extensive overview of the relationships has been given in reference (6).
5. A wide variety of relationships between the 1-octanol/water partition coefficient and other substance properties of relevance to environmental toxicology and chemistry have been established. As a consequence, the 1-octanol/water partition coefficient has evolved as a key parameter in the assessment of the environmental risk of chemicals as well as in the prediction of fate of chemicals in the environment.

**Scope**

6. The slow-stirring experiment is thought to reduce the formation of micro-droplets from 1-octanol droplets in the water phase. As a consequence, overestimation of the aqueous concentration due to test substance molecules associated to such droplets does not occur. Therefore, the slow-stirring method is particularly suitable for the determination of  $P_{OW}$  for substances with expected  $\log P_{OW}$  values of 5 and higher, for which the shake-flask method (2) is prone to yield erroneous results.

**▼M4**

## DEFINITION AND UNITS

7. The partition coefficient of a substance between water and a lipophilic solvent (1-octanol) characterizes the equilibrium distribution of the chemical between the two phases. The partition coefficient between water and 1-octanol ( $P_{OW}$ ) is defined as the ratio of the equilibrium concentrations of the test substance in 1-octanol saturated with water ( $C_O$ ) and water saturated with 1-octanol ( $C_W$ ).

$$P_{OW} = C_O/C_W$$

As a ratio of concentrations it is dimensionless. Most frequently it is given as the logarithm to the base 10 ( $\log P_{OW}$ ).  $P_{OW}$  is temperature dependent and reported data should include the temperature of the measurement.

## PRINCIPLE OF THE METHOD

8. In order to determine the partitioning coefficient, water, 1-octanol, and the test substance are equilibrated with each other at constant temperature. Then the concentrations of the test substance in the two phases are determined.
9. The experimental difficulties associated with the formation of micro-droplets during the shake-flask experiment can be reduced in the slow-stirring experiment proposed here. In the slow-stirring experiment, water, 1-octanol and the test substance are equilibrated in a thermostated stirred reactor. Exchange between the phases is accelerated by stirring. The stirring introduces limited turbulence which enhances the exchange between 1-octanol and water without micro-droplets being formed (1).

## APPLICABILITY OF THE TEST

10. Since the presence of substances other than the test substance might influence the activity coefficient of the test substance, the test substance should be tested as a pure substance. The highest purity commercially available should be employed for the 1-octanol/water partition experiment.
11. The present method applies to pure substances that do not dissociate or associate and that do not display significant interfacial activity. It can be applied to determine the 1-octanol/water partition ratio of such substances and of mixtures. When the method is used for mixtures, the 1-octanol/water partition ratios determined are conditional and depend on the chemical composition of the mixture tested and on the electrolyte composition employed as aqueous phase. Provided additional steps are taken, the method is also applicable to dissociating or associating compounds (paragraph 12).
12. Due to the multiple equilibria in water and 1-octanol involved in the 1-octanol/water partitioning of dissociating substances such as organic acids and phenols, organic bases, and organometallic substances, the 1-octanol/water partition ratio is a conditional constant strongly dependent on electrolyte composition (7)(8). Determination of the 1-octanol/water partition ratio therefore requires that pH and electrolyte composition be controlled in the experiment and reported. Expert judgement has to be employed in the evaluation of these partition ratios. Using the value of dissociation constant(s), suitable pH-values need to be selected, such that a partitioning ratio is determined for each ionization state. Non-complexing buffers must be used when testing organometallic compounds (8). Taking the existing knowledge on the aqueous chemistry (complexation constants, dissociation constants) into account, the experimental conditions should be chosen in such a manner that the speciation of the test substance in the aqueous phase can be estimated. The ionic strength should be identical in all experiments by employing a background electrolyte.

**▼M4**

13. Difficulties in the test may arise in conducting the test for substances with low water solubility or high  $P_{OW}$ , due to the fact that the concentrations in the water become very low such that their accurate determination is difficult. This Test Method provides guidance on how to deal with this problem.

## INFORMATION ON THE TEST SUBSTANCE

14. Chemical reagents should be of analytical grade or of higher purity. The use of non-labelled test substances with known chemical composition and preferably at least 99 % purity, or of radiolabelled test substances with known chemical composition and radiochemical purity, is recommended. In the case of short half-life tracers, decay corrections should be applied. In the case of radiolabelled test substances, a chemical specific analytical method should be employed to ensure that the measured radioactivity is directly related to the test substance.
15. An estimate of  $\log P_{OW}$  may be obtained by using commercially available software for estimation of  $\log P_{OW}$ , or by using the ratio of the solubilities in both solvents.
16. Before carrying out a slow-stirring experiment for determination of  $P_{OW}$ , the following information on the test substance should be available:

- (a) structural formula
- (b) suitable analytical methods for determination of the concentration of the substance in water and 1-octanol
- (c) dissociation constant(s) of ionisable substances (OECD Guideline 112 (9))
- (d) aqueous solubility (10)
- (e) abiotic hydrolysis (11)
- (f) ready biodegradability (12)
- (g) vapour pressure (13).

## DESCRIPTION OF THE METHOD

**Equipment and apparatus**

17. Standard laboratory equipment is required, in particular, the following:
- magnetic stirrers and Teflon coated magnetic stir bars are employed to stir the water phase;
  - analytical instrumentation, suitable for determination of the concentration of the test substance at the expected concentrations;
  - stirring-vessel with a tap at the bottom. Dependent on the estimate of  $\log P_{OW}$  and the Limit of Detection (LOD) of the test compound, the use of a reaction vessel of the same geometry larger than one litre has to be considered, so that a sufficient volume of water can be obtained for chemical extraction and analysis. This will result in higher concentrations in the water extract and thus a more reliable analytical determination. A table giving estimates of the minimum volume needed, the LOD of the compound, its estimated  $\log P_{OW}$  and its water solubility is given in Appendix 1. The table is based on the relationship between  $\log P_{OW}$  and the ratio between the solubilities in octanol and water, as presented by Pinsuwan *et al.* (14):

$$\log P_{OW} = 0,88 \log SR + 0,41$$

**▼ M4**

where

$$SR = S_{\text{oct}}/S_{\text{w}} \text{ (in molarity);}$$

and the relationship given by Lyman (15) for predicting water solubility. Water solubilities calculated with the equation given in Appendix 1 must be seen as a first estimate. It should be noted that the user is free to generate an estimate of water solubility by means of any relationship that is considered to better represent the relationship between hydrophobicity and solubility. For solid compounds, inclusion of melting point in the prediction of solubility is for instance recommended. In case a modified equation is used, it should be ascertained that the equation for calculation of solubility in octanol is still valid. A schematic drawing of a glass-jacketed stirring-vessel with a volume of ca. one litre is given in Appendix 2. The proportions of the vessel shown in Appendix 2 have proven favourable and should be maintained when apparatus of a different size is used;

— a means for keeping the temperature constant during the slow-stirring experiment is essential.

18. Vessels should be made from inert material such that adsorption to vessel surfaces is negligible.

#### **Preparation of the test solutions**

19. The  $P_{\text{OW}}$  determination should be carried out with the highest purity 1-octanol that is commercially available (at least + 99 %). Purification of 1-octanol by extraction with acid, base and water and subsequent drying is recommended. In addition, distillation can be used to purify 1-octanol. Purified 1-octanol is to be used to prepare standard solutions of the test substances. Water to be used in the  $P_{\text{OW}}$  determination should be glass or quartz distilled, or obtained from a purification system, or HPLC-grade water may be used. Filtration through a 0,22  $\mu\text{m}$  filter is required for distilled water, and blanks should be included to check that no impurities are in the concentrated extracts that may interfere with the test substance. If a glass fibre filter is used, it should be cleaned by baking for at least three hours at 400 °C.
20. Both solvents are mutually saturated prior to the experiment by equilibrating them in a sufficiently large vessel. This is accomplished by slow-stirring the two-phase system for two days.
21. An appropriate concentration of test substance is selected and dissolved in 1-octanol (saturated with water). The 1-octanol/water partition coefficient needs to be determined in dilute solutions in 1-octanol and water. Therefore the concentration of the test substance should not exceed 70 % of its solubility with a maximum concentration of 0,1 M in either phase (1). The 1-octanol solutions used for the experiment must be devoid of suspended solid test substance.
22. The appropriate amount of test substance is dissolved in 1-octanol (saturated with water). If the estimate of  $\log P_{\text{OW}}$  exceeds five, care has to be taken that the 1-octanol solutions used for the experiment are devoid of suspended solid test substance. To that end, the following procedure for chemicals with an estimated value of  $\log P_{\text{OW}} > 5$  is followed:

— the test substance is dissolved in 1-octanol (saturated with water);



**▼ M4**

- the solution is given sufficient time for the suspended solid substance to settle out. During the settling period, the concentration of the test substance is monitored;
  
- after the measured concentrations in the 1-octanol-solution have attained stable values, the stock solution is diluted with an appropriate volume of 1-octanol;
  
- the concentration of the diluted stock solution is measured. If the measured concentration is consistent with the dilution, the diluted stock solution can be employed in the slow-stirring experiment.

**Extraction and analysis of samples**

23. A validated analytical method should be used for the assay of test substance. The investigators have to provide evidence that the concentrations in the water saturated 1-octanol as well as in the 1-octanol saturated water phase during the experiment are above the method limit of quantification of the analytical procedures employed. Analytical recoveries of the test substance from the water phase and from the 1-octanol phase need to be established prior to the experiment in those cases for which extraction methods are necessary. The analytical signal needs to be corrected for blanks and care should be taken that no carry-over of analyte from one sample to another can occur.
  
24. Extraction of the water phase with an organic solvent and preconcentration of extract are likely to be required prior to analysis, due to rather low concentrations of hydrophobic test substances in the water phase. For the same reason it is necessary to reduce eventual blank concentrations. To that end, it is necessary to employ high purity solvents, preferably solvents for residue analysis. Moreover, working with carefully pre-cleaned (e.g. solvent washing or baking at elevated temperature) glassware can help to avoid cross-contamination.
  
25. An estimate of  $\log P_{OW}$  may be obtained from an estimation program or by expert judgment. If the value is higher than six then blank corrections and analyte carry-over need to be monitored closely. Similarly, if the estimate of  $\log P_{OW}$  exceeds six, the use of a surrogate standard for recovery correction is mandatory, so that high preconcentration factors can be reached. A number of software programs for the estimation of  $\log P_{OW}$  are commercially available<sup>(1)</sup>, e.g. Clog P (16), KOWWIN (17), ProLogP (18) and ACD log P (19). Descriptions of the estimation approaches can be found in references (20-22).
  
26. The limits of quantification (LOQ) for determination of the test substance in 1-octanol and water are established using accepted methods. As a rule of thumb, the method limit of quantification can be determined as the concentration in water or 1-octanol that produces a signal to noise ratio of ten. A suitable extraction and pre-concentration method should be selected and analytical recoveries should also be specified. A suitable pre-concentration factor is selected in order to obtain a signal of the required size upon analytical determination.

<sup>(1)</sup> This information is only given for the convenience of users. Other equivalent computer programmes may be used if they can be shown to produce the same results.

**▼ M4**

27. On the basis of the parameters of the analytical method and the expected concentrations, an approximate sample size required for an accurate determination of the compound concentration is determined. The use of water samples that are too small to obtain a sufficient analytical signal should be avoided. Also, the use of excessively large water samples should be avoided, since otherwise there might be too little water left for the minimum number of analyses required ( $n = 5$ ). In Appendix 1, the minimum sample volume is indicated as a function of the vessel volume, the LOD of the test substance and the solubility of the test substance.
28. Quantification of the test substances occurs by comparison with calibration curves of the respective compound. The concentrations in the samples analysed must be bracketed by concentrations of standards.
29. For test substances with a  $\log P_{OW}$  estimate higher than six a surrogate standard has to be spiked to the water sample prior to extraction in order to register losses occurring during extraction and pre-concentration of the water samples. For accurate recovery correction, the surrogates must have properties that are very close to, or identical with, those of the test substance. Preferably, (stable) isotopically-labelled analogues of the substances of interest (for example, perdeuterated or  $^{13}C$ -labelled) are used for this purpose. If the use of labelled stable isotopes, i.e.  $^{13}C$  or  $^2H$ , is not possible it should be demonstrated from reliable data in the LITERATURE that the physical-chemical properties of the surrogate are very close to those of the test substance. During liquid-liquid extraction of the water phase emulsions can form. They can be reduced by addition of salt and allowing the emulsion to settle overnight. Methods used for extracting and pre-concentrating the samples need to be reported.
30. Samples withdrawn from the 1-octanol phase may, if necessary, be diluted with a suitable solvent prior to analysis. Moreover, the use of surrogate standards for recovery correction is recommended for substances for which the recovery experiments demonstrated a high degree of variation in the recovery experiments (relative standard deviation  $> 10\%$ ).
31. The details of the analytical method need to be reported. This includes the method of extraction, pre-concentration and dilution factors, instrument parameters, calibration routine, calibration range, analytical recovery of the test substance from water, addition of surrogate standards for recovery correction, blank values, detection limits and limits of quantification.

**Performance of the Test***Optimal 1-octanol/water volume ratios*

32. When choosing the water and 1-octanol volumes, the LOQ in 1-octanol and water, the pre-concentration factors applied to the water samples, the volumes sampled in 1-octanol and water, and the expected concentrations should be considered. For experimental reasons, the volume of 1-octanol in the slow-stirring system should be chosen such that the 1-octanol layer is sufficiently thick ( $> 0,5$  cm) in order to allow for sampling of the 1-octanol phase without disturbing it.
33. Typical phase ratios used for the determinations of compounds with  $\log P_{OW}$  of 4,5 and higher are 20 to 50 ml of 1-octanol and 950 to 980 ml of water in a one litre vessel.

**▼ M4***Test conditions*

34. During the test the reaction vessel is thermostated to reduce temperature variation to below 1 °C. The assay should be performed at 25 °C.
35. The experimental system should be protected from daylight by either performing the experiment in a dark room or by covering the reaction vessel with aluminium foil.
36. The experiment should be performed in a dust-free (as far as possible) environment.
37. The 1-octanol-water system is stirred until equilibrium is attained. In a pilot experiment the length of the equilibration period is assessed by performing a slow-stirring experiment and sampling water and 1-octanol periodically. The sampling time points should be interspersed by a minimum period of five hours.
38. Each  $P_{OW}$  determination has to be performed employing at least three independent slow-stirring experiments.

*Determination of the equilibration time*

39. It is assumed that the equilibrium is achieved when a regression of the 1-octanol/water concentration ratio against time over a time span of four time points yields a slope that is not significantly different from zero at a p-level of 0,05. The minimum equilibration time is one day before sampling can be started. As a rule of thumb, sampling of substances with a log  $P_{OW}$  estimate of less than five can take place during days two and three. The equilibration might have to be extended for more hydrophobic compounds. For a compound with log  $P_{OW}$  of 8,23 (decachlorobiphenyl) 144 hours were sufficient for equilibration. Equilibrium is assessed by means of repeated sampling of a single vessel.

*Starting the experiment*

40. At the start of the experiment the reaction vessel is filled with 1-octanol-saturated water. Sufficient time should be allowed to reach the thermostated temperature.
41. The desired amount of test substance (dissolved in the required volume of 1-octanol saturated with water) is carefully added to the reaction vessel. This is a crucial step in the experiment, since turbulent mixing of the two phases has to be avoided. To that end, the 1-octanol phase can be pipetted slowly against the wall of the experimental vessel, close to the water surface. It will subsequently flow along the glass wall and form a film above the water phase. The decantation of 1-octanol directly into the flask should always be avoided; drops of 1-octanol should not be allowed to fall directly into the water.
42. After starting the stirring, the stirring rate should be increased slowly. If the stirring motors cannot be appropriately adjusted the use of a transformer should be considered. The stirring rate should be adjusted so that a vortex at the interface between water and 1-octanol of 0,5 to maximally 2,5 cm depth is created. The stirring rate should be reduced if the vortex depth of 2,5 cm is exceeded; otherwise micro-droplets may be formed from 1-octanol droplets in the water phase, leading to an overestimation of the concentration of the test substance in the water. The maximum stirring rate of 2,5 cm is recommended on the basis of the findings in the ring-test validation study (5). It is a compromise between achieving a rapid rate of equilibration, while limiting the formation of 1-octanol micro-droplets.

**▼ M4***Sampling and Sample Treatment*

43. The stirrer should be turned off prior to sampling and the liquids should be allowed to stop moving. After sampling is completed, the stirrer is started again slowly, as described above, and then the stirring rate is increased gradually.
  
44. The water phase is sampled from a stopcock at the bottom of the reaction vessel. Always discard the dead volume of water contained in the taps (approximately 5 ml in the vessel shown in the Appendix 2). The water in the taps is not stirred and therefore not in equilibrium with the bulk. Note the volume of the water samples, and make sure that the amount of test substance present in the discarded water is taken into account when setting up a mass balance. Evaporative losses should be minimized by allowing the water to flow quiescently into the separatory funnel, such that there is no disturbance of the water/1-octanol layer.
  
45. 1-Octanol samples are obtained by withdrawing a small aliquot (ca. 100  $\mu$ l) from the 1-octanol layer with a 100 microlitre all glass-metal syringe. Care should be taken not to disturb the boundary. The volume of the sampled liquid is recorded. A small aliquot is sufficient, since the 1-octanol sample will be diluted.
  
46. Unnecessary sample transfer steps should be avoided. To that end the sample volume should be determined gravimetrically. In case of water samples this can be achieved by collecting the water sample in a separatory funnel that contains already the required volume of solvent.

## DATA AND REPORTING

47. According to the present Test Method,  $P_{OW}$  is determined by performing three slow-stirring experiments (three experimental units) with the compound under investigation employing identical conditions. The regression used to demonstrate attainment of equilibrium should be based on the results of at least four determinations of  $C_O/C_W$  at consecutive time points. This allows for calculating variance as a measure of the uncertainty of the average value obtained by each experimental unit.
  
48. The  $P_{OW}$  can be characterized by the variance in the data obtained for each experimental unit. This information is employed to calculate the  $P_{OW}$  as the weighted average of the results of the individual experimental units. To do so, the inverse of the variance of the results of the experimental units is employed as weight. As a result, data with a large variation (expressed as the variance) and thus with lower reliability have less influence on the result than data with a low variance.
  
49. Analogously, the weighted standard deviation is calculated. It characterizes the repeatability of the  $P_{OW}$  measurement. A low value of the weighted standard deviation indicates that the  $P_{OW}$  determination was very repeatable within one laboratory. The formal statistical treatment of the data is outlined below.

▼ **M4****Treatment of the results***Demonstration of attainment of equilibrium*

50. The logarithm of the ratio of the concentration of the test substance in 1-octanol and water ( $\log(C_o/C_w)$ ) is calculated for each sampling time. Achievement of chemical equilibrium is demonstrated by plotting this ratio against time. A plateau in this plot that is based on at least four consecutive time points indicates that equilibrium has been attained, and that the compound is truly dissolved in 1-octanol. If not, the test needs to be continued until four successive time points yield a slope that is not significantly different from 0 at a p-level of 0,05, indicating that  $\log C_o/C_w$  is independent of time.

*Log  $P_{OW}$ -calculation*

51. The value of  $\log P_{OW}$  of the experimental unit is calculated as the weighted average value of  $\log C_o/C_w$  for the part of the curve of  $\log C_o/C_w$  vs. time, for which equilibrium has been demonstrated. The weighted average is calculated by weighting the data with the inverse of the variance so that the influence of the data on the final result is inversely proportional to the uncertainty in the data.

*Average log  $P_{OW}$* 

52. The average value of  $\log P_{OW}$  of different experimental units is calculated as the average of the results of the individual experimental units weighted with their respective variances.

The calculation is performed as follows:

$$\log P_{OW,Av} = (\sum w_i \times \log P_{OW,i}) \times (\sum w_i)^{-1}$$

where:

$\log P_{OW,i}$  = the  $\log P_{OW}$  value of the individual experimental unit i;

$\log P_{OW,Av}$  = the weighted average value of the individual  $\log P_{OW}$  determinations;

$w_i$  = the statistical weight assigned to the  $\log P_{OW}$  value of the experimental unit i.

The reciprocal of the variance of  $\log P_{OW,i}$  is employed as  $w_i$  ( $w_i = \text{var}(\log P_{OW,i})^{-1}$ )

53. The error of the average of  $\log P_{OW}$  is estimated as the repeatability of  $\log C_o/C_w$  determined during the equilibrium phase in the individual experimental units. It is expressed as the weighted standard deviation of  $\log P_{OW,Av}$  ( $\sigma_{\log P_{OW,Av}}$ ) which in turn is a measure of the error associated with  $\log P_{OW,Av}$ . The weighted standard deviation can be computed from the weighted variance ( $\text{var}_{\log P_{OW,Av}}$ ) as follows:

$$\text{var}_{\log P_{OW,Av}} = (\sum w_i \times (\log P_{OW,i} - \log P_{OW,Av})^2) \times (\sum w_i \times (n - 1))^{-1}$$

$$\sigma_{\log P_{OW,Av}} = (\text{var}_{\log P_{OW,Av}})^{0,5}$$

The symbol n stands for the number of experimental units.

**▼ M4****Test Report**

54. The test report should include the following information:

*Test substance:*

- common name, chemical name, CAS number, structural formula (indicating position of label when radiolabelled substance is used) and relevant physical-chemical properties (see paragraph 17)
- purity (impurities) of test substance
- label purity of labelled chemicals and molar activity (where appropriate)
- the preliminary estimate of  $\log P_{ow}$ , as well as the method used to derive the value.

*Test conditions:*

- dates of the performance of the studies
- temperature during the experiment
- volumes of 1-octanol and water at the beginning of the test
- volumes of withdrawn 1-octanol and water samples
- volumes of 1-octanol and water remaining in the test vessels
- description of the test vessels and stirring conditions (geometry of the stirring bar and of the test vessel, vortex height in mm, and when available: stirring rate) used
- analytical methods used to determine the test substance and the method limit of quantification
- sampling times
- the aqueous phase pH and the buffers used, when pH is adjusted for ionizable molecules
- number of replicates.

*Results:*

- repeatability and sensitivity of the analytical methods used
- determined concentrations of the test substance in 1-octanol and water as a function of time
- demonstration of mass balance
- temperature and standard deviation or the range of temperature during the experiment
- the regression of concentration ratio against time
- the average value  $\log P_{ow,Av}$  and its standard error
- discussion and interpretation of the results

**▼M4**

- examples of raw data figures of representative analysis (all raw data have to be stored in accordance with GLP standards), including recoveries of surrogates, and the number of levels used in the calibration (along with the criteria for the correlation coefficient of the calibration curve), and results of quality assurance/quality control (QA/QC)
- when available: validation report of the assay procedure (to be indicated among references).

*LITERATURE:*

- (1) De Bruijn JHM, Busser F, Seinen W, Hermens J. (1989). Determination of octanol/water partition coefficients with the 'slow-stirring' method. *Environ. Toxicol. Chem.* 8: 499-512.
- (2) Chapter A.8 of this Annex, Partition Coefficient.
- (3) Chapter A.8 of this Annex, Partition Coefficient.
- (4) OECD (2000). OECD Draft Guideline for the Testing of Chemicals: 122 Partition Coefficient (n-Octanol/Water): pH-Metric Method for Ionisable Substances. Paris.
- (5) Tolls J (2002). Partition Coefficient 1-Octanol/Water (Pow) Slow-Stirring Method for Highly Hydrophobic Chemicals, Validation Report. RIVM contract-Nrs 602730 M/602700/01.
- (6) Boethling RS, Mackay D (eds.) (2000). Handbook of property estimation methods for chemicals. Lewis Publishers Boca Raton, FL, USA.
- (7) Schwarzenbach RP, Gschwend PM, Imboden DM (1993). Environmental Organic Chemistry. Wiley, New York, NY.
- (8) Arnold CG, Widenhaupt A, David MM, Müller SR, Haderlein SB, Schwarzenbach RP (1997). Aqueous speciation and 1-octanol-water partitioning of tributyl- and triphenyltin: effect of pH and ion composition. *Environ. Sci. Technol.* 31: 2596-2602.
- (9) OECD (1981) OECD Guidelines for the Testing of Chemicals: 112 Dissociation Constants in Water. Paris.
- (10) Chapter A.6 of this Annex, Water Solubility.
- (11) Chapter C.7 of this Annex, Degradation – Abiotic Degradation Hydrolysis as a Function of pH.
- (12) Chapter C.4 — Part II – VII (Method A to F) of this Annex, Determination of 'Ready' Biodegradability.
- (13) Chapter A.4 of this Annex, Vapour Pressure.
- (14) Pinsuwan S, Li A and Yalkowsky S.H. (1995). Correlation of octanol/water solubility ratios and partition coefficients, *J. Chem. Eng. Data.* 40: 623-626.
- (15) Lyman WJ (1990). Solubility in water. In: Handbook of Chemical Property Estimation Methods: Environmental Behavior of Organic Compounds, Lyman WJ, Reehl WF, Rosenblatt DH, Eds. American Chemical Society, Washington, DC, 2-1 to 2-52.
- (16) Leo A, Weininger D (1989). Medchem Software Manual. Daylight Chemical Information Systems, Irvine, CA.
- (17) Meylan W (1993). SRC-LOGKOW for Windows. SRC, Syracuse, N.Y.
- (18) Compudrug L (1992). ProLogP. Compudrug, Ltd, Budapest.
- (19) ACD. ACD logP; Advanced Chemistry Development: Toronto, Ontario M5H 3V9, Canada, 2001.

**▼M4**

- (20) Lyman WJ (1990). Octanol/water partition coefficient. In Lyman WJ, Reehl WF, Rosenblatt DH, eds, *Handbook of chemical property estimation*, American Chemical Society, Washington, D.C.
- (21) Rekker RF, de Kort HM (1979). The hydrophobic fragmental constant: An extension to a 1 000 data point set. *Eur. J. Med. Chem. Chim. Ther.* 14: 479-488.
- (22) Jübermann O (1958). Houben-Weyl, ed, *Methoden der Organischen Chemie*: 386-390.



▼ **M4***Appendix 1***Spreadsheet for computation of minimum volumes of water required for detection of test substances of different log P<sub>ow</sub> values in aqueous phase**

Assumptions:

- Maximum volume of individual aliquots = 10 % of total volume; 5 aliquots = 50 % of total volume.
- Concentration of test substances =  $0,7 \times$  solubility in either phase. In case of lower concentrations, larger volumes would be required.
- Volume used for LOD determination = 100 ml.
- log P<sub>ow</sub> vs. log S<sub>w</sub> and log P<sub>ow</sub> vs. SR (S<sub>oct</sub>/S<sub>w</sub>) are reasonable representations of relationships for test substances.

*Estimation of S<sub>w</sub>*

log P <sub>ow</sub>	Equation	log S <sub>w</sub>	S <sub>w</sub> (mg/l)
4	$(-), 922 \times \log P_{ow} + 4,184$	0,496	3,133E+00
4,5	$(-), 922 \times \log P_{ow} + 4,184$	0,035	1,084E+00
5	$(-), 922 \times \log P_{ow} + 4,184$	-0,426	3,750E-01
5,5	$(-), 922 \times \log P_{ow} + 4,184$	-0,887	1,297E-01
6	$(-), 922 \times \log P_{ow} + 4,184$	-1,348	4,487E-02
6,5	$(-), 922 \times \log P_{ow} + 4,184$	-1,809	1,552E-02
7	$(-), 922 \times \log P_{ow} + 4,184$	-2,270	5,370E-03
7,5	$(-), 922 \times \log P_{ow} + 4,184$	-2,731	1,858E-03
8	$(-), 922 \times \log P_{ow} + 4,184$	-3,192	6,427E-04

*Estimation of S<sub>oct</sub>*

log P <sub>ow</sub>	Equation	S <sub>oct</sub> (mg/l)
4	$\log P_{ow} = 0,88 \log SR + 0,41$	3,763E+04
4,5	$\log P_{ow} = 0,88 \log SR + 0,42$	4,816E+04
5	$\log P_{ow} = 0,88 \log SR + 0,43$	6,165E+04
5,5	$\log P_{ow} = 0,88 \log SR + 0,44$	7,890E+04
6	$\log P_{ow} = 0,88 \log SR + 0,45$	1,010E+05
6,5	$\log P_{ow} = 0,88 \log SR + 0,46$	1,293E+05
7	$\log P_{ow} = 0,88 \log SR + 0,47$	1,654E+05
7,5	$\log P_{ow} = 0,88 \log SR + 0,48$	2,117E+05
8	$\log P_{ow} = 0,88 \log SR + 0,49$	2,710E+05

▼ **M4**

Total Mass test substance (mg)	Mass <sub>oct</sub> /Mass <sub>water</sub>	Mass <sub>H<sub>2</sub>O</sub> (mg)	Conc <sub>H<sub>2</sub>O</sub> (mg/l)	Mass <sub>oct</sub> (mg)	Conc <sub>oct</sub> (mg/l)
1 319	526	2,5017	2,6333	1 317	26 333
1 686	1 664	1,0127	1,0660	1 685	33 709
2 158	5 263	0,4099	0,4315	2 157	43 149
2 762	16 644	0,1659	0,1747	2 762	55 230
3 535	52 632	0,0672	0,0707	3 535	70 691
4 524	1664 36	0,0272	0,0286	4 524	90 480
5 790	5263 16	0,0110	0,0116	5 790	115 807
7 411	1 664 357	0,0045	0,0047	7 411	148 223
9 486	5 263 158	0,0018	0,0019	9 486	189 713

*Computation of volumes***Minimum volume required for H<sub>2</sub>O phase at each LOD concentration**

log K <sub>ow</sub>	LOD (micrograms/l)→	0,001	0,01	0,10	1,00	10
4		0,04	0,38	3,80	38	380
4,5		0,09	0,94	9,38	94	938
5		0,23	2,32	23,18	232	2 318
5,5		0,57	5,73	57,26	573	5 726
6		1,41	14,15	141	1 415	14 146
6,5		3,50	34,95	350	3 495	34 950
7		8,64	86,35	864	8 635	86 351
7,5		21,33	213	2 133	21 335	213 346
8		52,71	527	5 271	52 711	527 111
Volume used for LOD (l)	0,1					

*Key to Computations*

Represents < 10 % of total volume of aqueous phase, 1 litre equilibration vessel.

Represents < 10 % of total volume of aqueous phase, 2 litre equilibration vessel.

Represents < 10 % of total volume of aqueous phase, 5 litre equilibration vessel.

Represents < 10 % of total volume of aqueous phase, 10 litre equilibration vessel.

Exceeds 10 % of even the 10 liter equilibration vessel.

## ▼M4

Overview of volumes required, as a function of water solubility and Log P<sub>ow</sub>Minimum volume required for H<sub>2</sub>O phase at each LOD concentration (ml)

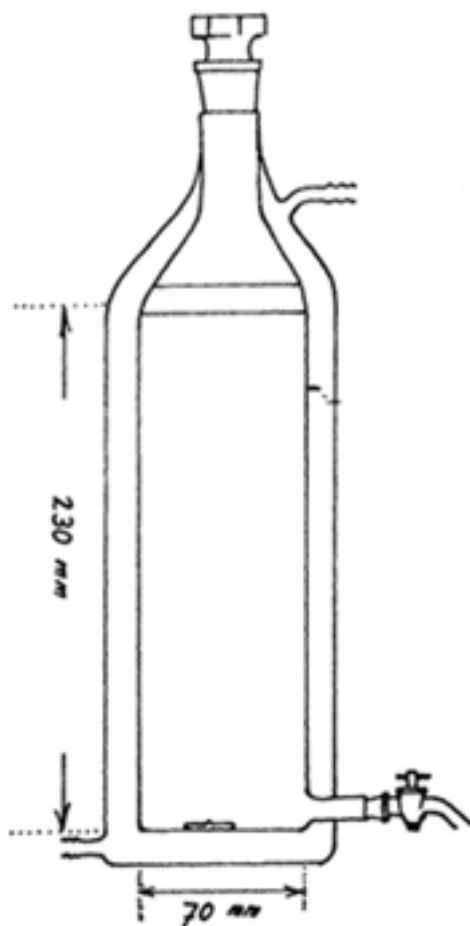
log P <sub>ow</sub>	S <sub>w</sub> (mg/l)	LOD (micrograms/l)→	0,001	0,01	0,10	1,00	10
4	10		0,01	0,12	1,19	11,90	118,99
	5		0,02	0,24	2,38	23,80	237,97
	3		0,04	0,40	3,97	39,66	396,62
4,5	1		0,12	1,19	11,90	118,99	1 189,86
	5		0,02	0,20	2,03	20,34	203,37
	2		0,05	0,51	5,08	50,84	508,42
5	1		0,10	1,02	10,17	101,68	1 016,83
	0,5		0,20	2,03	20,34	203,37	2 033,67
	1		0,09	0,87	8,69	86,90	869,01
5,5	0,5		0,17	1,74	17,38	173,80	1 738,02
	0,375		0,23	2,32	23,18	231,75	2 317,53
	0,2		0,43	4,35	43,45	434,51	4 345,05
6	0,4		0,19	1,86	18,57	185,68	1 856,79
	0,2		0,37	3,71	37,14	371,36	3 713,59
	0,1		0,74	7,43	74,27	742,72	7 427,17
6,5	0,05		1,49	14,85	148,54	1 485,43	14 854,35
	0,1		0,63	6,35	63,48	634,80	6 347,95
	0,05		1,27	12,70	126,96	1 269,59	12 695,91
7	0,025		2,54	25,39	253,92	2 539,18	25 391,82
	0,0125		5,08	50,78	507,84	5 078,36	50 783,64
	0,025		2,17	21,70	217,02	2 170,25	21 702,46
7,5	0,0125		4,34	43,40	434,05	4 340,49	43 404,93
	0,006		9,04	90,43	904,27	9 042,69	90 426,93
	0,003		18,09	180,85	1 808,54	18 085,39	180 853,86
8	0,006		7,73	77,29	772,89	7 728,85	77 288,50
	0,003		15,46	154,58	1 545,77	15 457,70	154 577,01
	0,0015		23,19	231,87	2 318,66	23 186,55	231 865,51
8,5	0,001		46,37	463,73	4 637,31	46 373,10	463 731,03
	0,002		19,82	198,18	1 981,77	19 817,73	198 177,33
	0,001		39,64	396,35	3 963,55	39 635,47	396 354,66
9	0,0005		79,27	792,71	7 927,09	79 270,93	792 709,32

▼ **M4**

log $P_{ow}$	$S_w$ (mg/l)	LOD (micrograms/l)→	0,001	0,01	0,10	1,00	10
	0,00025		158,54	1 585,42	15 854,19	158 541,86	1 585 418,63
8	0,001		33,88	338,77	3 387,68	33 876,77	338 767,72
	0,0005		67,75	677,54	6 775,35	67 753,54	677 535,44
	0,00025		135,51	1 355,07	13 550,71	135 507,09	1 355 070,89
	0,000125		271,01	2 710,14	27 101,42	271 014,18	2 710 141,77
Volume used for LOD (l)		0,1					

*Appendix 2*

**An example of glass-jacketed test vessel for the slow-stirring experiment for determination of  $P_{ow}$**





**PART B: METHODS FOR THE DETERMINATION OF TOXICITY AND OTHER HEALTH EFFECTS**

TABLE OF CONTENTS

	GENERAL INTRODUCTION
B.1 bis.	ACUTE ORAL TOXICITY — FIXED DOSE PROCEDURE
B.1 tris.	ACUTE ORAL TOXICITY — ACUTE TOXIC CLASS METHOD
B.2.	ACUTE INHALATION TOXICITY
B.3.	ACUTE TOXICITY (DERMAL)
B.4.	ACUTE TOXICITY: DERMAL IRRITATION/CORROSION
B.5.	ACUTE TOXICITY: EYE IRRITATION/CORROSION
B.6.	SKIN SENSITISATION
B.7.	REPEATED DOSE 28-DAY ORAL TOXICITY STUDY IN RODENTS
B.8.	SUBACUTE INHALATION TOXICITY: 28-DAY STUDY
B.9.	REPEATED DOSE (28 DAYS) TOXICITY (DERMAL)
B.10.	MUTAGENICITY — <i>IN VITRO</i> MAMMALIAN CHROMOSOME ABERRATION TEST
B.11.	MUTAGENICITY — <i>IN VIVO</i> MAMMALIAN BONE MARROW CHROMOSOME ABERRATION TEST
B.12.	MUTAGENICITY — <i>IN VIVO</i> MAMMALIAN ERYTHROCYTE MICRONUCLEUS TEST
B.13/14.	MUTAGENICITY: REVERSE MUTATION TEST USING BACTERIA
B.15.	MUTAGENICITY TESTING AND SCREENING FOR CARCINOGENICITY GENE MUTATION — <i>SACCHAROMYCES CEREVISIAE</i>
B.16.	MITOTIC RECOMBINATION — <i>SACCHAROMYCES CEREVISIAE</i>
B.17.	MUTAGENICITY — <i>IN VITRO</i> MAMMALIAN CELL GENE MUTATION TEST
B.18.	DNA DAMAGE AND REPAIR — UNSCHEDULED DNA SYNTHESIS — MAMMALIAN CELLS <i>IN VITRO</i>
B.19.	SISTER CHROMATID EXCHANGE ASSAY <i>IN VITRO</i>
B.20.	SEX-LINKED RECESSIVE LETHAL TEST IN <i>DROSOPHILA MELANOGASTER</i>

**▼B**

- B.21. *IN VITRO* MAMMALIAN CELL TRANSFORMATION TESTS
- B.22. RODENT DOMINANT LETHAL TEST
- B.23. MAMMALIAN SPERMATOGONIAL CHROMOSOME ABERRATION TEST
- B.24. MOUSE SPOT TEST
- B.25. MOUSE HERITABLE TRANSLOCATION
- B.26. SUB-CHRONIC ORAL TOXICITY TEST REPEATED DOSE 90 — DAY ORAL TOXICITY STUDY IN RODENTS
- B.27. SUB-CHRONIC ORAL TOXICITY TEST REPEATED DOSE 90 — DAY ORAL TOXICITY STUDY IN NON-RODENTS
- B.28. SUB-CHRONIC DERMAL TOXICITY STUDY 90-DAY REPEATED DERMAL DOSE STUDY USING RODENT SPECIES
- B.29. SUBCHRONIC INHALATION TOXICITY: 90-DAY STUDY
- B.30. CHRONIC TOXICITY STUDIES
- B.31. PRENATAL DEVELOPMENTAL TOXICITY STUDY
- B.32. CARCINOGENICITY STUDIES
- B.33. COMBINED CHRONIC TOXICITY/CARCINOGENICITY STUDIES
- B.34. ONE-GENERATION REPRODUCTION TOXICITY TEST
- B.35. TWO-GENERATION REPRODUCTION TOXICITY STUDY
- B.36. TOXICOKINETICS
- B.37. DELAYED NEUROTOXICITY OF ORGANOPHOSPHORUS SUBSTANCES FOLLOWING ACUTE EXPOSURE
- B.38. DELAYED NEUROTOXICITY OF ORGANOPHOSPHORUS SUBSTANCES 28 DAY REPEATED DOSE STUDY
- B.39. UNSCHEDULED DNA SYNTHESIS (UDS) TEST WITH MAMMALIAN LIVER CELLS *IN VIVO*
- B.40. *IN VITRO* SKIN CORROSION: TRANSCUTANEOUS ELECTRICAL RESISTANCE TEST (TER)
- B.40 BIS. *IN VITRO* SKIN CORROSION: HUMAN SKIN MODEL TEST

**▼ B**

- B.41. *IN VITRO* 3T3 NRU PHOTOTOXICITY TEST
- B.42. SKIN SENSITISATION: LOCAL LYMPH NODE ASSAY
- B.43. NEUROTOXICITY STUDY IN RODENTS
- B.44. SKIN ABSORPTION: *IN VIVO* METHOD
- B.45. SKIN ABSORPTION: *IN VITRO* METHOD
- B.46. *IN VITRO* SKIN IRRITATION: RECONSTRUCTED HUMAN EPIDERMIS TEST METHOD
- B.47. BOVINE CORNEAL OPACITY AND PERMEABILITY TEST METHOD FOR IDENTIFYING OCULAR CORROSIVES AND SEVERE IRRITANTS
- B.48. ISOLATED CHICKEN EYE TEST METHOD FOR IDENTIFYING OCULAR CORROSIVES AND SEVERE IRRITANTS
- B.49. *IN VITRO* MAMMALIAN CELL MICRONUCLEUS TEST
- B.50. SKIN SENSITISATION: LOCAL LYMPH NODE ASSAY: DA
- B.51. SKIN SENSITISATION: LOCAL LYMPH NODE ASSAY: BrdU-ELISA

**▼ M4**

- B.52. ACUTE INHALATION TOXICITY — ACUTE TOXIC CLASS METHOD

**▼ M5**

- B.53. DEVELOPMENTAL NEUROTOXICITY STUDY
- B.54. UTEROTROPHIC BIOASSAY IN RODENTS: A SHORT-TERM SCREENING TEST FOR OESTROGENIC PROPERTIES
- B.55. HERSHBERGER BIOASSAY IN RATS: A SHORT-TERM SCREENING ASSAY FOR (ANTI)ANDROGENIC PROPERTIES
- B.56. EXTENDED ONE-GENERATION REPRODUCTIVE TOXICITY STUDY
- B.57. H295R STEROIDOGENESIS ASSAY
- B.58. TRANSGENIC RODENT SOMATIC AND GERM CELL GENE MUTATION ASSAYS

**▼B****GENERAL INTRODUCTION****A. CHARACTERISATION OF THE TEST SUBSTANCE**

The composition of the test substance, including major impurities, and its relevant physico-chemical properties including stability, should be known prior to the initiation of any toxicity study.

The physico-chemical properties of the test substance provide important information for the selection of the route of administration, the design of each particular study and the handling and storage of the test substance.

The development of an analytical method for qualitative and quantitative determination of the test substance (including major impurities when possible) in the dosing medium and the biological material should precede the initiation of the study.

All information relating to the identification, the physico-chemical properties, the purity, and behaviour of the test substance should be included in the test report.

**B. ANIMAL CARE**

Stringent control of environmental conditions and proper animal care techniques are essential in toxicity testing.

**(i) Housing conditions**

The environmental conditions in the experimental animal rooms or enclosures should be appropriate to the test species. For rats, mice and guinea pigs, suitable conditions are a room temperature of  $22\text{ °C} \pm 3\text{ °C}$  with a relative humidity of 30 to 70 %; for rabbits the temperature should be  $20 \pm 3\text{ °C}$  with a relative humidity of 30 to 70 %.

Some experimental techniques are particularly sensitive to temperature effects and, in these cases, details of appropriate conditions are included in the description of the test method. In all investigations of toxic effects, the temperature and humidity should be monitored, recorded, and included in the final report of the study.

Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. Details of the lighting pattern should be recorded and included in the final report of the study.

Unless otherwise specified in the method, animals may be housed individually, or be caged in small groups of the same sex; for group caging, no more than five animals should be housed per cage.

In reports of animal experiments, it is important to indicate the type of caging used and the number of animals housed in each cage both during exposure to the chemical and any subsequent observation period.



**▼B**(ii) *Feeding conditions*

Diets should meet all the nutritional requirements of the species under test. Where test substances are administered to animals in their diet the nutritional value may be reduced by interaction between the substance and a dietary constituent. The possibility of such a reaction should be considered when interpreting the results of tests. Conventional laboratory diets may be used with an unlimited supply of drinking water. The choice of the diet may be influenced by the need to ensure a suitable admixture of a test substance when administered by this method.

Dietary contaminants which are known to influence the toxicity should not be present in interfering concentrations.

**C. ALTERNATIVE TESTING**

The European Union is committed to promoting the development and validation of alternative techniques which can provide the same level of information as current animal tests, but which use fewer animals, cause less suffering or avoid the use of animals completely.

Such methods, as they become available, must be considered wherever possible for hazard characterisation and consequent classification and labelling for intrinsic hazards and chemical safety assessment.

**D. EVALUATION AND INTERPRETATION**

When tests are evaluated and interpreted, limitations in the extent to which the results of animal and *in vitro* studies can be extrapolated directly to man must be considered and therefore, evidence of adverse effects in humans, where available, may be used for confirmation of testing results.

**E. LITERATURE REFERENCES**

Most of these methods are developed within the framework of the OECD programme for Testing Guidelines, and should be performed in conformity with the principles of Good Laboratory Practice, in order to ensure as wide as possible 'mutual acceptance of data'.

Additional information may be found in the references listed in the OECD guidelines and the relevant literature published elsewhere.

**▼B****B.1 bis. ACUTE ORAL TOXICITY — FIXED DOSE PROCEDURE****1. METHOD**

This test method is equivalent to OECD TG 420 (2001)

**1.1. INTRODUCTION**

Traditional methods for assessing acute toxicity use death of animals as an endpoint. In 1984, a new approach to acute toxicity testing was suggested by the British Toxicology Society based on the administration at a series of fixed dose levels (1). The approach avoided using death of animals as an endpoint, and relied instead on the observation of clear signs of toxicity at one of a series of fixed dose levels. Following UK (2) and international (3) *in vivo* validation studies the procedure was adopted as a testing method in 1992. Subsequently, the statistical properties of the Fixed Dose Procedure have been evaluated using mathematical models in a series of studies (4)(5)(6). Together, the *in vivo* and modelling studies have demonstrated that the procedure is reproducible, uses fewer animals and causes less suffering than the traditional methods and is able to rank substances in a similar manner to the other acute toxicity testing methods.

Guidance on the selection of the most appropriate test method for a given purpose can be found in the Guidance Document on Acute Oral Toxicity Testing (7). This guidance document also contains additional information on the conduct and interpretation of Testing Method B.1bis.

It is a principle of the method that in the main study only moderately toxic doses are used, and that administration of doses that are expected to be lethal should be avoided. Also, doses that are known to cause marked pain and distress, due to corrosive or severely irritant actions, need not be administered. Moribund animals, or animals obviously in pain or showing signs of severe and enduring distress shall be humanely killed, and are considered in the interpretation of the test results in the same way as animals that died on test. Criteria for making the decision to kill moribund or severely suffering animals, and guidance on the recognition of predictable or impending death, are the subject of a separate Guidance Document (8).

The method provides information on the hazardous properties and allows the substance to be ranked and classified according to the Globally Harmonised System (GHS) for the classification of chemicals which cause acute toxicity (9).

The testing laboratory should consider all available information on the test substance prior to conducting the study. Such information will include the identity and chemical structure of the substance; its physico-chemical properties; the results of any other *in vitro* or *in vivo* toxicity tests on the substance; toxicological data on structurally related substances; and the anticipated use(s) of the substance. This information is necessary to satisfy all concerned that the test is relevant for the protection of human health, and will help in the selection of an appropriate starting dose.

**▼ B**

## 1.2. DEFINITIONS

**Acute oral toxicity:** refers to those adverse effects occurring following oral administration of a single dose of a substance or multiple doses given within 24 hours.

**Delayed death:** means that an animal does not die or appear moribund within 48 hours but dies later during the 14-day observation period.

**Dose:** is the amount of test substance administered. Dose is expressed as weight of test substance per unit weight of test animal (e.g. mg/kg).

**Evident toxicity:** is a general term describing clear signs of toxicity following the administration of test substance (see (3) for examples) such that at the next highest fixed dose either severe pain and enduring signs of severe distress, moribund status (criteria are presented in the Humane Endpoints Guidance Document (8)), or probable mortality in most animals can be expected.

**GHS:** Globally Harmonised Classification System for Chemical Substances and Mixtures. A joint activity of OECD (human health and the environment), UN Committee of Experts on Transport of Dangerous Goods (physical-chemical properties) and ILO (hazard communication) and coordinated by the Interorganisation Programme for the Sound Management of Chemicals (IOMC).

**Impending death:** when moribund state or death is expected prior to the next planned time of observation. Signs indicative of this state in rodents could include convulsions, lateral position, recumbence and tremor. (See the Humane Endpoint Guidance Document (8) for more details).

**LD<sub>50</sub> (median lethal dose):** is a statistically derived single dose of a substance that can be expected to cause death in 50 % of animals when administered by the oral route. The LD<sub>50</sub> value is expressed in terms of weight of test substance per unit weight of test animal (mg/kg).

**Limit dose:** refers to a dose at an upper limitation on testing (2 000 or 5 000 mg/kg).

**Moribund status:** being in a state of dying or inability to survive, even if treated. (See the Humane Endpoint Guidance Document (8) for more details).

**Predictable death:** presence of clinical signs indicative of death at a known time in the future before the planned end of the experiment, for example: inability to reach water or food. (See the Humane Endpoint Guidance Document (8) for more details).

**▼B**

## 1.3. PRINCIPLE OF THE TEST METHOD

Groups of animals of a single sex are dosed in a stepwise procedure using the fixed doses of 5, 50, 300 and 2 000 mg/kg (exceptionally an additional fixed dose of 5 000 mg/kg may be considered, see Section 1.6.2). The initial dose level is selected on the basis of a sighting study as the dose expected to produce some signs of toxicity without causing severe toxic effects or mortality. Clinical signs and conditions associated with pain, suffering, and impending death, are described in detail in a separate OECD Guidance Document (8). Further groups of animals may be dosed at higher or lower fixed doses, depending on the presence or absence of signs of toxicity or mortality. This procedure continues until the dose causing evident toxicity or no more than one death is identified, or when no effects are seen at the highest dose or when deaths occur at the lowest dose.

## 1.4. DESCRIPTION OF THE TEST METHOD

1.4.1. **Selection of animal species**

The preferred rodent species is the rat, although other rodent species may be used. Normally females are used (7). This is because literature surveys of conventional LD<sub>50</sub> tests show that usually there is little difference in sensitivity between the sexes, but in those cases where differences are observed, females are generally slightly more sensitive (10). However, if knowledge of the toxicological or toxicokinetic properties of structurally related chemicals indicates that males are likely to be more sensitive then this sex should be used. When the test is conducted in males, adequate justification should be provided.

Healthy young adult animals of commonly used laboratory strains should be employed. Females should be nulliparous and non-pregnant. Each animal, at the commencement of its dosing, should be between eight and 12 weeks old and its weight should fall in an interval within  $\pm 20$  % of the mean weight of any previously dosed animals.

1.4.2. **Housing and feeding conditions**

The temperature of the experimental animal room should be 22 °C ( $\pm 3$  °C). Although the relative humidity should be at least 30 % and preferably not exceed 70 % other than during room cleaning the aim should be 50-60 %. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water. Animals may be group-caged by dose, but the number of animals per cage must not interfere with clear observations of each animal.

1.4.3. **Preparation of animals**

The animals are randomly selected, marked to permit individual identification, and kept in their cages for at least five days prior to the start of dosing to allow for acclimatisation to the laboratory conditions.

**▼B****1.4.4. Preparation of doses**

In general test substances should be administered in a constant volume over the range of doses to be tested by varying the concentration of the dosing preparation. Where a liquid end product or mixture is to be tested however, the use of the undiluted test substance, i.e. at a constant concentration, may be more relevant to the subsequent risk assessment of that substance, and is a requirement of some regulatory authorities. In either case, the maximum dose volume for administration must not be exceeded. The maximum volume of liquid that can be administered at one time depends on the size of the test animal. In rodents, the volume should not normally exceed 1 ml /100 g of body weight; however in the case of aqueous solutions 2 ml/100 g body weight can be considered. With respect to the formulation of the dosing preparation, the use of an aqueous solution/suspension/emulsion is recommended wherever possible, followed in order of preference by a solution/suspension/emulsion in oil (e.g. corn oil) and then possibly solution in other vehicles. For vehicles other than water the toxicological characteristics of the vehicle should be known. Doses must be prepared shortly prior to administration unless the stability of the preparation over the period during which it will be used is known and shown to be acceptable.

**1.5. PROCEDURE****1.5.1. Administration of doses**

The test substance is administered in a single dose by gavage using a stomach tube or a suitable intubation canula. In the unusual circumstance that a single dose is not possible, the dose may be given in smaller fractions over a period not exceeding 24 hours.

Animals should be fasted prior to dosing (e.g. with the rat, food but not water should be withheld over-night; with the mouse, food but not water should be withheld for three to four hours). Following the period of fasting, the animals should be weighed and the test substance administered. After the substance has been administered, food may be withheld for a further three to four hours in rats or one to two hours in mice. Where a dose is administered in fractions over a period of time, it may be necessary to provide the animals with food and water depending on the length of the period.

**1.5.2. Sighting study**

The purpose of the sighting study is to allow selection of the appropriate starting dose for the main study. The test substance is administered to single animals in a sequential manner following the flow-charts in Appendix 1. The sighting study is completed when a decision on the starting dose for the main study can be made (or if a death is seen at the lowest fixed dose).

The starting dose for the sighting study is selected from the fixed dose levels of 5, 50, 300 and 2 000 mg/kg as a dose expected to produce evident toxicity based, when possible, on evidence from *in vivo* and *in vitro* data from the same chemical and from structurally related chemicals. In the absence of such information, the starting dose will be 300 mg/kg.

A period of at least 24 hours will be allowed between the dosing of each animal. All animals should be observed for at least 14 days.

**▼B**

Exceptionally, and only when justified by specific regulatory needs, the use of an additional upper fixed dose level of 5 000 mg/kg may be considered (see Appendix 3). For reasons of animal welfare concern, testing of animals in GHS Category 5 ranges (2 000-5 000 mg/kg) is discouraged and should only be considered when there is a strong likelihood that the results of such a test have a direct relevance for protecting human or animal health or the environment.

In cases where an animal tested at the lowest fixed dose level (5 mg/kg) in the sighting study dies, the normal procedure is to terminate the study and assign the substance to GHS Category 1 (as shown in Appendix 1). However, if further confirmation of the classification is required, an optional supplementary procedure may be conducted, as follows. A second animal is dosed at 5 mg/kg. If this second animal dies, then GHS Category 1 will be confirmed and the study will be immediately terminated. If the second animal survives, then a maximum of three additional animals will be dosed at 5 mg/kg. Because there will be a high risk of mortality, these animals should be dosed in a sequential manner to protect animal welfare. The time interval between dosing each animal should be sufficient to establish that the previous animal is likely to survive. If a second death occurs, the dosing sequence will be immediately terminated and no further animals will be dosed. Because the occurrence of a second death (irrespective of the number of animals tested at the time of termination) falls into outcome A (two or more deaths), the classification rule of Appendix 2 at the 5 mg/kg fixed dose is followed (Category 1 if there are two or more deaths or Category 2 if there is no more than one death). In addition, Appendix 4 gives guidance on the classification in the EU system until the new GHS is implemented.

### 1.5.3. **Main study**

#### 1.5.3.1. *Numbers of animals and dose levels*

The action to be taken following testing at the starting dose level is indicated by the flowcharts in Appendix 2. One of three actions will be required; either stop testing and assign the appropriate hazard classification class, test at a higher fixed dose or test at a lower fixed dose. However, to protect animals, a dose level that caused death in the sighting study will not be revisited in the main study (see Appendix 2). Experience has shown that the most likely outcome at the starting dose level will be that the substance can be classified and no further testing will be necessary.

A total of five animals of one sex will normally be used for each dose level investigated. The five animals will be made up of one animal from the sighting study dosed at the selected dose level together with an additional four animals (except, unusually, if a dose level used on the main study was not included in the sighting study).

The time interval between dosing at each level is determined by the onset, duration, and severity of toxic signs. Treatment of animals at the next dose should be delayed until one is confident of survival of the previously dosed animals. A period of three or four days between dosing at each dose level is recommended, if needed, to allow for the observation of delayed toxicity. The time interval may be adjusted as appropriate, e.g. in case of inconclusive response.

**▼B**

When the use of an upper fixed dose of 5 000 mg/kg is considered, the procedure outlined in Appendix 3 should be followed (see also section 1.6.2).

**1.5.3.2. Limit test**

The limit test is primarily used in situations where the experimenter has information indicating that the test material is likely to be nontoxic, i.e., having toxicity only above regulatory limit doses. Information about the toxicity of the test material can be gained from knowledge about similar tested compounds or similar tested mixtures or products, taking into consideration the identity and percentage of components known to be of toxicological significance. In those situations where there is little or no information about its toxicity, or in which the test material is expected to be toxic, the main test should be performed.

Using the normal procedure, a sighting study starting dose of 2 000 mg/kg (or exceptionally 5 000 mg/kg) followed by dosing of a further four animals at this level serves as a limit test for this guideline.

**1.6. OBSERVATIONS**

Animals are observed individually after dosing at least once during the first 30 minutes, periodically during the first 24 hours, with special attention given during the first four hours, and daily thereafter, for a total of 14 days, except where they need to be removed from the study and humanely killed for animal welfare reasons or are found dead. However, the duration of observation should not be fixed rigidly. It should be determined by the toxic reactions, time of onset and length of recovery period, and may thus be extended when considered necessary. The times at which signs of toxicity appear and disappear are important, especially if there is a tendency for toxic signs to be delayed (11). All observations are systematically recorded, with individual records being maintained for each animal.

Additional observations will be necessary if the animals continue to display signs of toxicity. Observations should include changes in skin and fur, eyes and mucous membranes, and also respiratory, circulatory, autonomic and central nervous systems, and somatomotor activity and behaviour pattern. Attention should be directed to observations of tremors, convulsions, salivation, diarrhoea, lethargy, sleep and coma. The principles and criteria summarised in the Humane Endpoints Guidance Document should be taken into consideration (8). Animals found in a moribund condition and animals showing severe pain or enduring signs of severe distress should be humanely killed. When animals are killed for humane reasons or found dead, the time of death should be recorded as precisely as possible.

**1.6.1. Body weight**

Individual weights of animals should be determined shortly before the test substance is administered and at least weekly thereafter. Weight changes should be calculated and recorded. At the end of the test surviving animals are weighed and then humanely killed.

**▼ B**1.6.2. **Pathology**

All test animals (including those that die during the test or are removed from the study for animal welfare reasons) should be subjected to gross necropsy. All gross pathological changes should be recorded for each animal. Microscopic examination of organs showing evidence of gross pathology in animals surviving 24 or more hours after the initial dosing may also be considered because it may yield useful information.

2. **DATA**

Individual animal data should be provided. Additionally, all data should be summarised in tabular form, showing for each test group the number of animals used, the number of animals displaying signs of toxicity, the number of animals found dead during the test or killed for humane reasons, time of death of individual animals, a description and the time course of toxic effects and reversibility, and necropsy findings.

3. **REPORTING**3.1. **TEST REPORT**

The test report must include the following information, as appropriate:

Test substance:

- physical nature, purity, and, where relevant, physico-chemical properties (including isomerisation),
- identification data, including CAS number.

Vehicle (if appropriate):

- justification for choice of vehicle, if other than water.

Test animals:

- species/strain used,
- microbiological status of the animals, when known,
- number, age and sex of animals (including, where appropriate, a rationale for use of males instead of females),
- source, housing conditions, diet, etc.

Test conditions:

- details of test substance formulation, including details of the physical form of the material administered,
- details of the administration of the test substance including dosing volumes and time of dosing,
- details of food and water quality (including diet type/source, water source),
- the rationale for the selection of the starting dose.



**▼ B**

## Results:

- tabulation of response data and dose level for each animal (i.e. animals showing signs of toxicity including mortality, nature, severity and duration of effects),
- tabulation of body weight and body weight changes,
- individual weights of animals at the day of dosing, in weekly intervals thereafter, and at time of death or sacrifice,
- date and time of death if prior to scheduled sacrifice,
- time course of onset of signs of toxicity and whether these were reversible for each animal,
- necropsy findings and histopathological findings for each animal, if available.

Discussion and interpretation of results.

Conclusions.

## 4.

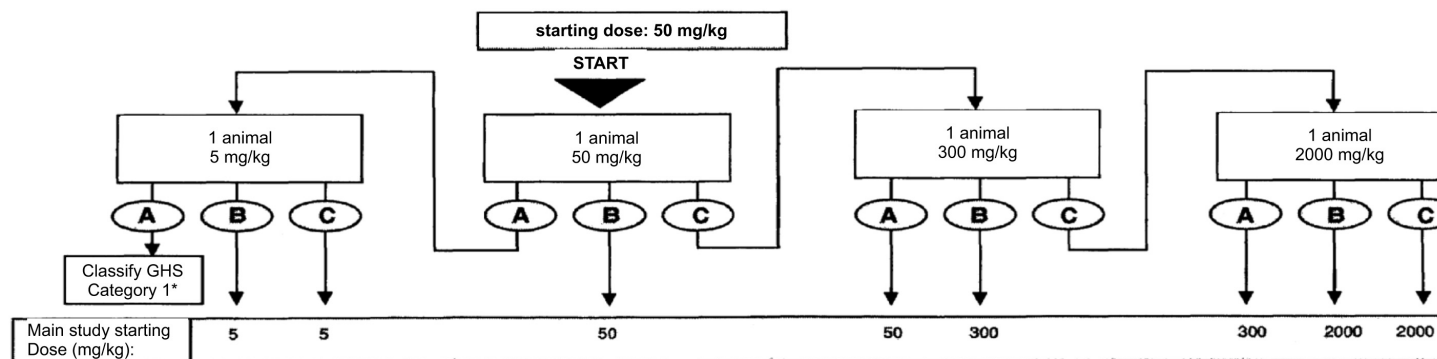
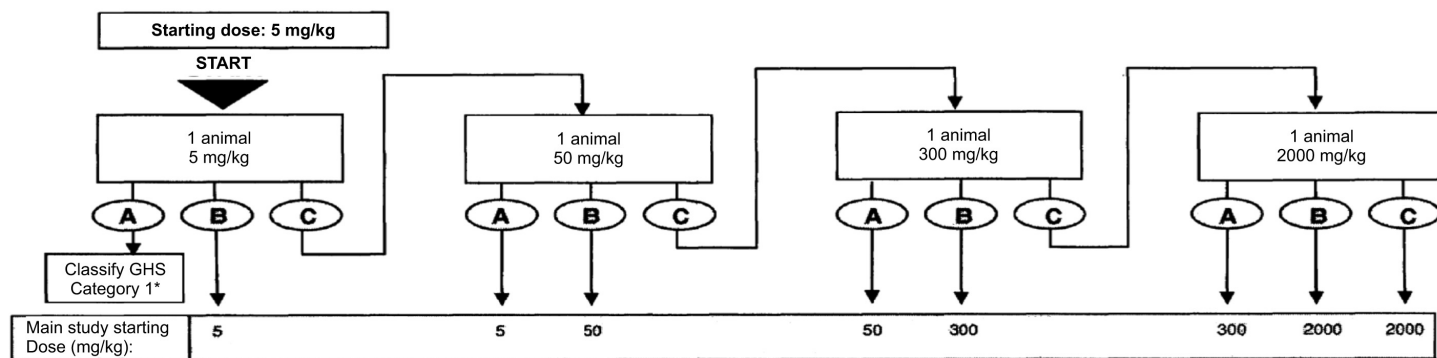
**REFERENCES**

- (1) British Toxicology Society Working Party on Toxicity (1984) Special report: a new approach to the classification of substances and preparations on the basis of their acute toxicity. *Human Toxicol.*, 3, p. 85-92.
- (2) Van den Heuvel, M.J., Dayan, A.D. and Shillaker, R.O (1987) Evaluation of the BTS approach to the testing of substances and preparations for their acute toxicity. *Human Toxicol.*, 6, p. 279-291.
- (3) Van den Heuvel, M.J., Clark, D.G., Fielder, R.J., Koundakjian, P.P., Oliver, G.J.A., Pelling, D., Tomlinson, N.J. and Walker, A.P (1990) The international validation of a fixed-dose procedure as an alternative to the classical LD<sub>50</sub> test. *Fd. Chem. Toxicol.* 28, p. 469-482.
- (4) Whitehead, A. and Curnow, R.N (1992) Statistical evaluation of the fixed-dose procedure. *Fd. Chem. Toxicol.*, 30, p. 313-324.
- (5) Stallard, N. and Whitehead, A (1995) Reducing numbers in the fixed-dose procedure. *Human Exptl. Toxicol.* 14, p. 315-323.
- (6) Stallard, N., Whitehead, A and Ridgeway, P. (2002) Statistical evaluation of the revised fixed dose procedure. *Hum. Exp. Toxicol.*, 21, p. 183-196.
- (7) OECD (2001) Guidance Document on Acute Oral Toxicity Testing. Environmental Health and Safety Monograph Series on Testing and Assessment No 24. Paris
- (8) OECD (2000) Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation. Environmental Health and Safety Monograph Series on Testing and Assessment No 19.

**▼B**

- (9) OECD (1998) Harmonised Integrated Hazard Classification for Human Health and Environmental Effects of Chemical Substances as endorsed by the 28th Joint Meeting of the Chemicals Committee and the Working Party on Chemicals in November 1998, Part 2, p. 11 [<http://webnet1.oecd.org/-oecd/pages/home/displaygeneral/0,3380, EN-documents-521-14-no-24-no-0,FF.html>].
- (10) Lipnick, R.L., Cotruvo, J.A., Hill, R.N., Bruce, R.D., Stitzel, K.A., Walker, A.P., Chu, I., Goddard, M., Segal, L., Springer, J.A. and Myers, R.C (1995) Comparison of the Up-and-Down, Conventional LD<sub>50</sub>, and Fixed-Dose Acute Toxicity Procedures. *Fd. Chem. Toxicol.* 33, p. 223-231.
- (11) Chan P.K and A. W Hayes (1994) Chapter 16 Acute Toxicity and Eye Irritation. In: *Principles and Methods of Toxicology*. 3rd Edition. A.W. Hayes, Editor. Raven Press Ltd. New York, USA.

FLOW CHART FOR THE SIGHTING STUDY



Outcome

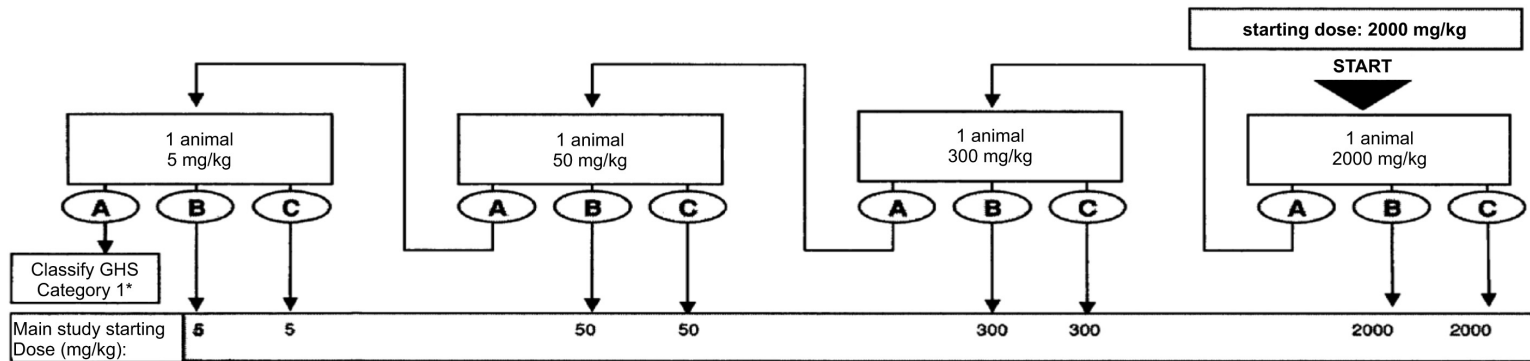
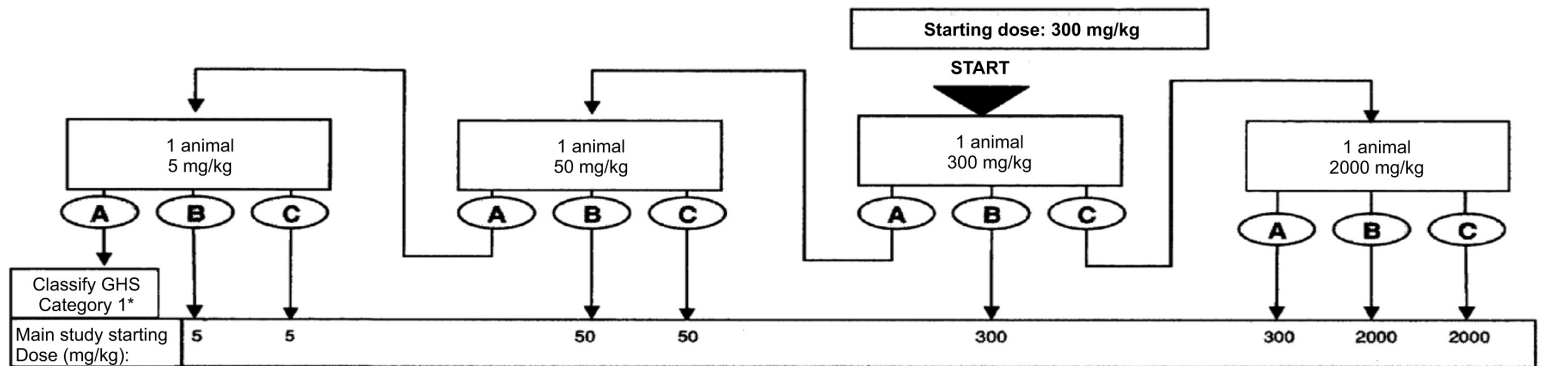
**A** death

**B** evident toxicity

**C** no evident toxicity and no death

\* for outcome **A** at 5 mg/kg there is an optional supplementary procedure to confirm the GHS classification: see section 1.5.2

▼B

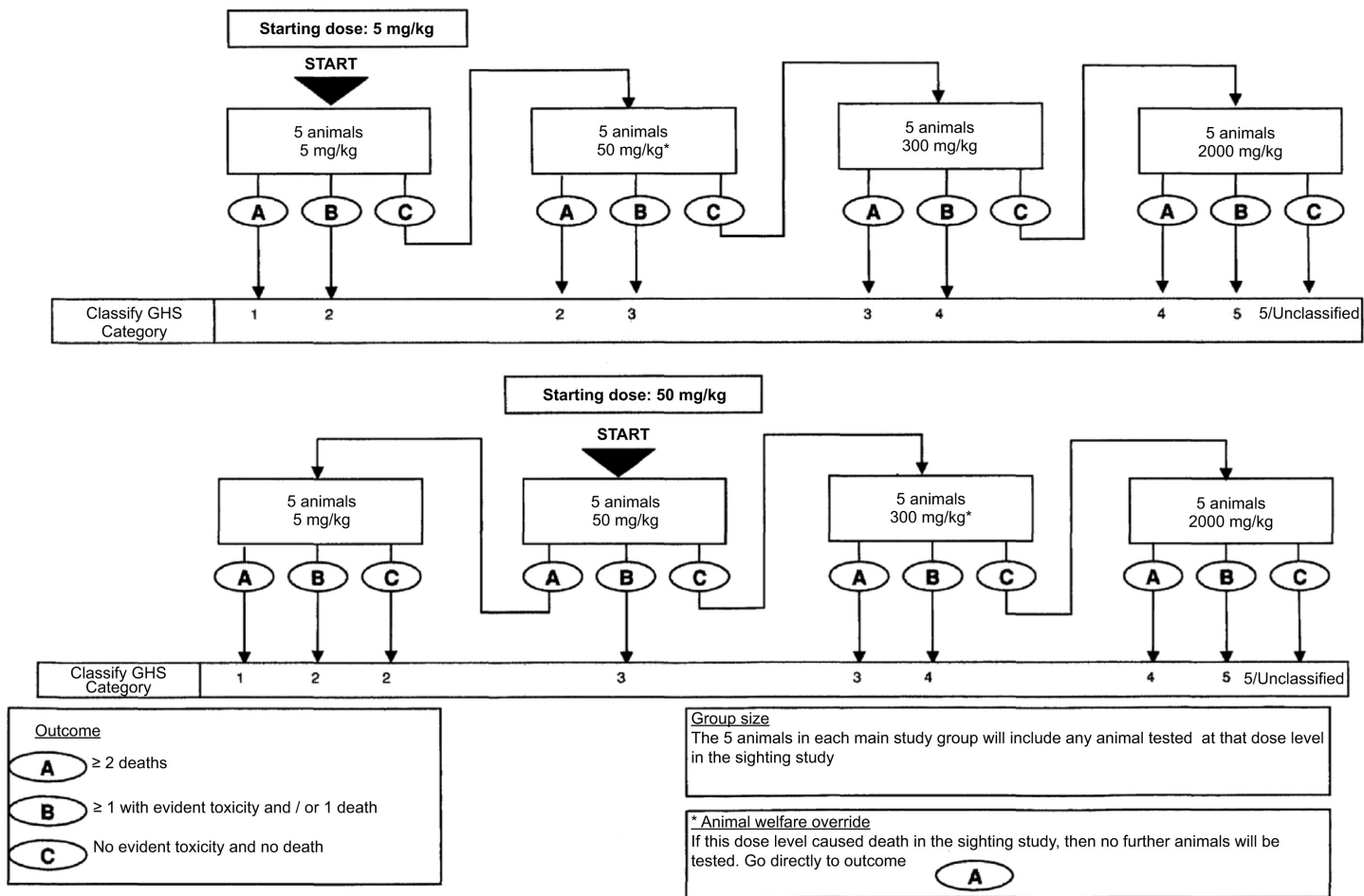


Outcome

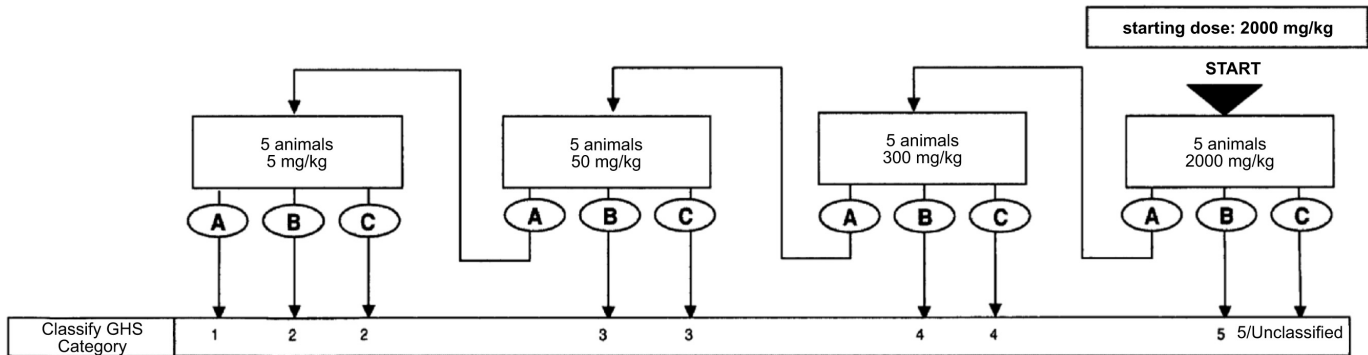
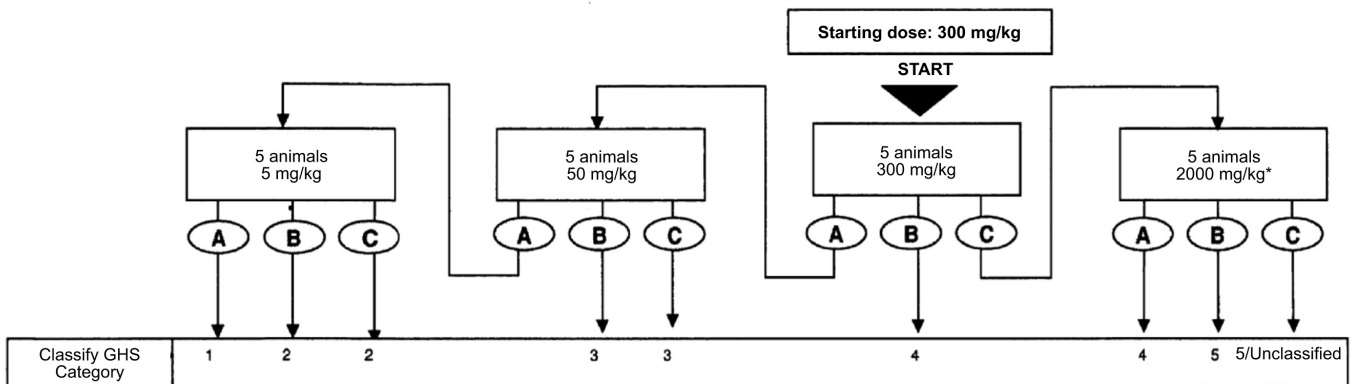
<b>A</b>	death
<b>B</b>	evident toxicity
<b>C</b>	no evident toxicity and no death

\* for outcome **A** at 5 mg/kg there is an optional supplementary procedure to confirm the GHS classification: see section 1.5.2

FLOW CHART FOR THE MAIN STUDY



▼B



<u>Outcome</u>	
<b>A</b>	≥ 2 deaths
<b>B</b>	≥ 1 with evident toxicity and / or 1 death
<b>C</b>	No evident toxicity and no death

Group size  
The 5 animals in each main study group will include any animal tested at that dose level in the sighting study

\* Animal welfare override  
If this dose level caused death in the sighting study, then no further animals will be tested. Go directly to outcome **A**

*Appendix 3***CRITERIA FOR CLASSIFICATION OF TEST SUBSTANCES WITH EXPECTED LD<sub>50</sub> VALUES EXCEEDING 2 000 MG/KG WITHOUT THE NEED FOR TESTING.**

Criteria for hazard Category 5 are intended to enable the identification of test substances which are of relatively low acute toxicity hazard but which, under certain circumstances may present a danger to vulnerable populations. These substances are anticipated to have an oral or dermal LD<sub>50</sub> in the range of 2 000-5 000 mg/kg or equivalent doses for other routes. Test substances could be classified in the hazard category defined by: 2 000 mg/kg < LD<sub>50</sub> < 5 000 mg/kg (Category 5 in the GHS) in the following cases:

- (a) if directed to this category by any of the testing schemes of Appendix 2, based on mortality incidences
  
- (b) if reliable evidence is already available that indicates the LD<sub>50</sub> to be in the range of Category 5 values; or other animal studies or toxic effects in humans indicate a concern for human health of an acute nature;
  
- (c) through extrapolation, estimation or measurement of data if assignment to a more hazardous class is not warranted; and
  - reliable information is available indicating significant toxic effects in humans, or
  
  - any mortality is observed when tested up to Category 4 values by the oral route, or
  
  - where expert judgement confirms significant clinical signs of toxicity, when tested up to Category 4 values, except for diarrhoea, piloerection or an ungroomed appearance, or
  
  - where expert judgement confirms reliable information indicating the potential for significant acute effects from the other animal studies.

**TESTING AT DOSES ABOVE 2 000 MG/KG**

Exceptionally, and only when justified by specific regulatory needs, the use of an additional upper fixed dose level of 5 000 mg/kg may be considered. Recognising the need to protect animal welfare, testing at 5 000 mg/kg is discouraged and should only be considered when there is a strong likelihood that the results of such a test would have a direct relevance for protecting animal or human health (9).

**Sighting study**

The decision rules governing the sequential procedure presented in Appendix 1 are extended to include a 5 000 mg/kg dose level. Thus, when a sighting study starting dose of 5 000 mg/kg is used outcome A (death) will require a second animal to be tested at 2 000 mg/kg; outcomes B and C (evident toxicity or no toxicity) will allow the selection of 5 000 mg/kg as the main study starting dose. Similarly, if a starting dose other than 5 000 mg/kg is used then testing will progress to 5 000 mg/kg in the event of outcomes B or C at 2 000 mg/kg; a subsequent 5 000 mg/kg outcome A will dictate a main study starting dose of 2 000 mg/kg and outcomes B and C will dictate a main study starting dose of 5 000 mg/kg.

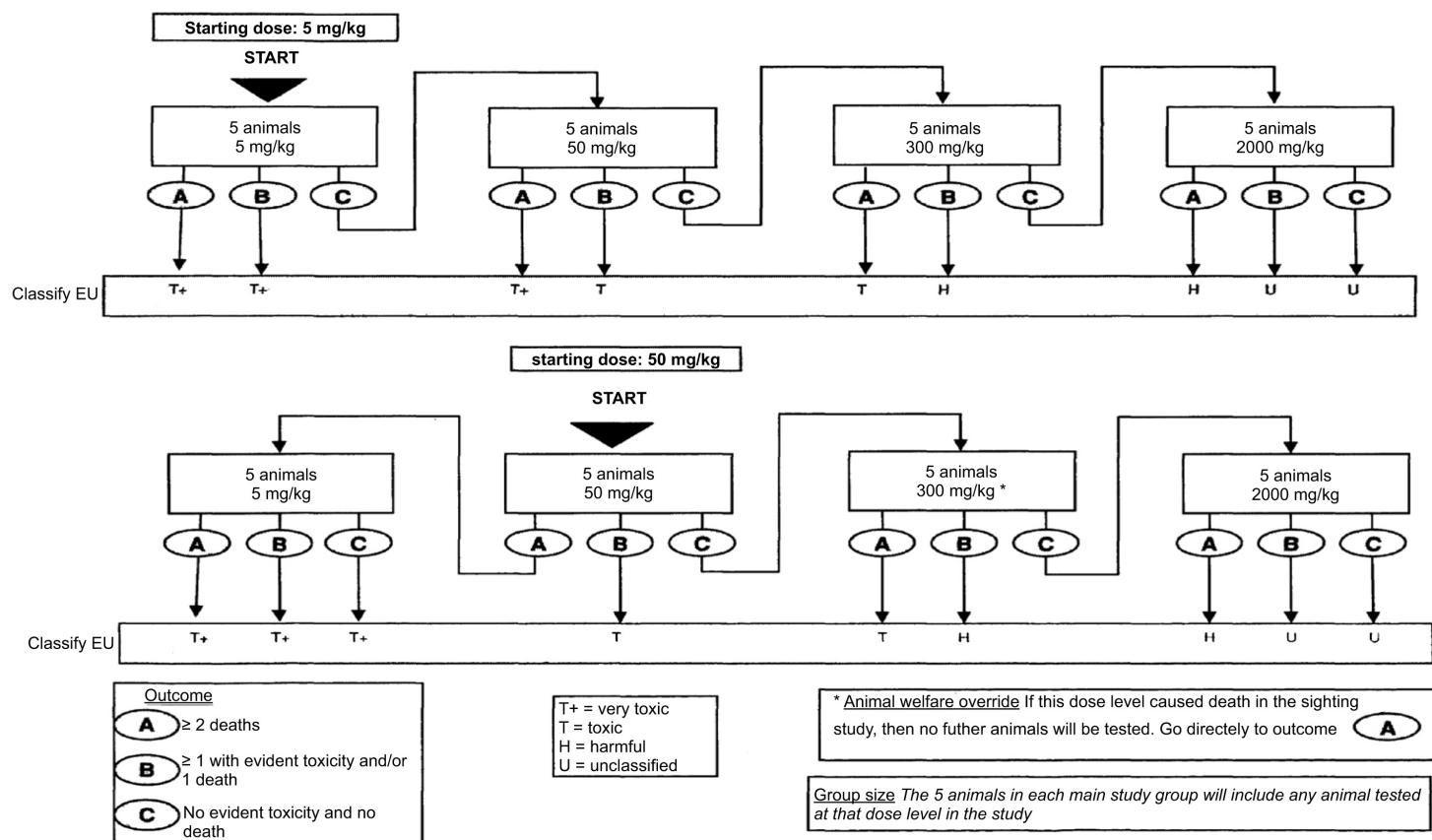
**▼B****Main study**

The decision rules governing the sequential procedure presented in Appendix 2 are extended to include a 5 000 mg/kg dose level. Thus, when a main study starting dose of 5 000 mg/kg is used, outcome A ( $\geq 2$  deaths) will require the testing of a second group at 2 000 mg/kg; outcome B (evident toxicity and/or  $\leq 1$  death) or C (no toxicity) will result in the substance being unclassified according to GHS. Similarly, if a starting dose other than 5 000 mg/kg is used then testing will progress to 5 000 mg/kg in the event of outcome C at 2 000 mg/kg; a subsequent 5 000 mg/kg outcome A will result in the substance being assigned to GHS Category 5 and outcomes B or C will lead to the substance being unclassified.

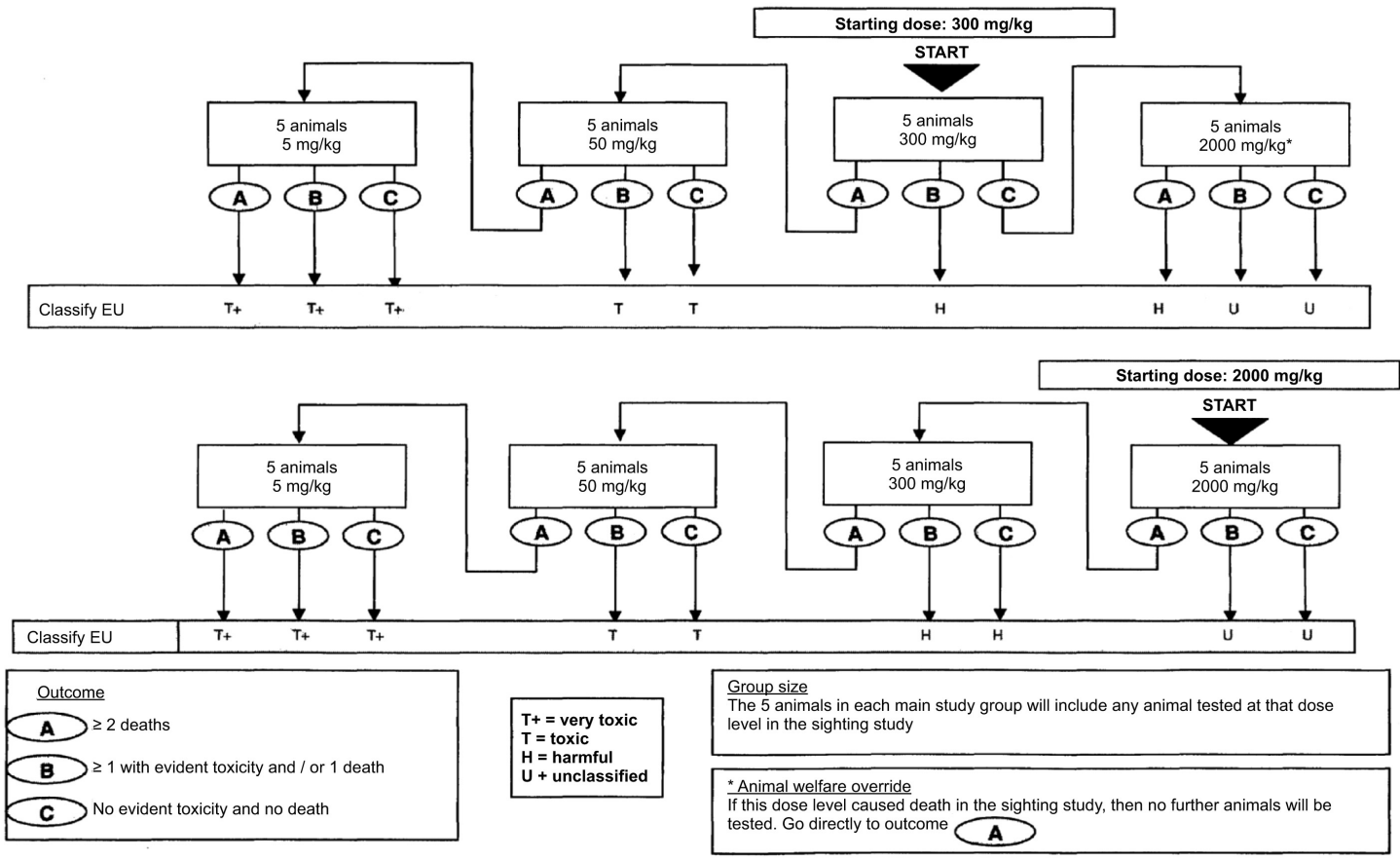


TEST METHOD B.1 bis

Guidance on classification according to the EU scheme to cover the transition period until full implementation of the Globally Harmonised Classification System (GHS) (taken from reference (8))



▼B



**▼B****B.1 tris. ACUTE ORAL TOXICITY — ACUTE TOXIC CLASS METHOD****1. METHOD**

This test method is equivalent to OECD TG 423 (2001)

**1.1. INTRODUCTION**

The acute toxic class method (1) set out in this test is a stepwise procedure with the use of three animals of a single sex per step. Depending on the mortality and/or the moribund status of the animals, on average two to four steps may be necessary to allow judgement on the acute toxicity of the test substance. This procedure is reproducible, uses very few animals and is able to rank substances in a similar manner to the other acute toxicity testing methods. The acute toxic class method is based on biometric evaluations (2)(3)(4)(5) with fixed doses, adequately separated to enable a substance to be ranked for classification purposes and hazard assessment. The method as adopted in 1996 was extensively validated *in vivo* against LD<sub>50</sub> data obtained from the literature, both nationally (6) and internationally (7).

Guidance on the selection of the most appropriate test method for a given purpose can be found in the Guidance Document on Acute Oral Toxicity Testing (8). This Guidance Document also contains additional information on the conduct and interpretation of testing method B.1tris.

Test substances, at doses that are known to cause marked pain and distress due to corrosive or severely irritant actions, need not be administered. Moribund animals, or animals obviously in pain or showing signs of severe and enduring distress shall be humanely killed, and are considered in the interpretation of the test results in the same way as animals that died on test. Criteria for making the decision to kill moribund or severely suffering animals, and guidance on the recognition of predictable or impending death, are the subject of a separate Guidance Document (9).

The method uses pre-defined doses and the results allow a substance to be ranked and classified according to the Globally Harmonised System for the classification of chemicals which cause acute toxicity (10).

In principle, the method is not intended to allow the calculation of a precise LD<sub>50</sub>, but does allow for the determination of defined exposure ranges where lethality is expected since death of a proportion of the animals is still the major endpoint of this test. The method allows for the determination of an LD<sub>50</sub> value only when at least two doses result in mortality higher than 0 % and lower than 100 %. The use of a selection of pre-defined doses, regardless of test substance, with classification explicitly tied to number of animals observed in different states improves the opportunity for laboratory to laboratory reporting consistency and repeatability.

**▼B**

The testing laboratory should consider all available information on the test substance prior to conducting the study. Such information will include the identity and chemical structure of the substance; its physico-chemical properties; the result of any other *in vivo* or *in vitro* toxicity tests on the substance; toxicological data on the structurally related substances; and the anticipated use(s) of the substance. This information is necessary to satisfy all concerned that the test is relevant for the protection of human health and will help in the selection of the most appropriate starting dose.

## 1.2. DEFINITIONS

**Acute oral toxicity:** refers to those adverse effects occurring following oral administration of a single dose of a substance or multiple doses given within 24 hours.

**Delayed death:** means that an animal does not die or appear moribund within 48 hours but dies later during the 14-day observation period.

**Dose:** is the amount of test substance administered. Dose is expressed as weight of test substance per unit weight of test animal (e.g. mg/kg).

**GHS:** Globally Harmonised Classification System for Chemical Substances and Mixtures. A joint activity of OECD (human health and the environment), UN Committee of Experts on Transport of Dangerous Goods (physical-chemical properties) and ILO (hazard communication) and coordinated by the Interorganisation Programme for the Sound Management of Chemicals (IOMC).

**Impending death:** when moribund state or death is expected prior to the next planned time of observation. Signs indicative of this state in rodents could include convulsions, lateral position, recumbence and tremor (See the Humane Endpoint Guidance Document (9) for more details).

**LD<sub>50</sub> (median lethal oral dose):** is a statistically derived single dose of a substance that can be expected to cause death in 50 % of animals when administered by the oral route. The LD<sub>50</sub> value is expressed in terms of weight of test substance per unit weight of test animal (mg/kg).

**Limit dose:** refers to a dose at an upper limitation on testing (2 000 or 5 000 mg/kg).

**Moribund status:** being in a state of dying or inability to survive, even if treated (See the Humane Endpoint Guidance Document (9) for more details).

**Predictable death:** presence of clinical signs indicative of death at a known time in the future before the planned end of the experiment; for example: inability to reach water or food. (See the Humane Endpoint Guidance Document (9) for more details).

**▼B**

## 1.3. PRINCIPLE OF THE TEST

It is the principle of the test that, based on a stepwise procedure with the use of a minimum number of animals per step, sufficient information is obtained on the acute toxicity of the test substance to enable its classification. The substance is administered orally to a group of experimental animals at one of the defined doses. The substance is tested using a stepwise procedure, each step using three animals of a single sex (normally females). Absence or presence of compound-related mortality of the animals dosed at one step will determine the next step, i.e.;

— no further testing is needed,

— dosing of three additional animals, with the same dose,

— dosing of three additional animals at the next higher or the next lower dose level.

Details of the test procedure are described in Appendix 1. The method will enable a judgement with respect to classifying the test substance to one of a series of toxicity classes defined by fixed LD<sub>50</sub> cut-off values.

## 1.4. DESCRIPTION OF THE METHOD

1.4.1. **Selection of animal species**

The preferred rodent species is the rat, although other rodent species may be used. Normally females are used (9). This is because literature surveys of conventional LD<sub>50</sub> tests show that, although there is little difference in sensitivity between the sexes, in those cases where differences are observed females are generally slightly more sensitive (11). However if knowledge of the toxicological or toxicokinetic properties of structurally related chemicals indicates that males are likely to be more sensitive, then this sex should be used. When the test is conducted in males, adequate justification should be provided.

Healthy young adult animals of commonly used laboratory strains should be employed. Females should be nulliparous and non-pregnant. Each animal, at the commencement of its dosing, should be between eight and 12 weeks old and its weight should fall in an interval within  $\pm 20$  % of the mean weight of any previously dosed animals.

1.4.2. **Housing and feeding conditions**

The temperature in the experimental animal room should be 22 °C ( $\pm 3$  °C). Although the relative humidity should be at least 30 % and preferably not exceed 70 % other than during room cleaning the aim should be 50-60 %. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water. Animals may be group-caged by dose, but the number of animals per cage must not interfere with clear observations of each animal.

**▼B****1.4.3. Preparation of animals**

The animals are randomly selected, marked to permit individual identification, and kept in their cages for at least five days prior to dosing to allow for acclimatisation to the laboratory conditions.

**1.4.4. Preparation of doses**

In general, test substances should be administered in a constant volume over the range of doses to be tested by varying the concentration of the dosing preparation. Where a liquid end product or mixture is to be tested however, the use of the undiluted test substance, i.e. at a constant concentration, may be more relevant to the subsequent risk assessment of that substance, and is a requirement of some regulatory authorities. In either case, the maximum dose volume for administration must not be exceeded. The maximum volume of liquid that can be administered at one time depends on the size of the test animal. In rodents, the volume should not normally exceed 1 ml/100 g of body weight: however in the case of aqueous solutions 2 ml/100 g body weight can be considered. With respect to the formulation of the dosing preparation, the use of an aqueous solution/suspension/emulsion is recommended wherever possible, followed in order of preference by a solution/suspension/emulsion in oil (e.g. corn oil) and then possibly solution in other vehicles. For vehicles other than water the toxicological characteristics of the vehicle should be known. Doses must be prepared shortly prior to administration unless the stability of the preparation over the period during which it will be used is known and shown to be acceptable.

**1.5. PROCEDURE****1.5.1. Administration of doses**

The test substance is administered in a single dose by gavage using a stomach tube or a suitable intubation canula. In the unusual circumstance that a single dose is not possible, the dose may be given in smaller fractions over a period not exceeding 24 hours.

Animals should be fasted prior to dosing (e.g. with the rat, food but not water should be withheld overnight, with the mouse, food but not water should be withheld for three or four hours). Following the period of fasting, the animals should be weighed and the test substance administered. After the substance has been administered, food may be withheld for a further three or four hours in rats or one or two hours in mice. Where a dose is administered in fractions over a period it may be necessary to provide the animals with food and water depending on the length of the period.

**1.5.2. Number of animals and dose levels**

Three animals are used for each step. The dose level to be used as the starting dose is selected from one of four fixed levels, 5, 50, 300 and 2 000 mg/kg body weight. The starting dose level should be that which is most likely to produce mortality in some of the dosed animals. The flowcharts of Appendix 1 describe the procedure that should be followed for each of the starting doses. In addition, Appendix 4 gives guidance on the classification in the EU system until the new GHS is implemented.

**▼B**

When available information suggests that mortality is unlikely at the highest starting dose level (2 000 mg/kg body weight), then a limit test should be conducted. When there is no information on a substance to be tested, for animal welfare reasons it is recommended to use the starting dose of 300 mg/kg body weight.

The time interval between treatment groups is determined by the onset, duration, and severity of toxic signs. Treatment of animals at the next dose should be delayed until one is confident of survival of the previously dosed animals.

Exceptionally, and only when justified by specific regulatory needs, the use of additional upper dose level of 5 000 mg/kg body weight may be considered (see Appendix 2). For reasons of animal welfare concern, testing of animals in GHS Category 5 ranges (2 000-5 000 mg/kg) is discouraged and should only be considered when there is a strong likelihood that the results of such a test would have a direct relevance for protecting human or animal health or the environment.

**1.5.3. Limit test**

The limit test is primarily used in situations where the experimenter has information indicating that the test material is likely to be non-toxic, i.e., having toxicity only above regulatory limit doses. Information about the toxicity of the test material can be gained from knowledge about similar tested compounds or similar tested mixtures or products, taking into consideration the identity and percentage of components known to be of toxicological significance. In those situations where there is little or no information about its toxicity, or in which the test material is expected to be toxic, the main test should be performed.

A limit test at one dose level of 2 000 mg/kg body weight may be carried out with six animals (three animals per step). Exceptionally a limit test at one dose level of 5 000 mg/kg may be carried out with three animals (see Appendix 2). If test substance-related mortality is produced, further testing at the next lower level may need to be carried out.

**1.6. OBSERVATIONS**

Animals are observed individually after dosing at least once during the first 30 minutes, periodically during the first 24 hours, with special attention given during the first four hours, and daily thereafter, for a total of 14 days, except where they need to be removed from the study and humanely killed for animal welfare reasons or are found dead. However, the duration of observation should not be fixed rigidly. It should be determined by the toxic reactions, time of onset and length of recovery period, and may thus be extended when considered necessary. The times at which signs of toxicity appear and disappear are important, especially if there is a tendency for toxic signs to be delayed (12). All observations are systematically recorded with individual records being maintained for each animal.

**▼B**

Additional observations will be necessary if the animals continue to display signs of toxicity. Observations should include changes in skin and fur, eyes and mucous membranes, and also respiratory, circulatory, autonomic and central nervous systems, and somatomotor activity and behaviour pattern. Attention should be directed to observations of tremors, convulsions, salivation, diarrhoea, lethargy, sleep and coma. The principles and criteria summarised in the Humane Endpoints Guidance Document (9) should be taken into consideration. Animals found in a moribund condition and animals showing severe pain or enduring signs of severe distress should be humanely killed. When animals are killed for humane reasons or found dead, the time of death should be recorded as precisely as possible.

**1.6.1. Body weight**

Individual weights of animals should be determined shortly before the test substance is administered, and at least weekly thereafter. Weight changes should be calculated and recorded. At the end of the test surviving animals are weighed and humanely killed.

**1.6.2. Pathology**

All test animals (including those that die during the test or are removed from the study for animal welfare reasons) should be subjected to gross necropsy. All gross pathological changes should be recorded for each animal. Microscopic examination of organs showing evidence of gross pathology in animals surviving 24 or more hours may also be considered because it may yield useful information.

**2. DATA**

Individual animal data should be provided. Additionally, all data should be summarised in tabular form, showing for each test group the number of animals used, the number of animals displaying signs of toxicity, the number of animals found dead during the test or killed for humane reasons, time of death of individual animals, a description and the time course of toxic effects and reversibility, and necropsy findings.

**3. REPORTING****3.1. Test report**

The test report must include the following information, as appropriate:

Test substance:

- physical nature, purity, and, where relevant, physico-chemical properties (including isomerisation),
- identification data, including CAS number.

Vehicle (if appropriate):

- justification for choice of vehicle, if other than water.

Test animals:

- species/strain used,



**▼ B**

- microbiological status of the animals, when known,
- number, age, and sex of animals (including, where appropriate, a rationale for the use of males instead of females),
- source, housing conditions, diet, etc.

## Test conditions:

- details of test substance formulation including details of the physical form of the material administered,
- details of the administration of the test substance including dosing volumes and time of dosing,
- details of food and water quality (including diet type/source, water source),
- the rationale for the selection of the starting dose.

## Results:

- tabulation of response data and dose level for each animal (i.e. animals showing signs of toxicity including mortality; nature, severity, and duration of effects),
- tabulation of body weight and body weight changes,
- individual weights of animals at the day of dosing, in weekly intervals thereafter, and at the time of death or sacrifice,
- date and time of death if prior to scheduled sacrifice,
- time course of onset of signs of toxicity, and whether these were reversible for each animal,
- necropsy findings and histopathological findings for each animal, if available.

## Discussion and interpretation of results.

## Conclusions.

4. **REFERENCES**

- (1) Roll R., Höfer-Bosse Th. And Kayser D (1986) New Perspectives in Acute Toxicity Testing of Chemicals. Toxicol. Lett., Suppl. 31, p. 86.
- (2) Roll R., Riebschläger M., Mischke U. and Kayser D (1989) Neue Wege zur Bestimmung der akuten Toxizität von Chemikalien. Bundesgesundheitsblatt 32, p. 336-341.
- (3) Diener W., Sichha L., Mischke U., Kayser D. and Schlede E (1994) The Biometric Evaluation of the Acute-Toxic-Class Method (Oral). Arch. Toxicol. 68, p. 559-610.
- (4) Diener W., Mischke U., Kayser D. and Schlede E., (1995) The Biometric Evaluation of the OECD Modified Version of the Acute-Toxic-Class Method (Oral). Arch. Toxicol. 69, p. 729-734.
- (5) Diener W., and Schlede E., (1999) Acute Toxicity Class Methods: Alterations to LD/LC<sub>50</sub> Tests. ALTEX 16, p. 129-134.
- (6) Schlede E., Mischke U., Roll R. and Kayser D., (1992). A National Validation Study of the Acute-Toxic- Class Method — An Alternative to the LD<sub>50</sub> Test. Arch. Toxicol. 66, 455-470.

**▼B**

- (7) Schlede E., Mischke U., Diener W. and Kayser D., (1994) The International Validation Study of the Acute-Toxic-Class Method (Oral). *Arch. Toxicol.* 69, p. 659-670.
- (8) OECD, (2001) Guidance Document on Acute Oral Toxicity Testing. Environmental Health and Safety Monograph Series on Testing and Assessment No 24. Paris.
- (9) OECD, (2000) Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation. Environmental Health and Safety Monograph Series on Testing and Assessment No 19.
- (10) OECD, (1998) Harmonised Integrated Hazard Classification System For Human Health And Environmental Effects Of Chemical Substances as endorsed by the 28th Joint Meeting of the Chemicals Committee and the Working Party on Chemicals in November 1998, Part 2, p. 11 [<http://webnet1.oecd.org/oecd/pages/home/displaygeneral/0,3380,EN-documents-521-14-no-24-no-0,FF.html>].
- (11) Lipnick R. L., Cotruvo, J.A., Hill R. N., Bruce R. D., Stitzel K. A., Walker A. P., Chu I.; Goddard M., Segal L., Springer J. A. and Myers R. C. (1995) Comparison of the Up-and Down, Conventional LD<sub>50</sub>, and Fixed Dose Acute Toxicity Procedures. *Fd. Chem. Toxicol* 33, p. 223-231.
- (12) Chan P.K. and A.W. Hayes. (1994). Chap. 16. Acute Toxicity and Eye Irritancy. *Principles and Methods of Toxicology*. Third Edition. A.W. Hayes, Editor. Raven Press, Ltd., New York, USA.

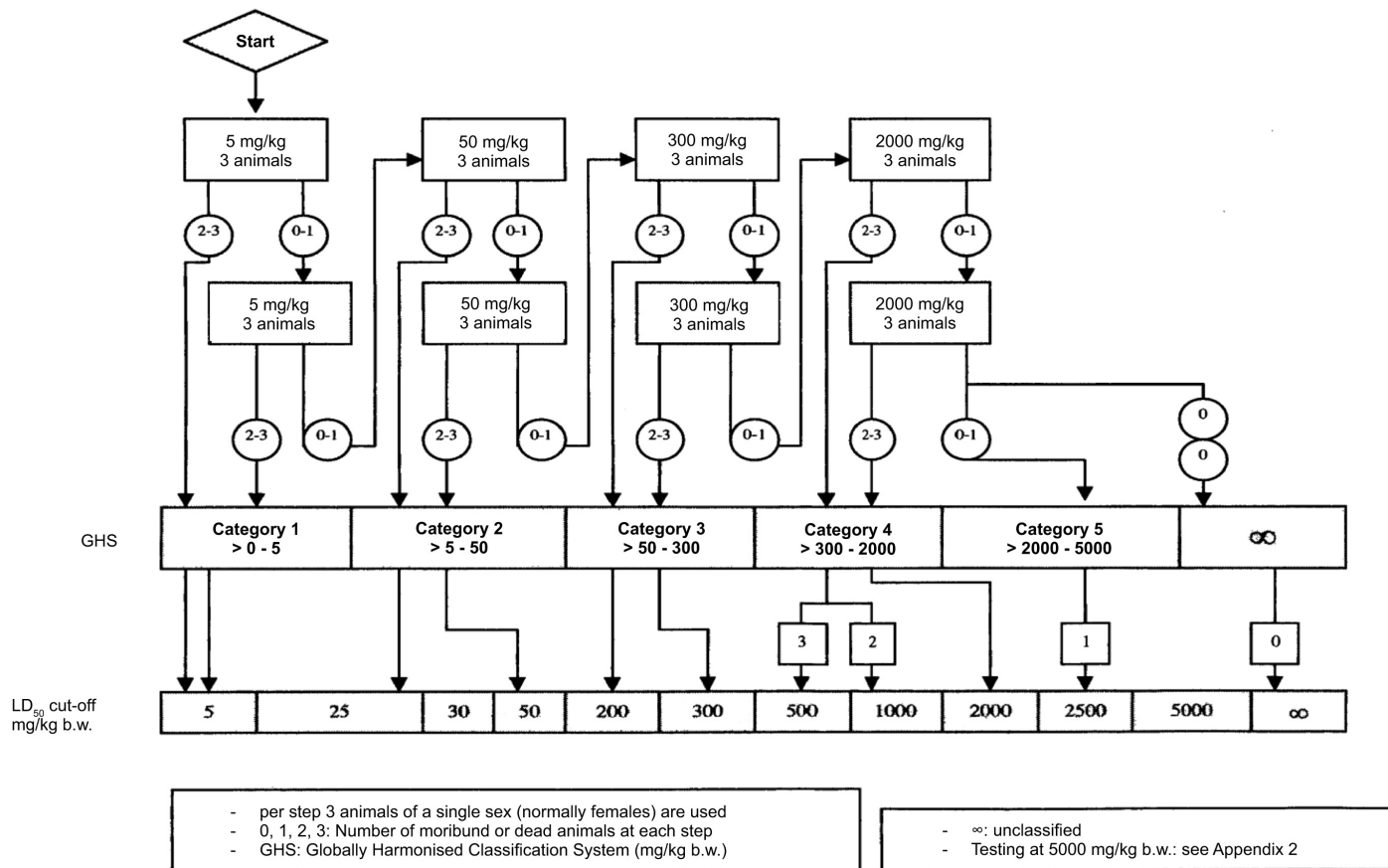
**▼B***Appendix 1***PROCEDURE TO BE FOLLOWED FOR EACH OF THE STARTING DOSES***GENERAL REMARKS*

For each starting dose, the respective testing schemes as included in this Appendix outline the procedure to be followed.

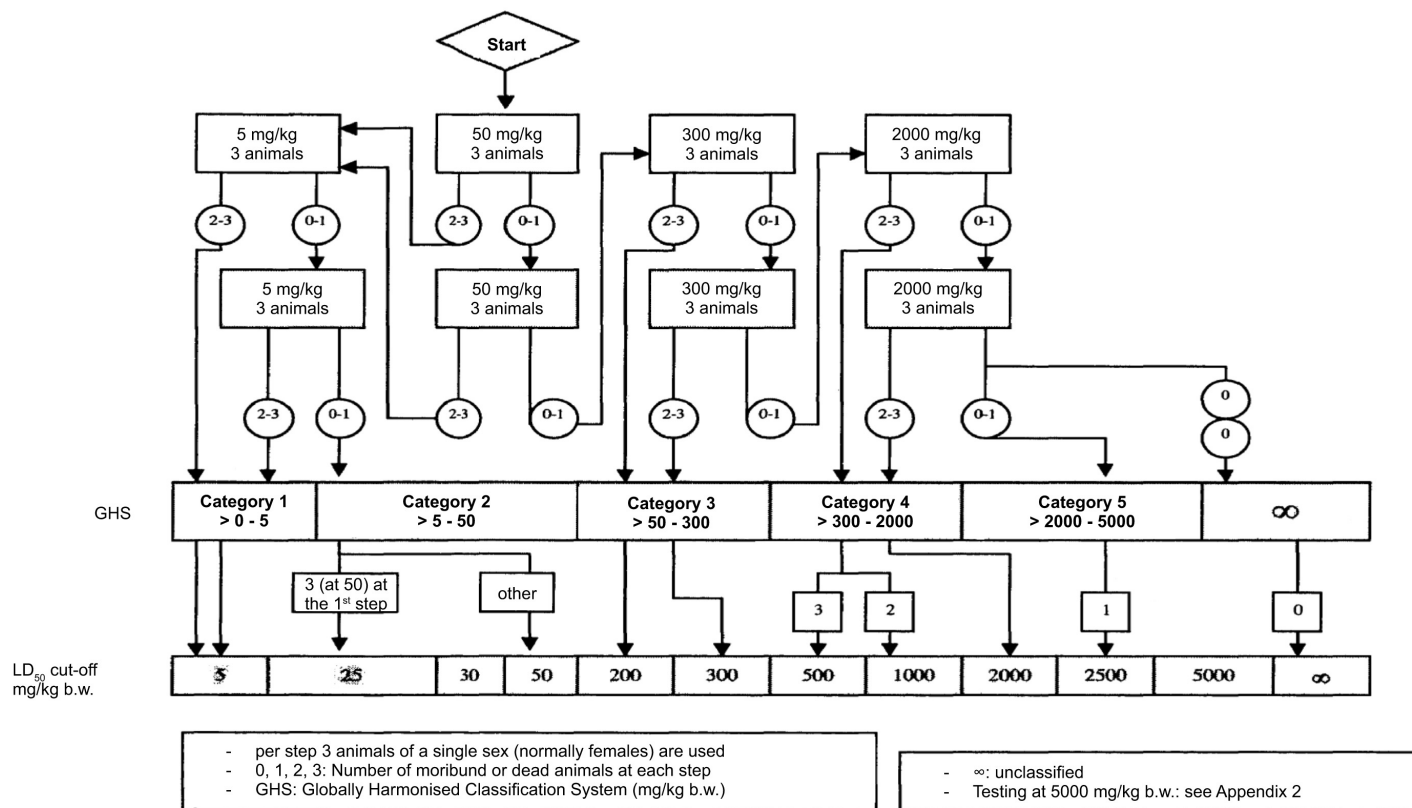
- Appendix 1 a: starting dose is 5 mg/kg bw,
- Appendix 1 b: starting dose is 50 mg/kg bw,
- Appendix 1 c: starting dose is: 300 mg/kg bw,
- Appendix 1 d: starting dose is: 2 000 mg/kg bw.

Depending on the number of humanely killed or dead animals, the test procedure follows the indicated arrows.

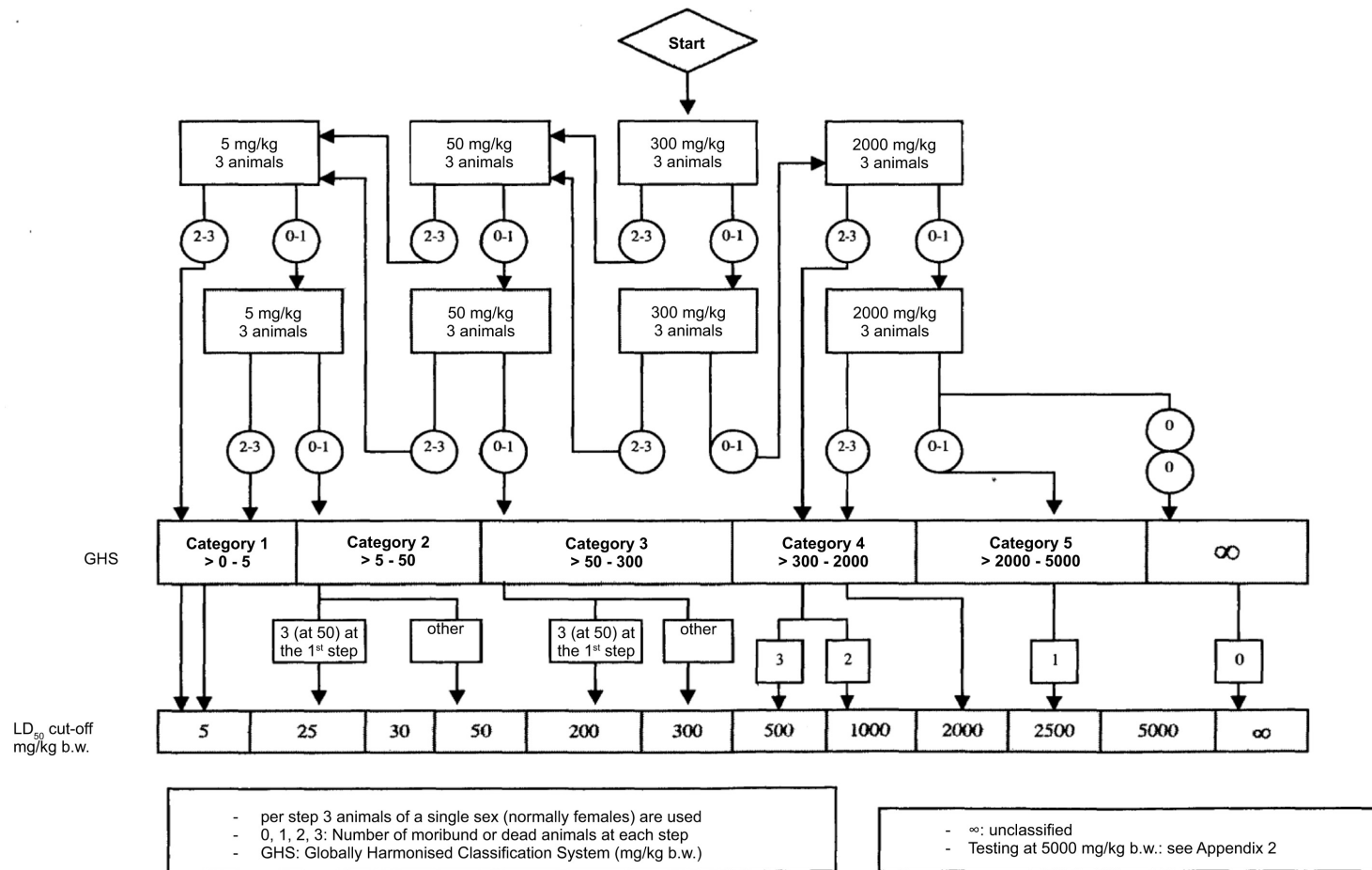
TEST PROCEDURE WITH A STARTING DOSE OF 5 MG/KG BODY WEIGHT



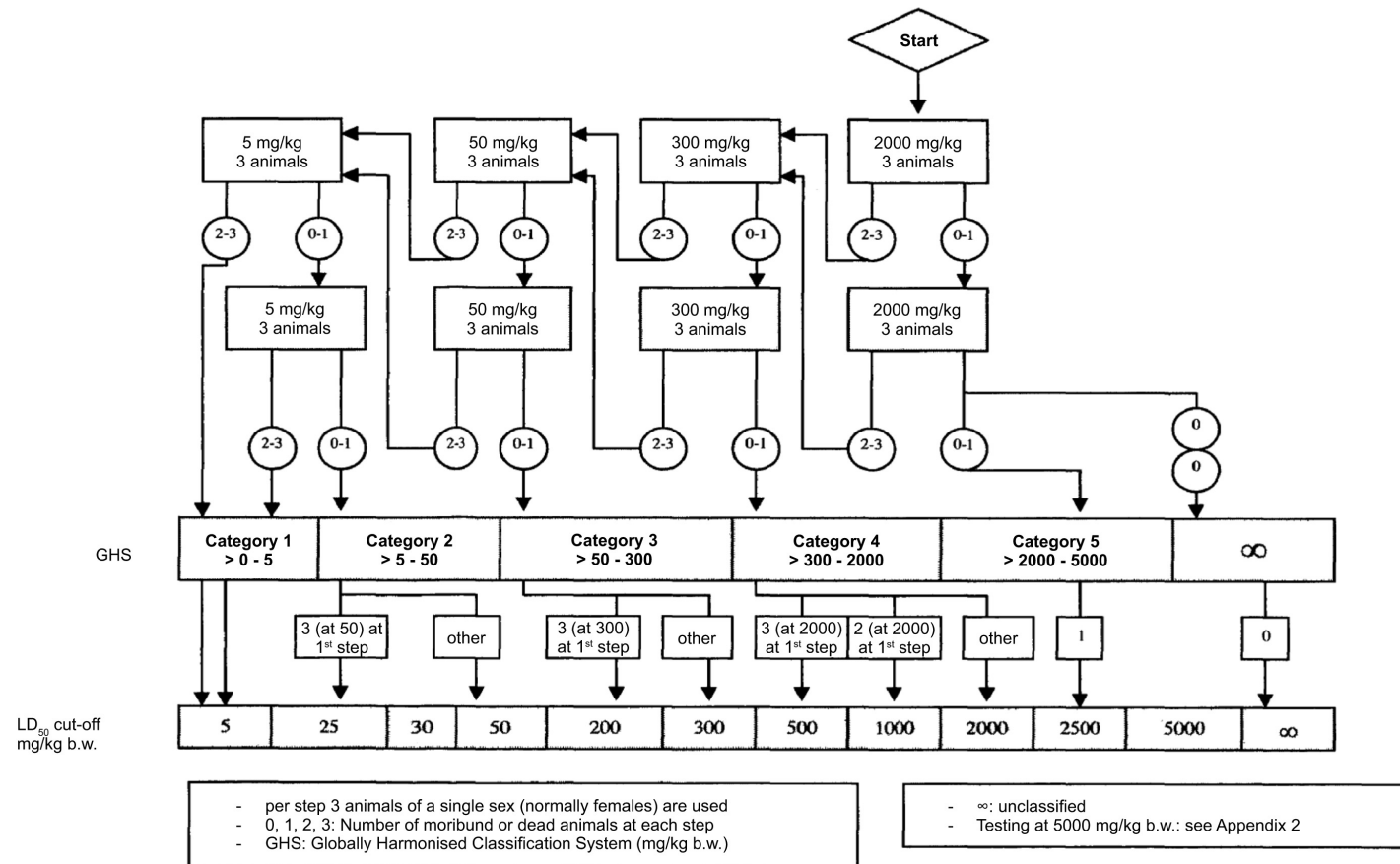
TEST PROCEDURE WITH A STARTING DOSE OF 50 MG/KG BODY WEIGHT



TEST PROCEDURE WITH A STARTING DOSE OF 300 MG/KG BODY WEIGHT



TEST PROCEDURE WITH A STARTING DOSE OF 2 000 MG/KG BODY WEIGHT



*Appendix 2***CRITERIA FOR CLASSIFICATION OF TEST SUBSTANCES WITH EXPECTED LD<sub>50</sub> VALUES EXCEEDING 2 000 MG/KG WITHOUT THE NEED FOR TESTING**

Criteria for hazard Category 5 are intended to enable the identification of test substances which are of relatively low acute toxicity hazard but which, under certain circumstances may present a danger to vulnerable populations. These substances are anticipated to have an oral or dermal LD<sub>50</sub> in the range of 2 000-5 000 mg/kg or equivalent doses for other routes. The test substance should be classified in the hazard category defined by: 2 000 mg/kg < LD<sub>50</sub> < 5 000 mg/kg (Category 5 in the GHS) in the following cases:

- (a) If directed to this category by any of the testing schemes of Appendix 1a-1d, based on mortality incidences;
- (b) if reliable evidence is already available that indicates the LD<sub>50</sub> to be in the range of Category 5 values; or other animal studies or toxic effects in humans indicate a concern for human health of an acute nature;
- (c) through extrapolation, estimation or measurement of data if assignment to a more hazardous class is not warranted; and
  - reliable information is available indicating significant toxic effects in humans, or
  - any mortality is observed when tested up to Category 4 values by the oral route, or
  - where expert judgement confirms significant clinical signs of toxicity, when tested up to Category 4 values, except for diarrhoea, piloerection or an ungroomed appearance, or
  - where expert judgement confirms reliable information indicating the potential for significant acute effects from the other animal studies.

**TESTING AT DOSES ABOVE 2 000 MG/KG**

Recognising the need to protect animal welfare, testing of animals in Category 5 (5 000 mg/kg) ranges is discouraged and should only be considered when there is a strong likelihood that results of such a test have a direct relevance for protecting human or animal health (10). No further testing should be conducted at higher dose levels.

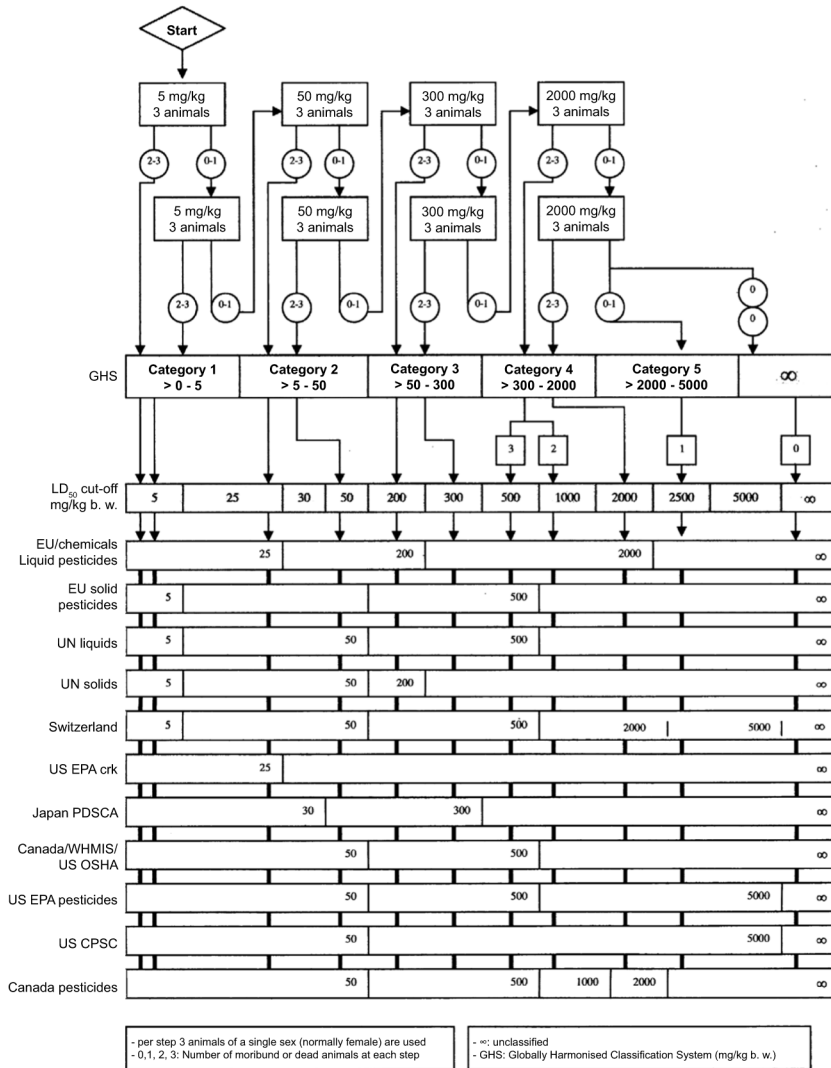
When testing is required a dose of 5 000 mg/kg, only one step (i.e. three animals) is required. If the first animal dosed dies, then dosing proceeds at 2 000 mg/kg in accordance with the flowcharts in Appendix 1. If the first animal survives, two further animals are dosed. If only one of the three animals dies, the LD<sub>50</sub> value is expected to exceed 5 000 mg/kg. If both animals die, then dosing proceeds at 2 000 mg/kg.



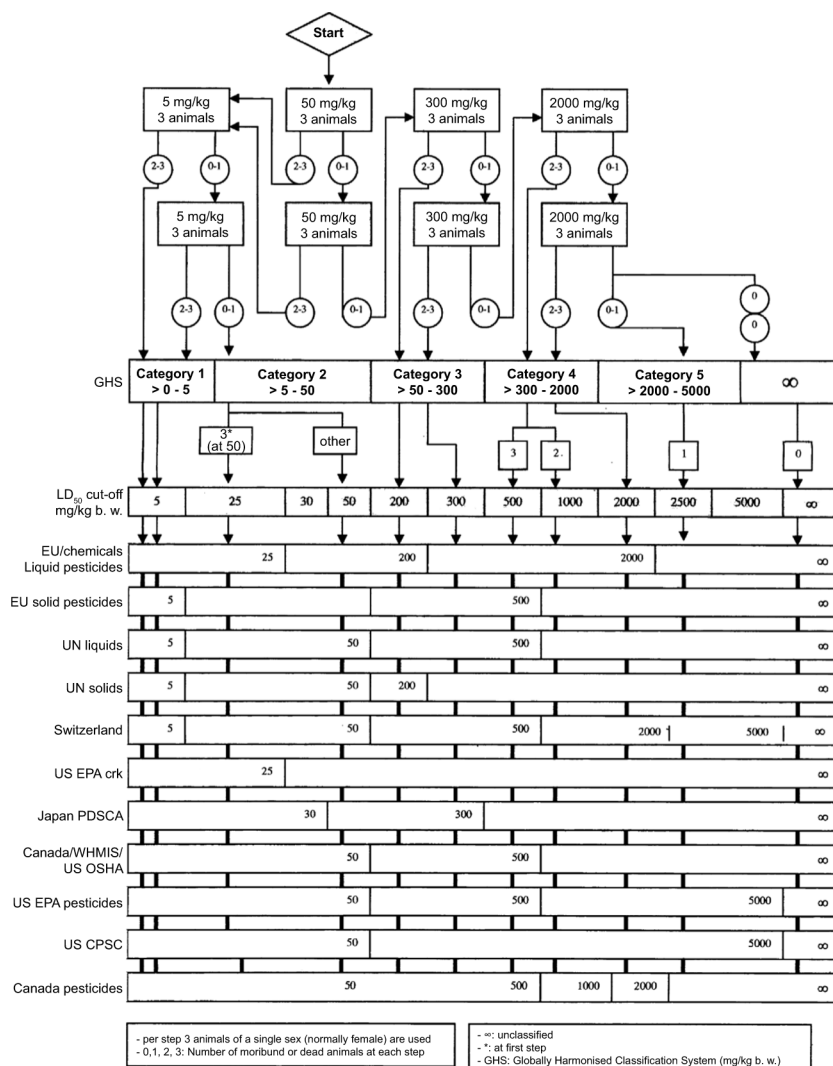
**▼B**

Appendix 3

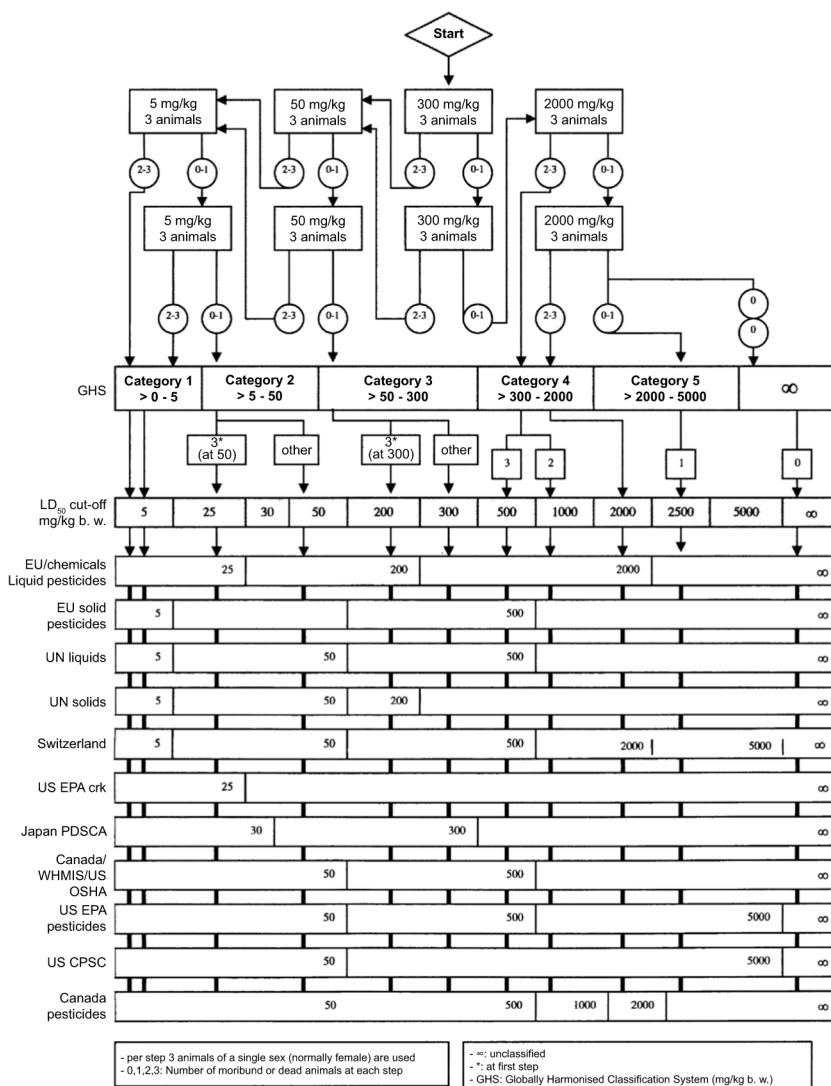
**TEST METHOD B.1 tris: Guidance on classification according to EU scheme to cover the transition period until full implementation of the Globally Harmonised Classification System (GHS) (taken from reference (8))**



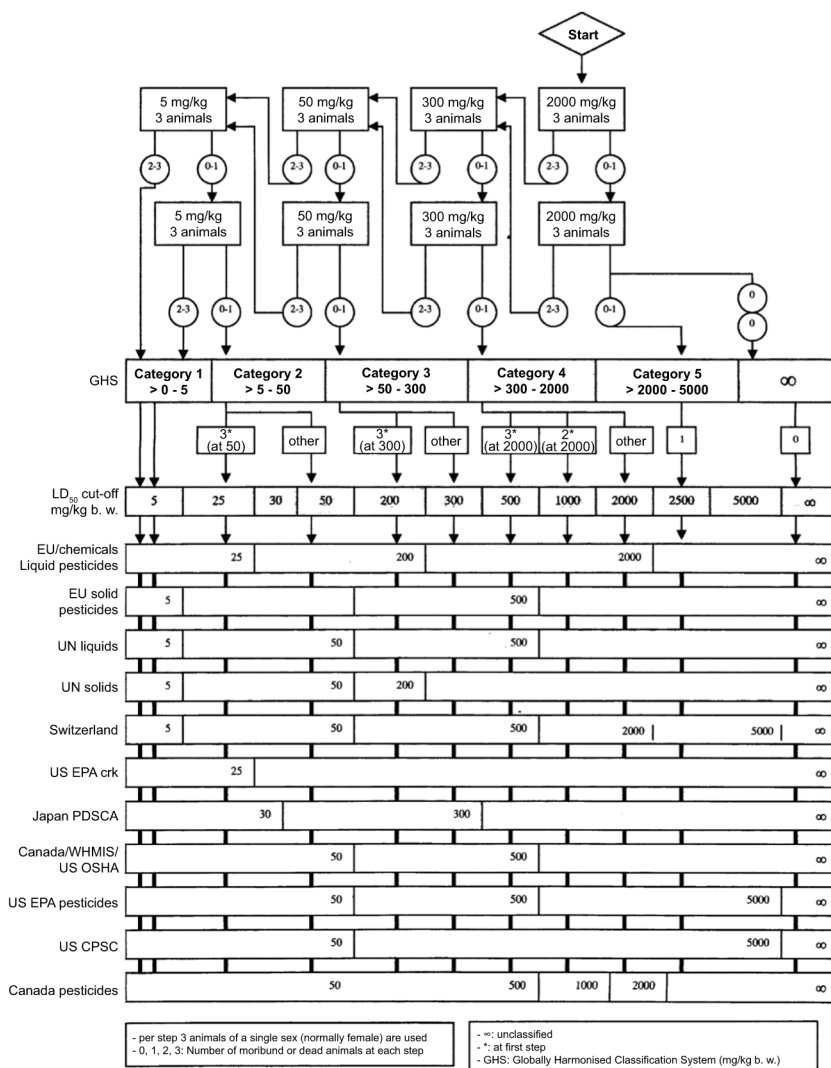
▼ B



**▼ B**



▼ B



**▼M4****B.2. ACUTE INHALATION TOXICITY**

## INTRODUCTION

1. This Test Method is equivalent to OECD Test Guideline 403 (2009) (1). The original acute inhalation Test Guideline 403 (TG 403) was adopted in 1981. This revised Test Method B.2 (as equivalent to the revised TG 403) has been designed to be more flexible, to reduce animal usage, and to fulfil regulatory needs. The revised Test Method features two study types: a Traditional LC<sub>50</sub> protocol and a Concentration × Time (C × t) protocol. Primary features of this Test Method are the ability to provide a concentration-response relationship ranging from non-lethal to lethal outcomes in order to derive a median lethal concentration (LC<sub>50</sub>), non-lethal threshold concentration (e.g. LC01), and slope, and to identify possible sex susceptibility. The C × t protocol should be used when there is a specific regulatory or scientific need that calls for the testing of animals over multiple time durations, such as for purposes of emergency response planning [e.g. deriving Acute Exposure Guideline Levels (AEGL), Emergency Response Planning Guidelines (ERPG), or Acute Exposure Threshold Levels (AETL) values], or for land-use planning.
2. Guidance on the conduct and interpretation of this Test Method studies can be found in the Guidance Document on Acute Inhalation Toxicity Testing (GD 39) (2).
3. Definitions used in the context of this Test Method are provided at the end of this chapter and in GD 39 (2).
4. This Test Method enables test chemical characterisation and quantitative risk assessment, and allows test chemicals to be ranked and classified according to Regulation (EC) No 1272/2008 (3). GD 39 (2) provides guidance in the selection of the appropriate Test Method for acute testing. When information on classification and labelling only is required, chapter B.52 of this Annex (4) is generally recommended [see GD 39 (2)]. This Test Method B.2 is not specifically intended for the testing of specialised materials, such as poorly soluble isometric or fibrous materials or manufactured nano-materials.

## INITIAL CONSIDERATIONS

5. Before considering testing in accordance with this Test Method all available information on the test chemical, including existing studies (e.g. chapter B.52 of this Annex (4)) whose data would support not doing additional testing should be considered by the testing laboratory in order to minimise animal usage. Information that may assist in the selection of the most appropriate species, strain, sex, mode of exposure and appropriate test concentrations include the identity, chemical structure, and physico-chemical properties of the test chemical; results of any *in vitro* or *in vivo* toxicity tests; anticipated uses and potential for human exposure; available (Q)SAR data and toxicological data on structurally related substances [see GD 39 (2)].

**▼M4**

6. Testing corrosive and/or irritating test chemicals at concentrations that are expected to cause severe pain and/or distress should be avoided to the extent possible. The corrosive/irritating potential should be evaluated by expert judgment using such evidence as human and animal experience (e.g. from repeat dose studies performed at non-corrosive/irritant concentrations), existing *in vitro* data (e.g. from chapters B.40, (5), B.40bis (6) of this Annex or OECD TG 435 (7)), pH values, information from similar substances or any other pertinent data, for the purpose of investigating whether further testing can be waived. For specific regulatory needs (e.g. for emergency planning purposes), this Test Method may be used for exposing animals to these materials because it provides the study director or principal investigator with control over the selection of target concentrations. However, the targeted concentrations should not induce severe irritation/corrosive effects, yet sufficient to extend the concentration-response curve to levels that reach the regulatory and scientific objective of the test. These concentrations should be selected on a case-by-case basis and justification for concentration selection should be provided [see GD 39 (2)].

**PRINCIPLE OF THE TEST**

7. This revised Test Method B.2 has been designed to obtain sufficient information on the acute toxicity of a test chemical to enable its classification and to provide lethality data (e.g. LC<sub>50</sub>, LC<sub>01</sub> and slope) for one or both sexes as needed for quantitative risk assessments. This Test Method offers two methods. The first method is a traditional protocol in which groups of animals are exposed to a limit concentration (limit test) or a series of concentrations in a stepwise procedure for a predetermined duration of usually 4 hours. Other durations of exposure may apply to serve specific regulatory purposes. The second method is a (C × t) protocol in which groups of animals are exposed to one (limit concentration) or a series of multiple concentrations over multiple durations.
8. Moribund animals or animals obviously in pain or showing signs of severe and enduring distress should be humanely killed and are considered in the interpretation of the test result in the same way as animals that died on test. Criteria for making the decision to kill moribund or severely suffering animals, and guidance on the recognition of predictable or impending death, are the subject of an OECD Guidance Document No 19 on Humane Endpoints (8).

**DESCRIPTION OF THE METHOD****Selection of animal species**

9. Healthy young adult animals of commonly used laboratory strains should be used. The preferred species is the rat and justification should be provided if other species are used.

**Preparation of animals**

10. Females should be nulliparous and non-pregnant. On the exposure day, animals should be young adults 8 to 12 weeks of age, and body weights should be within ± 20 % of the mean weight for each sex of any previously exposed animals of the same age. The animals are randomly selected and marked for individual identification. The animals are kept in their cages for at least 5 days prior to the start of the test to allow for acclimatisation to laboratory conditions. Animals should also be acclimatised to the test apparatus for a short period prior to testing, as this will lessen the stress caused by introduction to the new environment.

**▼ M4****Animal husbandry**

11. The temperature of the experimental animal maintenance room should be  $22 \pm 3$  °C. The relative humidity should ideally be maintained in the range of 30 to 70 %, though this may not be possible when using water as a vehicle. Before and after exposures, animals generally should be caged in groups by sex and concentration, but the number of animals per cage should not interfere with clear observation of each animal and should minimise losses due to cannibalism and fighting. When animals are to be exposed nose-only, it may be necessary for them to be acclimated to the restraining tubes. The restraining tubes should not impose undue physical, thermal, or immobilisation stress on the animals. Restraint may affect physiological endpoints such as body temperature (hyperthermia) and/or respiratory minute volume. If generic data are available to show that no such changes occur to any appreciable extent, then pre-adaptation to the restraining tubes is not necessary. Animals exposed whole-body to an aerosol should be housed individually during exposure to prevent them from filtering the test aerosol through the fur of their cage mates. Conventional and certified laboratory diets may be used, except during exposure, accompanied with an unlimited supply of municipal drinking water. Lighting should be artificial, the sequence being 12 hours light/12 hours dark.

**Inhalation chambers**

12. The nature of the test chemical and the objective of the test should be considered when selecting an inhalation chamber. The preferred mode of exposure is nose-only (which term includes head-only, nose-only or snout-only). Nose-only exposure is generally preferred for studies of liquid or solid aerosols and for vapours that may condense to form aerosols. Special objectives of the study may be better achieved by using a whole-body mode of exposure, but this should be justified in the study report. To ensure atmosphere stability when using a whole-body chamber, the total volume of the test animals should not exceed 5 % of the chamber volume. Principles of the nose-only and whole body exposure techniques and their particular advantages and disadvantages are described in GD 39 (2).

**EXPOSURE CONDITIONS****Administration of concentrations**

13. Nose-only exposures may be any duration up to 6 hours in rats. If mice are exposed nose-only, exposures generally should not exceed 4 hours. Justification should be provided if longer duration studies are needed [see GD 39 (2)]. Animals exposed to aerosols in whole-body chambers should be housed individually to prevent ingestion of test chemical due to grooming of cage mates. Feed should be withheld during the exposure period. Water may be provided throughout a whole-body exposure.
14. Animals are exposed to the test chemical as a gas, vapour, aerosol, or a mixture thereof. The physical state to be tested depends on the physico-chemical properties of the test chemical, the selected concentration, and/or the physical form most likely present during the handling and use of the test chemical. Hygroscopic and chemically reactive test chemicals should be tested under dry air conditions. Care should be taken to avoid generating explosive concentrations.

**▼ M4****Particle-size distribution**

15. Particle sizing should be performed for all aerosols and for vapours that may condense to form aerosols. To allow for exposure of all relevant regions of the respiratory tract, aerosols with mass median aerodynamic diameters (MMAD) ranging from 1 to 4 µm with a geometric standard deviation ( $\sigma_g$ ) in the range of 1,5 to 3,0 are recommended (2) (9) (10). Although a reasonable effort should be made to meet this standard, expert judgment should be provided if it cannot be achieved. For example, metal fumes may be smaller than this standard, and charged particles, fibres, and hygroscopic materials (which increase in size in the moist environment of the respiratory tract) may exceed this standard.

**Test chemical preparation in a vehicle**

16. A vehicle may be used to generate an appropriate concentration and particle size of the test chemical in the atmosphere. As a rule, water should be given preference. Particulate material may be subjected to mechanical processes to achieve the required particle size distribution, however, care should be taken to not decompose or alter the test chemical. In cases where mechanical processes are believed to have altered test chemical composition (e.g. extreme temperatures from excessive milling due to friction), the composition of the test chemical should be verified analytically. Adequate care should be taken to not contaminate the test chemical. It is not necessary to test non-friable granular materials which are purposefully formulated to be un-inhalable. An attrition test should be used to demonstrate that respirable particles are not produced when the granular material is handled. If an attrition test produces respirable substances, an inhalation toxicity test should be performed.

**Control animals**

17. A concurrent negative (air) control group is not necessary. When a vehicle other than water is used to assist in generating the test atmosphere, a vehicle control group should only be used when historical inhalation toxicity data are not available. If a toxicity study of a test chemical formulated in a vehicle reveals no toxicity, it follows that the vehicle is non-toxic at the concentration tested; thus, there is no need for a vehicle control.

**MONITORING OF EXPOSURE CONDITIONS****Chamber airflow**

18. The flow of air through the chamber should be carefully controlled, continuously monitored, and recorded at least hourly during each exposure. The monitoring of test atmosphere concentration (or stability) is an integral measurement of all dynamic parameters and provides an indirect means to control all relevant dynamic atmosphere generation parameters. Special consideration should be given to avoiding re-breathing in nose-only chambers in cases where airflow through the exposure system are inadequate to provide dynamic flow of test chemical atmosphere. There are prescribed methodologies that can be used to demonstrate that re-breathing does not occur under the selected operation conditions (2) (11). Oxygen concentration should be at least 19 % and carbon dioxide concentration should not exceed 1 %. If there is reason to believe that these standards cannot be met, oxygen and carbon dioxide concentrations should be measured.



**▼M4****Chamber temperature and relative humidity**

19. Chamber temperature should be maintained at  $22 \pm 3$  °C. Relative humidity in the animals' breathing zone, for both nose-only and whole-body exposures, should be monitored and recorded at least three times for durations of up to 4 hrs, and hourly for shorter durations. The relative humidity should ideally be maintained in the range of 30 to 70 %, but this may either be unattainable (e.g. when testing water based mixtures) or not measurable due to test chemical interference with the test method.

**Test chemical: Nominal concentration**

20. Whenever feasible, the nominal exposure chamber concentration should be calculated and recorded. The nominal concentration is the mass of generated test chemical divided by the total volume of air passed through the chamber system. The nominal concentration is not used to characterise the animals' exposure, but a comparison of the nominal concentration and the actual concentration gives an indication of the generation efficiency of the test system, and thus may be used to discover generation problems.

**Test chemical: Actual concentration**

21. The actual concentration is the test chemical concentration at the animals' breathing zone in an inhalation chamber. Actual concentrations can be obtained by specific methods (e.g. direct sampling, adsorptive or chemical reactive methods, and subsequent analytical characterisation) or by non-specific methods such as gravimetric filter analysis. The use of gravimetric analysis is acceptable only for single component powder aerosols or aerosols of low volatility liquids and should be supported by appropriate pre-study test chemical-specific characterisations. Multi-component powder aerosol concentration may also be determined by gravimetric analysis. However, this requires analytical data which demonstrate that the composition of airborne material is similar to the starting material. If this information is not available, a reanalysis of the test chemical (ideally in its airborne state) at regular intervals during the course of the study may be necessary. For aerosolised agents that may evaporate or sublime, it should be shown that all phases were collected by the method chosen. The target, nominal, and actual concentrations should be provided in the study report, but only actual concentrations are used in statistical analyses to calculate lethal concentration values.
22. One lot of the test chemical should be used, if possible, and the test sample should be stored under conditions that maintain its purity, homogeneity, and stability. Prior to the start of the study, there should be a characterisation of the test chemical, including its purity and, if technically feasible, the identity, and quantities of identified contaminants and impurities. This can be demonstrated by, but is not limited to, the following data: retention time and relative peak area, molecular weight from mass spectroscopy or gas chromatography analyses, or other estimates. Although the test sample's identity is not the responsibility of the test laboratory, it may be prudent for the test laboratory to confirm the sponsor's characterisation at least in a limited way (e.g. colour, physical nature, etc.).
23. The exposure atmosphere shall be held as constant as practicable and monitored continuously and/or intermittently depending on the method of analysis. When intermittent sampling is used, chamber atmosphere samples should be taken at least twice in a four hour study. If not feasible due to limited air flow rates or low concentrations, one sample may be collected over the entire exposure period. If marked sample-to-sample fluctuations occur, the next concentrations tested should use four samples per exposure.

**▼ M4**

Individual chamber concentration samples should not deviate from the mean concentration by more than  $\pm 10\%$  for gases and vapours or  $\pm 20\%$  for liquid or solid aerosols. Time to chamber equilibration ( $t_{95}$ ) should be calculated and recorded. The duration of an exposure spans the time that the test chemical is generated and this takes into account the times required to attain  $t_{95}$ . Guidance for estimating  $t_{95}$  can be found in GD 39 (2).

24. For very complex mixtures consisting of gases/vapours, and aerosols (e.g. combustion atmospheres and test chemicals propelled from purpose-driven end-use products/devices), each phase may behave differently in an inhalation chamber so at least one indicator substance (analyte), normally the principal active substance in the mixture, of each phase (gas/vapour and aerosol) should be selected. When the test chemical is a mixture, the analytical concentration should be reported for the mixture and not just for the active substance or the component (analyte). Additional information regarding actual concentrations can be found in GD 39 (2).

**Test chemical: Particle size distribution**

25. The particle size distribution of aerosols should be determined at least twice during each 4 hour exposure by using a cascade impactor or an alternative instrument such as an aerodynamic particle sizer. If equivalence of the results obtained by a cascade impactor or an alternative instrument can be shown, then the alternative instrument may be used throughout the study. A second device, such as a gravimetric filter or an impinger/gas bubbler, should be used in parallel to the primary instrument to confirm the collection efficiency of the primary instrument. The mass concentration obtained by particle size analysis should be within reasonable limits of the mass concentration obtained by filter analysis [see GD 39 (2)]. If equivalence can be demonstrated in the early phase of the study, then further confirmatory measurements may be omitted. For animal welfare reasons, measures should be taken to minimise inconclusive data which may lead to a need to repeat an exposure. Particle sizing should be performed for vapours if there is any possibility that vapour condensation may result in the formation of an aerosol, or if particles are detected in a vapour atmosphere with potential for mixed phases (see paragraph 15).

**PROCEDURE**

26. Two study types are described below: the Traditional protocol, and the  $C \times t$  protocol. Both protocols may include a sighting study, a main study, and/or a limit test (Traditional protocol) or testing at a limit concentration ( $C \times t$ ). If one sex is known to be more susceptible, the study director may choose to perform these studies using only the susceptible sex. If rodent species other than rats are exposed nose-only, maximum exposure durations may be adjusted to minimise species-specific distress. Before commencing, all available data should be considered in order to minimise animal usage. For example, data generated using chapter B.52 of this Annex (4) may eliminate the need for a sighting study, and may also demonstrate whether one sex is more susceptible [see GD 39 (2)].

**▼M4**

## TRADITIONAL PROTOCOL

**General considerations: Traditional protocol**

27. In a Traditional study, groups of animals are exposed to a test chemical for a fixed period of time (generally 4 hours) in either a nose-only or whole-body exposure chamber. Animals are exposed to either a limit concentration (limit test), or to at least three concentrations in a stepwise procedure (main study). A sighting study may precede a main study unless some information about the test chemical already exists, such as a previously performed B.52 study [see GD 39 (2)].

**Sighting study: Traditional protocol**

28. A sighting study is used to estimate test chemical potency, identify sex differences in susceptibility, and assist in selecting exposure concentration levels for the main study or limit test. When selecting concentration levels for the sighting study, all available information should be used including available (Q)SAR data and data for similar chemicals. No more than three males and three females should be exposed at each concentration (3 animals/sex may be needed to establish a sex difference). A sighting study may consist of a single concentration, but more concentrations may be tested if necessary. A sighting study should not test so many animals and concentrations that it resembles a main study. A previously performed B.52 study (4) may be used instead of a sighting study [see GD 39 (2)].

**Limit test: Traditional protocol**

29. A limit test is used when the test chemical is known or expected to be virtually non-toxic, i.e. eliciting a toxic response only above the regulatory limit concentration. In a limit test, a single group of three males and three females is exposed to the test chemical at a limit concentration. Information about the toxicity of the test chemical can be gained from knowledge about similar tested chemicals, taking into consideration the identity and percentage of components known to be of toxicological significance. In those situations where there is little or no information about its toxicity, or the test chemical is expected to be toxic, the main test should be performed.
30. The selection of limit concentrations usually depends on regulatory requirements. When Regulation (EC) No 1272/2008 is used, the limit concentrations for gases, vapours, and aerosols are 20 000 ppm, 20 mg/l and 5 mg/l, respectively (or the maximum attainable concentration) (3). It can be technically challenging to generate limit concentrations of some test chemicals, especially as vapours and aerosols. When testing aerosols, the primary goal should be to achieve a respirable particle size (MMAD of 1-4 µm). This is possible with most test chemicals at a concentration of 2 mg/l. Aerosol testing at greater than 2 mg/l should only be attempted if a respirable particle size can be achieved [see GD 39 (2)]. Regulation (EC) No 1272/2008 discourages testing in excess of a limit concentration for animal welfare reasons (3). The limit concentration should only be considered when there is a strong likelihood that results of such a test would have direct relevance for protecting human health (3), and justification provided in the study report. In the case of potentially explosive test chemicals, care should be taken to avoid conditions favourable for an explosion. To avoid an unnecessary use of animals, a test run without animals should be conducted prior to the limit test to ensure that the chamber conditions for a limit test can be achieved.

**▼M4**

31. If mortality or moribundity is observed at the limit concentration, the results of the limit test can serve as a sighting study for further testing at other concentrations (see main study). If a test chemical's physical or chemical properties make it impossible to attain a limit concentration, the maximum attainable concentration should be tested. If less than 50 % lethality occurs at the maximum attainable concentration, no further testing is necessary. If the limit concentration could not be attained, the study report should provide an explanation and supportive data. If the maximum attainable concentration of a vapour does not elicit toxicity, it may be necessary to generate the test chemical as a liquid aerosol.

**Main study: Traditional protocol**

32. A main study is typically performed using five males and five females (or 5 animals of the susceptible sex, if known) per concentration level, with at least three concentration levels. Sufficient concentration levels should be used to obtain a robust statistical analysis. The time interval between exposure groups is determined by the onset, duration, and severity of toxic signs. Exposure of animals at the next concentration level should be delayed until there is reasonable confidence of survival for previously tested animals. This allows the study director to adjust the target concentration for the next exposure group. Due to the dependence on sophisticated technologies, this may not always be practical in inhalation studies, so the exposure of animals at the next concentration level should be based on previous experience and scientific judgement. GD 39 (2) should be consulted when testing mixtures.

**CONCENTRATION × TIME (C × T) PROTOCOL****General considerations: C × t protocol**

33. A step-wise C × t study may be considered as an alternative to a Traditional protocol when assessing inhalation toxicity (12) (13) (14). This approach allows animals to be exposed to a test chemical at several concentration levels and for multiple time durations. All testing is performed in a nose-only chamber (whole-body chambers are not practical for this protocol). A flow diagram in Appendix 1 illustrates this protocol. A simulation analysis has shown that the Traditional protocol and the C × t protocol are both capable of yielding robust LC<sub>50</sub> values, but the C × t protocol is generally better at yielding robust LC<sub>01</sub> and LC<sub>10</sub> values (15).
34. A simulation analysis has demonstrated that using two animals per C × t interval (one per sex using both sexes, or two of the more susceptible sex) may generally be adequate when testing 4 concentrations and 5 exposure durations in a main study. Under some circumstances, the study director may elect to use two rats per sex per C × t interval (15). Using 2 animals per sex per concentration and time point may reduce bias and variability of the estimates, increase the estimation success rate, and improve confidence interval coverage. However, in case of an insufficient close fit to the data for estimation (when using one animal per sex or two animals of the more susceptible sex) a 5th exposure concentration may also suffice. Further guidance on the number of animals and concentrations to be used in a C × t study can be found in GD 39 (2).

**▼ M4****Sighting study: C × t protocol**

35. A sighting study is used to estimate test chemical potency and to assist in selecting exposure concentration levels for the main study. A sighting study using up to three animals/sex/concentration [for details see Appendix III of GD 39 (2)] may be needed to choose an appropriate starting concentration for the main study and to minimise the number of animals used. It may be necessary to use three animals per sex to establish a sex difference. These animals should be exposed for a single duration, generally 240 min. The feasibility of generating adequate test atmospheres should be assessed during technical pre-tests without animals. It is generally not necessary to perform a sighting study if mortality data are available from a B.52 study (4). When selecting the initial target concentration in a B.2 study, the study director should consider the mortality patterns observed in any available B.52 studies (4) for both sexes and for all concentrations tested [see GD 39 (2)].

**Initial Concentration: C × t protocol**

36. The initial concentration (Exposure Session I) (Appendix 1) will either be a limit concentration or a concentration selected by the study director based on the sighting study. Groups of 1 animal/sex are exposed to this concentration for multiple durations (e.g. 15, 30, 60, 120, or 240 minutes), resulting in a total number of 10 animals (called Exposure Session I) (Appendix 1).
  
37. The selection of limit concentrations usually depends on regulatory requirements. When Regulation (EC) No 1272/2008 is used, the limit concentrations for gases, vapours, and aerosols are 20 000 ppm, 20 mg/l and 5 mg/l, respectively (or the maximum attainable concentration) (3). It can be technically challenging to generate limit concentrations of some test chemicals, especially as vapours and aerosols. When testing aerosols, the goal should be to achieve a respirable particle size (i.e. an MMAD of 1-4 µm) at a limit concentration of 2 mg/l. This is possible with most test chemicals. Aerosol testing at greater than 2 mg/l should only be attempted if a respirable particle size can be achieved [see GD 39 (2)]. Regulation (EC) No 1272/2008 discourages testing in excess of a limit concentration for animal welfare reasons (3). Testing in excess of the limit concentration should only be considered when there is a strong likelihood that results of such a test would have direct relevance for protecting human health (3), justification should be provided in the study report. In the case of potentially explosive test chemicals, care should be taken to avoid conditions favourable for an explosion. To avoid an unnecessary use of animals, a test run without animals should be conducted prior to testing at the initial concentration to ensure that the chamber conditions for this concentration can be achieved.
  
38. If mortality or moribundity is observed at the initial concentration, the results at this concentration can serve as a starting point for further testing at other concentrations (see main study). When a test chemical's physical or chemical properties make it impossible to attain a limit concentration, the maximum attainable concentration should be tested. If less than 50 % lethality occurs at the maximum attainable concentration, no further testing is necessary. If the limit concentration could not be attained, the study report should provide an explanation and supportive data. If the maximum attainable concentration of a vapour does not elicit toxicity, it may be necessary to generate the test chemical as a liquid aerosol.

**▼ M4****Main study: C × t protocol**

39. The initial concentration (Exposure Session I) (Appendix 1) tested in the main study will either be a limit concentration or a concentration selected by the study director based on the sighting study. If mortality has been observed during or following Exposure Session I, the minimum exposure (C × t) which results in mortality will be taken as a guide to establish the concentration and periods of exposure for Exposure Session II. Each subsequent exposure session will depend on the previous session (see Appendix 1).
  
40. For many test chemicals the results obtained at the initial concentration, together with three additional exposure sessions with a smaller time grid (i.e. the geometric spacing of exposure periods as indicated by the factor between successive periods, generally  $\sqrt{2}$ ), will be sufficient to establish the C × t mortality relationship (15), but there may be some benefit to using a 5th exposure concentration [see Appendix 1 and GD 39 (2)]. For mathematical treatment of results for the C × t protocol, see Appendix 1.

**OBSERVATIONS**

41. The animals should be clinically observed frequently during the exposure period. Following exposure, clinical observations should be made at least twice on the day of exposure, or more frequently when indicated by the response of the animals to treatment, and at least once daily thereafter for a total of 14 days. The length of the observation period is not fixed, but should be determined by the nature and time of onset of clinical signs and length of the recovery period. The times at which signs of toxicity appear and disappear are important, especially if there is a tendency for signs of toxicity to be delayed. All observations are systematically recorded with individual records being maintained for each animal. Animals found in a moribund condition and animals showing severe pain and/or enduring signs of severe distress should be humanely killed for animal welfare reasons. Care should be taken when conducting examinations for clinical signs of toxicity that initial poor appearance and transient respiratory changes, resulting from the exposure procedure, are not mistaken for test chemical-related toxicity that would require premature killing of the animals. The principles and criteria summarised in the Guidance Document on Humane Endpoints (GD 19) should be taken into consideration (7). When animals are killed for humane reasons or found dead, the time of death should be recorded as precisely as possible.
  
42. Cage-side observations should include changes in the skin and fur, eyes and mucous membranes, and also respiratory, circulatory, autonomic and central nervous systems, and somatomotor activity and behaviour patterns. When possible, any differentiation between local and systemic effects should be noted. Attention should be directed to observations of tremors, convulsions, salivation, diarrhoea, lethargy, sleep and coma. The measurement of rectal temperature may provide supportive evidence of reflex bradypnea or hypo/hyperthermia related to treatment or confinement.

**Body weights**

43. Individual animal weights should be recorded once during the acclimatization period, on the day of exposure prior to exposure (day 0), and at least on days 1, 3 and 7 (and weekly thereafter), and at the time of death or euthanasia if exceeding day 1. Body weight is recognised as a critical indicator of toxicity so animals exhibiting a sustained decrement of  $\geq 20\%$ , compared to pre-study values, should be closely monitored. Surviving animals are weighed and humanely killed at the end of the post-exposure period.

**▼ M4****Pathology**

44. All test animals, including those which die during the test or are euthanised and removed from the study for animal welfare reasons, should be subjected to gross necropsy. If necropsy cannot be performed immediately after a dead animal is discovered, the animal should be refrigerated (not frozen) at temperatures low enough to minimise autolysis. Necropsies should be performed as soon as possible, normally within a day or two. All gross pathological changes should be recorded for each animal with particular attention to any changes in the respiratory tract.
45. Additional examinations included a priori by design may be considered to extend the interpretive value of the study, such as measuring lung weight of surviving rats, and/or providing evidence of irritation by microscopic examination of the respiratory tract. Examined organs may also include those showing evidence of gross pathology in animals surviving 24 or more hours, and organs known or expected to be affected. Microscopic examination of the entire respiratory tract may provide useful information for test chemicals that are reactive with water, such as acids and hygroscopic test chemicals.

**DATA AND REPORTING****Data**

46. Individual animal data on body weights and necropsy findings should be provided. Clinical observation data should be summarised in tabular form, showing for each test group the number of animals used, the number of animals displaying specific signs of toxicity, the number of animals found dead during the test or killed for humane reasons, time of death of individual animals, a description and time course of toxic effects and reversibility, and necropsy findings.

**Test report**

47. The test report should include the following information, as appropriate:

*Test animals and husbandry*

- Description of caging conditions, including: number (or change in number) of animals per cage, bedding material, ambient temperature and relative humidity, photoperiod, and identification of diet
- Species/strain used and justification for using a species other than the rat
- Number, age and sex of animals
- Method of randomisation
- Details of food and water quality (including diet type/source, water source)
- Description of any pre-test conditioning including diet, quarantine, and treatment for disease;

**▼ M4***Test chemical*

- Physical nature, purity and, where relevant, physico-chemical properties (including isomerisation)
- Identification data and Chemical Abstract Services (CAS) Registry Number, if known;

*Vehicle*

- Justification for use of vehicle and justification for choice of vehicle (if other than water)
- Historical or concurrent data demonstrating that the vehicle does not interfere with the outcome of the study;

*Inhalation chamber*

- Description of the inhalation chamber including dimensions and volume
  - Source and description of equipment used for the exposure of animals as well as generation of atmosphere
  - Equipment for measuring temperature, humidity, particle-size, and actual concentration
  - Source of air and treatment of air supplied/extracted and system used for conditioning
  - Methods used for calibration of equipment to ensure a homogeneous test atmosphere
  - Pressure difference (positive or negative)
  - Exposure ports per chamber (nose-only); location of animals in the system (whole-body)
  - Temporal homogeneity/stability of test atmosphere
  - Location of temperature and humidity sensors and sampling of test atmosphere in the chamber
  - Air flow rates, air flow rate/exposure port (nose-only), or animal load/chamber (whole-body)
  - Information about the equipment used to measure oxygen and carbon dioxide, if applicable
  - Time required to reach inhalation chamber equilibrium ( $t_{95}$ )
  - Number of volume changes per hour
  - Metering devices (if applicable);
- Exposure data*
- Rationale for target concentration selection in the main study
  - Nominal concentrations (total mass of test chemical generated into the inhalation chamber divided by the volume of air passed through the chamber)
  - Actual test chemical concentrations collected from the animals' breathing zone; for mixtures that produce heterogeneous physical forms (gases, vapours, aerosols), each may be analysed separately



**▼ M4**

- All air concentrations should be reported in units of mass (e.g. mg/l, mg/m<sup>3</sup>, etc.); units of volume (e.g. ppm, ppb, etc.) may also be reported parenthetically
- Particle size distribution, mass median aerodynamic diameter (MMAD), and geometric standard deviation ( $\sigma_g$ ), including their methods of calculation. Individual particle size analyses should be reported;

*Test conditions*

- Details of test chemical preparation, including details of any procedures used to reduce the particle size of solid materials or to prepare solutions of the test chemical. In cases where mechanical processes may have altered test chemical composition, include the results of analyses to verify the composition of the test chemical
- A description (preferably including a diagram) of the equipment used to generate the test atmosphere and to expose the animals to the test atmosphere
- Details of the chemical analytical method used and method validation (including efficiency of recovery of test chemical from the sampling medium)
- The rationale for the selection of test concentrations;

*Results*

- Tabulation of chamber temperature, humidity, and airflow
- Tabulation of chamber nominal and actual concentration data
- Tabulation of particle size data including analytical sample collection data, particle size distribution and calculations of the MMAD and  $\sigma_g$
- Tabulation of response data and concentration level for each animal (i.e. animals showing signs of toxicity including mortality, nature, severity, time of onset and duration of effects)
- Individual body weights of animals collected on study; date and time of death if prior to scheduled euthanasia, time course of onset of signs of toxicity and whether these were reversible for each animal
- Necropsy findings and histopathological findings for each animal, if available
- Lethality estimates (e.g. LC<sub>50</sub>, LD<sub>01</sub>) including 95 % confidence limits, and slope (if provided by the evaluation method)
- Statistical relation, including estimate for the exponent n ( $C \times t$  protocol). The name of the statistical software used should be provided;

**▼ M4***Discussion and interpretation of results*

- Particular emphasis should be made to the description of methods used to meet this Test Method's criteria, e.g. the limit concentration or the particle size
- The respirability of particles in light of the overall findings should be addressed, especially if the particle-size criteria could not be met
- An explanation should be provided if there was a need to humanely sacrifice animals in pain or showing signs of severe and enduring distress, based on the criteria in the OECD Guidance Document on Humane Endpoints (8)
- If testing with chapter B.52 of this Annex (4) was discontinued in favour of this Test Method B.2, justifications should be provided
- The consistency of methods used to determine nominal and actual concentrations, and the relation of actual concentration to nominal concentration should be included in the overall assessment of the study
- The likely cause of death and predominant mode of action (systemic versus local) should be addressed.

*LITERATURE:*

- (1) OECD (2009). Acute Inhalation Toxicity Testing. OECD Guideline for Testing of Chemicals No 403, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (2) OECD (2009). Guidance Document on Acute Inhalation Toxicity Testing. Environmental Health and Safety Monograph Series on Testing and Assessment No 39, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (3) Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC and amending Regulation (EC) No 1907/2006 (OJ L 353, 31.12.2008, p. 1).
- (4) Chapter B.52 of this Annex, Acute Inhalation Toxicity — Acute Toxic Class (ATC) Method.
- (5) Chapter B.40 of this Annex, In Vitro Skin Corrosion: Transcutaneous Electrical Resistance Test (TER).
- (6) Chapter B.40bis of this Annex, In Vitro Skin Corrosion: Human Skin Model Test.
- (7) OECD (2005), In Vitro Membrane Barrier Test Method For Skin Corrosion. OECD Guideline for Testing of Chemicals No 435, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (8) OECD (2000). Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation. Environmental Health and Safety Monograph Series on Testing and Assessment No 19, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (9) SOT (1992). Technical Committee of the Inhalation Specialty Section, Society of Toxicology (SOT). Recommendations for the Conduct of Acute Inhalation Limit Tests. Fund. Appl. Toxicol. 18: 321-327.
- (10) Phalen RF (2009). Inhalation Studies: Foundations and Techniques. (2nd Edition) Informa Healthcare, New York.

**▼ M4**

- (11) Pauluhn J and Thiel A (2007). A Simple Approach to Validation of Directed-Flow Nose-Only Inhalation Chambers. *J. Appl. Toxicol.* 27: 160-167.
- (12) Zwart JHE, Arts JM, ten Berge WF, Appelman LM (1992). Alternative Acute Inhalation Toxicity Testing by Determination of the Concentration-Time-Mortality Relationship: Experimental Comparison with Standard LC50 Testing. *Reg. Toxicol. Pharmacol.* 15: 278-290.
- (13) Zwart JHE, Arts JM, Klokman-Houweling ED, Schoen ED (1990). Determination of Concentration-Time-Mortality Relationships to Replace LC50 Values. *Inhal. Toxicol.* 2: 105-117.
- (14) Ten Berge WF and Zwart A (1989). More Efficient Use of Animals in Acute Inhalation Toxicity Testing. *J. Haz. Mat.* 21: 65-71.
- (15) OECD (2009). Performance Assessment: Comparison of 403 and  $C \times t$  Protocols via Simulation and for Selected Real Data Sets. Environmental Health and Safety Monograph Series on Testing and Assessment No 104, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (16) Finney DJ (1977). Probit Analysis, 3rd ed. Cambridge University Press, London/New York.

## DEFINITION

**Test chemical:** Any substance or mixture tested using this Test Method.

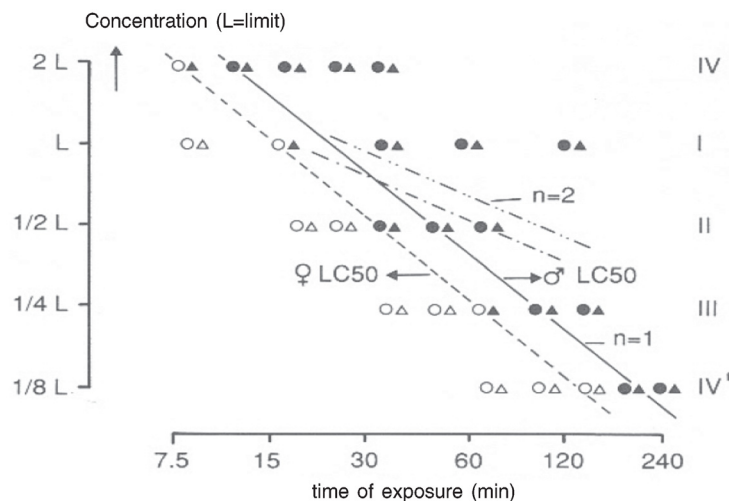
**▼ M4***Appendix 1***C × t Protocol**

1. A step-wise Concentration × Time (C × t) study may be considered as an alternative to the Traditional protocol for assessing inhalation toxicity (12) (13) (14). It should be performed preferentially when there is a specific regulatory or scientific need that calls for the testing of animals over multiple time durations such as for emergency response planning or land use planning. This approach usually begins with testing at a limit concentration (Exposure Session I) in which animals are exposed to a test chemical for five time durations (e.g. 15, 30, 60, 120 and 240 min) so that multiple durations of time will be obtained within one exposure session (see Figure 1). When Regulation (EC) No 1272/2008 is used, the limit concentrations for gases, vapours, and aerosols are 20 000 ppm, 20 mg/l, and 5 mg/l, respectively. These levels may only be exceeded if there is a regulatory or scientific need for testing at these levels (see paragraph 37 in the B.2 main text).
2. In situations where there is little or no information about the toxicity of a test chemical, a sighting study should be performed in which groups of no more than 3 animals per sex are exposed to target concentrations selected by the study director, generally for 240 min.
3. If a limit concentration is tested during Exposure Session I and less than 50 % mortality is observed, no additional testing is needed. If there is a regulatory or scientific need to establish the concentration/time/response relationship at higher levels than the indicated limit concentration, the next exposure should be carried out at a higher level such as at two times the limit concentration (i.e. 2L in Figure 1).
4. If toxicity is observed at the limit concentration, additional testing (main study) is necessary. These additional exposures are carried out either at lower concentrations (in Figure 1: Exposure Sessions II, III or IV') or at higher concentrations using shorter durations (in Figure 1: Exposure Session IV) using durations that are adapted and not as widely spaced.
5. The test (initial concentration and additional concentrations) is carried out using 1 animal/sex per concentration/time point or with 2 animals of the more susceptible sex per concentration/time point. Under some circumstances, the study director may elect to utilise 2 rats per sex per concentration/time point (or 4 animals of the susceptible sex per concentration/time point) (15). Using 2 animals per sex per concentration/time point generally reduces bias and variability of the estimates, increases the estimation success rate, and improves confidence interval coverage relative to the protocol as described here. Further details are provided in GD 39 (2).
6. Ideally, each exposure session is carried out on one day. This gives the opportunity to delay the next exposure until there is reasonable confidence of survival, and it allows the study director to adjust the target concentration and durations for the next exposure session. It is advised to start each exposure session with the group that will be exposed the longest, e.g. the 240-min exposure group, followed by the 120 minute exposure group, and so on. If, for example, animals in the 240 minute group are dying after 90 minutes or showing severe signs of toxicity (e.g. extreme changes in breathing pattern such as laboured breathing), it would not make sense to expose a group for 120 minutes because mortality would likely be 100 %. Thus the study director should select shorter exposure durations for that concentration (e.g. 90, 65, 45, 33 and 25 minutes).

▼ **M4**

7. The chamber concentration should be measured frequently to determine the time-weighted-average concentration for each exposure duration. Whenever possible, the time of death for each animal (rather than the exposure duration) should be used in the statistical analysis.
8. The results of the first four exposure sessions should be examined to identify a data gap in the concentration-time curve (see Figure 1). In case of an insufficient fit, an additional exposure (5th concentration) may be performed. Concentration and exposure durations for the 5th exposure should be chosen to cover this gap.
9. All exposure sessions (including the first Exposure Session) will be used to calculate the concentration-time-response relationship using Statistical Analysis (16). If possible, for each  $C \times t$  interval, the time-weighted average concentration and the duration of exposure until death (if death occurs during the exposure) should be used.

Figure 1

**Hypothetical illustration of a concentration-time-mortality relationship in rats**

Open symbols = survivors; closed symbols = dead animals

Triangles = females; circles = males

Solid line =  $LC_{50}$  values (range 7,5-240 min) for males with  $n = 1$

Dashed line =  $LC_{50}$  values (range 7,5-240 min) for females with  $n = 1$

Dotted lines = hypothetical  $LC_{50}$  values line for males and females if  $n$  had been equal to 2 (12).

Glossary

Concentration:

Time of exposure:

**▼ M4**

10. Below is an example of the stepwise procedure:

**Exposure Session I — Testing at the limit concentration (see Figure 1)**

- 1 animal/sex per concentration/time point; 10 animals in total <sup>(a)</sup>
- Target concentration <sup>(b)</sup> = limit concentration.
- Expose five groups of animals at this target concentration for durations of 15, 30, 60, 120 and 240 minutes, respectively.

↓

**Exposure Session II <sup>(c)</sup> — Main Study**

- 1 animal/sex per concentration/time point; 10 animals in total.
- Expose five groups of animals at a lower concentration <sup>(d)</sup> (1/2L) with slightly longer exposure durations (factor  $\sqrt{2}$  spaced; see Figure 1).

↓

**Exposure Session III — Main Study**

- 1 animal/sex per concentration/time point; 10 animals total.
- Expose five groups of animals at a lower concentration <sup>(d)</sup> (1/4L) with slightly longer exposure durations (factor  $\sqrt{2}$  spaced; see Figure 1).

↓

**Exposure Session IV' — Main Study**

- 1 animal/sex per concentration/time point; 10 animals total.
- Expose five groups of animals at a lower concentration <sup>(d)</sup> (1/8L) with slightly longer exposure durations (factor  $\sqrt{2}$  spaced; see Figure 1).

↓ or

<sup>(a)</sup> If no sex susceptibility information is available, rats of both sexes will be used, *i.e.* 1 animal/sex per concentration. Based on existing information, or if it becomes apparent during this exposure session that one sex is more susceptible, 10 animals of the susceptible sex will be used (2 animals per concentration/time point) at each concentration level during subsequent testing.

<sup>(b)</sup> When Regulation (EC) No 1272/2008 is used, the limit concentrations for gases, vapours, and aerosols are 20 000 ppm, 20 mg/l, and 5 mg/l, respectively. In case of expected toxicity or based on the results of the sighting study, lower starting concentrations should be chosen. In case of regulatory or scientific needs, higher concentrations may be used.

<sup>(c)</sup> Ideally, exposure of animals at the next concentration level should be delayed until there is reasonable confidence of survival for previously treated animals. This allows the study director to adjust the target concentration and durations for the next exposure session.

<sup>(d)</sup> The minimum dose (concentration x time) which resulted in mortality during testing at initial concentration (first exposure session) will be taken as a guide to establish the next combination of concentration and exposure durations. Typically, the concentration will be decreased two-fold (1/2L) and animals will be exposed over a new time range with a finer grid using a geometric division of exposure periods with a factor 1,4 ( $\sqrt{2}$ ; see reference 11) around the time according to the minimum lethal dose level (time x concentration) observed during the first exposure. In this figure (Figure 1), mortality in Exposure session I was first observed at 15 min; the durations during session II are therefore centred around 30 min, and are 15, 21 30, 42 and 60 min. After the first two exposures, it is strongly advised to plot the data in a similar figure as indicated above, and to check whether the relationship between concentration and time has an angle of 45 degrees ( $n = 1$ ) or if the concentration-time-response relationship is less steep (e.g.  $n = 2$ ) or steeper (e.g.  $n = 0,8$ ). In the latter cases it is strongly advised to adapt the next concentrations and durations accordingly.

**▼ M4****Exposure Session IV — Main Study**

- 1 animal/sex per concentration/time point; 10 animals total.
- Expose five groups of animals at a higher concentration<sup>(e)</sup> (2L) with slightly shorter exposure durations (factor  $\sqrt{2}$  spaced; see Figure 1).

**Mathematical treatment of results for the C × t protocol**

11. A C × t procedure with 4 or 5 exposure concentrations and five durations will yield 20 or 25 data points, respectively. With these data points, the C × t relationship can be calculated using statistical analysis (16):

*Equation 1:*

$$\text{Probit}(P) = b_0 + b_1 \ln C + b_2 \ln t$$

where C = concentration; t = exposure duration, or

*Equation 2:*

$$\text{Response} = f(C^n t)$$

where  $n = b_1/b_2$ .

Using equation 1, the LC<sub>50</sub> value can be calculated for a given time period (e.g. 4 hour, 1 hour, 30 minutes, or any time period within the range of time periods tested) using P = 5 (50 % response). Note that Haber's rule is only applicable when  $n = 1$ . The LC<sub>01</sub> can be calculated using P = 2,67.

---

<sup>(e)</sup> In certain cases it may be necessary to increase the concentration (2L) over a new time range with a still finer grid using a geometric division of exposure periods with a factor 1,4 ( $\sqrt{2}$ ) around the time according to the minimum lethal concentration level observed during the first exposure. The minimum exposure duration should preferably exceed 5 minutes; the maximum exposure duration should not exceed 8 hours.

**▼B****B.3. ACUTE TOXICITY (DERMAL)****1. METHOD****1.1. INTRODUCTION**

See General introduction Part B (A).

**1.2. DEFINITION**

See General introduction Part B (B).

**1.3. REFERENCE SUBSTANCES**

None.

**1.4. PRINCIPLE OF THE TEST METHOD**

The test substance is applied to the skin in graduated doses to several groups of experimental animals, one dose being used per group. Subsequently, observations of effects and deaths are made. Animals, which die during the test are necropsied and at the conclusion of the test surviving animals are necropsied.

Animals showing severe and enduring signs of distress and pain may need to be humanely killed. Dosing test substances in a way known to cause marked pain and distress due to corrosive or irritating properties need not be carried out.

**1.5. QUALITY CRITERIA**

None.

**1.6. DESCRIPTION OF THE TEST METHOD****1.6.1. Preparations**

The animals are kept in their experimental cages under the experimental housing and feeding conditions for at least five days prior to the experiment. Before the test, healthy young adult animals are randomised and assigned to the treatment groups. Approximately 24 hours before the test, fur should be removed by clipping or shaving from the dorsal area of the trunk of the animals. When clipping or shaving the fur, care must be taken to avoid abrading the skin which could alter its permeability. Not less than 10 % of the body surface should be clear for the application of the test substance. When testing solids, which may be pulverised if appropriate, the test substance should be moistened sufficiently with water or, where necessary, a suitable vehicle to ensure good contact with the skin. When a vehicle is used, the influence of the vehicle on penetration of skin by the test substance should be taken into account. Liquid test substances are generally used undiluted.

**1.6.2. Test conditions****1.6.2.1. Experimental animals**

The adult rat or rabbit may be used. Other species may be used but their use would require justification. Commonly used laboratory strains should be employed. For each sex, at the start of the test the range of weight variation in the animals used should not exceed  $\pm 20$  % of the appropriate mean value.



**▼ B**1.6.2.2. *Number and sex*

At least five animals are used at each dose level. They should all be of the same sex. If females are used, they should be nulliparous and non-pregnant. Where information is available demonstrating that a sex is markedly more sensitive, animals of this sex should be dosed.

*Note:* in acute toxicity tests with animals of a higher order than rodents, the use of smaller numbers should be considered. Doses should be carefully selected, and every effort should be made not to exceed moderately toxic doses. In such tests, administration of lethal doses of the test substance should be avoided.

1.6.2.3. *Dose levels*

These should be sufficient in number, at least three, and spaced appropriately to produce test groups with a range of toxic effects and mortality rates. Any irritant or corrosive effects should be taken into account when deciding on dose levels. The data should be sufficient to produce a dose/response curve and, where possible, permit an acceptable determination of the LD<sub>50</sub>.

1.6.2.4. *Limit test*

A limit test at one dose level of at least 2 000 mg/kg bodyweight may be carried out in a group of five male and five female animals, using the procedures described above. If compound-related mortality is produced, a full study may need to be considered.

1.6.2.5. *Observation period*

The observation period should be at least 14 days. However, the duration of observation should not be rigidly fixed. It should be determined by the toxic reactions, their rate of onset and the length of the recovery period; it may thus be extended when considered necessary. The time at which signs of toxicity appear and disappear, their duration and the time of death are important, especially if there is a tendency for deaths to be delayed.

1.6.3. **Procedure**

Animals should be caged individually. The test substance should be applied uniformly over an area, which is approximately 10 % of the total body surface area. With highly toxic substances the surface area covered may be less but as much of the area should be covered with a layer as thin and uniform as possible.

Test substances should be held in contact with the skin with a porous gauze dressing and non-irritating tape throughout a 24-hour exposure period. The test site should be further covered in a suitable manner to retain the gauze dressing and test substance and ensure that the animals cannot ingest the test substance. Restraint devices may be used to prevent the ingestion of the test substance but complete immobilisation is not a recommended method.

**▼B**

At the end of the exposure period, residual test substance should be removed, where practicable, using water or some other appropriate method of cleansing the skin.

Observations should be recorded systematically as they are made. Individual records should be maintained for each animal. Observations should be made frequently during the first day. A careful clinical examination should be made at least once each working day, other observations should be made daily with appropriate actions taken to minimise loss of animals to the study, e.g. necropsy or refrigeration of those animals found dead and isolation or sacrifice of weak or moribund animals.

Observations should include changes in fur, treated skin, eyes and mucous membranes, and also respiratory, circulatory, autonomic and central nervous systems, and somatomotor activity and behaviour pattern. Particular attention should be directed to observations of tremors, convulsions, salivation, diarrhoea, lethargy, sleep and coma. The time of death must be recorded as precisely as possible. Animals that die during the test and those surviving at the termination of the test are subjected to necropsy. All gross pathological changes should be recorded. Where indicated, tissues should be taken for histopathological examination.

*Assessment of toxicity in the other sex*

After completion of the study in one sex, at least one group of five animals of the other sex is dosed to establish that animals of this sex are not markedly more sensitive to the test substance. The use of fewer animals may be justified in individual circumstances. Where adequate information is available to demonstrate that animals of the sex tested are markedly more sensitive, testing in animals of the other sex may be dispensed with.

**2. DATA**

Data should be summarised in tabular form, showing for each test group the number of animals at the start of the test, time of death of individual animals, number of animals displaying other signs of toxicity, description of toxic effects and necropsy findings. Individual weights of animals should be determined and recorded shortly before the test substance is applied, weekly thereafter, and at death; changes in weight should be calculated and recorded when survival exceeds one day. Animals, which are humanely killed due to compound-related distress and pain are recorded as compound-related deaths. The LD<sub>50</sub> should be determined by a recognised method.

Data evaluation should include an evaluation of relationships, if any, between the animal's exposure to the test substance and the incidence and severity of all abnormalities, including behavioural and clinical abnormalities, gross lesions, body weight changes, mortality, and any other toxicological effects.

**3. REPORTING****3.1. TEST REPORT**

The test report shall, if possible, include the following information:

**▼ B**

- species, strain, source, environmental conditions, diet, etc.,
- test conditions (including method of skin cleansing and type of dressing: occlusive or not occlusive),
- dose levels (with vehicle, if used, and concentrations),
- sex of animals dosed,
- tabulation of response data by sex and dose level (i.e. number of animals that died or were killed during the test, number of animals showing signs of toxicity, number of animals exposed),
- time of death after dosing, reasons and criteria used for humane killing of animals,
- all observations,
- LD<sub>50</sub> value for the sex subjected to a full study, determined at 14 days with the method of determination specified,
- 95 % confidence interval for the LD<sub>50</sub> (where this can be provided),
- dose/mortality curve and slope where permitted by the method of determination,
- necropsy findings,
- any histopathological findings,
- results of any test on the other sex,
- discussion of results (particular attention should be given to the effect that humane killing of animals during the test may have on the calculated LD<sub>50</sub> value),
- interpretation of the results.

**3.2. EVALUATION AND INTERPRETATION**

See General introduction Part B (D).

**4. REFERENCES**

See General introduction Part B (E).

**▼B****B.4. ACUTE TOXICITY: DERMAL IRRITATION/CORROSION****1. METHOD**

This method is equivalent to the OECD TG 404 (2002).

**1.1. INTRODUCTION**

In the preparation of this updated method special attention was given to possible improvements in relation to animal welfare concerns and to the evaluation of all existing information on the test substance in order to avoid unnecessary testing in laboratory animals. This method includes the recommendation that prior to undertaking the described *in vivo* test for corrosion/irritation of the substance, a weight-of-the-evidence analysis be performed on the existing relevant data. Where insufficient data are available, they can be developed through application of sequential testing (1). The testing strategy includes the performance of validated and accepted *in vitro* tests and is provided as an Appendix to this method. In addition, where appropriate, the successive, instead of simultaneous, application of the three test patches to the animal in the initial *in vivo* test is recommended.

In the interest of both sound science and animal welfare, *in vivo* testing should not be undertaken until all available data relevant to the potential dermal corrosivity/irritation of the substance have been evaluated in a weight-of-the-evidence analysis. Such data will include evidence from existing studies in humans and/or laboratory animals, evidence of corrosivity/irritation of one or more structurally related substances or mixtures of such substances, data demonstrating strong acidity or alkalinity of the substance (2)(3), and results from validated and accepted *in vitro* or *ex vivo* tests (4)(5)-(5a). This analysis should decrease the need for *in vivo* testing for dermal corrosivity/irritation of substances for which sufficient evidence already exists from other studies as to those two endpoints.

A preferred sequential testing strategy, which includes the performance of validated and accepted *in vitro* or *ex vivo* tests for corrosion/irritation, is included as an Appendix to this Method. The strategy was developed at, and unanimously recommended by the participants of, an OECD workshop (6), and has been adopted as the recommended testing strategy in the Globally Harmonised System for the Classification of Chemical Substances (GHS) (7). It is recommended that this testing strategy be followed prior to undertaking *in vivo* testing. For new substances it is the recommended a stepwise testing approach for developing scientifically sound data on the corrosivity/irritation of the substance. For existing substances with insufficient data on dermal corrosion/irritation, the strategy should be used to fill missing data gaps. The use of a different testing strategy or procedure, or a decision not to use a stepwise testing approach, should be justified.

If a determination of corrosivity or irritation cannot be made using a weight-of-the-evidence analysis, consistent with the sequential testing strategy, an *in vivo* test should be considered (see Appendix).

**▼B**

## 1.2. DEFINITIONS

**Dermal irritation:** is the production of reversible damage of the skin following the application of a test substance for up to four hours.

**Dermal corrosion:** is the production of irreversible damage of the skin; namely, visible necrosis through the epidermis and into the dermis, following the application of a test substance for up to four hours. Corrosive reactions are typified by ulcers, bleeding, bloody scabs, and, by the end of observation at 14 days, by discoloration due to blanching of the skin, complete areas of alopecia, and scars. Histopathology should be considered to evaluate questionable lesions.

## 1.3. PRINCIPLE OF THE TEST METHOD

The substance to be tested is applied in a single dose to the skin of an experimental animal; untreated skin areas of the test animal serve as the control. The degree of irritation/corrosion is read and scored at specified intervals and is further described in order to provide a complete evaluation of the effects. The duration of the study should be sufficient to evaluate the reversibility or irreversibility of the effects observed.

Animals showing continuing signs of severe distress and/or pain at any stage of the test should be humanely killed, and the substance assessed accordingly. Criteria for making the decision to humanely kill moribund and severely suffering animals can be found in reference (8).

## 1.4. DESCRIPTION OF THE TEST METHOD

1.4.1. **Preparation for the *in vivo* test**1.4.1.1. *Selection of animal species*

The albino rabbit is the preferable laboratory animal and healthy young adult rabbits are used. A rationale for using other species should be provided.

1.4.1.2. *Preparation of the animals*

Approximately 24 hours before the test, fur should be removed by closely clipping the dorsal area of the trunk of the animals. Care should be taken to avoid abrading the skin, and only animals with healthy, intact skin should be used.

Some strains of rabbit have dense patches of hair that are more prominent at certain times of the year. Such areas of dense hair growth should not be used as test sites.

1.4.1.3. *Housing and feeding conditions*

Animals should be individually housed. The temperature of the experimental animal room should be 20 °C ( $\pm$  3 °C) for rabbits. Although the relative humidity should be at least 30 % and preferably not exceed 70 %, other than during room cleaning, the aim should be 50-60 %. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unrestricted supply of drinking water.

**▼B**1.4.2. **Test procedure**1.4.2.1. *Application of the test substance*

The test substance should be applied to a small area (approximately 6 cm<sup>2</sup>) of skin and covered with a gauze patch, which is held in place with non-irritating tape. In cases in which direct application is not possible (e.g. liquids or some pastes), the test substance should first be applied to the gauze patch, which is then applied to the skin. The patch should be loosely held in contact with the skin by means of a suitable semi-occlusive dressing for the duration of the exposure period. If the test substance is applied to the patch, it should be attached to the skin in such a manner that there is good contact and uniform distribution of the substance on the skin. Access by the animal to the patch and ingestion or inhalation of the test substance should be prevented.

Liquid test substances are generally used undiluted. When testing solids (which may be pulverised, if considered necessary), the test substance should be moistened with the smallest amount of water (or, where necessary, of another suitable vehicle) sufficient to ensure good skin contact. When vehicles other than water are used, the potential influence of the vehicle on irritation of the skin by the test substance should be minimal, if any.

At the end of the exposure period, which is normally four hours, residual test substance should be removed, where practicable, using water or an appropriate solvent without altering the existing response or the integrity of the epidermis.

1.4.2.2. *Dose level*

A dose of 0,5 ml. of liquid or 0,5 g of solid or paste is applied to the test site.

1.4.2.3. *Initial test (in vivo dermal irritation/corrosion test using one animal)*

It is strongly recommended that the *in vivo* test be performed initially using one animal, especially when the substance is suspected to have corrosion potential. This is in accordance with the sequential testing strategy (see Appendix 1).

When a substance has been judged to be corrosive on the basis of a weight-of-the-evidence analysis, no further animal testing is needed. For most substances suspected of being corrosive, further *in vivo* testing is normally not necessary. However, in those cases where additional data are felt warranted because of insufficient evidence, limited animal testing may be carried out using the following approach: up to three tests patches are applied sequentially to the animal. The first patch is removed after three minutes. If no serious skin reaction is observed, a second patch is applied and removed after one hour. If the observations at this stage indicate that exposure can humanely be allowed to extend to four hours, a third patch is applied and removed after four hours, and the response is graded.

If a corrosive effect is observed after any of the three sequential exposures, the test is immediately terminated. If a corrosive effect is not observed after the last patch is removed, the animal is observed for 14 days, unless corrosion develops at an earlier time point.

**▼B**

In those cases in which the test substance is not expected to produce corrosion but may be irritating, a single patch should be applied to one animal for four hours.

1.4.2.4. *Confirmatory test (in vivo dermal irritation test with additional animals)*

If a corrosive effect is not observed in the initial test, the irritant or negative response should be confirmed using up to two additional animals, each with one patch, for an exposure period of four hours. If an irritant effect is observed in the initial test, the confirmatory test may be conducted in a sequential manner, or by exposing two additional animals simultaneously. In the exceptional case, in which the initial test is not conducted, two or three animals may be treated with a single patch, which is removed after four hours. When two animals are used, if both exhibit the same response, no further testing is needed. Otherwise, the third animal is also tested. Equivocal responses may need to be evaluated using additional animals.

1.4.2.5. *Observation period*

The duration of the observation period should be sufficient to evaluate fully the reversibility of the effects observed. However, the experiment should be terminated at any time that the animal shows continuing signs of severe pain or distress. To determine the reversibility of effects, the animals should be observed up to 14 days after removal of the patches. If reversibility is seen before 14 days, the experiment should be terminated at that time.

1.4.2.6. *Clinical observations and grading of skin reactions*

All animals should be examined for signs of erythema and oedema, and the responses scored at 60 minutes, and then at 24, 48 and 72 hours after patch removal. For the initial test in one animal, the test site is also examined immediately after the patch has been removed. Dermal reactions are graded and recorded according to the grades in the Table below. If there is damage to skin which cannot be identified as irritation or corrosion at 72 hours, observations may be needed until day 14 to determine the reversibility of the effects. In addition to the observation of irritation, all local toxic effects, such as defatting of the skin, and any systemic adverse effects (e.g. effects on clinical signs of toxicity and body weight), should be fully described and recorded. Histopathological examination should be considered to clarify equivocal responses.

The grading of skin responses is necessarily subjective. To promote harmonisation in grading of skin response and to assist testing laboratories and those involved in making and interpreting the observations, the personnel performing the observations need to be adequately trained in the scoring system used (see Table below). An illustrated guide for grading skin irritation and other lesions could be helpful (9).

## 2. DATA

### 2.1. PRESENTATION OF RESULTS

Study results should be summarised in tabular form in the final test report and should cover all items listed in section 3.1.

**▼ B**

## 2.2. EVALUATION OF RESULTS

The dermal irritation scores should be evaluated in conjunction with the nature and severity of lesions, and their reversibility or lack of reversibility. The individual scores do not represent an absolute standard for the irritant properties of a material, as other effects of the test material are also evaluated. Instead, individual scores should be viewed as reference values, which need to be evaluated in combination with all other observations from the study.

Reversibility of dermal lesions should be considered in evaluating irritant responses. When responses such as alopecia (limited area), hyperkeratosis, hyperplasia and scaling, persist to the end of the 14-day observation period, the test substance should be considered an irritant.

3. **REPORTING**

## 3.1. TEST REPORT

The test report must include the following information:

Rationale for *in vivo* testing: weight-of-evidence analysis of pre-existing test data, including results from sequential testing strategy:

- description of relevant data available from prior testing,
- data derived at each stage of testing strategy,
- description of *in vitro* tests performed, including details of procedures, results obtained with test/reference substances,
- weight-of-the-evidence analysis for performing *in vivo* study.

Test substance:

- identification data (e.g. CAS number, source, purity, known impurities, lot number),
- physical nature and physicochemical properties (e.g. pH, volatility, solubility, stability),
- if mixture, composition and relative percentages of components.

Vehicle:

- identification, concentration (where appropriate), volume used,
- justification for choice of vehicle.

Test animals:

- species/strain used, rationale for using animals other than albino rabbit,
- number of animals of each sex,
- individual animal weights at start and conclusion of test,
- age at start of study,
- source of animals, housing conditions, diet, etc.



**▼B**

Test conditions:

- technique of patch site preparation,
- details of patch materials used and patching technique,
- details of test substance preparation, application, and removal.

Results:

- tabulation of irritation/corrosion response scores for each animal at all time points measured,
- descriptions of all lesions observed,
- narrative description of nature and degree of irritation or corrosion observed, and any histopathological findings,
- description of other adverse local (e.g. defatting of skin) and systemic effects in addition to dermal irritation or corrosion.
- Discussion of results

#### 4. REFERENCES

- (1) Barratt, M.D., Castell, J.V., Chamberlain, M., Combes, R.D., Dearden, J.C., Fentem, J.H., Gerner, I., Giuliani, A., Gray, T.J.B., Livingston, D.J., Provan, W.M., Rutten, F.A.J.J.L., Verhaar, H.J.M., Zbinden, P. (1995) The Integrated Use of Alternative Approaches for Predicting Toxic Hazard. ECVAM Workshop Report 8. ATLA 23, p. 410-429.
- (2) Young, J.R., How, M.J., Walker, A.P., Worth W.M.H. (1988) Classification as Corrosive or Irritant to Skin of Preparations Containing Acidic or Alkaline Substance Without Testing on Animals. Toxicology *In Vitro*, 2, p. 19-26.
- (3) Worth, A.P., Fentem, J.H., Balls, M., Botham, P.A., Curren, R.D., Earl, L.K., Esdaile, D.J., Liebsch, M. (1998) Evaluation of the proposed OECD Testing Strategy for skin corrosion. ATLA 26, p. 709-720.
- (4) ECETOC (1990) Monograph No 15, 'Skin Irritation', European Chemical Industry, Ecology and Toxicology Centre, Brussels.
- (5) Fentem, J.H., Archer, G.E.B., Balls, M., Botham, P.A., Curren, R.D., Earl, L.K., Esdaile, D.J., Holzhutter, H.G. and Liebsch, M. (1998) The ECVAM international validation study on in vitro tests for skin corrosivity. 2. Results and evaluation by the Management Team. Toxicology *In Vitro* 12, p. 483-524.
- (5a) Testing Method B.40 Skin Corrosion.
- (6) OECD (1996) OECD Test Guidelines Programme: Final Report of the OECD Workshop on Harmonisation of Validation and Acceptance Criteria for Alternative Toxicological Test Methods. Held in Solna, Sweden, 22-24 January 1996 (<http://www.oecd1.org/ehs/test/background.htm>).
- (7) OECD (1998) Harmonised Integrated Hazard Classification System for Human Health and Environmental Effects of Chemical Substances, as endorsed by the 28th Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, November 1998 (<http://www.oecd1.org/ehs/Class/-HCL6.htm>).

**▼B**

- (8) OECD (2000). Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation. OECD Environmental Health and Safety Publications. Series on Testing and Assessment No 19 (<http://www.oecd1.org/ehs/test/monos.htm>).
- (9) EPA (1990). Atlas of Dermal Lesions, (20T-2004). United States Environmental Protection Agency, Office of Pesticides and Toxic Substances, Washington, DC, August 1990.

[Available from OECD Secretariat upon request].

▼B

Table I

**GRADING OF SKIN REACTIONS****Erythema and Eschar formation**

No erythema .....	0
Very slight erythema (barely perceptible) .....	1
Well defined erythema .....	2
Moderate to severe erythema .....	3
Severe erythema (beef redness) to eschar formation preventing grading of erythema .....	4

Maximum possible: 4

**Oedema formation**

No oedema .....	0
Very slight oedema (barely perceptible) .....	1
Slight oedema (edges of area well defined by definite raising) .....	2
Moderate oedema (raised approximately 1 mm) .....	3
Severe oedema (raised more than 1 mm and extending beyond area of exposure) .....	4

Maximum possible: 4

Histopathological examination may be carried out to clarify equivocal responses.



## Appendix

### A Sequential Testing Strategy for Dermal Irritation and Corrosion

#### GENERAL CONSIDERATIONS

In the interest of sound science and animal welfare, it is important to avoid the unnecessary use of animals and to minimise any testing that is likely to produce severe responses in animals. All information on a substance relevant to its potential skin corrosivity/irritancy should be evaluated prior to considering *in vivo* testing. Sufficient evidence may already exist to classify a test substance as to its dermal corrosion or irritation potential without the need to conduct testing in laboratory animals. Therefore, utilising a weight-of-the-evidence analysis and a sequential testing strategy, will minimise the need for *in vivo* testing, especially if the substance is likely to produce severe reactions.

It is recommended that a weight-of-the-evidence analysis be used to evaluate existing information regarding the skin irritation and corrosion of substances to determine whether additional studies, other than *in vivo* dermal studies, should be performed to help characterise such potential. Where further studies are needed, it is recommended that the sequential testing strategy be utilised to develop the relevant experimental data. For substances which have no testing history, the sequential testing strategy should be utilised to develop the data set needed to evaluate its dermal corrosion/irritation potential. The testing strategy described in this Appendix was developed at an OECD workshop (1) and was later affirmed and expanded in the Harmonised Integrated Hazard Classification System for Human Health and Environmental Effects of Chemical Substances, as endorsed by the 28th Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, in November 1998 (2).

Although this sequential testing strategy is not an integral part of testing method B.4, it expresses the recommended approach for the determination of skin irritation/corrosion characteristics. This approach represents both best practice and an ethical benchmark for *in vivo* testing for skin irritation/corrosion. The testing method provides guidance for the conduct of the *in vivo* test and summarises the factors that should be addressed before initiating such a test. The strategy provides an approach for the evaluation of existing data on the skin irritation/corrosion properties of test substances and a tiered approach for the generation of relevant data on substances for which additional studies are needed, or for which no studies have been performed. It also recommends the performance of validated and accepted *in vitro* or *ex vivo* tests for skin corrosion/irritation under specific circumstances.

#### DESCRIPTION OF THE EVALUATION AND TESTING STRATEGY

Prior to undertaking tests as part of the sequential testing strategy (Figure), all available information should be evaluated to determine the need for *in vivo* skin testing. Although significant information might be gained from the evaluation of single parameters (e.g. extreme pH), the totality of existing information should be considered. All relevant data on the effects of the substance in question, or its analogues, should be evaluated in making a weight-of-the-evidence decision, and a rationale for the decision should be presented. Primary emphasis should be placed upon existing human and animal data on the substance, followed by the outcome of *in vitro* or *ex vivo* testing. *In vivo* studies of corrosive substances should be avoided whenever possible. The factors considered in the testing strategy include:

**▼ B**

*Evaluation of existing human and animal data (Step 1).* Existing human data, e.g. clinical or occupational studies and case reports, and/or animal test data, e.g. from single or repeated dermal exposure toxicity studies, should be considered first, because they provide information directly related to effects on the skin. Substances with known irritancy or corrosivity, and those with clear evidence of non-corrosivity or non-irritancy, need not be tested in *in vivo* studies.

*Analysis of structure activity relationships (SAR) (Step 2).* The results of testing of structurally related substances should be considered, if available. When sufficient human and/or animal data are available on structurally related substances or mixtures of such substances to indicate their skin corrosion/irritancy potential, it can be presumed that the test substance being evaluated will produce the same responses. In those cases, the test substance may not need to be tested. Negative data from studies of structurally related substances or mixtures of such substances do not constitute sufficient evidence of non-corrosivity/non-irritancy of a substance under the sequential testing strategy. Validated and accepted SAR approaches should be used to identify both dermal corrosion and irritation potential.

*Physicochemical properties and chemical reactivity (Step 3).* Substances exhibiting pH extremes such as  $\leq 2,0$  and  $\geq 11,5$  may have strong local effects. If extreme pH is the basis for identifying a substance as corrosive to skin, then its acid/alkali reserve (or buffering capacity) may also be taken into consideration (3)(4). If the buffering capacity suggests that a substance may not be corrosive to the skin, then further testing should be undertaken to confirm this, preferably by the use of a validated and accepted *in vitro* or *ex vivo* test (see steps 5 and 6).

*Dermal toxicity (Step 4).* If a chemical has proven to be very toxic by the dermal route, an *in vivo* dermal irritation/corrosion study may not be practicable because the amount of test substance normally applied could exceed the very toxic dose and, consequently result in the death or severe suffering of the animals. In addition, when dermal toxicity studies utilising albino rabbits have already been performed up to the limit dose level of 2 000 mg/kg body weight or higher, and no dermal irritation or corrosion has been seen, additional testing for skin irritation/corrosion may not be needed. A number of considerations should be borne in mind when evaluating acute dermal toxicity in previously performed studies. For example, reported information on dermal lesions may be incomplete. Testing and observations may have been made on a species other than the rabbit, and species may differ widely in sensitivity of their responses. Also the form of test substance applied to animals may not have been suitable for assessment of skin irritation/corrosion (e.g., dilution of substances for testing dermal toxicity (5)). However, in those cases in which well-designed and conducted dermal toxicity studies have been performed in rabbits, negative findings may be considered sufficient evidence that the substance is not corrosive or irritating.

*Results from in vitro or ex vivo tests (Steps 5 and 6).* Substances that have demonstrated corrosive or severe irritant properties in a validated and accepted *in vitro* or *ex vivo* test (6)(7) designed for the assessment of these specific effects, need not be tested in animals. It can be presumed that such substances will produce similar severe effects *in vivo*.

**▼B**

*In vivo test in rabbits (Steps 7 and 8).* Should a weight-of-the-evidence decision be made to conduct *in vivo* testing, it should begin with an initial test using one animal. If the results of this test indicate the substance to be corrosive to the skin, further testing should not be performed. If a corrosive effect is not observed in the initial test, the irritant or negative response should be confirmed using up to two additional animals for an exposure period of four hours. If an irritant effect is observed in the initial test, the confirmatory test may be conducted in a sequential manner, or by exposing the two additional animals simultaneously.

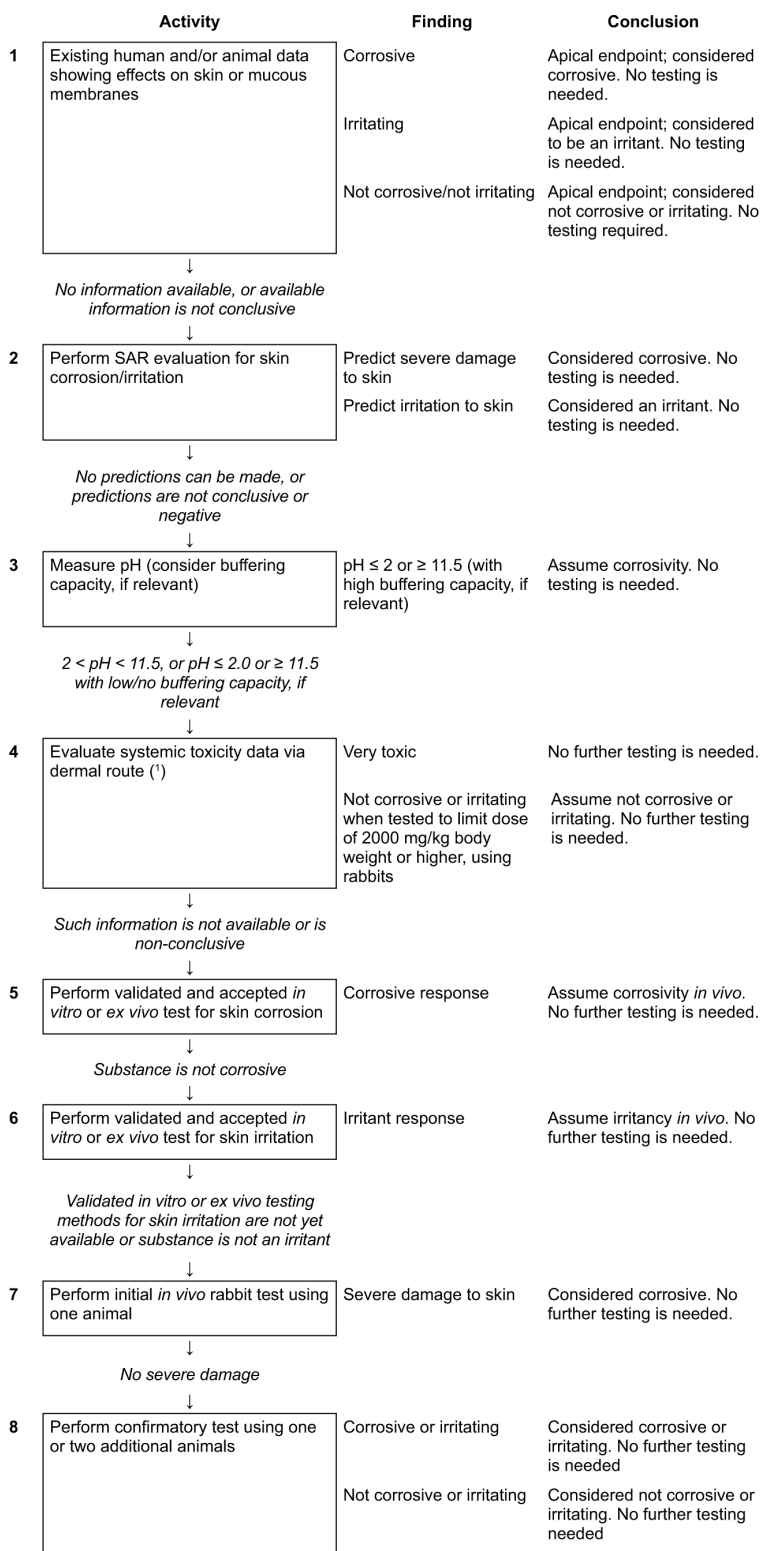
**REFERENCES**

- (1) OECD, (1996) Test Guidelines Programme: Final Report on the OECD Workshop on Harmonisation of Validation and Acceptance Criteria for Alternative Toxicological Test Methods. Held on Solna, Sweden, 22-24 January 1996 (<http://www1.oecd.org/ehs/test/background.htm>).
- (2) OECD, (1998) Harmonised Integrated Hazard Classification System for Human Health and Environmental Effects of Chemical Substances, as endorsed by the 28th Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, November 1998 (<http://www1.oecd.org/ehs/Class/HCL6.htm>).
- (3) Worth, A.P., Fentem J.H., Balls M., Botham P.A., Curren R.D., Earl L.K., Esdaile D.J., Liebsch M., (1998). An Evaluation of the Proposed OECD Testing Strategy for Skin Corrosion. *ATLA* 26, p. 709-720.
- (4) Young, J.R., How, M.J., Walker, A.P., Worth, W.M.H., (1988). Classification as Corrosive or Irritant to Skin of Preparations Containing Acidic or Alkaline Substances, Without Testing on Animals. *Toxicology In Vitro*, 2(1), p. 19-26.
- (5) Patil, S.M., Patrick, E., Maibach, H.I. (1996) Animal, Human, and In Vitro Test Methods for Predicting Skin Irritation, in: Francis N. Marzulli and Howard I. Maibach (editors): *Dermatotoxicology*. Fifth Edition ISBN 1-56032-356-6, Chapter 31, p. 411-436.
- (6) Testing Method B.40.
- (7) Fentem, J.H., Archer, G.E.B., Balls, M., Botham, P.A., Curren, R.D., Earl, L.K., Esdaile, D.J., Holzhutter, H.G. and Liebsch, M. (1998) The ECVAM international validation study on in vitro tests for skin corrosivity. 2. Results and evaluation by the Management Team. *Toxicology In Vitro* 12, p. 483-524.



Figure

## TESTING AND EVALUATION STRATEGY FOR DERMAL IRRITATION/CORROSION



<sup>(1)</sup> can be considered before Steps 2 and 3.

**▼B****B.5. ACUTE TOXICITY: EYE IRRITATION/CORROSION****1. METHOD**

This method is equivalent to the OECD TG 405 (2002)

**1.1. INTRODUCTION**

In the preparation of this updated method special attention was given to possible improvements through the evaluation of all existing information on the test substance in order to avoid unnecessary testing in laboratory animals and thereby address animal welfare concerns. This method includes the recommendation that prior to undertaking the described *in vivo* test for acute eye irritation/corrosion, a weight-of-the-evidence analysis be performed (1) on the existing relevant data. Where insufficient data are available, it is recommended that they be developed through application of sequential testing (2)(3). The testing strategy includes the performance of validated and accepted *in vitro* tests and is provided as an Appendix to the testing method. In addition, the use of an *in vivo* dermal irritation/corrosion test to predict eye corrosion prior to consideration of an *in vivo* eye test is recommended.

In the interest of both sound science and animal welfare, *in vivo* testing should not be considered until all available data relevant to the potential eye corrosivity/irritation of the substance has been evaluated in a weight-of-the-evidence analysis. Such data will include evidence from existing studies in humans and/or laboratory animals, evidence of corrosivity/irritation of one or more structurally related substances or mixtures of such substances, data demonstrating high acidity or alkalinity of the substance (4)(5), and results from validated and accepted *in vitro* or *ex vivo* tests for skin corrosion and irritation (6)(6a). The studies may have been conducted prior to, or as a result of, a weight-of-the-evidence analysis.

For certain substances, such an analysis may indicate the need for *in vivo* studies of the ocular corrosion/irritation potential of the substance. In all such cases, before considering the use of the *in vivo* eye test, preferably a study of the *in vivo* dermal effects of the substance should be conducted first and evaluated in accordance with testing method B.4 (7). The application of a weight-of-the-evidence analysis and the sequential testing strategy should decrease the need for *in vivo* testing for eye corrosivity/irritation of substances for which sufficient evidence already exists from other studies. If a determination of eye corrosion or irritation potential cannot be made using the sequential testing strategy, even after the performance of an *in vivo* study of dermal corrosion and irritation, an *in vivo* eye corrosion/irritation test may be performed.



**▼B**

A preferred sequential testing strategy, which includes the performance of validated *in vitro* or *ex vivo* tests for corrosion/irritation, is included in the Appendix to this testing method. The strategy was developed at, and unanimously recommended by the participants of, an OECD workshop (8), and has been adopted as the recommended testing strategy in the Globally Harmonised System for the Classification of Chemical Substances (GHS) (9). It is recommended that this testing strategy be followed prior to undertaking *in vivo* testing. For new substances it is the recommended stepwise testing approach for developing scientifically sound data on the corrosivity/irritation of the substance. For existing substances with insufficient data on skin and eye corrosion/irritation, the strategy should be used to fill missing data gaps. The use of a different testing strategy or procedure, or the decision not to use a stepwise testing approach, should be justified.

**1.2. DEFINITIONS**

**Eye irritation:** is the production of changes in the eye following the application of a test substance to the anterior surface of the eye, which are fully reversible within 21 days of application.

**Eye corrosion:** is the production of tissue damage in the eye, or serious physical decay of vision, following application of a test substance to the anterior surface of the eye, which is not fully reversible within 21 days of application.

**1.3. PRINCIPLE OF THE TEST METHOD**

The substance to be tested is applied in a single dose to one of the eyes of the experimental animal; the untreated eye serves as the control. The degree of eye irritation/corrosion is evaluated by scoring lesions of conjunctiva, cornea, and iris, at specific intervals. Other effects in the eye and adverse systemic effects are also described to provide a complete evaluation of the effects. The duration of the study should be sufficient to evaluate the reversibility or irreversibility of the effects.

Animals showing continuing signs of severe distress and/or pain at any stage of the test should be humanely killed, and the substance assessed accordingly. Criteria for making the decision to humanely kill moribund and severely suffering animals can be found in reference (10).

**1.4. DESCRIPTION OF THE TEST METHOD****1.4.1. Preparation for the *in vivo* test****1.4.1.1. Selection of species**

The albino rabbit is the preferable laboratory animal, and healthy young adult animals are used. A rationale for using other strains or species should be provided.

**1.4.1.2. Preparation of animals**

Both eyes of each experimental animal provisionally selected for testing should be examined within 24 hours before testing starts. Animals showing eye irritation, ocular defects, or pre-existing corneal injury should not be used.

**▼B**1.4.1.3. *Housing and feeding conditions*

Animals should be individually housed. The temperature of the experimental animal room should be 20 °C ( $\pm$  3 °C) for rabbits. Although the relative humidity should be at least 30 % and preferably not exceed 70 %, other than during room cleaning, the aim should be 50-60 %. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unrestricted supply of drinking water.

1.4.2. **Test procedure**1.4.2.1. *Application of the test substance*

The test substance should be placed in the conjunctival sac of one eye of each animal after gently pulling the lower lid away from the eyeball. The lids are then gently held together for about one second in order to prevent loss of the material. The other eye, which remains untreated, serves as a control.

1.4.2.2. *Irrigation*

The eyes of the test animals should not be washed for at least 24 hours following instillation of the test substance, except for solids (see Section 1.4.2.3.2), and in case of immediate corrosive or irritating effects. At 24 hours a washout may be used if considered appropriate.

Use of a satellite group of animals to investigate the influence of washing is not recommended unless it is scientifically justified. If a satellite group is needed, two rabbits should be used. Conditions of washing should be carefully documented, e.g. time of washing; composition and temperature of wash solution; duration, volume, and velocity of application.

1.4.2.3. *Dose level*1.4.2.3.1. *Testing of liquids*

For testing liquids, a dose of 0,1 ml is used. Pump sprays should not be used for instilling the substance directly into the eye. The liquid spray should be expelled and collected in a container prior to instilling 0,1 ml into the eye.

1.4.2.3.2. *Testing of solids*

When testing solids, pastes, and particulate substances, the amount used should have a volume of 0,1 ml or a weight of not more than 100 mg. The test material should be ground to a fine dust. The volume of solid material should be measured after gently compacting it, e.g. by tapping the measuring container. If the solid test substance has not been removed from the eye of the test animal by physiological mechanisms at the first observation time point of one hour after treatment, the eye may be rinsed with saline or distilled water.

**▼B**1.4.2.3.3. *Testing of aerosols*

It is recommended that all pump sprays and aerosols be collected prior to instillation into the eye. The one exception is for substances in pressurised aerosol containers, which cannot be collected due to vapourisation. In such cases, the eye should be held open, and the test substance administered to the eye in a simple burst of about one second, from a distance of 10 cm directly in front of the eye. This distance may vary depending on the pressure of the spray and its contents. Care should be taken not to damage the eye from the pressure of the spray. In appropriate cases, there may be a need to evaluate the potential for 'mechanical' damage to the eye from the force of the spray.

An estimate of the dose from an aerosol can be made by simulating the test as follows: the substance is sprayed on to weighing paper through an opening the size of a rabbit eye placed directly before the paper. The weight increase of the paper is used to approximate the amount sprayed into the eye. For volatile substances, the dose may be estimated by weighing a receiving container before and after removal of the test material.

1.4.2.4. *Initial test (in vivo eye irritation/corrosion test using one animal)*

As articulated in the sequential testing strategy (see Appendix 1), it is strongly recommended that the *in vivo* test be performed initially using one animal.

If the results of this test indicate the substance to be corrosive or a severe irritant to the eye using the procedure described, further testing for ocular irritancy should not be performed.

1.4.2.5. *Local anaesthetics*

Local anaesthetics may be used on a case-by-case basis. If the weight-of-the-evidence analysis indicates that the substance has the potential to cause pain, or initial testing shows that a painful reaction will occur, a local anaesthetic may be used prior to instillation of the test substance. The type, concentration, and dose of the local anaesthetic should be carefully selected to ensure that differences in reaction to the test substance will not result from its use. The control eye should be similarly anaesthetised.

1.4.2.6. *Confirmatory test (in vivo eye irritation test with additional animals)*

If a corrosive effect is not observed in the initial test, the irritant or negative response should be confirmed using up to two additional animals. If a severe irritant effect is observed in the initial test indicating a possible strong (irreversible) effect in the confirmatory testing, it is recommended that the confirmatory test be conducted in a sequential manner in one animal at a time, rather than exposing the two additional animals simultaneously. If the second animal reveals corrosive or severe irritant effects, the test is not continued. Additional animals may be needed to confirm weak or moderate irritant responses.

**▼B**1.4.2.7. *Observation period*

The duration of the observation period should be sufficient to evaluate fully the magnitude and reversibility of the effects observed. However, the experiment should be terminated at any time that the animal shows continuing signs of severe pain or distress (9). To determine reversibility of effects, the animals should be observed normally for 21 days post administration of the test substance. If reversibility is seen before 21 days, the experiment should be terminated at that time.

1.4.2.7.1. *Clinical observations and grading of eye reactions*

The eyes should be examined at one, 24, 48, and 72 hours after test substance application. Animals should be kept on test no longer than necessary once definitive information has been obtained. Animals showing continuing severe pain or distress should be humanely killed without delay, and the substance assessed accordingly. Animals with the following eye lesions post-instillation should be humanely killed: corneal perforation or significant corneal ulceration including staphyloma; blood in the anterior chamber of the eye; grade 4 corneal opacity which persists for 48 hours; absence of a light reflex (iridial response grade 2) which persists for 72 hours; ulceration of the conjunctival membrane; necrosis of the conjunctivae or nictitating membrane; or sloughing. This is because such lesions generally are not reversible

Animals that do not develop ocular lesions may be terminated not earlier than three days post instillation. Animals with mild to moderate lesions should be observed until the lesions clear, or for 21 days, at which time the study is terminated. Observations should be performed at seven, 14, and 21 days in order to determine the status of the lesions, and their reversibility or irreversibility.

The grades of ocular reaction (conjunctivae, cornea and iris) should be recorded at each examination (Table I). Any other lesions in the eye (e.g. pannus, staining) or adverse systemic effects should also be reported.

Examination of reactions can be facilitated by use of a binocular loupe, hand slit-lamp, biomicroscope, or other suitable device. After recording the observations at 24 hours, the eyes may be further examined with the aid of fluorescein.

The grading of ocular responses is necessarily subjective. To promote harmonisation of grading of ocular response and to assist testing laboratories and those involved in making and interpreting the observations, the personnel performing the observations need to be adequately trained in the scoring system used.

**2. DATA**

## 2.2. EVALUATION OF RESULTS

The ocular irritation scores should be evaluated in conjunction with the nature and severity of lesions, and their reversibility or lack of reversibility. The individual scores do not represent an absolute standard for the irritant properties of a material, as other effects of the test material are also evaluated. Instead, individual scores should be viewed as reference values and are only meaningful when supported by a full description and evaluation of all observations.

**▼ B**3. **REPORTING**

## 3.1. TEST REPORT

The test report must include the following information:

Rationale for *in vivo* testing: weight-of-the-evidence analysis of pre-existing test data, including results from sequential testing strategy

- description of relevant data available from prior testing,
- data derived in each step of testing strategy,
- description of *in vitro* tests performed, including details of procedures, results obtained with test/reference substances,
- description of *in vivo* dermal irritation/corrosion study performed, including results obtained,
- weight-of-the-evidence analysis for performing *in vivo* study.

Test substance:

- identification data (e.g. CAS number, source, purity, known impurities, lot number),
- physical nature and physicochemical properties (e.g. pH, volatility, solubility, stability, reactivity with water),
- in case of a mixture, composition and relative percentages of components,
- if local anaesthetic is used, identification, purity, type, dose, and potential interaction with test substance.

Vehicle:

- identification, concentration (where appropriate), volume used,
- justification for choice of vehicle.

Test animals:

- species/strain used, rationale for using animals other than albino rabbit,
- age of each animal at start of study,
- number of animals of each sex in test and control groups (if required),
- individual animal weights at start and conclusion of test,
- source, housing conditions, diet, etc.

Results:

- description of method used to score irritation at each observation time (e.g. hand slitlamp, biomicroscope, fluorescein),

**▼B**

- tabulation of irritant/corrosive response data for each animal at each observation time up to removal of each animal from the test,
- narrative description of the degree and nature of irritation or corrosion observed,
- description of any other lesions observed in the eye (e.g. vascularisation, pannus formation, adhesions, staining),
- description of non-ocular local and systemic adverse effects, and histopathological findings, if any.

Discussion of results.

### 3.2. INTERPRETATION OF THE RESULTS

Extrapolation of the results of eye irritation studies in laboratory animals to humans is valid only to a limited degree. In many cases the albino rabbit is more sensitive than humans to ocular irritants or corrosives.

Care should be taken in the interpretation of data to exclude irritation resulting from secondary infection.

## 4. REFERENCES

- (1) Barratt, M.D., Castell, J.V., Chamberlain, M., Combes, R.D., Dearden, J.C., Fentem, J.H., Gerner, I., Giuliani, A., Gray, T.J.B., Livingston, D.J., Provan, W.M., Rutten, F.A.J.J.L., Verhaar, H.J.M., Zbinden, P. (1995) The Integrated Use of Alternative Approaches for Predicting Toxic Hazard. ECVAM Workshop Report 8. ATLA 23, p. 410-429.
- (2) de Silva, O., Cottin, M., Dami, N., Roguet, R., Catroux, P., Toufic, A., Sicard, C., Dossou, K.G., Gerner, I., Schlede, E., Spielmann, H., Gupta, K.C., Hill, R.N., (1997) Evaluation of Eye Irritation Potential: Statistical Analysis and Tier Testing Strategies. Food Chem. Toxicol 35, p. 159-164.
- (3) Worth A.P. and Fentem J.H., (1999) A general approach for evaluating stepwise testing strategies ATLA 27, p. 161-177
- (4) Young, J.R., How, M.J., Walker, A.P., Worth W.M.H., (1988) Classification as Corrosive or Irritant to Skin of Preparations Containing Acidic or Alkaline Substance Without Testing on Animals. Toxicology *In Vitro*, 2, p. 19-26.
- (5) Neun, D.J. (1993) Effects of Alkalinity on the Eye Irritation Potential of Solutions Prepared at a Single pH. J. Toxicol. Cut. Ocular Toxicol. 12, p. 227-231.
- (6) Fentem, J.H., Archer, G.E.B., Balls, M., Botham, P.A., Curren, R.D., Earl, L.K., Edsaile, D.J., Holzhutter, H.G. and Liebsch, M. (1998) The ECVAM international validation study on in vitro tests for skin corrosivity. 2. Results and evaluation by the Management Team. Toxicology *In Vitro* 12, p. 483-524.
- (6a) Testing Method B.40 Skin Corrosion.
- (7) Testing method B.4. Acute toxicity: dermal irritation/corrosion.
- (8) OECD, (1996) OECD Test Guidelines Programme: Final Report of the OECD Workshop on Harmonisation of Validation and Acceptance Criteria for Alternative Toxicological Test Methods. Held in Solna, Sweden, 22-24 January 1996 (<http://www.oecd.org/ehs/test/background.htm>).

**▼B**

- (9) OECD, (1998) Harmonised Integrated Hazard Classification System for Human Health and Environmental Effects of Chemical Substances, as endorsed by the 28th Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, November 1998 (<http://www.oecd.org/ehs/Class/-HCL6.htm>).
- (10) OECD, (2000) Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation. OECD Environmental Health and Safety Publications. Series on Testing and Assessment No 19 (<http://www.oecd.org/ehs/test/monos.htm>).

**▼B***Table I***GRADING OF OCULAR LESIONS****Cornea**

Opacity: degree of density (readings should be taken from most dense area) (\*)

No ulceration or opacity .....	0
Scattered or diffuse areas of opacity (other than slight dulling of normal lustre); details of iris clearly visible .....	1
Easily discernible translucent area; details of iris slightly obscured ...	2
Nacrous area; no details of iris visible; size of pupil barely discernible	3
Opaque cornea; iris not discernible through the opacity .....	4

Maximum possible: 4

**NOTES**

(\*) The area of corneal opacity should be noted

**Iris**

Normal .....	0
Markedly deepened rugae, congestion, swelling, moderate circum-corneal hyperaemia; or injection; iris reactive to light (a sluggish reaction is considered to be an effect) .....	1
Hemorrhage, gross destruction, or no reaction to light .....	2

Maximum possible: 2

**Conjunctivae**

Redness (refers to palpebral and bulbar conjunctivae; excluding cornea and iris)

Normal .....	0
Some blood vessels hyperaemic (injected) .....	1
Diffuse, crimson colour; individual vessels not easily discernible ....	2
Diffuse beefy red .....	3

Maximum possible: 3

**Chemosis**

Swelling (refers to lids and/or nictating membranes)

Normal .....	0
Some swelling above normal .....	1
Obvious swelling, with partial eversion of lids .....	2
Swelling, with lids about half closed .....	3
Swelling, with lids more than half closed .....	4

Maximum possible: 4





## Appendix

### A Sequential Testing Strategy for Eye Irritation and Corrosion

#### GENERAL CONSIDERATIONS

In the interests of sound science and animal welfare, it is important to avoid the unnecessary use of animals, and to minimise testing that is likely to produce severe responses in animals. All information on a substance relevant to its potential ocular irritation/corrosivity should be evaluated prior to considering *in vivo* testing. Sufficient evidence may already exist to classify a test substance as to its eye irritation or corrosion potential without the need to conduct testing in laboratory animals. Therefore, utilising a weight-of-the-evidence analysis and sequential testing strategy will minimise the need for *in vivo* testing, especially if the substance is likely to produce severe reactions.

It is recommended that a weight-of-the-evidence analysis be used to evaluate existing information pertaining to eye irritation and corrosion of substances and to determine whether additional studies, other than *in vivo* eye studies, should be performed to help characterise such potential. Where further studies are needed, it is recommended that the sequential testing strategy be utilised to develop the relevant experimental data. For substances which have no testing history, the sequential testing strategy should be utilised to develop the data needed to evaluate its eye corrosion/irritation. The testing strategy described in this Appendix was developed at an OECD workshop (1). It was subsequently affirmed and expanded in the Harmonised Integrated Hazard Classification System for Human Health and Environmental Effects of Chemical Substances, as endorsed by the 28th Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, in November 1998 (2).

Although this testing strategy is not an integrated part of testing method B.5, it expresses the recommended approach for the determination of eye irritation/corrosion properties. This approach represents both best practice and an ethical benchmark for *in vivo* testing for eye irritation/corrosion. The testing method provides guidance for the conduct of the *in vivo* test and summarises the factors that should be addressed before considering such a test. The sequential testing strategy provides a weight-of-the-evidence approach for the evaluation of existing data on the eye irritation/corrosion properties of substances and a tiered approach for the generation of relevant data on substances for which additional studies are needed or for which no studies have been performed. The strategy includes the performance first of validated and accepted *in vitro* or *ex vivo* tests and then of testing method B.4 skin irritation/corrosion studies under specific circumstances (3)(4).

#### DESCRIPTION OF THE STEPWISE TESTING STRATEGY

Prior to undertaking tests as part of the sequential testing strategy (Figure), all available information should be evaluated to determine the need for *in vivo* eye testing. Although significant information might be gained from the evaluation of single parameters (e.g., extreme pH), the totality of existing information should be assessed. All relevant data on the effects of the substance in question, and its structural analogues, should be evaluated in making a weight-of-the-evidence decision, and a rationale for the decision should be presented. Primary emphasis should be placed upon existing human and animal data on the substance, followed by the outcome of *in vitro* or *ex vivo* testing. *In vivo* studies of corrosive substances should be avoided whenever possible. The factors considered in the testing strategy include:

**▼B**

*Evaluation of existing human and animal data (Step 1).* Existing human data, e.g. clinical and occupational studies, and case reports, and/or animal test data from ocular studies should be considered first, because they provide information directly related to effects on the eyes. Thereafter, available data from human and/or animal studies investigating dermal corrosion/irritation should be evaluated. Substances with known corrosivity or severe irritancy to the eye should not be instilled into the eyes of animals, nor should substances showing corrosive or irritant effects to the skin; such substances should be considered to be corrosive and/or irritating to the eyes as well. Substances with sufficient evidence of non-corrosivity and non-irritancy from previously performed ocular studies should also not be tested in *in vivo* eye studies.

*Analysis of structure activity relationships (SAR) (Step 2).* The results of testing of structurally related chemicals should be considered, if available. When sufficient human and/or animal data are available on structurally related substances or mixtures of such substances to indicate their eye corrosion/irritancy potential, it can be presumed that the test substance will produce the same responses. In those cases, the substance may not need to be tested. Negative data from studies of structurally related substances or mixtures of such substances do not constitute sufficient evidence of non-corrosivity/non-irritancy of a substance under the sequential testing strategy. Validated and accepted SAR approaches should be used to identify the corrosion and irritation potential for both dermal and ocular effects.

*Physicochemical properties and chemical reactivity (Step 3).* Substances exhibiting pH extremes such as  $\leq 2,0$  or  $\geq 11,5$  may have strong local effects. If extreme pH is the basis for identifying a substance as corrosive or irritant to the eye, then its acid/alkaline reserve (buffering capacity) may also be taken into consideration (5)(6). If the buffering capacity suggests that a substance may not be corrosive to the eye, then further testing should be undertaken to confirm this, preferably by the use of a validated and accepted *in vitro* or *ex vivo* test (see Section step 5 and 6).

*Consideration of other existing information (Step 4).* All available information on systemic toxicity via the dermal route should be evaluated at this stage. The acute dermal toxicity of the test substance should also be considered. If the test substance has been shown to be very toxic by the dermal route, it may not need to be tested in the eye. Although there is not necessarily a relationship between acute dermal toxicity and eye irritation/corrosion, it can be assumed that if an agent is very toxic via the dermal route, it will also exhibit high toxicity when instilled into the eye. Such data may also be considered between Steps 2 and 3.

*Results from *in vitro* or *ex vivo* tests (Steps 5 and 6).* Substances that have demonstrated corrosive or severe irritant properties in an *in vitro* or *ex vivo* test (7)(8) that has been validated and accepted for the assessment specifically of eye or skin corrosivity/irritation, need not be tested in animals. It can be presumed that such substances will produce similar severe effects *in vivo*. If validated and accepted *in vitro/ex vivo* tests are not available, one should bypass Steps 5 and 6 and proceed directly to Step 7.

*Assessment of *in vivo* dermal irritancy or corrosivity of the substance (Step 7).* When insufficient evidence exists with which to perform a conclusive weight-of-the-evidence analysis of the potential eye irritation/corrosivity of a substance based upon data from the studies listed above, the *in vivo* skin irritation/corrosion potential should be evaluated first, using testing method B.4 (4) and its accompanying Appendix (9). If the substance is shown to produce corrosion or severe skin irritation, it should be considered to be a corrosive eye irritant unless other information supports an alternative conclusion. Thus, an *in vivo* eye test would not need to be performed. If the substance is not corrosive or severely irritating to the skin, an *in vivo* eye test should be performed.

**▼B**

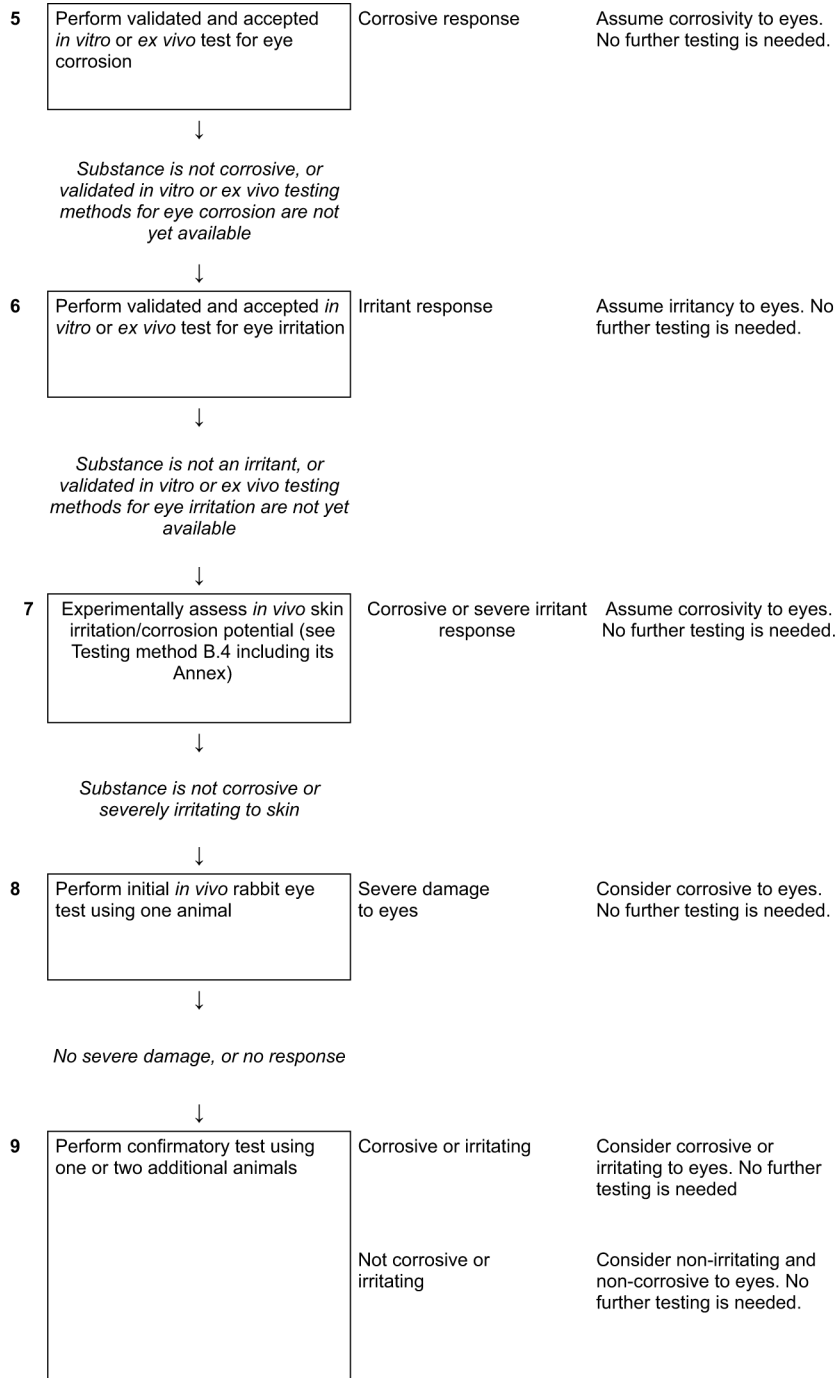
*In vivo* test in rabbits (Steps 8 and 9): *in vivo* ocular testing should begin with an initial test using one animal. If the results of this test indicate the substance to be a severe irritant or corrosive to the eyes, further testing should not be performed. If that test does not reveal any corrosive or severe irritant effects, a confirmatory test is conducted with two additional animals.

**REFERENCES**

- (1) OECD, (1996) OECD Test Guidelines Programme: Final Report of the OECD Workshop on Harmonisation of Validation and Acceptance Criteria for Alternative Toxicological Test Methods. Held in Solna, Sweden, 22-24 January 1996 (<http://www.oecd.org/ehs/test/background.htm>).
- (2) OECD, (1998) Harmonised Integrated Hazard Classification System for Human Health and Environmental Effects of Chemical Substances, as endorsed by the 28th Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, November 1998 (<http://www.oecd.org/ehs/Class/HCL6.htm>).
- (3) Worth, A.P. and Fentem J.H., (1999). A General Approach for Evaluating Stepwise Testing Strategies. *ATLA* 27, p. 161-177.
- (4) Testing method B.4. Acute Toxicity: dermal irritation/corrosion.
- (5) Young, J.R., How, M.J., Walker, A.P., Worth W.M.H., (1988) Classification as Corrosive or Irritant to Skin of Preparations Containing Acidic or Alkaline Substance Without Testing on Animals. *Toxicology In Vitro*, 27, p. 19-26.
- (6) Neun, D.J., (1993) Effects of Alkalinity on the Eye Irritation Potential of Solutions Prepared at a Single pH. *J. Toxicol. Cut. Ocular Toxicol.* 12, p. 227-231.
- (7) Fentem, J.H., Archer, G.E.B., Balls, M., Botham, P.A., Curren, R.D., Earl, L.K., Edsail, D.J., Holzhutter, H.G. and Liebsch, M., (1998) The ECVAM international validation study on *in vitro* tests for skin corrosivity. 2. Results and evaluation by the Management Team. *Toxicology In Vitro* 12, p. 483-524.
- (8) Testing Method B.40 Skin Corrosion.
- (9) Appendix to Testing method B.4: A Sequential Testing Strategy for Skin Irritation and Corrosion.



**▼ B**



**▼ B****B.6. SKIN SENSITISATION****1. METHOD****1.1. INTRODUCTION***Remarks:*

The sensitivity and ability of tests to detect potential human skin sensitisers are considered important in a classification system for toxicity relevant to public health.

There is no single test method which will adequately identify all substances with a potential for sensitising human skin and which is relevant for all substances.

Factors such as the physical characteristics of a substance, including its ability to penetrate the skin, must be considered in the selection of a test.

Two types of tests using guinea pigs have been developed: the adjuvant-type tests, in which an allergic state is potentiated by dissolving or suspending the test substance in Freund's Complete Adjuvant (FCA), and the non-adjuvant tests.

Adjuvant-type tests are likely to be more accurate in predicting a probable skin sensitising effect of a substance in humans than those methods not employing Freund's Complete Adjuvant and are thus the preferred methods.

The Guinea-Pig Maximisation Test (GPMT) is a widely used adjuvant-type test. Although several other methods can be used to detect the potential of a substance to provoke skin sensitisation reaction, the GPMT is considered to be the preferred adjuvant technique.

With many chemical classes, non-adjuvant tests (the preferred one being the Buehler test) are considered to be less sensitive.

In certain cases there may be good reasons for choosing the Buehler test involving topical application rather than the intradermal injection used in the Guinea-Pig Maximisation Test. Scientific justification should be given when the Buehler test is used.

The Guinea-Pig Maximisation Test (GPMT) and the Buehler test are described in this method. Other methods may be used provided that they are well-validated and scientific justification is given.

If a positive result is seen in a recognised screening test, a test substance may be designated as a potential sensitiser, and it may not be necessary to conduct a further guinea pig test. However, if a negative result is seen in such a test, the guinea pig test must be conducted using the procedure described in this test method.

See also General introduction Part B.

**▼B**

## 1.2. DEFINITIONS

**Skin sensitisation:** (allergic contact dermatitis) is an immunologically mediated cutaneous reaction to a substance. In the human, the responses may be characterised by pruritis, erythema, oedema, papules, vesicles, bullae or a combination of these. In other species the reactions may differ and only erythema and oedema may be seen.

**Induction exposure:** an experimental exposure of a subject to a test substance with the intention of inducing a hypersensitive state.

**Induction period:** a period of at least one week following an induction exposure during which a hypersensitive state may be developed.

**Challenge exposure:** an experimental exposure of a previously treated subject to a test substance following an induction period, to determine if the subject reacts in a hypersensitive manner.

## 1.3. REFERENCE SUBSTANCES

The sensitivity and reliability of the experimental technique used should be assessed every six months by use of substances, which are known to have mild-to-moderate skin sensitisation properties.

In a properly conducted test, a response of at least 30 % in an adjuvant test and at least 15 % in a non-adjuvant test should be expected for mild/moderate sensitisers.

The following substances are preferred.

CAS numbers	EINECS numbers	EINECS names	Common names
101-86-0	202-983-3	$\alpha$ -hexylcinnamaldehyde	$\alpha$ -hexylcinnamaldehyde
149-30-4	205-736-8	Benzothiazole-2-thiol (mercaptobenzothiazole)	kaptax
94-09-7	202-303-5	Benzocaine	norcaine

There may be circumstances where, given adequate justification other control substances meeting the above criteria may be used.

## 1.4. PRINCIPLE OF THE TEST METHOD

The test animals are initially exposed to the test substance by intradermal injections and/or epidermal application (induction exposure). Following a rest period of 10 to 14 days (induction period), during which an immune response may develop, the animals are exposed to a challenge dose. The extent and degree of skin reaction to the challenge exposure in the test animals is compared with that demonstrated by control animals which undergo sham treatment during induction and receive the challenge exposure.

## 1.5. DESCRIPTION OF THE TEST METHODS

If removal of the test substance is considered necessary, this should be achieved using water or an appropriate solvent without altering the existing response or the integrity of the epidermis.

**▼B**1.5.1. *Guinea-Pig Maximisation Test (GPMT)*

## 1.5.1.1. Preparations

Healthy young adult albino guinea pigs are acclimatised to the laboratory conditions for at least five days prior to the test. Before the test, animals are randomised and assigned to the treatment groups. Removal of hair is by clipping, shaving or possibly by chemical depilation, depending on the test method used. Care should be taken to avoid abrading the skin. The animals are weighed before the test commences and at the end of the test.

## 1.5.1.2. Test conditions

## 1.5.1.2.1. Test animals

Commonly used laboratory strains of albino guinea-pigs are used.

## 1.5.1.2.2. Number and sex

Male and/or female animals can be used. If females are used, they should be nulliparous and non-pregnant.

A minimum of 10 animals is used in the treatment group and at least five animals in the control group. When fewer than 20 test and 10 control guinea pigs have been used, and it is not possible to conclude that the test substance is a sensitiser, testing in additional animals to give a total of at least 20 test and 10 control animals is strongly recommended.

## 1.5.1.2.3. Dose levels

The concentration of the test substance used for each induction exposure should be well-tolerated systemically and should be the highest to cause mild-to-moderate skin irritation. The concentration used for the challenge exposure should be the highest non-irritant dose. The appropriate concentrations should be determined from a pilot study using two or three animals, if other information is not available. Consideration should be given to the use of FCA-treated animals for this purpose.

## 1.5.1.3. Procedure

## 1.5.1.3.1. Induction

## Day 0-treated group

Three pairs of intradermal injections of 0,1 ml volume are given in the shoulder region which is cleared of hair so that one of each pair lies on each side of the midline.

Injection 1: a 1:1 mixture (v/v) FCA/water or physiological saline.

Injection 2: the test substance in an appropriate vehicle at the selected concentration.

Injection 3: the test substance at the selected concentration formulated in a 1:1 mixture (v/v) FCA/water or physiological saline.

In injection 3, water soluble substances are dissolved in the aqueous phase prior to mixing with FCA. Liposoluble or insoluble substances are suspended in FCA prior to combining with the aqueous phase. The final concentration of test substance shall be equal to that used in injection 2.



**▼B**

Injections 1 and 2 are given close to each other and nearest the head, while 3 is given towards the caudal part of the test area.

Day 0-control group

Three pairs of intradermal injections of 0,1 ml volume are given in the same sites as in the treated animals.

Injection 1: a 1:1 mixture (v/v) FCA/water or physiological saline.

Injection 2: the undiluted vehicle.

Injection 3: a 50 % w/v formulation of the vehicle in a 1:1 mixture (v/v) FCA/water or physiological saline.

Day 5-7-treated and control groups

Approximately 24 hours before the topical induction application, if the substance is not a skin irritant, the test area, after close-clipping and/or shaving is treated with 0,5 ml of 10 % sodium lauryl sulphate in vaseline, in order to create a local irritation.

Day 6-8-treated group

The test area is again cleared of hair. A filter paper (2 × 4 cm) is fully-loaded with test substance in a suitable vehicle and applied to the test area and held in contact by an occlusive dressing for 48 hours. The choice of the vehicle should be justified. Solids are finely pulverised and incorporated in a suitable vehicle. Liquids can be applied undiluted, if appropriate.

Day 6-8-control group

The test area is again cleared of hair. The vehicle only is applied in a similar manner to the test area and held in contact by an occlusive dressing for 48 hours.

#### 1.5.1.3.2. Challenge

Day 20-22-treated and control groups

The flanks of treated and control animals are cleared of hair. A patch or chamber loaded with the test substance is applied to one flank of the animals and, when relevant, a patch or chamber loaded with the vehicle only may also be applied to the other flank. The patches are held in contact by an occlusive dressing for 24 hours.

#### 1.5.1.3.3. Observation and Grading: treated and control groups

- approximately 21 hours after removing the patch the challenge area is cleaned and closely-clipped and/or shaved and depilated if necessary;
- approximately three hours later (approximately 48 hours from the start of the challenge application) the skin reaction is observed and recorded according to the grades shown in the Appendix;
- approximately 24 hours after this observation a second observation (72 hours) is made and once again recorded.

Blind reading of test and control animals is encouraged.

If it is necessary to clarify the results obtained in the first challenge, a second challenge (i.e. a rechallenge), where appropriate with a new control group, should be considered approximately one week after the first one. A rechallenge may also be performed on the original control group.

**▼ B**

All skin reactions and any unusual findings, including systemic reactions, resulting from induction and challenge procedures should be observed and recorded according to the grading scale of Magnusson/Kligman (See Appendix). Other procedures, e.g. histopathological examination, the measurement of skin fold thickness, may be carried out to clarify doubtful reactions.

1.5.2. *Buehler test*

## 1.5.2.1. Preparations

Healthy young adult albino guinea-pigs are acclimatised to the laboratory conditions for at least five days prior to the test. Before the test, animals are randomised and assigned to the treatment groups. Removal of hair is by clipping, shaving or possibly by chemical depilation, depending on the test method used. Care should be taken to avoid abrading the skin. The animals are weighed before the test commences and at the end of the test.

## 1.5.2.2. Test conditions

## 1.5.2.2.1. Test animals

Commonly used laboratory strains of albino guinea-pigs are used.

## 1.5.2.2.2. Number and sex

Male and/or female animals can be used. If females are used, they should be nulliparous and non-pregnant.

A minimum of 20 animals is used in the treatment group and at least 10 animals in the control group.

## 1.5.2.2.3. Dose levels

The concentration of test substance used for each induction exposure should be the highest possible to produce a mild but not excessive irritation. The concentration used for the challenge exposure should be the highest non-irritating dose. If necessary, the appropriate concentration can be determined from a pilot study using two or three animals.

For water soluble test materials, it is appropriate to use water or a dilute non-irritating solution of surfactant as the vehicle. For other test materials 80 % ethanol/water is preferred for induction and acetone for challenge.

## 1.5.2.3. Procedure

## 1.5.2.3.1. Induction

## Day 0-treated group

One flank is cleared of hair (closely-clipped). The test patch system should be fully loaded with test substance in a suitable vehicle (the choice of the vehicle should be justified; liquid test substances can be applied undiluted, if appropriate).

The test patch system is applied to the test area and held in contact with the skin by an occlusive patch or chamber and a suitable dressing for six hours.

**▼B**

The test patch system must be occlusive. A cotton pad is appropriate and can be circular or square, but should approximate 4-6 cm<sup>2</sup>. Restraint using an appropriate restrainer is preferred to assure occlusion. If wrapping is used, additional exposures may be required.

Day 0-control group

One flank is cleared of hair (closely-clipped). The vehicle only is applied in a similar manner to that used for the treated group. The test patch system is held in contact with the skin by an occlusive patch or chamber and a suitable dressing for six hours. If it can be demonstrated that a sham control group is not necessary, a naive control group may be used.

Days 6-8 and 13-15-treated and control group

The same application as on day 0 is carried out on the same test area (cleared of hair if necessary) of the same flank on day 6-8, and again on day 13-15.

1.5.2.3.2. Challenge

Day 27-29-treated and control group

The untreated flank of treated and control animals is cleared of hair (closely-clipped). An occlusive patch or chamber containing the appropriate amount of test substance is applied, at the maximum non-irritant concentration, to the posterior untreated flank of treated and control animals.

When relevant, an occlusive patch or chamber with vehicle only is also applied to the anterior untreated flank of both treated and control animals. The patches or chambers are held in contact by a suitable dressing for six hours.

1.5.2.3.3. Observation and grading

- approximately 21 hours after removing the patch the challenge area is cleared of hair,
- approximately three hours later (approximately 30 hours after application of the challenge patch) the skin reactions are observed and recorded according to the grades shown in the Appendix,
- approximately 24 hours after the 30 hour observation (approximately 54 hours after application of the challenge patch) skin reactions are again observed and recorded.

Blind reading of the test and control animals is encouraged.

If it is necessary to clarify the results obtained in the first challenge, a second challenge (i.e. a rechallenge), where appropriate with a new control group, should be considered approximately one week after the first one. A rechallenge may also be performed on the original control group.

**▼B**

All skin reactions and any unusual findings, including systemic reactions, resulting from induction and challenge procedures should be observed and recorded according to the Magnusson/Kligman grading scale (See Appendix). Other procedures, e.g. histopathological examination, the measurement of skin fold thickness, may be carried out to clarify doubtful reactions.

**2. DATA (GPMT and Buehler test)**

Data should be summarised in tabular form, showing for each animal the skin reactions at each observation.

**3. REPORTING (GPMT and Buehler test)**

If a screening assay is performed before the guinea pig test the description or reference of the test (e.g. Mouse Ear Swelling Test (MEST)), including details of the procedure, must be given together with results obtained with the test and reference substances.

**Test report (GMPT and Buehler test)**

The test report shall, if possible, include the following information:

Test animals:

- strain of guinea-pig used,
- number, age and sex of animals,
- source, housing conditions, diet, etc.,
- individual weights of animals at the start of the test.

Test conditions:

- technique of patch site preparation,
- details of patch materials used and patching technique,
- result of pilot study with conclusion on induction and challenge concentrations to be used in the test,
- details of test substance preparation, application and removal,
- justification for choice of vehicle,
- vehicle and test substance concentrations used for induction and challenge exposures and the total amount of substance applied for induction and challenge.

Results:

- a summary of the results of the latest sensitivity and reliability check (see 1.3) including information on substance, concentration and vehicle used,

**▼B**

- on each animal including grading system,
- narrative description of the nature and degree effects observed,
- any histopathological findings.

Discussion of results.

Conclusions.

4. **REFERENCES**

This method is analogous to OECD TG 406.

**▼B**

*Appendix*

*TABLE*

**Magnusson/Kligman grading scale for the evaluation of challenge patch test reactions**

0 = no visible change

1 = discrete or patchy erythema

2 = moderate and confluent erythema

3 = intense erythema and swelling

**▼M4****B.7. REPEATED DOSE 28-DAY ORAL TOXICITY STUDY IN RODENTS**

## INTRODUCTION

1. This Test Method is equivalent to OECD Test Guideline 407 (2008). The original Test Guideline 407 was adopted in 1981. In 1995 a revised version was adopted, to obtain additional information from the animal used in the study, in particular on neurotoxicity and immunotoxicity.
2. In 1998, the OECD initiated a high-priority activity, to revise existing Test Guidelines and to develop new Test Guidelines for the screening and testing of potential endocrine disruptors (8). One element of the activity was to update the existing OECD guideline for 'repeated dose 28-day oral toxicity study in rodents' (TG 407) by parameters suitable to detect endocrine activity of test chemicals. This procedure underwent an extensive international program to test for the relevance and practicability of the additional parameters, the performance of these parameters for chemicals with (anti)oestrogenic, (anti)androgenic, and (anti)thyroid activity, the intra- and inter-laboratory reproducibility, and the interference of the new parameters with those required by the prior TG 407. The large amount of data thereby obtained has been compiled and evaluated in detail in a comprehensive OECD report (9). This updated Test Method B.7 (as equivalent to TG 407) is the outcome of the experience and results gained during the international test program. This Test Method allows certain endocrine mediated effects to be put into context with other toxicological effects.

## INITIAL CONSIDERATIONS AND LIMITATIONS

3. In the assessment and evaluation of the toxic characteristics of a chemical, the determination of oral toxicity using repeated doses may be carried out after initial information on toxicity has been obtained by acute toxicity testing. This Test Method is intended to investigate effects on a very broad variety of potential targets of toxicity. It provides information on the possible health hazards likely to arise from repeated exposure over a relatively limited period of time, including effects on the nervous, immune and endocrine systems. Regarding these particular endpoints, it should identify chemicals with neurotoxic potential, which may warrant further in-depth investigation of this aspect, and chemicals that interfere with thyroid physiology. It may also provide data on chemicals that affect the male and/or female reproductive organs in young adult animals and may give an indication of immunological effects.
4. The results from this Test Method B.7 should be used for hazard identification and risk assessment. The results obtained by the endocrine related parameters should be seen in the context of the 'OECD Conceptual Framework for Testing and Assessment of Endocrine Disrupting Chemicals' (11). The method comprises the basic repeated dose toxicity study that may be used for chemicals on which a 90-day study is not warranted (e.g. when the production volume does not exceed certain limits) or as a preliminary to a long-term study. The duration of exposure should be 28 days.

**▼ M4**

5. The international program conducted on the validation of parameters suitable to potentially detect endocrine activity of a test chemical showed that the quality of data obtained by this Test Method B.7 will depend much on the experience of the test laboratory. This relates specifically to the histopathological determination of cyclic changes in the female reproductive organs and to the weight determination of the small hormone dependent organs which are difficult to dissect. Guidance on histopathology has been developed (19). It is available on the OECD public website on Test Guidelines. It is intended to assist pathologists in their examinations and help increase the sensitivity of the assay. A variety of parameters were found to be indicative of endocrine-related toxicity and have been incorporated in the Test Method. Parameters for which insufficient data were available to prove usefulness or which showed only weak evidence in the validation programme of their ability to help in detection of endocrine disrupters are proposed as optional endpoints (see Appendix 2).
6. On the basis of data generated in the validation process, it must be emphasised that the sensitivity of this assay is not sufficient to identify all substances with (anti)androgenic or (anti)oestrogenic modes of action (9). The Test Method is not performed in a life-stage that is most sensitive to endocrine disruption. The Test Method nevertheless, during the validation process identified substances weakly and strongly affecting thyroid function, and strong and moderate endocrine active substances acting through oestrogen or androgen receptors, but in most cases failed to identify endocrine active substances that weakly affect oestrogen or androgen receptors. Thus it cannot be described as a screening assay for endocrine activity.
7. Consequently, the lack of effects related to these modes of action can not be taken as evidence for the lack of effects on the endocrine system. Regarding endocrine mediated effects, substance characterisation should not therefore be based on the results of this Test Method alone but should be used in a weight of evidence approach incorporating all existing data on a chemical to characterise potential endocrine activity. For this reason, regulatory decision making on endocrine activity (substance characterisation) should be a broadly based approach, not solely reliant on results from application of this test method.
8. It is acknowledged that all animal-based procedures will conform to local standards of animal care; the descriptions of care and treatment set forth below are minimal performance standards, and will be superseded by local regulations where more stringent. Further guidance of the humane treatment of animals is given by the OECD (14).
9. Definitions used are given in Appendix 1.

**PRINCIPLE OF THE TEST**

10. The test chemical is orally administered daily in graduated doses to several groups of experimental animals, one dose level per group for a period of 28 days. During the period of administration the animals are observed closely, each day for signs of toxicity. Animals that die or are euthanised during the test are necropsied and at the conclusion of the test surviving animals are euthanised and necropsied. A 28 day study provides information on the



**▼ M4**

effects of repeated oral exposure and can indicate the need for further longer term studies. It can also provide information on the selection of concentrations for longer term studies. The data derived from using the Test Method should allow for the characterisation of the test chemical toxicity, for an indication of the dose response relationship and the determination of the No-Observed Adverse Effect Level (NOAEL).

**DESCRIPTION OF THE METHOD****Selection of animal species**

11. The preferred rodent species is the rat, although other rodent species may be used. If the parameters specified within this Test Method B.7 are investigated in another rodent species a detailed justification should be given. Although it is biologically plausible that other species should respond to toxicants in a similar manner to the rat, the use of smaller species may result in increased variability due to technical challenges of dissecting smaller organs. In the international validation program for the detection of endocrine disruptors, the rat was the only species used. Young healthy adult animals of commonly used laboratory strains should be employed. Females should be nulliparous and non pregnant. Dosing should begin as soon as feasible after weaning, and, in any case, before the animals are nine weeks old. At the commencement of the study the weight variation of animals used should be minimal and not exceed  $\pm 20\%$  of the mean weight of each sex. When a repeated oral dose is conducted as a preliminary to a longer-term study, it is preferable that animals from the same strain and source should be used in both studies.

**Housing and feeding**

12. All procedures should conform to local standards of laboratory animal care. The temperature in the experimental animal room should be  $22\text{ }^{\circ}\text{C}$  ( $\pm 3\text{ }^{\circ}\text{C}$ ). Although the relative humidity should be at least 30 % and preferably not to exceed 70 % other than during room cleaning, the aim should be 50-60 %. Lighting should be artificial, the photoperiod being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water. The choice of diet may be influenced by the need to ensure a suitable admixture of a test chemical when administered by this method. Animals should be group housed in small groups of the same sex; animals may be housed individually if scientifically justified. For group caging, no more than five animals should be housed per cage.
13. The feed should be regularly analysed for contaminants. A sample of the diet should be retained until finalisation of the report.

**Preparation of animals**

14. Healthy young adult animals are randomly assigned to the control and treatment groups. Cages should be arranged in such a way that possible effects due to cage placement are minimised. The animals are identified uniquely and kept in their cages for at least five days prior to the start of the treatment study to allow for acclimatisation to the laboratory conditions.

**Preparation of doses**

15. The test chemical is administered by gavage or via the diet or drinking water. The method of oral administration is dependent on the purpose of the study, and the physical/chemical/toxico-kinetic properties of the test chemical.

**▼M4**

16. Where necessary, the test chemical is dissolved or suspended in a suitable vehicle. It is recommended that, wherever possible, the use of an aqueous solution/suspension be considered first, followed by consideration of a solution/suspension in oil (e.g. corn oil) and then by possible solution in other vehicles. For vehicles other than water the toxic characteristics of the vehicle must be known. The stability of the test chemical in the vehicle should be determined.

**PROCEDURE****Number and sex of animals**

17. At least 10 animals (five female and five male) should be used at each dose level. If interim euthanasia are planned, the number should be increased by the number of animals scheduled to be euthanised before the completion of the study. Consideration should be given to an additional satellite group of ten animals (five per sex) in the control and in the top dose group for observation of reversibility, persistence, or delayed occurrence of toxic effects, for at least 14 days post treatment.

**Dosage**

18. Generally, at least three test groups and a control group should be used, but if from assessment of other data, no effects would be expected at a dose of 1 000 mg/kg bw/d, a limit test may be performed. If there are no suitable data available, a range finding study (animals of the same strain and source) may be performed to aid the determination of the doses to be used. Except for treatment with the test chemical, animals in the control group should be handled in an identical manner to the test group subjects. If a vehicle is used in administering the test chemical, the control group should receive the vehicle in the highest volume used.
19. Dose levels should be selected taking into account any existing toxicity and (toxico-) kinetic data available for the test chemical or related chemicals. The highest dose level should be chosen with the aim of inducing toxic effects but not death or severe suffering. Thereafter, a descending sequence of dose levels should be selected with a view to demonstrating any dosage related response and no-observed-adverse effects at the lowest dose level (NOAEL). Two to four fold intervals are frequently optimal for setting the descending dose levels and addition of a fourth test group is often preferable to using very large intervals (e.g. more than a factor of 10) between dosages.
20. In the presence of observed general toxicity (e.g. reduced body weight, liver, heart, lung or kidney effects, etc.) or other changes that may not be toxic responses (e.g. reduced food intake, liver enlargement), observed effects on immune, neurological or endocrine sensitive endpoints should be interpreted with caution.

**Limit test**

21. If a test at one dose level of at least 1 000 mg/kg body weight/day or, for dietary or drinking water administration, an equivalent percentage in the diet, or drinking water (based upon body weight determinations), using the procedures described for this study, produces no observable toxic effects and if toxicity would not be expected based upon data from structurally related chemicals, then a full study using three dose levels may not be considered necessary. The limit test applies except when human exposure indicates the need for a higher dose level to be used.

**▼ M4****Administration of doses**

22. The animals are dosed with test chemical daily 7 days each week for a period of 28 days. When the test chemical is administered by gavage, this should be done in a single dose to the animals using a stomach tube or a suitable intubation cannula. The maximum volume of liquid that can be administered at one time depends on the size of the test animal. The volume should not exceed 1 ml/100 g body weight except in the case of aqueous solutions where 2 ml/100 g body weight may be used. Except for irritating or corrosive chemicals, which will normally reveal exacerbated effects with higher concentrations, variability in test volume should be minimized by adjusting the concentration to ensure a constant volume at all dose levels.
  
23. For chemicals administered via the diet or drinking water it is important to ensure that the quantities of the test chemical involved do not interfere with normal nutrition or water balance. When the test chemical is administered in the diet either a constant dietary concentration (ppm) or a constant dose level in terms of the animals' body weight may be used; the alternative used must be specified. For a chemical administered by gavage, the dose should be given at similar times each day, and adjusted as necessary to maintain a constant dose level in terms of animal body weight. Where a repeated dose study is used as a preliminary to a long term study, a similar diet should be used in both studies.

**Observations**

24. The observation period should be 28 days. Animals in a satellite group scheduled for follow-up observations should be kept for at least 14 days without treatment to detect delayed occurrence, or persistence of, or recovery from toxic effects.
  
25. General clinical observations should be made at least once a day, preferably at the same time(s) each day and considering the peak period of anticipated effects after dosing. The health condition of the animals should be recorded. At least twice daily, all animals are observed for morbidity and mortality.
  
26. Once before the first exposure (to allow for within-subject comparisons), and at least once a week thereafter, detailed clinical observations should be made in all animals. These observations should be made outside the home cage in a standard arena and preferably at the same time of day on each occasion. They should be carefully recorded, preferably using scoring systems, explicitly defined by the testing laboratory. Effort should be made to ensure that variations in the test conditions are minimal and that observations are preferably conducted by observers unaware of the treatment. Signs noted should include, but not be limited to, changes in skin, fur, eyes, mucous membranes, occurrence of secretions and excretions and autonomic activity (e.g. lacrimation, piloerection, pupil size, unusual respiratory pattern). Changes in gait, posture and response to handling as well as the presence of clonic or tonic movements, stereotypies (e.g. excessive grooming, repetitive circling) or bizarre behaviour (e.g. self-mutilation, walking backwards) should also be recorded (2).
  
27. In the fourth exposure week sensory reactivity to stimuli of different types (2) (e.g. auditory, visual and proprioceptive stimuli) (3)(4)(5), assessment of grip strength (6) and motor activity assessment (7) should be conducted. Further details of the procedures that could be followed are given in the respective references. However, alternative procedures than those referenced could be used.

**▼ M4**

28. Functional observations conducted in the fourth exposure week may be omitted when the study is conducted as a preliminary study to a subsequent subchronic (90-day) study. In that case, the functional observations should be included in this follow-up study. On the other hand, the availability of data on functional observations from the repeated dose study may enhance the ability to select dose levels for a subsequent subchronic study.
29. As an exception, functional observations may also be omitted for groups that otherwise reveal signs of toxicity to an extent that would significantly interfere with the functional test performance.
30. At necropsy, the oestrus cycle of all females could be determined (optional) by taking vaginal smears. These observations will provide information regarding the stage of oestrus cycle at the time of sacrifice and assist in histological evaluation of estrogen sensitive tissues [see guidance on histopathology (19)].

**Body weight and food/water consumption**

31. All animals should be weighed at least once a week. Measurements of food consumption should be made at least weekly. If the test chemical is administered via the drinking water, water consumption should also be measured at least weekly.

**Haematology**

32. The following haematological examinations should be made at the end of the test period: haematocrit, haemoglobin concentrations, erythrocyte count, reticulocytes, total and differential leucocyte count, platelet count and a measure of blood clotting time/potential. Other determinations that should be carried out, if the test chemical or its putative metabolites have or are suspected to have oxidising properties include methaemoglobin concentration and Heinz bodies.
33. Blood samples should be taken from a named site just prior to or as part of the procedure for euthanasia of the animals, and stored under appropriate conditions. Animals should be fasted overnight prior to euthanasia <sup>(1)</sup>.

**Clinical biochemistry**

34. Clinical biochemistry determinations to investigate major toxic effects in tissues and, specifically, effects on kidney and liver, should be performed on blood samples obtained of all animals just prior to or as part of the procedure for euthanasia of the animals (apart from those found moribund and/or euthanised prior to the termination of the study). Investigations of plasma or serum shall include sodium, potassium, glucose, total cholesterol, urea, creatinine, total protein and albumin, at least two enzymes indicative of hepatocellular effects (such as alanin aminotransferase, aspartate aminotransferase, alkaline phosphatase,  $\gamma$ -glutamyl trans-peptidase and glutamate dehydrogenase), and bile acids. Measurements of additional enzymes (of hepatic or other origin) and bilirubin may provide useful information under certain circumstances.
35. Optionally, the following urinalysis determinations could be performed during the last week of the study using timed urine volume collection; appearance, volume, osmolality or specific gravity, pH, protein, glucose and blood/blood cells.

<sup>(1)</sup> For a number of measurements in serum and plasma, most notably for glucose, overnight fasting would be preferable. The major reason for this preference is that the increased variability which would inevitably result from non-fasting, would tend to mask more subtle effects and make interpretation difficult. On the other hand, however, overnight fasting may interfere with the general metabolism of the animals and, particularly in feeding studies, may disturb the daily exposure to the test chemical. If overnight fasting is adopted, clinical biochemical determinations should be performed after the conduct of functional observations in week 4 of the study.

**▼ M4**

36. In addition, studies to investigate plasma or serum markers of general tissue damage should be considered. Other determinations that should be carried out, if the known properties of the test chemical may, or are suspected to, affect related metabolic profiles include calcium, phosphate, triglycerides, specific hormones, and cholinesterase. These need to be identified for chemicals in certain classes or on a case-by-case basis.
37. Although in the international evaluation of the endocrine related endpoints a clear advantage for the determination of thyroid hormones (T3, T4) and TSH could not be demonstrated, it may be helpful to retain plasma or serum samples to measure T3, T4 and TSH (optional) if there is an indication for an effect on the pituitary-thyroid axis. These samples may be frozen at  $-20^{\circ}$  for storage. The following factors may influence the variability and the absolute concentrations of the hormone determinations:
- time of sacrifice because of diurnal variation of hormone concentrations
  - method of sacrifice to avoid undue stress to the animals that may affect hormone concentrations
  - test kits for hormone determinations that may differ by their standard curves.

Definitive identification of thyroid-active chemicals is more reliable by histopathological analysis rather than hormone levels.

38. Plasma samples specifically intended for hormone determination should be obtained at a comparable time of the day. It is recommended that consideration should be given to T3, T4 and TSH determinations triggered based upon alterations of thyroid histopathology. The numerical values obtained when analysing hormone concentrations differ with various commercial assay kits. Consequently, it may not be possible to provide performance criteria based upon uniform historical data. Alternatively, laboratories should strive to keep control coefficients of variation below 25 for T3 and T4 and below 35 for TSH. All concentrations are to be recorded in ng/ml.
39. If historical baseline data are inadequate, consideration should be given to determination of haematological and clinical biochemistry variables before dosing commences or preferably in a set of animals not included in the experimental groups.

## PATHOLOGY

### Gross necropsy

40. All animals in the study shall be subjected to a full, detailed gross necropsy which includes careful examination of the external surface of the body, all orifices, and the cranial, thoracic and abdominal cavities and their contents. The liver, kidneys, adrenals, testes, epididymides, prostate + seminal vesicles with coagulating glands as a whole, thymus, spleen, brain and heart of all animals (apart from those found moribund and/or euthanised prior to the termination of the study) should be trimmed of any adherent tissue, as appropriate, and their wet weight taken as soon as possible after dissection to avoid drying. Care must be exercised when trimming the prostate complex to avoid puncture of the fluid filled seminal vesicles. Alternatively, seminal vesicles and prostate may be trimmed and weighed after fixation.

**▼ M4**

41. In addition, two other tissues could be optionally weighed as soon as possible after dissection, to avoid drying: paired ovaries (wet weight) and uterus, including cervix (guidance on removal and preparation of the uterine tissues for weight measurement is provided in OECD TG 440 (18)).
  
42. The thyroid weight (optional) could be determined after fixation. Trimming should also be done very carefully and only after fixation to avoid tissue damage. Tissue damage could compromise histopathology analysis.
  
43. The following tissues should be preserved in the most appropriate fixation medium for both the type of tissue and the intended subsequent histopathological examination (see paragraph 47): all gross lesions, brain (representative regions including cerebrum, cerebellum and pons), spinal cord, eye, stomach, small and large intestines (including Peyer's patches), liver, kidneys, adrenals, spleen, heart, thymus, thyroid, trachea and lungs (preserved by inflation with fixative and then immersion), gonads (testis and ovaries), accessory sex organs (uterus and cervix, epididymides, prostate + seminal vesicles with coagulating glands), vagina, urinary bladder, lymph nodes [besides the most proximal draining node another lymph node should be taken according to the laboratory's experience (15)], peripheral nerve (sciatic or tibial) preferably in close proximity to the muscle, skeletal muscle and bone, with bone marrow (section or, alternatively, a fresh mounted bone marrow aspirate). It is recommended that testes be fixed by immersion in Bouin's or modified Davidson's fixative (16) (17). The tunica albuginea must be gently and shallowly punctured at the both poles of the organ with a needle to permit rapid penetration of the fixative. The clinical and other findings may suggest the need to examine additional tissues. Also any organs considered likely to be target organs based on the known properties of the test chemical should be preserved.
  
44. The following tissues may give valuable indication for endocrine-related effects: Gonads (ovaries and testes), accessory sex organs (uterus including cervix, epididymides, seminal vesicles with coagulation glands, dorsolateral and ventral prostate), vagina, pituitary, male mammary gland, the thyroid and adrenal gland. Changes in male mammary glands have not been sufficiently documented but this parameter may be very sensitive to substances with oestrogenic action. Observation of organs/tissues that are not listed in paragraph 43 is optional (see Appendix 2).
  
45. The Guidance on histopathology (19) details extra information on dissection, fixation, sectioning and histopathology of endocrine tissues.
  
46. In the international test program some evidence was obtained that subtle endocrine effects by chemicals with a low potency for affecting sex hormone homeostasis may be identified by disturbance of the synchronisation of the oestrus cycle in different tissues and not so much by frank histopathological alterations in female sex organs. Although no definitive proof was obtained for such effects, it is recommended that evidence of possible asynchrony of the oestrus cycle should be taken into account in interpretation of the histopathology of the ovaries (follicular, thecal, and granulosa cells), uterus, cervix and vagina. If assessed, the stage of cycle as determined by vaginal smears could be included in this comparison as well.

**▼ M4****Histopathology**

47. Full histopathology should be carried out on the preserved organs and tissues of all animals in the control and high dose groups. These examinations should be extended to animals of all other dosage groups, if treatment-related changes are observed in the high dose group.
48. All gross lesions shall be examined.
49. When a satellite group is used, histopathology should be performed on tissues and organs identified as showing effects in the treated groups.

**DATA AND REPORTING****Data**

50. Individual data should be provided. Additionally, all data should be summarised in tabular form showing for each test group the number of animals at the start of the test, the number of animals found dead during the test or euthanised for humane reasons and the time of any death or euthanasia, the number showing signs of toxicity, a description of the signs of toxicity observed, including time of onset, duration, and severity of any toxic effects, the number of animals showing lesions, the type of lesions, their severity and the percentage of animals displaying each type of lesion.
51. When possible, numerical results should be evaluated by an appropriate and generally acceptable statistical method. Comparisons of the effect along a dose range should avoid the use of multiple t-tests. The statistical methods should be selected during the design of the study.
52. For quality control it is proposed that historical control data are collected and that for numerical data coefficients of variation are calculated, especially for the parameters linked with endocrine disrupter detection. These data can be used for comparison purposes when actual studies are evaluated.

**Test report**

53. The test report must include the following information:

*Test chemical:*

- physical nature, purity and physicochemical properties;
- identification data.

*Vehicle (if appropriate):*

- justification for choice of vehicle, if other than water.

*Test animals:*

- species/strain used;
- number, age and sex of animals;
- source, housing conditions, diet, etc.;
- individual weights of animals at the start of the test.
- justification for species if not rat

*Test conditions:*

- rationale for dose level selection;
- details of test chemical formulation/diet preparation, achieved concentration, stability and homogeneity of the preparation;

**▼ M4**

- details of the administration of the test chemical;
- conversion from diet/drinking water test chemical concentration (ppm) to the actual dose (mg/kg body weight/day), if applicable;
- details of food and water quality.

*Optional endpoints investigated*

- list of optional endpoints investigated

*Results:*

- body weight/body weight changes;
- food consumption, and water consumption, if applicable;
- toxic response data by sex and dose level, including signs of toxicity;
- nature, severity and duration of clinical observations (whether reversible or not);
- sensory activity, grip strength and motor activity assessments;
- haematological tests with relevant base-line values;
- clinical biochemistry tests with relevant base-line values;
- body weight at euthanasia and organ weight data;
- necropsy findings;
- a detailed description of all histopathological findings;
- absorption data if available;
- statistical treatment of results, where appropriate.

*Discussion of results**Conclusions*



**▼ M4***Appendix 1***DEFINITIONS**

**Androgenicity** is the capability of a chemical to act like a natural androgenic hormone (e.g. testosterone) in a mammalian organism.

**Antiandrogenicity** is the capability of a chemical to suppress the action of a natural androgenic hormone (e.g. testosterone) in a mammalian organism.

**Antioestrogenicity** is the capability of a chemical to suppress the action of a natural oestrogenic hormone (e.g. oestradiol 17 $\beta$ ) in a mammalian organism.

**Antithyroid activity** is the capability of a chemical to suppress the action of a natural thyroid hormone (e.g. T<sub>3</sub>) in a mammalian organism.

**Dosage** is a general term comprising of dose, its frequency and the duration of dosing.

**Dose** is the amount of test chemical administered. The dose is expressed as weight of test chemical per unit body weight of test animal per day (e.g. mg/kg body weight/day), or as a constant dietary concentration.

**Evident toxicity** is a general term describing clear signs of toxicity following administration of test chemical. These should be sufficient for hazard assessment and should be such that an increase in the dose administered can be expected to result in the development of severe toxic signs and probable mortality.

**NOAEL** is the abbreviation for no-observed-adverse-effect level. This is the highest dose level where no adverse treatment-related findings are observed due to treatment.

**Oestrogenicity** is the capability of a chemical to act like a natural oestrogenic hormone (e.g. oestradiol 17 $\beta$ ) in a mammalian organism.

**Test chemical:** Any substance or mixture tested using this Test Method.

**Thyroid activity** is the capability of a chemical to act like a natural thyroid hormone (e.g. T<sub>3</sub>) in a mammalian organism.

**Validation** is a scientific process designed to characterise the operational requirements and limitations of a test method and to demonstrate its reliability and relevance for a particular purpose.

▼ **M4***Appendix 2***Endpoints recommended for the detection of endocrine disrupters (EDs) in this Test Method B.7**

Mandatory endpoints	Optional endpoints
Weight	
<ul style="list-style-type: none"> <li>— Testes</li> <li>— Epididymides</li> <li>— Adrenals</li> <li>— Prostate + seminal vesicles with coagulating glands</li> </ul>	<ul style="list-style-type: none"> <li>— Ovaries</li> <li>— Uterus, including cervix</li> <li>— Thyroid</li> </ul>
Histopathology	
<ul style="list-style-type: none"> <li>— Gonads: <ul style="list-style-type: none"> <li>— Testes and</li> <li>— Ovaries</li> </ul> </li> <li>— Accessory sex organs: <ul style="list-style-type: none"> <li>— Epididymides,</li> <li>— Prostate + seminal vesicle with coagulating glands</li> <li>— Uterus, including cervix</li> </ul> </li> <li>— Adrenal</li> <li>— Thyroid</li> <li>— Vagina</li> </ul>	<ul style="list-style-type: none"> <li>— Vaginal smears</li> <li>— Male mammary glands</li> <li>— Pituitary</li> </ul>
Hormones measurement	
	<ul style="list-style-type: none"> <li>— Circulating levels of T3, T4</li> <li>— Circulating levels of TSH</li> </ul>

*LITERATURE:*

- (1) OECD (Paris, 1992). Chairman's Report of the Meeting of the ad hoc Working Group of Experts on Systemic Short-term and (Delayed) Neurotoxicity.
- (2) IPCS (1986). Principles and Methods for the Assessment of Neurotoxicity Associated with Exposure to Chemicals. Environmental Health Criteria Document No 60.
- (3) Tupper DE, Wallace RB (1980). Utility of the Neurologic Examination in Rats. *Acta Neurobiol. Exp.* 40: 999-1003.
- (4) Gad SC (1982). A Neuromuscular Screen for Use in Industrial Toxicology. *J. Toxicol Environ. Health* 9: 691-704.
- (5) Moser VC, McDaniel KM, Phillips PM (1991). Rat Strain and Stock Comparisons Using a Functional Observational Battery: Baseline Values and Effects of Amitraz. *Toxicol. Appl. Pharmacol.* 108: 267-283.
- (6) Meyer OA, Tilson HA, Byrd WC, Riley MT (1979). A Method for the Routine Assessment of Fore- and Hindlimb Grip Strength of Rats and Mice. *Neurobehav. Toxicol.* 1: 233-236.

**▼M4**

- (7) Crofton KM, Howard JL, Moser VC, Gill MW, Reiter LW, Tilson HA, MacPhail RC (1991). Interlaboratory Comparison of Motor Activity Experiments: Implication for Neurotoxicological Assessments. *Neurotoxicol. Teratol.* 13: 599-609.
- (8) OECD (1998). Report of the First Meeting of the OECD Endocrine Disrupter Testing and Assessment (EDTA) Task Force, 10th-11 March 1998, ENV/MC/CHEM/RA(98)5.
- (9) OECD. (2006). Report of the Validation of the Updated Test Guideline 407: Repeat Dose 28-day Oral Toxicity Study in Laboratory Rats. Series on Testing and Assessment No 59, ENV/JM/MONO(2006)26.
- (10) OECD (2002). Detailed Review Paper on the Appraisal of Test Methods for Sex Hormone Disrupting Chemicals. Series on Testing and Assessment No 21, ENV/JM/MONO(2002)8.
- (11) OECD (2012). Conceptual Framework for Testing and Assessment of Endocrine Disrupting Chemicals. [http://www.oecd.org/document/58/0,3343,fr\\_2649\\_37407\\_2348794\\_1\\_1\\_1\\_37407,00.html](http://www.oecd.org/document/58/0,3343,fr_2649_37407_2348794_1_1_1_37407,00.html)
- (12) OECD (2006). Final Summary report of the meeting of the Validation Management Group for mammalian testing. ENV/JM/TG/EDTA/M(2006)2.
- (13) OECD. Draft Summary record of the meeting of the Task Force on Endocrine Disrupters Testing and Assessment. ENV/JM/TG/EDTA/M(2006)3.
- (14) OECD (2000). Guidance document on the recognition, assessment and use of clinical signs as humane endpoints for experimental animals used in safety evaluation. Series on Testing and Assessment No 19. ENV/JM/MONO(2000)7.
- (15) Haley P, Perry R, Ennulat D, Frame S, Johnson C, Lapointe J-M, Nyska A, Snyder PW, Walker D, Walter G (2005). STP Position Paper: Best Practice Guideline for the Routine Pathology Evaluation of the Immune System. *Toxicol Pathol* 33: 404-407.
- (16) Hess RA, Moore BJ (1993). Histological Methods for the Evaluation of the Testis. In: *Methods in Reproductive Toxicology*, Chapin RE and Heindel JJ (eds). Academic Press: San Diego, CA, pp. 52-85.
- (17) Latendresse JR, Warbritton AR, Jonassen H, Creasy DM.(2002) Fixation of testes and eyes using a modified Davidson's fluid: comparison with Bouin's fluid and conventional Davidson's fluid. *Toxicol. Pathol.* 30, 524-533.
- (18) OECD (2007). OECD Guideline for Testing of Chemicals No 440: Uterotrophic Bioassay in Rodents: A short-term screening test for oestrogenic properties.
- (19) OECD (2009). Guidance Document 106 on Histologic evaluation of Endocrine and Reproductive Tests in Rodents ENV/JM/Mono(2009)11.

**▼ M4****B.8. SUBACUTE INHALATION TOXICITY: 28-DAY STUDY****SUMMARY**

This revised Test Method B.8 has been designed to fully characterise test chemical toxicity by the inhalation route following repeated exposure for a limited period of time (28 days), and to provide data for quantitative inhalation risk assessments. Groups of at least 5 male and 5 female rodents are exposed 6 hours per day for 28 days to a) the test chemical at three or more concentration levels, b) filtered air (negative control), and/or c) the vehicle (vehicle control). Animals are generally exposed 5 days per week but exposure for 7 days per week is also allowed. Males and females are always tested, but they may be exposed at different concentration levels if it is known that one sex is more susceptible to a given test chemical. This method allows the study director the flexibility to include satellite (reversibility) groups, bronchoalveolar lavage (BAL), neurologic tests, and additional clinical pathology and histopathological evaluations in order to better characterise the toxicity of a test chemical.

**INTRODUCTION**

1. This Test Method is equivalent to OECD Test Guideline 412 (2009). The original subacute inhalation Test Guideline 412 (TG 412) was adopted in 1981 (1). This Test Method B.8 (as equivalent to the revised TG 412) has been updated to reflect the state of science and to meet current and future regulatory needs.
2. This method enables the characterisation of adverse effects following repeated daily inhalation exposure to a test chemical for 28 days. The data derived from 28-day sub-acute inhalation toxicity studies can be used for quantitative risk assessments [if not followed by a 90-day subchronic inhalation toxicity study (Chapter B.29 of this Annex)]. The data can also provide information on the selection of concentrations for longer term studies such as the 90-day subchronic inhalation toxicity study. This test method is not specifically intended for the testing of nanomaterials. Definitions used in the context of this Test Method are provided at the end of this chapter and in the Guidance Document 39 (2).

**INITIAL CONSIDERATIONS**

3. All available information on the test chemical should be considered by the testing laboratory prior to conducting the study in order to enhance the quality of the study and minimize animal usage. Information that will assist in the selection of appropriate test concentrations might include the identity, chemical structure, and physico-chemical properties of the test chemical; results of any *in vitro* or *in vivo* toxicity tests; anticipated use(s) and potential for human exposure; available (Q)SAR data and toxicological data on structurally related chemicals; and data derived from acute inhalation toxicity testing. If neurotoxicity is expected or is observed in the course of the study, the study director may choose to include appropriate evaluations such as a functional observational battery (FOB) and measurement of motor activity. Although the timing of exposures relative to specific examinations may be critical, the performance of these additional activities should not interfere with the basic study design.

**▼M4**

4. Dilutions of corrosive or irritating test chemicals may be tested at concentrations that will yield the desired degree of toxicity [refer to GD 39 (2)]. When exposing animals to these materials, the targeted concentrations should be low enough to not cause marked pain and distress, yet sufficient to extend the concentration-response curve to levels that reach the regulatory and scientific objective of the test. These concentrations should be selected on a case-by-case basis, preferably based upon an adequately designed range-finding study that provides information regarding the critical endpoint, any irritation threshold, and the time of onset (see paragraphs 11-13). The justification for concentration selection should be provided.
5. Moribund animals or animals obviously in pain or showing signs of severe and enduring distress should be humanely killed. Moribund animals are considered in the same way as animals that die on test. Criteria for making the decision to kill moribund or severely suffering animals, and guidance on the recognition of predictable or impending death, are the subject of an OECD Guidance Document on Humane Endpoints (3).

**DESCRIPTION OF THE METHOD****Selection of Animal Species**

6. Healthy young adult rodents of commonly used laboratory strains should be employed. The preferred species is the rat. Justification should be provided if other species are used.

**Preparation of Animals**

7. Females should be nulliparous and non-pregnant. On the day of randomisation, animals should be young adults 7 to 9 weeks of age. Body weights should be within  $\pm 20\%$  of the mean weight for each sex. The animals are randomly selected, marked for individual identification, and kept in their cages for at least 5 days prior to the start of the test to allow for acclimatisation to laboratory conditions.

**Animal Husbandry**

8. Animals should be individually identified, if possible with subcutaneous transponders, to facilitate observations and avoid confusion. The temperature of the experimental animal maintenance room should be  $22 \pm 3$  °C. The relative humidity should ideally be maintained in the range of 30 to 70 %, though this may not be possible when using water as a vehicle. Before and after exposures, animals generally should be caged in groups by sex and concentration, but the number of animals per cage should not interfere with clear observation of each animal and should minimise losses due to cannibalism and fighting. When animals are to be exposed nose-only, it may be necessary for them to be acclimated to the restraining tubes. The restraining tubes should not impose undue physical, thermal, or immobilisation stress on the animals. Restraint may affect physiological endpoints such as body temperature (hyperthermia) and/or respiratory minute volume. If generic data are available to show that no such changes occur to any appreciable extent, then pre-adaptation to the restraining tubes is not necessary. Animals exposed whole-body to an aerosol should be housed individually during exposure to prevent them from filtering the test aerosol through the fur of their cage mates. Conventional and certified laboratory diets may be used, except during exposure, accompanied with an unlimited supply of municipal drinking water. Lighting should be artificial, the sequence being 12 hours light/12 hours dark.

**▼ M4****Inhalation Chambers**

9. The nature of the test chemical and the object of the test should be considered when selecting an inhalation chamber. The preferred mode of exposure is nose-only (which term includes head-only, nose-only, or snout-only). Nose-only exposure is generally preferred for studies of liquid or solid aerosols and for vapours that may condense to form aerosols. Special objectives of the study may be better achieved by using a whole-body mode of exposure, but this should be justified in the study report. To ensure atmosphere stability when using a whole-body chamber, the total 'volume' of the test animals should not exceed 5 % of the chamber volume. Principles of the nose-only and whole-body exposure techniques and their particular advantages and disadvantages are addressed in GD 39 (2).

## TOXICITY STUDIES

**Limit Concentrations**

10. Unlike with acute studies, there are no defined limit concentrations in 28-day sub-acute inhalation toxicity studies. The maximum concentration tested should consider: (1) the maximum attainable concentration, (2) the 'worst case' human exposure level, (3) the need to maintain an adequate oxygen supply, and/or (4) animal welfare considerations. In the absence of data-based limits, the acute limits of the Regulation (EC) No 1272/2008 (13) may be used (i.e. up to a maximum concentration of 5 mg/l for aerosols, 20 mg/l for vapours and 20 000 ppm for gases); refer to GD 39 (2). Justification should be provided if it is necessary to exceed these limits when testing gases or highly volatile test chemicals (e.g. refrigerants). The limit concentration should elicit unequivocal toxicity without causing undue stress to the animals or affecting their longevity (3).

**Range-Finding Study**

11. Before commencing with the main study, it may be necessary to perform a range-finding study. A range-finding study is more comprehensive than a sighting study because it is not limited to concentration selection. Knowledge learned from a range-finding study can lead to a successful main study. A range-finding study may, for example, provide technical information regarding analytical methods, particle sizing, discovery of toxic mechanisms, clinical pathology and histopathological data, and estimations of what may be NOAEL and MTC concentrations in a main study. The study director may choose to use the range-finding study to identify the threshold of respiratory tract irritation (e.g. with histopathology of the respiratory tract, pulmonary function testing, or bronchoalveolar lavage), the upper concentration which is tolerated without undue stress to the animals, and the parameters that will best characterise a test chemical's toxicity.
12. A range-finding study may consist of one or more concentration levels. No more than three males and three females should be exposed at each concentration level. A range-finding study should last a minimum of 5 days and generally no more than 14 days. The rationale for the selection of concentrations for the main study should be provided in the study report. The objective of the main study is to demonstrate a concentration-response relationship based on what is anticipated to be the most sensitive endpoint. The low concentration should ideally be a no-observed-adverse effect concentration while the high concentration should elicit unequivocal toxicity without causing undue stress to the animals or affecting their longevity (3).

**▼ M4**

13. When selecting concentration levels for the range-finding study, all available information should be considered including structure-activity relationships and data for similar chemicals (see paragraph 3). A range-finding study may verify/refute what are considered to be the most sensitive mechanistically based endpoints, e.g. cholinesterase inhibition by organophosphates, methaemoglobin formation by erythrocytotoxic agents, thyroidal hormones (T<sub>3</sub>, T<sub>4</sub>) for thyrotoxicants, protein, LDH, or neutrophils in bronchoalveolar lavage for innocuous poorly soluble particles or pulmonary irritant aerosols.

**Main Study**

14. The main sub-acute toxicity study generally consists of three concentration levels, and also concurrent negative (air) and/or vehicle controls as needed (see paragraph 17). All available data should be utilised to aid selection of appropriate exposure levels, including the results of systemic toxicity studies, metabolism and kinetics (particular emphasis should be given to avoiding high concentration levels which saturate kinetic processes). Each test group contains at least 10 rodents (5 male and 5 female) that are exposed to the test chemical for 6 hours per day on a 5 day per week basis for a period of 4 weeks (total study duration of 28 days). Animals may also be exposed 7 days per week (e.g. when testing inhaled pharmaceuticals). If one sex is known to be more susceptible to a given test chemical, the sexes may be exposed at different concentration levels in order to optimise the concentration-response as described in paragraph 15. If rodent species other than rats are exposed nose-only, maximum exposure durations may be adjusted to minimise species-specific distress. A rationale should be provided when using an exposure duration less than 6 hours/day, or when it is necessary to conduct a long duration (e.g. 22 hours/day) whole-body exposure study [refer to GD 39 (2)]. Feed should be withheld during the exposure period unless exposure exceeds 6 hours. Water may be provided throughout a whole-body exposure.
15. The target concentrations selected should identify the target organ(s) and demonstrate a clear concentration-response:
  - The high concentration level should result in toxic effects but not cause lingering signs or lethality which would prevent a meaningful evaluation.
  - The intermediate concentration level(s) should be spaced to produce a gradation of toxic effects between that of the low and high concentration.
  - The low concentration level should produce little or no evidence of toxicity.

**Satellite (Reversibility) Study**

16. A satellite (reversibility) study may be used to observe reversibility, persistence, or delayed occurrence of toxicity for a post-treatment period of an appropriate length, but no less than 14 days. Satellite (reversibility) groups consist of five males and five females exposed contemporaneously with the experimental animals in the main study. Satellite (reversibility) study groups should be exposed to the test chemical at the highest concentration level and there should be concurrent air and/or vehicle controls as needed (see paragraph 17).

**▼ M4****Control Animals**

17. Concurrent negative (air) control animals should be handled in a manner identical to the test group animals except that they are exposed to filtered air rather than test chemical. When water or another substance is used to assist in generating the test atmosphere, a vehicle control group, instead of a negative (air) control group, should be included in the study. Water should be used as the vehicle whenever possible. When water is used as the vehicle, the control animals should be exposed to air with the same relative humidity as the exposed groups. The selection of a suitable vehicle should be based on an appropriately conducted pre-study or historical data. If a vehicle's toxicity is not well known, the study director may choose to use both a negative (air) control and a vehicle control, but this is strongly discouraged. If historical data reveal that a vehicle is non-toxic, then there is no need for a negative (air) control group and only a vehicle control should be used. If a pre-study of a test chemical formulated in a vehicle reveals no toxicity, it follows that the vehicle is non-toxic at the concentration tested and this vehicle control should be used.

## EXPOSURE CONDITIONS

**Administration of Concentrations**

18. Animals are exposed to the test chemical as a gas, vapour, aerosol, or a mixture thereof. The physical state to be tested depends on the physico-chemical properties of the test chemical, the selected concentration, and/or the physical form most likely present during the handling and use of the test chemical. Hygroscopic and chemically reactive test chemicals should be tested under dry air conditions. Care should be taken to avoid generating explosive concentrations. Particulate material may be subjected to mechanical processes to decrease the particle size. Further guidance is provided in GD 39 (2).

**Particle-Size Distribution**

19. Particle sizing should be performed for all aerosols and for vapours that may condense to form aerosols. To allow for exposure of all relevant regions of the respiratory tract, aerosols with mass median aerodynamic diameters (MMAD) ranging from 1 to 3  $\mu\text{m}$  with a geometric standard deviation ( $\sigma_g$ ) in the range of 1,5 to 3,0 are recommended (4). Although a reasonable effort should be made to meet this standard, expert judgement should be provided if it cannot be achieved. For example, metal fume particles may be smaller than this standard, and charged particles and fibres may exceed it.

**Test chemical Preparation in a Vehicle**

20. Ideally, the test chemical should be tested without a vehicle. If it is necessary to use a vehicle to generate an appropriate test chemical concentration and particle size, water should be given preference. Whenever a test chemical is dissolved in a vehicle, its stability should be demonstrated.

## MONITORING OF EXPOSURE CONDITIONS

**Chamber Airflow**

21. The flow of air through the exposure chamber should be carefully controlled, continuously monitored, and recorded at least hourly during each exposure. The real-time monitoring of the test atmosphere concentration (or temporal stability) is an integral measurement of all dynamic parameters and provides an indirect means to control all relevant dynamic inhalation parameters. If the concentration is monitored real-time, the frequency of measurement of air flows may be reduced to one single



**▼M4**

measurement per exposure per day. Special consideration should be given to avoiding re-breathing in nose-only chambers. Oxygen concentration should be at least 19 % and carbon dioxide concentration should not exceed 1 %. If there is reason to believe that this standard cannot be met, oxygen and carbon dioxide concentrations should be measured. If measurements on the first day of exposure show that these gases are at proper levels, no further measurements should be necessary.

**Chamber Temperature and Relative Humidity**

22. Chamber temperature should be maintained at  $22 \pm 3$  °C. Relative humidity in the animals' breathing zone, for both nose-only and whole-body exposures, should be monitored continuously and recorded hourly during each exposure where possible. The relative humidity should preferably be maintained in the range of 30 to 70 %, but this may either be unattainable (e.g. when testing water based mixtures) or not measurable due to test chemical interference with the Test Method.

**Test chemical: Nominal Concentration**

23. Whenever feasible, the nominal exposure chamber concentration should be calculated and recorded. The nominal concentration is the mass of generated test chemical divided by the total volume of air passed through the inhalation chamber system. The nominal concentration is not used to characterise the animals' exposure, but a comparison of the nominal concentration and the actual concentration gives an indication of the generation efficiency of the test system, and thus may be used to discover generation problems.

**Test chemical: Actual Concentration**

24. The actual concentration is the test chemical concentration as sampled at the animals' breathing zone in an inhalation chamber. Actual concentrations can be obtained either by specific methods (e.g. direct sampling, adsorptive or chemical reactive methods, and subsequent analytical characterisation) or by non-specific methods such as gravimetric filter analysis. The use of gravimetric analysis is acceptable only for single component powder aerosols or aerosols of low volatility liquids and should be supported by appropriate pre-study test chemical-specific characterisations. Multi-component powder aerosol concentration may also be determined by gravimetric analysis. However, this requires analytical data which demonstrate that the composition of airborne material is similar to the starting material. If this information is not available, a reanalysis of the test chemical (ideally in its airborne state) at regular intervals during the course of the study may be necessary. For aerosolised agents that may evaporate or sublime, it should be shown that all phases were collected by the method chosen.
25. One lot of the test chemical should be used throughout the duration of the study, if possible, and the test sample should be stored under conditions that maintain its purity, homogeneity, and stability. Prior to the start of the study, there should be a characterisation of the test chemical including its purity and, if technically feasible, the identity, and quantities of identified contaminants and impurities. This can be demonstrated but is not limited by the following data: retention time and relative peak area, molecular weight from mass spectroscopy or gas chromatography analyses, or other estimates. Although the test sample's identity is not the responsibility of the test laboratory, it may be prudent for the test laboratory to confirm the sponsor's characterisation at least in a limited way (e.g. colour, physical nature, etc.).

**▼ M4**

26. The exposure atmosphere should be held as constant as practicable. A real-time monitoring device, such as an aerosol photometer for aerosols or a total hydrocarbon analyser for vapours may be used to demonstrate the stability of the exposure conditions. Actual chamber concentration should be measured at least 3 times during each exposure day for each exposure level. If not feasible due to limited air flow rates or low concentrations, one sample per exposure period is acceptable. Ideally, this sample should then be collected over the entire exposure period. Individual chamber concentration samples should deviate from the mean chamber concentration by no more than  $\pm 10\%$  for gases and vapours, and by no more than  $\pm 20\%$  for liquid or solid aerosols. Time to attain chamber equilibration ( $t_{95}$ ) should be calculated and reported. The duration of an exposure spans the time that the test chemical is generated. This takes into account the times required to attain chamber equilibration ( $t_{95}$ ) and decay. Guidance for estimating  $t_{95}$  can be found in GD 39 (2).
27. For very complex mixtures consisting of gases/vapours and aerosols (e.g. combustion atmospheres and test chemicals propelled from purpose-driven end-use products/devices), each phase may behave differently in an inhalation chamber. Therefore, at least one indicator substance (analyte), normally the principal active substance in the mixture, of each phase (gas/vapour and aerosol) should be selected. When the test chemical is a mixture, the analytical concentration should be reported for the total mixture, and not just for the active ingredient or the indicator substance (analyte). Additional information regarding actual concentrations can be found in GD 39 (2).

**Test chemical: Particle Size Distribution**

28. The particle size distribution of aerosols should be determined at least weekly for each concentration level by using a cascade impactor or an alternative instrument, such as an aerodynamic particle sizer (APS). If equivalence of the results obtained by a cascade impactor and the alternative instrument can be shown, then the alternative instrument may be used throughout the study.
29. A second device, such as a gravimetric filter or an impinger/gas bubbler, should be used in parallel to the primary instrument to confirm the collection efficiency of the primary instrument. The mass concentration obtained by particle size analysis should be within reasonable limits of the mass concentration obtained by filter analysis [see GD 39 (2)]. If equivalence can be demonstrated at all concentrations tested in the early phase of the study, then further confirmatory measurements may be omitted. For the sake of animal welfare, measures should be taken to minimise inconclusive data which may lead to a need to repeat a study.
30. Particle sizing should be performed for vapours if there is any possibility that vapour condensation may result in the formation of an aerosol, or if particles are detected in a vapour atmosphere with potential for mixed phases.

**OBSERVATIONS**

31. The animals should be clinically observed before, during and after the exposure period. More frequent observations may be indicated depending on the response of the animals during exposure. When animal observation is hindered by the use of animal restraint tubes, poorly lit whole body chambers, or opaque atmospheres, animals should be carefully observed after exposure. Observations before the next day's exposure can assess any reversibility or exacerbation of toxic effects.

**▼ M4**

32. All observations are recorded with individual records being maintained for each animal. When animals are killed for humane reasons or found dead, the time of death should be recorded as precisely as possible.
  
33. Cage-side observations should include changes in the skin and fur, eyes, and mucous membranes; changes in the respiratory and circulatory systems, changes in the nervous system, and changes in somatomotor activity and behaviour patterns. Attention should be directed to observations of tremors, convulsions, salivation, diarrhoea, lethargy, sleep, and coma. The measurement of rectal temperatures may provide supportive evidence of reflex bradypnea or hypo/hyperthermia related to treatment or confinement. Additional assessments may be included in the study protocol such as kinetics, biomonitoring, lung function, retention of poorly soluble materials that accumulate in lung tissue, and behavioural changes.

**BODY WEIGHTS**

34. Individual animal weights should be recorded shortly before the first exposure (day 0), twice weekly thereafter (for example: on Fridays and Mondays to demonstrate recovery over an exposure-free weekend or at a time interval to allow assessment of systemic toxicity), and at the time of death or euthanasia. If there are no effects in the first 2 weeks, body weights may be measured weekly for the remainder of the study. Satellite (reversibility) animals (if used) should continue to be weighed weekly throughout the recovery period. At study termination, all animals should be weighed shortly before sacrifice to allow for an unbiased calculated of organ to body weight ratios.

**FOOD AND WATER CONSUMPTION**

35. Food consumption should be measured weekly. Water consumption may also be measured.

**CLINICAL PATHOLOGY**

36. Clinical pathology assessments should be made for all animals, including control and satellite (reversibility) animals, when they are sacrificed. The time interval between the end of exposure and blood collection should be recorded, particularly when the reconstitution of the addressed endpoint is rapid. Sampling following the end of exposure is indicated for those parameters with a short plasma half-time (e.g. COHb, CHE, and MetHb).
  
37. Table 1 lists the clinical pathology parameters that are generally required for all toxicology studies. Urinalysis is not required on a routine basis, but may be performed when deemed useful based on expected or observed toxicity. The study director may choose to assess additional parameters in order to better characterise a test chemical's toxicity (e.g. cholinesterase, lipids, hormones, acid/base balance, methaemoglobin or Heinz bodies, creatine kinase, myeloid/erythroid ratio, troponins, arterial blood gases, lactate dehydrogenase, sorbitol dehydrogenase, glutamate dehydrogenase, and gamma glutamyl transpeptidase).

▼ **M4**

Table 1

**Standard Clinical Pathology Parameters**

Haematology	
Erythrocyte count	Total leukocyte count
Haematocrit	Differential leukocyte count
Haemoglobin concentration	Platelet count
Mean corpuscular haemoglobin	Clotting potential (select one):
Mean corpuscular volume	— Prothrombin time
Mean corpuscular haemoglobin concentration	— Clotting time
Reticulocytes	— Partial thromboplastin time
Clinical Chemistry	
Glucose (*)	Alanine aminotransferase
Total cholesterol	Aspartate aminotransferase
Triglycerides	Alkaline phosphatase
Blood urea nitrogen	Potassium
Total bilirubin	Sodium
Creatinine	Calcium
Total protein	Phosphorus
Albumin	Chloride
Globulin	
Urinalysis (optional)	
Appearance (colour and turbidity)	Total protein
Volume	Glucose
Specific gravity or osmolality	Blood/blood cells
pH	

(\*) Because a lengthy fasting period can introduce bias in glucose measurements for the treated versus control animals, the study director should determine whether it is appropriate to fast the animals. If a fasting period is used, it should be appropriate to the species used; for the rat this may be 16 h (overnight fasting). Determination of fasting glucose may be carried out after overnight fasting during the last exposure week, or after overnight fasting prior to necropsy (in the latter case together with all other clinical pathology parameters).

38. When there is evidence that the lower respiratory tract (i.e., the alveoli) is the primary site of deposition and retention, then bronchoalveolar lavage (BAL) may be the technique of choice to quantitatively analyse hypothesis-based dose-effect parameters focusing on alveolitis, pulmonary inflammation, and phospholipidosis. This allows for dose-response and time-course changes of alveolar injury to be suitably probed. The BAL fluid may be analysed for total and differential leukocyte counts, total protein, and lactate dehydrogenase. Other parameters that may be considered are those indicative of lysosomal injury, phospholipidosis, fibrosis, and irritant or allergic inflammation which may include the determination of pro-inflammatory cytokines/chemokines. BAL measurements generally complement the results from histopathology examinations but cannot replace them. Guidance on how to perform lung lavage can be found in GD 39 (2).

▼ **M4**

## GROSS PATHOLOGY AND ORGAN WEIGHTS

39. All test animals, including those which die during the test or are removed from the study for animal welfare reasons, should be subjected to complete exsanguination (if feasible) and gross necropsy. The time between the end of each animal's last exposure and their sacrifice should be recorded. If a necropsy cannot be performed immediately after a dead animal is discovered, the animal should be refrigerated (not frozen) at a temperature low enough to minimise autolysis. Necropsies should be performed as soon as possible, normally within a day or two. All gross pathological changes should be recorded for each animal with particular attention to any changes in the respiratory tract.
40. Table 2 lists the organs and tissues that should be preserved in a suitable medium during gross necropsy for histopathological examination. The preservation of the [bracketed] organs and tissues and any other organs and tissues is at the discretion of the study director. The **bolded** organs should be trimmed and weighed wet as soon as possible after dissection to avoid drying. The thyroid and epididymides should only be weighed if needed because trimming artefacts may hinder histopathological evaluation. Tissues and organs should be fixed in 10 % buffered formalin or another suitable fixative as soon as necropsy is performed, and no less than 24-48 hours prior to trimming depending on the fixative to be used.

Table 2

## Organs and Tissues Preserved During Gross Necropsy

<b>Adrenals</b>	Seminal vesicles
Bone marrow (and/or fresh aspirate)	Spinal cord (cervical, mid-thoracic, and lumbar)
<b>Brain</b> (including sections of cerebrum, cerebellum, and medulla/pons)	<b>Spleen</b>
[Eyes (retina, optic nerve) and eyelids]	Stomach
<b>Heart</b>	<b>Testes</b>
<b>Kidneys</b>	<b>Thymus</b>
Larynx (3 levels, 1 level to include the base of the epiglottis)	Thyroid
<b>Liver</b>	Trachea (at least 2 levels including 1 longitudinal section through the carina and 1 transverse section)
<b>Lung</b> (all lobes at one level, including main bronchi)	[Urinary bladder]
Lymph nodes from the hilar region of the lung, especially for poorly soluble particulate test chemicals, For more in depth examinations and/or studies with immunological focus, additional lymph nodes may be considered, e.g. those from the mediastinal, cervical/submandibular and/or auricular regions.	Uterus
Nasopharyngeal tissues (at least 4 levels; 1 level to include the nasopharyngeal duct and the Nasal Associated Lymphoid Tissue(NALT))	All gross lesions
Oesophagus	
[Olfactory bulb]	
<b>Ovaries</b>	

**▼ M4**

41. The lungs should be removed intact, weighed, and instilled with a suitable fixative at a pressure of 20-30 cm of water to ensure that lung structure is maintained (5). Sections should be collected for all lobes at one level, including main bronchi, but if lung lavage is performed, the unlavaged lobe should be sectioned at three levels (not serial sections).
42. At least 4 levels of the nasopharyngeal tissues should be examined, one of which should include the nasopharyngeal duct, (5, 6, 7, 8, 9) to allow adequate examination of the squamous, transitional (non-ciliated respiratory), respiratory (ciliated respiratory) and olfactory epithelium, and the draining lymphatic tissue (NALT; 10, 11). Three levels of the larynx should be examined, and one of these levels should include the base of the epiglottis (12). At least two levels of the trachea should be examined including one longitudinal section through the carina of the bifurcation of the extrapulmonary bronchi and one transverse section.

**HISTOPATHOLOGY**

43. A histopathological evaluation of all the organs and tissues listed in Table 2 should be performed for the control and high concentration groups, and for all animals which die or are sacrificed during the study. Particular attention should be paid to the respiratory tract, target organs, and gross lesions. The organs and tissues that have lesions in the high concentration group should be examined in all groups. The study director may choose to perform histopathological evaluations for additional groups to demonstrate a clear concentration response. When a satellite (reversibility) group is used, histopathological evaluation should be performed for all tissues and organs identified as showing effects in the treated groups. If there are excessive early deaths or other problems in the high exposure group that compromise the significance of the data, the next lower concentration should be examined histopathologically. An attempt should be made to correlate gross observations with microscopic findings.

**DATA AND REPORTING****Data**

44. Individual animal data on body weights, food consumption, clinical pathology, gross pathology, organ weights, and histopathology should be provided. Clinical observation data should be summarised in tabular form showing for each test group the number of animals used, the number of animals displaying specific signs of toxicity, the number of animals found dead during the test or killed for humane reasons, time of death of individual animals, a description and time course of toxic effects and reversibility, and necropsy findings. All results, quantitative and incidental, should be evaluated by an appropriate statistical method. Any generally accepted statistical method may be used and the statistical methods should be selected during the design of the study.

**Test Report**

45. The test report should include the following information, as appropriate:

*Test animals and husbandry*

- Description of caging conditions, including: number (or change in number) of animals per cage, bedding material, ambient temperature and relative humidity, photoperiod, and identification of diet.

**▼ M4**

- Species/strain used and justification for using a species other than the rat. Source and historical data may be provided, if they are from animals exposed under similar exposure, housing, and fasting conditions.
- Number, age, and sex of animals.
- Method of randomisation.
- Description of any pre-test conditioning including diet, quarantine, and treatment for disease.

*Test chemical*

- Physical nature, purity, and, where relevant, physico-chemical properties (including isomerisation).
- Identification data and Chemical Abstract Services (CAS) Registry Number, if known.

*Vehicle*

- Justification for use of vehicle and justification for choice of vehicle (if other than water).
- Historical or concurrent data demonstrating that the vehicle does not interfere with the outcome of the study.

*Inhalation chamber*

- Detailed description of the inhalation chamber including volume and a diagram.
- Source and description of equipment used for the exposure of animals as well as generation of the atmosphere.
- Equipment for measuring temperature, humidity, particle-size, and actual concentration.
- Source of air and system used for conditioning.
- Methods used for calibration of equipment to ensure a homogeneous test atmosphere.
- Pressure difference (positive or negative).
- Exposure ports per chamber (nose-only); location of animals in the chamber (whole-body).
- Stability of the test atmosphere.
- Location of temperature and humidity sensors and sampling of test atmosphere in the chamber.
- Treatment of air supplied/extracted.
- Air flow rates, air flow rate/exposure port (nose-only), or animal load/chamber (whole-body).
- Time to inhalation chamber equilibrium (t95).
- Number of volume changes per hour.
- Metering devices (if applicable).

*Exposure data*

- Rationale for target concentration selection in the main study.

**▼ M4**

- Nominal concentrations (total mass of test chemical generated into the inhalation chamber divided by the volume of air passed through the chamber).
- Actual test chemical concentrations collected from the animals' breathing zone; for mixtures that produce heterogeneous physical forms (gases, vapours, aerosols), each may be analysed separately.
- All air concentrations should be reported in units of mass (mg/l mg/m<sup>3</sup>, etc.) rather than in units of volume (ppm, ppb, etc.).
- Particle size distribution, mass median aerodynamic diameter (MMAD), and geometric standard deviation ( $\sigma_g$ ), including their methods of calculation. Individual particle size analyses should be reported.

*Test conditions*

- Details of test chemical preparation, including details of any procedures used to reduce the particle size of solids or to prepare solutions of the test chemical.
- A description (preferably including a diagram) of the equipment used to generate the test atmosphere and to expose the animals to the test atmosphere.
- Details of the equipment used to monitor chamber temperature, humidity, and chamber airflow (i.e. development of a calibration curve).
- Details of the equipment used to collect samples for determination of chamber concentration and particle size distribution.
- Details of the chemical analytical method used and method validation (including efficiency of recovery of test chemical from the sampling medium).
- Method of randomisation in assigning animals to test and control groups.
- Details of food and water quality (including diet type/source, water source).
- The rationale for the selection of test concentrations.

*Results*

- Tabulation of chamber temperature, humidity, and airflow.
- Tabulation of chamber nominal and actual concentration data.
- Tabulation of particle size data including analytical sample collection data, particle size distribution, and calculations of the MMAD and  $\sigma_g$ .
- Tabulation of response data and concentration level for each animal (i.e. animals showing signs of toxicity including mortality, nature, severity, time of onset, and duration of effects).



**▼M4**

- Tabulation of individual animal weights.
- Tabulation of food consumption
- Tabulation of clinical pathology data
- Necropsy findings and histopathological findings for each animal, if available.
- Tabulation of any other parameters measured

*Discussion and interpretation of results*

- Particular emphasis should be made to the description of methods used to meet the criteria of this Test Method, e.g. the limit concentration or the particle size.
- The respirability of particles in light of the overall findings should be addressed, especially if the particle-size criteria could not be met.
- The consistency of methods used to determine nominal and actual concentrations, and the relation of actual concentration to nominal concentration should be included in the overall assessment of the study.
- The likely cause of death and predominant mode of action (systemic versus local) should be addressed.
- An explanation should be provided if there was a need to humanely sacrifice animals in pain or showing signs of severe and enduring distress, based on the criteria in the OECD Guidance Document on Humane Endpoints (3).
- The target organ(s) should be identified.
- The NOAEL and LOAEL should be determined.

*LITERATURE:*

- (1) OECD (1981). Subchronic Inhalation Toxicity Testing, Original Test Guideline No 412, Environment Directorate, OECD, Paris.
- (2) OECD (2009). Guidance Document on Acute Inhalation Toxicity Testing, Environmental Health and Safety Monograph Series on Testing and Assessment No 39, ENV/JM/MONO(2009)28, OECD, Paris.
- (3) OECD (2000). Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation, Environmental Health and Safety Monograph Series on Testing and Assessment No 19, ENV/JM/MONO(2000)7, OECD, Paris.
- (4) Whalan JE and Redden JC (1994). Interim Policy for Particle Size and Limit Concentration Issues in Inhalation Toxicity Studies. Office of Pesticide Programs, United States Environmental Protection Agency.
- (5) Dungworth DL, Tyler WS, Plopper CE (1985). Morphological Methods for Gross and Microscopic Pathology (Chapter 9) in Toxicology of Inhaled Material, Witschi, H.P. and Brain, J.D. (eds), Springer Verlag Heidelberg, pp. 229-258.
- (6) Young JT (1981). Histopathological examination of the rat nasal cavity. *Fundam. Appl. Toxicol.* 1: 309-312.
- (7) Harkema JR (1990). Comparative pathology of the nasal mucosa in laboratory animals exposed to inhaled irritants. *Environ. Health Perspect.* 85: 231-238.

**▼M4**

- (8) Woutersen RA, Garderen-Hoetmer A, van Slootweg PJ, Feron VJ (1994). Upper respiratory tract carcinogenesis in experimental animals and in humans. In: Waalkes MP and Ward JM (eds) Carcinogenesis. Target Organ Toxicology Series, Raven Press, New York, 215-263.
- (9) Mery S, Gross EA, Joyner DR, Godo M, Morgan KT (1994). Nasal diagrams: A tool for recording the distribution of nasal lesions in rats and mice. *Toxicol. Pathol.* 22: 353-372.
- (10) Kuper CF, Koornstra PJ, Hamelers DMH, Biewenga J, Spit BJ, Duijvestijn AM, Breda Vriesman van PJC, Sminia T (1992). The role of nasopharyngeal lymphoid tissue. *Immunol. Today* 13: 219-224.
- (11) Kuper CF, Arts JHE, Feron VJ (2003). Toxicity to nasal-associated lymphoid tissue. *Toxicol. Lett.* 140-141: 281-285.
- (12) Lewis DJ (1981). Mitotic Indices of Rat Laryngeal Epithelia. *Journal of Anatomy* 132(3): 419-428.
- (13) Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006 (OJ L 353, 31.12.2008, p. 1).

▼ M4

*Appendix 1*

DEFINITION

**Test chemical:** Any substance or mixture tested using this Test Method.

**▼B****B.9. REPEATED DOSE (28 DAYS) TOXICITY (DERMAL)****1. METHOD****1.1. INTRODUCTION**

See General introduction Part B (A).

**1.2. DEFINITIONS**

See General introduction Part B (B).

**1.3. REFERENCE SUBSTANCES**

None.

**1.4. PRINCIPLE OF THE TEST METHOD**

The test substance is applied daily to the skin in graduated doses to several groups of experimental animals, one dose per group, for a period of 28 days. During the period of application, the animals are observed daily to detect signs of toxicity. Animals, which die during the test, are necropsied and at the conclusion of the test surviving animals are necropsied.

**1.5. QUALITY CRITERIA**

None.

**1.6. DESCRIPTION OF THE TEST METHOD****1.6.1. Preparations**

The animals are kept under the experimental housing and feeding conditions for at least five days prior to the test. Before the test, healthy young animals are randomised and assigned to the treatment and control groups. Shortly before testing, fur is clipped from the dorsal area of the trunk of the test animals. Shaving may be employed but it should be carried out approximately 24 hours before the test. Repeat clipping or shaving is usually needed at approximately weekly intervals. When clipping or shaving the fur, care must be taken to avoid abrading the skin. Not less than 10 % of the body surface area should be clear for the application of the test substance. The weight of the animal should be taken into account when deciding on the area to be cleared and on the dimensions of the covering. When testing solids, which may be pulverised if appropriate, the test substance should be moistened sufficiently with water or, where necessary, a suitable vehicle to ensure good contact with the skin. Liquid test substances are generally used undiluted. Daily application on a five to seven-day per week basis is used.

**1.6.2. Test conditions****1.6.2.1. Experimental animals**

The adult rat, rabbit or guinea-pig may be used. Other species may be used but their use would require justification.

**▼B**

At the commencement of the study, the range of weight variation in the animals used should not exceed  $\pm 20\%$  of the appropriate mean value.

1.6.2.2. *Number and sex*

At least 10 animals (five female and five male) with healthy skin should be used at each dose level. The females should be nulliparous and non-pregnant. If interim sacrifices are planned, the numbers should be increased by the number of animals scheduled to be sacrificed before the completion of the study. In addition, a satellite group of 10 animals (five animals per sex) may be treated with the high dose level for 28 days and observed for reversibility, persistence, or delayed occurrence of toxic effects for 14 days post-treatment. A satellite group of 10 control animals (five animals per sex) is also used.

1.6.2.3. *Dose levels*

At least three dose levels are required with a control or a vehicle control if a vehicle is used. The exposure period should be at least six hours per day. The application of the test substance should be made at similar times each day, and adjusted at intervals (weekly or bi-weekly) to maintain a constant dose level in terms of animal body-weight. Except for treatment with the test substance, animals in the control group should be handled in an identical manner to the test group subjects. Where a vehicle is used to facilitate dosing, the vehicle control group should be dosed in the same way as the treated groups, and receive the same amount as that received by the highest dose level group. The highest dose level should result in toxic effects but produce no, or few, fatalities. The lowest dose level should not produce any evidence or toxicity. Where there is a usable estimation of human exposure, the lowest level should exceed this. Ideally, the intermediate dose level should produce minimal observable toxic effects. If more than one intermediate dose is used the dose levels should be spaced to produce a gradation of toxic effects. In the low and intermediate groups and in the controls, the incidence of fatalities should be low in order to permit a meaningful evaluation of the results.

If application of the test substance produces severe skin irritation, the concentrations should be reduced and this may result in a reduction in, or absence of, other toxic effects at the high dose level. Moreover if the skin has been badly damaged it may be necessary to terminate the study and undertake a new study at lower concentrations.

1.6.2.4. *Limit test*

If a preliminary study at a dose level of 1 000 mg/kg, or a higher dose level related to possible human exposure where this is known, produces no toxic effects, further testing may not be considered necessary.

1.6.2.5. *Observation period*

The experimental animals should be observed daily for signs of toxicity. The time of death and the time at which signs of toxicity appear and disappear should be recorded.

**▼B****1.6.3. Procedure**

Animals should be caged individually. The animals are treated with the test substance, ideally on seven days per week, for a period of 28 days. Animals in any satellite groups scheduled for follow-up observations should be kept for a further 14 days without treatment to detect recovery from or persistence of toxic effects. Exposure time should be at least six hours per day.

The test substance should be applied uniformly over an area, which is approximately 10 % of the total body surface area. With highly toxic substances, the surface area covered may be less but as much of the area as possible should be covered with as thin and uniform a layer as possible.

During exposure the test substance is held in contact with the skin with porous gauze dressing and non-irritating tape. The test site should be further covered in a suitable manner to retain the gauze dressing and test substance and ensure that the animals cannot ingest the test substance. Restraint may be used to prevent the ingestion of the test substance but complete immobilisation is not a recommended method. As an alternative a 'collar protective device' may be used.

At the end of the exposure period, residual test substance should be removed, where practicable, using water or some other appropriate method of cleansing the skin.

All the animals should be observed daily and signs of toxicity recorded including the time of onset, their degree and duration. Observations should include changes in skin and fur, eyes and mucous membranes as well as respiratory, circulatory, autonomic and central nervous systems, somatomotor activity and behaviour pattern. Measurements should be made weekly of the animals' weight. It is also recommended that food consumption is measured weekly. Regular observation of the animals is necessary to ensure that animals are not lost from the study due to causes such as cannibalism, autolysis of tissues or misplacement. At the end of the study period, all survivors in the non-satellite treatment groups are necropsied. Moribund animals and animals in severe distress or pain should be removed when noticed, humanely killed and necropsied.

The following examinations shall be made at the end of the test on all animals including the controls:

- (1) haematology, including at least haematocrit, haemoglobin concentration, erythrocyte count, total and differential leucocyte count, and a measure of clotting potential;
- (2) clinical blood biochemistry including at least one parameter of liver and kidney function: serum alanine aminotransferase (formerly known as glutamic pyruvic transaminase), serum aspartate aminotransferase (formerly known as glutamic oxaloacetic transaminase), urea nitrogen, albumin, blood creatinine, total bilirubin and total serum protein;

Other determinations which may be necessary for an adequate toxicological evaluation include calcium, phosphorus, chloride, sodium, potassium, fasting glucose, analysis of lipids, hormones, acid/base balance, methaemoglobin and cholinesterase activity.

**▼B**

Additional clinical biochemistry may be employed, where necessary, to extend the investigation of observed effects.

**1.6.4. Gross necropsy**

All animals in the study should be subjected to a full gross necropsy. At least the liver, kidneys, adrenals, and testes should be weighed wet as soon as possible after dissection, to avoid drying. Organs and tissues, i.e. normal and treated skin, liver, kidney, spleen, testes, adrenals, heart, and target organs (that is those organs showing gross lesions or changes in size) should be preserved in a suitable medium for possible future histopathological examination.

**1.6.5. Histopathological examination**

In the high dose group and in the control group, histological examination should be performed on the preserved organs and tissues. Organs and tissues showing defects attributable to the test substance at the highest dosage level should be examined in all lower-dosage groups. Animals in the satellite group should be examined histologically with particular emphasis on those organs and tissues identified as showing effects in the other treated groups.

**2. DATA**

Data should be summarised in tabular form, showing for each test group the number of animals at the start of the test and the number of animals displaying each type of lesion.

All observed results should be evaluated by an appropriate statistical method. Any recognised statistical method may be used.

**3. REPORTING****3.1. TEST REPORT**

The test report shall, if possible, include the following information:

- animal data (species, strain, source, environmental conditions, diet, etc.),
- test conditions (including the type of dressing: occlusive or not-occlusive),
- dose levels (including vehicle, if used) and concentrations,
- no-effect level, where possible,
- toxic response data by sex and dose,
- time of death during the study or whether animals survived to termination,
- toxic or other effects,
- the time of observation of each abnormal sign and its subsequent course,
- food and body-weight data,
- haematological tests employed and results,

**▼B**

- clinical biochemistry tests employed and results,
- necropsy findings,
- a detailed description of all histopathological findings,
- statistical treatment of results where possible,
- discussion of the results,
- interpretation of the results.

**3.2. EVALUATION AND INTERPRETATION**

See General introduction Part B (D).

**4. REFERENCES**

See General introduction Part B (E).



**▼B****B.10. MUTAGENICITY — *IN VITRO* MAMMALIAN CHROMOSOME ABERRATION TEST****1. METHOD**

This method is a replicate of the OECD TG 473, *In Vitro* Mammalian Chromosome Aberration Test (1997).

**1.1. INTRODUCTION**

The purpose of the *in vitro* chromosomal aberration test is to identify agents that cause structural chromosome aberrations in cultured mammalian cells (1)(2)(3). Structural aberrations may be of two types, chromosome or chromatid. With the majority of chemical mutagens, induced aberrations are of the chromatid type, but chromosome-type aberrations also occur. An increase in polyploidy may indicate that a chemical has the potential to induce numerical aberrations. However, this method is not designed to measure numerical aberrations and is not routinely used for that purpose. Chromosome mutations and related events are the cause of many human genetic diseases and there is substantial evidence that chromosome mutations and related events causing alterations in oncogenes and tumour-suppressor genes of somatic cells are involved in cancer induction in humans and experimental animals.

The *in vitro* chromosome aberration test may employ cultures of established cell lines, cell strains or primary cell cultures. The cells used are selected on the basis of growth ability in culture, stability of the karyotype, chromosome number, chromosome diversity and spontaneous frequency of chromosome aberrations.

Tests conducted *in vitro* generally require the use of an exogenous source of metabolic activation. This metabolic activation system cannot mimic entirely the mammalian *in vivo* conditions. Care should be taken to avoid conditions which would lead to positive results which do not reflect intrinsic mutagenicity and may arise from changes in pH, osmolality or high levels of cytotoxicity (4)(5).

This test is used to screen for possible mammalian mutagens and carcinogens. Many compounds that are positive in this test are mammalian carcinogens; however, there is not a perfect correlation between this test and carcinogenicity. Correlation is dependent on chemical class and there is increasing evidence that there are carcinogens that are not detected by this test because they appear to act through mechanisms other than direct DNA damage.

See also General introduction Part B.

**1.2. DEFINITIONS**

**Chromatid-type aberration:** structural chromosome damage expressed as breakage of single chromatids or breakage and reunion between chromatids.

**Chromosome-type aberration:** structural chromosome damage expressed as breakage, or breakage and reunion, of both chromatids at an identical site.

**▼ B**

**Endoreduplication:** a process in which after an S period of DNA replication, the nucleus does not go into mitosis but starts another S period. The result is chromosomes with four, eight, 16, ...chromatids.

**Gap:** an achromatic lesion smaller than the width of one chromatid, and with minimum misalignment of the chromatids).

**Mitotic index:** the ratio of cells in metaphase divided by the total number of cells observed in a population of cells; an indication of the degree of proliferation of that population.

**Numerical aberration:** a change in the number of chromosomes from the normal number characteristic of the cells utilised.

**Polyploidy:** a multiple of the haploid chromosome number ( $n$ ) other than the diploid number (i.e.  $3n$ ,  $4n$  and so on).

**Structural aberration:** a change in chromosome structure detectable by microscopic examination of the metaphase stage of cell division, observed as deletions and fragments, intrachanges or interchanges.

### 1.3. PRINCIPLE OF THE TEST METHOD

Cell cultures are exposed to the test substance both with and without metabolic activation. At predetermined intervals after exposure of cell cultures to the test substance, they are treated with a metaphase-arresting substance (e.g. Colcemid® or colchicine), harvested, stained and metaphase cells are analysed microscopically for the presence of chromosome aberrations.

### 1.4. DESCRIPTION OF THE TEST METHOD

#### 1.4.1. Preparations

##### 1.4.1.1. Cells

A variety of cell lines, strains or primary cell cultures, including human cells, may be used (e.g. Chinese hamster fibroblasts, human or other mammalian peripheral blood lymphocytes).

##### 1.4.1.2. Media and culture conditions

Appropriate culture media and incubation conditions (culture vessels, CO<sub>2</sub> concentration, temperature and humidity) should be used in maintaining cultures. Established cell lines and strains should be checked routinely for stability in the modal chromosome number and the absence of mycoplasma contamination and should not be used if contaminated. The normal cell cycle time for the cells and culture conditions used should be known.

##### 1.4.1.3. Preparation of cultures

Established cell lines and strains: cells are propagated from stock cultures, seeded in culture medium at a density such that the cultures will not reach confluency before the time of harvest, and incubated at 37 °C.

**▼ B**

Lymphocytes: whole blood treated with an anti-coagulant (e.g. heparin) or separated lymphocytes obtained from healthy subjects are added to the culture medium containing a mitogen (e.g. phytohaemagglutinin) and incubated at 37 °C.

**1.4.1.4. *Metabolic activation***

Cells should be exposed to the test substance both in the presence and absence of an appropriate metabolic activation system. The most commonly used system is a cofactor-supplemented post-mitochondrial fraction (S9) prepared from the livers of rodents treated with enzyme-inducing agents such as: Aroclor 1254 (6)(7)(8)(9), or a mixture of phenobarbitone and  $\beta$ -naphthoflavone (10)(11)(12).

The post-mitochondrial fraction is usually used at concentrations in the range from 1-10 % v/v in the final test medium. The condition of a metabolic activation system may depend upon the class of chemical being tested. In some cases it may be appropriate to utilise more than one concentration of post-mitochondrial fraction.

A number of developments, including the construction of genetically engineered cell lines expressing specific activating enzymes, may provide the potential for endogenous activation. The choice of the cell lines used should be scientifically justified (e.g. by the relevance of the cytochrome P450 isoenzyme for the metabolism of the test substance).

**1.4.1.5. *Test substance/Preparation***

Solid test substances should be dissolved or suspended in appropriate solvents or vehicles and diluted if appropriate prior to treatment of the cells. Liquid test substances may be added directly to the test systems and/or diluted prior to treatment. Fresh preparations of the test substance should be employed unless stability data demonstrate the acceptability of storage.

**1.4.2. **Test conditions******1.4.2.1. *Solvent/vehicle***

The solvent/vehicle should not be suspected of chemical reaction with the test substance and should be compatible with the survival of the cells and the S9 activity. If other than well-known solvent/vehicles are used, their inclusion should be supported by data indicating their compatibility. It is recommended that wherever possible, the use of an aqueous solvent/vehicle be considered first. When testing water-unstable substances, the organic solvents used should be free of water. Water can be removed by adding a molecular sieve.

**1.4.2.2. *Exposure concentrations***

Among the criteria to be considered when determining the highest concentration are cytotoxicity, solubility in the test system and changes in pH or osmolality.

Cytotoxicity should be determined with and without metabolic activation in the main experiment using an appropriate indication of cell integrity and growth, such as degree of confluency, viable cell counts, or mitotic index. It may be useful to determine cytotoxicity and solubility in a preliminary experiment.

**▼ B**

At least three analysable concentrations should be used. Where cytotoxicity occurs, these concentrations should cover a range from the maximum to little or no toxicity; this will usually mean that the concentrations should be separated by no more than a factor between 2 and  $\sqrt{10}$ . At the time of harvesting, the highest concentration should show a significant reduction in degree of confluency, cell count or mitotic index, (all greater than 50 %). The mitotic index is only an indirect measure of cytotoxic/cytostatic effects and depends on the time after treatment. However, the mitotic index is acceptable for suspension cultures in which other toxicity measurements may be cumbersome and impractical. Information on cell cycle kinetics, such as average generation time (AGT), could be used as supplementary information. AGT, however, is an overall average that does not always reveal the existence of delayed subpopulations, and even slight increases in average generation time can be associated with very substantial delay in the time of optimal yield of aberrations.

For relatively non-cytotoxic substances, the maximum test concentration should be 5  $\mu\text{l/ml}$ , 5  $\text{mg/ml}$  or 0,01 M, whichever is the lowest.

For relatively insoluble substances that are not toxic at concentrations lower than the insoluble concentration, the highest dose used should be a concentration above the limit of solubility in the final culture medium at the end of the treatment period. In some cases (e.g. when toxicity occurs only at higher than the lowest insoluble concentration) it is advisable to test at more than one concentration with visible precipitation. It may be useful to assess solubility at the beginning and the end of the treatment, as solubility can change during the course of exposure in the test system due to presence of cells, S9, serum etc. Insolubility can be detected by using the unaided eye. The precipitate should not interfere with the scoring.

#### 1.4.2.3. *Negative and positive controls*

Concurrent positive and negative (solvent or vehicle) controls, both with and without metabolic activation, should be included in each experiment. When metabolic activation is used, the positive control chemical should be the one that requires activation to give a mutagenic response.

Positive controls should employ a known clastogen at exposure levels expected to give a reproducible and detectable increase over background, which demonstrates the sensitivity of the test system.

Positive control concentrations should be chosen so that the effects are clear but do not immediately reveal the identity of the coded slides to the reader. Examples of positive control substances include:

Metabolic Activation condition	Substance	CAS No	EINECS No
Absence of exogenous metabolic Activation	Methyl methanesulphonate	66-27-3	200-625-0
	Ethyl methanesulphonate	62-50-0	200-536-7

**▼B**

Metabolic Activation condition	Substance	CAS No	EINECS No
	Ethyl nitrosourea	759-73-9	212-072-2
	Mitomycin C	50-07-7	200-008-6
	4-Nitroquinoline- <i>N</i> -oxide	56-57-5	200-281-1
Presence of exogenous metabolic Activation	Benzo[ <i>a</i> ]pyrene	50-32-8	200-028-5
	Cyclophosphamide Cyclophosphamide monohydrate	50-18-0 6055-19-2	200-015-4

Other appropriate positive control substances may be used. The use of chemical class-related positive control chemicals should be considered, when available.

Negative controls, consisting of solvent or vehicle alone in the treatment medium, and treated in the same way as the treatment cultures, should be included for every harvest time. In addition, untreated controls should also be used unless there are historical control data demonstrating that no deleterious or mutagenic effects are induced by the chosen solvent.

#### 1.4.3. Procedure

##### 1.4.3.1. *Treatment with the test substance*

Proliferating cells are treated with the test substance in the presence and absence of a metabolic activation system. Treatment of lymphocytes should commence at about 48 hours after mitogenic stimulation.

1.4.3.2. Duplicate cultures should normally be used at each concentration, and are strongly recommended for negative/solvent control cultures. Where minimal variation between duplicate cultures can be demonstrated (13)(14), from historical data, it may be acceptable for single cultures to be used at each concentration.

Gaseous or volatile substances should be tested by appropriate methods, such as in sealed culture vessels (15)(16).

##### 1.4.3.3. *Culture harvest time*

In the first experiment, cells should be exposed to the test substance, both with and without metabolic activation, for three to six hours, and sampled at a time equivalent to about 1,5 normal cell cycle length after the beginning of treatment (12). If this protocol gives negative results both with and without activation, an additional experiment without activation should be done, with continuous treatment until sampling at a time equivalent to about 1,5 normal cell cycle lengths. Certain chemicals may be more readily detected by treatment/sampling times longer than 1,5 cycle lengths. Negative results with metabolic activation need to be confirmed on a case-by-case basis. In those cases where confirmation of negative results is not considered necessary, justification should be provided.

**▼B**1.4.3.4. *Chromosome preparation*

Cell cultures are treated with Colcemid® or colchicine usually for one to three hours prior to harvesting. Each cell culture is harvested and processed separately for the preparation of chromosomes. Chromosome preparation involves hypotonic treatment of the cells, fixation and staining.

1.4.3.5. *Analysis*

All slides, including those of positive and negative controls, should be independently coded before microscopic analysis. Since fixation procedures often result in the breakage of a proportion of metaphase cells with loss of chromosomes, the cells scored should therefore contain a number of centromeres equal to the modal number  $\pm 2$  for all cell types. At least 200 well spread metaphases should be scored per concentration and control, equally divided amongst the duplicates, if applicable. This number can be reduced when high number of aberrations is observed.

Though the purpose of the test is to detect structural chromosome aberrations, it is important to record polyploidy and endoreduplication when these events are seen.

**2. DATA**

## 2.1. TREATMENT OF RESULTS

The experimental unit is the cell, and therefore the percentage of cells with structural chromosome aberration(s) should be evaluated. Different types of structural chromosome aberrations should be listed with their numbers and frequencies for experimental and control cultures. Gaps are recorded separately and reported but generally not included in the total aberration frequency.

Concurrent measures of cytotoxicity for all treated and negative control cultures in the main aberration experiments should also be recorded.

Individual culture data should be provided. Additionally, all data should be summarised in tabular form.

There is no requirement for verification of a clear positive response. Equivocal results should be clarified by further testing preferably using modification of experimental conditions. The need to confirm negative results has been discussed in 1.4.3.3. Modification of study parameters to extend the range of conditions assessed should be considered in follow-up experiments. Study parameters that might be modified include the concentration spacing and the metabolic activation conditions.

## 2.2. EVALUATION AND INTERPRETATION OF RESULTS

There are several criteria for determining a positive result, such as a concentration-related increase or a reproducible increase in the number of cells with chromosome aberrations. Biological relevance of the results should be considered first. Statistical methods may be used as an aid in evaluating the test results (3)(13). Statistical significance should not be the only determining factor for a positive response.

**▼B**

An increase in the number of polyploid cells may indicate that the test substance has the potential to inhibit mitotic processes and to induce numerical chromosome aberrations. An increase in the number of cells with endoreduplicated chromosomes may indicate that the test substance has the potential to inhibit cell cycle progression (17)(18).

A test substance for which the results do not meet the above criteria is considered non-mutagenic in this system.

Although most experiments will give clearly positive or negative results, in rare cases the data set will preclude making a definite judgement about the activity of the test substance. Results may remain equivocal or questionable regardless of the number of times the experiment is repeated.

Positive results from the *in vitro* chromosome aberration test indicate that the test substance induces structural chromosome aberrations in cultured mammalian somatic cells. Negative results indicate that, under the test conditions, the test substance does not induce chromosome aberrations in cultured mammalian somatic cells.

### 3. **REPORTING**

#### TEST REPORT

The test report must include the following information:

Solvent/Vehicle:

- justification for choice of vehicle,
- solubility and stability of the test substance in solvent/vehicle, if known.

Cells:

- type and source of cells,
- karyotype features and suitability of the cell type used,
- absence of mycoplasma, if applicable,
- information on cell cycle length,
- sex of blood donors, whole blood or separated lymphocytes, mitogen used,
- number of passages, if applicable,
- methods for maintenance of cell culture, if applicable,
- modal number of chromosomes.

Test conditions:

- identity of metaphase arresting substance, its concentration and duration of cell exposure,
- rationale for selection of concentrations and number of cultures including, e.g. cytotoxicity data and solubility limitations, if available,
- composition of media, CO<sub>2</sub> concentration if applicable,

**▼ B**

- concentration of test substance,
- volume of vehicle and test substance added,
- incubation temperature,
- incubation time,
- duration of treatment,
- cell density at seeding, if appropriate,
- type and composition of metabolic activation system, including acceptability criteria,
- positive and negative controls,
- methods of slide preparation,
- criteria for scoring aberrations,
- number of metaphases analysed,
- methods for the measurements of toxicity,
- criteria for considering studies as positive, negative or equivocal.

## Results:

- signs of toxicity, e.g. degree of confluency, cell cycle data, cell counts, mitotic index,
- signs of precipitation,
- data on pH and osmolality of the treatment medium, if determined,
- definition for aberrations, including gaps,
- number of cells with chromosome aberrations and type of chromosome aberrations given separately for each treated and control culture,
- changes in ploidy if seen,
- dose-response relationship, where possible,
- statistical analyses, if any,
- concurrent negative (solvent/vehicle) and positive control data,
- historical negative (solvent/vehicle) and positive control data, with ranges, means and standard deviations.

Discussion of results.

Conclusions.

#### 4. REFERENCES

- (1) Evans, H. J., (1976) Cytological Methods for Detecting Chemical Mutagens. In: Chemical mutagens, Principles and Methods for their Detection, Vol. 4, Hollaender, A. (ed) Plenum Press, New York and London, p. 1-29.
- (2) Ishidate, M. Jr. and Sofuni, T., (1985). The *In Vitro* Chromosomal Aberration Test Using Chinese Hamster Lung (CHL) Fibroblast Cells in Culture. In: Progress in Mutation Research, Vol. 5, Ashby, J. et al., (Eds) Elsevier Science Publishers, Amsterdam-New York-Oxford, p. 427-432.



**▼B**

- (3) Galloway, S.M., Armstrong, M.J., Reuben, C., Colman, S., Brown, B., Cannon, C., Bloom, A.D., Nakamura, F., Ahmed, M., Duk, S., Rimpo, J., Margolin, G.H., Resnick, MA, Anderson, G. and Zeiger, E., (1978). Chromosome aberration and sister chromatic exchanges in Chinese hamster ovary cells: Evaluation of 108 chemicals. *Environ. Molec. Mutagen* 10 (suppl. 10), p. 1-175.
- (4) Scott, D., Galloway, S.M., Marshall, R.R., Ishidate, M.Jr., Brusick, D., Ashby, J. and Myhr, B.C., (1991) Genotoxicity under Extreme Culture Conditions. A report from ICPEMC Task Group 9. *Mutation Res*, 257, p. 147-204.
- (5) Morita, T., Nagaki, T., Fukuda, I. and Okumura, K., (1992). Clastogenicity of low pH to Various Cultured Mammalian Cells. *Mutation Res.*, 268, p. 297-305.
- (6) Ames, B.N., McCann, J. and Yamasaki, E., (1975). Methods for Detecting Carcinogens and Mutagens with the Salmonella/-Mammalian Microsome Mutagenicity Test. *Mutation Res.*, 31, p. 347-364.
- (7) Maron, D.M. and Ames, B.N., (1983) Revised Methods for the Salmonella Mutagenicity Test. *Mutation Res.*, 113, p. 173-215.
- (8) Natarajan, A.T., Tate, A.D., van Buul, P.P.W., Meijers, M. and de Vogel, N., (1976) Cytogenetic Effects of Mutagen/-Carcinogens after Activation in a Microsomal System in Vitro. I. Induction of Chromosome Aberrations and Sister Chromatid Exchange by Diethylnitrosamine (DEN) and Dimethylnitrosamine (DMN) in CHO Cells in the Presence of Rat-Liver Microsomes. *Mutation Res.*, 37, p. 83-90.
- (9) Matsuoka, A., Hayashi, M. and Ishidate, M. Jr., (1979) Chromosomal Aberration Tests on 29 Chemicals Combined with S9 Mix *In Vitro*. *Mutation Res.*, 66, p. 277-290.
- (10) Elliot, B.M., Combes, R.D., Elcombe, C.R., Gatehouse, D.G., Gibson, G.G., Mackay, J.M. and Wolf, R.C., (1992). Report of UK Environmental Mutagen Society Working Party. Alternatives to Aroclor 1254-induced S9 in *In Vitro* Genotoxicity Assays. *Mutagenesis*, 7, p. 175-177.
- (11) Matsushima, T., Sawamura, M., Hara, K. and Sugimura, T., (1976) A Safe Substitute for Polychlorinated Biphenyls as an Inducer of Metabolic Activation Systems. In: de Serres, F.J., Fouts, J.R., Bend, J.R. and Philpot, R.M. (eds) *In Vitro* Metabolic Activation in Mutagenesis Testing, Elsevier, North-Holland, p. 85-88.
- (12) Galloway, S.M., Aardema, M.J., Ishidate, M.Jr., Ivett, J.L., Kirkland, D.J., Morita, T., Mosesso, P., Sofuni, T., (1994) Report from Working Group on in *In Vitro* Tests for Chromosomal Aberrations. *Mutation Res.*, 312, p. 241-261.
- (13) Richardson, C., Williams, D.A., Allen, J.A., Amphlett, G., Chanter, D.O. and Phillips, B., (1989) Analysis of Data from *In Vitro* Cytogenetic Assays. In: Statistical Evaluation of Mutagenicity Test Data. Kirkland, D.J., (ed) Cambridge University Press, Cambridge, p. 141-154.
- (14) Soper, K.A. and Galloway, S.M., (1994) Replicate Flasks are not Necessary for *In Vitro* Chromosome Aberration Assays in CHO Cells. *Mutation Res.*, 312, p. 139-149.
- (15) Krahn, D.F., Barsky, F.C. and McCooey, K.T., (1982) CHO/HGPRT Mutation Assay: Evaluation of Gases and Volatile Liquids. In: Tice, R.R., Costa, D.L., Schaich, K.M. (eds). *Genotoxic Effects of Airborne Agents*. New York, Plenum, p. 91-103.

**▼B**

- (16) Zamora, P.O., Benson, J.M., Li, A.P. and Brooks, A.L., (1983) Evaluation of an Exposure System Using Cells Grown on Collagen Gels for Detecting Highly Volatile Mutagens in the CHO/HGPRT Mutation Assay. *Environmental Mutagenesis*, 5, p. 795-801.
- (17) Locke-Huhle, C., (1983) Endoreduplication in Chinese hamster cells during alpha-radiation induced G2 arrest. *Mutation Res.*, 119, p. 403-413.
- (18) Huang, Y., Change, C. and Trosko, J.E., (1983) Aphidicolin-induced endoreduplication in Chinese hamster cells. *Cancer Res.*, 43, p. 1362-1364.

**▼B****B.11. MUTAGENICITY — *IN VIVO* MAMMALIAN BONE MARROW CHROMOSOME ABERRATION TEST****1. METHOD**

This method is a replicate of the OECD TG 475, Mammalian Bone Marrow Chromosome Aberration Test (1997).

**1.1. INTRODUCTION**

The mammalian *in vivo* chromosome aberration test is used for the detection of structural chromosome aberrations induced by the test substance to the bone marrow cells of animals, usually rodents (1)-(2)(3)(4). Structural chromosome aberrations may be of two types, chromosome or chromatid. An increase in polyploidy may indicate that a chemical has the potential to induce numerical aberrations. With the majority of chemical mutagens, induced aberrations are of the chromatid-type, but chromosome-type aberrations also occur. Chromosome mutations and related events are the cause of many human genetic diseases and there is substantial evidence that chromosome mutations and related events causing alterations in oncogenes and tumour-suppressor genes are involved in cancer in humans and experimental systems.

Rodents are routinely used in this test. Bone marrow is the target tissue in this test, since it is a highly vascularised tissue, and it contains a population of rapidly cycling cells that can be readily isolated and processed. Other species and target tissues are not the subject of this method.

This chromosome aberration test is especially relevant to assessing mutagenic hazard in that it allows consideration of factors of *in vivo* metabolism, pharmacokinetics and DNA-repair processes although these may vary among species and among tissues. An *in vivo* test is also useful for further investigation of a mutagenic effect detected by *in vitro* test.

If there is evidence that the test substance, or a reactive metabolite, will not reach the target tissue, it is not appropriate to use this test.

See also General introduction Part B.

**1.2. DEFINITIONS**

**Chromatid-type aberration:** structural chromosome damage expressed as breakage of single chromatids or breakage and reunion between chromatids.

**Chromosome-type aberration:** structural chromosome damage expressed as breakage, or breakage and reunion, of both chromatids at an identical site.

**Endoreduplication:** a process in which after an S period of DNA replication, the nucleus does not go into mitosis but starts another S period. The result is chromosomes with four, eight, 16, ...chromatids.

**Gap:** an achromatic lesion smaller than the width of one chromatid, and with minimum misalignment of the chromatid(s).

**Numerical aberration:** a change in the number of chromosomes from the normal number characteristic of the cells utilised.

**▼ B**

**Polyploidy:** a multiple of the haploid chromosome number ( $n$ ) other than the diploid number (i.e.  $3n$ ,  $4n$  and so on).

**Structural aberration:** a change in chromosome structure detectable by microscopic examination of the metaphase stage of cell division, observed as deletions and fragments, intrachanges or interchanges.

### 1.3. PRINCIPLE OF THE TEST METHOD

Animals are exposed to the test substance by an appropriate route of exposure and are sacrificed at appropriate times after treatment. Prior to sacrifice, animals are treated with a metaphase arresting agent (e.g. Colcemid® or colchicine). Chromosome preparations are then made from the bone marrow cells and stained, and metaphase cells are analysed for chromosome aberrations.

### 1.4. DESCRIPTION OF THE TEST METHOD

#### 1.4.1. Preparations

##### 1.4.1.1. *Selection of animal species*

Rats, mice and Chinese hamsters are commonly used, although any appropriate mammalian species may be used. Commonly used laboratory strains of young healthy adult animals should be employed. At the commencement of the study, the weight variation of animals should be minimal and not exceed  $\pm 20\%$  of the mean weight of each sex.

##### 1.4.1.2. *Housing and feeding conditions*

General conditions referred in the General introduction to Part B are applied although the aim for humidity should be 50-60 %.

##### 1.4.1.3. *Preparation of the animals*

Healthy young adult animals are randomly assigned to the control and treatment groups. Cages should be arranged in such a way that possible effects due to cage placement are minimised. The animals are identified uniquely. The animals are acclimated to the laboratory conditions for at least five days.

##### 1.4.1.4. *Preparation of doses*

Solid test substances should be dissolved or suspended in appropriate solvents or vehicles and diluted, if appropriate, prior to dosing of the animals. Liquid test substances may be dosed directly or diluted prior to dosing. Fresh preparations of the test substance should be employed unless stability data demonstrate the acceptability of storage.

#### 1.4.2. Test conditions

##### 1.4.2.1. *Solvent/Vehicle*

The solvent/vehicle should not produce toxic effects at the dose levels used, and should not be suspected of chemical reaction with the test substance. If other than well-known solvents/vehicles are used, their inclusion should be supported with data indicating their compatibility. It is recommended that wherever possible, the use of an aqueous solvent/vehicle should be considered first.

**▼B**1.4.2.2. *Controls*

Concurrent positive and negative (solvent/vehicle) controls should be included for each sex in each test. Except for treatment with the test substance, animals in the control groups should be handled in an identical manner to the animals in the treated groups.

Positive controls should produce structural aberrations *in vivo* at exposure levels expected to give a detectable increase over background. Positive control doses should be chosen so that the effects are clear but do not immediately reveal the identity of the coded slides to the reader. It is acceptable that the positive control be administered by a route different from the test substance and sampled at only a single time. The use of chemical class related positive control chemicals may be considered, when available. Examples of positive control substances include:

Substance	CAS No	EINECS No
Ethyl methanesulphonate	62-50-0	200-536-7
Ethyl nitrosourea	759-73-9	212-072-2
Mitomycin C	50-07-7	200-008-6
Cyclophosphamide	50-18-0	200-015-4
Cyclophosphamide monohydrate	6055-19-2	
Triethylenemelamine	51-18-3	200-083-5

Negative controls, treated with solvent or vehicle alone, and otherwise treated in the same way as the treatment groups, should be included for every sampling time, unless acceptable inter-animal variability and frequencies of cells with chromosome aberrations are available from historical control data. If single sampling is applied for negative controls, the most appropriate time is the first sampling time. In addition, untreated controls should also be used unless there are historical or published control data demonstrating that no deleterious or mutagenic effects are induced by the chosen solvent/-vehicle.

## 1.5. PROCEDURE

1.5.1. **Number and sex of animals**

Each treated and control group must include at least five analysable animals per sex. If at the time of the study there are data available from studies in the same species and using the same route of exposure that demonstrate that there are no substantial differences in toxicity between sexes, then testing in a single sex will be sufficient. Where human exposure to chemicals may be sex-specific, as for example with some pharmaceutical agents, the test should be performed with animals of the appropriate sex.

1.5.2. **Treatment schedule**

Test substances are preferably administered as a single treatment. Test substances may also be administered as a split dose, i.e. two treatments on the same day separated by no more than a few hours, to facilitate administering a large volume of material. Other dose regimens should be scientifically justified.

**▼B**

Samples should be taken at two separate times following treatment on one day. For rodents, the first sampling interval is 1,5 normal cell cycle length (the latter being normally 12-18 hr) following treatment. Since the time required for uptake and metabolism of the test substance as well as its effect on cell cycle kinetics can affect the optimum time for chromosome aberration detection, a later sample collection 24 hr after the first sample time is recommended. If dose regimens of more than one day are used, one sampling time at 1,5 normal cell cycle lengths after the final treatment should be used.

Prior to sacrifice, animals are injected intraperitoneally with an appropriate dose of a metaphase arresting agent (e.g. Colcemid® or colchicine). Animals are sampled at an appropriate interval thereafter. For mice this interval is approximately three to five hours; for Chinese hamsters this interval is approximately four to five hours. Cells are harvested from the bone marrow and analysed for chromosome aberrations.

**1.5.3. Dose levels**

If a range finding study is performed because there are no suitable data available, it should be performed in the same laboratory, using the same species, strain, sex and treatment regimen to be used in the main study (5). If there is toxicity, three dose levels are used for the first sampling time. These dose levels should cover a range from the maximum to little or no toxicity. At the later sampling time only the highest dose needs to be used. The highest dose is defined as the dose producing signs of toxicity such that higher dose levels, based on the same dosing regimen, would be expected to produce lethality. Substances with specific biological activities at low non-toxic doses (such as hormones and mitogens) may be exceptions to the dose-setting criteria and should be evaluated on a case-by-case basis. The highest dose may also be defined as a dose that produces some indication of toxicity in the bone marrow (e.g. greater than 50 % reduction in mitotic index).

**1.5.4. Limit test**

If a test at one dose level of at least 2 000 mg/kg body weight using a single treatment, or as two treatments on the same day, produces no observable toxic effects, and if genotoxicity would not be expected based on data from structurally related substances, then a full study using three dose levels may not be considered necessary. For studies of a longer duration, the limit dose is 2 000 mg/kg/body weight/day for treatment up to 14 days, and 1 000 mg/kg/body weight/day for treatment longer than 14 days. Expected human exposure may indicate the need for a higher dose level to be used in the limit test.

**1.5.5. Administration of doses**

The test substance is usually administered by gavage using a stomach tube or a suitable intubation cannula, or by intraperitoneal injection. Other routes of exposure may be acceptable where they can be justified. The maximum volume of liquid that can be administered by gavage or injection at one time depends on the size of the test animal. The volume should not exceed 2 ml/100 g body weight. The use of volumes higher than these must be justified. Except for irritating or corrosive substances which will normally reveal exacerbated effects with higher concentrations, variability in test volume should be minimised by adjusting the concentration to ensure a constant volume at all dose levels.

**▼B****1.5.6. Chromosome preparation**

Immediately after sacrifice, bone marrow is obtained, exposed to hypotonic solution and fixed. The cells are then spread on slides and stained.

**1.5.7. Analysis**

The mitotic index should be determined as a measure of cytotoxicity in at least 1 000 cells per animal for all treated animals (including positive controls) and untreated negative control animals.

At least 100 cells should be analysed for each animal. This number could be reduced when high numbers of aberrations are observed. All slides, including those of positive and negative controls, should be independently coded before microscopic analysis. Since slide preparation procedures often result in the breakage of a proportion of metaphases with loss of chromosomes, the cells scored should therefore contain a number of centromeres equal to the number  $2n \pm 2$ .

**2. DATA****2.1. TREATMENT OF RESULTS**

Individual animal data should be presented in tabular form. The experimental unit is the animal. For each animal the number of cells scored, the number of aberrations per cell and the percentage of cells with structural chromosome aberration(s) should be evaluated. Different types of structural chromosome aberrations should be listed with their numbers and frequencies for treated and control groups. Gaps are recorded separately and reported but generally not included in the total aberration frequency. If there is no evidence for a difference in response between the sexes, the data from both sexes may be combined for statistical analysis.

**2.2. EVALUATION AND INTERPRETATION OF RESULTS**

There are several criteria for determining a positive result, such as a dose-related increase in the relative number of cells with chromosome aberrations or a clear increase in the number of cells with aberrations in a single dose group at a single sampling time. Biological relevance of the results should be considered first. Statistical methods may be used as an aid in evaluating the test results (6). Statistical significance should not be the only determining factor for a positive response. Equivocal results should be clarified by further testing preferably using a modification of experimental conditions.

An increase in polyploidy may indicate that the test substance has the potential to induce numerical chromosome aberrations. An increase in endoreduplication may indicate that the test substance has the potential to inhibit cell cycle progression (7)(8).

A test substance for which the results do not meet the above criteria is considered non-mutagenic in this test.

**▼B**

Although most experiments will give clearly positive or negative results, in rare cases the data set will preclude making a definite judgement about the activity of the test substance. Results may remain equivocal or questionable regardless of the number of experiments performed.

Positive results from the *in vivo* chromosome aberration test indicate that a substance induces chromosome aberrations in the bone marrow of the species tested. Negative results indicate that, under the test conditions, the test substance does not induce chromosome aberrations in the bone marrow of the species tested.

The likelihood that the test substance or its metabolites reach the general circulation or specifically the target tissue (e.g. systemic toxicity) should be discussed.

**3. REPORTING****3.1. TEST REPORT**

The test report must include the following information:

Solvent/Vehicle:

- justification for choice of vehicle,
- solubility and stability of the test substance in solvent/vehicle, if known,

Test animals:

- species/strain used,
- number, age and sex of animals,
- source, housing conditions, diet, etc.,
- individual weight of the animals at the start of the test, including body weight range, mean and standard deviation for each group,

Test conditions:

- positive and negative (vehicle/solvent) controls,
- data from range-finding study, if conducted,
- rationale for dose level selection,
- details of test substance preparation,
- details of the administration of the test substance,
- rationale for route of administration,
- methods for verifying that the test substance reached the general circulation or target tissue, if applicable,
- conversion from diet/drinking water test substance concentration (ppm) to the actual dose (mg/kg body weight/day), if applicable,
- details of food and water quality,
- detailed description of treatment and sampling schedules,
- methods for measurements of toxicity,



**▼B**

- identity of metaphase arresting substance, its concentration and duration of treatment,
- methods of slide preparation,
- criteria for scoring aberrations,
- number of cells analysed per animal,
- criteria for considering studies as positive, negative or equivocal.

## Results:

- signs of toxicity,
- mitotic index,
- type and number of aberrations, given separately for each animal,
- total number of aberrations per group with means and standard deviations,
- number of cells with aberrations per group with means and standard deviations,
- changes in ploidy, if seen,
- dose-response relationship, where possible,
- statistical analyses, if any,
- concurrent negative control data,
- historical negative control data with ranges, means and standard deviations,
- concurrent positive control data.

## Discussion of the results.

## Conclusions.

4. **REFERENCES**

- (1) Adler, I.D., (1984) Cytogenetic Tests in Mammals. In: Mutagenicity Testing: a Practical Approach. S. Venitt and J.M. Parry (Eds). IRL Press, Oxford, Washington D.C., p. 275-306.
- (2) Preston, R.J., Dean, B.J., Galloway, S., Holden, H., McFee, A.F. and Shelby, M. (1987). Mammalian *In Vivo* Cytogenetic Assays: Analysis of Chromosome Aberrations in Bone Marrow Cells. *Mutation Res.*, 189, p. 157-165.
- (3) Richold, M., Chandley, A., Ashby, J., Gatehouse, D.G., Bootman, J. and Henderson, L., (1990) *In Vivo* Cytogenetic Assays. In: D.J. Kirkland (Ed.) Basic Mutagenicity Tests, UKEMS Recommended Procedures. UKEMS Sub-Committee on Guidelines for Mutagenicity Testing. Report. Part I revised. Cambridge University Press, Cambridge, New York, Port Chester, Melbourne, Sydney, p. 115-141.
- (4) Tice, R.R., Hayashi, M., MacGregor, J.T., Anderson, D., Blakey, D.H., Holden, H.E., Kirsch-Volders, M., Oleson Jr., F.B., Pacchierotti, F., Preston, R.J., Romagna, F., Shimada, H., Sutou, S. and Vannier, B., (1994) Report from the Working Group on the *In Vivo* Mammalian Bone Marrow Chromosomal Aberration Test. *Mutation Res.*, 312, p. 305-312.
- (5) Fielder, R.J., Allen, J.A., Boobis, A.R., Botham, P.A., Doe, J., Esdaile, D.J., Gatehouse, D.G., Hodson-Walker, G., Morton, D.B., Kirkland, D.J. and Richold, M. (1992). Report of British Toxicology Society/UK Environmental Mutagen Society Working Group: Dose setting in *In vivo* Mutagenicity Assays. *Mutagenesis*, 7, p. 313-319.

**▼B**

- (6) Lovell, D.P., Anderson, D., Albanese, R., Amphlett, G.E., Clare, G., Ferguson, R., Richold, M., Papworth, D.G. and Savage, J.R.K. (1989). Statistical Analysis of *In Vivo* Cytogenetic Assays. In: UKEMS Sub-Committee on Guidelines for Mutagenicity Testing. Report Part III. Statistical Evaluation of Mutagenicity Test Data. D.J. Kirkland, (Ed.) Cambridge University Press, Cambridge. p. 184-232.
- (7) Locke-Huhle, C., (1983) Endoreduplication in Chinese hamster cells during alpha-radiation induced G2 arrest. Mutation Res. 119, p. 403-413.
- (8) Huang, Y., Change, C. and Trosko, J.E., (1983) Aphidicolin-induced endoreduplication in Chinese hamster cells. Cancer Res., 43, p. 1362-1364.

**▼ B****B.12. MUTAGENICITY — *IN VIVO* MAMMALIAN ERYTHROCYTE MICRONUCLEUS TEST****1. METHOD**

This method is a replicate of the OECD TG 474, Mammalian Erythrocyte Micronucleus Test (1997).

**1.1. INTRODUCTION**

The mammalian *in vivo* micronucleus test is used for the detection of damage induced by the test substance to the chromosomes or the mitotic apparatus of erythroblasts by analysis of erythrocytes as sampled in bone marrow and/or peripheral blood cells of animals, usually rodents.

The purpose of the micronucleus test is to identify substances that cause cytogenetic damage, which results in the formation of micronuclei containing lagging chromosome fragments or whole chromosomes.

When a bone marrow erythroblast develops into a polychromatic erythrocyte, the main nucleus is extruded; any micronucleus that has been formed may remain behind in the otherwise anucleated cytoplasm. Visualisation of micronuclei is facilitated in these cells because they lack a main nucleus. An increase in the frequency of micronucleated polychromatic erythrocytes in treated animals is an indication of induced chromosome damage.

The bone marrow of rodents is routinely used in this test since polychromatic erythrocytes are produced in that tissue. The measurement of micronucleated immature (polychromatic) erythrocytes in peripheral blood is equally acceptable in any species in which the inability of the spleen to remove micronucleated erythrocytes has been demonstrated, or which has shown an adequate sensitivity to detect agents that cause structural or numerical chromosome aberrations. Micronuclei can be distinguished by a number of criteria. These include identification of the presence or absence of a kinetochore or centromeric DNA in the micronuclei. The frequency of micronucleated immature (polychromatic) erythrocytes is the principal endpoint. The number of mature (normochromatic) erythrocytes in the peripheral blood that contain micronuclei among a given number of mature erythrocytes can also be used as the endpoint of the assay when animals are treated continuously for four weeks or more.

This mammalian *in vivo* micronucleus test is especially relevant to assessing mutagenic hazard in that it allows consideration of factors of *in vivo* metabolism, pharmacokinetics and DNA-repair processes although these may vary among species, among tissues and among genetic endpoints. An *in vivo* assay is also useful for further investigation of a mutagenic effect detected by an *in vitro* system.

If there is evidence that the test substance, or a reactive metabolite, will not reach the target tissue, it is not appropriate to use this test.

See also General introduction Part B.

**▼ B**

## 1.2. DEFINITIONS

**Centromere (Kinetochore):** region(s) of a chromosome with which spindle fibers are associated during cell division, allowing orderly movement of daughter chromosomes to the poles of the daughter cells.

**Micronuclei:** small nuclei, separate from and additional to the main nuclei of cells, produced during telophase of mitosis (meiosis) by lagging chromosome fragments or whole chromosomes.

**Normochromatic erythrocyte:** mature erythrocyte that lacks ribosomes and can be distinguished from immature, polychromatic erythrocytes by stains selective for ribosomes.

**Polychromatic erythrocyte:** immature erythrocyte, in an intermediate stage of development, that still contains ribosomes and therefore can be distinguished from mature, normochromatic erythrocytes by stains selective for ribosomes.

## 1.3. PRINCIPLE OF THE TEST METHOD

Animals are exposed to the test substance by an appropriate route. If bone marrow is used, the animals are sacrificed at appropriate times after treatment, the bone marrow extracted, and preparations made and stained (1)(2)(3)(4)(5)(6)(7). When peripheral blood is used, the blood is collected at appropriate times after treatment and smear preparations are made and stained (4)(8)(9)(10). For studies with peripheral blood, as little time as possible should elapse between the last exposure and cell harvest. Preparations are analysed for the presence of micronuclei.

## 1.4. DESCRIPTION OF THE TEST METHOD

## 1.4.1. Preparations

1.4.1.1. *Selection of animal species*

Mice or rats are recommended if bone marrow is used, although any appropriate mammalian species may be used. When peripheral blood is used, mice are recommended. However, any appropriate mammalian species may be used provided it is a species in which the spleen does not remove micronucleated erythrocytes or a species, which has shown an adequate sensitivity to detect agents that cause structural or numerical chromosome aberrations. Commonly used laboratory strains of young healthy animals should be employed. At the commencement of the study, the weight variation of animals should be minimal and not exceed  $\pm 20$  % of the mean weight of each sex.

1.4.1.2. *Housing and feeding conditions*

General conditions referred in the General introduction to Part B are applied although the aim for humidity should be 50-60 %.

**▼ B**1.4.1.3. *Preparation of the animals*

Healthy young adult animals are randomly assigned to the control and treatment groups. The animals are identified uniquely. The animals are acclimated to the laboratory conditions for at least five days. Cages should be arranged in such a way that possible effects due to cage placement are minimised.

1.4.1.4. *Preparation of doses*

Solid test substances should be dissolved or suspended in appropriate solvents or vehicles and diluted, if appropriate, prior to dosing of the animals. Liquid test substances may be dosed directly or diluted prior to dosing. Fresh preparations of the test substance should be employed unless stability data demonstrate the acceptability of storage.

1.4.2. **Test conditions**1.4.2.1. *Solvent/Vehicle*

The solvent/vehicle should not produce toxic effects at the dose levels used, and should not be suspected of chemical reaction with the test substance. If other than well-known solvents/vehicles are used, their inclusion should be supported with reference data indicating their compatibility. It is recommended that wherever possible, the use of an aqueous solvent/vehicle should be considered first.

1.4.2.2. *Controls*

Concurrent positive and negative (solvent/vehicle) controls should be included for each sex in each test. Except for treatment with the test substance, animals in the control groups should be handled in an identical manner to animals of the treatment groups.

Positive controls should produce micronuclei *in vivo* at exposure levels expected to give a detectable increase over background. Positive control doses should be chosen so that the effects are clear but do not immediately reveal the identity of the coded slides to the reader. It is acceptable that the positive control be administered by a route different from the test substance and sampled at only a single time. In addition, the use of chemical class-related positive control chemicals may be considered, when available. Examples of positive control substances include:

Substance	CAS No	EINECS No
Ethyl methanesulphonate	62-50-0	200-536-7
N-ethyl-N-nitrosourea	759-73-9	212-072-2
Mitomycin C	50-07-7	200-008-6
Cyclophosphamide	50-18-0	200-015-4
Cyclophosphamide monohydrate	6055-19-2	
Triethylenemelamine	51-18-3	200-083-5

**▼B**

Negative controls, treated with solvent or vehicle alone, and otherwise treated in the same way as the treatment groups should be included for every sampling time, unless acceptable inter-animal variability and frequencies of cells with micronuclei are demonstrated by historical control data. If single sampling is applied for negative controls, the most appropriate time is the first sampling time. In addition, untreated controls should also be used unless there are historical or published control data demonstrating that no deleterious or mutagenic effects are induced by the chosen solvent/-vehicle.

If peripheral blood is used, a pre-treatment sample may also be acceptable as a concurrent negative control, but only in the short peripheral blood studies (e.g. 1-3 treatment(s)) when the resulting data are in the expected range for the historical control.

## 1.5. PROCEDURE

### 1.5.1. **Number and sex of animals**

Each treated and control group must include at least five analysable animals per sex (11). If at the time of the study there are data available from studies in the same species and using the same route of exposure that demonstrate that there are no substantial differences between sexes in toxicity, then testing in a single sex will be sufficient. Where human exposure to chemicals may be sex-specific, as for example with some pharmaceutical agents, the test should be performed with animals of the appropriate sex.

### 1.5.2. **Treatment schedule**

No standard treatment schedule (i.e. one, two or more treatments at 24 h intervals) can be recommended. The samples from extended dose regimens are acceptable as long as a positive effect has been demonstrated for this study or, for a negative study, as long as toxicity has been demonstrated or the limit dose has been used, and dosing continued until the time of sampling. Test substances may also be administered as a split dose, i.e., two treatments on the same day separated by no more than a few hours, to facilitate administering a large volume of material.

The test may be performed in two ways:

- (a) animals are treated with the test substance once. Samples of bone marrow are taken at least twice, starting not earlier than 24 hours after treatment, but not extending beyond 48 hours after treatment with appropriate intervals between samples. The use of sampling times earlier than 24 hours after treatment should be justified. Samples of peripheral blood are taken at least twice, starting not earlier than 36 hours after treatment, with appropriate intervals following the first sample, but not extending beyond 72 hours. When a positive response is recognised at one sampling time, additional sampling is not required.

**▼B**

- (b) if two or more daily treatments are used (e.g. two or more treatments at 24 hour intervals), samples should be collected once between 18 and 24 hours following the final treatment for the bone marrow and once between 36 and 48 hours following the final treatment for the peripheral blood (12);

Other sampling times may be used in addition, when relevant.

**1.5.3. Dose levels**

If a range finding study is performed because there are no suitable data available, it should be performed in the same laboratory, using the same species, strain, sex, and treatment regimen to be used in the main study (13). If there is toxicity, three dose levels are used for the first sampling time. These dose levels should cover a range from the maximum to little or no toxicity. At the later sampling time only the highest dose needs to be used. The highest dose is defined as the dose producing signs of toxicity such that higher dose levels, based on the same dosing regimen, would be expected to produce lethality. Substances with specific biological activities at low non-toxic doses (such as hormones and mitogens) may be exceptions to the dose-setting criteria and should be evaluated on a case-by-case basis. The highest dose may also be defined as a dose that produces some indication of toxicity in the bone marrow (e.g. a reduction in the proportion of immature erythrocytes among total erythrocytes in the bone marrow or peripheral blood).

**1.5.4. Limit test**

If a test at one dose level of at least 2 000 mg/kg body weight using a single treatment, or as two treatments on the same day, produces no observable toxic effects, and if genotoxicity would not be expected based upon data from structurally related substances, then a full study using three dose levels may not be considered necessary. For studies of a longer duration, the limit dose is 2 000 mg/kg/body weight/day for treatment up to 14 days, and 1 000 mg/kg/body weight/day for treatment longer than 14 days. Expected human exposure may indicate the need for a higher dose level to be used in the limit test.

**1.5.5. Administration of doses**

The test substance is usually administered by gavage using a stomach tube or a suitable intubation cannula, or by intraperitoneal injection. Other routes of exposure may be acceptable where they can be justified. The maximum volume of liquid that can be administered by gavage or injection at one time depends on the size of the test animal. The volume should not exceed 2 ml/100 g body weight. The use of volumes higher than these must be justified. Except for irritating or corrosive substances, which will normally reveal exacerbated effects with higher concentrations, variability in test volume should be minimised by adjusting the concentration to ensure a constant volume at all dose levels.

**▼ B****1.5.6. Bone marrow/blood preparation**

Bone marrow cells are usually obtained from the femurs or tibias immediately following sacrifice. Commonly, cells are removed from femurs or tibias, prepared and stained using established methods. Peripheral blood is obtained from the tail vein or other appropriate blood vessel. Blood cells are immediately stained supravivally (8)-(9)(10) or smear preparations are made and then stained. The use of a DNA specific stain (e.g. acridine orange (14) or Hoechst 33258 plus pyronin-Y (15)) can eliminate some of the artifacts associated with using a non DNA specific stain. This advantage does not preclude the use of conventional stains (e.g. Giemsa). Additional systems (e.g. cellulose columns to remove nucleated cells (16)) can also be used provided that these systems have been shown to adequately work for micronucleus preparation in the laboratory.

**1.5.7. Analysis**

The proportion of immature among total (immature + mature) erythrocytes is determined for each animal by counting a total of at least 200 erythrocytes for bone marrow and 1 000 erythrocytes for peripheral blood (17). All slides, including those of positive and negative controls, should be independently coded before microscopic analysis. At least 2 000 immature erythrocytes per animal are scored for the incidence of micronucleated immature erythrocytes. Additional information may be obtained by scoring mature erythrocytes for micronuclei. When analysing slides, the proportion of immature erythrocytes among total erythrocytes should not be less than 20 % of the control value. When animals are treated continuously for four weeks or more, at least 2 000 mature erythrocytes per animal can also be scored for the incidence of micronuclei. Systems for automated analysis (image analysis and cell suspensions flow cytometry) are acceptable alternatives to manual evaluation if appropriately justified and validated.

**2. DATA****2.1. TREATMENT OF RESULTS**

Individual animal data should be presented in tabular form. The experimental unit is the animal. The number of immature erythrocytes scored, the number of micronucleated immature erythrocytes, and the number of immature among total erythrocytes should be listed separately for each animal analysed. When animals are treated continuously for four weeks or more, the data on mature erythrocytes should also be given if it is collected. The proportion of immature among total erythrocytes and, if considered applicable, the percentage of micronucleated erythrocytes is given for each animal. If there is no evidence for a difference in response between the sexes, the data from both sexes may be combined for statistical analysis.



**▼B****2.2. EVALUATION AND INTERPRETATION OF RESULTS**

There are several criteria for determining a positive result, such as a dose-related increase in the number of micronucleated cells or a clear increase in the number of micronucleated cells in a single dose group at a single sampling time. Biological relevance of the results should be considered first. Statistical methods may be used as an aid in evaluating the test results (18)(19). Statistical significance should not be the only determining factor for a positive response. Equivocal results should be clarified by further testing preferably using a modification of experimental conditions.

A test substance for which the results do not meet the above criteria is considered non-mutagenic in this test.

Although most experiments will give clearly positive or negative results, in rare cases the data set will preclude making a definite judgement about the activity of the test substance. Results may remain equivocal or questionable regardless of the number of times the experiment is repeated.

Positive results in the micronucleus test indicate that the substance induces micronuclei which are the result of chromosomal damage or damage to the mitotic apparatus in the erythroblasts of the test species. Negative results indicate that, under the test conditions, the test substance does not produce micronuclei in the immature erythrocytes of the test species.

The likelihood that the test substance or its metabolites reach the general circulation or specifically the target tissue (e.g. systemic toxicity) should be discussed.

**3. REPORTING****TEST REPORT**

The test report should include the following information:

**Solvent/Vehicle:**

- justification for choice of vehicle,
- solubility and stability of the test substance in solvent/vehicle, if known.

**Test animals:**

- species/strain used,
- number, age and sex of animals,
- source, housing conditions, diet, etc.,
- individual weight of the animals at the start of the test, including body weight range, mean and standard deviation for each group.

**Test conditions:**

- positive and negative (vehicle/solvent) control data,
- data from range-finding study, if conducted,

**▼ B**

- rationale for dose level selection,
- details of test substance preparation,
- details of the administration of the test substance,
- rationale for route of administration,
- methods for verifying that the test substance reached the general circulation or target tissue, if applicable,
- conversion from diet/drinking water test substance concentration (ppm) to the actual dose (mg/kg body weight/day), if applicable,
- details of food and water quality,
- detailed description of treatment and sampling schedules,
- methods of slide preparation,
- methods for measurements of toxicity,
- criteria for scoring micronucleated immature erythrocytes,
- number of cells analysed per animal,
- criteria for considering studies as positive, negative or equivocal.

## Results:

- signs of toxicity,
- proportion of immature erythrocytes among total erythrocytes,
- number of micronucleated immature erythrocytes, given separately for each animal,
- mean  $\pm$  standard deviation of micronucleated immature erythrocytes per group,
- dose-response relationship, where possible,
- statistical analyses and methods applied,
- concurrent and historical negative control data,
- concurrent positive control data.
- Discussion of the results.
- Conclusions.

4. **REFERENCES**

- (1) Heddle, J.A., (1973) A Rapid *In Vivo* Test for Chromosomal Damage, *Mutation Res.*, 18, p. 187-190.
- (2) Schmid, W., (1975) The Micronucleus Test, *Mutation Res.*, 31, p. 9-15.
- (3) Heddle, J.A., Salamone, M.F., Hite, M., Kirkhart, B., Mavourmin, K., MacGregor, J.G. and Newell, G.W. (1983). The Induction of Micronuclei as a Measure of Genotoxicity. *Mutation Res.*, 123, p. 61-118.
- (4) Mavourmin, K.H., Blakey, D.H., Cimino, M.C., Salamone, M.F. and Heddle, J.A., (1990) The *In Vivo* Micronucleus Assay in Mammalian Bone Marrow and Peripheral Blood. A report of the U.S. Environmental Protection Agency Gene-Tox Program, *Mutation Res.*, 239, p. 29-80.

**▼B**

- (5) MacGregor, J.T., Schlegel, R. Choy, W.N., and Wehr, C.M., (1983) Micronuclei in Circulating Erythrocytes: A Rapid Screen for Chromosomal Damage During Routine Toxicity Testing in Mice. In: 'Developments in Science and Practice of Toxicology', Ed. A.W. Hayes, R.C. Schnell and T.S. Miya, Elsevier, Amsterdam, p. 555-558.
- (6) MacGregor, J.T., Heddle, J.A. Hite, M., Margolin, G.H., Ramel, C., Salamone, M.F., Tice, R.R., and Wild, D., (1987) Guidelines for the Conduct of Micronucleus Assays in Mammalian Bone Marrow Erythrocytes. *Mutation Res.*, 189, p. 103-112.
- (7) MacGregor, J.T., Wehr, C.M., Henika, P.R., and Shelby, M.E., (1990) The *in vivo* Erythrocyte Micronucleus Test: Measurement at Steady State Increases Assay Efficiency and Permits Integration with Toxicity Studies. *Fund. Appl. Toxicol.*, 14, p. 513-522.
- (8) Hayashi, M., Morita, T., Kodama, Y., Sofuni, T. and Ishidate, M. Jr., (1990) The Micronucleus Assay with Mouse Peripheral Blood Reticulocytes Using Acridine Orange-Coated Slides. *Mutation Res.*, 245, p. 245-249.
- (9) The Collaborative Study Group for the Micronucleus Test, (1992) Micronucleus Test with Mouse Peripheral Blood Erythrocytes by Acridine Orange Supravital Staining: The Summary Report of the 5th Collaborative Study by CSGMT/-JEMMS. MMS. *Mutation Res.*, p. 278, 83-98.
- (10) The Collaborative Study Group for the Micronucleus Test (CSGMT/JEMMS. MMS: The Mammalian Mutagenesis Study Group of the Environmental Mutagen Society of Japan), (1995) Protocol recommended for the short-term mouse peripheral blood micronucleus test. *Mutagenesis*, 10, p. 153-159.
- (11) Hayashi, M., Tice, R.R., MacGregor, J.T., Anderson, D., Blackey, D.H., Kirsch-Volders, M., Oleson, Jr. F. B., Pacchierotti, F., Romagna, F., Shimada, H., Sutou, S. and Vannier, B., (1994) *In Vivo* Rodent Erythrocyte Micronucleus Assay. *Mutation Res.*, 312, p. 293-304.
- (12) Higashikuni, N. and Sutou, S., (1995) An optimal, generalised sampling time of  $30 \pm 6$  h after double dosing in the mouse peripheral blood micronucleus test. *Mutagenesis*, 10, p. 313-319.
- (13) Fielder, R.J., Allen, J.A., Boobis, A.R., Botham, P.A., Doe, J., Esdaile, D.J., Gatehouse, D.G., Hodson-Walker, G., Morton, D.B., Kirkland, D.J. and Rochold, M., (1992) Report of British Toxicology Society/UK Environmental Mutagen Society Working Group: Dose Setting in *In Vivo* Mutagenicity Assays. *Mutagenesis*, 7, p. 313-319.
- (14) Hayashi, M., Sofuni, T. and Ishidate, M. Jr., (1983) An Application of Acridine Orange Fluorescent Staining to the Micronucleus Test. *Mutation Res.*, 120, p. 241-247.
- (15) MacGregor, J.T., Wehr, C.M. and Langlois, R.G., (1983) A Simple Fluorescent Staining Procedure for Micronuclei and RNA in Erythrocytes Using Hoechst 33258 and Pyronin Y. *Mutation Res.*, 120, p. 269-275.
- (16) Romagna, F. and Staniforth, C.D., (1989) The automated bone marrow micronucleus test. *Mutation Res.*, 213, p. 91-104.
- (17) Gollapudi, B. and McFadden, L.G., (1995) Sample size for the estimation of polychromatic to normochromatic erythrocyte ratio in the bone marrow micronucleus test. *Mutation Res.*, 347, p. 97-99.

**▼B**

- (18) Richold, M., Ashby, J., Bootman, J., Chandley, A., Gatehouse, D.G. and Henderson, L., (1990) *In Vivo* Cytogenetics Assay. In: D.J. Kirkland (Ed.) Basic Mutagenicity tests. UKEMS Recommended Procedures. UKEMS Sub-Committee on Guidelines for Mutagenicity Testing. Report, Part I, revised. Cambridge University Press, Cambridge, New York, Port Chester, Melbourne, Sydney, p. 115-141.
- (19) Lovell, D.P., Anderson, D., Albanese, R., Amphlett, G.E., Clare, G., Ferguson, R., Richold, M., Papworth, D.G. and Savage, J.R.K., (1989) Statistical Analysis of *In Vivo* Cytogenetic Assays. In: D.J. Kirkland (Ed.) Statistical Evaluation of Mutagenicity Test Data. UKEMS Sub-Committee on Guidelines for Mutagenicity Testing. Report, Part III. Cambridge University Press, Cambridge, New York, Port Chester, Melbourne, Sydney, p. 184-232.

**▼B****B.13/14. MUTAGENICITY: REVERSE MUTATION TEST USING BACTERIA****1. METHOD**

This method is a replicate of the OECD TG 471, Bacterial Reverse Mutation Test (1997).

**1.1. INTRODUCTION**

The bacterial reverse mutation test uses amino-acid requiring strains of *Salmonella typhimurium* and *Escherichia coli* to detect point mutations, which involve substitution, addition or deletion of one or a few DNA base pairs (1)(2)(3). The principle of this bacterial reverse mutation test is that it detects mutations, which revert mutations present in the test strains and restore the functional capability of the bacteria to synthesise an essential amino acid. The revertant bacteria are detected by their ability to grow in the absence of the amino-acid required by the parent test strain.

Point mutations are the cause of many human genetic diseases and there is substantial evidence that point mutations in oncogenes and tumour-suppressor genes of somatic cells are involved in tumour formation in humans and experimental animals. The bacterial reverse mutation test is rapid, inexpensive and relatively easy to perform. Many of the test strains have several features that make them more sensitive for the detection of mutations including responsive DNA sequences at the reversion sites, increased cell permeability to large molecules and elimination of DNA repair systems or enhancement of error-prone DNA repair processes. The specificity of the test strains can provide some useful information on the types of mutations that are induced by genotoxic agents. A very large data base of results for a wide variety of structures is available for bacterial reverse mutation tests and well-established methodologies have been developed for testing chemicals with different physico-chemical properties, including volatile compounds.

See also General introduction Part B.

**1.2. DEFINITIONS**

**A reverse mutation test** in either *Salmonella typhimurium* or *Escherichia coli* detects mutation in an amino-acid requiring strain (histidine or tryptophan, respectively) to produce a strain independent of an outside supply of amino-acid.

**Base pair substitution mutagens** are agents that cause a base change in DNA. In a reversion test this change may occur at the site of the original mutation, or at a second site in the bacterial genome.

**Frameshift mutagens** are agents that cause the addition or deletion of one or more base pairs in the DNA, thus changing the reading frame in the RNA.

**▼ B**

## 1.3. INITIAL CONSIDERATIONS

The bacterial reverse mutation test utilises prokaryotic cells, which differ from mammalian cells in such factors as uptake, metabolism, chromosome structure and DNA repair processes. Tests conducted *in vitro* generally require the use of an exogenous source of metabolic activation. *In vitro* metabolic activation systems cannot mimic entirely the mammalian *in vivo* conditions. The test therefore does not provide direct information on the mutagenic and carcinogenic potency of a substance in mammals.

The bacterial reverse mutation test is commonly employed as an initial screen for genotoxic activity and, in particular, for point mutation-inducing activity. An extensive database has demonstrated that many chemicals that are positive in this test also exhibit mutagenic activity in other tests. There are examples of mutagenic agents, which are not detected by this test; reasons for these shortcomings can be ascribed to the specific nature of the endpoint detected, differences in metabolic activation, or differences in bio-availability. On the other hand, factors, which enhance the sensitivity of the bacterial reverse mutation test can lead to an overestimation of mutagenic activity.

The bacterial reverse mutation test may not be appropriate for the evaluation of certain classes of chemicals, for example highly bactericidal compounds (e.g. certain antibiotics) and those which are thought (or known) to interfere specifically with the mammalian cell replication system (e.g. some topoisomerase inhibitors and some nucleoside analogues). In such cases, mammalian mutation tests may be more appropriate.

Although many compounds that are positive in this test are mammalian carcinogens, the correlation is not absolute. It is dependent on chemical class and there are carcinogens that are not detected by this test because they act through other, non-genotoxic, mechanisms or mechanisms absent in bacterial cells.

## 1.4. PRINCIPLE OF THE TEST METHOD

Suspensions of bacterial cells are exposed to the test substance in the presence and in the absence of an exogenous metabolic activation system. In the plate incorporation method, these suspensions are mixed with an overlay agar and plated immediately onto minimal medium. In the preincubation method, the treatment mixture is incubated and then mixed with an overlay agar before plating onto minimal medium. For both techniques, after two or three days of incubation, revertant colonies are counted and compared to the number of spontaneous revertant colonies on solvent control plates.

Several procedures for performing the bacterial reverse mutation test have been described. Among those commonly used are the plate incorporation method (1)(2)(3)(4), the preincubation method (2)(3)(5)(6)(7)(8), the fluctuation method (9)(10), and the suspension method (11). Modifications for the testing of gases or vapours have been described (12).

**▼ B**

The procedures described in the method pertain primarily to the plate incorporation and preincubation methods. Either of them is acceptable for conducting experiments both with and without metabolic activation. Some substances may be detected more efficiently using the preincubation method. These substances belong to chemical classes that include short chain aliphatic nitrosamines, divalent metals, aldehydes, azo-dyes and diazo compounds, pyrrolizidine alkaloids, allyl compounds and nitro compounds (3). It is also recognised that certain classes of mutagens are not always detected using standard procedures such as the plate incorporation method or preincubation method. These should be regarded as 'special cases' and it is strongly recommended that alternative procedures should be used for their detection. The following 'special cases' could be identified (together with examples of procedures that could be used for their detection): azo-dyes and diazo compounds (3)(5)(6)(13), gases and volatile chemicals (12)-(14)(15)(16) and glycosides (17)(18). A deviation from the standard procedure needs to be scientifically justified.

## 1.5. DESCRIPTION OF THE TEST METHOD

1.5.1. **Preparations**1.5.1.1. *Bacteria*

Fresh cultures of bacteria should be grown up to the late exponential or early stationary phase of growth (approximately  $10^9$  cells per ml). Cultures in late stationary phase should not be used. It is essential that the cultures used in the experiment contain a high titre of viable bacteria. The titre may be demonstrated either from historical control data on growth curves, or in each assay through the determination of viable cell numbers by a plating experiment.

The recommended incubation temperature is 37 °C.

At least five strains of bacteria should be used. These should include four strains of *S. typhimurium* (TA 1535; TA 1537 or TA97a or TA97; TA98; and TA100) that have been shown to be reliable and reproducibly responsive between laboratories. These four *S. typhimurium* strains have GC base pairs at the primary reversion site and it is known that may not detect certain oxidising mutagens, cross-linking agents and hydrazines. Such substances may be detected by *E. coli* WP2 strains or *S. typhimurium* TA102 (19), which have an AT base pair at the primary reversion site. Therefore the recommended combination of strains is:

— *S. typhimurium* TA1535, and

— *S. typhimurium* TA1537 or TA97 or TA97a, and

— *S. typhimurium* TA98, and

— *S. typhimurium* TA100, and

— *E. coli* WP2 uvrA, or *E. coli* WP2 uvrA (pKM101), or *S. typhimurium* TA102.

In order to detect cross-linking mutagens it may be preferable to include TA102 or to add a DNA repair-proficient strain of *E. coli* [e.g. *E. coli* WP2 or *E. coli* WP2 (pKM101)]

**▼B**

Established procedures for stock culture preparation, marker verification and storage should be used. The amino-acid requirement for growth should be demonstrated for each frozen stock culture preparation (histidine for *S. typhimurium* strains, and tryptophan for *E. coli* strains). Other phenotypic characteristics should be similarly checked, namely: the presence or absence of R-factor plasmids where appropriate [i.e. ampicillin resistance in strains TA98, TA100 and TA97a or TA97, WP2 uvrA and WP2 uvrA (pKM101), and ampicillin + tetracycline resistance in strain TA102]; the presence of characteristic mutations (i.e. rfa mutation in *S. typhimurium* through sensitivity to crystal violet, and uvrA mutation in *E. coli* or uvrB mutation in *S. typhimurium*, through sensitivity to ultraviolet light) (2)(3). The strains should also yield spontaneous revertant colony plate counts within the frequency ranges expected from the laboratory's historical control data and preferably within the range reported in the literature.

1.5.1.2. *Medium*

An appropriate minimal agar (e.g. containing Vogel-Bonner minimal medium E and glucose), and an overlay agar containing histidine and biotin or tryptophan to allow for a few cell divisions, is used (1)-(2)(9).

1.5.1.3. *Metabolic activation*

Bacteria should be exposed to the test substance both in the presence and absence of an appropriate metabolic activation system. The most commonly used system is a cofactor-supplemented post-mitochondrial fraction (S9) prepared from the livers of rodents treated with enzyme-inducing agents such as Aroclor 1254 (1)(2) or a combination of Phenobarbitone and  $\beta$ -naphthoflavone (18)(20)(21). The post-mitochondrial fraction is usually used at concentrations in the range from 5 to 30 % v/v in the S9-mix. The choice and condition of a metabolic activation system may depend upon the class of chemical being tested. In some cases, it may be appropriate to utilise more than one concentration of post-mitochondrial fraction. For azo-dyes and diazo-compounds, using a reductive metabolic activation system may be more appropriate (6)(13).

1.5.1.4. *Test substance/Preparation*

Solid test substances should be dissolved or suspended in appropriate solvents or vehicles and diluted if appropriate prior to treatment of the bacteria. Liquid test substances may be added directly to the test systems and/or diluted prior to treatment. Fresh preparations should be employed unless stability data demonstrate the acceptability of storage.

The solvent/vehicle should not be suspected of chemical reaction with the test substance and should be compatible with the survival of the bacteria and the S9 activity (22). If other than well-known solvent/vehicles are used, their inclusion should be supported by data indicating their compatibility. It is recommended that wherever possible, the use of an aqueous solvent/vehicle be considered first. When testing water-unstable substances, the organic solvents used should be free of water.



**▼B****1.5.2. Test conditions**1.5.2.1. *Test strains (see 1.5.1.1)*1.5.2.2. *Exposure concentration*

Amongst the criteria to be taken into consideration when determining the highest amount of the test substance to be used are the cytotoxicity and the solubility in the final treatment mixture.

It may be useful to determine toxicity and insolubility in a preliminary experiment. Cytotoxicity may be detected by a reduction in the number of revertant colonies, a clearing or diminution of the background lawn, or the degree of survival of treated cultures. The cytotoxicity of a substance may be altered in the presence of metabolic activation systems. Insolubility should be assessed as precipitation in the final mixture under the actual test conditions and evident to the unaided eye.

The recommended maximum test concentration for soluble non-cytotoxic substances is 5 mg/plate or 5 µl/plate. For non-cytotoxic substances that are not soluble at 5 mg/plate or 5 µl/plate, one or more concentrations tested should be insoluble in the final treatment mixture. Test substances that are cytotoxic already below 5 mg/plate or 5 µl/plate should be tested up to a cytotoxic concentration. The precipitate should not interfere with the scoring.

At least five different analysable concentrations of the test substance should be used with approximately half log (i.e.  $\sqrt{10}$ ) intervals between test points for an initial experiment. Smaller intervals may be appropriate when a concentration-response is being investigated. Testing above the concentration of 5 mg/plate or 5 µl/plate may be considered when evaluating substances containing substantial amounts of potentially mutagenic impurities.

1.5.2.3. *Negative and positive controls*

Concurrent strain-specific positive and negative (solvent or vehicle) controls, both with and without metabolic activation, should be included in each assay. Positive control concentrations that demonstrate the effective performance of each assay should be selected.

For assays employing a metabolic activation system, the positive control reference substance(s) should be selected on the basis of the type of bacteria strains used.

The following substances are examples of suitable positive controls for assays with metabolic activation:

CA numbers	EINECS numbers	Names
781-43-1	212-308-4	9,10-dimethylanthracene
57-97-6	200-359-5	7,12-dimethylbenz[ <i>a</i> ]anthracene
50-32-8	200-028-5	benzo[ <i>a</i> ]pyrene
613-13-8	210-330-9	2-aminoanthracene

**▼B**

CA numbers	EINECS numbers	Names
50-18-0		cyclophosphamide
6055-19-2	200-015-4	cyclophosphamide monohydrate

The following substance is a suitable positive control for the reductive metabolic activation method:

CA numbers	EINECS numbers	Names
573-58-0	209-358-4	Congo Red

2-Aminoanthracene should not be used as the sole indicator of the efficacy of the S9-mix. If 2-aminoanthracene is used, each batch of S9 should also be characterised with a mutagen that requires metabolic activation by microsomal enzymes, e.g. benzo[a]pyrene, dimethylbenzanthracene.

The following substances are examples of strain-specific positive controls for assays performed without exogenous metabolic activation system:

CAS numbers	EINECS numbers	Names	Strain
26628-22-8	247-852-1	Sodium azide	TA 1535 and TA 100
607-57-8	210-138-5	2-nitrofluorene	TA 98
90-45-9	201-995-6	9-aminoacridine	TA 1537, TA 97 and TA 97a
17070-45-0	241-129-4	ICR 191	TA 1537, TA 97 and TA 97a
80-15-9	201-254-7	Cumene hydroperoxide	TA 102
50-07-7	200-008-6	Mitomycin C	WP2 uvrA and TA102
70-25-7	200-730-1	N-ethyl-N-nitro-N-nitrosoguanidine	WP2, WP2uvrA and WP2uvrA(pKM101)
56-57-5	200-281-1	4-nitroquinoline-1-oxide	WP2, WP2uvrA and WP2uvrA(pKM101)
3688-53-7		Furylfuramide (AF2)	plasmid-containing strains

Other appropriate positive control reference substances may be used. The use of chemical class-related positive control chemicals should be considered, when available.

Negative controls, consisting of solvent or vehicle alone, without test substance, and otherwise treated in the same way as the treatment groups, should be included. In addition, untreated controls should also be used unless there are historical control data demonstrating that no deleterious or mutagenic effects are induced by the chosen solvent.

**▼B****1.5.3. Procedure**

For the plate incorporation method (1)(2)(3)(4), without metabolic activation, usually 0,05 ml or 0,1 ml of the test solutions, 0,1 ml of fresh bacterial culture (containing approximately  $10^8$  viable cells) and 0,5 ml of sterile buffer are mixed with 2,0 ml of overlay agar. For the assay with metabolic activation, usually 0,5 ml of metabolic activation mixture containing an adequate amount of post-mitochondrial fraction (in the range from 5 to 30 % v/v in the metabolic activation mixture) are mixed with the overlay agar (2,0 ml), together with the bacteria and test substance/test solution. The contents of each tube are mixed and poured over the surface of a minimal agar plate. The overlay agar is allowed to solidify before incubation.

For the preincubation method (2)(3)(5)(6), the test substance/test solution is preincubated with the test strain (containing approximately  $10^8$  viable cells) and sterile buffer or the metabolic activation system (0,5 ml) usually for 20 min. or more at 30-37 °C prior to mixing with the overlay agar and pouring onto the surface of a minimal agar plate. Usually, 0,05 or 0,1 ml of test substance/test solution, 0,1 ml of bacteria, and 0,5 ml of S9-mix or sterile buffer are mixed with 2,0 ml of overlay agar. Tubes should be aerated during pre-incubation by using a shaker.

For an adequate estimate of variation, triplicate plating should be used at each dose level. The use of duplicate plating is acceptable when scientifically justified. The occasional loss of a plate does not necessarily invalidate the assay.

Gaseous or volatile substances should be tested by appropriate methods, such as in sealed vessels (12)(14)(15)(16).

**1.5.4. Incubation**

All plates in a given assay should be incubated at 37 °C for 48-72 hours. After the incubation period, the number of revertant colonies per plate is counted.

**2. DATA****2.1. TREATMENT OF RESULTS**

Data should be presented as the number of revertant colonies per plate. The number of revertant colonies on both negative (solvent control, and untreated control if used) and positive control plates should also be given. Individual plate counts, the mean number of revertant colonies per plate and the standard deviation should be presented for the test substance and positive and negative (untreated and/or solvent) controls.

There is no requirement for verification of a clear positive response. Equivocal results should be clarified by further testing preferably using a modification of experimental conditions. Negative results need to be confirmed on a case-by-case basis. In those cases where confirmation of negative results is not considered necessary, justification should be provided. Modification of study parameters to extend the range of conditions assessed should be considered in follow-up experiments. Study parameters that might be modified include the concentration spacing, the method of treatment (plate-incorporation or liquid pre-incubation), and metabolic activation conditions.

**▼B****2.2. EVALUATION AND INTERPRETATION OF RESULTS**

There are several criteria for determining a positive result, such as a concentration-related increase over the range tested and/or a reproducible increase at one or more concentrations in the number of revertant colonies per plate in at least one strain with or without metabolic activation system (23). Biological relevance of the results should be considered first. Statistical methods may be used as an aid in evaluating the test results (24). However, statistical significance should not be the only determining factor for a positive response.

A test substance for which the results do not meet the above criteria is considered non-mutagenic in this test.

Although most experiments will give clearly positive or negative results, in rare cases the data set will preclude making a definite judgement about the activity of the test substance. Results may remain equivocal or questionable regardless of the number of times the experiment is repeated.

Positive results from the bacterial reverse mutation test indicate that the substance induces point mutations by base substitutions or frameshifts in the genome of either *Salmonella typhimurium* and/or *Escherichia coli*. Negative results indicate that under the test conditions, the test substance is not mutagenic in the tested species.

**3. REPORTING****TEST REPORT**

The test report must include the following information:

Solvent/Vehicle:

- justification for choice of solvent/vehicle,
- solubility and stability of the test substance in solvent/vehicle, if known.

Strains:

- strains used,
- number of cells per culture,
- strain characteristics.

Test conditions:

- amount of test substance per plate (mg/plate or µl/plate) with rationale for selection of dose and number of plates per concentration,
- media used,
- type and composition of metabolic activation system, including acceptability criteria,
- treatment procedures.

Results:

- signs of toxicity,
- signs of precipitation,
- individual plate counts,

**▼B**

- the mean number of revertant colonies per plate and standard deviation,
- dose-response relationship, where possible,
- statistical analyses, if any,
- concurrent negative (solvent/vehicle) and positive control data, with ranges, means and standard deviations,
- historical negative (solvent/vehicle) and positive control data with ranges, means and standard deviations.

Discussion of results.

Conclusions.

#### 4. REFERENCES

- (1) Ames, B.N., McCann, J. and Yamasaki E., (1975) Methods for Detecting Carcinogens and Mutagens with the Salmonella/-Mammalian-Microsome Mutagenicity Test. *Mutation Res.*, 31, p. 347-364.
- (2) Maron, D.M. and Ames, B.N., (1983) Revised Methods for the Salmonella Mutagenicity Test. *Mutation Res.*, 113, p. 173-215.
- (3) Gatehouse, D., Haworth, S., Cebula, T., Gocke, E., Kier, L., Matsushima, T., Melcion, C., Nohmi, T., Venitt, S. and Zeiger, E., (1994) Recommendations for the Performance of Bacterial Mutation Assays. *Mutation Res.*, 312, p. 217-233.
- (4) Kier, L.D., Brusick D.J., Auletta, A.E., Von Halle, E.S., Brown, M.M., Simmon, V.F., Dunkel, V., McCann, J., Mortelmans, K., Prival, M., Rao, T.K. and Ray V., (1986) The *Salmonella typhimurium*/Mammalian Microsomal Assay: A Report of the U.S. Environmental Protection Agency Gene-Tox Program. *Mutation Res.*, 168, p. 69-240.
- (5) Yahagi, T., Degawa, M., Seino, Y.Y., Matsushima, T., Nagao, M., Sugimura, T. and Hashimoto, Y., (1975) Mutagenicity of Carcinogen Azo Dyes and their Derivatives, *Cancer Letters*, 1, p. 91-96.
- (6) Matsushima, M., Sugimura, T., Nagao, M., Yahagi, T., Shirai, A. and Sawamura, M., (1980) Factors Modulating Mutagenicity Microbial Tests. In: *Short-term Test Systems for Detecting Carcinogens*. Ed. Norpoth K.H. and Garner, R.C., Springer, Berlin-Heidelberg-New York. p. 273-285.
- (7) Gatehouse, D.G., Rowland, I.R., Wilcox, P., Callender, R.D. and Foster, R., (1980) Bacterial Mutation Assays. In: *Basic Mutagenicity Tests: UKEMS Part 1 Revised*. Ed. D.J. Kirkland, Cambridge University Press, p. 13-61.
- (8) Aeschacher, H.U., Wolleb, U. and Porchet, L., (1987) Liquid Preincubation Mutagenicity Test for Foods. *J. Food Safety*, 8, p. 167-177.
- (9) Green, M.H.L., Muriel, W.J. and Bridges, B.A., (1976) Use of a simplified fluctuation test to detect low levels of mutagens. *Mutation Res.*, 38, p. 33-42.
- (10) Hubbard, S.A., Green, M.H.L., Gatehouse, D. and J.W. Bridges (1984) The Fluctuation Test in Bacteria. In: *Handbook of Mutagenicity Test Procedures*. 2nd Edition. Ed. Kilbey, B.J., Legator, M., Nichols, W. and Ramel, C., Elsevier, Amsterdam-New York-Oxford, p. 141-161.

▼B

- (11) Thompson, E.D. and Melampy, P.J., (1981) An Examination of the Quantitative Suspension Assay for Mutagenesis with Strains of *Salmonella typhimurium*. Environmental Mutagenesis, 3, p. 453-465.
- (12) Araki, A., Noguchi, T., Kato, F. and T. Matsushima (1994) Improved Method for Mutagenicity Testing of Gaseous Compounds by Using a Gas Sampling Bag. Mutation Res., 307, p. 335-344.
- (13) Prival, M.J., Bell, S.J., Mitchell, V.D., Reipert, M.D. and Vaughn, V.L., (1984) Mutagenicity of Benzidine and Benzidine-Congener Dyes and Selected Monoazo Dyes in a Modified Salmonella Assay. Mutation Res., 136, p. 33-47.
- (14) Zeiger, E., Anderson, B.E., Haworth, S., Lawlor, T. and Mortelmans, K., (1992) Salmonella Mutagenicity Tests. V. Results from the Testing of 311 Chemicals. Environ. Mol. Mutagen., 19, p. 2-141.
- (15) Simmon, V., Kauhanen, K. and Tardiff, R.G., (1977) Mutagenic Activity of Chemicals Identified in Drinking Water. In Progress in Genetic Toxicology, D. Scott, B. Bridges and F. Sobels (Eds.) Elsevier, Amsterdam, p. 249-258.
- (16) Hughes, T.J., Simmons, D.M., Monteith, I.G. and Claxton, L.D., (1987) Vaporisation Technique to Measure Mutagenic Activity of Volatile Organic Chemicals in the Ames/-Salmonella Assay. Environmental Mutagenesis, 9, p. 421-441.
- (17) Matsushima, T., Matsumoto, A., Shirai, M., Sawamura, M., and Sugimura, T., (1979) Mutagenicity of the Naturally Occurring Carcinogen Cycasin and Synthetic Methylazoxy Methane Conjugates in *Salmonella typhimurium*. Cancer Res., 39, p. 3780-3782.
- (18) Tamura, G., Gold, C., Ferro-Luzzi, A. and Ames, B.N., (1980) Fecalase: A Model for Activation of Dietary Glycosides to Mutagens by Intestinal Flora. Proc. Natl. Acad. Sci. USA, 77, p. 4961-4965.
- (19) Wilcox, P., Naidoo, A., Wedd, D.J. and Gatehouse, D.G., (1990) Comparison of *Salmonella typhimurium* TA 102 with *Escherichia coli* WP2 Tester strains. Mutagenesis, 5, p. 285-291.
- (20) Matsushima, T., Sawamura, M., Hara, K. and Sugimura, T., (1976) A Safe Substitute for Polychlorinated Biphenyls as an Inducer or Metabolic Activation Systems. In: 'In vitro metabolic Activation in Mutagenesis Testing' Eds. F.J. de Serres et al. Elsevier, North Holland, p. 85-88.
- (21) Elliot, B.M., Combes, R.D., Elcombe, C.R., Gatehouse, D.G., Gibson, G.G., Mackay, J.M. and Wolf, R.C., (1992) Alternatives to Aroclor 1254-induced S9 in *in vitro* Genotoxicity Assays. Mutagenesis, 7, p. 175-177.
- (22) Maron, D., Katzenellenbogen, J. and Ames, B.N., (1981) Compatibility of Organic Solvents with the Salmonella/-Microsome Test. Mutation Res., p. 88343-350.
- (23) Claxton, L.D., Allen, J., Auletta, A., Mortelmans, K., Nestmann, E. and Zeiger, E., (1987) Guide for the *Salmonella typhimurium*/Mammalian Microsome Tests for Bacterial Mutagenicity. Mutation Res. 189, p. 83-91.

**▼B**

- (24) Mahon, G.A.T., Green, M.H.L., Middleton, B., Mitchell, I., Robinson, W.D. and Tweats, D.J., (1989) Analysis of Data from Microbial Colony Assays. In: UKEMS Sub-Committee on Guidelines for Mutagenicity Testing. Part II. Statistical Evaluation of Mutagenicity Test Data. Ed. Kirkland, D.J., Cambridge University Press, p. 28-65.

**▼B****B.15. MUTAGENICITY TESTING AND SCREENING FOR CARCINOGENICITY GENE MUTATION — *SACCHAROMYCES CEREVISIAE*****1. METHOD****1.1. INTRODUCTION**

See General introduction Part B.

**1.2. DEFINITION**

See General introduction Part B.

**1.3. REFERENCE SUBSTANCES**

None.

**1.4. PRINCIPLE OF THE TEST METHOD**

A variety of haploid and diploid strains of the yeast *Saccharomyces cerevisiae* can be used to measure the production of gene mutations induced by chemical agents with and without metabolic activation.

Forward mutation systems in haploid strains, such as the measurement of mutation from red, adenine-requiring mutants (*ade-1*, *ade-2*) to double adenine-requiring white mutants and selective systems such as the induction of resistance to canavaine and cycloheximide, have been utilised.

The most extensively validated reverse mutation system involves the use of the haploid strain XV 185-14C which carries the ochre nonsense mutations *ade 2-1*, *arg 4-17*, *lys 1-1* and *trp 5-48*, which are reversible by base substitution mutagens that induce site specific mutations or ochre suppressor mutations. XV 185-14C also carries the *his 1-7* marker, a missense mutation reverted mainly by second site mutations, and the marker *hom 3-10* which is reverted by frameshift mutagens.

In diploid strains of *S. cerevisiae* the only extensively used strain is D<sub>7</sub> which is homozygous for *ilv 1-92*.

**1.5. QUALITY CRITERIA**

None.

**1.6. DESCRIPTION OF THE TEST METHOD***Preparations*

Solutions of test chemicals and control should be prepared just prior to testing, using an appropriate vehicle. In the case of organic compounds, which are not water soluble, not more than a 2 % solution v/v of organic solvents such as ethanol, acetone or dimethylsulphoxide (DMSO) should be used. The final concentration of the vehicle should not significantly affect cell viability and growth characteristics.



**▼B****Metabolic activation**

Cells should be exposed to test chemicals both in the presence and absence of an appropriate exogenous metabolic activation system.

The most commonly used system is a co-factor supplemented post-mitochondrial fraction from the livers of rodents pre-treated with enzyme inducing agents. The use of other species, tissues, post-mitochondrial fractions, or procedures may also be appropriate for metabolic activation.

*Test conditions***Tester strains**

The haploid strain XV 185-14C and the diploid strain D<sub>7</sub> are the most used in gene mutation studies. Other strains may also be appropriate.

**Media**

Appropriate culture media are used for the determination of survival and mutant numbers.

**Use of negative and positive controls**

Positive, untreated and solvent controls should be performed concurrently. Appropriate positive control chemicals should be used for each specific mutational endpoint.

**Exposure concentration**

At least five adequately spaced concentrations of the test substance should be used. For toxic substances, the highest concentration tested should not reduce survival below 5 to 10 %. Relatively water-insoluble substances should be tested up to their limit of solubility, using appropriate procedures. For freely water-soluble non-toxic substances, the upper concentration should be determined on a case by case basis.

**Incubation conditions**

The plates are incubated four to seven days at 28 to 30 °C in the dark.

**Spontaneous mutation frequencies**

Sub-cultures should be used with spontaneous mutation frequencies within the accepted normal range.

**▼ B****Number of replicates**

At least three replicate plates should be used per concentration for the assay of prototrophs produced by gene mutation and for cell viability. In the case of experiments using markers such as *hom* 3-10 with a low mutation rate, the number of plates used must be increased to provide statistically relevant data.

*Procedure*

Treatment of *S. cerevisiae* strains is usually performed in a liquid test procedure involving either stationary or growing cells. Initial experiments should be carried out on growing cells:  $1-5 \times 10^7$  cells/ml are exposed to the test chemical for up to 18 hours at 28 to 37 °C with shaking; an adequate amount of metabolic activation system is added during treatment when appropriate. At the end of the treatment, cells are centrifuged, washed and seeded upon an appropriate culture medium. After incubation, plates are scored for survival and the induction of gene mutation. If the first experiment is negative, then a second experiment should be carried out using stationary phase cells. If the first experiment is positive it is confirmed in an appropriate independent experiment.

**2. DATA**

Data should be presented in tabular form indicating the number of colonies counted, number of mutants, survival and mutant frequency. All results should be confirmed in an independent experiment. The data should be evaluated using appropriate statistical methods.

**3. REPORTING****3.1. TEST REPORT**

The test report shall, if possible, contain the following information:

- strain used,
- test conditions: stationary phase or growing cells, compositions of media, incubation temperature and duration, metabolic activation system,
- treatment conditions: exposure levels, procedure and duration of treatment, treatment temperature, positive and negative controls,
- number of colonies counted, number of mutants, survival and mutant frequency, dose/response relationship if applicable, statistical evaluation of data,
- discussion of results,
- interpretation of results.

**3.2. EVALUATION AND INTERPRETATION**

See General introduction Part B.

**4. REFERENCES**

See General introduction Part B.

**▼B****B.16. MITOTIC RECOMBINATION — SACCHAROMYCES  
CEREVISIAE****1. METHOD****1.1. INTRODUCTION**

See General introduction Part B.

**1.2. DEFINITION**

See General introduction Part B.

**1.3. REFERENCE SUBSTANCES**

None.

**1.4. PRINCIPLE OF THE TEST METHOD**

Mitotic recombination in *Saccharomyces cerevisiae* can be detected between genes (or more generally between a gene and its centromere) and within genes. The former event is called mitotic crossing-over and generates reciprocal products whereas the latter event is most frequently non-reciprocal and is called gene conversion. Crossing-over is generally assayed by the production of recessive homozygous colonies or sectors produced in a heterozygous strain, whereas gene conversion is assayed by the production of prototrophic revertants produced in an auxotrophic heteroallelic strain carrying two different defective alleles of the same gene. The most commonly used strains for the detection of mitotic gene conversion are D<sub>4</sub> (heteroallelic at *ade 2* and *trp 5*) D<sub>7</sub> (heteroallelic at *trp 5*) BZ<sub>34</sub> (heteroallelic at *arg 4*) and JD1 (heteroallelic at *his 4* and *trp 5*). Mitotic crossing-over producing red and pink homozygous sectors can be assayed in D<sub>5</sub> or in D<sub>7</sub> (which also measures mitotic gene conversion and reverse mutation at *ilv 1-92*) both strains being heteroallelic for complementing alleles of *ade 2*.

**1.5. QUALITY CRITERIA**

None.

**1.6. DESCRIPTION OF THE TEST METHOD***Preparations*

Solutions of test chemicals and control or reference compounds should be prepared just prior to testing, using an appropriate vehicle. With organic compounds that are water insoluble not more than a 2 % solution v/v of organic solvents such as ethanol, acetone or dimethylsulphoxide (DMSO) should be used. The final concentration of the vehicle should not significantly affect cell viability and growth characteristics.

*Metabolic activation*

Cells should be exposed to test chemicals both in the presence and absence of an appropriate exogenous metabolic activation system. The system most commonly used is a co-factor supplemented post-mitochondrial fraction from the livers of rodents pre-treated with enzyme inducing agents. The use of other species, tissues, post-mitochondrial fractions, or procedures may also be appropriate for metabolic activation.

**▼B***Test conditions*

## Tester strains

The most frequently used strains are the diploids D<sub>4</sub>, D<sub>5</sub>, D<sub>7</sub> and JD1. The use of other strains may be appropriate.

## Media

Appropriate culture media are used for the determination of survival and the frequency of mitotic recombination.

## Use of negative and positive controls

Positive, untreated and solvent controls should be performed concurrently. Appropriate positive control chemicals should be used for each specific recombination endpoint.

## Exposure concentrations

At least five adequately spaced concentrations of the test substance should be used. Among the factors to be taken into consideration are cytotoxicity and solubility. The lowest concentration must have no effect on cell viability. For toxic chemicals, the highest concentration tested should not reduce survival below 5 to 10 %. Relatively water-insoluble chemicals should be tested up to the limit of solubility using appropriate procedures. For freely water-soluble non-toxic substances the upper concentration should be determined on a case by case basis.

Cells may be exposed to test chemicals in either the stationary phase or during growth for periods of up to 18 hours. However, for long treatment times cultures should be microscopically inspected for spore formation, the presence of which invalidates the test.

## Incubation conditions

The plates are incubated in the dark for four to seven days at 28 to 30 °C. Plates used for the assay of red and pink homozygous sectors produced by mitotic crossing-over should be kept in a refrigerator (about 4 °C) for a further one to two days before scoring to allow for the development of the appropriate pigmented colonies.

## Spontaneous mitotic recombination frequencies

Sub-cultures should be used with spontaneous mitotic recombination mutation frequencies within the accepted normal range.

**▼B****Number of replicates**

A minimum of three replicate plates should be used per concentration for the assay of prototrophs produced by mitotic gene conversion and for viability. In the case of the assay of recessive homozygosis produced by mitotic crossing-over, the plate number should be increased to provide an adequate number of colonies.

*Procedures*

Treatment of *S. cerevisiae* strains is usually performed in a liquid test procedure involving either stationary or growing cells. Initial experiments should be done on growing cells.  $1-5 \times 10^7$  celles/ml are exposed to the test chemical for up to 18 hours at 28 to 37 °C with shaking; an adequate amount of metabolic activation system is added during treatment when appropriate.

At the end of the treatment, cells are centrifuged, washed and seeded upon appropriate culutre medium. After incubation plates are scored for survival and the induction of mitotic recombination.

If the first experiment is negative, then a second experiment should be carried out using stationary phase cells. If the first experiment is positive it is confirmed in an independent experiment.

**2. DATA**

Data should be presented in tabular form indicating the number of colonies counted, the number of recombinants, survival and the frequency of recombinants.

Results should be confirmed in an independent experiment.

The data should be evaluated using appropriate statistical methods.

**3. REPORTING****3.1. TEST REPORT**

The test report shall, if possible, contain the following information:

- strain used,
- test conditions: stationary phase or growing cells, composition of media, incubation temperature and duration, metabolic activation system,
- treatment conditions: exposure concentration, procedure and duration of treatment, treatment temperature, positive and negative controls,
- number of colonies counted, number of recombinants; survival and recombination frequency, dose/response relationship if applicable, statistical evaluation of data,
- discussion of the results,
- interpretation of the results.

**▼B**

3.2. EVALUATION AND INTERPRETATION

See General introduction Part B.

4. **REFERENCES**

See General introduction Part B.

**▼B****B.17. MUTAGENICITY — *IN VITRO* MAMMALIAN CELL GENE MUTATION TEST****1. METHOD**

This method is a replicate of the OECD TG 476, *In Vitro* Mammalian Cell Gene Mutation Test (1997).

**1.1. INTRODUCTION**

The *in vitro* mammalian cell gene mutation test can be used to detect gene mutations induced by chemical substances. Suitable cell lines include L5178Y mouse lymphoma cells, the CHO, CHO-AS52 and V79 lines of Chinese hamster cells, and TK6 human lymphoblastoid cells (1). In these cell lines the most commonly-used genetic endpoints measure mutation at thymidine kinase (TK) and hypoxanthine-guanine phosphoribosyl transferase (HPRT), and a transgene of xanthine-guanine phosphoribosyl transferase (XPRT). The TK, HPRT and XPRT mutation tests detect different spectra of genetic events. The autosomal location of TK and XPRT may allow the detection of genetic events (e.g. large deletions) not detected at the HPRT locus on X-chromosomes (2)(3)(4)(5)(6).

In the *in vitro* mammalian cell gene mutation test, cultures of established cell lines or cell strains can be used. The cells used are selected on the basis of growth ability in culture and stability of the spontaneous mutation frequency.

Tests conducted *in vitro* generally require the use of an exogenous source of metabolic activation. This metabolic activation system cannot mimic entirely the mammalian *in vivo* conditions. Care should be taken to avoid conditions, which would lead to results not reflecting intrinsic mutagenicity. Positive results, which do not reflect intrinsic mutagenicity may arise from changes in pH, osmolality or high levels of cytotoxicity (7).

This test is used to screen for possible mammalian mutagens and carcinogens. Many compounds that are positive in this test are mammalian carcinogens; however, there is not a perfect correlation between this test and carcinogenicity. Correlation is dependent on chemical class and there is increasing evidence that there are carcinogens that are not detected by this test because they appear to act through other, non genotoxic mechanisms or mechanisms absent in bacterial cells (6).

See also General introduction Part B.

**1.2. DEFINITIONS**

**Forward mutation:** a gene mutation from the parental type to the mutant form which gives rise to an alteration or a loss of the enzymatic activity of the function of the encoded protein.

**Base pair substitution mutagens:** substances, which cause substitution of one or several base pairs in the DNA.

**Frameshift mutagens:** Substances, which cause the addition or deletion of single or multiple base pairs in the DNA molecule.

**▼ B**

**Phenotypic expression time:** a period during which unaltered gene products are depleted from newly mutated cells.

**Mutant frequency:** the number of mutant cells observed divided by the number of viable cells.

**Relative total growth:** increase in cell number over time compared to a control population of cells; calculated as the product of suspension growth relative to the negative control times cloning efficiency relative to negative control.

**Relative suspension growth:** increase in cell number over the expression period relative to the negative control.

**Viability:** the cloning efficiency of the treated cells at the time of plating in selective conditions after the expression period.

**Survival:** the cloning efficiency of the treated cells when plated at the end of the treatment period; survival is usually expressed in relation to the survival of the control cell population.

### 1.3. PRINCIPLE OF THE TEST METHOD

Cells deficient in thymidine kinase (TK) due to the mutation  $TK^{+/-} \rightarrow TK^{-/-}$  are resistant to the cytotoxic effects of the pyrimidine analogue trifluorothymidine (TFT). Thymidine kinase proficient cells are sensitive to TFT, which causes the inhibition of cellular metabolism and halts further cell division. Thus mutant cells are able to proliferate in the presence of TFT, whereas normal cells, which contain thymidine kinase, are not. Similarly, cells deficient in HPRT or XPRT are selected by resistance to 6-thioguanine (TG) or 8-azaguanine (AG). The properties of the test substance should be considered carefully if a base analogue or a compound related to the selective agent is tested in any of the mammalian cell gene mutation tests. For example, any suspected selective toxicity by the test substance for mutant and non-mutant cells should be investigated. Thus, performance of the selection system/agent must be confirmed when testing chemicals structurally related to the selective agent (8).

Cells in suspension or monolayer culture are exposed to the test substance, both with and without metabolic activation, for a suitable period of time and subcultured to determine cytotoxicity and to allow phenotypic expression prior to mutant selection (9)(10)(11)(12)(13). Cytotoxicity is usually determined by measuring the relative cloning efficiency (survival) or relative total growth of the cultures after the treatment period. The treated cultures are maintained in growth medium for a sufficient period of time, characteristic of each selected locus and cell type, to allow near-optimal phenotypic expression of induced mutations. Mutant frequency is determined by seeding known numbers of cells in medium containing the selective agent to detect mutant cells and in medium without selective agent to determine the cloning efficiency (viability). After a suitable incubation time, colonies are counted. The mutant frequency is derived from the number of mutant colonies in selective medium and the number of colonies in non-selective medium.



**▼B**

## 1.4. DESCRIPTION OF THE TEST METHOD

1.4.1. **Preparations**1.4.1.1. *Cells*

A variety of cell types are available for use in this test including subclones of L5178Y, CHO, CHO-AS52, V79 or TK6 cells. Cell types used in this test should have a demonstrated sensitivity to chemical mutagens, a high cloning efficiency and a stable spontaneous mutant frequency. Cells should be checked for mycoplasma contamination and should not be used if contaminated.

The test should be designed to have a predetermined sensitivity and power. The number of cells, cultures and concentrations of test substance used should reflect these defined parameters (14). The minimal number of viable cells surviving treatment and used at each stage in the test should be based on the spontaneous mutation frequency. A general guide is to use a cell number, which is at least 10 times the inverse of the spontaneous mutation frequency. However, it is recommended to utilise at least  $10^6$  cells. Adequate historical data on the cell system used should be available to indicate consistent performance of the test.

1.4.1.2. *Media and culture conditions*

Appropriate culture media, and incubation conditions (culture vessels, temperature, CO<sub>2</sub> concentration, and humidity) should be used. Media should be chosen according to the selective systems and cell type used in the test. It is particularly important that culture conditions should be chosen that ensure optimal growth of cells during the expression period and colony forming ability of both mutant and non-mutant cells.

1.4.1.3. *Preparation of cultures*

Cell are propagated from stock cultures, seeded in culture medium and incubated at 37 °C. Prior to use in this test, cultures may need to be cleansed of pre-existing mutant cells.

1.4.1.4. *Metabolic activation*

Cells should be exposed to the test substance both in the presence and absence of an appropriate metabolic activation system. The most commonly used system is a cofactor-supplemented post-mitochondrial fraction (S9) prepared from the livers of rodents treated with enzyme-inducing agents such as Aroclor 1254 (15)(16)(17)(18) or a combination of phenobarbitone and β-naphthoflavone (19)(20).

The post-mitochondrial fraction is usually used at concentrations in the range from 1-10 % v/v in the final test medium. The choice and condition of a metabolic activation system may depend upon the class of chemical being tested. In some cases it may be appropriate to utilise more than one concentration of post-mitochondrial fraction.

**▼B**

A number of developments, including the construction of genetically engineered cell lines expressing specific activating enzymes, may provide the potential for endogenous activation. The choice of the cell lines used should be scientifically justified (e.g. by the relevance of the cytochrome P450 isoenzyme for the metabolism of the test substance).

**1.4.1.5. Test substance/Preparation**

Solid test substances should be dissolved or suspended in appropriate solvents or vehicles and diluted if appropriate prior to treatment of the cells. Liquid test substances may be added directly to the test systems and/or diluted prior to treatment. Fresh preparations of the test substance should be employed unless stability data demonstrate the acceptability of storage.

**1.4.2. Test conditions****1.4.2.1. Solvent/Vehicle**

The solvent/vehicle should not be suspected of chemical reaction with the test substance and should be compatible with the survival of the cells and the S9 activity. If other than well-known solvent/vehicles are used, their inclusion should be supported by data indicating their compatibility. It is recommended that wherever possible, the use of an aqueous solvent/vehicle be considered first. When testing water-unstable substances, the organic solvents used should be free of water. Water can be removed by adding a molecular sieve.

**1.4.2.2. Exposure concentrations**

Among the criteria to be considered when determining the highest concentration are cytotoxicity, solubility in the test system and changes in pH or osmolality.

Cytotoxicity should be determined with and without metabolic activation in the main experiment using an appropriate indication of cell integrity and growth, such as relative cloning efficiency (survival) or relative total growth. It may be useful to determine cytotoxicity and solubility in a preliminary experiment.

At least four analysable concentrations should be used. Where there is cytotoxicity, these concentrations should cover a range from the maximum to little or no toxicity; this will usually mean that the concentration levels should be separated by no more than a factor between 2 and  $\sqrt{10}$ . If the maximum concentration is based on cytotoxicity then it should result in approximately 10-20 % (but not less than 10 %) relative survival (relative cloning efficiency) or relative total growth. For relatively non-cytotoxic substances, the maximum test concentration should be 5 mg/ml, 5 µl/ml, or 0,01 M, whichever is the lowest.

**▼B**

Relatively insoluble substances should be tested up to or beyond their limit of solubility under culture conditions. Evidence of insolubility should be determined in the final treatment medium to which cells are exposed. It may be useful to assess solubility at the beginning and the end of the treatment, as solubility can change during the course of exposure in the test system due to presence of cells, S9, serum, etc. Insolubility can be detected by using the unaided eye. The precipitate should not interfere with the scoring.

1.4.2.3. *Controls*

Concurrent positive and negative (solvent or vehicle) controls, both with and without metabolic activation should be included in each experiment. When metabolic activation is used, the positive control chemical should be the one that requires activation to give a mutagenic response.

Examples of positive control substances include:

Metabolic activation condition	Locus	Substance	CAS No	EINECS No
Absence of exogenous metabolic activation	HPRT	Ethyl methanesulphonate	62-50-0	200-536-7
		Ethyl nitrosourea	759-73-9	212-072-2
	TK (small and large colonies)	Methyl methanesulphonate	66-27-3	200-625-0
	XPRT	Ethyl methanesulphonate	62-50-0	200-536-7
		Ethyl nitrosourea	759-73-9	212-072-2
	Presence of exogenous metabolic activation	HPRT	3-Methylcholanthrene	56-49-5
N-Nitrosodimethylamine			62-75-9	200-549-8
7,12-Dimethylbenzanthracene			57-97-6	200-359-5
TK (small and large colonies)		Cyclophosphamide	50-18-0	200-015-4
		Cyclophosphamide monohydrate	6055-19-2	
		Benzo[ <i>a</i> ]pyrene	50-32-8	200-028-5
		3-Methylcholanthrene	56-49-5	200-276-5
XPRT		N-Nitrosodimethylamine (for high levels of S-9)	62-75-9	200-549-8
		Benzo[ <i>a</i> ]pyrene	50-32-8	200-028-5

Other appropriate positive control reference substances may be used, e.g. if a laboratory has a historical data base on 5-Bromo 2'-deoxyuridine (CAS n. 59-14-3, EINECS n. 200-415-9), this reference substance could be used as well. The use of chemical class-related positive control chemicals should be considered, when available.

**▼B**

Negative controls, consisting of solvent or vehicle alone in the treatment medium, and treated in the same way as the treatment groups, should be included. In addition, untreated controls should also be used unless there are historical control data demonstrating that no deleterious or mutagenic effects are induced by the chosen solvent.

**1.4.3. Procedure****1.4.3.1. *Treatment with the test substance***

Proliferating cells should be exposed to the test substance both with and without metabolic activation. Exposure should be for a suitable period of time (usually three to six hours is effective). Exposure time may be extended over one or more cell cycles.

Either duplicate or single treated cultures may be used at each concentration tested. When single cultures are used, the number of concentrations should be increased to ensure an adequate number of cultures for analysis (e.g. at least eight analysable concentrations). Duplicate negative (solvent) control cultures should be used.

Gaseous or volatile substances should be tested by appropriate methods, such as in sealed culture vessels (21)(22).

**1.4.3.2. *Measurement of survival, viability and mutant frequency***

At the end of the exposure period, cells are washed and cultured to determine survival and to allow for expression of the mutant phenotype. Measurement of cytotoxicity by determining the relative cloning efficiency (survival) or relative total growth of the cultures is usually initiated after the treatment period.

Each locus has a defined minimum time requirement to allow near optimal phenotypic expression of newly induced mutants (HPRT and XPRT require at least six to eight days, and TK at least two days). Cells are grown in medium with and without selective agent(s) for determination of numbers of mutants and cloning efficiency, respectively. The measurement of viability (used to calculate mutant frequency) is initiated at the end of the expression time by plating in non-selective medium.

If the test substance is positive in the L5178Y TK<sup>+/−</sup> test, colony sizing should be performed on at least one of the test cultures (the highest positive concentration) and on the negative and positive controls. If the test substance is negative in the L5178Y TK<sup>+/−</sup> test, colony sizing should be performed on the negative and positive controls. In studies using TK6TK<sup>+/−</sup>, colony sizing may also be performed.

**▼B****2. DATA****2.1. TREATMENT OF RESULTS**

Data should include cytotoxicity and viability determination, colony counts and mutant frequencies for the treated and control cultures. In the case of a positive response in the L5178Y TK<sup>+/−</sup> test, colonies are scored using the criteria of small and large colonies on at least one concentration of the test substance (highest positive concentration) and on the negative and positive control. The molecular and cytogenetic nature of both large and small colony mutants has been explored in detail (23)(24). In the TK<sup>+/−</sup> test, colonies are scored using the criteria of normal growth (large) and slow growth (small) colonies (25). Mutant cells that have suffered the most extensive genetic damage have prolonged doubling times and thus form small colonies. This damage typically ranges in scale from the losses of the entire gene to karyotypically visible chromosome aberrations. The induction of small colony mutants has been associated with chemicals that induce gross chromosome aberrations (26). Less seriously affected mutant cells grow at rates similar to the parental cells and form large colonies.

Survival (relative cloning efficiencies) or relative total growth should be given. Mutant frequency should be expressed as number of mutant cells per number of surviving cells.

Individual culture data should be provided. Additionally, all data should be summarised in tabular form.

There is no requirement for verification of a clear positive response. Equivocal results should be clarified by further testing preferably using modification of experimental conditions. Negative results need to be confirmed on a case-by-case basis. In those cases where confirmation of negative results is not considered necessary, justification should be provided. Modification of study parameters to extend the range of conditions assessed should be considered in follow-up experiments for either equivocal or negative results. Study parameters that might be modified include the concentration spacing and the metabolic activation conditions.

**2.2. EVALUATION AND INTERPRETATION OF RESULTS**

There are several criteria for determining a positive result, such as a concentration-related increase or a reproducible increase in mutant frequency. Biological relevance of the results should be considered first. Statistical methods may be used as an aid in evaluating the test results. Statistical significance should not be the only determining factor for a positive response.

A test substance for which the results do not meet the above criteria is considered non-mutagenic in this system.

**▼B**

Although most studies will give clearly positive or negative results, in rare cases the data set will preclude making a definite judgement about the activity of the test substance. Results may remain equivocal or questionable regardless of the number of times the experiment is repeated.

Positive results from the *in vitro* mammalian cell gene mutation test indicate that the test substance induces gene mutations in the cultured mammalian cells used. A positive concentration response that is reproducible is most meaningful. Negative results indicate that, under the test conditions, the test substance does not induce gene mutations in the cultured mammalian cells used.

**3. REPORTING****TEST REPORT**

The test report must include the following information:

**Solvent/Vehicle:**

- justification for choice of vehicle/solvent,
- solubility and stability of the test substance in solvent/vehicle, if known,

**Cells:**

- type and source of cells,
- number of cell cultures,
- number of cell passages, if applicable,
- methods for maintenance of cell culture, if applicable,
- absence of mycoplasma.

**Test conditions:**

- rationale for selection of concentrations and number of cultures including, e.g. cytotoxicity data and solubility limitations, if available,
- composition of media, CO<sub>2</sub> concentration,
- concentration of test substance,
- volume of vehicle and test substance added,
- incubation temperature,
- incubation time,
- duration of treatment,
- cell density during treatment,
- type and composition of metabolic activation system, including acceptability criteria,
- positive and negative controls,

**▼B**

- length of expression period (including number of cells seeded, and subcultures and feeding schedules, if appropriate),
- selective agents,
- criteria for considering tests as positive, negative or equivocal,
- methods used to enumerate numbers of viable and mutant cells.
- definition of colonies of which size and type are considered (including criteria for 'small' and 'large' colonies, as appropriate).

## Results:

- signs of toxicity,
- signs of precipitation,
- data on pH and osmolality during the exposure to the test substance, if determined,
- colony size if scored for at least negative and positive controls,
- laboratory's adequacy to detect small colony mutants with the L5178Y TK<sup>±</sup> system, where appropriate,
- dose-response relationship, where possible,
- statistical analyses, if any,
- concurrent negative (solvent/vehicle) and positive control data,
- historical negative (solvent/vehicle) and positive control data with ranges, means and standard deviations,
- mutant frequency.

## Discussion of results.

## Conclusions.

4. **REFERENCES**

- (1) Moore, M.M., DeMarini, D.M., DeSerres, F.J. and Tindall, K.R. (Eds.), (1987) Banbury Report 28: Mammalian Cell Mutagenesis, Cold Spring Harbor Laboratory, New York.
- (2) Chu, E.H.Y. and Malling, H.V., (1968) Mammalian Cell Genetics. II. Chemical Induction of Specific Locus Mutations in Chinese Hamster Cells *In Vitro*, Proc. Natl. Acad. Sci., USA, 61, p. 1306-1312.
- (3) Liber, H.L. and Thilly, W.G., (1982) Mutation Assay at the Thymidine Kinase Locus in Diploid Human Lymphoblasts. Mutation Res., 94, p. 467-485.
- (4) Moore, M.M., Harington-Brock, K., Doerr, C.L. and Dearfield, K.L., (1989) Differential Mutant Quantitation at the Mouse Lymphoma TK and CHO HGPRT Loci. Mutagenesis, 4, p. 394-403.

**▼B**

- (5) Aaron, C.S. and Stankowski, Jr.L.F., (1989) Comparison of the AS52/XPRT and the CHO/HPRT Assays: Evaluation of Six Drug Candidates. *Mutation Res.*, 223, p. 121-128.
- (6) Aaron, C.S., Bolcsfoldi, G., Glatt, H.R., Moore, M., Nishi, Y., Stankowski, Jr.L.F., Theiss, J. and Thompson, E., (1994) Mammalian Cell Gene Mutation Assays Working Group Report. Report of the International Workshop on Standardisation of Genotoxicity Test Procedures. *Mutation Res.*, 312, p. 235-239.
- (7) Scott, D., Galloway, S.M., Marshall, R.R., Ishidate, M., Brusick, D., Ashby, J. and Myhr, B.C., (1991) Genotoxicity Under Extreme Culture Conditions. A report from ICPEMC Task Group 9. *Mutation Res.*, 257, p. 147-204.
- (8) Clive, D., McCuen, R., Spector, J.F.S., Piper, C. and Mavourmin, K.H., (1983) Specific Gene Mutations in L5178Y Cells in Culture. A Report of the U.S. Environmental Protection Agency Gene-Tox Program. *Mutation Res.*, 115, p. 225-251.
- (9) Li, A.P., Gupta, R.S., Heflich, R.H. and Wasson, J.S., (1988) A Review and Analysis of the Chinese Hamster Ovary/Hypoxanthine Guanine Phosphoribosyl Transferase System to Determine the Mutagenicity of Chemical Agents: A Report of Phase III of the U.S. Environmental Protection Agency Gene-Tox Program. *Mutation Res.*, 196, p. 17-36.
- (10) Li, A.P., Carver, J.H., Choy, W.N., Hsie, A.W., Gupta, R.S., Loveday, K.S., O'Neill, J.P., Riddle, J.C., Stankowski, L.F. Jr. and Yang, L.L., (1987) A Guide for the Performance of the Chinese Hamster Ovary Cell/Hypoxanthine-Guanine Phosphoribosyl Transferase Gene Mutation Assay. *Mutation Res.*, 189, p. 135-141.
- (11) Liber, H.L., Yandell, D.W. and Little, J.B., (1989) A Comparison of Mutation Induction at the TK and HPRT Loci in Human Lymphoblastoid Cells: Quantitative Differences are Due to an Additional Class of Mutations at the Autosomal TK Locus. *Mutation Res.*, 216, p. 9-17.
- (12) Stankowski, L.F. Jr., Tindall, K.R. and Hsie, A.W., (1986) Quantitative and Molecular Analyses of Ethyl Methanesulphonate -and ICR 191- Induced Molecular Analyses of Ethyl Methanesulphonate -and ICR 191-Induced Mutation in AS52 Cells. *Mutation Res.*, 160, p. 133-147.
- (13) Turner, N.T., Batson, A.G. and Clive, D., (1984) Procedures for the L5178Y/TK<sup>+/+</sup> — TK<sup>+/+</sup> Mouse Lymphoma Cell Mutagenicity Assay. In: Kilbey, B.J. et al (eds.) *Handbook of Mutagenicity Test Procedures*, Elsevier Science Publishers, New York, p. 239-268.
- (14) Arlett, C.F., Smith, D.M., Clarke, G.M., Green, M.H.L., Cole, J., McGregor, D.B. and Asquith, J.C., (1989) Mammalian Cell Gene Mutation Assays Based upon Colony Formation. In: *Statistical Evaluation of Mutagenicity Test Data*, Kirkland, D.J., Ed., Cambridge University Press, p. 66-101.
- (15) Abbondandolo, A., Bonatti, S., Corti, G., Fiorio, R., Loprieno, N. and Mazzaccaro, A., (1977) Induction of 6-Thioguanine-Resistant Mutants in V79 Chinese Hamster Cells by Mouse-Liver Microsome-Activated Dimethylnitrosamine. *Mutation Res.*, 46, p. 365-373.
- (16) Ames, B.N., McCann, J. and Yamasaki, E., (1975) Methods for Detecting Carcinogens and Mutagens with the Salmonella/Mammalian-Microsome Mutagenicity Test. *Mutation Res.*, 31, p. 347-364.



**▼B**

- (17) Clive, D., Johnson, K.O., Spector, J.F.S., Batson, A.G. and Brown M.M.M., (1979) Validation and Characterisation of the L5178Y/TK<sup>+/-</sup> Mouse Lymphoma Mutagen Assay System. *Mutat. Res.*, 59, p. 61-108.
- (18) Maron, D.M. and Ames, B.N., (1983) Revised Methods for the Salmonella Mutagenicity Test. *Mutation Res.*, 113, p. 173-215.
- (19) Elliott, B.M., Combes, R.D., Elcombe, C.R., Gatehouse, D.G., Gibson, G.G., Mackay, J.M. and Wolf, R.C., (1992) Alternatives to Aroclor 1254-Induced S9 in *In Vitro* Genotoxicity Assays. *Mutagenesis*, 7, p. 175-177.
- (20) Matsushima, T., Sawamura, M., Hara, K. and Sugimura, T., (1976) A Safe Substitute for Polychlorinated Biphenyls as an Inducer of Metabolic Activation Systems. In: *In vitro* Metabolic Activation in Mutagenesis Testing, de Serres, F.J., Fouts, J.R., Bend, J.R. and Philpot, R.M. (eds), Elsevier, North-Holland, p. 85-88.
- (21) Krahn, D.F., Barsky, F.C. and McCooey, K.T., (1982) CHO/HGPRT Mutation Assay: Evaluation of Gases and Volatile Liquids. In: Tice, R.R., Costa, D.L., Schaich, K.M. (eds). *Genotoxic Effects of Airborne Agents*. New York, Plenum, p. 91-103.
- (22) Zamora, P.O., Benson, J.M., Li, A.P. and Brooks, A.L., (1983) Evaluation of an Exposure System Using Cells Grown on Collagen Gels for Detecting Highly Volatile Mutagens in the CHO/HGPRT Mutation Assay. *Environmental Mutagenesis*, 5, p. 795-801.
- (23) Applegate, M.L., Moore, M.M., Broder, C.B., Burrell, A. and Hozier, J.C., (1990) Molecular Dissection of Mutations at the Heterozygous Thymidine Kinase Locus in Mouse Lymphoma Cells. *Proc. Natl. Acad. Sci. USA*, 87, p. 51-55.
- (24) Moore, M.M., Clive, D., Hozier, J.C., Howard, B.E., Batson, A.G., Turner, N.T. and Sawyer, J., (1985) Analysis of Trifluorothymidine-Resistant (TFT<sup>r</sup>) Mutants of L5178Y/TK<sup>+/-</sup> Mouse Lymphoma Cells. *Mutation Res.*, 151, p. 161-174.
- (25) Yandell, D.W., Dryja, T.P. and Little, J.B., (1990) Molecular Genetic Analysis of Recessive Mutations at a Heterozygous Autosomal Locus in Human Cells. *Mutation Res.*, 229, p. 89-102.
- (26) Moore, M.M. and Doerr, C.L., (1990) Comparison of Chromosome Aberration Frequency and Small-Colony TK-Deficient Mutant Frequency in L5178Y/TK<sup>+/-</sup> - 3.7.2C Mouse Lymphoma Cells. *Mutagenesis*, 5, p. 609-614.

**▼B****B.18. DNA DAMAGE AND REPAIR — UNSCHEDULED DNA SYNTHESIS — MAMMALIAN CELLS *IN VITRO*****1. METHOD****1.1. INTRODUCTION**

See General introduction Part B.

**1.2. DEFINITION**

See General introduction Part B.

**1.3. REFERENCE SUBSTANCES**

None.

**1.4. PRINCIPLE OF THE TEST METHOD**

The Unscheduled DNA Synthesis (UDS) test measures the DNA repair synthesis after excision and removal of a stretch of DNA containing the region of damage induced by chemical and physical agents. The test is based on the incorporation of tritium labelled thymidine ( $^3\text{H-TdR}$ ) into the DNA of mammalian cells, which are not in the S phase of the cell cycle. The uptake of  $^3\text{H-TdR}$  may be determined by autoradiography or by liquid scintillation counting (LSC) of DNA from the treated cells. Mammalian cells in culture, unless primary rat hepatocytes are used, are treated with the test agent with and without an exogenous metabolic activation system. UDS may also be measured in *in vivo* systems.

**1.5. QUALITY CRITERIA**

None.

**1.6. DESCRIPTION OF THE TEST METHOD***Preparations*

Test chemicals and control or reference substances should be prepared in growth medium or dissolved or suspended in appropriate vehicles and then further diluted in growth medium for use in the assay. The final concentration of the vehicle should not affect cell viability.

Primary cultures of rat hepatocytes, human lymphocytes or established cell lines (e.g. human diploid fibroblasts) may be used in the assay.

Cells should be exposed to the test chemical both in the presence and absence of an appropriate metabolic activation system.

*Test conditions***Number of cultures**

At least two cell cultures for autoradiography and six cultures (or less if scientifically justified) for LSC UDS determinations are necessary for each experimental point.

**▼ B****Use of negative and positive controls**

Concurrent positive and negative (untreated and/or vehicle) controls with and without metabolic activation should be included in each experiment.

Examples of positive controls for the rat hepatocyte assay include 7,12- dimethylbenzathracene (7,12- DMBA) or 2-acetylaminofluorene (2-AAF). In the case of established cell lines 4-nitroquinoline-N-oxide (4-NQO) is an example of a positive control for both the autoradiographic and LSC assays performed without metabolic activation; N-dimethylnitrosamine is an example of a positive control compound when metabolic activation systems are used.

**Exposure concentrations**

Multiple concentrations of the test substance over a range adequate to define the response should be used. The highest concentration should elicit some cytotoxic effects. Relatively water-insoluble compounds should be tested up to the limit of solubility. For freely water-soluble non-toxic chemicals, the upper test chemical concentration should be determined on a case-by-case basis.

**Cells**

Appropriate growth media, CO<sub>2</sub> concentration, temperature and humidity should be used in maintaining cultures. Established cell lines should be periodically checked for *Mycoplasma* contamination.

**Metabolic activation**

A metabolic activation system is not used with primary hepatocyte cultures. Established cell lines and lymphocytes are exposed to test substance both in the presence and absence of an appropriate metabolic activation system.

*Procedure***Preparation of cultures**

Established cell lines are generated from stock cultures (e.g. by trypsinisation or by shaking off), seeded in culture vessels at appropriate density, and incubated at 37 °C.

Short-term cultures of rat hepatocytes are established by allowing freshly dissociated hepatocytes in an appropriate medium to attach themselves to the growing surface.

Human lymphocyte cultures are set up using appropriate techniques.

**▼B****Treatment of the cultures with the test substance****Primary rat hepatocytes**

Freshly isolated rat hepatocytes are treated with the test substance in a medium containing  $^3\text{H}$ -TdR for an appropriate length of time. At the end of the treatment period, medium should be drained off the cells, which are then rinsed fixed and dried. Slides should be dipped in autoradiographic emulsion (alternative stripping film may be used), exposed, developed, stained and counted.

**Established cell lines and lymphocytes**

Autoradiographic techniques: cell cultures are exposed to the test substance for appropriate durations followed by treatment with  $^3\text{H}$ -TdR. The times will be governed by the nature of the substance, the activity of metabolising systems and the type of cells. To detect the peak of UDS,  $^3\text{H}$ -TdR should be added either simultaneously with the test substance or within a few minutes after exposure to the test substance. The choice between these two procedures will be influenced by possible interactions between test substance and  $^3\text{H}$ -TdR. In order to discriminate between UDS and semi-conservative DNA replication, the latter can be inhibited, for example, by the use of an arginine-deficient medium, low serum content or by hydroxyurea in the culture medium.

LSC measurements of UDS: prior to treatment with test substance, entry of cells into S-phase should be blocked as described above; cells should then be exposed to test chemical as described for autoradiography. At the end of the incubation period, DNA should be extracted from the cells and the total DNA content, and the extent of  $^3\text{H}$ -TdR, incorporation determined.

It should be noted that, where human lymphocytes are used in the above techniques, the suppression of semi-conservative DNA replication is unnecessary in unstimulated cultures.

*Analysis**Autoradiographic determinations*

In determining UDS in cells in culture, S-phase nuclei are not counted. At least 50 cells per concentration should be counted. Slides should be coded before counting. Several widely separated random fields should be counted on each slide. The amount of  $^3\text{H}$ -TdR incorporation in the cytoplasm should be determined by counting three nucleus-sized areas in the cytoplasm of each cell counted.

**LSC determinations**

An adequate number of cultures should be used at each concentration and in the controls in LSC UDS determinations.

**▼ B**

All results should be confirmed in an independent experiment.

**2. DATA**

Data should be presented in tabular form.

**2.1. AUTORADIOGRAPHIC DETERMINATIONS**

The extent of  $^3\text{H}$ -TdR incorporation in the cytoplasm and the number of grains found over the cell nucleus should be recorded separately.

Mean, median and mode may be used to describe the distribution of the extent of  $^3\text{H}$ -TdR incorporation in the cytoplasm and the number of grains per nucleus.

**2.2. LSC DETERMINATIONS**

For LSC determinations,  $^3\text{H}$ -TdR incorporation should be reported as dpm/ $\mu\text{g}$  DNA. The mean dpm/ $\mu\text{g}$  DNA with standard deviation may be used to describe the distribution of incorporation.

Data should be evaluated using appropriate statistical methods.

**3. REPORTING****3.1. TEST REPORT**

The test report shall, if possible, contain the following information:

- cells used, density and passage number at time of treatment, number of cell cultures,
- methods used for maintenance of cell cultures including medium, temperature and  $\text{CO}_2$  concentration,
- test substance, vehicle, concentrations and rationale for selection of concentrations used in the assay,
- details of metabolic activation systems,
- treatment schedule,
- positive and negative controls,
- autoradiographic technique used,
- procedures used to block entry of cells into S-phase,
- procedures used for DNA extraction and determination of total DNA content in LSC determination,
- dose/response relationship, where possible,
- statistical evaluation,
- discussion of results,
- interpretation of results.

**▼B**

3.2. VALUATION AND INTERPRETATION

See General introduction Part B.

4. **REFERENCES**

See General introduction Part B.

**▼B****B.19. SISTER CHROMATID EXCHANGE ASSAY *IN VITRO*****1. METHOD****1.1. INTRODUCTION**

See General introduction Part B.

**1.2. DEFINITION**

See General introduction Part B.

**1.3. REFERENCE SUBSTANCES**

None.

**1.4. PRINCIPLE OF THE TEST METHOD**

The Sister Chromatid Exchange (SCE) assay is a short-term test for the detection of reciprocal exchanges of DNA between two sister chromatids of a duplicating chromosome. SCEs represent the interchange of DNA replication products at apparently homologous loci. The exchange process presumably involves DNA breakage and reunion, although little is known about its molecular basis. Detection of SCEs requires some means of differentially labelling sister chromatids and this can be achieved by incorporation of bromodeoxyuridine (BrdU) into chromosomal DNA for two cell cycles.

Mammalian cells *in vitro* are exposed to the test chemical with and without a mammalian exogenous metabolic activation system, if appropriate, and cultured for two rounds of replication in BrdU-containing medium. After treatment with a spindle inhibitor (e.g. colchicine) to accumulate cells in a metaphase-like stage of mitosis (c-metaphase), cells are harvested and chromosome preparations are made.

**1.5. QUALITY CRITERIA**

None.

**1.6. DESCRIPTION OF THE TEST METHOD****1.6.1. Preparations**

- Primary cultures, (human lymphocytes) or established cell lines (e.g. Chinese hamster ovary cells) may be used in the assay. Cell lines should be checked for *Mycoplasma* contamination,
- appropriate culture media and incubation conditions (e.g. temperature, culture vessels, CO<sub>2</sub> concentration and humidity) should be used,
- test substances may be prepared in culture media or dissolved or suspended in appropriate vehicles prior to treatment of the cells. The final concentration of a vehicle in the culture system should not significantly affect, cell viability or growth rate and effects on SCE frequency should be monitored by a solvent control,

**▼B**

- cells should be exposed to the test substance both in the presence and absence of an exogenous mammalian metabolic activation system. Alternatively, where cell types with intrinsic metabolic activity are used, the rate and nature of the activity should be appropriate to the chemical class being tested.

1.6.2. *Test conditions***Number of cultures**

At least duplicate cultures should be used for each experimental point.

**Use of negative and positive controls**

Positive controls, using both a direct acting compound and a compound requiring metabolic activation should be included in each experiment; a vehicle control should also be used.

The following are examples of substances which might be used as positive controls:

- direct acting compound:
  - ethylmethanesulphonate,
- indirect acting compound:
  - cyclophosphamide.

When appropriate, an additional positive control of the same chemical class as the chemical under test may be included.

**Exposure concentrations**

At least three adequately spaced concentrations of the test substance should be used. The highest concentration should give rise to a significant toxic effect but must still allow adequate cell replication to occur. Relatively water-insoluble substances should be tested up to the limit of solubility using appropriate procedures. For freely water-soluble non-toxic substances the upper test substance concentration should be determined on a case-by-case basis.

1.6.3. *Procedure***Preparation of cultures**

Established cell lines are generated from stock cultures (e.g. by trypsinisation or by shaking off), seeded in culture vessels at appropriate density and incubated at 37 °C. For monolayer cultures, the number of cells per culture vessel should be adjusted so that the cultures are not much more than 50 % confluent at the time of harvest. Alternatively, cells may be used in suspension culture. Human lymphocyte cultures are set up from heparinized blood, using appropriate techniques, and incubated at 37 °C.



**▼ B****Treatment**

Cells in an exponential stage of growth are exposed to the test substance for a suitable period of time; in most cases one to two hours may be effective, but the treatment time may be extended up to two complete cell cycles in certain cases. Cells without sufficient intrinsic metabolic activity should be exposed to the test chemical in the presence and absence of an appropriate metabolic activation system. At the end of the exposure period, cells are washed free of test substance and cultured for two rounds of replication in the presence of BrdU. As an alternative procedure cells may be exposed simultaneously to the test chemical and BrdU for the complete culture time of two cell cycles.

Human lymphocyte cultures are treated while they are in a semisynchronous condition.

Cells are analysed in their second post-treatment division, ensuring that the most sensitive cell cycle stages have been exposed to the chemical. All cultures to which BrdU is added should be handled in darkness or in dim light from incandescent lamps up to the harvesting of cells in order to minimise photolysis of BrdU-containing DNA.

**Harvesting of cells**

Cell cultures are treated with a spindle inhibitor (e.g. colchicine) one to four hours prior to harvesting. Each culture is harvested and processed separately for the preparation of chromosomes.

**Chromosome preparation and staining**

Chromosome preparations are made by standard cytogenetic techniques. Staining of slides to show SCEs can be performed by several techniques, (e.g. the fluorescence plus Giemsa method).

**Analysis**

The number of cells analysed should be based on the spontaneous control frequency of SCE. Usually, at least 25 well-spread metaphases per culture are analysed for SCEs. Slides are coded before analysis. In human lymphocytes only metaphases containing 46 centromeres are analysed. In established cell lines only metaphases containing  $\pm 2$  centromeres of the modal number are analysed. It should be stated whether or not centromeric switch of label is, scored as an SCE. The results should be confirmed in an independent experiment.

**2. DATA**

Data should be presented in tabular form. The number of SCEs for each metaphase and the number of SCEs per chromosome for each metaphase should be listed separately for all treated and control cultures.

The data should be evaluated using appropriate statistical methods.

**▼B****3. REPORTING****3.1. TEST REPORT**

The test report shall, if possible, contain the following information:

- cells used, methods of maintenance of cell culture,
- test conditions: composition of media, CO<sub>2</sub> concentration, concentration of test substance, vehicle used, incubation temperature, treatment time, spindle inhibitor used, its concentration and the duration of treatment with it, type of mammalian activation system used, positive and negative controls,
- number of cell cultures per experimental point,
- details of the technique used for slide preparation,
- number of metaphases analysed (data given separately for each culture),
- mean number of SCE per cell and per chromosome (data given separately for each culture),
- criteria for scoring SCE,
- rationale for dose selection,
- dose-response relationship, if applicable,
- statistical evaluation,
- discussion of results,
- interpretation of results.

**3.2. EVALUATION AND INTERPRETATION**

See General introduction Part B.

**4. REFERENCES**

See General introduction Part B.

**▼B****B.20. SEX-LINKED RECESSIVE LETHAL TEST IN *DROSOPHILA MELANOGASTER*****1. METHOD****1.1. INTRODUCTION**

See General introduction Part B.

**1.2. DEFINITION**

See General introduction Part B.

**1.3. REFERENCE SUBSTANCES**

None.

**1.4. PRINCIPLES OF THE TEST METHOD**

The sex-linked recessive lethal (SLRL) test using *Drosophila melanogaster* detects the occurrence of mutations, both point mutations and small deletions, in the germ line of the insect. This test is a forward mutation assay capable of screening for mutations at about 800 loci on the X-chromosome; this represents about 80 % of all X-chromosomal loci. The X-chromosome represents approximately one-fifth of the entire haploid genome.

Mutations in the X-chromosome of *Drosophila melanogaster* are phenotypically expressed in males carrying the mutant gene. When the mutation is lethal in the hemizygous condition, its presence is inferred from the absence of one class of male offspring out of the two that are normally produced by a heterozygous female. The SLRL test takes advantage of these facts by means of specially marked and arranged chromosomes.

**1.5. QUALITY CRITERIA**

None.

**1.6. DESCRIPTION OF THE TEST METHOD***Preparations***Stocks**

Males of a well-defined wild-type stock and females of the Muller-5 stock may be used. Other appropriately marked female stocks with multiple inverted X-chromosomes may also be used.

**Test substance**

Test substances should be dissolved in water. Substances, which are insoluble in water may be dissolved or suspended in appropriate vehicles (e.g. a mixture of ethanol and Tween-60 or 80), then diluted in water or saline prior to administration. Dimethylsulphoxide (DMSO) should be avoided as a vehicle.

**▼B****Number of animals**

The test should be designed with a predetermined sensitivity and power. The spontaneous mutant frequency observed in the appropriate control will influence strongly the number of treated chromosomes that must be analysed.

**Route of administration**

Exposure may be oral, by injection or by exposure to gases or vapours. Feeding of the test substance may be done in sugar solution. When necessary, substances may be dissolved in a 0,7 % NaCl solution and injected into the thorax or abdomen.

**Use of negative and positive controls**

Negative (vehicle) and positive controls should be included. However, if appropriate laboratory historical control data are available, no concurrent controls are needed.

**Exposure levels**

Three exposure levels should be used. For a preliminary assessment one exposure level of the test substance may be used, that exposure level being either the maximum tolerated concentration or that producing some indication of toxicity. For non-toxic substances exposure to the maximum practicable concentration should be used.

**Procedure**

Wild-type males (three to five days old) are treated with the test substance and mated individually to an excess of virgin females from the Muller-5 stock or from another appropriately marked (with multiple inverted X-chromosomes) stock. The females are replaced with fresh virgins every two to three days to cover the entire germ cell cycle. The offspring of these females are scored for lethal effects corresponding to the effects on mature sperm, mid or late-stage spermatids, early spermatids, spermatocytes and spermatogonia at the time of treatment.

Heterozygous F<sub>1</sub> females from the above crosses are allowed to mate individually (i.e. one female per vial) with their brothers. In the F<sub>2</sub> generation, each culture is scored for the absence of wild-type males. If a culture appears to have arisen from an F<sub>1</sub> female carrying a lethal in the parental X-chromosome (i.e. no males with the treated chromosome are observed) daughters of that female with the same genotype should be tested to ascertain whether the lethality is repeated in the next generation.

**2. DATA**

Data should be tabulated to show the number of X-chromosomes tested, the number of non-fertile males and the number of lethal chromosomes at each exposure concentration and for each mating period for each male treated. Numbers of clusters of different sizes per male should be reported. These results should be confirmed in a separate experiment.

**▼B**

Appropriate statistical methods should be used in evaluation sex-linked recessive lethal tests. Clustering of recessive lethals originating from one male should be considered and evaluated in an appropriate statistical manner.

**3. REPORTING****3.1. TEST REPORT**

The test report shall, if possible, contain the following information:

- stock: *Drosophila* stocks or strains used, age of insects, number of males treated, number of sterile males, number of F2 cultures established, number of F2 cultures without progeny, number of chromosomes carrying a lethal detected at each germ cell stage,
- criteria for establishing the size of treated groups,
- test conditions. detailed description of treatment and sampling schedule, exposure levels, toxicity data, negative (solvent) and positive controls, if appropriate,
- criteria for scoring lethal mutations,
- exposure/effect relationship where possible,
- evaluation of data,
- discussion of results,
- interpretation of results.

**3.2. EVALUATION AND INTERPRETATION**

See General introduction Part B.

**4. REFERENCES**

See General introduction Part B.

**▼ B****B.21. IN VITRO MAMMALIAN CELL TRANSFORMATION TESTS****1. METHOD****1.1. INTRODUCTION**

See General introduction Part B.

**1.2. DEFINITION**

See General introduction Part B.

**1.3. REFERENCE SUBSTANCES**

None.

**1.4. PRINCIPLE OF THE TEST METHOD**

Mammalian cell culture systems may be used to detect phenotypic changes *in vitro* induced by chemical substances associated with malignant transformation *in vivo*. Widely used cells include C3H10T<sub>1/2</sub>, 3T3, SHE, Fischer rat and the tests rely on changes in cell morphology, focus formation or changes in anchorage dependence in semi-solid agar. Less widely used systems exist which detect other physiological or morphological changes in cells following exposure to carcinogenic chemicals. None of the *in vitro* test endpoints has an established mechanistic link with cancer. Some of the test systems are capable of detecting tumour promoters. Cytotoxicity may be determined by measuring the effect of the test material on colony-forming abilities (cloning efficiency) or growth rates of the cultures. The measurement of cytotoxicity is to establish that exposure to the test chemical has been toxicologically relevant but cannot be used to calculate transformation: frequency in all assays since some may involve prolonged incubation and/or replating.

**1.5. QUALITY CRITERIA**

None.

**1.6. DESCRIPTION OF THE TEST METHOD***Preparations***C e l l s**

A variety of cell lines or primary cells are available depending on the transformation test being used. The investigator must ensure that the cells in the test being performed exhibit the appropriate phenotypic change after exposure to known carcinogens and that the test, in the investigator's laboratory, is of proven and documented validity and reliability.

**M e d i u m**

Media and experimental conditions should be used that are appropriate to the transformation assay in use.

**▼B****Test substance**

Test substances may be prepared in culture media or dissolved or suspended in appropriate vehicles prior to treatment of the cells. The final concentration of the vehicle in the culture system should not affect cell viability, growth rate or transformation incidence.

**Metabolic activation**

Cells should be exposed to the test substance both in the presence and absence of an appropriate metabolic activation system. Alternatively, when cell types are used that possess intrinsic metabolic activity, the nature of the activity should be known to be appropriate to the chemical class being tested.

*Test conditions***Use of negative and positive controls**

Positive controls, using both a direct-acting compound and a compound requiring metabolic activation should be included in each experiment; a negative (vehicle) control should also be used.

The following are examples of substances, which might be used as positive controls:

## — Direct-acting chemicals:

— Ethylmethanesulphonate,

—  $\beta$ -propiolactone,

## — Compounds requiring metabolic activation:

— 2-acetylaminofluorene,

— 4-dimethylaminoazobenzene,

— 7,12-dimethylbenzanthracene.

When appropriate, an additional positive control of the same chemical class as the compound under test should be included.

**Exposure concentrations**

Several concentrations of the test substance should be used. These concentrations should yield a concentration-related toxic effect, the highest concentration producing a low level of survival and the survival in the lowest concentration being approximately the same as that in the negative control. Relatively water-insoluble substances should be tested up to the limit of solubility using appropriate procedures. For freely water-soluble non-toxic substances the upper test substance concentration should be determined on a case-by-case basis.

**▼ B***Procedure*

Cells should be exposed for a suitable period of time depending on the test system in use, and this may involve re-dosing accompanied by a change of medium (and if necessary, fresh metabolic activation mixture) if exposure is prolonged. Cells without sufficient intrinsic metabolic activity should be exposed to the test substance in the presence and absence of an appropriate metabolic activation system. At the end of the exposure period, cells are washed free of test substance and cultured under conditions appropriate for the appearance of the transformed phenotype being monitored and the incidence of transformation determined. All results are confirmed in an independent experiment.

**2. DATA**

Data should be presented in tabular form and may take a variety of forms according to the assay being used e.g. plate counts, positive plates or numbers of transformed cells. Where appropriate, survival should be expressed as a percentage of control levels and transformation frequency expressed as the number of transformants per number of survivors. Data should be evaluated using appropriate statistical methods.

**3. REPORTING****3.1. TEST REPORT**

The test report shall, if possible, contain the following information:

- cell type used, number of cell cultures, methods for maintenance of cell cultures,
- test conditions: concentration of test substance, vehicle used, incubation time, duration and frequency of treatment, cell density during treatment, type of exogenous metabolic activation system used, positive and negative controls, specification of phenotype being monitored, selective system used (if appropriate), rationale for dose selection,
- method used to enumerate viable and transformed cells,
- statistical evaluation,
- discussion of results,
- interpretation of results.

**3.2. EVALUATION AND INTERPRETATION**

See General introduction Part B.

**4. REFERENCES**

See General introduction Part B.



**▼B****B.22. RODENT DOMINANT LETHAL TEST****1. METHOD****1.1. INTRODUCTION**

See General introduction Part B.

**1.2. DEFINITION**

See General introduction Part B.

**1.3. REFERENCE SUBSTANCES**

None.

**1.4. PRINCIPLE OF THE TEST METHOD**

Dominant lethal effects cause embryonic or foetal death. Induction of dominant lethals by exposure to a chemical substance indicates that the substance has affected germinal tissue of the test species. It is generally accepted that dominant lethals are due to chromosomal damage (structural and numerical anomalies). Embryonic death if females are treated may also be the result of toxic effects.

Generally, male animals are exposed to the test compound and mated to untreated virgin females. The various germ cell stages can be tested separately by the use of sequential mating intervals. The increase of dead implants per female in the treated group over the dead implants per female in the control group reflects the post-implantational loss. Pre-implantational loss can be estimated based on corpora lutea counts or by comparing the total implants per female in treated and control groups. The total dominant lethal effect is the sum of pre- and post-implantational loss. The calculation of the total dominant lethal effect is based on comparison of the live implants per female in the test group to the live implants per female in the control group. A reduction in the number of implants at certain intervals may be the result of cell killing (i.e. of spermatozoa and/or spermatogonia).

**1.5. QUALITY CRITERIA**

None.

**1.6. DESCRIPTION OF THE TEST METHOD***Preparations*

When possible, test substances should be dissolved or suspended in isotonic saline. Chemicals insoluble in water may be dissolved or suspended in appropriate vehicles. The vehicle used should neither interfere with the test chemical nor produce toxic effects. Fresh preparations of the test chemical should be employed.

**▼ B***Test conditions*

## Route of administration

The test compound should generally be administered only once. Based on toxicological information a repeated treatment schedule can be employed. The usual routes of administration are oral intubation or intraperitoneal injection. Other routes of administration may be appropriate.

## Experimental animals

Rats or mice are recommended as the test species. Healthy fully sexually mature animals are randomised and assigned to treatment and control groups.

## Number and sex

An adequate number of treated males should be used, taking into account the spontaneous variation of the biological character being evaluated. The number chosen should be based on the pre-determined sensitivity of detection and power of significance. For example in a typical test, the number of males in each dose group should be sufficient to provide between 30 and 50 pregnant females per mating interval.

## Use of negative and positive controls

Generally concurrent positive and negative (vehicle) controls should be included in each experiment. When acceptable positive control results are available from experiments conducted recently in the same laboratory these results can be used instead of a concurrent positive control. Positive control substances should be used at an appropriate low dose (e.g. MMS, intraperitoneally, at 10 mg/-kilogram) to demonstrate the test sensitivity.

## Dose levels

Normally, three dose levels should be used. The high dose should produce signs of toxicity or reduced fertility in the treated animals. In certain cases a single high dose level may be sufficient.

## Limit test

Non-toxic substances should be tested at 5 g/kilogram on a single administration or at 1 g/kilogram/day on repeated administration.

*Procedure*

Several treatment schedules are available. Single administration of the test substance is the most widely used. Other treatment schedules may be used.

Individual males are mated sequentially to one or two untreated virgin females at appropriate intervals after treatment. Females should be left with the males for at least the duration of one oestrous cycle or until mating has occurred as determined by the presence of sperm in the vagina or by the presence of a vaginal plug.

**▼B**

The number of matings following treatment is governed by the treatment schedule and should ensure that all germ cell stages are sampled after treatment.

Females are sacrificed in the second half of pregnancy and uterine contents are examined to determine the number of dead and live implants. The ovaries may be examined to determine the number of corpora lutea.

**2. DATA**

Data should be tabulated to show the number of males, the number of pregnant females, and the number of non-pregnant females. Results of each mating, including the identity of each male and female, should be reported individually. For each female, week of mating, dose level received by the males, the frequencies of live implants and of dead implants should be recorded.

The calculation of the total dominant lethal effect is based on comparison of the live implants per female in the test group to the live implants per female in the control group. The ratio of dead to live implants from the treated group compared to the same ratio from the control group is analysed to indicate the post-implantation loss.

If the data are recorded as early and late deaths, the tables should make that clear. If pre-implantation loss is estimated, it should be reported. Pre-implantation loss can be calculated as a discrepancy between the number of corpora lutea and the number of implants or as a reduction in the average number of implants per uterus in comparison with control matings.

Data are evaluated using appropriate statistical methods.

**3. REPORTING****3.1. TEST REPORT**

The test report shall, if possible, contain the following information:

- species, strain, age and weights of animals used, number of animals of each sex in experimental and control groups,
- test substance, vehicle, dose levels tested and rationale for dose selection, negative and positive controls, toxicity data,
- route and treatment schedule,
- mating schedule,
- method used to determine that mating has occurred,
- time of sacrifice,
- criteria for scoring dominant lethals,
- dose/response relationship, if applicable,

**▼B**

- statistical evaluation,
- discussion of results,
- interpretation of results.

3.2. **EVALUATION AND INTERPRETATION**

See General introduction Part B.

4. **REFERENCES**

See General introduction Part B.

**▼B****B.23. MAMMALIAN SPERMATOGENIAL CHROMOSOME ABERRATION TEST****1. METHOD**

This method is a replicate of the OECD TG 483, Mammalian Spermatogonial Chromosome Aberration Test (1997).

**1.1. INTRODUCTION**

The purpose of the *in vivo* mammalian spermatogonial chromosome aberration test is to identify those substances that cause structural chromosome aberrations in mammalian spermatogonial cells (1)(2)(3)(4)(5). Structural aberrations may be of two types, chromosome or chromatid. With the majority of chemical mutagens, induced aberrations are of the chromatid type, but chromosome-type aberrations also occur. This method is not designed to measure numerical aberrations and is not routinely used for that purpose. Chromosome mutations and related events are the cause of many human genetic diseases.

This test measures chromosome events in spermatogonial germ cells and is, therefore, expected to be predictive of induction of inheritable mutations in germ cells.

Rodents are routinely used in this test. This *in vivo* cytogenetic test detects chromosome aberrations in spermatogonial mitoses. Other target cells are not the subject of this method.

To detect chromatid-type aberrations in spermatogonial cells, the first mitotic cell division following treatment should be examined before these lesions are lost in subsequent cell divisions. Additional information from treated spermatogonial stem cells can be obtained by meiotic chromosome analysis for chromosome-type aberrations at diakinesis-metaphase I when the treated cells become spermatocytes.

This *in vivo* test is designed to investigate whether somatic cell mutagens are also active in germ cells. In addition, the spermatogonial test is relevant to assessing mutagenicity hazard in that it allows consideration of factors of *in vivo* metabolism, pharmacokinetics and DNA-repair processes.

A number of generations of spermatogonia are present in the testis with a spectrum of sensitivity to chemical treatment. Thus, the aberrations detected represent an aggregate response of treated spermatogonial cell populations, with the more numerous differentiated spermatogonial cells predominating. Depending on their position within the testis, different generations of spermatogonia may or may not be exposed to the general circulation, because of the physical and physiological Sertoli cell barrier and the blood-testis barrier.

If there is evidence that the test substance, or a reactive metabolite, will not reach the target tissue, it is not appropriate to use this test.

See also General introduction Part B.

**▼ B**

## 1.2. DEFINITIONS

**Chromatid-type aberration:** structural chromosome damage expressed as breakage of single chromatids or breakage and reunion between chromatids.

**Chromosome-type aberration:** structural chromosome damage expressed as breakage, or breakage and reunion, of both chromatids at an identical site.

**Gap:** an achromatic lesion smaller than the width of one chromatid, and with minimum misalignment of the chromatids.

**Numerical aberration:** a change in the number of chromosomes from the normal number characteristic of the animals utilised.

**Polyploidy:** a multiple of the haploid chromosome number ( $n$ ) other than the diploid number (i.e.  $3n$ ,  $4n$  and so on).

**Structural aberration:** a change in chromosome structure detectable by microscopic examination of the metaphase stage of cell division, observed as deletions, intrachanges or interchanges.

## 1.3. PRINCIPLE OF THE TEST METHOD

Animals are exposed to the test substance by an appropriate route of exposure and are sacrificed at appropriate times after treatment. Prior to sacrifice, animals are treated with a metaphase-arresting substance (e.g. Colcemid® or colchicine). Chromosome preparations are then made from germ cells and stained, and metaphase cells are analysed for chromosome aberrations.

## 1.4. DESCRIPTION OF THE TEST METHOD

## 1.4.1. Preparations

1.4.1.1. *Selection of animal species*

Male Chinese hamsters and mice are commonly used. However, males of other appropriate mammalian species may be used. Commonly used laboratory strains of young healthy adult animals should be employed. At the commencement of the study the weight variation of animals should be minimal and not exceed  $\pm 20$  % of the mean weight.

1.4.1.2. *Housing and feeding conditions*

General conditions referred in the General introduction to Part B are applied although the aim for humidity should be 50-60 %.

1.4.1.3. *Preparation of the animals*

Healthy young adult males are randomly assigned to the control and treatment groups. Cages should be arranged in such a way that possible effects due to cage placement are minimised. The animals are identified uniquely. The animals are acclimated to the laboratory conditions for at least five days prior to the start of the study.

**▼ B**1.4.1.4. *Preparation of doses*

Solid test substances should be dissolved or suspended in appropriate solvents or vehicles and diluted, if appropriate, prior to dosing of the animals. Liquid test substances may be dosed directly or diluted prior to dosing. Fresh preparations of the test substance should be employed unless stability data demonstrate the acceptability of storage.

1.4.2. **Test conditions**1.4.2.1. *Solvent/Vehicle*

The solvent/vehicle should not produce toxic effects at the dose levels used and should not be suspected of chemical reaction with the test substance. If other than well-known solvents/vehicles are used, their inclusion should be supported by data indicating their compatibility. It is recommended that wherever possible, the use of an aqueous solvent/vehicle should be considered first.

1.4.2.2. *Controls*

Concurrent positive and negative (solvent/vehicle) controls should be included in each test. Except for treatment with the test substance, animals in the control groups should be handled in an identical manner to animals in the treated groups.

Positive controls should produce structural chromosome aberrations *in vivo* in spermatogonial cells when administered at exposure levels expected to give a detectable increase over background.

Positive control doses should be chosen so that the effects are clear but do not immediately reveal the identity of the coded slides to the reader. It is acceptable that the positive control be administered by a route different from the test substance and sampled at only a single time. In addition, the use of chemical class-related positive control chemicals may be considered, when available. Examples of positive control substances include:

Substance	CAS No	Einecs No
Cyclophosphamide	50-18-0	200-015-4
Cyclophosphamide monohydrate	6055-19-2	
Cyclohexylamine	108-91-8	203-629-0
Mitomycin C	50-07-7	200-008-6
Monomeric acrylamide	79-06-1	201-173-7
Triethylenemelamine	51-18-3	200-083-5

Negative controls, treated with solvent or vehicle alone, and otherwise treated in the same way as the treatment groups, should be included for every sampling time, unless acceptable inter-animal variability and frequency of cells with chromosome aberrations are demonstrated by historical control data. In addition, untreated controls should also be used unless there are historical or published control data demonstrating that no deleterious or mutagenic effects are induced by the chosen solvent/vehicle.

**▼ B**

## 1.5. PROCEDURE

1.5.1. **Number of animals**

Each treated and control group must include at least five analysable males.

1.5.2. **Treatment schedule**

Test substances are preferably administered once or twice (i.e. as a single treatment or as two treatments). Test substances may also be administered as a split dose, i.e. two treatments on the same day separated by no more than a few hours, to facilitate administering a large volume of material. Other dose regimens should be scientifically justified.

In the highest dose group two sampling times after treatment are used. Since cell cycle kinetics can be influenced by the test substance, one early and one late sampling time are used around 24 and 48 hours after treatment. For doses other than the highest dose, a sampling time of 24 hours or 1,5 cell cycle length after treatment should be taken, unless another sampling time is known to be more appropriate for detection of effects (6).

In addition, other sampling times may be used. For example in the case of chemicals, which may induce chromosome lagging, or may exert S-independent effects, earlier sampling times may be appropriate (1).

The appropriateness of a repeated treatment schedule needs to be identified on a case-by-case basis. Following a repeated treatment schedule the animals should then be sacrificed 24 hours (1,5 cell cycle length) after the last treatment. Additional sampling times may be used where appropriate.

Prior to sacrifice, animals are injected intraperitoneally with an appropriate dose of a metaphase arresting substance (e.g. Colcemid® or colchicine). Animals are sampled at an appropriate interval thereafter. For mice this interval is approximately three to five hours, for Chinese hamsters this interval is approximately four to five hours.

1.5.3. **Dose levels**

If a range finding study is performed because there are no suitable data available, it should be performed in the same laboratory, using the same species, strain and treatment regimen to be used in the main study (7). If there is toxicity, three dose levels are used for the first sampling time. These dose levels should cover a range from the maximum to little or no toxicity. At the later sampling time only the highest dose needs to be used. The highest dose is defined as the dose producing signs of toxicity such that higher dose levels, based on the same dosing regimen, would be expected to produce lethality.



**▼B**

Substances with specific biological activities at low non-toxic doses (such as hormones and mitogens) may be exceptions to the dose-setting criteria and should be evaluated on a case-by-case basis. The highest dose may also be defined as a dose that produces some indication of toxicity in the spermatogonial cells (e.g. a reduction in the ratio of spermatogonial mitoses to first and second meiotic metaphases; this reduction should not exceed 50 %).

**1.5.4. Limit test**

If a test at one dose level of at least 2000 mg/kg body weight/day using a single treatment, or as two treatments on the same day, produces no observable toxic effects, and if genotoxicity would not be expected based upon data from structurally related substances, then a full study using three dose levels may not be considered necessary. Expected human exposure may indicate the need for a higher dose level to be used in the limit test.

**1.5.5. Administration of doses**

The test substance is usually administered by gavage using a stomach tube or a suitable intubation cannula, or by intraperitoneal injection. Other routes of exposure may be acceptable where they can be justified. The maximum volume of liquid that can be administered by gavage or injection at one time depends on the size of the test animal. The volume should not exceed 2 ml/100 g body weight. The use of volumes higher than these must be justified. Except for irritating or corrosive substances, which will normally reveal exacerbated effects with higher concentrations, variability in test volume should be minimised by adjusting the concentration to ensure a constant volume at all dose levels.

**1.5.6. Chromosome preparation**

Immediately after sacrifice, cell suspensions are obtained from one or both testes, exposed to hypotonic solution and fixed. The cells are then spread on slides and stained.

**1.5.7. Analysis**

For each animal at least 100 well-spread metaphase should be analysed (i.e. a minimum of 500 metaphases per group). This number could be reduced when high numbers of aberrations are observed. All slides, including those of positive and negative controls, should be independently coded before microscopic analysis. Since fixation procedures often result in the breakage of a proportion of metaphases with loss of chromosomes, the cells scored should contain a number of centromeres equal to the number  $2n \pm 2$ .

**2. DATA****2.1. TREATMENT OF RESULTS**

Individual animal data should be presented in a tabular form. The experimental unit is the animal. For each individual animal the number of cells with structural chromosome aberrations and the number of chromosome aberrations per cell should be evaluated. Different types of structural chromosome aberrations should be listed with their numbers and frequencies for treated and control groups. Gaps are recorded separately and reported but generally not included in the total aberration frequency.

**▼B**

If mitosis as well as meiosis is observed, the ratio of spermatogonial mitoses to first and second meiotic metaphases should be determined as a measure of cytotoxicity for all treated and negative control animals in a total sample of 100 dividing cells per animal to establish a possible cytotoxic effect. If only mitosis is observed, the mitosis index should be determined in at least 1 000 cells for each animal.

**2.2. EVALUATION AND INTERPRETATION OF RESULTS**

There are several criteria for determining a positive result, such as a dose-related increase in the relative number of cells with chromosome aberrations or a clear increase in the number of cells with aberrations in a single dose at a single sampling time. Biological relevance of the results should be considered first. Statistical methods may be used as an aid in evaluating the test results (8). Statistical significance should not be the only determining factor for a positive response. Equivocal results should be clarified by further testing preferably using a modification of experimental conditions.

A test substance for which the results do not meet the above criteria is considered non-mutagenic in this test.

Although most experiments will give clearly positive or negative results, in rare cases the data set will preclude making a definite judgement about the activity of the test substance. Results may remain equivocal or questionable regardless of the number of times the experiment is repeated.

Positive results from the *in vivo* spermatogonial chromosome aberration test indicate that the test substance induces structural chromosome aberrations in the germ cells of the species tested. Negative results indicate that, under the test conditions, the test substance does not induce chromosome aberrations in the germ cells of the species tested.

The likelihood that the test substance or its metabolites reach the target tissue should be discussed.

**3. REPORTING****TEST REPORT**

The test report must include the following information:

Solvent/Vehicle:

- justification for choice of vehicle,
- solubility and stability of the test substance in solvent/vehicle, if known.

Test animals:

- species/strain used,
- number and age of animals,
- source, housing conditions, diet, etc.,
- individual weight of the animals at the start of the test, including body weight range, mean and standard deviation for each group.

**▼B**

## Test conditions:

- data from range finding study, if conducted,
- rationale for dose level selection,
- rationale for route of administration,
- details of test substance preparation,
- details of the administration of the test substance,
- rationale for sacrifice times,
- conversion from diet/drinking water test substance concentration (ppm) to the actual dose (mg/kg body weight/day), if applicable.
- details of food and water quality,
- detailed description of treatment and sampling schedules,
- methods for measurement of toxicity,
- identity of metaphase arresting substance, its concentration and duration of treatment,
- methods of slide preparation,
- criteria for scoring aberrations,
- number of cells analysed per animal,
- criteria for considering studies as positive, negative or equivocal.

## Results:

- signs of toxicity,
- mitotic index,
- ratio of spermatogonial mitoses cells to first and second meiotic metaphases,
- type and number of aberrations, given separately for each animal,
- total number of aberrations per group,
- number of cells with aberrations per group,
- dose-response relationship, if possible,
- statistical analyses, if any,
- concurrent negative control data,
- historical negative control data with ranges, means and standard deviations,
- concurrent positive control data,
- changes in ploidy, if seen.

## Discussion of results.

## Conclusions.

**▼B**

4.

**REFERENCES**

- (1) Adler, I.D., (1986) Clastogenic Potential in Mouse Spermatogonia of Chemical Mutagens Related to their Cell-Cycle Specifications. In: Genetic Toxicology of Environmental Chemicals, Part B: Genetic Effects and Applied Mutagenesis, Ramel, C., Lambert, B. and Magnusson, J. (Eds.) Liss, New York, p. 477-484.
- (2) Adler, I.D., (1984) Cytogenetic tests in Mammals. In: Mutagenicity Testing: a Practical Approach. Ed. S. Venitt and J.M. Parry. IRL Press, Oxford, Washington DC, p. 275-306.
- (3) Evans, E.P., Breckon, G. and Ford, C.E., (1964) An Air-Drying Method for Meiotic Preparations from Mammalian Testes. Cytogenetics and Cell Genetics, 3, p. 289-294.
- (4) Richold, M., Ashby, J., Chandley, A., Gatehouse, D.G. and Henderson, L., (1990) *In Vivo* Cytogenetic Assays, In: D.J. Kirkland (Ed.) Basic Mutagenicity Tests, UKEMS Recommended Procedures. UKEMS Subcommittee on Guidelines for Mutagenicity Testing. Report. Part I revised. Cambridge University Press, Cambridge, New York, Port Chester, Melbourne, Sydney, p. 115-141.
- (5) Yamamoto, K. and Kikuchi, Y., (1978) A New Method for Preparation of Mammalian Spermatogonial Chromosomes. Mutation Res., 52, p. 207-209.
- (6) Adler, I.D., Shelby M.D., Bootman, J., Favor, J., Generoso, W., Pacchierotti, F., Shibuya, T. and Tanaka N., (1994) International Workshop on Standardisation of Genotoxicity Test Procedures. Summary Report of the Working Group on Mammalian Germ Cell Tests. Mutation Res., 312, p. 313-318.
- (7) Fielder, R.J., Allen, J.A., Boobis, A.R., Botham, P.A., Doe, J., Esdaile, D.J., Gatehouse, D.G., Hodson-Walker, G., Morton, D.B., Kirkland, D.J. and Richold, M., (1992) Report of British Toxicology Society/UK Environmental Mutagen Society Working group: Dose setting in *In Vivo* Mutagenicity Assays. Mutagenesis, 7, p. 313-319.
- (8) Lovell, D.P., Anderson, D., Albanese, R., Amphlett, G.E., Clare, G., Ferguson, R., Richold, M., Papworth, D.G. and Savage, J.R.K., (1989) Statistical Analysis of *In Vivo* Cytogenetic Assays In: D.J. Kirkland (Ed.) Statistical Evaluation of Mutagenicity Test Data. UKEMS Sub-Committee on Guidelines for Mutagenicity Testing, report, Part III. Cambridge University Press, Cambridge, New York, Port Chester, Melbourne, Sydney, p. 184-232.

**▼B****B.24. MOUSE SPOT TEST****1. METHOD****1.1. INTRODUCTION**

See General introduction Part B.

**1.2. DEFINITION**

See General introduction Part B.

**1.3. REFERENCE SUBSTANCES**

None.

**1.4. PRINCIPLE OF THE TEST METHOD**

This is an *in vivo* test in mice in which developing embryos are exposed to the chemicals. The target cells in the developing embryos are melanoblasts, and the target genes are those which control the pigmentation of the coat hairs. The developing embryos are heterozygous for a number of these coat colour genes. A mutation in, or loss of (by a variety of genetic events), the dominant allele of such a gene in a melanoblast results in the expression of the recessive phenotype in its descendant cells, constituting a spot of changed colour in the coat of the resulting mouse. The number of offspring with these spots, mutations, are scored and their frequency is compared with that among offspring resulting from embryos treated with the solvent only. The mouse spot test detects presumed somatic mutations in foetal cells.

**1.5. QUALITY CRITERIA**

None.

**1.6. DESCRIPTION OF THE TEST METHOD***Preparations*

When possible, test substances are dissolved or suspended in isotonic saline. Chemicals insoluble in water are dissolved or suspended in appropriate vehicles. The vehicle used should neither interfere with the test chemical nor produce toxic effects. Fresh preparations of the test chemical should be used.

*Experimental animals*

Mice of the T strain (nonagouti, a/a; chinchilla, pink eye, c<sup>h</sup>p/c<sup>h</sup>p; brown, b/b; dilute, short ear, d se/d se; piebald spotting, s/s) are mated either with the HT strain (pallid, nonagouti, brachypody, pa a bp/pa a bp; leaden fuzzy, ln fz/ln fz; pearl pe/pe) or with C57BL (nonagouti, a/a). Other appropriate crosses such as between NMRI (nonagouti, a/a; albino, c/c) and DBA (nonagouti, a/a; brown, b/b; dilute d/d) may be used provided they produce nonagouti offspring.

**▼B****Number and sex**

Sufficient pregnant females are treated to provide an appropriate number of surviving offspring for each dose level used. The appropriate sample size is governed by the number of spots observed in the treated mice and the scale of the control data. A negative result is acceptable only when at least 300 offspring from females treated with the highest dose have been scored.

**Use of negative and positive controls**

Concurrent control data from mice treated with the vehicle only (negative controls) should be available. Historical control data from the same laboratory may be pooled to increase the sensitivity of the test provided they are homogeneous. Positive control data recently obtained in the same laboratory from treatment with a chemical known to show mutagenicity by this test should be available if no mutagenicity of the test chemical is detected.

**Route of administration**

The usual routes of administration are oral intubation or intraperitoneal injection of the pregnant females. Treatment by inhalation or other routes of administration are used when appropriate.

**Dose levels**

At least two dose levels are used including one showing signs of toxicity or reduced litter size. For non-toxic chemicals exposure to the maximum practicable dose should be used.

*Procedure*

A single treatment is normally given on day 8,9 or 10 of pregnancy, counting as day 1 the day on which the vaginal plug is first observed. These days correspond to 7,25, 8,25 and 9,25 days after conception. Successive treatments over these days may be used.

**Analysis**

The offspring are coded and scored for spots between three and four weeks after birth. Three classes of spots are distinguished:

- (a) white spots within 5 mm of the mid-ventralline which are presumed to result from cell killing (WMVS);
- (b) yellow, agouti-like, spots associated with mammae, genitalia, throat, axillary and inguinal areas and on the mid-forehead, which are presumed to result from misdifferentiation (MDS); and
- (c) pigmented and white spots randomly distributed on the coat which are presumed to result from somatic mutations (RS).

**▼ B**

All three classes are scored but only the last, RS, is of genetic relevance. Problems of distinguishing between MDS and RS may be solved by fluorescence microscopy of sample hairs.

Obvious gross morphological abnormalities of the offspring should be noted.

**2. DATA**

The data are presented as the total number of offspring scored and the number having one or more presumed somatic mutation spots. Treatment and negative control data are compared by appropriate methods. Data are also presented on a per-litter basis.

**3. REPORTING****3.1. TEST REPORT**

The test report shall, if possible, contain the following information:

- the strains used in the cross,
- the number of pregnant females in the experimental and control groups,
- the average litter size in the experimental and control groups at birth and at weaning,
- the dose level(s) of the test chemical,
- the solvent used,
- the day of pregnancy of which treatment was given,
- the route of treatment,
- the total number of offspring scored, and the number with WMVS, MDS and RS in the experimental and control groups,
- gross morphological abnormalities,
- dose/response relationship of RS when possible,
- statistical evaluation,
- discussion of results,
- interpretation of results.

**3.2. EVALUATION AND INTERPRETATION**

See General introduction Part B.

**4. REFERENCES**

See General introduction Part B.

**▼B****B.25. MOUSE HERITABLE TRANSLOCATION****1. METHOD****1.1. INTRODUCTION**

See General introduction Part B.

**1.2. DEFINITION**

See General introduction Part B.

**1.3. REFERENCE SUBSTANCES**

None.

**1.4. PRINCIPLE OF THE TEST METHOD**

The mouse heritable translocation test detects structural and numerical chromosome changes in mammalian germ cells as recovered in first generation progeny. The types of chromosome changes detected are reciprocal translocations and, if female progeny are included, X-chromosome loss. Carriers of translocations and XO-females show reduced fertility which is used to select F<sub>1</sub> progeny for cytogenetic analysis. Complete sterility is caused by certain types of translocations (X-autosome and c-t type). Translocations are cytogenetically observed in meiotic cells at diakinesis-metaphase I of male individuals, either F<sub>1</sub> males or male offspring of F<sub>1</sub> females. The XO-females are cytogenetically identified by the presence of only 39 chromosomes in bone marrow mitoses.

**1.5. QUALITY CRITERIA**

None.

**1.6. DESCRIPTION OF THE TEST METHOD***Preparations*

The test chemicals are dissolved in isotonic saline. If insoluble they are dissolved or suspended in appropriate vehicles. Freshly prepared solutions of the test compound are employed. If a vehicle is used to facilitate dosing, it must not interfere with the test compound or produce toxic effects.

*Route of administration*

Routes of administration are usually oral intubation or intraperitoneal injection. Other routes of administration may be appropriate.

*Experimental animals*

For the ease of breeding and cytological verification these experiments are performed with mice. No specific mouse strain is required. However, the average litter-size of the strain should be greater than eight and be relatively constant.

Healthy sexually mature animals are used.



**▼B****Number of animals**

The number of animals necessary depends upon the spontaneous translocation frequency and the minimal rate of induction required for a positive result.

The test is usually performed by analyses of male F<sub>1</sub> progeny. At least 500 male F<sub>1</sub> progeny should be tested per dose group. If female F<sub>1</sub> progeny are included, 300 males and 300 females are required.

**Use of negative and positive controls**

Adequate control data, derived from concurrent and historic control should be available. When acceptable positive control results are available from experiments conducted recently in the same laboratory these results can be used instead of a concurrent positive control.

**Dose levels**

One dose level is tested, usually the highest dose associated with the production of minimal toxic effects, but without affecting reproductive behaviour or survival. To establish a dose/response relationship two additional lower doses are required. For non-toxic chemicals exposure to the maximum practicable dose should be used.

*Procedure***Treatment and mating**

Two treatment schedules are available. Single administration of the test substance is most widely used. Administration of the test substance on seven days per week for 35 days may also be used. The number of matings following treatment is governed by the treatment schedule and should ensure that all treated germ cell stages are sampled. At the end of the mating period females are caged individually. When females give birth, the date, litter size and sex of progeny are recorded. All male progeny are weaned and all female progeny are discarded unless they are included in the experiment.

**Testing for translocation heterozygosity**

One of two possible methods is used:

- fertility testing of F<sub>1</sub> progeny and subsequent verification of possible translocation carriers by cytogenetic analysis,
- cytogenetic analysis of all male F<sub>1</sub> progeny without prior selection by fertility testing.

**(a) Fertility testing**

Reduced fertility of an F<sub>1</sub> individual can be established by litter size observation and/or analysis of uterine contents of female mates.

Criteria for determining normal and reduced fertility must be established for the mouse strain used.

**▼B**

Litter size observation: F<sub>1</sub> males to be tested are caged individually with females either from the same experiment or from the colony. Cages are inspected daily beginning 18 days after mating. Litter size and sex of the F<sub>2</sub> progeny are recorded at birth and litters are discarded thereafter. If female F<sub>1</sub> progeny are tested the F<sub>2</sub> progeny of small litters are kept for further testing. Female translocation carriers are verified by cytogenetic analysis of a translocation in any of their male offspring. XO-females are recognised by the change in sex ratio among their progeny from 1:1 to 1:2 males *versus* females. In a sequential procedure, normal F<sub>1</sub> animals are eliminated from further testing if the first F<sub>2</sub> litter reaches or exceeds a predetermined normal value, otherwise a second or third F<sub>2</sub> litter is observed.

F<sub>1</sub> animals that cannot be classified as normal after observation of up to three F<sub>2</sub> litters are either tested further by analysis of uterine contents of female mates or directly subjected to cytogenetic analysis.

Analysis of uterine contents: the reduction in litter size of translocation carriers is due to embryonic death so that a high number of dead implants is indicative of the presence of a translocation in the animal under test. F<sub>1</sub> males to be tested are mated to two to three females each. Conception is established by daily inspection for vaginal plugs in the morning. Females are sacrificed 14 to 16 days later and living and dead implants in their uteri are recorded.

(b) Cytogenetic analysis

Testes preparations are made by the air-drying technique. Translocation carriers are identified by the presence of multivalent configurations at diakinesis-metaphase I in primary spermatocytes. Observation of at least two cells with multivalent association constitutes the required evidence that the tested animal is a translocation carrier.

If no breeding selection has been performed all F<sub>1</sub> males are inspected cytogenetically. A minimum of 25 diakinesis-metaphase I cells per male must be scored microscopically. Examination of mitotic metaphases, in spermatogonia or bone-marrow, is required in F<sub>1</sub> males with small testes and meiotic breakdown before diakinesis or from F<sub>1</sub> female XO suspects. The presence of an unusually long and/or short chromosome in each of 10 cells is evidence for a particular male sterile translocation (c-t type). Some X-autosome translocations that cause male sterility may only be identified by banding analysis of mitotic chromosomes. The presence of 39 chromosomes in all of 10 mitoses is evidence for an XO condition in a female.

2. **DATA**

Data are presented in tabular form.

The mean litter size and sex ratio from parental matings at birth and weaning are reported for each mating interval.

**▼B**

For fertility assessment of F<sub>1</sub> animals, the mean litter size of all normal matings and the individual litter sizes of F<sub>1</sub> translocation carriers are presented. For analysis of uterine contents, the mean number of living and dead implants of normal matings and the individual numbers of living and dead implants for each mating of F<sub>1</sub> translocation carriers are reported.

For cytogenetic analysis of diakinesis-metaphase I, the numbers of types of multivalent configurations and the total number of cells are listed for each translocation carrier.

For sterile F<sub>1</sub> individuals, the total number of matings and the duration of the mating period are reported. Testes weights and cytogenetic analysis details are given.

For XO females, the mean litter size, sex ratio of F<sub>1</sub> progeny and cytogenetic analysis results are reported.

Where possible F<sub>1</sub> translocation carriers are preselected by fertility tests, the tables have to include information on how many of these were confirmed translocation heterozygotes.

Data from negative controls and the positive control experiments are reported.

**3. REPORTING****3.1. TEST REPORT**

The test report shall, if possible, contain the following information:

- strain of mice, age of animals, weights of treated animals,
- numbers of parental animals of each sex in experimental and control groups,
- test conditions, detailed description of treatment, dose levels, solvents, mating schedule,
- number and sex of offspring per female, number and sex of offspring raised for translocation analysis,
- time and criteria of translocation analysis,
- number and detailed description of translocation carriers, including breeding data and uterine content data, if applicable;
- cytogenetic procedures and details of microscopic analysis, preferably with pictures,
- statistical evaluation,
- discussion of results,
- interpretation of results.

**▼B**

3.2. EVALUATION AND INTERPRETATION

See General introduction Part B.

4. **REFERENCES**

See General introduction Part B.

**▼B****B.26. SUB-CHRONIC ORAL TOXICITY TEST REPEATED DOSE 90  
– DAY ORAL TOXICITY STUDY IN RODENTS****1. METHOD**

This sub-chronic oral toxicity test method is a replicate of the OECD TG 408 (1998).

**1.1. INTRODUCTION**

In the assessment and evaluation of the toxic characteristics of a chemical, the determination of sub-chronic oral toxicity using repeated doses may be carried out after initial information on toxicity has been obtained from acute or repeated dose 28-day toxicity tests. The 90-day study provides information on the possible health hazards likely to arise from repeated exposure over a prolonged period of time covering post-weaning maturation and growth well into adulthood. The study will provide information on the major toxic effects, indicate target organs and the possibility of accumulation, and can provide an estimate of a no-observed-adverse-effect level of exposure which can be used in selecting dose levels for chronic studies and for establishing safety criteria for human exposure.

The method places additional emphasis on neurological endpoints and gives an indication of immunological and reproductive effects. The need for careful clinical observations of the animals, so as to obtain as much information as possible, is also stressed. This study should allow for the identification of chemicals with the potential to cause neurotoxic, immunological or reproductive organ effects, which may warrant further in-depth investigation.

See also General introduction Part B.

**1.2. DEFINITIONS**

**Dose:** is the amount of test substance administered. Dose is expressed as weight (g, mg) or as weight of test substance per unit weight of test animal (e.g. mg/kg), or as constant dietary concentrations (ppm).

**Dosage:** is a general term comprising of dose, its frequency and the duration of dosing.

**NOAEL:** is the abbreviation for no-observed-adverse-effect level and is the highest dose level where no adverse treatment-related findings are observed.

**1.3. PRINCIPLE OF THE TEST METHOD**

The test substance is orally administered daily in graduated doses to several groups of experimental animals, one dose level per group for a period of 90 days. During the period of administration the animals are observed closely for signs of toxicity. Animals, which die or are killed during the test are necropsied and, at the conclusion of the test, surviving animals are also killed and necropsied.

**▼B**

## 1.4. DESCRIPTION OF THE TEST METHOD

1.4.1. **Preparations of animals**

Healthy animals, which have been acclimated to laboratory conditions for at least five days and have not been subjected to previous experimental procedures, should be used. The test animals should be characterised as to species, strain, source, sex, weight and/or age. Animals should be randomly assigned to the control and treatment groups. Cages should be arranged in such a way that possible effects due to cage placement are minimised. Each animal should be assigned a unique identification number.

1.4.2. **Preparations of doses**

The test substance is administered by gavage or via the diet or drinking water. The method of oral administration is dependent on the purpose of the study, and the physical/chemical properties of the test material.

Where necessary, the test substance is dissolved or suspended in a suitable vehicle. It is recommended that, wherever possible, the use of an aqueous solution/suspension be considered first, followed by consideration of a solution/emulsion in oil (e.g. corn oil) and then by possible solution in other vehicles. For vehicles other than water the toxic characteristics of the vehicle must be known. The stability of the test substance under the conditions of administration should be determined.

1.4.3. **Test conditions**1.4.3.1. *Experimental animals*

The preferred species is the rat, although other rodent species, e.g. the mouse, may be used. Commonly used laboratory strains of young healthy adult animals should be employed. The females should be nulliparous and non-pregnant. Dosing should begin as soon as possible after weaning and, in any case, before the animals are nine weeks old. At the commencement of the study the weight variation of animals used should be minimal and not exceed  $\pm 20\%$  of the mean weight of each sex. Where the study is conducted as a preliminary to a long term chronic toxicity study, animals from the same strain and source should be used in both studies.

1.4.3.2. *Number and sex*

At least 20 animals (10 female and 10 male) should be used at each dose level. If interim kills are planned, the number should be increased by the number of animals scheduled to be killed before the completion of the study. Based on previous knowledge of the chemical or a close analogue, consideration should be given to including an additional satellite group of ten animals (five per sex) in the control and in the top dose group for observation, after the treatment period, of reversibility or persistence of any toxic effects. The duration of this post-treatment period should be fixed appropriately with regard to the effects observed.

**▼B**1.4.3.3. *Dose levels*

At least three dose levels and a concurrent control shall be used, except where a limit test is conducted (see 1.4.3.4). Dose levels may be based on the results of repeated dose or range finding studies and should take into account any existing toxicological and toxicokinetic data available for the test substance or related materials. Unless limited by the physical-chemical nature or biological effects of the test substance, the highest dose level should be chosen with the aim to induce toxicity but not death or severe suffering. A descending sequence of dose levels should be selected with a view to demonstrating any dosage related response and a no-observed-adverse-effect level (NOAEL) at the lowest dose level. Two to four-fold intervals are frequently optimal for setting the descending dose levels and addition of a fourth test group is often preferable to using very large intervals (e.g. more than a factor of about 6-10) between dosages.

The control group shall be an untreated group or a vehicle-control group if a vehicle is used in administering the test substance. Except for treatment with the test substance, animals in the control group should be handled in an identical manner to those in the test groups. If a vehicle is used, the control group shall receive the vehicle in the highest volume used. If a test substance is administered in the diet, and causes reduced dietary intake, then a pair-fed control group may be useful in distinguishing between reductions due to palatability or toxicological alterations in the test model.

Consideration should be given to the following characteristics of the vehicle and other additives, as appropriate: effects on the absorption, distribution, metabolism, or retention of the test substance; effects on the chemical properties of the test substance which may alter its toxic characteristics; and effects on the food or water consumption or the nutritional status of the animals.

1.4.3.4. *Limit test*

If a test at one dose level, equivalent to at least 1 000 mg/kg body weight/day, using the procedures described for this study, produces no-observed-adverse-effects and if toxicity would not be expected based upon data from structurally related substances, then a full study using three dose levels may not be considered necessary. The limit test applies except when human exposure indicates the need for a higher dose level to be used.

1.5. **PROCEDURE**1.5.1. **Administration of doses**

The animals are dosed with the test substance daily seven days each week for a period of 90 days. Any other dosing regime, e.g. five days per week, needs to be justified. When the test substance is administered by gavage, this should be done in a single dose to the animals using a stomach tube or a suitable intubation cannula. The maximum volume of liquid that can be administered at one time depends on the size of the test animal. The volume should not exceed 1 ml/100 g body weight, except in the case of aqueous solutions where 2 ml/100 g body weight may be used. Except for irritating or corrosive substances, which will normally reveal exacerbated effects with higher concentrations, variability in test volume should be minimised by adjusting the concentration to ensure a constant volume at all dose levels.

**▼B**

For substances administered via the diet or drinking water it is important to ensure that the quantities of the test substance involved do not interfere with normal nutrition or water balance. When the test substance is administered in the diet either a constant dietary concentration (ppm) or a constant dose level in terms of the animal's body weight may be used; the alternative used must be specified. For a substance administered by gavage, the dose should be given at similar times each day, and adjusted as necessary to maintain a constant dose level in terms of animal body weight. Where a 90-day study is used as a preliminary to a long term chronic toxicity study, a similar diet should be used in both studies.

**1.5.2. Observations**

The observation period should be at least 90 days. Animals in a satellite group scheduled for follow-up observations should be kept for an appropriate period without treatment to detect persistence of, or recovery from toxic effects.

General clinical observations should be made at least once a day, preferably at the same time(s) each day, taking into consideration the peak period of anticipated effects after dosing. The clinical condition of the animals should be recorded. At least twice daily, usually at the beginning and end of each day, all animals are inspected for signs of morbidity and mortality.

At least once prior to the first exposure (to allow for within-subject comparisons), and once a week thereafter, detailed clinical observations should be made in all animals. These observations should be made outside the home cage, preferably in a standard arena and at similar times on each occasion. They should be carefully recorded, preferably using scoring systems, explicitly defined by the testing laboratory. Effort should be made to ensure that variations in the observation conditions are minimal. Signs noted should include, but not be limited to, changes in skin, fur, eyes, mucous membranes, occurrence of secretions and excretions and autonomic activity (e.g. lacrimation, pilo-erection, pupil size, unusual respiratory pattern). Changes in gait, posture and response to handling as well as the presence of clonic or tonic movements, stereotypes (e.g. excessive grooming, repetitive circling) or bizarre behaviour (e.g. self-mutilation, walking backwards) should also be recorded (1).

Ophthalmological examination, using an ophthalmoscope or equivalent suitable equipment, should be made prior to the administration of the test substance and at the termination of the study, preferably in all animals but at least in the high dose and control groups. If changes in the eyes are detected all animals should be examined.

Towards the end of the exposure period and in any case not earlier than in week 11, sensory reactivity to stimuli of different types (1) (e.g. auditory, visual and proprioceptive stimuli) (2), (3), (4), assessment of grip strength (5) and motor activity assessment (6) should be conducted. Further details of the procedures that could be followed are given in the respective references. However, alternative procedures than those referenced could also be used.

Functional observations conducted towards the end of the study may be omitted when data on functional observations are available from other studies and the daily clinical observations did not reveal any functional deficits.



**▼B**

Exceptionally, functional observations may also be omitted for groups that otherwise reveal signs of toxicity to an extent that would significantly interfere with the functional test performance.

1.5.2.1. *Body weight and food/water consumption*

All animals should be weighed at least once a week. Measurements of food consumption should be made at least weekly. If the test substance is administered via the drinking water, water consumption should also be measured at least weekly. Water consumption may also be considered for dietary or gavage studies during which drinking activity may be altered.

1.5.2.2. *Haematology and clinical biochemistry*

Blood samples should be taken from a named site and stored, if applicable, under appropriate conditions. At the end of the test period, samples are collected just prior to or as part of the procedure for killing the animals.

The following haematological examinations should be made at the end of the test period and when any interim blood samples may have been collected: haematocrit, haemoglobin concentration, erythrocyte count, total and differential leukocyte count, platelet count and a measure of blood clotting time/potential.

Clinical biochemistry determinations to investigate major toxic effects in tissues and, specifically, effects on kidney and liver, should be performed on blood samples obtained from each animal just prior to or as part of the procedure for killing the animals (apart from those found moribund and/or intercurrently killed). In a similar manner to haematological investigations, interim sampling for clinical biochemical tests may be performed. Overnight fasting of the animals prior to blood sampling is recommended<sup>(1)</sup>. Determinations in plasma or serum should include sodium, potassium, glucose, total cholesterol, urea, blood urea nitrogen, creatinine, total protein and albumin, and more than two enzymes indicative of hepatocellular effects (such as alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, gamma glutamyl transpeptidase, and sorbitol dehydrogenase). Measurements of additional enzymes (of hepatic or other origin) and bile acids, which may provide useful information under certain circumstances, may also be included.

Optionally, the following urinalysis determinations could be performed during the last week of the study using timed urine volume collection: appearance, volume, osmolality or specific gravity, pH, protein, glucose and blood/blood cells.

In addition, studies to investigate serum markers of general tissue damage should be considered. Other determinations that should be carried out if the known properties of the test substance may, or are suspected to, affect related metabolic profiles include calcium, phosphorus, fasting triglycerides, specific hormones, methaemoglobin and cholinesterase. These need to be identified for chemicals in certain classes or on a case-by-case basis.

<sup>(1)</sup> For a number of measurements in serum and plasma, most notably for glucose, overnight fasting would be preferable. The major reason for this preference is that the increased variability which would inevitably result from non-fasting, would tend to mask more subtle effects and make interpretation difficult. On the other hand, however, overnight fasting may interfere with the general metabolism of the animals and, particularly in feeding studies, may disturb the daily exposure to the test substance. If overnight fasting is adopted, clinical biochemical determinations should be performed after the conduct of functional observations of the study.

**▼B**

Overall, there is a need for a flexible approach, depending on the species and the observed and/or expected effect from a given substance.

If historical baseline data are inadequate, consideration should be given as to whether haematological and clinical biochemistry variables need to be determined before dosing commences; it is generally not recommended that this data be generated before treatment (7).

#### 1.5.2.3. *Gross necropsy*

All animals in the study shall be subjected to a full, detailed gross necropsy which includes careful examination of the external surface of the body, all orifices, and the cranial, thoracic and abdominal cavities and their contents. The liver, kidneys, adrenals, testes, epididymides, uterus, ovaries, thymus, spleen, brain and heart of all animals (apart from those found moribund and/or intercurrently killed) should be trimmed of any adherent tissue, as appropriate, and their wet weight taken as soon as possible after dissection to avoid drying.

The following tissues should be preserved in the most appropriate fixation medium for both the type of tissue and the intended subsequent histopathological examination: all gross lesions, brain (representative regions including cerebrum, cerebellum and medulla/pons), spinal cord (at three levels: cervical, mid-thoracic and lumbar), pituitary, thyroid, parathyroid, thymus, oesophagus, salivary glands, stomach, small and large intestines (including Peyer's patches), liver, pancreas, kidneys, adrenals, spleen, heart, trachea and lungs (preserved by inflation with fixative and then immersion), aorta, gonads, uterus, accessory sex organs, female mammary gland, prostate, urinary bladder, gall bladder (mouse), lymph nodes (preferably one lymph node covering the route of administration and another one distant from the route of administration to cover systemic effects), peripheral nerve (sciatic or tibial) preferably in close proximity to the muscle, a section of bone marrow (and/or a fresh bone marrow aspirate), skin and eyes (if changes were observed during ophthalmological examinations). The clinical and other findings may suggest the need to examine additional tissues. Also any organs considered likely to be target organs based on the known properties of the test substance should be preserved.

#### 1.5.2.4. *Histopathology*

Full histopathology should be carried out on the preserved organs and tissues of all animals in the control and high dose groups. These examinations should be extended to animals of all other dosage groups, if treatment-related changes are observed in the high dose group.

All gross lesions should be examined.

When a satellite group is used, histopathology should be performed on tissues and organs identified as showing effects in the treated groups.

**▼ B****2. DATA AND REPORTING****2.1. DATA**

Individual data should be provided. Additionally, all data should be summarised in tabular form showing for each test group the number of animals at the start of the test, the number of animals found dead during the test or killed for humane reasons and the time of any death or humane kill, the number showing signs of toxicity, a description of the signs of toxicity observed, including time of onset, duration, and severity of any toxic effects, the number of animals showing lesions, the type of lesions and the percentage of animals displaying each type of lesion.

When applicable, numerical results should be evaluated by an appropriate and generally acceptable statistical method. The statistical methods and the data to be analysed should be selected during the design of the study.

**2.2. TEST REPORT**

The test report must include the following information:

**2.2.1. Test substance:**

- physical nature, purity and physico-chemical properties,
- identification data,
- vehicle (if appropriate): justification for choice of vehicle, if other than water.

**2.2.2. Test species:**

- species and strain used,
- number, age and sex of animals,
- source, housing conditions, diet etc.,
- individual weights of animals at the start of the test.

**2.2.3. Test conditions:**

- rationale for dose level selection,
- details of test substance formulation/diet preparation, achieved concentration, stability and homogeneity of the preparation,
- details of the administration of the test substance,
- actual doses (mg/kg body weight/day), and conversion factor from diet/drinking water test substance concentration (ppm) to the actual dose, if applicable,
- details of food and water quality.

**2.2.4. Results:**

- body weight and body weight changes,
- food consumption, and water consumption, if applicable,
- toxic response data by sex and dose level, including signs of toxicity,

**▼ B**

- nature, severity and duration of clinical observations (whether reversible or not),
- results of ophthalmological examination,
- sensory activity, grip strength and motor activity assessments (when available),
- haematological tests with relevant base-line values,
- clinical biochemistry tests with relevant base-line values,
- terminal body weight, organ weights and organ/body weight ratios,
- necropsy findings,
- a detailed description of all histopathological findings,
- absorption data if available,
- statistical treatment of results, where appropriate,

Discussion of results.

Conclusions.

## 3.

**REFERENCES**

- (1) IPCS (1986) Principles and Methods for the Assessment of Neurotoxicity Associated with Exposure to Chemicals. Environmental Health Criteria Document No 60.
- (2) Tupper, D.E., Wallace, R.B., (1980) Utility of the Neurologic Examination in Rats. *Acta Neurobiol. Exp.*, 40, p. 999-1003.
- (3) Gad, S.C., (1982) A Neuromuscular Screen for Use in Industrial Toxicology. *J.Toxicol. Environ. Health*, 9, p. 691-704.
- (4) Moser, V.C., Mc Daniel, K.M., Phillips, P.M., (1991) Rat Strain and Stock Comparisons Using a Functional Observational Battery: Baseline Values and Effects of Amitraz. *Toxicol. Appl. Pharmacol.*, 108, p. 267-283.
- (5) Meyer O.A., Tilson H.A., Byrd W.C., Riley M.T., (1979) A Method for the Routine Assessment of Fore- and Hind- limb grip Strength of Rats and Mice. *Neurobehav. Toxicol.*, 1, p. 233-236.
- (6) Crofton K.M., Howard J.L., Moser V.C., Gill M.W., Reiter L.W., Tilson H.A., MacPhail R.C., (1991) Interlaboratory Comparison of Motor Activity Experiments: Implication for Neurotoxicological Assessments. *Neurotoxicol. Teratol.*, 13, p. 599-609.
- (7) Weingand K., Brown G., Hall R. *et al.*, (1996) 'Harmonisation of Animal Clinic Pathology Testing in Toxicity and Safety Studies', *Fundam. & Appl. Toxicol.*, 29, p. 198-201.

**▼B****B.27. SUB-CHRONIC ORAL TOXICITY TEST REPEATED DOSE 90-DAY ORAL TOXICITY STUDY IN NON-RODENTS****1. METHOD**

This sub-chronic oral toxicity test method is a replicate of the OECD TG 409 (1998).

**1.1. INTRODUCTION**

In the assessment and evaluation of the toxic characteristics of a chemical, the determination of sub-chronic oral toxicity using repeated doses may be carried out after initial information on toxicity has been obtained from acute or repeated dose 28-day toxicity tests. The 90-day study provides information on the possible health hazards likely to arise from repeated exposure over a period of rapid growth and into young adulthood. The study will provide information on the major toxic effects, indicate target organs and the possibility of accumulation, and can provide an estimate of a no-observed-adverse-effect level of exposure which can be used in selecting dose levels for chronic studies and for establishing safety criteria for human exposure.

The test method allows for the identification in non-rodent species of adverse effects of chemical exposure and should only be used:

- where effects observed in other studies indicate a need for clarification/characterisation in a second, non-rodent species, or
- where toxicokinetic studies indicate that the use of a specific non-rodent species is the most relevant choice of laboratory animal, or
- where other specific reasons justify the use of a non-rodent species.

See also General introduction Part B.

**1.2. DEFINITIONS**

**Dose:** is the amount of test substance administered. Dose is expressed as weight (g, mg) or as weight of test substance per unit weight of test animal (e.g. mg/kg), or as constant dietary concentrations (ppm).

**Dosage:** is a general term comprising of dose, its frequency and the duration of dosing.

**NOAEL:** is the abbreviation for no-observed-adverse-effect level and is the highest dose level where no adverse treatment-related findings are observed.

**1.3. PRINCIPLE OF THE TEST METHOD**

The test substance is orally administered daily in graduated doses to several groups of experimental animals, one dose level per group for a period of 90 days. During the period of administration the animals are observed closely for signs of toxicity. Animals, which die or are killed during the test are necropsied and at the conclusion of the test surviving animals are also killed and necropsied.

**▼B**

## 1.4. DESCRIPTION OF THE TEST METHOD

1.4.1. **Selection of animal species**

The commonly used non-rodent species is the dog, which should be of a defined breed; the beagle is frequently used. Other species, e.g. swine, mini-pigs, may also be used. Primates are not recommended and their use should be justified. Young, healthy animals should be employed, and in the case of the dog, dosing should begin preferably at four to six months and not later than nine months of age. Where the study is conducted as a preliminary to a long-term chronic toxicity study, the same species/breed should be used in both studies.

1.4.2. **Preparation of animals**

Healthy young animals, which have been acclimated to laboratory conditions and have not been subjected to previous experimental procedures, should be used. The duration of acclimatisation will depend upon the selected test species and their source. At least five days for dogs or purpose bred swine from a resident colony and at least two weeks for these animals if from external sources are recommended. The test animals should be characterised as to species, strain, source, sex, weight and/or age. Animals should be randomly assigned to the control and treatment groups. Cages should be arranged in such a way that possible effects due to cage placement are minimised. Each animal should be assigned a unique identification number.

1.4.3. **Preparations of doses**

The test substance may be administered in the diet or in the drinking water, by gavage or in capsules. The method of oral administration is dependent on the purpose of the study, and the physical-chemical properties of the test material.

Where necessary, the test substance is dissolved or suspended in a suitable vehicle. It is recommended that, wherever possible, the use of an aqueous solution/suspension be considered first, followed by consideration of a solution/emulsion in oil (e.g. corn oil) and then by possible solution in other vehicles. For vehicles other than water the toxic characteristics of the vehicle must be known. The stability of the test substance under the conditions of administration should be determined.

## 1.5. PROCEDURE

1.5.1. **Number and sex of animals**

At least eight animals (four female and four male) should be used at each dose level. If interim kills are planned, the number should be increased by the number of animals scheduled to be killed before the completion of the study. The number of animals at the termination of the study must be adequate for a meaningful evaluation of toxic effects. Based on previous knowledge of the substance or a close analogue, consideration should be given to including an additional satellite group of eight animals (four per sex) in control and in top dose group for observation after the treatment period of reversibility or persistence of any toxic effects. The duration of this post-treatment period should be fixed appropriately with regard to the effects observed.

**▼ B****1.5.2. Dosage**

At least three dose levels and a concurrent control shall be used, except where a limit test is conducted (see 1.5.3). Dose levels may be based on the results of repeated dose or range finding studies and should take into account any existing toxicological and toxicokinetic data available for the test compound or related materials. Unless limited by the physical-chemical nature or biological effects of the test substance, the highest dose level should be chosen with the aim to induce toxicity but not death or severe suffering. A descending sequence of dose levels should be selected with a view to demonstrating any dosage related response and a no-observed-adverse-effect level (NOAEL) at the lowest dose level. Two to fourfold intervals are frequently optimal for setting the descending dose levels and addition of a fourth test group is often preferable to using very large intervals (e.g. more than a factor of about 6–10) between dosages.

The control group shall be an untreated group or a vehicle-control group if a vehicle is used in administering the test substance. Except for treatment with the test substance, animals in the control group should be handled in an identical manner to those in the test groups. If a vehicle is used, the control group shall receive the vehicle in the highest volume used. If a test substance is administered in the diet, and causes reduced dietary intake, then a pair-fed control group may be useful in distinguishing between reductions due to palatability or toxicological alterations in the test model.

Consideration should be given to the following characteristics of the vehicle and other additives, as appropriate: effects on the absorption, distribution, metabolism, or retention of the test substance; effects on the chemical properties of the test substance which may alter its toxic characteristics; and effects on the food or water consumption or the nutritional status of the animals.

**1.5.3. Limit test**

If a test at one dose level, equivalent to at least 1 000 mg/kg body weight/day, using the procedures described for this study, produces no-observed-adverse-effects and if toxicity would not be expected based upon data from structurally related substances, then a full study using three dose levels may not be considered necessary. The limit test applies except when human exposure indicates the need for a higher dose level to be used.

**1.5.4. Administration of doses**

The animals are dosed with the test substance daily seven days each week for a period of 90 days. Any other dosing regime, e.g. five days per week, needs to be justified. When the test substance is administered by gavage, this should be done in a single dose to the animals using a stomach tube or a suitable intubation cannula. The maximum volume of liquid that can be administered at one time depends on the size of the test animal. Normally the volume should be kept as low as possible. Except for irritating or corrosive substances which will normally reveal exacerbated effects with higher concentrations, variability in test volume should be minimised by adjusting the concentration to ensure a constant volume at all dose levels.

**▼B**

For substances administered via the diet or drinking water it is important to ensure that the quantities of the test substance involved do not interfere with normal nutrition or water balance. When the test substance is administered in the diet either a constant dietary concentration (ppm) or a constant dose level in terms of the animal's body weight may be used; any alternative used must be specified. For a substance administered by gavage or by capsule, the dose should be given at similar times each day, and adjusted as necessary to maintain a constant dose level in terms of animal body weight. Where the 90 day study is used as a preliminary to a long term chronic toxicity study, a similar diet should be used in both studies.

**1.5.5. Observations**

The observation period should be at least 90 days. Animals in a satellite group scheduled for follow-up observations should be kept for an appropriate period without treatment to detect persistence of, or recovery from toxic effects.

General clinical observations should be made at least once a day, preferably at the same time(s) each day, taking into consideration the peak period of anticipated effects after dosing. The clinical condition of the animals should be recorded. At least twice daily, usually at the beginning and end of each day, all animals should be inspected for signs of morbidity and mortality.

At least once prior to the first exposure (to allow for within-subject comparisons), and once a week thereafter, detailed clinical observations should be made in all animals. These observations should be made, where practical outside the home cage in a standard arena and preferably at similar times on each occasion. Effort should be made to ensure that variations in the observation conditions are minimal. Signs of toxicity should be carefully recorded, including time of onset, degree and duration. Observations should include, but not be limited to, changes in skin, fur, eyes, mucous membranes, occurrence of secretions and excretions and autonomic activity (e.g. lacrimation, pilo-erection, pupil size, unusual respiratory pattern). Changes in gait, posture and response to handling as well as the presence of clonic or tonic movements, stereotypes (e.g. excessive grooming, repetitive circling) or any bizarre behaviour should also be recorded.

Ophthalmological examination, using an ophthalmoscope or equivalent suitable equipment, should be made prior to the administration of the test substance and at the termination of the study, preferably in all animals but at least in the high dose and control groups. If treatment related changes in the eyes are detected all animals should be examined.

**1.5.5.1. *Body weight and food/water consumption***

All animals should be weighed at least once a week. Measurements of food consumption should be made at least weekly. If the test substance is administered via the drinking water, water consumption should also be measured at least weekly. Water consumption may also be considered for dietary or gavage studies during which drinking activity may be altered.



**▼B**1.5.5.2. *Haematology and clinical biochemistry*

Blood samples should be taken from a named site and stored, if applicable, under appropriate conditions. At the end of the test period, samples are collected just prior to or as part of the procedure for killing the animals.

Haematology, including haematocrit, haemoglobin concentration, erythrocyte count, total and differential leukocyte count, platelet count and a measure of clotting potential such as clotting time, prothrombin time, or thromboplastin time should be investigated at the start of the study, then either at monthly intervals or midway through the test period and finally at the end of the test period.

Clinical biochemistry determinations to investigate major toxic effects in tissues and, specifically, effects on kidney and liver, should be performed on blood samples obtained from all animals at the start, then either at monthly intervals or midway through the test and finally at the end of the test period. Test areas, which should be considered are electrolyte balance, carbohydrate metabolism, and liver and kidney function. The selection of specific tests will be influenced by observations on the mode of action of the test substance. Animals should be fasted for a period appropriate to the species prior to blood sampling. Suggested determinations include calcium, phosphorus, chloride, sodium, potassium, fasting glucose, alanine aminotransferase, aspartate aminotransferase, ornithine decarboxylase, gamma glutamyl transpeptidase, urea nitrogen, albumin, blood creatinine, total bilirubin and total serum protein measurements.

Urinalysis determinations should be performed at least at the start, then midway and finally at the end of the study using timed urine volume collection. Urinalysis determinations include appearance, volume, osmolality or specific gravity, pH, protein, glucose and blood/blood cells. Additional parameters may be employed where necessary to extend the investigation of observed effect(s).

In addition, studies to investigate markers of general tissue damage should be considered. Other determinations, which may be necessary for an adequate toxicological evaluation include analyses of lipids, hormones, acid/base balance, methaemoglobin, and cholinesterase inhibition. Additional clinical biochemistry may be employed where necessary to extend the investigation of observed effects. These need to be identified for chemicals in certain classes or on a case-by-case basis.

Overall, there is a need for a flexible approach, depending on the species and the observed and/or expected effect from a given substance.

1.5.5.3. *Gross necropsy*

All animals in the study shall be subjected to a full, detailed gross necropsy which includes careful examination of the external surface of the body, all orifices, and the cranial, thoracic and abdominal cavities and their contents. The liver with gall bladder, kidneys, adrenals, testes, epididymides, ovaries, uterus, thyroid (with parathyroids), thymus, spleen, brain and heart of all animals (apart from those found moribund and/or inter-currently killed) should be trimmed of any adherent tissue, as appropriate, and their wet weight taken as soon as possible after dissection to avoid drying.

**▼B**

The following tissues should be preserved in the most appropriate fixation medium for both the type of tissue and the intended subsequent histopathological examination: all gross lesions, brain (representative regions including cerebrum, cerebellum and medulla/pons), spinal cord (at three levels: cervical, mid-thoracic and lumbar), pituitary, eyes, thyroid, parathyroid, thymus, oesophagus, salivary glands, stomach, small and large intestines (including Peyer's patches), liver, gall bladder, pancreas, kidneys, adrenals, spleen, heart, trachea and lungs, aorta, gonads, uterus, accessory sex organs, female mammary gland, prostate, urinary bladder, lymph nodes (preferably one lymph node covering the route of administration and another one distant from the route of administration to cover systemic effects), peripheral nerve (sciatic or tibial) preferably in close proximity to the muscle, a section of bone marrow (and/or a fresh bone marrow aspirate) and skin. The clinical and other findings may suggest the need to examine additional tissues. Also any organs considered likely to be target organs based on the known properties of the test substance should be preserved.

**1.5.5.4. Histopathology**

Full histopathology should be carried out on the preserved organs and tissues in at least all animals in control and high dose group. The examination should be extended to animals of all other dosage groups, if treatment-related changes are observed in the high dose group.

All gross lesions should be examined.

When a satellite group is used, histopathology should be performed on tissues and organs identified as showing effects in the treated groups.

**2. DATA AND REPORTING****2.1. DATA**

Individual data should be provided. Additionally, all data should be summarised in tabular form showing for each test group the number of animals at the start of the test, the number of animals found dead during the test or killed for humane reasons and the time of any death or humane kill, the number showing signs of toxicity, a description of the signs of toxicity observed, including time of onset, duration, and severity of any toxic effects, the number of animals showing lesions, the type of lesions and the percentage of animals displaying each type of lesion.

When applicable, numerical results should be evaluated by an appropriate and generally acceptable statistical method. The statistical methods and the data to be analysed should be selected during the design of the study.

**2.2. TEST REPORT**

The test report must include the following information:

**▼ B**

- 2.2.1. **Test substance:**
- physical nature, purity and physico-chemical properties,
  - identification data,
  - vehicle (if appropriate): justification for choice of vehicle, if other than water.
- 2.2.2. **Test species:**
- species and strain used,
  - number, age and sex of animals,
  - source, housing conditions, diet etc.,
  - individual weights of animals at the start of the test.
- 2.2.3. **Test conditions:**
- rationale for dose level selection,
  - details of test substance formulation/diet preparation, achieved concentration, stability and homogeneity of the preparation,
  - details of the administration of the test substance,
  - actual doses (mg/kg body weight/day), and conversion factor from diet/drinking water test substance concentration (ppm) to the actual dose, if applicable,
  - details of food and water quality.
- 2.2.4. **Results:**
- body weight/body weight changes,
  - food consumption, and water consumption, if applicable,
  - toxic response data by sex and dose level, including signs of toxicity,
  - nature, severity and duration of clinical observations (whether reversible or not),
  - ophthalmological examination,
  - haematological tests with relevant base-line values,
  - clinical biochemistry tests with relevant base-line values,
  - terminal body weight, organ weights and organ/body weight ratios,
  - necropsy findings,
  - a detailed description of all histopathological findings,
  - absorption data if available,
  - statistical treatment of results, where appropriate.
- Discussion of results.
- Conclusions.

**▼B****B.28. SUB-CHRONIC DERMAL TOXICITY STUDY 90-DAY REPEATED DERMAL DOSE STUDY USING RODENT SPECIES****1. METHOD****1.1. INTRODUCTION**

See General introduction Part B.

**1.2. DEFINITIONS**

See General introduction Part B.

**1.3. REFERENCE SUBSTANCES**

None.

**1.4. PRINCIPLE OF THE TEST METHOD**

The test substance is applied daily to the skin in graduated doses to several groups of experimental animals, one dose per group for a period of 90 days. During the period of application the animals are observed daily to detect signs of toxicity. Animals, which die during the test are necropsied, and at the conclusion of the test surviving animals are necropsied.

**1.5. QUALITY CRITERIA**

None.

**1.6. DESCRIPTION OF THE TEST METHOD****1.6.1. Preparations**

The animals are kept under the experimental housing and feeding conditions for at least five days prior to the test. Before the test healthy young animals are randomised and assigned to the treated and control groups. Shortly before testing fur is clipped from the dorsal area of the trunk of the test animals. Shaving may be employed but it should be carried out approximately 24 hours before the test. Repeat clipping or shaving is usually needed at approximately weekly intervals. When clipping or shaving the fur, care must be taken to avoid abrading the skin. Not less than 10 % of the body surface area should be clear for the application of the test substance. The weight of the animal should be taken into account when deciding on the area to be cleared and on the dimensions of the covering. When testing solids, which may be pulverised if appropriate, the test substance should be moistened sufficiently with water or, where necessary, a suitable vehicle to ensure good contact with the skin. Liquid test substances are generally used undiluted. Daily application on a five to seven-day per week basis is used.

**1.6.2. Test conditions****1.6.2.1. Experimental animals**

The adult rat, rabbit or guinea pig may be used. Other species may be used but their use would require justification. At the commencement of the test the range of the weight variation should be  $\pm 20$  % of the mean weight. Where a sub-chronic dermal study is conducted as a preliminary to a long-term study, the same species and strain should be used in both studies.

**▼B**

## 1.6.2.2. Number and sex

At least 20 animals (10 female and 10 male) with healthy skin should be used at each dose level. The females should be nulliparous and non-pregnant. If interim sacrifices are planned the number should be increased by the number of animals scheduled to be sacrificed before the completion of the study. In addition, a satellite group of 20 animals (10 animals per sex) may be treated with the high-dose level for 90 days and observed for reversibility, persistence, or delayed occurrence of toxic effects for 28 days post-treatment.

## 1.6.2.3. Dose levels

At least three dose levels are required with a control or a vehicle control if a vehicle is used. The exposure period should be at least six hours per day. The application of the test substance should be made at similar times each day, and the amount of substance applied adjusted at intervals (weekly or bi-weekly) to maintain a constant dose level in terms of animal body weight. Except for treatment with the test substance, animals in the control group should be handled in an identical manner to the test group subjects. Where a vehicle is used to facilitate dosing, the vehicle control group should be dosed in the same way as the treated groups, and receive the same amount as that received by the highest dose level group. The highest dose level should result in toxic effects but produce no, or few, fatalities. The lowest dose level should not produce any evidence of toxicity. Where there is a usable estimation of human exposure the lowest level should exceed this. Ideally, the intermediate dose level should produce minimal observable toxic effects. If more than one intermediate dose is used, the dose levels should be spaced to produce a gradation of toxic effects. In the low and intermediate groups, and in the controls, the incidence of fatalities should be low, in order to permit a meaningful evaluation of the results.

If application of the test substance produces severe skin irritation the concentrations should be reduced and this may result in a reduction in, or absence of, other toxic effects at the high-dose level. If the skin has been badly damaged it may be necessary to terminate the study and undertake a new study at lower concentrations.

## 1.6.3. Limit test

If a preliminary study at a dose level of 1 000 mg/kilograms, or a higher dose level related to possible human exposure where this is known, produces no toxic effects, further testing may not be considered necessary.

## 1.6.4. Observation period

The experimental animals should be observed daily for signs of toxicity. The time of death and the time at which signs of toxicity appear and disappear should be recorded.

**▼B**

## 1.6.5. Procedure

Animals should be caged individually. The animals are treated with the test substance, ideally on seven days per week, for a period of 90 days.

Animals in any satellite groups scheduled for follow-up observations should be kept for a further 28 days without treatment to detect recovery from, or persistence of, toxic effects. Exposure time should be six hours per day.

The test substance should be applied uniformly over an area, which is approximately 10 % of the total body surface area. With highly toxic substances, the surface area covered may be less but as much of the area should be covered with as thin and uniform a film as possible.

During exposure the test substance is held in contact with the skin with a porous gauze dressing and non-irritating tape. The test site should be further covered in a suitable manner to retain the gauze dressing and test substance and ensure that the animals cannot ingest the test substance. Restrainers may be used to prevent the ingestion of the test substance but complete immobilisation is not a recommended method.

At the end of the exposure period, residual test substance should be removed where practicable using water or some other appropriate method of cleansing the skin.

All the animals should be observed daily and signs of toxicity recorded, including the time of onset, their degree and duration. Cageside observations should include changes in skin and fur, eyes and mucous membranes, as well as respiratory, circulatory, autonomic and central nervous systems, somatomotor activity and behavior pattern. Measurements should be made of food consumption weekly and the animals weighed weekly. Regular observations of the animals are necessary to ensure that animals are not lost from the study due to causes such as cannibalism, autolysis of tissues or misplacement. At the end of the study period all survivors in the non-satellite treatment groups are necropsied. Moribund animals should be removed and necropsied when noticed.

The following examinations are customarily made on all animals including the controls:

- (a) ophthalmological examination, using an ophthalmoscope or equivalent suitable equipment, should be made prior to exposure to the test substance and at the termination of the study, preferably in all animals but at least in the high-dose and control groups. If changes in the eyes are detected all animals should be examined.
- (b) haematology, including haematocrit, haemoglobin concentration, erythrocyte count, total and differential leucocyte count, and a measure of clotting potential, such as clotting time, prothrombin time, thromboplastin time, or platelet count, should be investigated at the end of the test period.

**▼B**

- (c) clinical biochemistry determination on blood should be carried out at the end of the test period. Test areas, which are considered appropriate to all studies are electrolyte balance, carbohydrate metabolism, liver and kidney function. The selection of specific tests will be influenced by observations on the mode of action of the substance. Suggested determinations are calcium, phosphorus, chloride, sodium, potassium, fasting glucose (with period of fasting appropriate to the species), serum glutamic pyruvic transaminase <sup>(1)</sup>, serum glutamic oxaloacetic transaminase <sup>(2)</sup>, ornithine decarboxylase, gamma glutamyl transpeptidase, urea nitrogen, albumin, blood creatinine, total bilirubin and total serum protein measurements. Other determinations which may be necessary for an adequate toxicological evaluation include analyses of lipids, hormones, acid/base balance, methaemoglobin and cholinesterase activity. Additional clinical biochemistry may be employed, where necessary, to extend the investigation of observed effects.
- (d) urinalysis is not required on a routine basis but only when there is an indication based on expected or observed toxicity.

If historical baseline data are inadequate, consideration should be given to determination of haem a to logical and clinical biochemistry parameters before dosing commences.

#### Gross necropsy

All animals should be subjected to a full gross necropsy which includes examination of the external surface of the body, all orifices, and the cranial, thoracic and abdominal cavities and their contents. The liver, kidneys, adrenals and testes must be weighed wet as soon as possible after dissection to avoid drying. The following organs and tissues should be preserved in a suitable medium for possible future histopathological examination: all gross lesions, brain — including sections of medulla/pons, cerebellar cortex and cerebral cortex, pituitary, thyroid/parathyroid, any thymic tissue, (trachea), lungs, heart, aorta, salivary glands, liver, spleen, kidneys, adrenals, pancreas, gonads, uterus, accessory genital organs, gall bladder (if present), oesophagus, stomach, duodenum, jejunum, ileum, caecum, colon, rectum, urinary bladder, representative lymph node, (female mammary gland), (thigh musculature), peripheral nerve, (eyes), (sternum with bone marrow), (femur — including articular surface), (spinal cord at three levels — cervical, mid-thoracic and lumbar), and (exorbital lachrymal glands). The tissues mentioned between brackets need only be examined if indicated by signs of toxicity or target organ involvement.

#### Histopathological examination

- (a) Full histopathology should be carried out on normal and treated skin and on organs and tissues of animals in the control and high-dose groups.
- (b) all gross lesions should be examined.

<sup>(1)</sup> Now known as serum alanine aminotransferase.

<sup>(2)</sup> Now known as serum aspartate aminotransferase.

**▼B**

- (c) target organs in other dose groups should be examined.
- (d) where rats are used, lungs of animals in the low- and intermediate-dose groups should be subjected to histopathological examination for evidence of infection, since this provides a convenient assessment of the state of health of the animals. Further histopathological examination may not be required routinely on the animals in these groups, but must always be carried out in organs, which show evidence of lesions in the high-dose group.
- (e) when a satellite group is used, histopathology should be performed on tissues and organs identified as showing effects in the other treated groups.

**2. DATA**

Data should be summarised in tabular form, showing for each test group the number of animals at the start of the test, the number of animals showing lesions, the type of lesions and the percentage of animals displaying each type of lesion. Results should be evaluated by an appropriate statistical method. Any recognised statistical method may be used.

**3. REPORTING****3.1. TEST REPORT**

The test report shall, if possible, contain the following information:

- species, strain, source, environmental conditions, diet,
- test conditions,
- dose levels (including vehicle, if used) and concentrations,
- toxic response data by sex and dose,
- no-effect level, where possible,
- time of death during the study or whether animals survived to termination,
- description of toxic or other effects,
- the time of observation of each abnormal sign and its subsequent course,
- food and bodyweight data,
- ophthalmological findings,
- haematological tests employed and all results,
- clinical biochemistry tests employed and all results (including results of any urinalysis),
- necropsy findings,



**▼B**

- a detailed description of all histopathological findings,
- statistical treatment of results where possible,
- discussion of the results,
- interpretation of the results.

3.2. EVALUATION AND INTERPRETATION

See General introduction Part B.

4. **REFERENCES**

See General introduction Part B.

**▼ M4****B.29. SUBCHRONIC INHALATION TOXICITY: 90-DAY STUDY****SUMMARY**

This revised Test Method B.29 has been designed to fully characterise test chemical toxicity by the inhalation route for a subchronic duration (90 days), and to provide robust data for quantitative inhalation risk assessments. Groups of 10 male and 10 female rodents are exposed 6 hours per day during a 90 day (13 week) period to a) the test chemical at three or more concentration levels, b) filtered air (negative control), and/or c) the vehicle (vehicle control). Animals are generally exposed 5 days per week but exposure for 7 days per week is also allowed. Males and females are always tested, but they may be exposed at different concentration levels if it is known that one sex is more susceptible to a given test chemical. This method allows the study director the flexibility to include satellite (reversibility) groups, interim sacrifices, bronchoalveolar lavage (BAL), neurologic tests, and additional clinical pathology and histopathological evaluations in order to better characterise the toxicity of a test chemical.

**INTRODUCTION**

1. This Test Method is equivalent to OECD Test Guideline 413 (2009). The original subchronic inhalation Test Guideline 413 (TG 413) was adopted in 1981 (1). This Test Method B.29 (as equivalent to the revised TG 413 (2009)) has been updated to reflect the state of the science and to meet current and future regulatory needs.
2. Subchronic inhalation toxicity studies are primarily used to derive regulatory concentrations for assessing worker risk in occupation settings. They are also used to assess human residential, transportation, and environmental risk. This method enables the characterisation of adverse effects following repeated daily inhalation exposure to a test chemical for 90 days (approximately 10 % of the lifespan of a rat). The data derived from subchronic inhalation toxicity studies can be used for quantitative risk assessments and for the selection of concentrations for chronic studies. This test method is not specifically intended for the testing of nanomaterials. Definitions used in the context of this Test Method are provided at the end of this chapter and in the Guidance Document (GD) 39 (2).

**INITIAL CONSIDERATIONS**

3. All available information on the test chemical should be considered by the testing laboratory prior to conducting the study in order to enhance the quality of the study and minimise animal usage. Information that will assist in the selection of appropriate test concentrations might include the identity, chemical structure, and physico-chemical properties of the test chemical; results of any *in vitro* or *in vivo* toxicity tests; anticipated use(s) and potential for human exposure; available (Q)SAR data and toxicological data on structurally related chemicals; and data derived from other repeated exposure studies. If neurotoxicity is expected or is observed in the course of the study, the study director may choose to include appropriate evaluations such as a functional observational battery (FOB) and measurement of motor activity. Although the timing of exposures relative to specific examinations may be critical, the performance of these additional activities should not interfere with the basic study design.

**▼M4**

4. Dilutions of corrosive or irritating test chemicals may be tested at concentrations that will yield the desired degree of toxicity. Please refer to GD 39 (2) for further information. When exposing animals to these materials, the targeted concentrations should be low enough to not cause marked pain and distress, yet sufficient to extend the concentration-response curve to levels that reach the regulatory and scientific objective of the test. These concentrations should be selected on a case-by-case basis, preferably based upon an adequately designed range-finding study that provides information regarding the critical endpoint, any irritation threshold, and the time of onset (see paragraphs 11-13). The justification for concentration selection should be provided.
  
5. Moribund animals or animals obviously in pain or showing signs of severe and enduring distress should be humanely killed. Moribund animals are considered in the same way as animals that die on test. Criteria for making the decision to kill moribund or severely suffering animals, and guidance on the recognition of predictable or impending death, are the subject of an OECD Guidance Document on Humane Endpoints (3).

**DESCRIPTION OF THE METHOD****Selection of Animal Species**

6. Healthy young adult rodents of commonly used laboratory strains should be employed. The preferred species is the rat. Justification should be provided if other species are used.

**Preparation of Animals**

7. Females should be nulliparous and non-pregnant. On the day of randomisation, animals should be young adults 7 to 9 weeks of age. Body weights should be within  $\pm 20$  % of the mean weight for each sex. The animals are randomly selected, marked for individual identification, and kept in their cages for at least 5 days prior to the start of the test to allow for acclimatization to laboratory conditions.

**Animal Husbandry**

8. Animals should be individually identified, preferably with subcutaneous transponders, to facilitate observations and avoid confusion. The temperature of the experimental animal maintenance room should be  $22 \pm 3$  °C. The relative humidity should ideally be maintained in the range of 30 to 70 %, though this may not be possible when using water as a vehicle. Before and after exposures, animals generally should be caged in groups by sex and concentration, but the number of animals per cage should not interfere with clear observation of each animal and should minimise losses due to cannibalism and fighting. When animals are to be exposed nose-only, it may be necessary for them to be acclimated to the restraining tubes. The restraining tubes should not impose undue physical, thermal, or immobilisation stress on the animals. Restraint may affect physiological endpoints such as body temperature (hyperthermia) and/or respiratory minute volume. If generic data are available to show that no such changes occur to any appreciable extent, then pre-adaptation to the restraining tubes is not necessary. Animals exposed whole-body to an aerosol should be housed individually during exposure to prevent them from filtering the test aerosol through the fur of their cage mates. Conventional and certified laboratory diets may be used, except during exposure, accompanied with an unlimited supply of municipal drinking water. Lighting should be artificial, the sequence being 12 hours light/12 hours dark.

**▼ M4****Inhalation Chambers**

9. The nature of the test chemical and the object of the test should be considered when selecting an inhalation chamber. The preferred mode of exposure is nose-only (which term includes head-only, nose-only, or snout-only). Nose-only exposure is generally preferred for studies of liquid or solid aerosols and for vapours that may condense to form aerosols. Special objectives of the study may be better achieved by using a whole-body mode of exposure, but this should be justified in the study report. To ensure atmosphere stability when using a whole-body chamber, the total volume of the test animals should not exceed 5 % of the chamber volume. Principles of the nose-only and whole body exposure techniques and their particular advantages and disadvantages are addressed in GD 39 (2).

## TOXICITY STUDIES

**Limit Concentrations**

10. Unlike with acute studies, there are no defined limit concentrations in subchronic inhalation toxicity studies. The maximum concentration tested should consider: 1) the maximum attainable concentration, 2) the 'worst case' human exposure level, 3) the need to maintain an adequate oxygen supply, and/or 4) animal welfare considerations. In the absence of data-based limits, the acute limits of Regulation (EC) No 1272/2008 (13) may be used (i.e. up to a maximum concentration of 5 mg/l for aerosols, 20 mg/l for vapours, and 20 000 ppm for gases); refer to GD 39 (2). Justification should be provided if it is necessary to exceed these limits when testing gases or highly volatile test chemicals (e.g. refrigerants). The limit concentration should elicit unequivocal toxicity without causing undue stress to the animals or affecting their longevity (3).

**Range-Finding Study**

11. Before commencing with the main study, it is generally necessary to perform a range-finding study. A range-finding study is more comprehensive than a sighting study because it is not limited to concentration selection. Knowledge learned from a range-finding study can lead to a successful main study. A range-finding study may, for example, provide technical information regarding analytical methods, particle sizing, discovery of toxic mechanisms, clinical pathology and histopathological data, and estimations of what may be NOAEL and MTC concentrations in a main study. The study director may choose to use the range-finding study to identify the threshold of respiratory tract irritation (e.g. with histopathology of the respiratory tract, pulmonary function testing, or bronchoalveolar lavage), the upper concentration which is tolerated without undue stress to the animals, and the parameters that will best characterise a test chemical's toxicity.
12. A range-finding study may consist of one or more concentration levels. Depending on the endpoints chosen, three to six males and three to six females should be exposed at each concentration level. A range-finding study should last a minimum of 5 days and generally no more than 28 days. The rationale for the selection of concentrations for the main study should be provided in the study report. The objective of the main study is to demonstrate a concentration-response relationship based on what is anticipated to be the most sensitive endpoint. The low concentration should ideally be a no-observed-adverse effect concentration while the high concentration should elicit unequivocal toxicity without causing undue stress to the animals or affecting their longevity (3).

**▼ M4**

13. When selecting concentration levels for the range-finding study, all available information should be considered including structure-activity relationships and data for similar chemicals (see paragraph 3). A range-finding study may verify/refute what are considered to be the most sensitive mechanistically based endpoints, e.g. cholinesterase inhibition by organophosphates, methaemoglobin formation by erythrocytotoxic agents, thyroidal hormones (T<sub>3</sub>, T<sub>4</sub>) for thyrotoxicants, protein, LDH, or neutrophils in bronchoalveolar lavage for innocuous poorly soluble particles or pulmonary irritant aerosols.

**Main Study**

14. The main subchronic toxicity study generally consists of three concentration levels, and also concurrent negative (air) and/or vehicle controls as needed (see paragraph 18). All available data should be utilised to aid selection of appropriate exposure levels, including the results of systemic toxicity studies, metabolism and kinetics (particular emphasis should be given to avoiding high concentration levels which saturate kinetic processes). Each test group contains 10 male and 10 female rodents that are exposed to the test chemical for 6 hours per day on a 5 day per week basis for a period of 13 weeks (total study duration of at least 90 days). Animals may also be exposed 7 days per week (e.g. when testing inhaled pharmaceuticals). If one sex is known to be more susceptible to a given test chemical, the sexes may be exposed at different concentration levels in order to optimise the concentration-response as described in paragraph 15. If rodent species other than rats are exposed nose-only, maximum exposure durations may be adjusted to minimise species-specific distress. A rationale should be provided when using an exposure duration less than 6 hours/day, or when it is necessary to conduct a long duration (e.g. 22 hours/day) whole-body exposure study (refer to GD 39) (2). Feed should be withheld during the exposure period unless exposure exceeds 6 hours. Water may be provided throughout a whole-body exposure.
15. The target concentrations selected should identify the target organ(s) and demonstrate a clear concentration-response:
  - The high concentration level should result in toxic effects but not cause lingering signs or lethality which would prevent a meaningful evaluation.
  - The intermediate concentration level(s) should be spaced to produce a gradation of toxic effects between that of the low and high concentration.
  - The low concentration level should produce little or no evidence of toxicity.

**Interim Sacrifices**

16. If interim sacrifices are planned, the number of animals at each exposure level should be increased by the number to be sacrificed before study completion. The rationale for using interim sacrifices should be provided, and statistical analyses should properly account for them.

**▼ M4****Satellite (Reversibility) Study**

17. A satellite (reversibility) study may be used to observe reversibility, persistence, or delayed occurrence of toxicity for a post-treatment period of an appropriate length, but no less than 14 days. Satellite (reversibility) groups consist of 10 males and 10 females exposed contemporaneously with the experimental animals in the main study. Satellite (reversibility) study groups should be exposed to the test chemical at the highest concentration level and there should be concurrent air and/or vehicle controls as needed (see paragraph 18).

**Control Animals**

18. Concurrent negative (air) control animals should be handled in a manner identical to the test group animals except that they are exposed to filtered air rather than test chemical. When water or another substance is used to assist in generating the test atmosphere, a vehicle control group, instead of a negative (air) control group, should be included in the study. Water should be used as the vehicle whenever possible. When water is used as the vehicle, the control animals should be exposed to air with the same relative humidity as the exposed groups. The selection of a suitable vehicle should be based on an appropriately conducted pre-study or historical data. If a vehicle's toxicity is not well known, the study director may choose to use both a negative (air) control and a vehicle control, but this is strongly discouraged. If historical data reveal that a vehicle is non-toxic, then there is no need for a negative (air) control group and only a vehicle control should be used. If a pre-study of a test chemical formulated in a vehicle reveals no toxicity, it follows that the vehicle is non-toxic at the concentration tested and this vehicle control should be used.

**EXPOSURE CONDITIONS****Administration of Concentrations**

19. Animals are exposed to the test chemical as a gas, vapour, aerosol, or a mixture thereof. The physical state to be tested depends on the physico-chemical properties of the test chemical, the selected concentrations, and/or the physical form most likely present during the handling and use of the test chemical. Hygroscopic and chemically reactive test chemicals should be tested under dry air conditions. Care should be taken to avoid generating explosive concentrations. Particulate materials may be subjected to mechanical processes to decrease the particle size. Further guidance is provided in GD 39 (2).

**Particle-Size Distribution**

20. Particle sizing should be performed for all aerosols and for vapours that may condense to form aerosols. To allow for exposure of all relevant regions of the respiratory tract, aerosols with mass median aerodynamic diameters (MMAD) ranging from 1 to 3  $\mu\text{m}$  with a geometric standard deviation ( $\sigma_g$ ) in the range of 1,5 to 3,0 are recommended (4). Although a reasonable effort should be made to meet this standard, expert judgement should be provided if it cannot be achieved. For example, metal fume particles will be smaller than this standard, and charged particles and fibres may exceed it.

**▼M4****Test chemical Preparation in a Vehicle**

21. Ideally, the test chemical should be tested without a vehicle. If it is necessary to use a vehicle to generate an appropriate test chemical concentration and particle size, water should be given preference. Whenever a test chemical is dissolved in a vehicle, its stability should be demonstrated.

**MONITORING OF EXPOSURE CONDITIONS****Chamber Airflow**

22. The flow of air through the exposure chamber should be carefully controlled, continuously monitored, and recorded at least hourly during each exposure. The real-time monitoring of the test atmosphere concentration (or temporal stability) is an integral measurement of all dynamic parameters and provides an indirect means to control all relevant dynamic inhalation parameters. If the concentration is monitored real-time, the frequency of measurement of air flows may be reduced to one single measurement per exposure per day. Special consideration should be given to avoiding rebreathing in nose-only chambers. Oxygen concentration should be at least 19 % and carbon dioxide concentration should not exceed 1 %. If there is reason to believe that this standard cannot be met, oxygen and carbon dioxide concentrations should be measured. If measurements on the first day of exposure show that these gases are at proper levels, no further measurements should be necessary.

**Chamber Temperature and Relative Humidity**

23. Chamber temperature should be maintained at  $22 \pm 3$  °C. Relative humidity in the animals' breathing zone, for both nose-only and whole-body exposures, should be monitored continuously and recorded hourly during each exposure where possible. The relative humidity should preferably be maintained in the range of 30 to 70 %, but this may either be unattainable (e.g. when testing water based mixtures) or not measurable due to test chemical interference with the Test Method.

**Test chemical: Nominal Concentration**

24. Whenever feasible, the nominal exposure chamber concentration should be calculated and recorded. The nominal concentration is the mass of generated test chemical divided by the total volume of air passed through the inhalation chamber system. The nominal concentration is not used to characterise the animals' exposure, but a comparison of the nominal concentration and the actual concentration gives an indication of the generation efficiency of the test system, and thus may be used to discover generation problems.

**Test chemical: Actual Concentration**

25. The actual concentration is the test chemical concentration as sampled at the animals' breathing zone in an inhalation chamber. Actual concentrations can be obtained either by specific methods (e.g. direct sampling, adsorptive or chemical reactive methods, and subsequent analytical characterisation) or by non-specific methods such as gravimetric filter analysis. The use of gravimetric analysis is acceptable only for single component powder aerosols or aerosols of low volatility liquids and should be supported by appropriate pre-study test chemical-specific characterisations. Multi-component powder

**▼M4**

aerosol concentration may also be determined by gravimetric analysis. However, this requires analytical data which demonstrate that the composition of airborne material is similar to the starting material. If this information is not available, a reanalysis of the test chemical (ideally in its airborne state) at regular intervals during the course of the study may be necessary. For aerosolised agents that may evaporate or sublimate, it should be shown that all phases were collected by the method chosen.

26. One lot of the test chemical should be used throughout the duration of the study, if possible, and the test sample should be stored under conditions that maintain its purity, homogeneity, and stability. Prior to the start of the study, there should be a characterisation of the test chemical, including its purity and, if technically feasible, the identity, and quantities of identified contaminants and impurities. This can be demonstrated by, but is not limited to, the following data: retention time and relative peak area, molecular weight from mass spectroscopy or gas chromatography analyses, or other estimates. Although the test sample's identity is not the responsibility of the test laboratory, it may be prudent for the test laboratory to confirm the sponsor's characterisation at least in a limited way (e.g. colour, physical nature, etc.).
  
27. The exposure atmosphere should be held as constant as practicable. A real-time monitoring device, such as an aerosol photometer for aerosols or a total hydrocarbon analyser for vapours, may be used to demonstrate the stability of the exposure conditions. Actual chamber concentration should be measured at least 3 times during each exposure day for each exposure level. If not feasible due to limited air flow rates or low concentrations, one sample per exposure period is acceptable. Ideally, this sample should then be collected over the entire exposure period. Individual chamber concentration samples should deviate from the mean chamber concentration by no more than  $\pm 10\%$  for gases and vapours, and by no more than  $\pm 20\%$  for liquid or solid aerosols. Time to attain chamber equilibration ( $t_{95}$ ) should be calculated and reported. The duration of an exposure spans the time that the test chemical is generated. This takes into account the times required to attain chamber equilibration ( $t_{95}$ ) and decay. Guidance for estimating  $t_{95}$  can be found in GD 39 (2).
  
28. For very complex mixtures consisting of gases/vapours and aerosols (e.g. combustion atmospheres and test chemicals propelled from purpose-driven end-use products/devices), each phase may behave differently in an inhalation chamber. Therefore, at least one indicator substance (analyte), normally the principal active ingredient in the mixture, of each phase (gas/vapour and aerosol) should be selected. When the test chemical is a mixture, the analytical concentration should be reported for the total mixture, and not just for the active ingredient or the indicator substance (analyte). Additional information regarding actual concentrations can be found in GD 39 (2).

**Test chemical: Particle Size Distribution**

29. The particle size distribution of aerosols should be determined at least weekly for each concentration level by using a cascade impactor or an alternative instrument such as an aerodynamic particle sizer (APS). If equivalence of the results obtained by a cascade impactor and the alternative instrument can be shown, then the alternative instrument may be used throughout the study.



**▼ M4**

30. A second device, such as a gravimetric filter or an impinger/gas bubbler, should be used in parallel to the primary instrument to confirm the collection efficiency of the primary instrument. The mass concentration obtained by particle size analysis should be within reasonable limits of the mass concentration obtained by filter analysis [see GD 39 (2)]. If equivalence can be demonstrated at all concentrations tested in the early phase of the study, then further confirmatory measurements may be omitted. For the sake of animal welfare, measures should be taken to minimise inconclusive data which may lead to a need to repeat a study.
31. Particle sizing should be performed for vapours if there is any possibility that vapour condensation may result in the formation of an aerosol, or if particles are detected in a vapour atmosphere with potential for mixed phases.

**OBSERVATIONS**

32. The animals should be clinically observed before, during, and after the exposure period. More frequent observations may be indicated depending on the response of the animals during exposure. When animal observation is hindered by the use of animal restraint tubes, poorly lit whole body chambers, or opaque atmospheres, animals should be carefully observed after exposure. Observations before the next day's exposure can assess any reversibility or exacerbation of toxic effects.
33. All observations are recorded with individual records being maintained for each animal. When animals are killed for humane reasons or found dead, the time of death should be recorded as precisely as possible.
34. Cage-side observations should include changes in the skin and fur, eyes, and mucous membranes; changes in the respiratory and circulatory systems; changes in the nervous system; and changes in somatomotor activity and behaviour patterns. Attention should be directed to observations of tremors, convulsions, salivation, diarrhoea, lethargy, sleep, and coma. The measurement of rectal temperatures may provide supportive evidence of reflex bradypnea or hypo/hyperthermia related to treatment or confinement. Additional assessments may be included in the study protocol such as kinetics, biomonitoring, lung function, retention of poorly soluble materials that accumulate in lung tissue, and behavioural changes.

**BODY WEIGHTS**

35. Individual animal weights should be recorded shortly before the first exposure (day 0), twice weekly thereafter (for example: on Fridays and Mondays to demonstrate recovery over an exposure-free weekend, or at a time interval to allow assessment of systemic toxicity), and at the time of death or euthanasia. If there are no effects in the first 4 weeks, body weights may be measured weekly for the remainder of the study. Satellite (reversibility) animals (if used) should continue to be weighed weekly throughout the recovery period. At study termination, all animals should be weighed shortly before sacrifice to allow for an unbiased calculated of organ to body weight ratios.

**FOOD AND WATER CONSUMPTION**

36. Food consumption should be measured weekly. Water consumption may also be measured.

▼ **M4**

## CLINICAL PATHOLOGY

37. Clinical pathology assessments should be made for all animals, including controls and satellite (reversibility) animals, when they are sacrificed. The time interval between the end of exposure and blood collection should be recorded, particularly when the reconstitution of the addressed endpoint is rapid. Sampling following the end of exposure is indicated for those parameters with a short plasma half-time (e.g. COHb, CHE, and MetHb).
38. Table 1 lists the clinical pathology parameters that are generally required for all toxicology studies. Urinalysis is not required on a routine basis, but may be performed when deemed useful based on expected or observed toxicity. The study director may choose to assess additional parameters in order to better characterise a test chemical's toxicity (e.g. cholinesterase, lipids, hormones, acid/base balance, methaemoglobin or Heinz bodies, creatine kinase, myeloid/erythroid ratio, troponins, arterial blood gases, lactate dehydrogenase, sorbital dehydrogenase, glutamate dehydrogenase, and gamma glutamyl transpeptidase).

Table 1

## Standard Clinical Pathology Parameters

Haematology	
Erythrocyte count	Total leukocyte count
Haematocrit	Differential leukocyte count
Haemoglobin concentration	Platelet count
Mean corpuscular haemoglobin	Clotting potential (select one):
Mean corpuscular volume	— Prothrombin time
Mean corpuscular haemoglobin concentration	— Clotting time
Reticulocytes	— Partial thromboplastin time
Clinical Chemistry	
Glucose (*)	Alanine aminotransferase
Total cholesterol	Aspartate aminotransferase
Triglycerides	Alkaline phosphatase
Blood urea nitrogen	Potassium
Total bilirubin	Sodium
Creatinine	Calcium
Total protein	Phosphorus
Albumin	Chloride
Globulin	
Urinalysis (optional)	
Appearance (colour and turbidity)	Total protein
Volume	Glucose
Specific gravity or osmolality	Blood/blood cells
pH	

(\*) Because a lengthy fasting period can introduce bias in glucose measurements for the treated versus control animals, the study director should determine whether it is appropriate to fast the animals. If a fasting period is used, it should be appropriate to the species used; for the rat this may be 16 h (overnight fasting). Determination of fasting glucose may be carried out after overnight fasting during the last exposure week, or after overnight fasting prior to necropsy (in the latter case together with all other clinical pathology parameters).

**▼ M4**

39. When there is evidence that the lower respiratory tract (i.e. the alveoli) is the primary site of deposition and retention, then bronchoalveolar lavage (BAL) may be the technique of choice to quantitatively analyse hypothesis-based dose-effect parameters focusing on alveolitis, pulmonary inflammation, and phospholipidosis. This allows for dose-response and time-course changes of alveolar injury to be suitably probed. The BAL fluid may be analysed for total and differential leukocyte counts, total protein, and lactate dehydrogenase. Other parameters that may be considered are those indicative of lysosomal injury, phospholipidosis, fibrosis, and irritant or allergic inflammation which may include the determination of pro-inflammatory cytokines/chemokines. BAL measurements generally complement the results from histopathology examinations but cannot replace them. Guidance on how to perform lung lavage can be found in GD 39 (2).

## OPHTHALMOLOGICAL EXAMINATION

40. Using an ophthalmoscope or an equivalent device, ophthalmological examinations of the fundus, refractive media, iris, and conjunctivae should be performed for all animals prior to the administration of the test chemical, and for all high concentration and control groups at termination. If changes in the eyes are detected, all animals in the other groups should be examined including the satellite (reversibility) group.

## GROSS PATHOLOGY AND ORGAN WEIGHTS

41. All test animals, including those which die during the test or are removed from the study for animal welfare reasons, should be subjected to complete exsanguination (if feasible) and gross necropsy. The time between the end of each animal's last exposure and its sacrifice should be recorded. If a necropsy cannot be performed immediately after a dead animal is discovered, the animal should be refrigerated (not frozen) at a temperature low enough to minimise autolysis. Necropsies should be performed as soon as possible, normally within a day or two. All gross pathological changes should be recorded for each animal with particular attention to any changes in the respiratory tract.
42. Table 2 lists the organs and tissues that should be preserved in a suitable medium during gross necropsy for histopathological examination. The preservation of the [bracketed] organs and tissues and any other organs and tissues is at the discretion of the study director. The **bolded** organs should be trimmed and weighed wet as soon as possible after dissection to avoid drying. The thyroid and epididymides should only be weighed if needed because trimming artefacts may hinder histopathological evaluation. Tissues and organs should be fixed in 10 % buffered formalin or another suitable fixative as soon as necropsy is performed, and no less than 24-48 hours prior to trimming depending on the fixative to be used.

▼ **M4**

Table 2

**Organs and Tissues Preserved During Gross Necropsy**

<b>Adrenals</b>	Oesophagus
Aorta	[Olfactory bulb]
Bone marrow (and/or fresh aspirate)	<b>Ovaries</b>
<b>Brain</b> (including sections of cerebrum, cerebellum, and medulla/pons)	Pancreas
Caecum	Parathyroids
Colon	Peripheral nerve (sciatic or tibial, preferably close to muscle)
Duodenum	Pituitary
<b>[Epididymides]</b>	Prostate
[Eyes (retina, optic nerve) and eyelids]	Rectum
Femur and stifle joint	Salivary glands
Gallbladder (where present)	Seminal vesicles
[Harderian glands]	Skin
<b>Heart</b>	Spinal cord (cervical, mid-thoracic, and lumbar)
Ileum	<b>Spleen</b>
Jejunum	Sternum
<b>Kidneys</b>	Stomach
[Lacrimal glands (extraorbital)]	Teeth
Larynx (3 levels including the base of the epiglottis)	<b>Testes</b>
<b>Liver</b>	<b>Thymus</b>
<b>Lung</b> (all lobes at one level, including main bronchi)	<b>Thyroids</b>
Lymph nodes from the hilar region of the lung, especially for poorly soluble particulate test chemicals. For more in depth examinations and/or studies with immunological focus, additional lymph nodes may be considered, e.g. those from the mediastinal, cervical/submandibular and/or auricular regions.	[Tongue]
Lymph nodes (distal from the portal-of-entry)	Trachea (at least 2 levels including 1 longitudinal section through the carina and 1 transverse section)
Mammary gland (female)	[Ureter]
Muscle (thigh)	[Urethra]
Nasopharyngeal tissues (at least 4 levels; 1 level to include the nasopharyngeal duct and the Nasal Associated Lymphoid Tissue (NALT))	Urinary bladder
	<b>Uterus</b>
	Target organs
	All gross lesions and masses

**▼ M4**

43. The lungs should be removed intact, weighed, and instilled with a suitable fixative at a pressure of 20-30 cm of water to ensure that lung structure is maintained (5). Sections should be collected for all lobes at one level, including main bronchi, but if lung lavage is performed, the unlavaged lobe should be sectioned at three levels (not serial sections).
44. At least 4 levels of the nasopharyngeal tissues should be examined, one of which should include the nasopharyngeal duct (5) (6) (7) (8) (9) to allow adequate examination of the squamous, transitional (non-ciliated respiratory), respiratory (ciliated respiratory) and olfactory epithelium, and the draining lymphatic tissue (NALT) (10) (11). Three levels of the larynx should be examined, and one of these levels should include the base of the epiglottis (12). At least two levels of the trachea should be examined including one longitudinal section through the carina of the bifurcation of the extrapulmonary bronchi and one transverse section.

**HISTOPATHOLOGY**

45. A histopathological evaluation of all the organs and tissues listed in Table 2 should be performed for the control and high concentration groups, and for all animals which die or are sacrificed during the study. Particular attention should be paid to the respiratory tract, target organs, and gross lesions. The organs and tissues that have lesions in the high concentration group should be examined in all groups. The study director may choose to perform histopathological evaluations for additional groups to demonstrate a clear concentration response. When a satellite (reversibility) group is used, histopathological evaluation should be performed for all tissues and organs identified as showing effects in the treated groups. If there are excessive early deaths or other problems in the high exposure group that compromise the significance of the data, the next lower concentration should be examined histopathologically. An attempt should be made to correlate gross observations with microscopic findings.

**DATA AND REPORTING****Data**

46. Individual animal data on body weights, food consumption, clinical pathology, gross pathology, organ weights, and histopathology should be provided. Clinical observation data should be summarised in tabular form showing for each test group the number of animals used, the number of animals displaying specific signs of toxicity, the number of animals found dead during the test or killed for humane reasons, time of death of individual animals, a description and time course of toxic effects and reversibility, and necropsy findings. All results, quantitative and incidental, should be evaluated by an appropriate statistical method. Any generally accepted statistical method may be used and the statistical methods should be selected during the design of the study.

**Test Report**

47. The test report should include the following information, as appropriate:

*Test animals and husbandry*

- Description of caging conditions, including: number (or change in number) of animals per cage, bedding material, ambient temperature and relative humidity, photoperiod, and identification of diet.

**▼ M4**

- Species/strain used and justification for using a species other than the rat. Source and historical data may be provided, if they are for animals exposed under similar exposure, housing, and fasting conditions.
- Number, age, and sex of animals.
- Method of randomisation.
- Description of any pre-test conditioning including diet, quarantine, and treatment for disease.

*Test chemical*

- Physical nature, purity, and, where relevant, physico-chemical properties (including isomerisation).
- Identification data and Chemical Abstract Services (CAS) Registry Number, if known.

*Vehicle*

- Justification for use of vehicle and justification for choice of vehicle (if other than water).
- Historical or concurrent data demonstrating that the vehicle does not interfere with the outcome of the study.

*Inhalation chamber*

- Detailed description of the inhalation chamber including volume and a diagram.
- Source and description of equipment used for the exposure of animals as well as generation of atmosphere.
- Equipment for measuring temperature, humidity, particle-size, and actual concentration.
- Source of air and system used for conditioning.
- Methods used for calibration of equipment to ensure a homogeneous test atmosphere.
- Pressure difference (positive or negative).
- Exposure ports per chamber (nose-only); location of animals in the chamber (whole-body).
- Stability of the test atmosphere.
- Location of temperature and humidity sensors and sampling of test atmosphere in the chamber.
- Treatment of air supplied/extracted.
- Air flow rates, air flow rate/exposure port (nose-only), or animal load/chamber (whole-body).
- Time to inhalation chamber equilibrium (t95).
- Number of volume changes per hour.
- Metering devices (if applicable).

*Exposure data*

- Rationale for target concentration selection in the main study.

**▼ M4**

- Nominal concentrations (total mass of test chemical generated into the inhalation chamber divided by the volume of air passed through the chamber).
- Actual test chemical concentrations collected from the animals' breathing zone; for mixtures that produce heterogeneous physical forms (gases, vapours, aerosols), each may be analysed separately.
- All air concentrations should be reported in units of mass (mg/l, mg/m<sup>3</sup>, etc.) rather than in units of volume (ppm, ppb, etc.).
- Particle size distribution, mass median aerodynamic diameter (MMAD), and geometric standard deviation ( $\sigma_g$ ), including their methods of calculation. Individual particle size analyses should be reported.

*Test conditions*

- Details of test chemical preparation, including details of any procedures used to reduce the particle size of solid materials or to prepare solutions of the test chemical.
- A description (preferably including a diagram) of the equipment used to generate the test atmosphere and to expose the animals to the test atmosphere.
- Details of the equipment used to monitor chamber temperature, humidity, and chamber airflow (i.e. development of a calibration curve).
- Details of the equipment used to collect samples for determination of chamber concentration and particle size distribution.
- Details of the chemical analytical method used and method validation (including efficiency of recovery of test chemical from the sampling medium).
- Method of randomisation in assigning animals to test and control groups.
- Details of food and water quality (including diet type/source, water source).
- The rationale for the selection of test concentrations.

*Results*

- Tabulation of chamber temperature, humidity, and airflow.
- Tabulation of chamber nominal and actual concentration data.
- Tabulation of particle size data including analytical sample collection data, particle size distribution, and calculations of the MMAD and  $\sigma_g$ .
- Tabulation of response data and concentration level for each animal (i.e. animals showing signs of toxicity including mortality, nature, severity, time of onset, and duration of effects).

**▼M4**

- Tabulation of individual animal weights.
- Tabulation of food consumption
- Tabulation of clinical pathology data
- Necropsy findings and histopathological findings for each animal, if available.

*Discussion and interpretation of results*

- Particular emphasis should be made to the description of methods used to meet the criteria of this Test Method, e.g. the limit concentration or the particle size.
- The respirability of particles in light of the overall findings should be addressed, especially if the particle-size criteria could not be met.
- The consistency of methods used to determine nominal and actual concentrations, and the relation of actual concentration to nominal concentration should be included in the overall assessment of the study.
- The likely cause of death and predominant mode of action (systemic versus local) should be addressed.
- An explanation should be provided if there was a need to humanely sacrifice animals in pain or showing signs of severe and enduring distress, based on the criteria in the OECD Guidance Document on Humane Endpoints (3).
- The target organ(s) should be identified.
- The NOAEL and LOAEL should be determined.

*LITERATURE:*

- (1) OECD (1981). *Subchronic Inhalation Toxicity Testing*, Original Test Guideline No 413, Environment Directorate, OECD, Paris.
- (2) OECD (2009). *Guidance Document on Acute Inhalation Toxicity Testing*, Environmental Health and Safety Monograph Series on Testing and Assessment No 39, ENV/JM/MONO(2009)28, OECD, Paris.
- (3) OECD (2000). *Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation*, Environmental Health and Safety Monograph Series on Testing and Assessment No 19, ENV/JM/MONO(2000)7, OECD, Paris.
- (4) Whalan.E and Redden JC (1994). *Interim Policy for Particle Size and Limit Concentration Issues in Inhalation Toxicity Studies*. Office of Pesticide Programs, United States Environmental Protection Agency.
- (5) Dungworth DL, Tyler WS, Plopper CE (1985). *Morphological Methods for Gross and Microscopic Pathology (Chapter 9) in Toxicology of Inhaled Material*, Witschi, H.P. and Brain, J.D. (eds), Springer Verlag Heidelberg, pp. 229-258.
- (6) Young JT (1981). *Histopathological examination of the rat nasal cavity*. *Fundam. Appl. Toxicol.* 1: 309-312.
- (7) Harkema JR (1990). *Comparative pathology of the nasal mucosa in laboratory animals exposed to inhaled irritants*. *Environ. Health Perspect.* 85: 231-238.



**▼M4**

- (8) Woutersen RA, Garderen-Hoetmer A, van Slootweg PJ, Feron VJ (1994). Upper respiratory tract carcinogenesis in experimental animals and in humans. In: Waalkes MP and Ward JM (eds) Carcinogenesis. Target Organ Toxicology Series, Raven Press, New York, 215-263.
- (9) Mery S, Gross EA, Joyner DR, Godo M, Morgan KT (1994). Nasal diagrams: A tool for recording the distribution of nasal lesions in rats and mice. *Toxicol. Pathol.* 22: 353-372.
- (10) Kuper CF, Koornstra PJ, Hamelers DMH, Biewenga J, Spit BJ, Duijvestijn AM, Breda Vriesman van PJC, Sminia T (1992). The role of nasopharyngeal lymphoid tissue. *Immunol. Today* 13: 219-224.
- (11) Kuper CF, Arts JHE, Feron VJ (2003). Toxicity to nasal-associated lymphoid tissue. *Toxicol. Lett.* 140-141: 281-285.
- (12) Lewis DJ (1981). Mitotic Indices of Rat Laryngeal Epithelia. *Journal of Anatomy* 132(3): 419-428.
- (13) Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006 (OJ L 353, 31.12.2008, p. 1).

▼ M4

*Appendix 1*

DEFINITION

**Test chemical:** Any substance or mixture tested using this Test Method.

**▼M4****B.30. CHRONIC TOXICITY STUDIES**

## INTRODUCTION

1. This Test Method is equivalent to OECD Test Guideline (TG) 452 (2009). The original TG 452 was adopted in 1981. Development of this revised Test Method B.30 was considered necessary in order to reflect recent developments in the field of animal welfare and regulatory requirements (1) (2) (3) (4). The updating of this Test Method B.30 has been carried out in parallel with revisions of Chapter B.32 of this Annex, Carcinogenicity Studies, and Chapter B.33 of this Annex, Combined Chronic Toxicity/Carcinogenicity studies, with the objective of obtaining additional information from the animals used in the study and providing further detail on dose selection. This Test Method is designed to be used in the testing of a broad range of chemicals, including pesticides and industrial chemicals.
  
2. The majority of chronic toxicity studies are carried out in rodent species, and this Test Method is intended therefore to apply primarily to studies carried out in these species. Should such studies be required in non-rodent species, the principles and procedures outlined in this Test Method, together with those outlined in Chapter B.27 of this Annex, Repeated Dose 90-day Oral Toxicity Study in Non-Rodents (5), may also be applied, with appropriate modifications, as outlined in the OECD Guidance Document No 116 on the Design and Conduct of Chronic Toxicity and Carcinogenicity Studies (6).
  
3. The three main routes of administration used in chronic toxicity studies are oral, dermal and inhalation. The choice of the route of administration depends on the physical and chemical characteristics of the test chemical and the predominant route of exposure of humans. Additional information on choice of route of exposure is provided in the OECD Guidance Document No 116 (6).
  
4. This Test Method focuses on exposure via the oral route, the route most commonly used in chronic toxicity studies. While long-term chronic toxicity studies involving exposure via the dermal or inhalation routes may also be necessary for human health risk assessment and/or may be required under certain regulatory regimes, both routes of exposure involve considerable technical complexity. Such studies will need to be designed on a case-by-case basis, although the Test Method outlined here for the assessment and evaluation of chronic toxicity by oral administration could form the basis of a protocol for inhalation and/or dermal studies, with respect to recommendations for treatment periods, clinical and pathology parameters, etc. OECD Guidance is available on the administration of test chemicals by the inhalation (6) (7) and dermal routes (6). Chapter B.8 of this Annex (8) and Chapter B.29 of this Annex (9), together with the OECD Guidance Document on acute inhalation testing (7), should be specifically consulted in the design of longer term studies involving exposure via the inhalation route. Chapter B.9 of this Annex (10) should be consulted in the case of testing carried out by the dermal route.

**▼ M4**

5. The chronic toxicity study provides information on the possible health hazards likely to arise from repeated exposure over a considerable part of the lifespan of the species used. The study will provide information on the toxic effects of the test chemical; indicate target organs and the possibility of accumulation. It can also provide an estimate of the no-observed-adverse effect level which can be used for establishing safety criteria for human exposure. The need for careful clinical observations of the animals, so as to obtain as much information as possible, is also stressed.
  
6. The objectives of studies covered by this Test Method include:
  - The identification of the chronic toxicity of a test chemical;
  
  - The identification of target organs;
  
  - Characterisation of the dose-response relationship;
  
  - Identification of a no-observed-adverse-effect level (NOAEL) or point of departure for establishment of a Benchmark Dose (BMD);
  
  - The prediction of chronic toxicity effects at human exposure levels;
  
  - Provision of data to test hypotheses regarding mode of action (6).

## INITIAL CONSIDERATIONS

7. In the assessment and evaluation of the toxicological characteristics of a test chemical, all available information on the test chemical should be considered by the testing laboratory prior to conducting the study, in order to focus the design of the study to more efficiently test for chronic toxicity potential and to minimize animal usage. Information that will assist in the study design includes the identity, chemical structure, and physico-chemical properties of the test chemical; any information on the mode of action; results of any *in vitro* or *in vivo* toxicity tests; anticipated use(s) and potential for human exposure; available (Q)SAR data and toxicological data on structurally-related chemicals; available toxicokinetic data (single dose and also repeat dose kinetics where available) and data derived from other repeated exposure studies. The determination of chronic toxicity should only be carried out after initial information on toxicity has been obtained from repeated dose 28-day and/or 90-day toxicity tests. A phased testing approach to chronic toxicity testing should be considered as part of the overall assessment of the potential adverse health effects of a particular test chemical (11) (12) (13) (14).
  
8. The statistical methods most appropriate for the analysis of results, given the experimental design and objectives, should be established before commencing the study. Issues to consider include whether the statistics should include adjustment for survival and analysis in the event of premature termination of one or more groups. Guidance on the appropriate statistical analyses and key references to internationally accepted statistical methods are given in Guidance Document No 116 (6), and also in Guidance Document No 35 on the analysis and evaluation of chronic toxicity and carcinogenicity studies (15).

**▼ M4**

9. In conducting a chronic toxicity study, the guiding principles and considerations outlined in the OECD Guidance Document No 19 on the recognition, assessment, and use of clinical signs as humane endpoints for experimental animals used in safety evaluation (16), in particular paragraph 62 thereof, should always be followed. This paragraph states that *'In studies involving repeated dosing, when an animal shows clinical signs that are progressive, leading to further deterioration in condition, an informed decision as to whether or not to humanely kill the animal should be made. The decision should include consideration as to the value of the information to be gained from the continued maintenance of that animal on study relative to its overall condition. If a decision is made to leave the animal on test, the frequency of observations should be increased, as needed. It may also be possible, without adversely affecting the purpose of the test, to temporarily stop dosing if it will relieve the pain or distress, or reduce the test dose.'*
10. Detailed guidance on and discussion of the principles of dose selection for chronic toxicity and carcinogenicity studies can be found in Guidance Document No 116 (6), as well as two International Life Sciences Institute publications (17) (18). The core dose selection strategy is dependent on the primary objective or objectives of the study (paragraph 6). In selecting appropriate dose levels, a balance should be achieved between hazard screening on the one hand and characterisation of low-dose responses and their relevance on the other. This is particularly relevant in the situation where a combined chronic toxicity and carcinogenicity study (Chapter B.33 of this Annex) is to be carried out (paragraph 11).
11. Consideration should be given to carrying out a combined chronic toxicity and carcinogenicity study (Chapter B.33 of this Annex), rather than separate execution of a chronic toxicity study (this Test Method B.30) and carcinogenicity study (Chapter B.32 of this Annex). The combined test provides greater efficiency in terms of time and cost compared to conducting two separate studies, without compromising the quality of the data in either the chronic phase or the carcinogenicity phase. Careful consideration should however be given to the principles of dose selection (paragraphs 9 and 20-25) when undertaking a combined chronic toxicity and carcinogenicity study (Chapter B.33 of this Annex), and it is also recognised that separate studies may be required under certain regulatory frameworks.
12. Definitions used in the context of this Test Method can be found at the end of this chapter and in the Guidance Document No 116 (6).

**PRINCIPLE OF THE TEST**

13. The test chemical is administered daily in graduated doses to several groups of experimental animals, normally for a period of 12 months, although longer or shorter durations may also be chosen depending on regulatory requirements (see paragraph 33). This duration is chosen to be sufficiently long to allow any effects of cumulative toxicity to become manifest, without the confounding effects of geriatric changes. Deviations from exposure duration of 12 months should be justified, particularly in the case of shorter durations. The test chemical is normally administered by the oral route although testing by the inhalation or dermal route may also be appropriate. The study design may also include one or more interim kills, e.g. at 3 and 6 months, and additional groups of animals may be included to accommodate this (see paragraph 19). During the period of administration the animals are observed closely for signs of toxicity. Animals which die or are killed during the test are necropsied and, at the conclusion of the test, surviving animals are killed and necropsied.

**▼M4**

## DESCRIPTION OF THE METHOD

**Selection of animal species**

14. This Test Method primarily covers assessment and evaluation of chronic toxicity in rodents (see paragraph 2) although it is recognised that similar studies in non-rodents may be required under certain regulatory regimes. The choice of species should be justified. The design and conduct of chronic toxicity studies in non-rodent species, when required, should be based on the principles outlined in this Test Method together with those in Chapter B.27 of this Annex, Repeated Dose 90-day Oral Toxicity Study in Non-Rodents (5). Additional information on choice of species and strain is provided in Guidance Document No 116 (6).
  
15. In this Test Method, the preferred rodent species is the rat, although other rodent species, e.g. the mouse, may be used. Rats and mice have been preferred experimental models because of their relatively short life span, their widespread use in pharmacological and toxicological studies, their susceptibility to tumour induction, and the availability of sufficiently characterised strains. As a consequence of these characteristics, a large amount of information is available on their physiology and pathology. Young healthy adult animals of commonly used laboratory strains should be employed. The chronic toxicity study should be carried out in animals from the same strain and source as those used in preliminary toxicity study(ies) of shorter duration. The females should be nulliparous and non-pregnant.

**Housing and feeding conditions**

16. Animals may be housed individually, or be caged in small groups of the same sex; individual housing should be considered only if scientifically justified (19) (20) (21). Cages should be arranged in such a way that possible effects due to cage placement are minimised. The temperature in the experimental animal room should be 22 °C ( $\pm$  3 °C). Although the relative humidity should be at least 30 % and preferably not exceed 70 % other than during room cleaning, the aim should be 50-60 %. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water. The diet should meet all the nutritional requirements of the species tested and the content of dietary contaminants including but not limited to pesticide residues, persistent organic pollutants, phytoestrogens, heavy metals and mycotoxins, that might influence the outcome of the test, should be as low as possible. Analytical information on the nutrient and dietary contaminant levels should be generated periodically, at least at the beginning of the study and when there is a change in the batch used, and should be included in the final report. Analytical information on the drinking water used in the study should similarly be provided. The choice of diet may be influenced by the need to ensure a suitable admixture of a test chemical and to meet the nutritional requirements of the animals when the test chemical is administered by the dietary route.

**Preparation of animals**

17. Healthy animals, which have been acclimated to laboratory conditions for at least 7 days and have not been subjected to previous experimental procedures, should be used. In the case of rodents, dosing of the animals should begin as soon as possible after weaning and acclimatisation and preferably before the animals are 8 weeks old. The test animals should be characterised as to species, strain, source, sex, weight and age. At the commencement of the study, the weight variation for each sex of animals used should be minimal and not exceed  $\pm$  20 % of the mean weight of all

**▼ M4**

the animals within the study, separately for each sex. Animals should be randomly assigned to the control and treatment groups. After randomisation, there should be no significant differences in mean body weights between groups within each sex. If there are statistically significant differences, then the randomisation step should be repeated, if possible. Each animal should be assigned a unique identification number, and permanently marked with this number by tattooing, microchip implant, or other suitable method.

**PROCEDURE****Number and sex of animals**

18. Both sexes should be used. A sufficient number of animals should be used so that at the end of the study enough animals in every group are available for thorough biological and statistical evaluation. For rodents, at least 20 animals per sex per group should normally be used at each dose level, while for non-rodents a minimum of 4 per sex per group is recommended. In studies involving mice, additional animals may be needed in each dose group to conduct all required haematological determinations.

**Provision for interim kills, satellite groups and sentinel animals**

19. The study may make provision for interim kills (at least 10 animals/sex/group), e.g. at 6 months, to provide information on progression of toxicological changes and mechanistic information, if scientifically justified. Where such information is already available from previous repeat dose toxicity studies on the test chemical, interim kills may not be scientifically justified. Satellite groups may also be included to monitor the reversibility of any toxicological changes induced by the test chemical under investigation; these will normally be restricted to the highest dose level of the study plus control. An additional group of sentinel animals (typically 5 animals per sex) may also be included for monitoring of disease status, if necessary, during the study (22). If interim kills or inclusion of satellite or sentinel groups are planned, the number of animals included in the study design should be increased by the number of animals scheduled to be killed before the completion of the study. These animals should normally undergo the same observations, including body weight, food/water consumption, haematological and clinical biochemistry measurements and pathological investigations as the animals in the chronic toxicity phase of the main study, although provision may also be made (in the interim kill groups) for measurements to be restricted to specific, key measures such as neurotoxicity or immunotoxicity.

**Dose groups and dosage**

20. Guidance on all aspects of dose selection and dose level spacing is provided in Guidance Document No 116 (6). At least three dose levels and a concurrent control should be used, except where a limit test is conducted (see paragraph 27). Dose levels will generally be based on the results of shorter-term repeated dose or range finding studies and should take into account any existing toxicological and toxicokinetic data available for the test chemical or related chemicals.

**▼ M4**

21. Unless limited by the physical-chemical nature or biological effects of the test chemical, the highest dose level should normally be chosen to identify the principal target organs and toxic effects while avoiding suffering, severe toxicity, morbidity, or death. While taking into account the factors outlined in paragraph 22 below, the highest dose level should be chosen to elicit evidence of toxicity, as evidenced by, for example, depression of body weight gain (approximately 10 %).
22. However, dependent on the objectives of the study (see paragraph 6), a top dose lower than the dose providing evidence of toxicity may be chosen, e.g. if a dose elicits an adverse effect of concern that nonetheless has little impact on lifespan or body weight. The top dose should not exceed 1 000 mg/kg body weight/day (limit dose, see paragraph 27).
23. Dose levels and dose level spacing may be selected to establish a dose-response and a NOAEL or other intended outcome of the study, e.g. a BMD (see paragraph 25) at the lowest dose level. Factors that should be considered in the placement of lower doses include the expected slope of the dose-response curve, the doses at which important changes may occur in metabolism or mode of toxic action, where a threshold is expected, or where a point of departure for low-dose extrapolation is expected.
24. The dose level spacing selected will depend on the characteristics of the test chemical, and cannot be prescribed in this Test Method, but two to four fold intervals frequently provide good test performance when used for setting the descending dose levels and addition of a fourth test group is often preferable to using very large intervals (e.g. more than a factor of about 6-10) between dosages. In general the use of factors greater than 10 should be avoided, and should be justified if used.
25. As outlined further in Guidance Document No 116 (6), points to be considered in dose selection include:
  - Known or suspected nonlinearities or inflection points in the dose-response;
  - Toxicokinetics, and dose ranges where metabolic induction, saturation, or nonlinearity between external and internal doses does or does not occur;
  - Precursor lesions, markers of effect, or indicators of the operation of key underlying biological processes;
  - Key (or suspected) aspects of mode of action, such as doses at which cytotoxicity begins to arise, hormone levels are perturbed, homeostatic mechanisms are overwhelmed, etc.;
  - Regions of the dose-response curve where particularly robust estimation is needed, e.g. in the range of the anticipated BMD or a suspected threshold;
  - Consideration of anticipated human exposure levels.



**▼M4**

26. The control group shall be an untreated group or a vehicle-control group if a vehicle is used in administering the test chemical. Except for treatment with the test chemical, animals in the control group should be handled in an identical manner to those in the test groups. If a vehicle is used, the control group shall receive the vehicle in the highest volume used among the dose groups. If a test chemical is administered in the diet, and causes significantly reduced dietary intake due to the reduced palatability of the diet, an additional pair-fed control group may be useful, to serve as a more suitable control.
  
27. If it can be anticipated, based on information from preliminary studies, that a test at one dose level, equivalent to at least 1 000 mg/kg body weight/day, using the procedures described for this study, is unlikely to produce adverse effects and if toxicity would not be expected based upon data from structurally related chemicals, then a full study using three dose levels may not be considered necessary. A limit of 1 000 mg/kg body weight/day may apply except when human exposure indicates the need for a higher dose level to be used.

**Preparation of doses and administration of test chemical**

28. The test chemical is normally administered orally, via the diet or drinking water, or by gavage. Additional information on routes and methods of administration is provided in Guidance Document No 116 (6). The route and method of administration is dependent on the purpose of the study, the physical/chemical properties of the test chemical, its bioavailability and the predominant route and method of exposure of humans. A rationale should be provided for the chosen route and method of administration. In the interests of animal welfare, oral gavage should normally be selected only for those agents for which this route and method of administration reasonably represent potential human exposure (e.g. pharmaceuticals). For dietary or environmental chemicals including pesticides, administration is typically via the diet or drinking water. However, for some scenarios, e.g. occupational exposure, administration via other routes may be more appropriate.
  
29. Where necessary, the test chemical is dissolved or suspended in a suitable vehicle. Consideration should be given to the following characteristics of the vehicle and other additives, as appropriate: effects on the absorption, distribution, metabolism, or retention of the test chemical; effects on the chemical properties of the test chemical which may alter its toxic characteristics; and effects on the food or water consumption or the nutritional status of the animals. It is recommended that, wherever possible, the use of an aqueous solution/suspension be considered first, followed by consideration of a solution/emulsion in oil (e.g. corn oil) and then by possible solution in other vehicles. For vehicles other than water, the toxic characteristics of the vehicle should be known. Information should be available on the stability of the test chemical and the homogeneity of dosing solutions or diets (as appropriate) under the conditions of administration (e.g. diet).
  
30. For chemicals administered via the diet or drinking water it is important to ensure that the quantities of the test chemical involved do not interfere with normal nutrition or water balance. In long-term toxicity studies using dietary administration, the concentration of the test chemical in the feed should not normally exceed an upper limit of 5 % of the total diet, in order to avoid nutritional imbalances. When the test chemical is administered in the diet, either a constant dietary concentration (mg/kg diet or ppm) or a constant dose level in terms of the animal's body weight (mg/kg body weight), calculated on a weekly basis, may be used. The alternative used should be specified.

**▼M4**

31. In the case of oral administration, the animals are dosed with the test chemical daily (seven days each week), normally for a period of 12 months (see also paragraph 33), although a longer duration may be required depending on regulatory requirements. Any other dosing regime, e.g. five days per week, needs to be justified. In the case of dermal administration, animals are normally treated with the test chemical for at least 6 hours per day, 7 days per week, as specified in Chapter B.9 of this Annex (10), for a period of 12 months. Exposure by the inhalation route is carried out for 6 hours per day, 7 days per week, but exposure for 5 days per week may also be used, if justified. The period of exposure will normally be for a period of 12 months. If rodent species other than rats are exposed nose-only, maximum exposure durations may be adjusted to minimise species-specific distress. A rationale should be provided when using an exposure duration of less than 6 hours per day. See also Chapter B.8 of this Annex (8).
  
32. When the test chemical is administered by gavage to the animals this should be done using a stomach tube or a suitable intubation cannula, at similar times each day. Normally a single dose will be administered once daily, where for example a chemical is a local irritant, it may be possible to maintain the daily dose-rate by administering it as a split dose (twice a day). The maximum volume of liquid that can be administered at one time depends on the size of the test animal. The volume should be kept as low as practical, and should not normally exceed 1 ml/100 g body weight for rodents (22). Variability in test volume should be minimised by adjusting the concentration to ensure a constant volume at all dose levels. Potentially corrosive or irritant chemicals are the exception, and need to be diluted to avoid severe local effects. Testing at concentrations that are likely to be corrosive or irritant to the gastrointestinal tract should be avoided.

**Duration of study**

33. While this Test Method primarily is designed as a 12 month chronic toxicity study, the study design also allows for and can be applied to either shorter (e.g. 6 or 9 months) or longer (e.g. 18 or 24 months) duration studies, depending on the requirements of particular regulatory regimes or for specific mechanistic purposes. Deviations from an exposure duration of 12 months should be justified, particularly in the case of shorter durations. Satellite groups included to monitor the reversibility of any toxicological changes induced by the test chemical under investigation should be maintained without dosing for a period not less than 4 weeks and not more than one third of the total study duration after cessation of exposure. Further guidance, including consideration of survival in the study, is provided in Guidance Document No 116 (6).

**OBSERVATIONS**

34. All animals should be checked for morbidity or mortality, usually at the beginning and end of each day, including at weekends and holidays. General clinical observations should be made at least once a day, preferably at the same time(s) each day, taking into consideration the peak period of anticipated effects after dosing in the case of gavage administration.

**▼ M4**

35. Detailed clinical observations should be made on all animals at least once prior to the first exposure (to allow for within-subject comparisons), at the end of the first week of the study and monthly thereafter. The protocol for observations should be arranged such that variations between individual observers are minimised and independent of test group. These observations should be made outside the home cage, preferably in a standard arena and at similar times on each occasion. They should be carefully recorded, preferably using scoring systems, explicitly defined by the testing laboratory. Efforts should be made to ensure that variations in the observation conditions are minimal. Signs noted should include, but not be limited to, changes in skin, fur, eyes, mucous membranes, occurrence of secretions and excretions and autonomic activity (e.g. lacrimation, piloerection, pupil size, and unusual respiratory pattern). Changes in gait, posture and response to handling as well as the presence of clonic or tonic movements, stereotypies (e.g. excessive grooming, repetitive circling) or bizarre behaviour (e.g. self-mutilation, walking backwards) should also be recorded (24).
  
36. Ophthalmological examination, using an ophthalmoscope or other suitable equipment, should be carried out on all animals prior to the first administration of the test chemical. At the termination of the study, this examination should be preferably conducted in all animals but at least in the high dose and control groups. If treatment-related changes in the eyes are detected, all animals should be examined. If structural analysis or other information suggests ocular toxicity, then the frequency of ocular examination should be increased.
  
37. For chemicals where previous repeated dose 28-day and/or 90-day toxicity tests indicated the potential to cause neurotoxic effects, sensory reactivity to stimuli of different types (24) (e.g. auditory, visual and proprioceptive stimuli) (25), (26), (27), assessment of grip strength (28) and motor activity assessment (29) may optionally be conducted before commencement of the study and at 3 month periods after study initiation up to and including 12 months, as well as at study termination (if longer than 12 months). Further details of the procedures that could be followed are given in the respective references. However, alternative procedures than those referenced could also be used.
  
38. For chemicals where previous repeated dose 28-day and/or 90-day toxicity tests indicated the potential to cause immunotoxic effects, further investigations of this endpoint may optionally be conducted at termination.

**Body weight, food/water consumption and food efficiency**

39. All animals should be weighed at the start of treatment, at least once a week for the first 13 weeks, and at least monthly thereafter. Measurements of food consumption and food efficiency should be made at least weekly for the first 13 weeks and at least monthly thereafter. Water consumption should be measured at least weekly for the first 13 weeks and at least monthly thereafter when the chemical is administered in drinking water. Water consumption measurements should also be considered for studies in which drinking activity is altered.

**▼ M4****Haematology and clinical biochemistry**

40. In studies involving rodents, haematological examinations should be carried out in at least 10 male and 10 female animals per group, at 3, 6, and 12 months, as well as at study termination (if longer than 12 months), using the same animals throughout. In mice, satellite animals may be needed in order to conduct all required haematological determinations (see paragraph 18). In non-rodent studies, samples will be taken from smaller numbers of animals (e.g. 4 animals per sex and per group in dog studies), at interim sampling times and at termination as described for rodents. Measurements at 3 months, either in rodents or non-rodents, need not be conducted if no effect was seen on haematological parameters in a previous 90 day study carried out at comparable dose levels. Blood samples should be taken from a named site, for example by cardiac puncture or from the retro-orbital sinus, under anaesthesia.
  
41. The following list of parameters should be investigated (30): Total and differential leukocyte count, erythrocyte count, platelet count, haemoglobin concentration, haematocrit (packed cell volume), mean corpuscular volume (MCV), mean corpuscular haemoglobin (MCH), mean corpuscular haemoglobin concentration (MCHC), prothrombin time, and activated partial thromboplastin time. Other hematology parameters such as Heinz bodies or other atypical erythrocyte morphology or methaemoglobin may be measured as appropriate depending on the toxicity of the test chemical. Overall, a flexible approach should be adopted, depending on the observed and/or expected effect from a given test chemical. If the test chemical has an effect on the haematopoietic system, reticulocyte counts and bone marrow cytology may also be indicated, although these need not be routinely conducted.
  
42. Clinical biochemistry determinations to investigate major toxic effects in tissues and, specifically, effects on kidney and liver, should be performed on blood samples obtained from at least 10 male and 10 female animals per group at the same time intervals as specified for the haematological investigations, using the same animals throughout. In mice, satellite animals may be needed in order to conduct all required clinical biochemistry determinations. In non-rodent studies, samples will be taken from smaller numbers of animals (e.g. 4 animals per sex and per group in dog studies), at interim sampling times and at termination as described for rodents. Measurements at 3 months, either in rodents or non-rodents, need not be conducted if no effect was seen on clinical biochemistry parameters in a previous 90 day study carried out at comparable dose levels. Overnight fasting of the animals (with the exception of mice) prior to blood sampling is recommended. The following list of parameters should be investigated (30): glucose, urea (urea nitrogen), creatinine, total protein, albumin, calcium, sodium, potassium, total cholesterol, at least two appropriate tests for hepatocellular evaluation (alanine aminotransferase, aspartate aminotransferase, glutamate dehydrogenase, total bile acids) (31), and at least two appropriate tests for hepatobiliary evaluation (alkaline phosphatase, gamma glutamyl transferase, 5'-nucleotidase, total bilirubin, total bile acids) (31). Other clinical chemistry parameters such as fasting triglycerides, specific hormones and cholinesterase may be measured as appropriate, depending on the toxicity of the test chemical. Overall, there is a need for a flexible approach, depending on the observed and/or expected effect from a given test chemical.

▼ **M4**

43. Urinalysis determinations should be performed on at least 10 male and 10 female animals per group on samples collected at the same intervals as for haematology and clinical chemistry. Measurements at 3 months need not be conducted if no effect was seen on urinalysis in a previous 90 day study carried out at comparable dose levels. The following list of parameters was included in an expert recommendation on clinical pathology studies (30): appearance, volume, osmolality or specific gravity, pH, total protein, and glucose. Other determinations include ketone, urobilinogen, bilirubin, and occult blood. Further parameters may be employed where necessary to extend the investigation of observed effect(s).
44. It is generally considered that baseline haematological and clinical biochemistry variables are needed before treatment for dog studies, but need not be determined in rodent studies (30). However, if historical baseline data (see paragraph 50) are inadequate, consideration should be given to generating such data.

**Pathology***Gross necropsy*

45. All animals in the study shall normally be subjected to a full, detailed gross necropsy which includes careful examination of the external surface of the body, all orifices, and the cranial, thoracic and abdominal cavities and their contents. However provision may also be made (in the interim kill or satellite groups) for measurements to be restricted to specific, key measures such as neurotoxicity or immunotoxicity (see paragraph 19). These animals need not be subjected to necropsy and the subsequent procedures described in the following paragraphs. Sentinel animals may require necropsy on a case-by-case basis, at the discretion of the study director.
46. Organ weights should be collected from all animals, other than those excluded by the latter part of paragraph 45. The adrenals, brain, epididymides, heart, kidneys, liver, ovaries, spleen, testes, thyroid (weighed post-fixation, with parathyroids), and uterus of all animals (apart from those found moribund and/or intercurrently killed) should be trimmed of any adherent tissue, as appropriate, and their wet weight taken as soon as possible after dissection to prevent drying. In a study using mice, weighing of the adrenal glands is optional.
47. The following tissues should be preserved in the most appropriate fixation medium for both the type of tissue and the intended subsequent histopathological examination (32) (tissues in square brackets are optional):

all gross lesions	heart	pancreas	stomach (forestomach, glandular stomach)
adrenal gland	ileum	parathyroid gland	[teeth]
aorta	jejunum	peripheral nerve	testis
brain (including sections of cerebrum, cerebellum, and medulla/pons)	kidney	pituitary	thymus
caecum	lacrimal gland (exorbital)	prostate	thyroid
cervix	liver	rectum	[tongue]

▼ **M4**

coagulating gland	lung	salivary gland	trachea
colon	lymph nodes (both superficial and deep)	seminal vesicle	urinary bladder
duodenum	mammary gland (obligatory for females and, if visibly dissectable, from males)	skeletal muscle	uterus (including cervix)
epididymis	[upper respiratory tract, including nose, turbinates, and paranasal sinuses]	skin	[ureter]
eye (including retina)	oesophagus	spinal cord (at three levels: cervical, mid-thoracic, and lumbar)	[urethra]
[femur with joint]	[olfactory bulb]	spleen	vagina
gall bladder (for species other than rat)	ovary	[sternum],	section of bone marrow and/or a fresh bone marrow aspirate
Harderian gland			

In the case of paired organs, e.g. kidney, adrenal, both organs should be preserved. The clinical and other findings may suggest the need to examine additional tissues. Also any organs considered likely to be target organs based on the known properties of the test chemical should be preserved. In studies involving the dermal route of administration, the list of organs as set out for the oral route should be preserved, and specific sampling and preservation of the skin from the site of application is essential. In inhalation studies, the list of preserved and examined tissues from the respiratory tract should follow the recommendations of Chapters B.8 of this Annex (8) and Chapter B.29 of this Annex (9). For other organs/tissues (and in addition to the specifically preserved tissues from the respiratory tract) the list of organs as set out for the oral route should be examined.

*Histopathology*

48. Guidance is available on best practices in the conduct of toxicological pathology studies (32). The minimum histopathological examinations should be:

- all tissues from the high dose and control groups;
- all tissues from animals dying or killed during the study;
- all tissues showing macroscopic abnormalities;
- target tissues, or tissues which showed treatment-related changes in the high dose group, from all animals in all other dose groups;
- in the case of paired organs, e.g. kidney, adrenal, both organs should be examined.

**▼ M4****DATA AND REPORTING****Data**

49. Individual animal data should be provided for all parameters evaluated. Additionally, all data should be summarised in tabular form showing for each test group the number of animals at the start of the test, the number of animals found dead during the test or killed for humane reasons and the time of any death or humane kill, the number showing signs of toxicity, a description of the signs of toxicity observed, including time of onset, duration, and severity of any toxic effects, the number of animals showing lesions, the type of lesions and the percentage of animals displaying each type of lesion. Summary data tables should provide the means and standard deviations (for continuous test data) of animals showing toxic effects or lesions, in addition to the grading of lesions.
50. Historical control data may be valuable in the interpretation of the results of the study, e.g. in the case when there are indications that the data provided by the concurrent controls are substantially out of line when compared to recent data from control animals from the same test facility/colony. Historical control data, if evaluated, should be submitted from the same laboratory and relate to animals of the same age and strain generated during the five years preceding the study in question.
51. When applicable, numerical results should be evaluated by an appropriate and generally acceptable statistical method. The statistical methods and the data to be analysed should be selected during the design of the study (paragraph 8). Selection should make provision for survival adjustments, if needed.

**Test Report**

52. The test report should include the following information:

*Test chemical:*

- physical nature, purity, and physicochemical properties;
- identification data;
- source of chemical;
- batch number;
- certificate of chemical analysis

*Vehicle (if appropriate):*

- justification for choice of vehicle (if other than water).

*Test animals:*

- species/strain used and justification for choice made;
- number, age, and sex of animals at start of test;

**▼ M4**

- source, housing conditions, diet, etc.;
- individual weights of animals at the start of the test.

*Test conditions:*

- rationale for route of administration and dose selection;
- when applicable, the statistical methods used to analyse the data;
- details of test chemical formulation/diet preparation;
- analytical data on achieved concentration, stability and homogeneity of the preparation;
- route of administration and details of the administration of the test chemical;
- for inhalation studies, whether nose only or whole body;
- actual doses (mg/kg body weight/day), and conversion factor from diet/drinking water test chemical concentration (mg/kg or ppm) to the actual dose, if applicable;
- details of food and water quality.

*Results (summary tabulated data and individual animal data should be presented):*

- survival data;
- body weight/body weight changes;
- food consumption, calculations of food efficiency, if made, and water consumption if applicable;
- toxic response data by sex and dose level, including signs of toxicity;
- nature, incidence (and, if scored, severity), and duration of clinical observations ((whether transitory or permanent);
- ophthalmological examination;
- haematological tests;
- clinical biochemistry tests;
- urinalysis tests;
- outcome of any investigations of neurotoxicity or immunotoxicity;
- terminal body weight;
- organ weights (and their ratios, if applicable);
- necropsy findings;
- a detailed description of all treatment-related histopathological findings;
- absorption data if available;



**▼M4**

*Statistical treatment of results, as appropriate*

*Discussion of results including:*

- Dose: response relationships
- Consideration of any mode of action information
- Discussion of any modelling approaches
- BMD, NOAEL or LOAEL determination
- Historical control data
- Relevance for humans

*Conclusions*

*LITERATURE:*

- (1) OECD (1995). Report of the Consultation Meeting on Sub-chronic and Chronic Toxicity/Carcinogenicity Testing (Rome, 1995), internal working document, Environment Directorate, OECD, Paris.
- (2) Combes RD, Gaunt I, Balls M (2004). A Scientific and Animal Welfare Assessment of the OECD Health Effects Test Guidelines for the Safety Testing of Chemicals under the European Union REACH System. ATLA 32: 163-208.
- (3) Barlow SM, Greig JB, Bridges JW *et al.* (2002). Hazard identification by methods of animal-based toxicology. Food. Chem. Toxicol. 40, 145-191.
- (4) Chhabra RS, Bucher JR, Wolfe M, Portier C (2003). Toxicity characterization of environmental chemicals by the US National Toxicology Programme: an overview. Int. J. Hyg. Environ. Health 206: 437-445.
- (5) Chapter B.27 of this Annex, Sub-chronic Oral Toxicity Test Repeated Dose 90-day Oral Toxicity Study in Non-Rodents.
- (6) OECD (2012). Guidance Document on the Design and Conduct of Chronic Toxicity and Carcinogenicity Studies, Supporting Test Guidelines 451, 452 and 453 — Second edition. Series on Testing and Assessment No 116, available on the OECD public website for Test Guideline at [www.oecd.org/env/testguidelines](http://www.oecd.org/env/testguidelines).
- (7) OECD (2009). Guidance Document on Acute Inhalation Toxicity Testing, Series on Testing and Assessment No 39, ENV/JM/MONO(2009)28, OECD, Paris.
- (8) Chapter B.8 of this Annex, Subacute Inhalation Toxicity: 28-Day Study.
- (9) Chapter B.29 of this Annex, Subchronic Inhalation Toxicity: 90-Day Study.
- (10) Chapter B.9 of this Annex, Repeated Dose (28 Days) Toxicity (Dermal).
- (11) Carmichael NG, Barton HA, Boobis AR *et al.* (2006). Agricultural Chemical Safety Assessment: A Multisector Approach to the Modernization of Human Safety Requirements. Critical Reviews in Toxicology 36: 1-7.
- (12) Barton HA, Pastoor TP, Baetcke T *et al.* (2006). The Acquisition and Application of Absorption, Distribution, Metabolism, and Excretion (ADME) Data in Agricultural Chemical Safety Assessments. Critical Reviews in Toxicology 36: 9-35.
- (13) Doe JE, Boobis AR, Blacker A *et al.* (2006). A Tiered Approach to Systemic Toxicity Testing for Agricultural Chemical Safety Assessment. Critical Reviews in Toxicology 36: 37-68.

▼ **M4**

- (14) Cooper RL, Lamb JS, Barlow SM *et al.* (2006). A Tiered Approach to Life Stages Testing for Agricultural Chemical Safety Assessment. *Critical Reviews in Toxicology* 36: 69-98.
- (15) OECD (2002). Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies, Series on Testing and Assessment No 35 and Series on Pesticides No 14, ENV/JM/MONO(2002)19, OECD, Paris.
- (16) OECD (2000). Guidance Document on the recognition, assessment, and use of clinical signs as humane endpoints for experimental animals used in safety evaluation, No 19, ENV/JM/MONO(2000)7, OECD, Paris.
- (17) Rhomberg LR, Baetcke K, Blancato J, Bus J, Cohen S, Conolly R, Dixit R, Doe J, Ekelman K, Fenner-Crisp P, Harvey P, Hattis D, Jacobs A, Jacobson-Kram D, Lewandowski T, Liteplo R, Pelkonen O, Rice J, Somers D, Turturro A, West W, Olin S (2007). Issues in the Design and Interpretation of Chronic Toxicity and Carcinogenicity Studies in Rodents: Approaches to Dose Selection *Crit Rev. Toxicol.* 37 (9): 729 - 837.
- (18) ILSI (International Life Sciences Institute) (1997). Principles for the Selection of Doses in Chronic Rodent Bioassays. Foran JA (Ed.). ILSI Press, Washington, DC.
- (19) Directive 2010/63/EU of the European parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes (OJ L 276, 20.10.2010, p. 33).
- (20) National Research Council, 1985. Guide for the care and use of laboratory animals. NIH Publication No 86-23. Washington D.C., US. Dept. of Health and Human Services.
- (21) GV-SOLAS (Society for Laboratory Animal Science, Gesellschaft für Versuchstierkunde, 1988). Publication on the Planning and Structure of Animal Facilities for Institutes Performing Animal Experiments. ISBN 3-906255-04-2.
- (22) GV-SOLAS (Society for Laboratory Animal Science, Gesellschaft für Versuchstierkunde, 2006). Microbiological monitoring of laboratory animals in various housing systems.
- (23) Diehl K-H, Hull R, Morton D, Pfister R, Rabemampianina Y, Smith D, Vidal J-M, van de Vorstenbosch C. 2001. A good practice guide to the administration of substances and removal of blood, including routes and volumes. *Journal of Applied Toxicology* 21:15-23.
- (24) IPCS (1986). Principles and Methods for the Assessment of Neurotoxicity Associated with Exposure to Chemicals. Environmental Health Criteria Document No 60.
- (25) Tupper DE, Wallace RB (1980). Utility of the Neurologic Examination in Rats. *Acta Neurobiol. Exp.* 40: 999-1003.
- (26) Gad SC (1982). A Neuromuscular Screen for Use in Industrial Toxicology. *J. Toxicol. Environ. Health* 9: 691-704.
- (27) Moser VC, McDaniel KM, Phillips PM (1991). Rat Strain and Stock Comparisons Using a Functional Observational Battery: Baseline Values and Effects of Amitraz. *Toxicol. Appl. Pharmacol.* 108: 267-283.
- (28) Meyer OA, Tilson HA, Byrd WC, Riley MT (1979). A Method for the Routine Assessment of Fore- and Hind-limb Grip Strength of Rats and Mice. *Neurobehav. Toxicol.* 1: 233-236.
- (29) Crofton KM, Howard JL, Moser VC, Gill MW, Reiter LW, Tilson HA, MacPhail RC (1991). Interlaboratory Comparison of Motor Activity Experiments: Implication for Neurotoxicological Assessments. *Neurotoxicol. Teratol.* 13: 599-609.

**▼M4**

- (30) Weingand K, Brown G, Hall R *et al.* (1996). Harmonisation of Animal Clinical Pathology Testing in Toxicity and Safety Studies. *Fundam. & Appl. Toxicol.* 29: 198-201.
- (31) EMEA (draft) document 'Non-clinical guideline on drug-induced hepatotoxicity' (Doc. Ref. EMEA/CHMP/SWP/a50115/2006).
- (32) Crissman JW, Goodman DG, Hildebrandt PK *et al.* (2004). Best Practices Guideline: Toxicological Histopathology. *Toxicologic Pathology* 32: 126-131.

▼ M4

*Appendix 1*

DEFINITION

**Test chemical:** Any substance or mixture tested using this Test Method.

**▼B****B.31. PRENATAL DEVELOPMENTAL TOXICITY STUDY****1. METHOD**

This method is a replicate of OECD TG 414 (2001).

**1.1. INTRODUCTION**

This method for developmental toxicity testing is designed to provide general information concerning the effects of prenatal exposure on the pregnant test animal and on the developing organism in utero; this may include assessment of maternal effects as well as death, structural abnormalities, or altered growth in the foetus. Functional deficits, although an important part of development, are not an integral part of this test method. They may be tested for in a separate study or as an adjunct to this study using the test method for developmental neurotoxicity. For information on testing for functional deficiencies and other postnatal effects the test method for the two-generation reproductive toxicity study and the developmental neurotoxicity study should be consulted as appropriate.

This test method may require specific adaptation in individual cases on the basis of specific knowledge on e.g. physicochemical or toxicological properties of the test substance. Such adaptation is acceptable, when convincing scientific evidence suggests that the adaptation will lead to a more informative test. In such a case, this scientific evidence should be carefully documented in the study report.

**1.2. DEFINITIONS**

**Developmental toxicology:** the study of adverse effects on the developing organism that may result from exposure prior to conception, during prenatal development, or postnatally to the time of sexual maturation. The major manifestations of developmental toxicity include 1) death of the organism, 2) structural abnormality, 3) altered growth, and 4) functional deficiency. Developmental toxicology was formerly often referred to as teratology.

**Adverse effect:** any treatment-related alteration from baseline that diminishes an organism's ability to survive, reproduce or adapt to the environment. Concerning developmental toxicology, taken in its widest sense it includes any effect which interferes with normal development of the conceptus, both before and after birth.

**Altered growth:** an alteration in offspring organ or body weight or size.

**Alterations (anomalies):** structural alterations in development that include both malformations and variations (28).

**Malformation/Major abnormality:** structural change considered detrimental to the animal (may also be lethal) and is usually rare.

**▼ B**

**Variation/Minor abnormality:** structural change considered to have little or no detrimental effect on the animal; may be transient and may occur relatively frequently in the control population.

**Conceptus:** the sum of derivatives of a fertilised ovum at any stage of development from fertilisation until birth including the extra-embryonic membranes as well as the embryo or foetus.

**Implantation (nidation):** attachment of the blastocyst to the epithelial lining of the uterus, including its penetration through the uterine epithelium, and its embedding in the endometrium.

**Embryo:** the early or developing stage of any organism, especially the developing product of fertilisation of an egg after the long axis appears and until all major structures are present.

**Embryotoxicity:** detrimental to the normal structure, development, growth, and/or viability of an embryo.

**Foetus:** the unborn offspring in the post-embryonic period.

**Foetotoxicity:** detrimental to the normal structure, development, growth, and/or viability of a foetus.

**Abortion:** the premature expulsion from the uterus of the products of conception: of the embryo or of a nonviable foetus.

**Resorption:** a conceptus which, having implanted in the uterus, subsequently died and is being, or has been resorbed.

**Early resorption:** evidence of implantation without recognisable embryo/foetus

**Late resorption:** dead embryo or foetus with external degenerative changes

**NOAEL:** abbreviation for no-observed-adverse-effect level and is the highest dose or exposure level where no adverse treatment-related findings are observed.

### 1.3. REFERENCE SUBSTANCE

None.

### 1.4. PRINCIPLE OF THE TEST METHOD

Normally, the test substance is administered to pregnant animals at least from implantation to one day prior to the day of scheduled kill, which should be as close as possible to the normal day of delivery without risking loss of data resulting from early delivery. The test method is not intended to examine solely the period of organogenesis, (e.g. days 5-15 in the rodent, and days 6-18 in the rabbit) but also effects from preimplantation, when appropriate, through the entire period of gestation to the day before caesarean section. Shortly before caesarean section, the females are killed, the uterine contents are examined, and the foetuses are evaluated for externally visible anomalies and for soft tissue and skeletal changes.

**▼B**

## 1.5. DESCRIPTION OF THE TEST METHOD

1.5.1. **Selection of animal species**

It is recommended that testing be performed in the most relevant species, and that laboratory species and strains which are commonly used in prenatal developmental toxicity testing be employed. The preferred rodent species is the rat and the preferred non-rodent species is the rabbit. Justification should be provided if another species is used.

1.5.2. **Housing and feeding conditions**

The temperature in the experimental animal room should be 22 °C ( $\pm 3^\circ$ ) for rodents and 18 °C ( $\pm 3^\circ$ ) for rabbits. Although the relative humidity should be at least 30 % and preferably not exceed 70 % other than during room cleaning, the aim should be 50-60 %. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water.

Mating procedures should be carried out in cages suitable for the purpose. While individual housing of mated animals is preferred, group housing in small numbers is also acceptable.

1.5.3. **Preparation of the animals**

Healthy animals, which have been acclimated to laboratory conditions for at least five days and have not been subjected to previous experimental procedures, should be used. The test animals should be characterised as to species, strain, source, sex, weight and/or age. The animals of all test groups should, as nearly as practicable, be of uniform weight and age. Young adult nulliparous female animals should be used at each dose level. The females should be mated with males of the same species and strain, and the mating of siblings should be avoided. For rodents day 0 of gestation is the day on which a vaginal plug and/or sperm are observed; for rabbits day 0 is usually the day of coitus or of artificial insemination, if this technique is used. Mated females should be assigned in an unbiased manner to the control and treatment groups. Cages should be arranged in such a way that possible effects due to cage placement are minimised. Each animal should be assigned a unique identification number. Mated females should be assigned in an unbiased manner to the control and treatment groups, and if the females are mated in batches, the animals in each batch should be evenly distributed across the groups. Similarly, females inseminated by the same male should be evenly distributed across the groups.

## 1.6. PROCEDURE

1.6.1. **Number and sex of animals**

Each test and control group should contain a sufficient number of females to result in approximately 20 female animals with implantation sites at necropsy. Groups with fewer than 16 animals with implantation sites may be inappropriate. Maternal mortality does not necessarily invalidate the study providing it does not exceed approximately 10 %.

**▼B****1.6.2. Preparation of doses**

If a vehicle or other additive is used to facilitate dosing, consideration should be given to the following characteristics: effects on the absorption, distribution, metabolism, and retention or excretion of the test substance; effects on the chemical properties of the test substance which may alter its toxic characteristics; and effects on the food or water consumption or the nutritional status of the animals. The vehicle should neither be developmentally toxic nor have effects on reproduction.

**1.6.3. Dosage**

Normally, the test substance should be administered daily from implantation (e.g. day 5 post mating) to the day prior to scheduled caesarean section. If preliminary studies, when available, do not indicate a high potential for preimplantation loss, treatment may be extended to include the entire period of gestation, from mating to the day prior to scheduled kill. It is well known that inappropriate handling or stress during pregnancy can result in prenatal loss. To guard against prenatal loss from factors which are not treatment related, unnecessary handling of pregnant animals as well as stress from outside factors such as noise should be avoided.

At least three dose levels and a concurrent control should be used. Healthy animals should be assigned in an unbiased manner to the control and treatment groups. The dose levels should be spaced to produce a gradation of toxic effects. Unless limited by the physical/chemical nature or biological properties of the test substance, the highest dose should be chosen with the aim to induce some developmental and/or maternal toxicity (clinical signs or a decrease in body weight) but not death or severe suffering. At least one intermediate dose level should produce minimal observable toxic effects. The lowest dose level should not produce any evidence of either maternal or developmental toxicity. A descending sequence of dose levels should be selected with a view to demonstrating any dosage-related response and no-observed-adverse-effect level (NOAEL). Two- to four-fold intervals are frequently optimal for setting the descending dose levels, and the addition of a fourth test group is often preferable to using very large intervals (e.g. more than a factor of 10) between dosages. Although establishment of a maternal NOAEL is the goal, studies which do not establish such a level may also be acceptable (1).

Dose levels should be selected taking into account any existing toxicity data as well as additional information on metabolism and toxicokinetics of the test substance or related materials. This information will also assist in demonstrating the adequacy of the dosing regimen.

A concurrent control group should be used. This group should be a sham-treated control group or a vehicle-control group if a vehicle is used in administering the test substance. All groups should be administered the same volume of either test substance or vehicle. Animals in the control group(s) should be handled in an identical manner to test group animals. Vehicle control groups should receive the vehicle in the highest amount used (as in the lowest treatment group).



**▼B****1.6.4. Limit test**

If a test at one dose level of at least 1 000 mg/kg body weight/day by oral administration, using the procedures described for this study, produces no observable toxicity in either pregnant animals or their progeny and if an effect would not be expected based upon existing data (e.g. from structurally and/or metabolically related compounds), then a full study using three dose levels may not be considered necessary. Expected human exposure may indicate the need for a higher oral dose level to be used in the limit test. For other types of administration, such as inhalation or dermal application, the physico-chemical properties of the test substance often may indicate and limit the maximum attainable level of exposure (for example, dermal application should not cause severe local toxicity).

**1.6.5. Administration of doses**

The test substance or vehicle is usually administered orally by intubation. If another route of administration is used, the tester should provide justification and reasoning for its selection, and appropriate modifications may be necessary (2)(3)(4). The test substance should be administered at approximately the same time each day.

The dose to individual animals should normally be based on the most recent individual body weight determination. However, caution should be exercised when adjusting the dose during the last trimester of pregnancy. Existing data should be used for dose selection to prevent excess maternal toxicity. However, if excess toxicity is noted in the treated dams, those animals should be humanely killed. If several pregnant animals show signs of excess toxicity, consideration should be given to terminating that dose group. When the substance is administered by gavage, this should preferably be given as a single dose to the animals using a stomach tube or a suitable intubation canula. The maximum volume of liquid that can be administered at one time depends on the size of the test animal. The volume should not exceed 1 ml/100 g body weight, except in the case of aqueous solutions where 2 ml/100 g body weight may be used. When corn oil is used as a vehicle, the volume should not exceed 0.4 ml/100 g body weight. Variability in test volume should be minimised by adjusting the concentrations to ensure a constant volume across all dose levels.

**1.6.6. Observations of the dams**

Clinical observations should be made and recorded at least once a day, preferably at the same time(s) each day taking into consideration the peak period of anticipated effects after dosing. The condition of the animals should be recorded including mortality, moribundity, pertinent behavioural changes, and all signs of overt toxicity.

**1.6.7. Body weight and food consumption**

Animals should be weighed on day 0 of gestation or no later than day 3 of gestation if time-mated animals are supplied by an outside breeder, on the first day of dosing, at least every three days during the dosing period and on the day of scheduled kill.

**▼B**

Food consumption should be recorded at three-day intervals and should coincide with days of body weight determination.

**1.6.8. Post-mortem examination**

Females should be killed one day prior to the expected day of delivery. Females showing signs of abortion or premature delivery prior to scheduled kill should be killed and subjected to a thorough macroscopic examination.

At the time of termination or death during the study, the dam should be examined macroscopically for any structural abnormalities or pathological changes. Evaluation of the dams during caesarean section and subsequent foetal analyses should be conducted preferably without knowledge of treatment group in order to minimise bias.

**1.6.9. Examination of uterine contents**

Immediately after termination or as soon as possible after death, the uteri should be removed and the pregnancy status of the animals ascertained. Uteri that appear non gravid should be further examined (e.g. by ammonium sulphide staining for rodents and Salewski staining or a suitable alternative method for rabbits) to confirm the non-pregnant status (5).

Gravid uteri including the cervix should be weighed. Gravid uterine weights should not be obtained from animals found dead during the study.

The number of corpora lutea should be determined for pregnant animals.

The uterine contents should be examined for numbers of embryonic or foetal deaths and viable foetuses. The degree of resorption should be described in order to estimate the relative time of death of the conceptus (see Section 1.2).

**1.6.10. Examination of foetuses**

The sex and body weight of each foetus should be determined.

Each foetus should be examined for external alterations (6).

Foetuses should be examined for skeletal and soft tissue alterations (e.g. variations and malformations or anomalies) (7)(8)(9)(10)(11)(12)(13)(14)(15)(16)(17)(18)(19)(20)(21)(22)(23)(24). Categorisation of foetal alterations is preferable but not required. When categorisation is done, the criteria for defining each category should be clearly stated. Particular attention should be paid to the reproductive tract which should be examined for signs of altered development.

For rodents, approximately one-half of each litter should be prepared and examined for skeletal alterations. The remainder should be prepared and examined for soft tissue alterations, using accepted or appropriate serial sectioning methods or careful gross dissection techniques.

**▼B**

For non-rodents, e.g. rabbits, all foetuses should be examined for both soft tissue and skeletal alterations. The bodies of these foetuses are evaluated by careful dissection for soft tissue alterations, which may include procedures to further evaluate internal cardiac structure (25). The heads of one-half of the foetuses examined in this manner should be removed and processed for evaluation of soft tissue alterations (including eyes, brain, nasal passages and tongue), using standard serial sectioning methods (26) or an equally sensitive method. The bodies of these foetuses and the remaining intact foetuses should be processed and examined for skeletal alterations, utilising the same methods as described for rodents.

**2. DATA****2.1. TREATMENT OF RESULTS**

Data shall be reported individually for the dams as well as for their offspring and summarised in tabular form, showing for each test group and each generation the number of animals at the start of the test, the number of animals found dead during the test or killed for humane reasons, the time of any death or humane kill, the number of pregnant females, the number of animals showing signs of toxicity, a description of the signs of toxicity observed, including time of onset, duration, and severity of any toxic effects, the types of embryo/foetal observations, and all relevant litter data.

Numerical results should be evaluated by an appropriate statistical method using the litter as the unit for data analysis. A generally accepted statistical method should be used; the statistical methods should be selected as part of the design of the study and should be justified. Data from animals that do not survive to the scheduled kill should also be reported. These data may be included in group means where relevant. Relevance of the data obtained from such animals, and therefore inclusion or exclusion from any group mean(s), should be justified and judged on an individual basis.

**2.2. EVALUATION OF RESULTS**

The findings of the Prenatal Developmental Toxicity Study should be evaluated in terms of the observed effects. The evaluation will include the following information:

- maternal and embryo/foetal test results, including the evaluation of the relationship, or lack thereof, between the exposure of the animals to the test substance and the incidence and severity of all findings,
  
- criteria used for categorising foetal external, soft tissue, and skeletal alterations if categorisation has been done,

**▼ B**

- when appropriate, historical control data to enhance interpretation of study results,
- the numbers used in calculating all percentages or indices,
- adequate statistical analysis of the study findings, when appropriate, which should include sufficient information on the method of analysis, so that an independent reviewer/statistician can re-evaluate and reconstruct the analysis.

In any study which demonstrates the absence of any toxic effects, further investigations to establish absorption and bioavailability of the test substance should be considered.

### 2.3. INTERPRETATION OF RESULTS

A prenatal developmental toxicity study will provide information on the effects of repeated exposure to a substance during pregnancy on the dams and on the intrauterine development of their progeny. The results of the study should be interpreted in conjunction with the findings from subchronic, reproduction, toxicokinetic and other studies. Since emphasis is placed both on general toxicity in terms of maternal toxicity and on developmental toxicity endpoints, the results of the study will allow to a certain extent for the discrimination between developmental effects occurring in the absence of general toxicity and those which are only induced at levels that are also toxic to the maternal animal (27).

## 3. REPORTING

### 3.1. TEST REPORT

The test report must include the following specific information:

Test substance:

- physical nature and, where relevant, physiochemical properties,
- identification including CAS number if known/established,
- purity.

Vehicle (if appropriate):

- justification for choice of vehicle, if other than water.

Test animals:

- species and strain used,
- number and age of animals,
- source, housing conditions, diet, etc.,
- individual weights of animals at the start of the test.

Test conditions:

- rationale for dose level selection,
- details of test substance formulation/diet preparation, achieved concentration, stability and homogeneity of the preparation,

**▼ B**

- details of the administration of the test substance,
- conversion from diet/drinking water test substance concentration (ppm) to the actual dose (mg/kg body weight/day), if applicable,
- environmental conditions,
- details of food and water quality.

## Results:

Maternal toxic response data by dose, including but not limited to:

- the number of animals at the start of the test, the number of animals surviving, the number pregnant, and the number aborting, number of animals delivering early,
- day of death during the study or whether animals survived to termination,
- data from animals that do not survive to the scheduled kill should be reported but not included in the inter-group statistical comparisons,
- day of observation of each abnormal clinical sign and its subsequent course,
- body weight, body weight change and gravid uterine weight, including, optionally, body weight change corrected for gravid uterine weight,
- food consumption and, if measured, water consumption,
- necropsy findings, including uterine weight,
- NOAEL values for maternal and developmental effects should be reported.

Developmental endpoints by dose for litters with implants, including:

- number of corpora lutea,
- number of implantations, number and percent of live and dead fetuses and resorptions,
- number and percent of pre- and post-implantation losses.

Developmental endpoints by dose for litters with live fetuses, including:

- number and percent of live offspring,
- sex ratio,
- foetal body weight, preferably by sex and with sexes combined,
- external, soft tissue, and skeletal malformations and other relevant alterations,
- criteria for categorisation if appropriate,

**▼B**

- total number and percent of fetuses and litters with any external, soft tissue, or skeletal alteration, as well as the types and incidences of individual anomalies and other relevant alterations.

Discussion of results.

Conclusions.

#### 4. REFERENCES

- (1) Kavlock R.J. et al., (1996) A Simulation Study of the Influence of Study Design on the Estimation of Benchmark Doses for Developmental Toxicity. *Risk Analysis* 16; p. 399-410.
- (2) Kimmel, C.A. and Francis, E.Z., (1990) Proceedings of the Workshop on the Acceptability and Interpretation of Dermal Developmental Toxicity Studies. *Fundamental and Applied Toxicology* 14; p. 386-398.
- (3) Wong, B.A., et al., (1997) Developing Specialised Inhalation Exposure Systems to Address Toxicological Problems. *CIIT Activities* 17; p. 1-8.
- (4) US Environmental Protection Agency, (1985) Subpart E-Specific Organ/Tissue Toxicity, 40 CFR 798.4350: Inhalation Developmental Toxicity Study.
- (5) Salewski, E., (1964) Faerbermethode zum Makroskopischen Nachweis von Implantations Stellen am Uterusder Ratte. *Naunyn-Schmeidebergs Archiv fur Pharmakologie und Experimentelle Pathologie*, 247, 367.
- (6) Edwards, J.A., (1968) The external Development of the Rabbit and Rat Embryo. In *Advances in Teratology*. D.H.M. Woolam (ed.) Vol. 3. Academic Press, NY.
- (7) Inouye, M. (1976) Differential Staining of Cartilage and Bone in Fetal Mouse Skeleton by Alcian Blue and Alizarin Red S. *Congenital Anomalies* 16, p. 171-173.
- (8) Igarashi, E. et al., (1992) Frequency Of Spontaneous Axial Skeletal Variations Detected by the Double Staining Technique for Ossified and Cartilaginous Skeleton in Rat Foetuses. *Congenital Anomalies* 32, p. 381-391.
- (9) Kimmel, C.A. et al. (1993) Skeletal Development Following Heat Exposure in the Rat. *Teratology*, 47 p.229-242.
- (10) Marr, M.C. et al. (1988) Comparison of Single and Double Staining for Evaluation of Skeletal Development: The Effects of Ethylene Glycol (EG) in CD Rats. *Teratology* 37; 476.
- (11) Barrow, M.V. and Taylor, W.J. (1969) A Rapid Method for Detecting Malformations in Rat Foetuses. *Journal of Morphology*, 127, p. 291-306.
- (12) Fritz, H. (1974) Prenatal Ossification in Rabbits ss Indicative of Foetal Maturity. *Teratology*, 11 p. 313-320.
- (13) Gibson, J.P. et al. (1966) Use of the Rabbit in Teratogenicity Studies. *Toxicology and Applied Pharmacology*, 9, p. 398-408.
- (14) Kimmel, C.A. and Wilson, J.G. (1973) Skeletal Deviation in Rats: Malformations or Variations? *Teratology*, 8, p. 309-316.
- (15) Marr, M.C. et al. (1992) Developmental Stages of the CD (Sprague-Dawley) Rat Skeleton after Maternal Exposure to Ethylene Glycol. *Teratology*, 46, p. 169-181.

**▼B**

- (16) Monie, I.W. et al. (1965) Dissection Procedures for Rat Foetuses Permitting Alizarin Red Staining of Skeleton and Histological Study of Viscera. *Supplement to Teratology Workshop Manual*, p. 163-173.
- (17) Spark, C. and Dawson, A.B. (1928) The Order and Time of appearance of Centers of Ossification in the Fore and Hind Limbs of the Albino Rat, with Special Reference to the Possible Influence of the Sex Factor. *American Journal of Anatomy*, 41, p. 411-445.
- (18) Staples, R.E. and Schnell, V.L. (1964) Refinements in Rapid Clearing Technique in the KOH-Alizarin Red S Method for Fetal Bone. *Stain Technology*, 39, p. 61-63.
- (19) Strong, R.M. (1928) The Order Time and Rate of Ossification of the Albino Rat (*Mus Norvegicus Albinus*) Skeleton. *American Journal of Anatomy*, 36, p. 313-355.
- (20) Stuckhardt, J.L. and Poppe, S.M. (1984) Fresh Visceral Examination of Rat and Rabbit Foetuses Used in Teratogenicity Testing. *Teratogenesis, Carcinogenesis, and Mutagenesis*, 4, p. 181-188.
- (21) Walker, D.G. and Wirtschafter, Z.T. (1957) *The Genesis of the Rat Skeleton*. Thomas, Springfield, IL.
- (22) Wilson, J.G. (1965) Embryological Considerations in Teratology. In *Teratology: Principles and Techniques*, Wilson J.G. and Warkany J. (eds). University of Chicago, Chicago, IL, p. 251-277.
- (23) Wilson, J.G. and Fraser, F.C. (eds). (1977) *Handbook of Teratology*, Vol. 4. Plenum, NY.
- (24) Varnagy, L. (1980) Use of Recent Fetal Bone Staining Techniques in the Evaluation of Pesticide Teratogenicity. *Acta Vet. Acad. Sci. Hung*, 28, p. 233-239.
- (25) Staples, R.E. (1974) Detection of visceral Alterations in Mammalian Foetuses. *Teratology*, 9, p. 37-38.
- (26) Van Julsingha, E.B. and C.G. Bennett (1977) A Dissecting Procedure for the Detection of Anomalies in the Rabbit Foetal Head. In: *Methods in Prenatal Toxicology* Neubert, D., Merker, H.J. and Kwasigroch, T.E. (eds.). University of Chicago, Chicago, IL, p. 126-144.
- (27) US Environmental Protection Agency (1991) Guidelines for Developmental Toxicity Risk Assessment. *Federal Register*, 56, p. 63798-63826.
- (28) Wise, D.L. et al. (1997) Terminology of Developmental Abnormalities in Common Laboratory Mammals (Version 1) *Teratology*, 55, p. 249-292.

**▼ M4****B.32. CARCINOGENICITY STUDIES**

## INTRODUCTION

1. This Test Method is equivalent to OECD Test Guideline (TG) 451 (2009). The original TG 451 on Carcinogenicity Studies was adopted in 1981. Development of this revised Test Method B.32 was considered necessary, in order to reflect recent developments in the field of animal welfare and regulatory requirements (2) (3) (4) (5) (6). The updating of this Test Method B.32 has been carried out in parallel with revisions of Chapter B.30 of this Annex, Chronic Toxicity Studies, and Chapter B.33, of this Annex, Combined Chronic Toxicity\Carcinogenicity Studies, and with the objective of obtaining additional information from the animals used in the study and providing further detail on dose selection. This Test Method B.32 is designed to be used in the testing of a broad range of chemicals, including pesticides and industrial chemicals. It should be noted however that some details and requirements may differ for pharmaceuticals (see International Conference on Harmonisation (ICH) Guidance S1B on Testing for Carcinogenicity of Pharmaceuticals).
2. The majority of carcinogenicity studies are carried out in rodent species, and this Test Method is intended therefore to apply primarily to studies carried out in these species. Should such studies be required in non-rodent species, the principles and procedures outlined in this Test Method together with those outlined in Chapter B.27 of this Annex, Repeated Dose 90-day Oral Toxicity Study in Non-Rodents (6), should be applied, with appropriate modifications. Further guidance is available in the OECD Guidance Document No 116 on the Design and Conduct of Chronic Toxicity and Carcinogenicity Studies (7).
3. The three main routes of administration used in carcinogenicity studies are oral, dermal and inhalation. The choice of the route of administration depends on the physical and chemical characteristics of the test chemical and the predominant route of exposure of humans. Additional information on choice of route of exposure is provided in Guidance Document No 116 (7).
4. This Test Method focuses on exposure via the oral route, the route most commonly used in carcinogenicity studies. While carcinogenicity studies involving exposure via the dermal or inhalation routes may also be necessary for human health risk assessment and/or may be required under certain regulatory regimes, both routes of exposure involve considerable technical complexity. Such studies will need to be designed on a case-by-case basis, although the Test Method outlined here for the assessment and evaluation of carcinogenicity by oral administration could form the basis of a protocol for inhalation and/or dermal studies, with respect to recommendations for treatment periods, clinical and pathology parameters, etc. OECD Guidance is available on the administration of test chemicals by the dermal (7), and inhalation routes (7) (8). Chapter B.8 of this Annex (9) and Chapter B.29 of this Annex (10), together with the OECD Guidance Document on acute inhalation testing (8), should be specifically consulted in the design of longer term studies involving exposure via the inhalation route. Chapter B.9 of this Annex (11) should be consulted in the case of testing carried out by the dermal route.



**▼ M4**

5. The carcinogenicity study provides information on the possible health hazards likely to arise from repeated exposure for a period lasting up to the entire lifespan of the species used. The study will provide information on the toxic effects of the test chemical including potential carcinogenicity, and may indicate target organs and the possibility of accumulation. It can provide an estimate of the no-observed-adverse effect level for toxic effects and, in the case of non-genotoxic carcinogens, for tumour responses, which can be used for establishing safety criteria for human exposure. The need for careful clinical observations of the animals, so as to obtain as much information as possible, is also stressed.
  
6. The objectives of carcinogenicity studies covered by this Test Method include:
  - The identification of the carcinogenic properties of a test chemical, resulting in an increased incidence of neoplasms, increased proportion of malignant neoplasms or a reduction in the time to appearance of neoplasms, compared with concurrent control groups;
  
  - The identification of target organ(s) of carcinogenicity;
  
  - The identification of the time to appearance of neoplasms;
  
  - Characterisation of the tumour dose-response relationship;
  
  - Identification of a no-observed-adverse-effect level (NOAEL) or point of departure for establishment of a Benchmark Dose (BMD);
  
  - Extrapolation of carcinogenic effects to low dose human exposure levels;
  
  - Provision of data to test hypotheses regarding mode of action (2) (7) (12) (13) (14) (15).

## INITIAL CONSIDERATIONS

7. In the assessment and evaluation of the potential carcinogenicity of a test chemical, all available information on the test chemical should be considered by the testing laboratory prior to conducting the study, in order to focus the design of the study to more efficiently test for carcinogenic potential and to minimise animal usage. Information on, and consideration of, the mode of action of a suspected carcinogen (2) (7) (12) (13) (14) (15) is particularly important, since the optimal design may differ depending on whether the test chemical is a known or suspected genotoxic carcinogen. Further guidance on mode of action considerations can be found in Guidance Document No 116 (7).
  
8. Information that will assist in the study design includes the identity, chemical structure, and physico-chemical properties of the test chemical; results of any *in vitro* or *in vivo* toxicity tests including genotoxicity tests; anticipated use(s) and potential for human exposure; available (Q)SAR data, mutagenicity/genotoxicity, carcinogenicity and other toxicological data on structurally-related chemicals; available toxicokinetic data (single dose and also repeat dose kinetics where available) and data derived from other repeated exposure studies. Assessment of carcinogenicity should be

**▼ M4**

carried out after initial information on toxicity has been obtained from repeated dose 28-day and/or 90-day toxicity tests. Short-term cancer initiation-promotion tests could also provide useful information. A phased testing approach to carcinogenicity testing should be considered as part of the overall assessment of the potential adverse health effects of a particular test chemical (16) (17) (18) (19).

9. The statistical methods most appropriate for the analysis of results, given the experimental design and objectives, should be established before commencing the study. Issues to consider include whether the statistics should include adjustment for survival, analysis of cumulative tumour risks relative to survival duration, analysis of the time to tumour and analysis in the event of premature termination of one or more groups. Guidance on the appropriate statistical analyses and key references to internationally accepted statistical methods are given in Guidance Document No 116 (7), and also in Guidance Document No 35 on the analysis and evaluation of chronic toxicity and carcinogenicity studies (20).
  
10. In conducting a carcinogenicity study, the guiding principles and considerations outlined in the OECD Guidance Document No 19 on the recognition, assessment, and use of clinical signs as humane endpoints for experimental animals used in safety evaluation (21), in particular paragraph 62 thereof, should always be followed. This paragraph states that *'In studies involving repeated dosing, when an animal shows clinical signs that are progressive, leading to further deterioration in condition, an informed decision as to whether or not to humanely kill the animal should be made. The decision should include consideration as to the value of the information to be gained from the continued maintenance of that animal on study relative to its overall condition. If a decision is made to leave the animal on test, the frequency of observations should be increased, as needed. It may also be possible, without adversely affecting the purpose of the test, to temporarily stop dosing if it will relieve the pain or distress, or reduce the test dose.'*
  
11. Detailed guidance on and discussion of the principles of dose selection for chronic toxicity and carcinogenicity studies can be found in Guidance Document No 116 (7) as well as two International Life Sciences Institute publications (22) (23). The core dose selection strategy is dependent on the primary objective or objectives of the study (paragraph 6). In selecting appropriate dose levels, a balance should be achieved between hazard screening on the one hand and characterisation of low-dose responses and their relevance on the other. This is particularly relevant in the situation where a combined chronic toxicity and carcinogenicity study (Chapter B.33 of this Annex) is to be carried out (paragraph 12).
  
12. Consideration should be given to carrying out a combined chronic toxicity and carcinogenicity study (Chapter B.33 of this Annex), rather than separate execution of a chronic toxicity study (Chapter B.30 of this Annex) and carcinogenicity study (this Test Method B.32). The combined test provides greater efficiency in terms of time and cost compared to conducting two separate studies, without compromising the quality of the data in either the chronic phase or the carcinogenicity phase. Careful consideration should however be given to the principles of dose selection (paragraphs 11 and 22-25) when undertaking a combined chronic toxicity and carcinogenicity study (Chapter B.33 of this Annex), and it is also recognised that separate studies may be required under certain regulatory frameworks.

**▼ M4**

13. Definitions used in the context of this Test Method can be found at the end of this chapter and in the Guidance Document No 116 (7).

**PRINCIPLE OF THE TEST**

14. The test chemical is administered daily in graduated doses to several groups of test animals for the majority of their life span, normally by the oral route. Testing by the inhalation or dermal route may also be appropriate. The animals are observed closely for signs of toxicity and for the development of neoplastic lesions. Animals which die or are killed during the test are necropsied and, at the conclusion of the test, surviving animals are killed and necropsied.

**DESCRIPTION OF METHOD****Selection of animal species**

15. This Test Method primarily covers assessment and evaluation of carcinogenicity in rodents (paragraph 2). The use of non-rodent species may be considered when available data suggest that they are more relevant for the prediction of health effects in humans. The choice of species should be justified. The preferred rodent species is the rat, although other rodent species, e.g. the mouse, may be used. Although the use of the mouse in carcinogenicity testing may have limited utility (24) (25) (26), under some current regulatory programmes carcinogenicity testing in the mouse is still required unless it is determined that such a study is not scientifically necessary. Rats and mice have been preferred experimental models because of their relatively short life span, their widespread use in pharmacological and toxicological studies, their susceptibility to tumour induction, and the availability of sufficiently characterised strains. As a consequence of these characteristics, a large amount of information is available on their physiology and pathology. Additional information on choice of species and strain is provided in Guidance Document No 116 (7).
  
16. Young healthy adult animals of commonly used laboratory strains should be employed. The carcinogenicity study should preferably be carried out in animals from the same strain and source as those used in preliminary toxicity study(ies) of shorter duration although, if animals from this strain and source are known to present problems in achieving the normally accepted criteria of survival for long-term studies [see Guidance Document No 116 (7)], consideration should be given to using a strain of animal that has an acceptable survival rate for the long-term study. The females should be nulliparous and non-pregnant.

**Housing and feeding**

17. Animals may be housed individually, or be caged in small groups of the same sex; individual housing should be considered only if scientifically justified (27) (28) (29). Cages should be arranged in such a way that possible effects due to cage placement are minimised. The temperature in the experimental animal room should be 22 °C ( $\pm$  3 °C). Although the relative humidity should be at least 30 % and preferably not exceed 70 % other than during room cleaning, the aim should be 50-60 %. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water. The diet should meet all the nutritional requirements of the species tested and the content of dietary contaminants, including but not limited to pesticide residues, persistent organic pollutants, phytoestrogens,

**▼M4**

heavy metals and mycotoxins, that might influence the outcome of the test, should be as low as possible. Analytical information on the nutrient and dietary contaminant levels should be generated periodically, at least at the beginning of the study and when there is a change in the batch used, and should be included in the final report. Analytical information on the drinking water used in the study should similarly be provided. The choice of diet may be influenced by the need to ensure a suitable admixture of a test chemical and to meet the nutritional requirements of the animals when the test chemical is administered by the dietary route.

**Preparation of animals**

18. Healthy animals, which have been acclimated to laboratory conditions for at least 7 days and have not been subjected to previous experimental procedures, should be used. In the case of rodents, dosing of the animals should begin as soon as possible after weaning and acclimatisation and preferably before the animals are 8 weeks old. The test animals should be characterised as to species, strain, source, sex, weight and age. At the commencement of the study, the weight variation for each sex of animal used should be minimal and not exceed  $\pm 20\%$  of the mean weight of all the animals within the study, separately for each sex. Animals should be randomly assigned to the control and treatment groups. After randomisation, there should be no significant differences in mean body weights between groups within each sex. If there are statistically significant differences, then the randomisation step should be repeated, if possible. Each animal should be assigned a unique identification number, and permanently marked with this number by tattooing, microchip implant, or other suitable method.

## PROCEDURE

**Number and sex of animals**

19. Both sexes should be used. A sufficient number of animals should be used so that a thorough biological and statistical evaluation is possible. Each dose group and concurrent control group should therefore contain at least 50 animals of each sex. Depending on the aim of the study, it may be possible to increase the statistical power of the key estimates by differentially allocating animals unequally to the various dose groups, with more than 50 animals in the low dose groups; e.g. to estimate the carcinogenic potential at low doses. However it should be recognised that a moderate increase in group size will provide relatively little increase in statistical power of the study. Further information on statistical design of the study and choice of dose levels to maximise statistical power is provided in Guidance Document No 116 (7).

**Provision for interim kills and satellite (sentinel) groups**

20. The study may make provision for interim kills, e.g. at 12 months, to provide information on progression of neoplastic changes and mechanistic information, if scientifically justified. Where such information is already available from previous repeat dose toxicity studies on the test chemical, interim kills may not be scientifically justified. If interim kills are included in the study design, the number of animals in each dose group scheduled for an interim kill will normally be 10 animals per sex, and the total number of animals included in the study design should be increased by the number of animals scheduled to be killed before the completion of the study. An additional group of sentinel animals (typically 5 animals per sex) may be included for monitoring of disease status, if necessary, during the study (30). Further guidance is provided in Guidance Document No 116 (7).

**▼M4****Dose groups and dosage**

21. Guidance on all aspects of dose selection and dose level spacing is provided in Guidance Document No 116 (7). At least three dose levels and a concurrent control should be used. Dose levels will generally be based on the results of shorter-term repeated dose or range finding studies and should take into account any existing toxicological and toxicokinetic data available for the test chemical or related chemicals.
22. Unless limited by the physical-chemical nature or biological effects of the test chemical, the highest dose level should be chosen to identify the principal target organs and toxic effects while avoiding suffering, severe toxicity, morbidity, or death. While taking into account the factors outlined in paragraph 23 below, the highest dose level should normally be chosen to elicit evidence of toxicity, as evidenced by, for example, depression of body weight gain (approximately 10 %). However, dependent on the objectives of the study (see paragraph 6), a top dose lower than the dose providing evidence of toxicity may be chosen, e.g. if a dose elicits an adverse effect of concern that nonetheless has little impact on lifespan or body weight.
23. Dose levels and dose level spacing may be selected to establish a dose-response and, depending on the mode of action of the test chemical, a NOAEL or other intended outcome of the study, e.g. a BMD (see paragraph 25) at the lowest dose level. Factors that should be considered in the placement of lower doses include the expected slope of the dose-response curve, the doses at which important changes may occur in metabolism or mode of toxic action, where a threshold is expected, or where a point of departure for low-dose extrapolation is expected.
24. The dose level spacing selected will depend on the characteristics of the test chemical, and cannot be prescribed in this Test Method, but two to four fold intervals frequently provide good test performance for setting the descending dose levels and addition of a fourth test group is often preferable to using very large intervals (e.g. more than a factor of about 6-10) between dosages. In general, the use of factors greater than 10 should be avoided, and should be justified if used.
25. As discussed further in Guidance Document No 116 (7), points to be considered in dose selection include:
  - Known or suspected nonlinearities or inflection points in the dose-response;
  - Toxicokinetics, and dose ranges where metabolic induction, saturation, or nonlinearity between external and internal doses does or does not occur;
  - Precursor lesions, markers of effect, or indicators of the operation of key underlying biological processes;
  - Key (or suspected) aspects of mode of action, such as doses at which cytotoxicity begins to arise, hormone levels are perturbed, homeostatic mechanisms are overwhelmed, etc.;

**▼ M4**

— Regions of the dose–response curve where particularly robust estimation is needed, e.g. in the range of the anticipated BMD or a suspected threshold;

— Consideration of anticipated human exposure levels.

26. The control group shall be an untreated group or a vehicle-control group if a vehicle is used in administering the test chemical. Except for treatment with the test chemical, animals in the control group should be handled in an identical manner to those in the test groups. If a vehicle is used, the control group shall receive the vehicle in the highest volume used among the dose groups. If a test chemical is administered in the diet, and causes significantly reduced dietary intake due to the reduced palatability of the diet, an additional pair-fed control group may be useful, to serve as a more suitable control.

**Preparation of doses and administration of test chemical**

27. The test chemical is normally administered orally, via the diet or drinking water, or by gavage. Additional information on routes and methods of administration is provided in Guidance Document No 116 (7). The route and method of administration is dependent on the purpose of the study, the physical-chemical properties of the test chemical, its bioavailability and the predominant route and method of exposure of humans. A rationale should be provided for the chosen route and method of administration. In the interest of animal welfare, oral gavage should normally be selected only for those agents, for which this route and method of administration reasonably represent potential human exposure (e.g. pharmaceuticals). For dietary or environmental chemicals including pesticides, administration is typically via the diet or drinking water. However, for some scenarios, e.g. occupational exposure, administration via other routes may be more appropriate.
28. Where necessary, the test chemical is dissolved or suspended in a suitable vehicle. Consideration should be given to the following characteristics of the vehicle and other additives, as appropriate: effects on the absorption, distribution, metabolism, or retention of the test chemical; effects on the chemical properties of the test chemical which may alter its toxic characteristics; and effects on the food or water consumption or the nutritional status of the animals. It is recommended that, wherever possible, the use of an aqueous solution/suspension be considered first, followed by consideration of a solution/emulsion in oil (e.g. corn oil) and then by possible solution in other vehicles. For vehicles other than water, the toxic characteristics of the vehicle should be known. Information should be available on the stability of the test chemical and the homogeneity of dosing solutions or diets (as appropriate) under the conditions of administration (e.g. diet).
29. For chemicals administered via the diet or drinking water it is important to ensure that the quantities of the test chemical involved do not interfere with normal nutrition or water balance. In long-term toxicity studies using dietary administration, the concentration of the test chemical in the feed should not normally exceed an upper limit of 5 % of the total diet, in order to avoid nutritional imbalances. When the test chemical is administered in the diet, either a constant dietary concentration (mg/kg diet or ppm) or a constant dose level in terms of the animal's body weight (mg/kg body weight), calculated on a weekly basis, may be used. The alternative used should be specified.

**▼M4**

30. In the case of oral administration, the animals are dosed with the test chemical daily (seven days per week), normally for a period of 24 months for rodents (see also paragraph 32). Any other dosing regime, e.g. five days per week, needs to be justified. In the case of dermal administration, animals are normally treated with the test chemical for at least 6 hours per day, 7 days per week, as specified in Chapter B.9 of this Annex (11), for a period of 24 months. Exposure by the inhalation route is carried out for 6 hours per day, 7 days per week, but exposure for 5 days per week may also be used, if justified. The period of exposure will normally be for a period of 24 months. If rodent species other than rats are exposed nose-only, maximum exposure durations may be adjusted to minimise species-specific distress. A rationale should be provided when using an exposure duration of less than 6 hours per day. See also Chapter B.8 of this Annex (9).
31. When the test chemical is administered by gavage to the animals, this should be done using a stomach tube or a suitable intubation cannula, at similar times each day. Normally a single dose will be administered once daily; where for example a chemical is a local irritant, it may be possible to maintain the daily dose-rate by administering it as a split dose (twice a day). The maximum volume of liquid that can be administered at one time depends on the size of the test animal. The volume should be kept as low as practical, and should not normally exceed 1 ml/100g body weight for rodents (31). Variability in test volume should be minimised by adjusting the concentration to ensure a constant volume at all dose levels. Potentially corrosive or irritant chemicals are the exception, and need to be diluted to avoid severe local effects. Testing at concentrations that are likely to be corrosive or irritant to the gastrointestinal tract should be avoided.

**Duration of study**

32. The duration of the study will normally be 24 months for rodents, representing the majority of the normal life span of the animals to be used. Shorter or longer study durations may be used, dependent on the lifespan of the strain of the animal species in the study, but should be justified. For specific strains of mice, e.g. AKR/J, C3H/J or C57BL/6J strains a duration of 18 months may be more appropriate. The following provides some guidance on duration, termination of the study and survival; further guidance, including consideration of the acceptability of a negative carcinogenicity relative to survival in the study, is provided in the OECD Guidance Document No 116 on the Design and Conduct of Chronic Toxicity and Carcinogenicity Studies (7).

— Termination of the study should be considered when the number of survivors in the lower dose groups or the control group falls below 25 per cent.

— In the case where only the high dose group dies prematurely due to toxicity, this should not trigger termination of the study.

— Survival of each sex should be considered separately.

— The study should not be extended beyond the point when the data available from the study are no longer sufficient to enable a statistically valid evaluation to be made.

**▼ M4**

## OBSERVATIONS

33. All animals should be checked for morbidity or mortality, usually at the beginning and the end of each day, including at weekends and holidays. Animals should additionally be checked once a day for specific signs of toxicological relevance, taking into consideration the peak period of anticipated effects after dosing in the case of gavage administration. Particular attention should be paid to tumour development; and the time of tumour onset, location, dimensions, appearance, and progression of each grossly visible or palpable tumour should be recorded.

*Body weight, food/water consumption and food efficiency*

34. All animals should be weighed at the start of treatment, at least once a week for the first 13 weeks and at least monthly thereafter. Measurements of food consumption and food efficiency should be made at least weekly for the first 13 weeks and at least monthly thereafter. Water consumption should be measured at least weekly for the first 13 weeks and at least monthly thereafter when the test chemical is administered in drinking water. Water consumption measurements should also be considered for studies in which drinking activity is altered.

*Haematology, clinical biochemistry and other measurements*

35. In order to maximise the information obtained from the study, especially for mode of action considerations, blood samples may be taken for haematology and clinical biochemistry, and this at the discretion of the study director. Urinalysis may also be appropriate. Further guidance on the value of taking such samples as part of a carcinogenicity study is provided in Guidance Document No 116 (7). If considered appropriate, blood sampling for haematological and clinical chemistry determinations and urinalysis may be conducted as part of an interim kill (paragraph 20) and at study termination on a minimum of 10 animals per sex per group. Blood samples should be taken from a named site, for example by cardiac puncture or from the retro-orbital sinus under anaesthesia, and stored, if applicable, under appropriate conditions. Blood smears may also be prepared for examination, particularly if bone marrow appears to be the target organ, although the value of such examination for the assessment of carcinogenic/oncogenic potential has been questioned (32).

## PATHOLOGY

*Gross necropsy*

36. All animals in the study except sentinel animals (see paragraph 20) and other satellite animals should be subjected to a full, detailed gross necropsy which includes careful examination of the external surface of the body, all orifices, and the cranial, thoracic and abdominal cavities and their contents. Sentinel animals and other satellite animals may require necropsy on a case-by-case basis, at the discretion of the study director. Organ weights are not normally part of a carcinogenesis study, since geriatric changes and, at later stages, the development of tumours confounds the usefulness of organ weight data. They may, however, be critical to performing a weight of evidence evaluation and especially for mode of action considerations. If they are part of a satellite study, they should be collected at no later than one year after initiation of the study.



## ▼M4

37. The following tissues should be preserved in the most appropriate fixation medium for both the type of tissue and the intended subsequent histopathological examination (33) (tissues in square brackets are optional):

all gross lesions	heart	pancreas	stomach (forestomach, glandular stomach)
adrenal gland	ileum	parathyroid gland	[teeth]
aorta	jejunum	peripheral nerve	testis
brain (including sections of cerebrum, cerebellum, and medulla/pons)	kidney	pituitary	thymus
caecum	lacrimal gland (exorbital)	prostate	thyroid
cervix	liver	rectum	[tongue]
coagulating gland	lung	salivary gland	trachea
colon	lymph nodes (both superficial and deep)	seminal vesicle	urinary bladder
duodenum	mammary gland (obligatory for females and, if visibly dissectable, from males)	skeletal muscle	uterus (including cervix)
epididymis	[upper respiratory tract, including nose, turbinates, and paranasal sinuses]	skin	[ureter]
eye (including retina)	oesophagus	spinal cord (at three levels: cervical, mid-thoracic, and lumbar)	[urethra]
[femur with joint]	[olfactory bulb]	spleen	vagina
gall bladder (for species other than rat)	ovary	[sternum],	section of bone marrow and/or a fresh bone marrow aspirate
Harderian gland			

In the case of paired organs, e.g. kidney, adrenal, both organs should be preserved. The clinical and other findings may suggest the need to examine additional tissues. Also, any organs considered likely to be target organs based on the known properties of the test chemical should be preserved. In studies involving the dermal route of administration, the list of organs as set out for the oral route should be preserved, and specific sampling and preservation of the skin from the site of application is essential. In inhalation studies, the list of preserved and examined tissues from the respiratory tract should follow the recommendations of Chapters B.8 and B.29 of this Annex. For other organs/tissues (and in addition to the specifically preserved tissues from the respiratory tract) the list of organs as set out for the oral route should be examined.

**▼ M4***Histopathology*

38. Guidance is available on best practices in the conduct of toxicological pathology studies (33). The minimum tissues examined should be:
- All tissues from the high dose and control groups;
  - All tissues of animals dying or killed during the study;
  - All tissues showing macroscopic abnormalities including tumours;
  - When treatment-related histopathological changes are observed in the high dose group, those same tissues are to be examined from all animals in all other dose groups;
  - In the case of paired organs, e.g. kidney, adrenal, both organs should be examined.

## DATA AND REPORTING

**Data**

39. Individual animal data should be provided for all parameters evaluated. Additionally, all data should be summarised in tabular form showing for each test group the number of animals at the start of the test, the number of animals found dead during the test or killed for humane reasons and the time of any death or humane kill, the number showing signs of toxicity, a description of the signs of toxicity observed, including time of onset, duration, and severity of any toxic effects, the number of animals showing lesions, the type of lesions and the percentage of animals displaying each type of lesion. Summary data tables should provide the means and standard deviations (for continuous test data) of animals showing toxic effects or lesions, in addition to the grading of lesions.
40. Historical control data may be valuable in the interpretation of the results of the study, e.g. in the case when there are indications that the data provided by the concurrent controls are substantially out of line when compared to recent data from control animals from the same test facility/colony. Historical control data, if evaluated, should be submitted from the same laboratory and relate to animals of the same age and strain generated during the five years preceding the study in question.
41. When applicable, numerical results should be evaluated by an appropriate and generally acceptable statistical method. The statistical methods and the data to be analysed should be selected during the design of the study (paragraph 9). Selection should make provision for survival adjustments, if needed.

**Test report**

42. The test report should include the following information:

*Test chemical:*

- physical nature, purity, and physicochemical properties;
- identification data;

**▼ M4**

- source of chemical;
- batch number;
- certificate of chemical analysis;

*Vehicle (if appropriate):*

- justification for choice of vehicle (if other than water);

*Test animals:*

- species/strain used and justification for choice made;
- number, age, and sex of animals at start of test;
- source, housing conditions, diet, etc.;
- individual weights of animals at the start of the test;

*Test conditions:*

- rationale for route of administration and dose selection;
- when applicable, the statistical methods used to analyse the data;
- details of test chemical formulation/diet preparation.
- analytical data on achieved concentration, stability and homogeneity of the preparation;
- route of administration and details of the administration of the test chemical;
- for inhalation studies, whether nose only or whole body;
- actual doses (mg/kg body weight/day), and conversion factor from diet/drinking water test chemical concentration (mg/kg or ppm) to the actual dose, if applicable;
- details of food and water quality;

*Results (summary tabulated data and individual animal data should be presented)**General*

- survival data;
- body weight/body weight changes;
- food consumption, calculations of food efficiency, if made, and water consumption, if applicable;
- toxicokinetic data (if available);
- ophthalmoscopy (if available);
- haematology (if available);
- clinical chemistry (if available);

**▼ M4***Clinical findings*

- Signs of toxicity;
- Incidence (and, if scored, severity) of any abnormality;
- Nature, severity, and duration of clinical observations (whether transitory or permanent);

*Necropsy data*

- Terminal body weight;
- Organ weights and their ratios, if applicable;
- Necropsy findings; Incidence and severity of abnormalities;

*Histopathology*

- Non neoplastic histopathological findings,;
- Neoplastic histopathological findings;
- Correlation between gross and microscopic findings;
- Detailed description of all treatment-related histopathological findings including severity gradings;
- Report of any peer review of slides;

*Statistical treatment of results, as appropriate**Discussion of results including*

- Discussion of any modelling approaches;
- Dose-response relationships;
- Historical control data;
- Consideration of any mode of action information;
- BMD, NOAEL or LOAEL determination;
- Relevance for humans;

*Conclusions**LITERATURE:*

- (1) OECD (1995). Report of the Consultation Meeting on Sub-chronic and Chronic Toxicity/Carcinogenicity Testing (Rome, 1995), internal working document, Environment Directorate, OECD, Paris.
- (2) EPA (2005). Guidelines for Carcinogen Risk Assessment Risk Assessment Forum U.S. Environmental Protection Agency Washington, DC.
- (3) Combes RD, Gaunt, I, Balls M (2004). A Scientific and Animal Welfare Assessment of the OECD Health Effects Test Guidelines for the Safety Testing of Chemicals under the European Union REACH System. ATLA 32: 163-208.
- (4) Barlow SM, Greig JB, Bridges JW et al (2002). Hazard identification by methods of animal-based toxicology. Food. Chem. Toxicol. 40: 145-191.
- (5) Chhabra RS, Bucher JR, Wolfe M, Portier C (2003). Toxicity characterization of environmental chemicals by the US National Toxicology Programme: an overview. Int. J. Hyg. Environ. Health 206: 437-445.

**▼M4**

- (6) Chapter B.27 of this Annex, Sub-chronic Oral Toxicity Test Repeated Dose 90-day Oral Toxicity Study in Non-Rodents.
- (7) OECD (2012). Guidance Document on the Design and Conduct of Chronic Toxicity and Carcinogenicity Studies, Supporting Test Guidelines 451, 452 and 453 — Second edition. Series on Testing and Assessment No 116, available on the OECD public website for Test Guideline at [www.oecd.org/env/testguidelines](http://www.oecd.org/env/testguidelines).
- (8) OECD (2009). Guidance Document on Acute Inhalation Toxicity Testing, Series on Testing and Assessment No 39, ENV/JM/MONO(2009)28, OECD, Paris.
- (9) Chapter B.8 of this Annex, Subacute Inhalation Toxicity: 28-Day Study.
- (10) Chapter B.29 of this Annex, Subchronic Inhalation Toxicity: 90-Day Study.
- (11) Chapter B.9 of this Annex, Repeated Dose (28 Days) Toxicity (Dermal).
- (12) Boobis AR, Cohen SM, Dellarco V, McGregor D, Meek ME, Vickers C, Willcocks D, Farland W (2006). IPCS Framework for analyzing the Relevance of a Cancer Mode of Action for Humans. *Crit. Rev. in Toxicol*, 36: 793-801.
- (13) Cohen SM, Meek ME, Klaunig JE, Patton DE, and Fenner-Crisp PA (2003). The human relevance of information on carcinogenic Modes of Action: An Overview. *Crit. Rev. Toxicol*. 33: 581-589.
- (14) Holsapple MP, Pitot HC, Cohen SN, Boobis AR, Klaunig JE, Pastoor T, Dellarco VL, Dragan YP (2006). Mode of Action in Relevance of Rodent Liver Tumors to Human Cancer Risk. *Toxicol. Sci*. 89: 51-56.
- (15) Meek EM, Bucher JR, Cohen SM, Dellarco V, Hill RN, Lehman-McKemmon LD, Longfellow DG, Pastoor T, Seed J, Patton DE (2003). A Framework for Human Relevance analysis of Information on Carcinogenic Modes of Action. *Crit. Rev. Toxicol*. 33: 591-653.
- (16) Carmichael NG, Barton HA, Boobis AR et al (2006). Agricultural Chemical Safety Assessment: A Multisector Approach to the Modernization of Human Safety Requirements. *Critical Reviews in Toxicology* 36: 1-7.
- (17) Barton HA, Pastoor TP, Baetcke T et al (2006). The Acquisition and Application of Absorption, Distribution, Metabolism, and Excretion (ADME) Data in Agricultural Chemical Safety Assessments. *Critical Reviews in Toxicology* 36: 9-35.
- (18) Doe JE, Boobis AR, Blacker A et al (2006). A Tiered Approach to Systemic Toxicity Testing for Agricultural Chemical Safety Assessment. *Critical Reviews in Toxicology* 36: 37-68.
- (19) Cooper RL, Lamb JS, Barlow SM et al (2006). A Tiered Approach to Life Stages Testing for Agricultural Chemical Safety Assessment. *Critical Reviews in Toxicology* 36: 69-98.
- (20) OECD (2002). Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies, Series on Testing and Assessment No 35 and Series on Pesticides No 14, ENV/JM/MONO(2002)19, OECD, Paris.
- (21) OECD (2000). Guidance Document on the recognition, assessment, and use of clinical signs as humane endpoints for experimental animals used in safety evaluation, Series on Testing and Assessment No 19, ENV/JM/MONO(2000)7, OECD, Paris.

▼ **M4**

- (22) Rhomberg LR, Baetcke K, Blancato J, Bus J, Cohen S, Conolly R, Dixit R, Doe J, Ekelman K, Fenner-Crisp P, Harvey P, Hattis D, Jacobs A, Jacobson-Kram D, Lewandowski T, Liteplo R, Pelkonen O, Rice J, Somers D, Turturro A, West, W, Olin S(2007). Issues in the Design and Interpretation of Chronic Toxicity and Carcinogenicity Studies in Rodents: Approaches to Dose Selection Crit Rev. Toxicol. 37 (9): 729 – 837.
- (23) ILSI (International Life Sciences Institute) (1997). Principles for the Selection of Doses in Chronic Rodent Bioassays. Foran JA (Ed.). ILSI Press, Washington, DC.
- (24) Griffiths SA, Parkinson C, McAuslane JAN and Lumley CE (1994). The utility of the second rodent species in the carcinogenicity testing of pharmaceuticals. The Toxicologist 14(1):214.
- (25) Usui T, Griffiths SA and Lumley CE (1996). The utility of the mouse for the assessment of the carcinogenic potential of pharmaceuticals. In D'Arcy POF & Harron DWG (eds). Proceedings of the Third International Conference on Harmonisation. Queen's University Press, Belfast. pp 279-284.
- (26) Carmichael NG, Enzmann H, Pate I, Waechter F (1997). The Significance of Mouse Liver Tumor Formation for Carcinogenic Risk Assessment: Results and Conclusions from a Survey of 10 Years of Testing by the Agrochemical Industry. Environ Health Perspect. 105:1196-1203.
- (27) Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes (OJ L 276, 20.10.2010, p. 33).
- (28) National Research Council, 1985. Guide for the care and use of laboratory animals. NIH Publication No 86-23. Washington, D.C., US Dept. of Health and Human Services.
- (29) GV-SOLAS (Society for Laboratory Animal Science, Gesellschaft für Versuchstierkunde, 1988). Publication on the Planning and Structure of Animal Facilities for Institutes Performing Animal Experiments. ISBN 3-906255-04-2.
- (30) GV-SOLAS (Society for Laboratory Animal Science, Gesellschaft für Versuchstierkunde, 2006). Microbiological monitoring of laboratory animals in various housing systems.
- (31) Diehl K-H, Hull R, Morton D, Pfister R, Rabemampianina Y, Smith D, Vidal J-M, van de Vorstenbosch C. (2001). A good practice guide to the administration of substances and removal of blood, including routes and volumes. Journal of Applied Toxicology 21:15-23.
- (32) Weingand K, *et al.* (1996). Harmonization of Animal Clinical Pathology Testing in Toxicity and Safety Studies. Fund. Appl. Toxicol. 29: 198-201.
- (33) Crissman J, Goodman D, Hildebrandt P, *et al.* (2004). Best Practices Guideline: Toxicological Histopathology. Toxicologic Pathology 32: 126-131.

▼ M4

*Appendix 1*

DEFINITION

**Test chemical:** Any substance or mixture tested using this Test Method.

**▼M4****B.33. COMBINED CHRONIC TOXICITY/CARCINOGENICITY STUDIES**

## INTRODUCTION

1. This Test Method is equivalent to OECD Test Guideline (TG) 453 (2009). The original TG 453 was adopted in 1981. Development of this updated Test Method B.33 was considered necessary, in order to reflect recent developments in the field of animal welfare and regulatory requirements (1) (2) (3) (4) (5). The updating of this Test Method B.33 has been carried out in parallel with revisions of Chapter B.32 of this Annex, Carcinogenicity Studies, and Chapter B.30 of this Annex, Chronic Toxicity Studies, with the objective of obtaining additional information from the animals used in the study and providing further detail on dose selection. This Test Method is designed to be used in the testing of a broad range of chemicals, including pesticides and industrial chemicals. It should be noted however that some details and requirements may differ for pharmaceuticals [see International Conference on Harmonisation (ICH) Guidance S1B on Testing for Carcinogenicity of Pharmaceuticals].
2. The majority of chronic toxicity and carcinogenicity studies are carried out in rodent species and this Test Method is intended therefore to apply primarily to studies carried out in these species. Should such studies be required in non-rodent species, the principles and procedures outlined may also be applied, with appropriate modifications, together with those outlined in Chapter B.27 of this Annex, Repeated Dose 90-day Oral Toxicity Study in Non-Rodents (6), as outlined in the OECD Guidance Document No 116 on the Design and Conduct of Chronic Toxicity and Carcinogenicity Studies (7).
3. The three main routes of administration used in chronic toxicity/carcinogenicity studies are oral, dermal and inhalation. The choice of the route of administration depends on the physical and chemical characteristics of the test chemical and the predominant route of exposure of humans. Additional information on choice of route of exposure is provided in Guidance Document No 116 (7).
4. This Test Method focuses on exposure via the oral route, the route most commonly used in chronic toxicity and carcinogenicity studies. While long-term studies involving exposure via the dermal or inhalation routes may also be necessary for human health risk assessment and/or may be required under certain regulatory regimes, both routes of exposure involve considerable technical complexity. Such studies will need to be designed on a case-by-case basis, although the Test Method outlined here for the assessment and evaluation of chronic toxicity and carcinogenicity by oral administration could form the basis of a protocol for inhalation and/or dermal studies, with respect to recommendations for treatment periods, clinical and pathology parameters, etc. OECD Guidance is available on the administration of test chemicals by the inhalation (7) (8) and dermal routes (7). Chapter B.8 of this Annex (9) and Chapter B.29 of this Annex (10), together with the OECD Guidance Document on acute inhalation testing (8), should be specifically consulted in the design of longer term studies involving exposure via the inhalation route. Chapter B.9 of this Annex (11) should be consulted in the case of testing carried out by the dermal route.



**▼M4**

5. The combined chronic toxicity/carcinogenicity study provides information on the possible health hazards likely to arise from repeated exposure for a period lasting up to the entire lifespan of the species used. The study will provide information on the toxic effects of the test chemical, including potential carcinogenicity, indicate target organs and the possibility of accumulation. It can provide an estimate of the no-observed-adverse effect level for toxic effects and, in the case of non-genotoxic carcinogens, for tumour responses, which can be used for establishing safety criteria for human exposure. The need for careful clinical observations of the animals, so as to obtain as much information as possible, is also stressed.
6. The objectives of chronic toxicity/carcinogenicity studies covered by this Test Method include:
  - The identification of the carcinogenic properties of a test chemical, resulting in an increased incidence of neoplasms, increased proportion of malignant neoplasms or a reduction in the time to appearance of neoplasms, compared with concurrent control groups;
  - The identification of the time to appearance of neoplasms;
  - The identification of the chronic toxicity of the test chemical;
  - The identification of target organ(s) of chronic toxicity and carcinogenicity,
  - Characterisation of the dose:response relationship,
  - Identification of a no-observed-adverse-effect level (NOAEL) or point of departure for establishment of a Benchmark Dose (BMD),
  - Extrapolation of carcinogenic effects to low dose human exposure levels,
  - Prediction of chronic toxicity effects at human exposure levels,
  - Provision of data to test hypotheses regarding mode of action (2) (7) (12) (13) (14) (15).

**INITIAL CONSIDERATIONS**

7. In the assessment and evaluation of the potential carcinogenicity and chronic toxicity of a test chemical, all available information on the test chemical should be considered by the testing laboratory prior to conducting the study, in order to focus the design of the study to more efficiently test for its toxicological properties and to minimise animal usage. Information on, and consideration of, the mode of action of a suspected carcinogen (2) (7) (12) (13) (14) (15) is particularly important, since the optimal design may differ depending on whether the test chemical is a known or suspected genotoxic carcinogen. Further guidance on mode of action considerations can be found in Guidance Document No 116 (7).
8. Information that will assist in the study design includes the identity, chemical structure, and physico-chemical properties of the test chemical; any information on the mode of action; results of any *in vitro* or *in vivo* toxicity tests including genotoxicity tests; anticipated use(s) and potential for human exposure; available (Q)SAR data, mutagenicity/genotoxicity, carcinogenicity and other toxicological data on structurally-related chemicals; available toxicokinetic data (single dose and also repeat dose kinetics where available) and data derived from other repeated exposure studies. The determination of chronic toxicity/carcinogenicity should only be

**▼ M4**

carried out after initial information on toxicity has been obtained from repeated dose 28-day and/or 90-day toxicity tests. Short-term cancer initiation-promotion tests could also provide useful information. A phased testing approach to carcinogenicity testing should be considered as part of the overall assessment of the potential adverse health effects of a particular test chemical (16) (17) (18) (19).

9. The statistical methods most appropriate for the analysis of results, given the experimental design and objectives, should be established before commencing the study. Issues to consider include whether the statistics should include adjustment for survival, analysis of cumulative tumour risks relative to survival duration, analysis of the time to tumour and analysis in the event of premature termination of one or more groups. Guidance on the appropriate statistical analyses and key references to internationally accepted statistical methods are given in Guidance Document No 116 (7), and also in Guidance Document No 35 on the analysis and evaluation of chronic toxicity and carcinogenicity studies (20).
  
10. In conducting a carcinogenicity study, the guiding principles and considerations outlined in the OECD Guidance Document on the recognition, assessment, and use of clinical signs as humane endpoints for experimental animals used in safety evaluation (21), in particular paragraph 62 thereof, should always be followed. This paragraph states that *'In studies involving repeated dosing, when an animal shows clinical signs that are progressive, leading to further deterioration in condition, an informed decision as to whether or not to humanely kill the animal should be made. The decision should include consideration as to the value of the information to be gained from the continued maintenance of that animal on study relative to its overall condition. If a decision is made to leave the animal on test, the frequency of observations should be increased, as needed. It may also be possible, without adversely affecting the purpose of the test, to temporarily stop dosing if it will relieve the pain or distress, or reduce the test dose.'*
  
11. Detailed guidance on and discussion of the principles of dose selection for chronic toxicity and carcinogenicity studies can be found in Guidance Document No 116 (7), as well as two International Life Sciences Institute publications (22) (23). The core dose selection strategy is dependent on the primary objective or objectives of the study (paragraph 6). In selecting appropriate dose levels, a balance should be achieved between hazard screening on the one hand and characterisation of low-dose responses and their relevance on the other. This is particularly relevant in the case of this combined chronic toxicity and carcinogenicity study.
  
12. Consideration should be given to carrying out this combined chronic toxicity and carcinogenicity study, rather than separate execution of a chronic toxicity study (Chapter B.30 of this Annex) and carcinogenicity study (Chapter B.32 of this Annex). The combined test provides greater efficiency in terms of time and cost, and some reduction in animal use, compared to conducting two separate studies, without compromising the quality of the data in either the chronic phase or the carcinogenicity phase. Careful consideration should however be given to the principles of dose selection (paragraphs 11 and 22-26) when undertaking a combined chronic toxicity and carcinogenicity study, and it is also recognised that separate studies may be required under certain regulatory frameworks. Further guidance on the design of the combined chronic toxicity and carcinogenicity study in order to achieve maximum efficiency of the study in terms of possibilities for reduction in numbers of animals used as well as via the streamlining of the various experimental procedures can be found in Guidance Document No 116 (7).

**▼M4**

13. Definitions used in the context of this Test Method can be found at the end of this chapter and in Guidance Document No 116 (7).

**PRINCIPLE OF THE TEST**

14. The study design consists of two parallel phases, a chronic phase and a carcinogenicity phase (for duration see paragraphs 34 and 35, respectively). The test chemical is normally administered by the oral route although testing by the inhalation or dermal route may also be appropriate. For the chronic phase, the test chemical is administered daily in graduated doses to several groups of test animals, one dose level per group, normally for a period of 12 months, although longer or shorter durations may also be chosen depending on regulatory requirements (see paragraph 34). This duration is chosen to be sufficiently long to allow any effects of cumulative toxicity to become manifest, without the confounding effects of geriatric changes. The study design may also include one or more interim kills, e.g. at 3 and 6 months, and additional groups of animals may be included to accommodate this (see paragraph 20). For the carcinogenicity phase, the test chemical is administered daily to several groups of test animals for a major portion of their life span. The animals in both phases are observed closely for signs of toxicity and for the development of neoplastic lesions. Animals which die or are killed during the test are necropsied and, at the conclusion of the test, surviving animals are killed and necropsied.

**DESCRIPTION OF THE METHOD****Selection of animal species**

15. This Test Method primarily covers assessment and evaluation of chronic toxicity and carcinogenicity in rodents (paragraph 2). The use of non-rodent species may be considered when available data suggest that they are more relevant for the prediction of health effects in humans. The choice of species should be justified. The preferred rodent species is the rat, although other rodent species, e.g. the mouse, may be used. Although the use of the mouse in carcinogenicity testing may have limited utility (24) (25) (26), under some current regulatory programmes carcinogenicity testing in the mouse is still required unless it is determined that such a study is not scientifically necessary. Rats and mice have been preferred experimental models because of their relatively short life span, their widespread use in pharmacological and toxicological studies, their susceptibility to tumour induction, and the availability of sufficiently characterised strains. As a consequence of these characteristics, a large amount of information is available on their physiology and pathology. The design and conduct of chronic toxicity/carcinogenicity studies in non-rodent species, when required, should be based on the principles outlined in this Test Method together with those in Chapter B.27 of this Annex, Repeated Dose 90-day Oral Toxicity Study in Non-Rodents (6). Additional information on choice of species and strain is provided in Guidance Document No 116 (7).
16. Young healthy adult animals of commonly used laboratory strains should be employed. The combined chronic toxicity/carcinogenicity study should be carried out in animals from the same strain and source as those used in preliminary toxicity study(ies) of shorter duration, although, if animals from this strain and source are known to present problems in achieving the normally accepted criteria of survival for long-term studies [see Guidance Document No 116 (7)], consideration should be given to using a strain of animal that has a acceptable survival rate for the long-term study. The females should be nulliparous and non-pregnant.

**▼M4****Housing and feeding conditions**

17. Animals may be housed individually, or be caged in small groups of the same sex; individual housing should be considered only if scientifically justified (27) (28) (29). Cages should be arranged in such a way that possible effects due to cage placement are minimised. The temperature in the experimental animal room should be 22 °C ( $\pm$  3 °C). Although the relative humidity should be at least 30 % and preferably not exceed 70 % other than during room cleaning, the aim should be 50-60 %. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water. The diet should meet all the nutritional requirements of the species tested and the content of dietary contaminants, including but not limited to pesticide residues, persistent organic pollutants, phytoestrogens, heavy metals and mycotoxins, that might influence the outcome of the test, should be as low as possible. Analytical information on the nutrient and dietary contaminant levels should be generated periodically, at least at the beginning of the study and when there is a change in the batch used, and should be included in the final report. Analytical information on the drinking water used in the study should similarly be provided. The choice of diet may be influenced by the need to ensure a suitable admixture of a test chemical and to meet the nutritional requirements of the animals when the test chemical is administered by the dietary route.

**Preparation of animals**

18. Healthy animals, which have been acclimated to laboratory conditions for at least 7 days and have not been subjected to previous experimental procedures, should be used. In the case of rodents, dosing of the animals should begin as soon as possible after weaning and acclimatisation and preferably before the animals are 8 weeks old. The test animals should be characterised as to species, strain, source, sex, weight and age. At the commencement of the study, the weight variation for each sex of animals used should be minimal and not exceed  $\pm$  20 % of the mean weight of all the animals within the study, separately for each sex. Animals should be randomly assigned to the control and treatment groups. After randomisation, there should be no significant differences in mean body weights between groups within each sex. If there are statistically significant differences, then the randomisation step should be repeated, if possible. Each animal should be assigned a unique identification number, and permanently marked with this number by tattooing, microchip implant, or other suitable method.

**PROCEDURE****Number and sex of animals**

19. Both sexes should be used. A sufficient number of animals should be used so that a thorough biological and statistical evaluation is possible. For rodents, each dose group (as outlined in paragraph 22) and concurrent control group intended for the carcinogenicity phase of the study should therefore contain at least 50 animals of each sex. Depending on the aim of the study, it may be possible to increase the statistical power of the key estimates by differentially allocating animals unequally to the various dose groups, with more than 50 animals in the low dose groups, e.g. to estimate the carcinogenic potential in low doses. However it should be recognised that a moderate increase in group size will provide relatively little increase in statistical power of the study. Each dose group (as outlined in paragraph 22) and concurrent control group intended for the chronic toxicity phase of the study should contain at least 10 animals of each sex, in the case of rodents. It should be noted that this number is lower than in the chronic

**▼M4**

toxicity study (Chapter B.30 of this Annex). The interpretation of the data from the reduced number of animals per group in the chronic toxicity phase of this combined study will however be supported by the data from the larger number of animals in the carcinogenicity phase of the study. In studies involving mice, additional animals may be needed in each dose group of the chronic toxicity phase, to conduct all required haematological determinations. Further information on statistical design of the study and choice of dose levels to maximise statistical power is provided in Guidance Document No 116 (7).

**Provision for interim kills, satellite group and sentinel animals**

20. The study may make provision for interim kills, e.g. at 6 months for the chronic toxicity phase, to provide information on progression of non-neoplastic changes and mechanistic information, if scientifically justified. Where such information is already available from previous repeat dose toxicity studies on the test chemical, interim kills may not be scientifically justified. The animals used in the chronic toxicity phase of the study, normally of 12 months duration (paragraph 34) provide interim kill data for the carcinogenicity phase of the study, thus achieving a reduction in the number of animals used overall. Satellite groups may also be included in the chronic toxicity phase of the study, to monitor the reversibility of any toxicological changes induced by the test chemical under investigation. These may be restricted to the highest dose level of the study plus control. An additional group of sentinel animals (typically 5 animals per sex) may be included for monitoring of disease status, if necessary, during the study (30). Further guidance on study design to include interim kills, satellite and sentinel animals, while minimising the number of animals used overall is provided in Guidance Document No 116 (7).
  
21. If satellite animals and/or interim kills are included in the study design, the number of animals in each dose group included for this purpose will normally be 10 animals per sex, and the total number of animals included in the study design should be increased by the number of animals scheduled to be killed before the completion of the study. Interim kill and satellite animals should normally undergo the same observations, including body weight, food/water consumption, haematological and clinical biochemistry measurements and pathological investigations as the animals in the chronic toxicity phase of the main study, although provision may also be made (in the interim kill groups) for measurements to be restricted to specific, key measures such as neurotoxicity or immunotoxicity.

**Dose groups and dosage**

22. Guidance on all aspects of dose selection and dose level spacing is provided in Guidance Document No 116 (7). At least three dose levels and a concurrent control should be used, for both the chronic and carcinogenicity phases. Dose levels will generally be based on the results of shorter-term repeated dose or range finding studies and should take into account any existing toxicological and toxicokinetic data available for the test chemical or related chemicals.

**▼ M4**

23. For the chronic toxicity phase of the study, a full study using three dose levels may not be considered necessary, if it can be anticipated that a test at one dose level, equivalent to at least 1 000 mg/kg body weight/day, is unlikely to produce adverse effects. This should be based on information from preliminary studies and a consideration that toxicity would not be expected, based upon data from structurally related chemicals. A limit of 1 000 mg/kg body weight/day may apply except when human exposure indicates the need for a higher dose level to be used.
  
24. Unless limited by the physical-chemical nature or biological effects of the test chemical, the highest dose level should be chosen to identify the principal target organs and toxic effects while avoiding suffering, severe toxicity, morbidity, or death. The highest dose level should be normally chosen to elicit evidence of toxicity, as evidenced by, for example, depression of body weight gain (approximately 10 %). However, dependent on the objectives of the study (see paragraph 6), a top dose lower than the dose providing evidence of toxicity may be chosen, e.g. if a dose elicits an adverse effect of concern, which nonetheless has little impact on lifespan or body weight.
  
25. Dose levels and dose level spacing may be selected to establish a dose-response and, depending on the mode of action of the test chemical, a NOAEL or other intended outcome of the study, e.g. a BMD (see paragraph 27). Factors that should be considered in the placement of lower doses include the expected slope of the dose-response curve, the doses at which important changes may occur in metabolism or mode of toxic action, where a threshold is expected, or where a point of departure for low-dose extrapolation is expected. In conducting a combined carcinogenicity/chronic toxicity study, the primary objective will be to obtain information for carcinogenicity risk assessment purposes, and information on chronic toxicity will normally be a subsidiary objective. This should be borne in mind when selecting dose levels and dose level spacing for the study.
  
26. The dose level spacing selected will depend on the objectives of the study and the characteristics of the test chemical, and cannot be prescribed in detail in this Test Method, but two to four fold intervals frequently provide good test performance when used for setting the descending dose levels and addition of a fourth test group is often preferable to using very large intervals (e.g. more than a factor of about 6-10) between dosages. In general the use of factors greater than 10 should be avoided, and should be justified if used.
  
27. As outlined further in Guidance Document No 116 (7), points to be considered in dose selection include:
  - Known or suspected nonlinearities or inflection points in the dose-response;
  
  - Toxicokinetics, and dose ranges where metabolic induction, saturation, or nonlinearity between external and internal doses does or does not occur;
  
  - Precursor lesions, markers of effect, or indicators of the operation of key underlying biological processes;

**▼M4**

- Key (or suspected) aspects of mode of action, such as doses at which cytotoxicity begins to arise, hormone levels are perturbed, homeostatic mechanisms are overwhelmed, etc.;
  - Regions of the dose–response curve where particularly robust estimation is needed, e.g. in the range of the anticipated BMD or a suspected threshold;
  - Consideration of anticipated human exposure levels, especially in the choice of mid and low doses.
28. The control group shall be an untreated group or a vehicle-control group if a vehicle is used in administering the test chemical. Except for treatment with the test chemical, animals in the control group should be handled in an identical manner to those in the test groups. If a vehicle is used, the control group shall receive the vehicle in the highest volume used among the dose groups. If a test chemical is administered in the diet, and causes significantly reduced dietary intake due to the reduced palatability of the diet, an additional pair-fed control group may be useful, to serve as a more suitable control.

**Preparation of doses and administration of test chemical**

29. The test chemical is normally administered orally, via the diet or drinking water, or by gavage. Additional information on routes and methods of administration is provided in Guidance Document No 116 (7). The route and method of administration is dependent on the purpose of the study, the physical/chemical properties of the test chemical, its bioavailability, and the predominant route and method of exposure of humans. A rationale should be provided for the chosen route and method of administration. In the interests of animal welfare, oral gavage should normally be selected only for those agents for which this route and method of administration reasonably represent potential human exposure (e.g. pharmaceuticals). For dietary or environmental chemicals including pesticides, administration is typically via the diet or drinking water. However, for some scenarios, e.g. occupational exposure, administration via other routes may be more appropriate.
30. Where necessary, the test chemical is dissolved or suspended in a suitable vehicle. Consideration should be given to the following characteristics of the vehicle and other additives, as appropriate: effects on the absorption, distribution, metabolism, or retention of the test chemical; effects on the chemical properties of the test chemical which may alter its toxic characteristics; and effects on the food or water consumption or the nutritional status of the animals. It is recommended that, wherever possible, the use of an aqueous solution/suspension be considered first, followed by consideration of a solution/emulsion in oil (e.g. corn oil) and then by possible solution in other vehicles. For vehicles other than water, the toxic characteristics of the vehicle should be known. Information should be available on the stability of the test chemical and the homogeneity of dosing solutions or diets (as appropriate) under the conditions of administration (e.g. diet).
31. For chemicals administered via the diet or drinking water it is important to ensure that the quantities of the test chemical involved do not interfere with normal nutrition or water balance. In long-term toxicity studies using dietary administration, the concentration of the test chemical in the feed should not normally exceed an upper limit of 5 % of the total diet, in order to avoid nutritional imbalances. When the test chemical is administered in the diet, either a constant dietary concentration (mg/kg diet or ppm), or a constant dose level in terms of the animal's body weight (mg/kg body weight), calculated on a weekly basis, may be used. The alternative used should be specified.

**▼ M4**

32. In the case of oral administration, the animals are dosed with the test chemical daily (seven days each week) for a period of 12 months (chronic phase) or 24 months (carcinogenicity phase), see also paragraphs 33 and 34. Any other dosing regime, e.g. five days per week, needs to be justified. In the case of dermal administration, animals are normally treated with the test chemical for at least 6 hours per day, 7 days per week, as specified in Chapter B.9 of this Annex (11), for a period of 12 months (chronic phase) or 24 months (carcinogenicity phase). Exposure by the inhalation route is carried out for 6 hours per day, 7 days per week, but exposure for 5 days per week may also be used, if justified. The period of exposure will normally be for a period of 12 months (chronic phase) or 24 months (carcinogenicity phase). If rodent species other than rats are exposed nose-only, maximum exposure durations may be adjusted to minimise species-specific distress. A rationale should be provided when using an exposure duration of less than 6 hours per day. See also Chapter B.8 of this Annex (9).
33. When the test chemical is administered by gavage to the animals this should be done using a stomach tube or a suitable intubation cannula, at similar times each day. Normally a single dose will be administered once daily, where for example a chemical is a local irritant, it may be possible to maintain the daily dose-rate by administering it as a split dose (twice a day). The maximum volume of liquid that can be administered at one time depends on the size of the test animal. The volume should be kept as low as practical, and should not normally exceed 1 ml/100g body weight for rodents (31). Variability in test volume should be minimised by adjusting the concentration to ensure a constant volume at all dose levels. Potentially corrosive or irritant chemicals are the exception, and need to be diluted to avoid severe local effects. Testing at concentrations that are likely to be corrosive or irritant to the gastrointestinal tract should be avoided.

**Duration of study**

34. The period of dosing and duration of the chronic phase of this study is normally 12 months, although the study design also allows for and can be applied to either shorter (e.g. 6 or 9 months) or longer (e.g. 18 or 24 months) duration studies, depending on the requirements of particular regulatory regimes or for specific mechanistic purposes. Deviations from an exposure duration of 12 months should be justified, particularly in the case of shorter durations. All dose groups allocated to this phase will be terminated at the designated time for evaluation of chronic toxicity and non-neoplastic pathology. Satellite groups included to monitor the reversibility of any toxicological changes induced by the test chemical under investigation should be maintained without dosing for a period not less than 4 weeks and not more than one third of the total study duration after cessation of exposure.
35. The duration of the carcinogenicity phase of this study will normally be 24 months for rodents, representing the majority of the normal life span of the animals to be used. Shorter or longer study durations may be used, dependent on the lifespan of the strain of the animal species in the study, but should be justified. For specific strains of mice, e.g. AKR/J, C3H/J or C57BL/6J strains a duration of 18 months may be more appropriate. The following provides some guidance on duration, termination of the study and survival; further guidance, including consideration of the acceptability of a negative carcinogenicity study relative to survival in the study, is provided in Guidance Document No 116 (7):



**▼M4**

- Termination of the study should be considered when the number of survivors in the lower dose groups or the control group falls below 25 per cent.
  
- In the case where only the high dose group dies prematurely due to toxicity, this should not trigger termination of the study.
  
- Survival of each sex should be considered separately.
  
- The study should not be extended beyond the point when the data available from the study are no longer sufficient to enable a statistically valid evaluation to be made.

**OBSERVATIONS (CHRONIC TOXICITY PHASE)**

36. All animals should be checked for morbidity or mortality, usually at the beginning and end of each day, including at weekends and holidays. General clinical observations should be made at least once a day, preferably at the same time(s) each day, taking into consideration the peak period of anticipated effects after dosing in the case of gavage administration.
  
37. Detailed clinical observations should be made on all animals at least once prior to the first exposure (to allow for within-subject comparisons), at the end of the first week of the study and monthly thereafter. The protocol for observations should be arranged such that variations between individual observers are minimised and independent of test group. These observations should be made outside the home cage, preferably in a standard arena and at similar times on each occasion. They should be carefully recorded, preferably using scoring systems, explicitly defined by the testing laboratory. Efforts should be made to ensure that variations in the observation conditions are minimal. Signs noted should include, but not be limited to, changes in skin, fur, eyes, mucous membranes, occurrence of secretions and excretions and autonomic activity (e.g. lacrimation, piloerection, pupil size, unusual respiratory pattern). Changes in gait, posture and response to handling as well as the presence of clonic or tonic movements, stereotypies (e.g. excessive grooming, repetitive circling) or bizarre behaviour (e.g. self-mutilation, walking backwards) should also be recorded (32).
  
38. Ophthalmological examination, using an ophthalmoscope or other suitable equipment, should be carried out on all animals prior to the first administration of the test chemical. At the termination of the study, this examination should be preferably conducted in all animals but at least in the high dose and control groups. If treatment-related changes in the eyes are detected, all animals should be examined. If structural analysis or other information suggests ocular toxicity, then the frequency of ocular examination should be increased.
  
39. For chemicals where previous repeated dose 28-day and/or 90-day toxicity tests indicated the potential to cause neurotoxic effects, sensory reactivity to stimuli of different types (32) (e.g. auditory, visual and proprioceptive stimuli) (33) (34) (35), assessment of grip strength (36) and motor activity assessment (37) may optionally be conducted before commencement of the study and at 3 month periods after study initiation up to and including 12 months, as well as at study termination (if longer than 12 months). Further details of the procedures that could be followed are given in the respective references. However, alternative procedures than those referenced could also be used.

**▼ M4**

40. For chemicals where previous repeated dose 28-day and/or 90-day toxicity tests indicated the potential to cause immunotoxic effects, further investigations of this endpoint may optionally be conducted at termination.

*Body weight, food/water consumption and food efficiency*

41. All animals should be weighed at the start of treatment, at least once a week for the first 13 weeks and at least monthly thereafter. Measurements of food consumption and food efficiency should be made at least weekly for the first 13 weeks and at least monthly thereafter. Water consumption should be measured at least weekly for the first 13 weeks and at least monthly thereafter when the test chemical is administered in drinking water. Water consumption measurements should also be considered for studies in which drinking activity is altered.

*Haematology and clinical biochemistry*

42. In studies involving rodents, haematological examinations should be carried out on all study animals (10 male and 10 female animals per group) at 3, 6, and 12 months, as well as at study termination (if longer than 12 months). In mice, satellite animals may be needed in order to conduct all required haematological determinations (see paragraph 19). In non-rodent studies, samples will be taken from smaller numbers of animals (e.g. 4 animals per sex and per group in dog studies), at interim sampling times and at termination as described for rodents. Measurements at 3 months, either in rodents or non-rodents, need not be conducted if no effect was seen on haematological parameters in a previous 90 day study carried out at comparable dose levels. Blood samples should be taken from a named site, for example by cardiac puncture or from the retro-orbital sinus, under anaesthesia.
43. The following list of parameters should be investigated (38): total and differential leukocyte count, erythrocyte count, platelet count, haemoglobin concentration, haematocrit (packed cell volume), mean corpuscular volume (MCV), mean corpuscular haemoglobin (MCH), mean corpuscular haemoglobin concentration (MCHC), prothrombin time, and activated partial thromboplastin time. Other hematology parameters such as Heinz bodies or other atypical erythrocyte morphology or methaemoglobin may be measured as appropriate depending on the toxicity of the test chemical. Overall, a flexible approach should be adopted, depending on the observed and/or expected effect from a given test chemical. If the test chemical has an effect on the haematopoietic system, reticulocyte counts and bone marrow cytology may also be indicated, although these need not be routinely conducted.
44. Clinical biochemistry determinations to investigate major toxic effects in tissues and, specifically, effects on kidney and liver, should be performed on blood samples obtained from all study animals (10 male and 10 female animals per group), at the same time intervals as specified for the haematological investigations. In mice, satellite animals may be needed in order to conduct all required clinical biochemistry determinations. In non-rodent studies, samples will be taken from smaller numbers of animals (e.g. 4 animals per sex and per group in dog studies), at interim sampling times and at termination as described for rodents. Measurements at 3 months, either in rodents or non-rodents, need not be conducted if no effect was seen on clinical biochemistry parameters in a previous 90 day study carried

**▼ M4**

out at comparable dose levels. Overnight fasting of the animals (with the exception of mice) prior to blood sampling is recommended<sup>(1)</sup>. The following list of parameters should be investigated (38): glucose, urea (urea nitrogen), creatinine, total protein, albumin, calcium, sodium, potassium, total cholesterol, at least two appropriate tests for hepatocellular evaluation (alanine aminotransferase, aspartate aminotransferase, glutamate dehydrogenase, total bile acids) (39), and at least two appropriate tests for hepatobiliary evaluation (alkaline phosphatase, gamma glutamyl transferase, 5'-nucleotidase, total bilirubin, total bile acids) (39). Other clinical chemistry parameters such as fasting triglycerides, specific hormones and cholinesterase may be measured as appropriate, depending on the toxicity of the test chemical. Overall, there is a need for a flexible approach, depending on the observed and/or expected effect from a given test chemical.

45. Urinalysis determinations should be performed on all study animals (10 male and 10 female animals per group), on samples collected at the same intervals as for haematology and clinical chemistry. Measurements at 3 months need not be conducted if no effect was seen on urinalysis in a previous 90 day study carried out at comparable dose levels. The following list of parameters was included in an expert recommendation on clinical pathology studies (38): appearance, volume, osmolality or specific gravity, pH, total protein, and glucose. Other determinations include ketone, urobilinogen, bilirubin, and occult blood. Further parameters may be employed where necessary to extend the investigation of observed effect(s).
  
46. It is generally considered that baseline haematological and clinical biochemistry variables need be determined before treatment for dog studies, but need not be determined in rodent studies (38). However, if historical baseline data (see paragraph 58) are inadequate, consideration should be given to generating such data.

**PATHOLOGY***Gross necropsy*

47. All animals in the study shall be normally subjected to a full, detailed gross necropsy which includes careful examination of the external surface of the body, all orifices, and the cranial, thoracic and abdominal cavities and their contents. However provision may also be made (in the interim kill or satellite groups) for measurements to be restricted to specific, key measures such as neurotoxicity or immunotoxicity (see paragraph 21). These animals need not be subjected to necropsy and the subsequent procedures described in the following paragraphs. Sentinel animals may require necropsy on a case-by-case basis, at the discretion of the study director.

---

<sup>(1)</sup> For a number of measurements in serum and plasma, most notably for glucose, overnight fasting is preferable. The major reason for this preference is that the increased variability which would inevitably result from non-fasting, would tend to mask more subtle effects and make interpretation difficult. However it should be noted that overnight fasting may interfere with the general metabolism of the animals and, particularly in feeding studies, may disrupt the daily exposure to the test chemical. All animals should be assessed in the same physiological condition and preferably detailed or neurological assessments should therefore be scheduled for a different day than clinical biochemistry sampling.

▼ **M4**

48. Organ weights should be collected from all animals, other than those excluded by the latter part of paragraph 47. The adrenals, brain, epididymides, heart, kidneys, liver, ovaries, spleen, testes, thyroid (weighed post-fixation, with parathyroids), and uterus of all animals (apart from those found moribund and/or intercurrently killed) should be trimmed of any adherent tissue, as appropriate, and their wet weight taken as soon as possible after dissection to prevent drying.
49. The following tissues should be preserved in the most appropriate fixation medium for both the type of tissue and the intended subsequent histopathological examination (40) (tissues in square brackets are optional):

all gross lesions	heart	pancreas	stomach (forestomach, glandular stomach)
adrenal gland	ileum	parathyroid gland	[teeth]
aorta	jejunum	peripheral nerve	testis
brain (including sections of cerebrum, cerebellum, and medulla/pons)	kidney	pituitary	thymus
caecum	lacrimal gland (exorbital)	prostate	thyroid
cervix	liver	rectum	[tongue]
coagulating gland	lung	salivary gland	trachea
colon	lymph nodes (both superficial and deep)	seminal vesicle	urinary bladder
duodenum	mammary gland (obligatory for females and, if visibly dissectable, from males)	skeletal muscle	uterus (including cervix)
epididymis	[upper respiratory tract, including nose, turbinates, and paranasal sinuses]	skin	[ureter]
eye (including retina)	oesophagus	spinal cord (at three levels: cervical, mid-thoracic, and lumbar)	[urethra]
[femur with joint]	[olfactory bulb]	spleen	vagina
gall bladder (for species other than rat)	ovary	[sternum],	section of bone marrow and/or a fresh bone marrow aspirate
Harderian gland			

**▼M4**

In the case of paired organs, e.g. kidney, adrenal, both organs should be preserved. The clinical and other findings may suggest the need to examine additional tissues. Also any organs considered likely to be target organs based on the known properties of the test chemical should be preserved. In studies involving the dermal route of administration, the list of organs as set out for the oral route should be examined, and specific sampling and preservation of the skin from the site of application is necessary. In inhalation studies, the list of preserved and examined tissues from the respiratory tract should follow the recommendations of Chapters B.8 of this Annex (9) and Chapter B.29 of this Annex (10). For other organs/tissues (and in addition to the specifically preserved tissues from the respiratory tract) the list of organs as set out for the oral route should be examined.

*Histopathology*

50. Guidance is available on best practices in the conduct of toxicological pathology studies (40). The minimum histopathological examinations should be:
- all tissues from the high dose and control groups;
  - all tissues from animals dying or killed during the study;
  - all tissues showing macroscopic abnormalities;
  - target tissues, or tissues which showed treatment-related changes in the high dose group, from all animals in all other dose groups,
  - in the case of paired organs, e.g. kidney, adrenal, both organs should be examined.

**OBSERVATIONS (CARCINOGENICITY PHASE)**

51. All animals should be checked for morbidity or mortality, usually at the beginning and the end of each day, including at weekends and holidays. Animals should additionally be checked once a day for specific signs of toxicological relevance. In the case of gavage studies, animals should be checked in the period immediately following dosing. Particular attention should be paid to tumour development; and the time of tumour onset, location, dimensions, appearance, and progression of each grossly visible or palpable tumour should be recorded.
52. All animals should be weighed at the start of treatment, at least once a week for the first 13 weeks and at least monthly thereafter. Measurements of food consumption and food efficiency should be made at least weekly for the first 13 weeks and at least monthly thereafter. Water consumption should be measured at least weekly for the first 13 weeks and at least monthly thereafter when the test chemical is administered in drinking water. Water consumption measurements should also be considered for studies in which drinking activity is altered.

▼ **M4***Haematology, clinical biochemistry and other measurements*

53. In order to maximise the information obtained from the study, especially for mode of action considerations, blood samples may be taken for haematology and clinical biochemistry, although this is at the discretion of the study director. Urinalysis may also be appropriate. Data on the animals used in the chronic toxicity phase of the study, normally of 12 months duration (paragraph 34) will provide information on these parameters. Further guidance on the value of taking such samples as part of a carcinogenicity study is provided in Guidance Document No 116 (7). If blood samples are taken, these should be collected at the end of the test period, just prior to or as part of the procedure for killing the animals. They should be taken from a named site, for example by cardiac puncture or from the retro-orbital sinus, under anaesthesia. Blood smears may also be prepared for examination, particularly if bone marrow appears to be the target organ, although the value of such examination of blood smears in the carcinogenicity phase for the assessment of carcinogenic/oncogenic potential has been questioned (38).

## PATHOLOGY

*Gross necropsy*

54. All animals in the study except sentinel animals and other satellite animals (see paragraph 20) shall be subjected to a full, detailed gross necropsy which includes careful examination of the external surface of the body, all orifices, and the cranial, thoracic and abdominal cavities and their contents. Sentinel animals and other satellite animals may require necropsy on a case-by-case basis, at the discretion of the study director. Organ weights are not normally part of a carcinogenesis study, since geriatric changes and, at later stages, the development of tumours confounds the usefulness of organ weight data. They may, however, be critical to performing a weight of evidence evaluation and especially for mode of action considerations. If they are part of a satellite study, they should be collected at no later than one year after initiation of the study.
55. The following tissues should be preserved in the most appropriate fixation medium for both the type of tissue and the intended subsequent histopathological examination (40) (tissues in square brackets are optional):

all gross lesions	heart	pancreas	stomach (forestomach, glandular stomach)
adrenal gland	ileum	parathyroid gland	[teeth]
aorta	jejunum	peripheral nerve	testis
brain (including sections of cerebrum, cerebellum, and medulla/pons)	kidney	pituitary	thymus
caecum	lacrimal gland (exorbital)	prostate	thyroid
cervix	liver	rectum	[tongue]
coagulating gland	lung	salivary gland	trachea
colon	lymph nodes (both superficial and deep)	seminal vesicle	urinary bladder

▼ **M4**

duodenum	mammary gland (obligatory for females and, if visibly dissectable, from males)	skeletal muscle	uterus (including cervix)
epididymis	[upper respiratory tract, including nose, turbinates, and paranasal sinuses]	skin	[ureter]
eye (including retina)	oesophagus	spinal cord (at three levels: cervical, mid-thoracic, and lumbar)	[urethra]
[femur with joint]	[olfactory bulb]	spleen	vagina
gall bladder (for species other than rat)	ovary	[sternum],	section of bone marrow and/or a fresh bone marrow aspirate
Harderian gland			

In the case of paired organs, e.g. kidney, adrenal, both organs should be preserved. The clinical and other findings may suggest the need to examine additional tissues. Also, any organs considered likely to be target organs based on the known properties of the test chemical should be preserved. In studies involving the dermal route of administration, the list of organs as set out for the oral route should be examined, and specific sampling and preservation of the skin from the site of application is necessary. In inhalation studies, the list of preserved and examined tissues from the respiratory tract should follow the recommendations of Chapters B.8 of this Annex (8) and Chapter B.29 of this Annex (9). For other organs/tissues (and in addition to the specifically preserved tissues from the respiratory tract) the list of organs as set out for the oral route should be examined.

*Histopathology*

56. Guidance is available on best practices in the conduct of toxicological pathology studies (40). The minimum tissues examined should be:

- All tissues from the high dose and control groups;
- All tissues of animals dying or killed during the study;
- All tissues showing macroscopic abnormalities including tumours;
- When treatment-related histopathological changes are observed in the high dose group, those same tissues are to be examined from all animals in all other dose groups,
- In the case of paired organs, e.g. kidney, adrenal, both organs should be examined.

**▼ M4****DATA AND REPORTING (CARCINOGENICITY AND CHRONIC TOXICITY)****Data**

57. Individual animal data should be provided for all parameters evaluated. Additionally, all data should be summarised in tabular form showing for each test group the number of animals at the start of the test, the number of animals found dead during the test or killed for humane reasons and the time of any death or humane kill, the number showing signs of toxicity, a description of the signs of toxicity observed, including time of onset, duration, and severity of any toxic effects, the number of animals showing lesions, the type of lesions and the percentage of animals displaying each type of lesion. Summary data tables should provide the means and standard deviations (for continuous test data) of animals showing toxic effects or lesions, in addition to the grading of lesions.
58. Historical control data may be valuable in the interpretation of the results of the study, e.g. in the case when there are indications that the data provided by the concurrent controls are substantially out of line when compared to recent data from control animals from the same test facility/colony. Historical control data, if evaluated, should be submitted from the same laboratory, relate to animals of the same age and strain, generated during the five years preceding the study in question.
59. When applicable, numerical results should be evaluated by an appropriate and generally acceptable statistical method. The statistical methods and the data to be analysed should be selected during the design of the study (paragraph 9). Selection should make provision for survival adjustments, if needed.
60. The test report should include the following information:

*Test chemical:*

- physical nature, purity, and physicochemical properties;
- identification data;
- source of chemical;
- batch number;
- certificate of chemical analysis.

*Vehicle (if appropriate):*

- justification for choice of vehicle (if other than water).

*Test animals:*

- species/strain used and justification for choice made;
- number, age, and sex of animals at start of test;
- source, housing conditions, diet, etc.;
- individual weights of animals at the start of the test.



**▼ M4***Test conditions:*

- rationale for route of administration and dose selection;
- when applicable, the statistical methods used to analyse the data;
- details of test chemical formulation/diet preparation;
- analytical data on achieved concentration, stability and homogeneity of the preparation;
- route of administration and details of the administration of the test chemical;
- for inhalation studies, whether nose only or whole body;
- actual doses (mg/kg body weight/day), and conversion factor from diet/drinking water test chemical concentration (mg/kg or ppm) to the actual dose, if applicable;
- details of food and water quality.

*Results (summary tabulated data and individual animal data should be presented):**General*

- Survival data;
- Body weight/body weight changes;
- Food consumption, calculations of food efficiency, if made, and water consumption if applicable;
- Toxicokinetic data if available;
- Ophthalmoscopy (if available)
- Haematology (if available)
- Clinical chemistry (if available)

*Clinical findings*

- Signs of toxicity;
- Incidence (and, if scored, severity) of any abnormality;
- Nature, severity, and duration of clinical observations (whether transitory or permanent);

*Necropsy data*

- Terminal body weight;
- Organ weights and their ratios, if applicable;
- Necropsy findings; Incidence and severity of abnormalities.

*Histopathology*

- Non neoplastic histopathological findings,
- Neoplastic histopathological findings,

**▼M4**

- Correlation between gross and microscopic findings
- Detailed description of all treatment-related histopathological findings including severity gradings;
- Report of any peer review of slides

*Statistical treatment of results, as appropriate*

*Discussion of results including:*

- Discussion of any modelling approaches
- Dose:response relationships
- Historical control data
- Consideration of any mode of action information
- BMD, NOAEL or LOAEL determination
- Relevance for humans

*Conclusions*

*LITERATURE:*

- (1) OECD (1995). Report of the Consultation Meeting on Sub-chronic and Chronic Toxicity/Carcinogenicity Testing (Rome, 1995), internal working document, Environment Directorate, OECD, Paris.
- (2) EPA (2005). Guidelines for Carcinogen Risk Assessment Risk Assessment Forum U.S. Environmental Protection Agency Washington, DC.
- (3) Combes RD, Gaunt I, Balls M (2004). A Scientific and Animal Welfare Assessment of the OECD Health Effects Test Guidelines for the Safety Testing of Chemicals under the European Union REACH System. ATLA 32: 163-208
- (4) Barlow SM, Greig JB, Bridges JW et al (2002). Hazard identification by methods of animal-based toxicology. Food. Chem. Toxicol. 40: 145-191
- (5) Chhabra RS, Bucher JR, Wolfe M, Portier C (2003). Toxicity characterization of environmental chemicals by the US National Toxicology Programme: an overview. Int. J. Hyg. Environ. Health 206: 437-445
- (6) Chapter B.27 of this Annex, Sub-Chronic Oral Toxicity Test Repeated Dose 90 — Day Oral Toxicity Study In Non-Rodents.
- (7) OECD (2012). Guidance Document on the Design and Conduct of Chronic Toxicity and Carcinogenicity Studies, Supporting Test Guidelines 451, 452 and 453 — Second edition. Series on Testing and Assessment No 116, available on the OECD public website for Test Guideline at [www.oecd.org/env/testguidelines](http://www.oecd.org/env/testguidelines).
- (8) OECD (2009). Guidance Document on Acute Inhalation Toxicity Testing. Series on Testing and Assessment No 39, ENV/JM/MONO(2009)28, OECD, Paris.
- (9) Chapter B.8 of this Annex. Subacute Inhalation Toxicity: 28-Day Study.
- (10) Chapter B.29 of this Annex, Subchronic Inhalation Toxicity: 90-Day Study.
- (11) Chapter B.9 of this Annex, Repeated Dose (28 Days) Toxicity (Dermal).

## ▼M4

- (12) Boobis AR, Cohen SM, Dellarco V, McGregor D, Meek ME, Vickers C, Willcocks D, Farland W (2006). IPCS Framework for analyzing the Relevance of a Cancer Mode of Action for Humans. *Crit. Rev. in Toxicol*, 36: 793-801.
- (13) Cohen SM, Meek ME, Klaunig JE, Patton DE, Fenner-Crisp PA (2003). The human relevance of information on carcinogenic Modes of Action: An Overview. *Crit. Rev. Toxicol*. 33: 581-589.
- (14) Holsapple MP, Pitot HC, Cohen SN, Boobis AR, Klaunig JE, Pastoor T, Dellarco VL, Dragan YP (2006). Mode of Action in Relevance of Rodent Liver Tumors to Human Cancer Risk. *Toxicol. Sci*. 89: 51-56.
- (15) Meek EM, Bucher JR, Cohen SM, Dellarco V, Hill RN, Lehman-McKemmon LD, Longfellow DG, Pastoor T, Seed J, Patton DE (2003). A Framework for Human Relevance analysis of Information on Carcinogenic Modes of Action. *Crit. Rev. Toxicol*. 33: 591-653.
- (16) Carmichael NG, Barton HA, Boobis AR *et al.* (2006). Agricultural Chemical Safety Assessment: A Multisector Approach to the Modernization of Human Safety Requirements. *Crit. Rev. Toxicol*. 36, 1-7.
- (17) Barton HA, Pastoor TP, Baetcke T *et al.* (2006). The Acquisition and Application of Absorption, Distribution, Metabolism, and Excretion (ADME) Data in Agricultural Chemical Safety Assessments. *Crit. Rev. Toxicol*. 36: 9-35.
- (18) Doe JE, Boobis AR, Blacker A *et al.* (2006). A Tiered Approach to Systemic Toxicity Testing for Agricultural Chemical Safety Assessment. *Crit. Rev. Toxicol*. 36: 37-68.
- (19) Cooper RL, Lamb JS, Barlow SM *et al.* (2006). A Tiered Approach to Life Stages Testing for Agricultural Chemical Safety Assessment. *Crit. Rev. Toxicol*. 36: 69-98.
- (20) OECD (2002). Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies, Series on Testing and Assessment No 35 and Series on Pesticides No 14, ENV/JM/MONO(2002)19, OECD, Paris.
- (21) OECD (2000). Guidance Document on the recognition, assessment, and use of clinical signs as humane endpoints for experimental animals used in safety evaluation, Series on Testing and Assessment No 19, ENV/JM/MONO(2000)7, OECD, Paris.
- (22) Rhomberg LR, Baetcke K, Blancato J, Bus J, Cohen S, Conolly R, Dixit R, Doe J, Ekelman K, Fenner-Crisp P, Harvey P, Hattis D, Jacobs A, Jacobson-Kram D, Lewandowski T, Liteplo R, Pelkonen O, Rice J, Somers D, Turturro A, West W, Olin S (2007). Issues in the Design and Interpretation of Chronic Toxicity and Carcinogenicity Studies in Rodents: Approaches to Dose Selection *Crit Rev. Toxicol*. 37 (9): 729 – 837.
- (23) ILSI (International Life Sciences Institute) (1997). Principles for the Selection of Doses in Chronic Rodent Bioassays. Foran JA (Ed.). ILSI Press, Washington, DC.
- (24) Griffiths SA, Parkinson C, McAuslane JAN and Lumley CE (1994). The utility of the second rodent species in the carcinogenicity testing of pharmaceuticals. *The Toxicologist* 14(1):214.
- (25) Usui T, Griffiths SA and Lumley CE (1996). The utility of the mouse for the assessment of the carcinogenic potential of pharmaceuticals. In D'Arcy POF & Harron DWG (eds). *Proceedings of the Third International Conference on Harmonisation*. Queen's University Press, Belfast. pp 279-284.
- (26) Carmichael NG, Enzmann H, Pate I, Waechter F (1997). The Significance of Mouse Liver Tumor Formation for Carcinogenic Risk Assessment: Results and Conclusions from a Survey of 10 Years of Testing by the Agrochemical Industry. *Environ Health Perspect* 105:1196-1203.

▼ **M4**

- (27) Directive 2010/63/EU of the European parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes (OJ L 276, 20.10.2010, p. 33).
- (28) National Research Council, 1985. Guide for the care and use of laboratory animals. NIH Publication No 86-23. Washington D.C., US. Dept. of Health and Human Services.
- (29) GV-SOLAS (Society for Laboratory Animal Science, Gesellschaft für Versuchstierkunde, December, 1989). Publication on the Planning and Structure of Animal Facilities for Institutes Performing Animal Experiments. ISBN 3-906255-06-9.
- (30) GV-SOLAS (Society for Laboratory Animal Science, Gesellschaft für Versuchstierkunde, 2006). Microbiological monitoring of laboratory animals in various housing systems.
- (31) Diehl K-H, Hull R, Morton D, Pfister R, Rabemampianina Y, Smith D, Vidal J-M, van de Vorstenbosch C. (2001). A good practice guide to the administration of substances and removal of blood, including routes and volumes. *Journal of Applied Toxicology*, 21: 15-23.
- (32) IPCS (1986). Principles and Methods for the Assessment of Neurotoxicity Associated with Exposure to Chemicals. Environmental Health Criteria Document No 60.
- (33) Tupper DE, Wallace RB (1980). Utility of the Neurologic Examination in Rats. *Acta Neurobiol. Exp.* 40: 999-1003.
- (34) Gad SC (1982). A Neuromuscular Screen for Use in Industrial Toxicology. *J. Toxicol. Environ. Health* 9: 691-704.
- (35) Moser VC, McDaniel KM, Phillips PM (1991). Rat Strain and Stock Comparisons Using a Functional Observational Battery: Baseline Values and Effects of Amitraz. *Toxicol. Appl. Pharmacol.* 108: 267-283.
- (36) Meyer OA, Tilson HA, Byrd WC, Riley MT (1979). A Method for the Routine Assessment of Fore- and Hind-limb Grip Strength of Rats and Mice. *Neurobehav. Toxicol.* 1: 233-236.
- (37) Crofton KM, Howard JL, Moser VC, Gill MW, Reiter LW, Tilson HA, MacPhail RC (1991). Interlaboratory Comparison of Motor Activity Experiments: Implication for Neurotoxicological Assessments. *Neurotoxicol. Teratol.* 13: 599-609.
- (38) Weingand K, Brown G, Hall R *et al.* (1996). Harmonisation of Animal Clinical Pathology Testing in Toxicity and Safety Studies. *Fundam. & Appl. Toxicol.* 29: 198-201.
- (39) EMEA (draft) document 'Non-clinical guideline on drug-induced hepatotoxicity' (Doc. Ref. EMEA/CHMP/SWP/a50115/2006).
- (40) Crissman JW, Goodman DG, Hildebrandt PK *et al.* (2004). Best Practices Guideline: Toxicological Histopathology. *Toxicologic Pathology* 32: 126-131.

▼ M4

*Appendix 1*

DEFINITION

**Test chemical:** Any substance or mixture tested using this Test Method.

**▼B****B.34. ONE-GENERATION REPRODUCTION TOXICITY TEST****1. METHOD****1.1. INTRODUCTION**

See General introduction Part B.

**1.2. DEFINITIONS**

See General introduction Part B.

**1.3. REFERENCE SUBSTANCES**

None.

**1.4. PRINCIPLE OF THE TEST METHOD**

The test substance is administered in graduated doses to several groups of males and females. Males should be dosed during growth and for at least one complete spermatogenic cycle (approximately 56 days in the mouse and 70 days in the rat) in order to elicit any adverse effects on spermatogenesis by the test substance.

Females of the parental (P) generation should be dosed for at least two complete oestrous cycles in order to elicit adverse effects on oestrus by the test substance. The animals are then mated. The test substance is administered to both sexes during the mating period and thereafter only to females during pregnancy and for the duration of the nursing period. For administration by inhalation the method will require modification.

**1.5. QUALITY CRITERIA**

None.

**1.6. DESCRIPTION OF THE TEST METHOD****1.6.1. Preparations**

Before the test, healthy young adult animals are randomised and assigned to the treated and control groups. The animals are kept under the experimental housing and feeding conditions for at least five days prior to the test. It is recommended that the test substance be administered in the diet or drinking water. Other routes of administration are also acceptable. All animals should be dosed by the same method during the appropriate experimental period. If a vehicle or other additives are used to facilitate dosing, they should be known not to produce toxic effects. Dosing should be on a seven-day per week basis.

**▼B**

## 1.6.2. Experimental animals

## Selection of species

The rat or mouse are the preferred species. Healthy animals, not subjected to previous experimental procedures, should be used. Strains with low fecundity should not be used. The test animals should be characterized as to species, strain, sex, weight and/or age.

For an adequate assessment of fertility, both males and females should be studied. All test and control animals should be weaned before dosing begins.

## Number and sex

Each treated and control group should contain a sufficient number of animals to yield about 20 pregnant females at or near term.

The objective is to produce enough pregnancies and offspring to assure a meaningful evaluation of the potential of the substance to affect fertility, pregnancy and maternal behaviour in P generation animals and suckling, growth and development of the F<sub>1</sub> offspring from conception to weaning.

## 1.6.3. Test conditions

Food and water should be provided ad libitum. Near parturition, pregnant females should be caged separately in delivery or maternity cages and may be provided with nesting materials.

## 1.6.3.1. Dose levels

At least three treated groups and a control group should be used. If a vehicle is used in administering the test substance, the control group should receive the vehicle in the highest volume used. If a test substance causes reduced dietary intake or utilisation, then the use of a paired fed control group may be considered necessary. Ideally, unless limited by the physical/chemical nature or biological effects of the test substance, the highest dose level should induce toxicity but not mortality in the parental (P) animals. The intermediate dose(s) should induce minimal toxic effects attributable to the test substance, and the low dose should not induce any observable adverse effects on the parents or offspring. When administered by gavage or capsule the dosage given to each animal should be based on the individual animal's body weight and adjusted weekly for changes in body weight. For females during pregnancy, dosages may be based on the body weight at day 0 or 6 of the pregnancy, if desired.

## 1.6.3.2. Limit test

In the case of substances of low toxicity, if a dose level of at least 1 000 mg/kilogram produces no evidence of interference with reproductive performance, studies at other dose levels may not be considered necessary. If a preliminary study at the high-dose level, with definite evidence of maternal toxicity, shows no adverse effects on fertility, studies at other dose levels may not be considered necessary.

**▼B**1.6.3.3. *Performance of the test***Experimental schedules**

Daily dosing of the parental (P) males should begin when they are about five to nine weeks of age, after they have been weaned and acclimatised for at least five days. In rats, dosing is continued for 10 weeks prior to the mating period (for mice, eight weeks). Males should be killed and examined either at the end of the mating period or, alternatively, males may be retained on the test diet for the possible production of a second litter and should be killed and examined at some time before the end of the study. For parental (P) females dosing should begin after at least five days of acclimatisation and continue for at least two weeks prior to mating. Daily dosing of the p females should continue throughout the three-week mating period, pregnancy and up to the weaning of the F<sub>1</sub> offspring. Consideration should be given to modification of the dosing schedule based on other available information on the test substance, such as induction of metabolism or bioaccumulation.

**Mating procedure**

Either 1:1 (one male to one female) or 1:2 (one male to two females) mating may be used in reproduction toxicity studies.

Based on 1:1 mating, one female should be placed with the same male until pregnancy occurs or three weeks have elapsed. Each morning the females should be examined for presence of sperm or vaginal plugs. Day 0 of pregnancy is defined as the day a vaginal plug or sperm is found.

Those pairs that fail to mate should be evaluated to determine the cause of the apparent infertility.

This may involve such procedures as providing additional opportunities to mate with other proven sires or dams, microscopic examination of the reproductive organs, and examination of the oestrous cycle or spermatogenesis.

**Litter sizes**

Animals dosed during the fertility study are allowed to litter normally and rear their progeny to the stage of weaning without standardisation of litters.

Where standardisation is done, the following procedure is suggested. Between day 1 and day 4 after birth, the size of each litter may be adjusted by eliminating extra pups by selection to yield, as nearly as possible, four males and four females per litter.

Whenever the number of male or female pups prevents having four of each sex per litter, partial adjustment (for example, five males and three females) is acceptable. Adjustments are not applicable for litters of less than eight pups.



**▼B**

## 1.6.4. Observations

Throughout the test period, each animal should be observed at least once daily. Pertinent behavioural changes, signs of difficult or prolonged parturition, and all signs of toxicity, including mortality, should be recorded. During pre-mating and mating periods, food consumption may be measured daily. After parturition and during lactation, food consumption measurements (and water consumption measurements when the test substance is administered in the drinking water) should be made on the same day as the weighing of the litter. P males and females should be weighed on the first day of dosing and weekly thereafter. These observations should be reported individually for each adult animal.

The duration of gestation should be calculated from day 0 of pregnancy. Each litter should be examined as soon as possible after delivery to establish the number and sex of pups, still births, live births and the presence of gross anomalies.

Dead pups and pups sacrificed at day 4 should be preserved and studied for possible defects. Live pups should be counted and litters weighed on the morning after birth and on days 4 and 7 and weekly thereafter until the termination of the study, when animals should be weighed individually.

Physical or behavioural abnormalities observed in the dams or offspring should be recorded.

## 1.6.5. Pathology

## 1.6.5.1. Necropsy

At the time of sacrifice or death during the study the animals of the P generation should be examined macroscopically for any structural abnormalities or pathological changes, with special attention being paid to the organs of the reproductive system. Dead or moribund pups should be examined for defects.

## 1.6.5.2. Histopathology

The ovaries, uterus, cervix, vagina, testes, epididymes, seminal vesicles, prostate, coagulating gland, pituitary gland and target organ(s) of all P animals should be preserved for microscopic examination. In the event that these organs have not been examined in other multiple-dose studies, they should be microscopically examined in all high-dose and control animals and animals which die during the study where practicable.

Organs showing abnormalities in these animals should then be examined in all other P animals. In these instances, microscopic examination should be made of all tissues showing gross pathological changes. As suggested under mating procedures, reproductive organs of animals suspected of infertility may be subjected to microscopic examination.

**▼ B****2. DATA**

Data may be summarised in tabular form, showing for each test group the number of animals at the start of the test, the number of fertile males, the number of pregnant females, the types of changes and the percentage of animals displaying each type of change.

When possible, numerical results should be evaluated by an appropriate statistical method. Any generally accepted statistical method may be used.

**3. REPORTING****TEST REPORT**

The test report shall, if possible, contain the following information:

- species/strain used,
- toxic response data by sex and dose, including fertility, gestation and viability,
- time of death during the study or whether animals survived to time of scheduled sacrifice or to termination of the study,
- table presenting the weights of each litter, the mean: pup weights and the individual weights of the pups at termination,
- toxic or other effects on reproduction, offspring and postnatal growth,
- the day of observation of each abnormal sign and its subsequent course,
- bodyweight data for P animals,
- necropsy findings,
- a detailed description of all microscopic findings,
- statistical treatment of results, where appropriate,
- discussion of the results,
- interpretation of the results.

**3.2. EVALUATION AND INTERPRETATION**

See General introduction Part B.

**4. REFERENCES**

See General introduction Part B.

**▼B****B.35. TWO-GENERATION REPRODUCTION TOXICITY STUDY****1. METHOD**

This method is a replicate of the OECD TG 416 (2001).

**1.1. INTRODUCTION**

This method for two-generation reproduction testing is designed to provide general information concerning the effects of a test substance on the integrity and performance of the male and female reproductive systems, including gonadal function, the oestrus cycle, mating behaviour, conception, gestation, parturition, lactation, and weaning, and the growth and development of the offspring. The study may also provide information about the effects of the test substance on neonatal morbidity, mortality, and preliminary data on prenatal and postnatal developmental toxicity and serve as a guide for subsequent tests. In addition to studying growth and development of the F1 generation, this test method is also intended to assess the integrity and performance of the male and female reproductive systems as well as growth and development of the F2 generation. For further information on developmental toxicity and functional deficiencies, either additional study segments can be incorporated into this protocol, consulting the methods for developmental toxicity and/or developmental neurotoxicity as appropriate, or these endpoints could be studied in separate studies, using the appropriate test methods.

**1.2. PRINCIPLE OF THE TEST METHOD**

The test substance is administered in graduated doses to several groups of males and females. Males of the P generation should be dosed during growth and for at least one complete spermatogenic cycle (approximately 56 days in the mouse and 70 days in the rat) in order to elicit any adverse effects on spermatogenesis. Effects on sperm are determined by a number of sperm parameters (e.g. sperm morphology and motility) and in tissue preparation and detailed histopathology. If data on spermatogenesis are available from a previous repeated dose study of sufficient duration, e.g. a 90-day study, males of the P generation need not be included in the evaluation. It is recommended, however, that samples or digital recordings of sperm of the P generation are saved, to enable later evaluation. Females of the P generation should be dosed during growth and for several complete oestrus cycles in order to detect any adverse effects on oestrus cycle normality by the test substance. The test substance is administered to parental (P) animals during their mating, during the resulting pregnancies, and through the weaning of their F1 offspring. At weaning the administration of the substance is continued to F1 offspring during their growth into adulthood, mating and production of an F2 generation, until the F2 generation is weaned.

Clinical observations and pathological examinations are performed on all animals for signs of toxicity with special emphasis on effects on the integrity and performance of the male and female reproductive systems and on the growth and development of the offspring.

**▼B**

## 1.3. DESCRIPTION OF THE TEST METHOD

1.3.1. **Selection of animal species**

The rat is the preferred species for testing. If other species are used, justification should be given and appropriate modifications will be necessary. Strains with low fecundity or well-known high incidence of developmental defects should not be used. At the commencement of the study, the weight variation of animals used should be minimal and not exceed 20 % of the mean weight of each sex.

1.3.2. **Housing and feeding conditions**

The temperature in the experimental animal room should be 22 °C ( $\pm$  3°). Although the relative humidity should be at least 30 % and preferably not exceed 70 % other than during room cleaning, the aim should be 50-60 %. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water. The choice of diet may be influenced by the need to ensure a suitable admixture of a test substance when administered by this method.

Animals may be housed individually or be caged in small groups of the same sex. Mating procedures should be carried out in cages suitable for the purpose. After evidence of copulation, mated females shall be single-caged in delivery or maternity cages. Mated rats may also be kept in small groups and separated one or two days prior to parturition. Mated animals shall be provided with appropriate and defined nesting materials when parturition is near.

1.3.3. **Preparation of animals**

Healthy young animals, which have been acclimated to laboratory conditions for at least five days and have not been subjected to previous experimental procedures, should be used. The test animals should be characterised as to species, strain, source, sex, weight and/or age. Any sibling relationships among the animals should be known so that mating of siblings is avoided. The animals should be randomly assigned to the control and treated groups (stratification by body weight is recommended). Cages should be arranged in such a way that possible effects due to cage placement are minimised. Each animal should be assigned a unique identification number. For the P generation, this should be done before dosing starts. For the F1 generation, this should be done at weaning for animals selected for mating. Records indicating the litter of origin should be maintained for all selected F1 animals. In addition, individual identification of pups as soon after birth as possible is recommended when individual weighing of pups or any functional tests are considered.

Parental (P) animals shall be about five to nine weeks old at the start of dosing. The animals of all test groups shall, as nearly as practicable, be of uniform weight and age.

**▼B**

## 1.4. PROCEDURE

1.4.1. **Number and sex of animals**

Each test and control group should contain a sufficient number of animals to yield preferably not less than 20 pregnant females at or near parturition. For substances that cause undesirable treatment related effects (e.g. sterility, excessive toxicity at the high dose) this may not be possible. The objective is to produce enough pregnancies to assure a meaningful evaluation of the potential of the substance to affect fertility, pregnancy and maternal behaviour and suckling, growth and development of the F1 offspring from conception to maturity, and the development of their offspring (F2) to weaning. Therefore, failure to achieve the desired number of pregnant animals (i.e. 20) does not necessarily invalidate the study and should be evaluated on a case-by-case basis.

1.4.2. **Preparation of doses**

It is recommended that the test substance be administered orally (by diet, drinking water or gavage) unless another route of administration (e.g. dermal or inhalation) is considered more appropriate.

Where necessary, the test substance is dissolved or suspended in a suitable vehicle. It is recommended that, wherever possible, the use of an aqueous solution/suspension be considered first, followed by consideration of a solution/emulsion in oil (e.g. corn oil) and then by possible solution in other vehicles. For vehicles other than water, the toxic characteristics of the vehicle must be known. The stability of the test substance in the vehicle should be determined.

1.4.3. **Dosage**

At least three dose levels and a concurrent control shall be used. Unless limited by the physical-chemical nature or biological effects of the test substance, the highest dose level should be chosen with the aim to induce toxicity but not death or severe suffering. In case of unexpected mortality, studies with a mortality rate of less than approximately 10 % in the parental (P) animals would normally still be acceptable. A descending sequence of dose levels should be selected with a view to demonstrating any dosage related effect and no-observed-adverse-effects levels (NOAEL). Two to four fold intervals are frequently optimal for setting the descending dose levels and addition of a fourth test group is often preferable to using very large intervals (e.g. more than a factor of 10) between dosages. For the dietary studies the dose interval should be not more than three fold. Dose levels should be selected taking into account any existing toxicity data, especially results from repeated dose studies. Any available information on metabolism and kinetics of the test compound or related materials should also be considered. In addition, this information will also assist in demonstrating the adequacy of the dosing regimen.

**▼B**

The control group shall be an untreated group or a vehicle-control group if a vehicle is used in administering the test substance. Except for treatment with the test substance, animals in the control group should be handled in an identical manner to the test group subjects. If a vehicle is used, the control group shall receive the vehicle in the highest volume used. If a test substance is administered in the diet, and causes reduced dietary intake or utilisation, then the use of a pair-fed control group may be considered necessary. Alternatively data from controlled studies designed to evaluate the effects of decreased food consumption on reproductive parameters may be used in lieu of a concurrent pair-fed control group.

Consideration should be given to the following characteristics of vehicle and other additives: effects on the absorption, distribution, metabolism, or retention of the test substance; effects on the chemical properties of the test substance which may alter its toxic characteristics; and effects on the food or water consumption or the nutritional status of the animals.

**1.4.4. Limit test**

If an oral study at one dose level of at least 1 000 mg/kg body weight/day or, for dietary or drinking water administration, an equivalent percentage in the diet or drinking water using the procedures described for this study, produces no observable toxic effects in either parental animals or their offspring and if toxicity would not be expected based upon data from structurally and/or metabolically related compounds, then a full study using several dose levels may not be considered necessary. The limit test applies except when human exposure indicates the need for a higher oral dose level to be used. For other types of administration, such as inhalation or dermal application, the physical-chemical properties of the test substance, such as solubility, often may indicate and limit the maximum attainable level of exposure.

**1.4.5. Administration of doses**

The animals should be dosed with the test substance on a 7-days per week basis. The oral route of administration (diet, drinking water, or gavage) is preferred. If another route of administration is used, justification shall be provided, and appropriate modifications may be necessary. All animals shall be dosed by the same method during the appropriate experimental period. When the test substance is administered by gavage, this should be done using a stomach tube. The volume of liquid administered at one time should not exceed 1 ml/100 g body weight (0,4 ml/100 g body weight is the maximum for corn oil), except in the case of aqueous solutions where 2 ml/100 g body weight may be used. Except for irritant or corrosive substances, which will normally reveal exacerbated effects with higher concentrations, variability in test volume should be minimised by adjusting the concentration to ensure a constant volume at all dose levels. In gavage studies, the pups will normally only receive test substance indirectly through the milk, until direct dosing commences for them at weaning. In diet or drinking water studies, the pups will additionally receive test substance directly when they commence eating for themselves during the last week of the lactation period.

**▼ B**

For substances administered via the diet or drinking water, it is important to ensure that the quantities of the test substance involved do not interfere with normal nutrition or water balance. When the test substance is administered in the diet either a constant dietary concentration (ppm) or a constant dose level in terms of the body weight of the animal may be used; the alternative used must be specified. For a substance administered by gavage, the dose should be given at similar times each day, and adjusted at least weekly to maintain a constant dose level in terms of animal body weight. Information regarding placental distribution should be considered when adjusting the gavage dose based on weight.

**1.4.6. Experimental schedules**

Daily dosing of the parental (P) males and females shall begin when they are five to nine weeks old. Daily dosing of the F1 males and females shall begin at weaning; it should be kept in mind that in cases of test substance administration via diet or drinking water, direct exposure of the F1 pups to the test substance may already occur during the lactation period. For both sexes (P and F1), dosing shall be continued for at least 10 weeks before the mating period. Dosing is continued in both sexes during the two week mating period. Males should be humanely killed and examined when they are no longer needed for assessment of reproductive effects. For parental (P) females, dosing should continue throughout pregnancy and up to the weaning of the F1 offspring. Consideration should be given to modifications in the dosing schedule based on available information on the test substance, including existing toxicity data, induction of metabolism or bioaccumulation. The dose to each animal should normally be based on the most recent individual body weight determination. However, caution should be exercised when adjusting the dose during the last trimester of pregnancy.

Treatment of the P and F1 males and females shall continue until termination. All P and F1 adult males and females should be humanely killed when they are no longer needed for assessment of reproductive effects. F1 offspring not selected for mating and all F2 offspring should be humanely killed after weaning.

**1.4.7. Mating procedure****1.4.7.1. Parental (P) mating**

For each mating, each female shall be placed with a single male from the same dose level (1:1 mating) until copulation occurs or two weeks have elapsed. Each day, the females shall be examined for presence of sperm or vaginal plugs. Day 0 of pregnancy is defined as the day a vaginal plug or sperm are found. In case pairing is unsuccessful, re-mating of females with proven males of the same group could be considered. Mating pairs should be clearly identified in the data. Mating of siblings should be avoided.

**▼B**1.4.7.2. *F1 mating*

For mating the F1 offspring, at least one male and one female should be selected at weaning from each litter for mating with other pups of the same dose level but different litter, to produce the F2 generation. Selection of pups from each litter should be random when no significant differences are observed in body weight or appearance between the litter mates. In case these differences are observed, the best representatives of each litter should be selected. Pragmatically, this is best done on a body weight basis but it may be more appropriate on the basis of appearance. The F1 offspring should not be mated until they have attained full sexual maturity.

Pairs without progeny should be evaluated to determine the apparent cause of the infertility. This may involve such procedures as additional opportunities to mate with other proven sires or dams, microscopic examination of the reproductive organs, and examination of the oestrous cycles or spermatogenesis.

1.4.7.3. *Second mating*

In certain instances, such as treatment-related alterations in litter size or the observation of an equivocal effect in the first mating, it is recommended that the P or F1 adults be remated to produce a second litter. It is recommended to remate females or males, which have not produced a litter with proven breeders of the opposite sex. If production of a second litter is deemed necessary in either generation, animals should be remated approximately one week after weaning of the last litter.

1.4.7.4. *Litter size*

Animals shall be allowed to litter normally and rear their offspring to weaning. Standardisation of litter sizes is optional. When standardisation is done, the method used should be described in detail.

## 1.5. OBSERVATIONS

1.5.1. **Clinical observations**

A general clinical observation should be made each day and, in the case of gavage dosing its timing should take into account the anticipated peak period of effects after dosing. Behavioural changes, signs of difficult or prolonged parturition and all signs of toxicity should be recorded. An additional, more detailed examination of each animal should be conducted on at least a weekly basis and could conveniently be performed on an occasion when the animal is weighed. Twice daily, during the weekend once daily when appropriate, all animals should be observed for morbidity and mortality.



**▼ B****1.5.2. Body weight and food/water consumption of parent animals**

Parental animals (P and F1) shall be weighed on the first day of dosing and at least weekly thereafter. Parental females (P and F1) shall be weighed at a minimum on gestation days 0, 7, 14, and 20 or 21, and during lactation on the same days as the weighing of litters and on the day the animals are killed. These observations should be reported individually for each adult animal. During the pre-mating and gestation periods food consumption shall be measured weekly at a minimum. Water consumption shall be measured weekly at a minimum if the test substance is administered in the water.

**1.5.3. Oestrus cycle**

Estrous cycle length and normality are evaluated in P and F1 females by vaginal smears prior to mating, and optionally during mating, until evidence of mating is found. When obtaining vaginal/cervical cells, care should be taken to avoid disturbance of mucosa and subsequently, the induction of pseudopregnancy (1).

**1.5.4. Sperm parameters**

For all P and F1 males at termination, testis and epididymis weight shall be recorded and one of each organ reserved for histopathological examination (see Section 1.5.7, 1.5.8.1). Of a subset of at least 10 males of each group of P and F1 males, the remaining testes and epididymides should be used for enumeration of homogenisation-resistant spermatids and cauda epididymal sperm reserves, respectively. For this same subset of males, sperm from the cauda epididymides or vas deferens should be collected for evaluation of sperm motility and sperm morphology. If treatment-related effects are observed or when there is evidence from other studies of possible effects on spermatogenesis, sperm evaluation should be conducted in all males in each dose group; otherwise enumeration may be restricted to control and high-dose P and F1 males.

The total number of homogenisation-resistant testicular spermatids and cauda epididymal sperm should be enumerated (2)(3). Cauda sperm reserves can be derived from the concentration and volume of sperm in the suspension used to complete the qualitative evaluations, and the number of sperm recovered by subsequent mincing and/or homogenising of the remaining cauda tissue. Enumeration should be performed on the selected subset of males of all dose groups immediately after killing the animals unless video or digital recordings are made, or unless the specimens are frozen and analysed later. In these instances, the controls and high dose group may be analysed first. If no treatment-related effects (e.g. effects on sperm count, motility, or morphology) are seen the other dose groups need not be analysed. When treatment-related effects are noted in the high-dose group, then the lower dose groups should also be evaluated.

**▼ B**

Epididymal (or ductus deferens) sperm motility should be evaluated or video taped immediately after sacrifice. Sperm should be recovered while minimising damage, and diluted for motility analysis using acceptable methods (4). The percentage of progressively motile sperm should be determined either subjectively or objectively. When computer-assisted motion analysis is performed (5)(6)(7)(8)(9)(10) the derivation of progressive motility relies on user-defined thresholds for average path velocity and straightness or linear index. If samples are videotaped (11) or the images are otherwise recorded at the time of necropsy, subsequent analysis of only control and high-dose P and F1 males may be performed unless treatment-related effects are observed; in that case, the lower dose groups should also be evaluated. In the absence of a video or digital image, all samples in all treatment groups should be analysed at necropsy.

A morphological evaluation of an epididymal (or *vas deferens*) sperm sample should be performed. Sperm (at least 200 per sample) should be examined as fixed, wet preparations (12) and classified as either normal or abnormal. Examples of morphologic sperm abnormalities would include fusion, isolated heads, and misshapen heads and/or tails. Evaluation should be performed on the selected subset of males of all dose groups either immediately after killing the animals, or, based on the video or digital recordings, at a later time. Smears, once fixed, can also be read at a later time. In these instances, the controls and high dose group may be analysed first. If no treatment-related effects (e.g. effects on sperm morphology) are seen the other dose groups need not be analysed. When treatment-related effects are noted in the high-dose group, then the lower dose groups should also be evaluated.

If any of the above sperm evaluation parameters have already been examined as part of a systemic toxicity study of at least 90 days, they need not necessarily be repeated in the two-generation study. It is recommended, however, that samples or digital recordings of sperm of the P generation are saved, to enable later evaluation, if necessary.

**1.5.5. Offspring**

Each litter should be examined as soon as possible after delivery (lactation day 0) to establish the number and sex of pups, stillbirths, live births, and the presence of gross anomalies. Pups found dead on day 0, if not macerated, should preferably be examined for possible defects and cause of death and preserved. Live pups should be counted and weighed individually at birth (lactation day 0) or on day 1, and on regular weigh days thereafter, e.g. on days 4, 7, 14, and 21 of lactation. Physical or behavioural abnormalities observed in the dams or offspring should be recorded.

**▼ B**

Physical development of the offspring should be recorded mainly by body weight gain. Other physical parameters (e.g. ear and eye opening, tooth eruption, hair growth) may give supplementary information, but these data should preferably be evaluated in the context of data on sexual maturation (e.g. age and body weight at vaginal opening or balano-preputial separation) (13). Functional investigations (e.g. motor activity, sensory function, reflex ontogeny) of the F1 offspring before and/or after weaning, particularly those related to sexual maturation, are recommended if such investigations are not included in separate studies. The age of vaginal opening and preputial separation should be determined for F1 weanlings selected for mating. Anogenital distance should be measured at postnatal day 0 in F2 pups if triggered by alterations in F1 sex ratio or timing of sexual maturation.

Functional observations may be omitted in groups that otherwise reveal clear signs of adverse effects (e.g. significant decrease in weight gain, etc.). If functional investigations are made, they should not be done on pups selected for mating.

**1.5.6. Gross necropsy**

At the time of termination or death during the study, all parental animals (P and F1), all pups with external abnormalities or clinical signs, as well as one randomly selected pup/sex/litter from both the F1 and F2 generation, shall be examined macroscopically for any structural abnormalities or pathological changes. Special attention should be paid to the organs of the reproductive system. Pups that are humanely killed in a moribund condition and dead pups, when not macerated, should be examined for possible defects and/or cause of death and preserved.

The uteri of all primiparous females should be examined, in a manner which does not compromise histopathological evaluation, for the presence and number of implantation sites.

**1.5.7. Organ weights**

At the time of termination, body weight and the weight of the following organs of all P and F1 parental animals shall be determined (paired organs should be weighed individually):

- uterus, ovaries,
- testes, epididymides (total and cauda),
- prostate,
- seminal vesicles with coagulating glands and their fluids and prostate (as one unit),
- brain, liver, kidneys, spleen, pituitary, thyroid and adrenal glands and known target organs.

Terminal body weights should be determined for F1 and F2 pups that are selected for necropsy. The following organs from the one randomly selected pup/sex/litter (see Section 1.5.6) shall be weighed: Brain, spleen and thymus.

**▼B**

Gross necropsy and organ weight results should be assessed in context with observations made in other repeated dose studies, when feasible.

1.5.8. **Histopathology**

1.5.8.1. *Parental animals*

The following organs and tissues of parental (P and F1) animals, or representative samples thereof, shall be fixed and stored in a suitable medium for histopathological examination.

- Vagina, uterus with cervix, and ovaries (preserved in appropriate fixative),
- one testis (preserved in Bouin's or comparable fixative), one epididymis, seminal vesicles, prostate, and coagulating gland,
- previously identified target organ(s) from all P and F1 animals selected for mating.

Full histopathology of the preserved organs and tissues listed above should be performed for all high dose and control P and F1 animals selected for mating. Examination of the ovaries of the P animals is optional. Organs demonstrating treatment-related changes should also be examined in the low- and mid-dose groups to aid in the elucidation of the NOAEL. Additionally, reproductive organs of the low- and mid-dose animals suspected of reduced fertility, e.g. those that failed to mate, conceive, sire, or deliver healthy offspring, or for which oestrus cyclicity or sperm number, motility, or morphology were affected, should be subjected to histopathological evaluation. All gross lesions such as atrophy or tumours shall be examined.

Detailed testicular histopathological examination (e.g. using Bouin's fixative, paraffin embedding and transverse sections of 4-5 µm thickness) should be conducted in order to identify treatment-related effects such as retained spermatids, missing germ cell layers or types, multinucleated giant cells or sloughing of spermatogenic cells into the lumen (14). Examination of the intact epididymis should include the caput, corpus, and cauda, which can be accomplished by evaluation of a longitudinal section. The epididymis should be evaluated for leukocyte infiltration, change in prevalence of cell types, aberrant cell types, and phagocytosis of sperm. PAS and haematoxylin staining may be used for examination of the male reproductive organs.

The postlactational ovary should contain primordial and growing follicles as well as the large corpora lutea of lactation. Histopathological examination should detect qualitative depletion of the primordial follicle population. A quantitative evaluation of primordial follicles should be conducted for F1 females; the number of animals, ovarian section selection, and section sample size should be statistically appropriate for the evaluation procedure used. Examination should include enumeration of the number of primordial follicles, which can be combined with small growing follicles, for comparison of treated and control ovaries (15)(16)(17)(18)(19).

**▼B**1.5.8.2. *Weanlings*

Grossly abnormal tissue and target organs from all pups with external abnormalities or clinical signs, as well as from the one randomly selected pup/sex/litter from both the F1 and F2 generation which have not been selected for mating, shall be fixed and stored in a suitable medium for histopathological examination. Full histopathological characterisation of preserved tissue should be performed with special emphasis on the organs of the reproductive system.

**2. DATA**

## 2.1. TREATMENT OF RESULTS

Data shall be reported individually and summarised in tabular form, showing for each test group and each generation the number of animals at the start of the test, the number of animals found dead during the test or killed for humane reasons, the time of any death or humane kill, the number of fertile animals, the number of pregnant females, the number of animals showing signs of toxicity, a description of the signs of toxicity observed, including time of onset, duration, and severity of any toxic effects, the types of parental and offspring observations, the types of histopathological changes, and all relevant litter data.

Numerical results should be evaluated by an appropriate, generally accepted statistical method; the statistical methods should be selected as part of the design of the study and should be justified. Dose-response statistical models may be useful for analysing data. The report should include sufficient information on the method of analysis and the computer program employed, so that an independent reviewer/statistician can re-evaluate and reconstruct the analysis.

## 2.2. EVALUATION OF RESULTS

The findings of this two-generation reproduction toxicity study should be evaluated in terms of the observed effects including necropsy and microscopic findings. The evaluation will include the relationship, or lack thereof, between the dose of the test substance and the presence or absence, incidence and severity of abnormalities, including gross lesions, identified target organs, affected fertility, clinical abnormalities, affected reproductive and litter performance, body weight changes, effects on mortality and any other toxic effects. The physico-chemical properties of the test substance, and when available, toxicokinetics data should be taken into consideration when evaluating test results.

A properly conducted reproduction toxicity test should provide a satisfactory estimation of a no-effect level and an understanding of adverse effects on reproduction, parturition, lactation, postnatal development including growth and sexual development.

**▼ B**

## 2.3. INTERPRETATION OF RESULTS

A two-generation reproduction toxicity study will provide information on the effects of repeated exposure to a substance during all phases of the reproductive cycle. In particular, the study provides information on the reproductive parameters, and on development, growth, maturation and survival of offspring. The results of the study should be interpreted in conjunction with the findings from subchronic, prenatal developmental and toxicokinetic and other available studies. The results of this study can be used in assessing the need for further testing of a chemical. Extrapolation of the results of the study to man is valid to a limited degree. They are best used to provide information on no-effect-levels and permissible human exposure (20)(21)(22)(23).

3. **REPORTING**

## 3.1. TEST REPORT

The test report must include the following information:

Test substance:

- physical nature and, where relevant, physicochemical properties,
- identification data,
- purity.

Vehicle (if appropriate):

- justification for choice of vehicle if other than water.

**Test animals:**

- species/strain used,
- number, age and sex of animals,
- source, housing conditions, diet, nesting materials, etc.,
- individual weights of animals at the start of the test.

**Test conditions:**

- rationale for dose level selection,
- details of test substance formulation/diet preparation, achieved concentrations,
- stability and homogeneity of the preparation,
- details of the administration of the test substance,
- conversion from diet/drinking water test substance concentration (ppm) to the achieved dose (mg/kg body weight/day), if applicable,
- details of food and water quality.

**▼ B**

## Results:

- food consumption, and water consumption if available, food efficiency (body weight gain per gram of food consumed), and test material consumption for P and F1 animals, except for the period of cohabitation and for at least the last third of lactation,
- absorption data (if available),
- body weight data for P and F1 animals selected for mating,
- litter and pup weight data,
- body weight at sacrifice and absolute and relative organ weight data for the parental animals,
- nature, severity and duration of clinical observations (whether reversible or not),
- time of death during the study or whether animals survived to termination,
- toxic response data by sex and dose, including indices of mating, fertility, gestation, birth, viability, and lactation; the report should indicate the numbers used in calculating these indices,
- toxic or other effects on reproduction, offspring, post-natal growth, etc.,
- necropsy findings,
- detailed description of all histopathological findings,
- number of P and F1 females cycling normally and cycle length,
- total cauda epididymal sperm number, percent progressively motile sperm, percent morphologically normal sperm, and percent of sperm with each identified abnormality,
- time-to-mating, including the number of days until mating,
- gestation length,
- number of implantations, corpora lutea, litter size,
- number of live births and post-implantation loss,
- number of pups with grossly visible abnormalities, if determined the number of runts should be reported,
- data on physical landmarks in pups and other post natal developmental data, physical landmarks evaluated should be justified,
- data on functional observations in pups and adults, as applicable,
- statistical treatment of results, where appropriate.

**▼B**

Discussion of results.

Conclusions, including NOAEL values for maternal and offspring effects.

#### 4. REFERENCES

- (1) Sadleir, R.M.F.S., (1979) Cycles and Seasons, In: *Reproduction in Mammals: I. Germ Cells and Fertilisation*, C.R. Auston and R.V. Short (eds.), Cambridge, New York.
- (2) Gray, L.E. et al., (1989) A Dose-Response Analysis of Methoxychlor-Induced Alterations of Reproductive Development and Function in the Rat. *Fundamental and Applied Toxicology*, 12, p. 92-108.
- (3) Robb, G.W. et al., (1978). Daily Sperm Production and Epididymal Sperm Reserves of Pubertal and Adult Rats. *Journal of Reproduction and Fertility* 54:103-107.
- (4) Klinefelter, G.R. et al., (1991) The Method of Sperm Collection Significantly Influences Sperm Motion Parameters Following Ethane Dimethanesulfonate Administration in the Rat. *Reproductive Toxicology*, 5, p. 39-44.
- (5) Seed, J. et al., (1996). Methods for Assessing Sperm Motility, Morphology, and Counts in the Rat, Rabbit, and Dog: a Consensus Report. *Reproductive Toxicology*, 10(3), p. 237-244.
- (6) Chapin, R.E. et al., (1992) Methods for Assessing Rat Sperm Motility. *Reproductive Toxicology*, 6, p. 267-273
- (7) Klinefelter, G.R. et al., (1992) Direct Effects of Ethane Dimethanesulphonate on Epididymal Function in Adult Rats: an *In Vitro* Demonstration. *Journal of Andrology*, 13, p. 409-421.
- (8) Slott, V.L. et al., (1991) Rat Sperm Motility Analysis: Methodologic Considerations. *Reproductive Toxicology*, 5, p. 449-458.
- (9) Slott, V.L. and Perreault, S.D., (1993) Computer-Assisted Sperm Analysis of Rodent Epididymal Sperm Motility Using the Hamilton-Thorn Motility Analyzer. In: *Methods in Toxicology*, Part A., Academic, Orlando, Florida, p. 319-333.
- (10) Toth, G.P. et al., (1989) The Automated Analysis of Rat Sperm Motility Following Subchronic Epichlorhydrin Administration: Methodologic and Statistical Considerations. *Journal of Andrology*, 10, p. 401-415.
- (11) Working, P.K. and M. Hurtt, (1987) Computerised Videomicrographic Analysis of Rat Sperm Motility. *Journal of Andrology*, 8, p. 330-337.
- (12) Linder, R.E. et al., (1992) Endpoints of Spermatotoxicity in the Rat After Short Duration Exposures to Fourteen Reproductive Toxicants. *Reproductive Toxicology*, 6, p. 491-505.
- (13) Korenbrot, C.C. et al., (1977) Preputial Separation as an External Sign of Pubertal Development in the Male Rat. *Biological Reproduction*, 17, p. 298-303.
- (14) Russell, L.D. et al., (1990) Histological and Histopathological Evaluation of the Testis, Cache River Press, Clearwater, Florida.
- (15) Heindel, J.J. and R.E. Chapin, (eds.) (1993) Part B. Female Reproductive Systems, *Methods in Toxicology*, Academic, Orlando, Florida.
- (16) Heindel, J.J. et al., (1989) Histological Assessment of Ovarian Follicle Number in Mice As a Screen of Ovarian Toxicity. In: *Growth Factors and the Ovary*, A.N. Hirshfield (ed.), Plenum, New York, p. 421-426.
- (17) Manson, J.M. and Y.J. Kang, (1989) Test Methods for Assessing Female Reproductive and Developmental Toxicology. In: *Principles and Methods of Toxicology*, A.W. Hayes (ed.), Raven, New York.



**▼B**

- (18) Smith, B.J. et al., (1991) Comparison of Random and Serial Sections in Assessment of Ovarian Toxicity. *Reproductive Toxicology*, 5, p. 379-383.
- (19) Heindel, J.J., (1999) Oocyte Quantitation and Ovarian Histology. In: *An Evaluation and Interpretation of Reproductive Endpoints for Human Health Risk Assessment*, G. Daston, and C.A. Kimmel, (eds.), ILSI Press, Washington, DC.
- (20) Thomas, J. A., (1991) Toxic Responses of the Reproductive System. In: *Casarett and Doull's Toxicology*, M.O. Amdur, J. Doull, and C.D. Klaassen (eds.), Pergamon, New York.
- (21) Zenick, H. and E.D. Clegg, (1989) Assessment of Male Reproductive Toxicity: A Risk Assessment Approach. In: *Principles and Methods of Toxicology*, A.W. Hayes (ed.), Raven Press, New York.
- (22) Palmer, A.K., (1981) In: *Developmental Toxicology*, Kimmel, C.A. and J. Buelke-Sam (eds.), Raven Press, New York.
- (23) Palmer, A.K., (1978) In *Handbook of Teratology*, Vol. 4, J.G. Wilson and F.C. Fraser (eds.), Plenum Press, New York.

**▼M4****B.36. TOXICOKINETICS**

## INTRODUCTION

1. This Test Method is equivalent to OECD TG 417 (2010). Studies examining the toxicokinetics (TK) of a test chemical are conducted to obtain adequate information on its absorption, distribution, biotransformation (i.e. metabolism) and excretion, to aid in relating concentration or dose to the observed toxicity, and to aid in understanding its mechanism of toxicity. TK may help to understand the toxicology studies by demonstrating that the test animals are systemically exposed to the test chemical and by revealing which are the circulating moieties (parent chemical/metabolites). Basic TK parameters determined from these studies will also provide information on the potential for accumulation of the test chemical in tissues and/or organs and the potential for induction of biotransformation as a result of exposure to the test chemical.
2. TK data can contribute to the assessment of the adequacy and relevance of animal toxicity data for extrapolation to human hazard and/or risk assessment. Additionally, toxicokinetic studies may provide useful information for determining dose levels for toxicity studies (linear vs. non-linear kinetics), route of administration effects, bioavailability, and issues related to study design. Certain types of TK data can be used in physiologically based toxicokinetic (PBTK) model development.
3. There are important uses for metabolite/TK data such as suggesting possible toxicities and modes of action and their relation to dose level and route of exposure. In addition, metabolism data can provide information useful for assessing the toxicological significance of exposures to exogenously produced metabolites of the test chemical.
4. Adequate toxicokinetic data will be helpful to support the further acceptability and applicability of quantitative structure-activity relationships, read-across or grouping approaches in the safety evaluation of chemicals. Kinetics data may also be used to evaluate the toxicological relevance of other studies (e.g. *in vivo/in vitro*).
5. Unless another route of administration is mentioned (see in particular paragraphs 74-78), this Test Method is applicable to oral administration of the test chemical.

## INITIAL CONSIDERATIONS

6. Regulatory systems have different requirements and needs regarding the measurement of endpoints and parameters related to toxicokinetics for different classes of chemicals (e.g. pesticides, biocides, industrial chemicals). Unlike most Test Methods this Test Method describes toxicokinetics testing, which involves multiple measurements and endpoints. In the future, several new Test Methods, and/or guidance document(s), may be developed to describe each endpoint separately and in more detail. In the case of this Test Method, which tests or assessments are conducted is specified by the requirements and/or needs of each regulatory system.

**▼ M4**

7. There are numerous studies that might be performed to evaluate the TK behaviour of a test chemical for regulatory purposes. However, depending on particular regulatory needs or situations, not all of these possible studies may be necessary for the evaluation of a test chemical. Flexibility, taking into consideration the characteristics of the test chemical being investigated, is needed in the design of toxicokinetic studies. In some cases, only a certain set of questions may need to be explored in order to address test chemical-associated hazard and risk concerns. In some situations, TK data can be collected as part of the evaluation in other toxicology studies. For other situations, additional and/or more extensive TK studies may be necessary, depending on regulatory needs and/or if new questions arise as part of test chemical evaluation.
  
8. All available information on the test chemical and relevant metabolites and analogues should be considered by the testing laboratory prior to conducting the study in order to enhance study quality and avoid unnecessary animal use. This could include data from other relevant Test Methods (*in vivo* studies, *in vitro* studies, and/or *in silico* evaluations). Physicochemical properties, such as octanol-water partition coefficient (expressed as  $\log P_{OW}$ ), pKa, water solubility, vapour pressure, and molecular weight of a chemical may be useful for study planning and interpretation of results. They can be determined using appropriate methods as described in the relevant Test Methods.

**LIMITATIONS**

9. This Test Method is not designed to address special circumstances, such as the pregnant or lactating animal and offspring, or to evaluate potential residues in exposed food-producing animals. However, the data obtained from a B.36 study can provide background information to guide the design of specific studies for these investigations. This Test Method is not intended for the testing of nanomaterials. A report on preliminary review of OECD Test Guidelines for their applicability to nanomaterials indicates that TG 417 (equivalent to this Test Method B.36) may not apply to nanomaterials (1).

**DEFINITIONS**

10. Definitions used for the purpose of this Test Method are provided in Appendix.

**ANIMAL WELFARE CONSIDERATIONS**

11. Guidance on humane treatment of animals is available in OECD Guidance Document (GD) 19 (2). It is recommended that OECD GD 19 be consulted for all *in vivo* and *in vitro* studies described in this Test Method.

**DESCRIPTION OF THE METHODS****Pilot Studies**

12. The use of pilot studies is recommended and encouraged for the selection of experimental parameters for the toxicokinetics studies (e.g. metabolism, mass balance, analytical procedures, dose-finding, exhalation of CO<sub>2</sub>, etc.). Characterisation of some of these parameters may not necessitate the use of radiolabelled chemicals.

**▼ M4****Animal Selection***Species*

13. The animal species (and strain) used for TK testing should preferably be the same as that used in other toxicological studies performed with the test chemical of interest. Normally, the rat should be used as it has been used extensively for toxicological studies. The use of other or additional species may be warranted if critical toxicology studies demonstrate evidence of significant toxicity in these species or if their toxicity/toxicokinetics is shown to be more relevant to humans. Justification should be provided for the selection of the animal species and its strain.
  
14. Unless mentioned otherwise, this Test Method refers to the rat as the test species. Certain aspects of the method might have to be modified for the use of other test species.

*Age and Strain*

15. Young healthy adult animals (normally 6-12 weeks at the time of dosing) should be used (see also paragraphs 13 and 14). Justification should be provided for the use of animals that are not young adults. All animals should be of similar age at the outset of the study. The weight variation of individual animals should not exceed  $\pm 20\%$  of the mean weight of the test group. Ideally, the strain used should be the same as that used in deriving the toxicological database for the test chemical.

*Number and Sex of Animals*

16. A minimum of four animals of one sex should be used for each dose tested. Justification should be provided for the sex of the animals used. The use of both sexes (four males and four females) should be considered if there is evidence to support significant sex-related differences in toxicity.

*Housing and feeding conditions*

17. Animals should generally be housed individually during the testing period. Group housing might be justified in special circumstances. Lighting should be artificial, the sequence being 12 h light/12 h dark. The temperature of the experimental animal room should be 22 °C ( $\pm 3$  °C) and the relative humidity 30-70 %. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water.

**Test Chemical**

18. A radiolabelled test chemical using  $^{14}\text{C}$  should be used for all mass balance and metabolite identification aspects of the study; however, if it can be demonstrated that:

— mass balance and metabolite identification can be adequately evaluated using the unlabelled test chemical,

— the analytical specificity and sensitivity of the method used with non-radioactive test chemical is equal to or greater than that which could be obtained with the radiolabelled test chemical,

**▼M4**

then a radiolabelled test chemical does not need to be used. Furthermore, other radioactive and stable isotopes may be used, particularly if the element is responsible for or is a part of the toxic portion of the test chemical. If possible, the radiolabel should be located in a core portion of the molecule which is metabolically stable (it is not exchangeable, is not removed metabolically as CO<sub>2</sub>, and does not become part of the one-carbon pool of the organism). Labelling of multiple sites or specific regions of the molecule may be necessary to follow the metabolic fate of the test chemical.

19. The radiolabelled and non-radiolabelled test chemicals should be analysed using appropriate methods to establish purity and identity. The radio-purity of the radioactive test chemical should be the highest attainable for a particular test chemical (ideally it should be greater than 95 %) and reasonable effort should be made to identify impurities present at or above 2 %. The purity, along with the identity and proportion of any impurities which have been identified, should be reported. Individual regulatory programmes may choose to provide additional guidance to assist in the definition and specifications of test chemicals composed of mixtures and methods for determination of purity.

**Dose Selection***Pilot Study*

20. Usually a single oral dose is sufficient for the pilot study. The dose should be non-toxic, but high enough to allow for metabolite identification in excreta (and plasma, if appropriate) as well as to meet the stated purpose of the pilot study as noted in paragraph 12 of this Test Method.

*Main Studies*

21. For the main studies, a minimum of two doses is preferred since information gathered from at least two dose groups may aid in dose setting in other toxicity studies, and help in the dose-response assessment of already available toxicity tests.
22. Where two doses are administered, both doses should be high enough to allow for metabolite identification in excreta (and plasma, if appropriate). Information from available toxicity data should be considered for dose selection. If information is not available (e.g. from acute oral toxicity studies recording clinical signs of toxicity, or from repeated dose toxicity studies) a value for the higher dose that is below the LD<sub>50</sub> (oral and dermal routes) or LC<sub>50</sub> (inhalation route) estimate or below the lower value of the acute toxicity range estimate may be considered. The lower dose should be some fraction of the higher dose.
23. If only one dose level is investigated, ideally the dose should be high enough to allow for metabolite identification in excreta (and plasma, if appropriate), while not producing apparent toxicity. A rationale should be provided as to why no second dose level has been included.
24. If the effect of dose on kinetic processes needs to be established, two doses may not be sufficient and at least one dose should be high enough so as to saturate these processes. If the area under the plasma concentration-time curve (AUC) is not linear between two dose levels used in the main study, this is a strong indication that saturation of one or more of the kinetic processes is occurring somewhere between the two dose levels.

**▼ M4**

25. For test chemicals of low toxicity, a maximum dose of 1 000 mg/kg body weight (oral and dermal routes) should be used (if administration is by the inhalation route, refer to Chapter B.2 of this Annex for guidance; typically this dose would not exceed 2 mg/l). Chemical-specific considerations may necessitate a higher dose depending on regulatory needs. Dose selection should always be justified.
  
26. Single dose toxicokinetic and tissue distribution data may be adequate to determine the potential for accumulation and/or persistence. However in some circumstances repeated dose administration may be needed (i) to address more fully the potential for accumulation and/or persistence or changes in TK (i.e. for instance, enzyme induction and inhibition), or (ii) as required by the applicable regulatory system. In studies involving repeated dosing, while repeated low dose administration is usually sufficient, under certain circumstances repeated high dose administration may also be necessary (see also paragraph 57).

**Administration of Test Chemical**

27. The test chemical should be dissolved or suspended homogeneously in the same vehicle employed for the other oral gavage toxicity studies performed with the test chemical, if such vehicle information is available. Rationale for the choice of vehicle should be provided. The choice of the vehicle and the volume of dosing should be considered in the design of the study. The customary method of administration is by gavage; however, administration by gelatine capsule or as a dietary mixture may be advantageous in specific situations (in both cases, justification should be given). Verification of the actual dose administered to each animal should be provided.
  
28. The maximum volume of liquid to be administered by oral gavage at one time depends on the size of the test animals, the type of dose vehicle, and whether or not feed is withheld prior to administration of the test chemical. The rationale for administering or restricting food prior to dosing should be provided. Normally the volume should be kept as low as practical for either aqueous or non-aqueous vehicles. Dose volumes should not normally exceed 10 ml/kg body weight for rodents. Volumes of vehicles used for more lipophilic test chemicals might start at 4 ml/kg body weight. For repeated dosing, when daily fasting would be contraindicated, lower dose volumes (e.g. 2-4 ml/kg body weight) should be considered. Where possible, consideration may be given to the use of a dose volume consistent with that administered in other oral gavage studies for a test chemical.
  
29. Intravenous (IV) administration of the test chemical and measurement of the test chemical in blood and/or excreta may be used to establish bioavailability or relative oral absorption. For the IV study, a single dose (usually equivalent to but not to exceed the lower oral dose – see dose selection) of test chemical is administered using an appropriate vehicle. This material should be administered in a suitable volume (e.g. 1 ml/kg bw) at the chosen site of administration to at least four animals of the appropriate sex (both sexes might be used, if warranted, see paragraph 16). A fully dissolved or suspended dose preparation is necessary for IV administration of the test chemical. The vehicle for IV administration should not interfere with blood integrity or blood flow. If the test chemical is infused, the

**▼ M4**

infusion rate should be reported and standardised between animals, provided an infusion pump is used. Anaesthesia should be used if one cannulates the jugular vein (for administration of test chemical and/or collection of blood) or if one uses the femoral artery for administration. Due consideration should be given to the type of anaesthesia as it may have effects on toxicokinetics. Animals should be allowed to recover adequately before administration of the test chemical plus the vehicle.

30. Other routes of administration, such as dermal and inhalation, (see paragraphs 74-78) may be applicable for certain test chemicals, considering their physico-chemical properties and the expected human use or exposure.

**Measurements***Mass Balance*

31. Mass balance is determined by summation of the percent of the administered (radioactive) dose excreted in urine, faeces, and expired air, and the percent present in tissues, residual carcass, and cage wash (see paragraph 46). Generally, total recoveries of administered test chemical (radioactivity) in the order of > 90 % are considered to be adequate.

*Absorption*

32. An initial estimation of absorption can be achieved by excluding the percentage of dose in the gastro-intestinal (GI) tract and/or faeces from the mass balance determination. For the calculation of percent absorption, see paragraph 33. For investigation of excreta, see paragraphs 44-49. If the exact extent of absorption following oral dosing cannot be established from mass balance studies (e.g. where greater than 20 % of the administered dose is present in faeces), further investigations might be necessary. These studies could comprise either 1) oral administration of test chemical and measurement of test chemical in bile or 2) oral and IV administration of test chemical and measurement of net test chemical present in urine plus expired air plus carcass by each of the two routes. In either study design, measurement of radioactivity is conducted as a surrogate method for chemical-specific analysis of test chemical plus metabolites.
33. If a biliary excretion study is undertaken, the oral route of administration is typically used. In this study, the bile ducts of at least four animals of the appropriate sex (or of both sexes, if warranted) should be cannulated and a single dose of the test chemical should be administered. Following administration of the test chemical, excretion of radioactivity/test chemical in bile should be monitored as long as necessary to estimate the percentage of the administered dose that is excreted via this route, which can be used to directly calculate the extent of oral absorption, as follows:

$$\text{Percent absorption} = \frac{(\text{amount in bile} + \text{urine} + \text{expired air} + \text{carcass without GI tract contents})}{\text{amount administered}} \times 100$$

34. With some classes of test chemical, direct secretion of the absorbed dose can occur across intestinal membranes. In such cases the measurement of % dose in faeces following an oral dose in the bile duct cannulated rat is not considered to be representative of the unabsorbed dose. It is recommended that where intestinal secretion is thought to occur then the % dose absorbed be based on the absorption calculated from a comparison of the excretion following the oral versus IV route (intact or bile duct cannulated rat) (see paragraph 35). It is also recommended that where quantification of the intestinal secretion is considered necessary, excretion in the bile duct cannulated rat following IV dose administration be measured.

▼ **M4***Bioavailability*

35. Bioavailability can be determined from plasma/blood kinetics of the oral and IV groups, as described in paragraphs 50-52, by specific chemical analysis of the test chemical and/or relevant metabolite(s), therefore not requiring radiolabelled test chemical. The calculation of bioavailability (F) of the test chemical or relevant metabolite(s) can then be made as follows:

$$F = (AUC_{\text{exp}}/AUC_{\text{IV}}) \times (\text{Dose}_{\text{IV}}/\text{Dose}_{\text{exp}})$$

where AUC is the area under the plasma concentration-time curve, and exp is the experimental route (oral, dermal or via inhalation).

36. For use in risk assessment of systemic effects, bioavailability of the toxic component is in general preferred over the percent absorption when comparing systemic concentrations from animal studies with analogous biomonitoring data from worker exposure studies. The situation may become more complex if doses are in the non-linear range so it is important that toxicokinetic screening determines doses in the linear range.

*Tissue Distribution*

37. Knowledge of tissue distribution of a test chemical and/or its metabolites is important for the identification of target tissues, and understanding of the underlying mechanisms of toxicity, and in order to get information on the potential for test chemical and metabolite accumulation and persistence. The percent of the total (radioactive) dose in tissues as well as residual carcass should at a minimum be measured at the termination of the excretion experiment (e.g. typically up to 7 days post dose or less depending on test chemical specific behaviour). When no test chemical is detected in tissues at study termination (e.g. because the test chemical might have been eliminated before study termination due to a short half-life), care should be taken in order to prevent misinterpretation of the data. In this type of situation, tissue distribution should be investigated at the time of test chemical (and/or metabolite) peak plasma/blood concentration ( $T_{\text{max}}$ ) or peak rate of urinary excretion, as appropriate (see paragraph 38). Furthermore, tissue collection at additional time points may be needed to determine tissue distribution of the test chemical and/or its metabolites, to evaluate time dependency (if appropriate), to aid in establishing mass balance, and/or as required by a competent authority. Tissues that should be collected include liver, fat, GI tract, kidney, spleen, whole blood, residual carcass, target organ tissues and any other tissues (e.g. thyroid, erythrocytes, reproductive organs, skin, eye (particularly in pigmented animals) of potential significance in the toxicological evaluation of the test chemical. Analysis of additional tissues at the same time points should be considered to maximise utilisation of animals and in the event that target organ toxicity is observed in sub-chronic or chronic toxicity studies. The (radioactive) residue concentration and tissue-to-plasma (blood) ratios should also be reported.
38. The evaluation of tissue distribution at additional time points such as the time of peak plasma/blood concentration (e.g.  $T_{\text{max}}$ ) or the peak rate of urinary excretion, obtained from the respective plasma/blood kinetic or excretion experiments, may also be needed or required by a competent authority. This information can be useful for understanding toxicity and the potential for test chemical and metabolite accumulation and persistence. Justification for sample selection should be provided; samples for analysis generally should be the same as those above (see paragraph 37).



**▼ M4**

39. Quantification of radioactivity for tissue distribution studies can be performed using organ dissection, homogenisation, combustion and/or solubilisation, followed by liquid scintillation counting (LSC) of trapped residues. Certain techniques, currently at various stages of development, e.g. Quantitative whole-body autoradiography and receptor microscopic autoradiography, may prove useful in determining the distribution of a test chemical in organs and/or tissues (3) (4).
  
40. For routes of exposure other than oral, specific tissues should be collected and analysed, such as lungs in inhalation studies and skin in dermal studies. See paragraphs 74-78.

*Metabolism*

41. Excreta (and plasma, if appropriate) should be collected for identification and quantitation of unchanged test chemical and metabolites as described under paragraphs 44-49. Pooling of excreta to facilitate metabolite identification within a given dose group is acceptable. Profiling of metabolites from each time period is recommended. However, if lack of sample and/or radioactivity precludes this, pooling of urine and faeces across several time points is acceptable but pooling across sexes or doses is not acceptable. Appropriate qualitative and quantitative methods should be used to assay urine, faeces, expired radioactivity from treated animals, and bile if appropriate.
  
42. Reasonable efforts should be made to identify all metabolites present at 5 % or greater of the administered dose and to provide a metabolic scheme for the test chemical. Test chemicals which have been characterised in excreta as comprising 5 % or greater of the administered dose should be identified. Identification refers to the exact structural determination of components. Typically, identification is accomplished either by co-chromatography of the metabolite with known standards using two dissimilar systems or by techniques capable of positive structural identification such as mass spectrometry, nuclear magnetic resonance (NMR), etc. In the case of co-chromatography, chromatographic techniques utilising the same stationary phase with two different solvent systems are not considered to be an adequate two-method verification of metabolite identity, since the methods are not independent. Identification by co-chromatography should be obtained using two dissimilar, analytically independent systems such as reverse and normal phase thin layer chromatography (TLC) and high performance liquid chromatography (HPLC). Provided that the chromatographic separation is of suitable quality, then additional confirmation by spectroscopic means is not necessary. Alternatively, unambiguous identification can also be obtained using methods providing structural information such as: liquid chromatography/mass spectrometry (LC-MS), or liquid chromatography/tandem mass spectrometry (LC-MS/MS), gas chromatography/mass spectrometry (GC-MS), and NMR spectrometry.
  
43. If identification of metabolites at 5 % or greater of the administered dose is not possible, a justification/explanation should be provided in the final report. It might be appropriate to identify metabolites representing less than 5 % of the administered dose to gain a better understanding of the metabolic pathway for hazard and/or risk assessment of the test chemical. Structural confirmation should be provided whenever possible. This may include profiling in plasma or blood or other tissues.

**▼M4***Excretion*

44. The rate and extent of excretion of the administered dose should be determined by measuring the percent recovered (radioactive) dose from urine, faeces and expired air. These data will also assist in establishing mass balance. The quantities of test chemical (radioactivity) eliminated in the urine, faeces, and expired air should be determined at appropriate time intervals (see paragraphs 47-49). Repeated dose experiments should be properly designed to allow for collection of excretion data to meet the objectives described in the paragraph 26. This will allow for comparison to single dose experiments.
45. If a pilot study has shown that no significant amount of test chemical (radioactivity) (according to paragraph 49) is excreted in expired air, then expired air does not need to be collected in the definitive study.
46. Each animal is to be placed in a separate metabolic unit for collection of excreta (urine, faeces and expired air). At the end of each collection period (see paragraphs 47-49), the metabolic units should be rinsed with appropriate solvent (this is known as the 'cage wash') to ensure maximum recovery of the test chemical (radioactivity). Collection of excreta should be terminated at 7 days, or after at least 90 % of the administered dose has been recovered, whichever occurs first.
47. The total quantities of test chemical (radioactivity) in urine are to be determined for at least two time points on day 1 of collection, one of which should be at 24 h post dosing, and daily thereafter until study termination. The selection of more than two sampling points on day one (e.g. at 6, 12 and 24 h) is encouraged. The results of pilot studies should be analysed for information on alternate or additional time points for collection. A rationale should be provided for the collection schedules.
48. The total quantities of test chemical (radioactivity) in faeces should be determined on a daily basis beginning at 24 h post-dosing until study termination, unless pilot studies suggest alternate or additional time points for collection. A rationale should be provided for alternative collection schedules.
49. The collection of expired CO<sub>2</sub> and other volatile materials may be discontinued in a given study experiment when less than 1 % of the administered dose is found in the exhaled air during a 24-h collection period.

**Time Course Studies***Plasma/Blood Kinetics*

50. The purpose of these studies is to obtain estimates of basic TK parameters [e.g. C<sub>max</sub>, T<sub>max</sub>, half-life (t<sub>1/2</sub>), AUC] for the test chemical. These studies may be conducted at one dose or, more likely, at two or more doses. Dose setting should be determined by the nature of the experiment and/or the issue being addressed. Kinetic data may be needed to resolve issues such as test chemical bioavailability and/or to clarify the effect of dose on clearance (e.g. to clarify whether clearance is saturated in a dose-dependent fashion).
51. For these studies a minimum of four animals of one sex per dose group should be used. Justification should be provided for the sex of the animals used. The use of both sexes (four males and four females) should be considered if there is evidence to support significant sex-related differences in toxicity.

**▼ M4**

52. Following administration of the test chemical (radiolabelled), blood samples should be obtained from each animal at suitable time points using appropriate sampling methodology. The volume and number of blood samples which can be obtained per animal might be limited by potential effects of repeated sampling on animal health/physiology and/or the sensitivity of the analytical method. Samples should be analysed for each individual animal. In some circumstances (e.g. metabolite characterisation), it might be necessary to pool samples from more than one animal. Pooled samples should be clearly identified and an explanation for pooling provided. If a radiolabelled test chemical is used, analysis of total radioactivity present might be adequate. If so, total radioactivity should be analyzed in whole blood and plasma or plasma and red blood cells to allow calculation of the blood/plasma ratio. In other circumstances, more specific investigations requiring the identification of parent compound and/or metabolites, or to assess protein binding might be necessary.

*Other Tissue Kinetics*

53. The purpose of these studies is to obtain time course information to address questions related to issues such as toxic mode of action, bioaccumulation and bio-persistence via determination of levels of test chemical in various tissues. The selection of tissues and the number of time points evaluated will depend on the issue to be addressed and the toxicological database for the test chemical. The design of these additional tissue kinetics studies should take into account information gathered as described in paragraphs 37-40. These studies might involve single or repeated dosing. A detailed rationale for the approach used should be provided.
54. Reasons for performing other tissue kinetic studies might include:
- Evidence of extended blood half-life, suggesting possible accumulation of test chemical in various tissues, or
  - interest in seeing if a steady state level has been achieved in specific tissues (e.g. in repeated dosing studies, even though an apparent blood steady state level of test chemical may have been achieved, there may be interest in ascertaining that a steady state level has also been attained in target tissues).
55. For these types of time-course studies, an appropriate oral dose of test chemical should be administered to a minimum of four animals per dose per time point and the time course of distribution monitored in selected tissues. Only one sex may be used, unless gender specific toxicity is observed. Whether total radioactivity or parent chemical and/or metabolites are analysed will also depend on the issue being addressed. Assessment of tissue distribution should be made using appropriate techniques.

*Enzyme Induction/Inhibition*

56. Studies addressing the possible effects of enzyme induction/inhibition or biotransformation of test chemical under study may be needed under one or more of the following cases:
- (1) Available evidence indicates a relationship between biotransformation of test chemical and enhanced toxicity;
  - (2) The available toxicity data indicate a non-linear relationship between dose and metabolism;

**▼ M4**

- (3) The results of metabolite identification studies show identification of a potentially toxic metabolite that might have been produced by an enzyme pathway induced by the test chemical;
  - (4) In explaining effects which are postulated to be linked to enzyme induction phenomena;
  - (5) If toxicologically significant alterations in the metabolic profile of the test chemical are observed through either *in vitro* or *in vivo* experiments with different species or conditions, characterisation of the enzyme(s) involved may be needed (e.g. Phase I enzymes such as isoenzymes of the Cytochrome P450-dependent mono-oxygenase system, Phase II enzymes such as isoenzymes of sulfotransferase or uridine diphosphate glucuronosyl transferase, or any other relevant enzymes). This information might be used to evaluate the pertinence of species to species extrapolations.
57. Appropriate study protocols to evaluate test chemical related changes in TK, suitably validated and justified should be used. Example study designs consist of repeated dosing with unlabelled test chemical, followed by a single radiolabelled dose on day 14, or repeated dosing with radiolabelled test chemical and sampling at days 1, 7 and 14 for determination of metabolite profiles. Repeated dosing with radiolabelled test chemical may also provide information on bioaccumulation (see paragraph 26).

**SUPPLEMENTAL APPROACHES**

58. Supplemental approaches beyond the *in vivo* experiments described in this Test Method may provide useful information on the absorption, distribution, metabolism or elimination of a test chemical in certain species.

**Use of *in vitro* information**

59. Several questions concerning the metabolism of the test chemical may be addressed in *in vitro* studies using appropriate test systems. Freshly isolated or cultured hepatocytes and subcellular fractions (e.g. microsomes and cytosol or S9 fraction) from liver may be used to study possible metabolites. Local metabolism in the target organ, e.g. lung, may be of interest for risk assessment. For these purposes, microsomal fractions of target tissues may be useful. Studies with microsomes may be useful to address potential gender and life-stage differences and characterise enzyme parameters ( $K_m$  and  $V_{max}$ ) which can aid in the assessment of dose dependency of metabolism in relation to exposure levels. In addition microsomes may be useful to identify the specific microsomal enzymes involved in the metabolism of the test chemical which can be relevant in species extrapolation (see also paragraph 56). The potential for induction of biotransformation can also be examined by using liver subcellular fractions (e.g. microsomes and cytosol) of animals pre-treated with the test chemical of interest, *in vitro* via hepatocyte induction studies or from specific cell lines expressing relevant enzymes. In certain circumstances and under appropriate conditions, subcellular fractions coming from human tissues might be considered for use in determining potential species differences in biotransformation. The results from *in vitro* investigations may also have utility in the development of PBTK models (5).

**▼ M4**

60. *In vitro* dermal absorption studies may provide supplemental information to characterise absorption (6).
61. Primary cell cultures from liver cells and fresh tissue slices may be used to address similar questions as with liver microsomes. In certain cases, it may be possible to answer specific questions using cell lines with defined expression of the relevant enzyme or engineered cell lines. In certain cases, it may be useful to study the inhibition and induction of specific cytochrome P450 isozymes (e.g. CYP1A1, 2E1, 1A2, and others) and/or phase II enzymes by the parent compound using *in vitro* studies. Information obtained may have utility for similarly structured compounds.

**Use of Toxicokinetic Data from Toxicity Studies as Complementary Information**

62. Analysis of blood, tissue and/or excreta samples obtained during the conduct of any other toxicity studies can provide data on bioavailability, changes in plasma concentration in time (AUC,  $C_{max}$ ), bioaccumulation potential, clearance rates, and gender or life-stage changes in metabolism and kinetics.
63. Consideration of the study design can be used to answer questions relating to: saturation of absorption, biotransformation or excretion pathways at higher dose levels; the operation of new metabolic pathways at higher doses and the limitation of toxic metabolites to higher doses.
64. Other hazard assessment considerations could include issues such as:
- Age-related sensitivity due to differences in the status of the blood-brain barrier, the kidney and/or detoxification capacities;
  - Sub-population sensitivity due to differences in biotransformation capacities or other TK differences;
  - Extent of exposure of the foetus by transplacental transfer of chemicals or of the newborn through lactation.

**Use of Toxicokinetic Modelling**

65. Toxicokinetic models may have utility for various aspects of hazard and risk assessment as for example in the prediction of systemic exposure and internal tissue dose. Furthermore specific questions on mode of action may be addressed, and these models can provide a basis for extrapolation across species, routes of exposure, dosing patterns, and for human risk assessment. Data useful for developing PBTK models for a test chemical in any given species include (1) partition coefficients, (2) biochemical constants and physiological parameters, (3) route-specific absorption parameters and 4) *in vivo* kinetic data for model evaluation [e.g. clearance parameters for relevant (> 10 %) excretion pathways,  $K_m$  and  $V_{max}$  for metabolism]. The experimental data used in model development should be generated with scientifically sound methods and the model results validated. Test chemical- and species-specific parameters such as absorption rates, blood-tissue partitioning and metabolic rate constants are often determined to facilitate development of non-compartmental or physiologically-based models (7).

**▼ M4**

## DATA AND REPORTING

66. It is recommended that the study report include a table of contents.

**Body of the Report**

67. The body of the report should include information covered by this Test Method organised into sections and paragraphs as follows:

*Summary*

68. This section of the study report should include a summary of the study design and a description of methods used. It should also highlight the key findings regarding mass balance, the nature and magnitude of metabolites, tissue residue, rate of clearance, bioaccumulation potential, sex differences, etc. The summary should be presented in sufficient detail to permit evaluation of the findings.

*Introduction*

69. This section of the report should include the study objectives, rationale and design, as well as, appropriate references and any background history.

*Materials and Methods*

70. This section of the report should include detailed descriptions of all pertinent information including:

## (a) Test Chemical

This subsection should include identification of the test chemical: chemical name, molecular structure, qualitative and quantitative determination of its chemical composition, chemical purity and whenever possible, type and quantities of any impurities. It should also include information on physical/chemical properties including physical state, colour, gross solubility and/or partition coefficient, stability, and if appropriate, corrosivity. If applicable, information on isomers should be provided. If the test chemical is radiolabelled, information on the following should be included in this subsection: the type of radio-nuclide, position of label, specific activity, and radiochemical purity.

The type or description of any vehicle, diluents, suspending agents, and emulsifiers or other materials used in administering the test chemical should be stated.

## (b) Test Animals

This subsection should include information on the test animals, including selection and justification for species, strain, and age at study initiation, sex as well as body weight, health status, and animal husbandry.

## (c) Methods

This subsection should include details of the study design and methodology used. It should include a description of:

- (1) Justification for any modification of route of exposure and exposure conditions, if applicable;

**▼ M4**

- (2) Justification for selection of dose levels;
- (3) Description of pilot studies used in the experimental design of the follow-up studies, if applicable. Pilot study supporting data should be submitted;
- (4) How the dosing solution was prepared and the type of solvent or vehicle, if any, used;
- (5) Number of treatment groups and number of animals per group;
- (6) Dosage levels and volume (and specific activity of the dose when radioactivity is used);
- (7) Route(s) and methods of administration;
- (8) Frequency of dosing;
- (9) Fasting period (if used);
- (10) Total radioactivity per animal;
- (11) Animal handling;
- (12) Sample collection and handling;
- (13) Analytical methods used for separation, quantitation and identification of metabolites;
- (14) Limit of detection for the employed methods;
- (15) Other experimental measurements and procedures employed (including validation of methods for metabolite analysis).

**(d) Statistical Analysis**

If statistical analysis is used to analyse the study findings, then sufficient information on the method of analysis and the computer program employed should be included, so that an independent reviewer/statistician can re-evaluate and reconstruct the analysis.

In the case of systems modelling studies such as PBTK, presentation of models should include a full description of the model to allow independent reconstruction and validation of the model (see paragraph 65 and Appendix: Definitions).

*Results*

71. All data should be summarised and tabulated with appropriate statistical evaluation and described in the text of this section. Radioactivity counting data should be summarised and presented as appropriate for the study, typically as microgram or milligram equivalents per mass of sample, although other units may be used. This section should include graphic illustrations of the findings, reproduction of representative chromatographic and spectrometric data, metabolite identification/quantification and proposed metabolic pathways including molecular structure of metabolites. In addition the following information is to be included in this section, if applicable:

- (1) Quantity and percent recovery of radioactivity in urine, faeces, expired air, and urine and faeces cage wash.
  - For dermal studies, also include data on test chemical recovery from treated skin, skin washes, and residual radioactivity in the skin covering apparatus and metabolic unit as well as results of the dermal washing study. For further discussion, see paragraphs 74-77.

**▼M4**

- For inhalation studies, also include data on recovery of test chemical from lungs and nasal tissues (8). For further discussion, see paragraph 78.
- (2) Tissue distribution reported as percent of administered dose and concentration (microgram equivalents per gram of tissue), and tissue-to-blood or tissue-to-plasma ratios;
- (3) Material balance developed from each study involving the assay of body tissues and excreta;
- (4) Plasma concentrations and toxicokinetic parameters (bioavailability, AUC, C<sub>max</sub>, T<sub>max</sub>, clearance, half-life) after administration by the relevant route(s) of exposure;
- (5) Rate and extent of absorption of the test chemical after administration by the relevant route(s) of exposure;
- (6) Quantities of the test chemical and metabolites (reported as percent of the administered dose) collected in excreta;
- (7) Reference to appendix data which contain individual animal data for all measurement endpoints (e.g. dose administration, percent recovery, concentrations, TK parameters, etc.);
- (8) A figure with the proposed metabolic pathways and the molecular structures of the metabolites.

*Discussion and Conclusions*

72. In this section the author(s) should:

- (1) Provide a proposed metabolic pathway based on the results of the metabolism and disposition of the test chemical;
- (2) Discuss any potential species and sex differences regarding the disposition and/or biotransformation of the test chemical;
- (3) Tabulate and discuss the identification and magnitude of metabolites, rates of clearance, bioaccumulation potential, and level of tissue residues of parent, and/or metabolite(s), as well as possible dose-dependent changes in TK parameters, as appropriate;
- (4) Integrate into this section any relevant TK data obtained in the course of conducting toxicity studies;
- (5) Provide a concise conclusion that can be supported by the findings of the study;
- (6) Add Sections (as needed or appropriate).

73. Additional sections should be used to include supporting bibliographic information, tables, figures, appendices, etc.



**▼ M4**

## ALTERNATIVE ROUTES OF EXPOSURE

**Dermal***Dermal Treatment*

74. This section provides specific information on the investigation of the toxicokinetics of the test chemical by the dermal route. For dermal absorption, chapter B.44 of this Annex [Skin absorption: *in vivo* method (9)] should be consulted. For other endpoints such as distribution and metabolism, this Test Method B.36 can be used. One or more dose levels for the test chemical should be used in the dermal treatment. The test chemical (e.g. neat, diluted or formulated material containing the test chemical which is applied to the skin) should be the same (or a realistic surrogate) as that to which humans or other potential target species might be exposed. The dose level(s) should be selected in accordance with paragraphs 20-26 of this Test Method. Factors that could be taken into consideration in dermal dose selection include expected human exposure and/or doses at which toxicity was observed in other dermal toxicity studies. The dermal dose(s) should be dissolved, if necessary, in a suitable vehicle and applied in a volume adequate to deliver the doses. Shortly before testing, fur should be clipped from the dorsal area of the trunk of the test animals. Shaving may be employed, but it should be carried out approximately 24 h before the test. When clipping or shaving the fur, care should be taken to avoid abrading the skin, which could alter its permeability. Approximately 10 % of the body surface should be cleared for application of the test chemical. With highly toxic chemicals, the surface area covered may be less than approximately 10 %, but as much of the area as possible is to be covered with a thin and uniform film. The same treatment surface area should be used for all dermal test groups. The dosed areas are to be protected with a suitable covering which is secured in place. The animals should be housed separately.
75. A dermal washing study should be conducted to assess the amount of the applied dose of the test chemical that may be removed from the skin by washing the treated skin area with a mild soap and water. This study can also aid in establishing mass balance when the test chemical is administered by the dermal route. For this dermal washing study, a single dose of the test chemical should be applied to two animals. Dose level selection is in accordance with paragraph 23 of this Test Method (also see paragraph 76 for discussion of skin contact time). The amounts of test chemical recovered in the washes should be determined to assess the effectiveness of removal of the test chemical by the washing procedure.
76. Unless precluded by corrosiveness, the test chemical should be applied and kept on the skin for a minimum of 6 h. At the time of removal of the covering, the treated area should be washed following the procedure as outlined in the dermal washing study (see paragraph 75). Both covering and the washes should be analysed for residual test chemical. At the termination of the studies, each animal should be humanely killed in accordance with (2), and the treated skin removed. An appropriate section of treated skin should be analysed to determine residual test chemical (radioactivity).
77. For the toxicokinetic assessment of pharmaceuticals, different procedures, in accordance with the appropriate regulatory system, may be needed.

**▼ M4****Inhalation**

78. A single concentration (or more if needed) of test chemical should be used. The concentration(s) should be selected in accordance with paragraphs 20-26 of this Test Method. Inhalation treatments are to be conducted using a 'nose-cone' or 'head-only' apparatus to prevent absorption by alternate routes of exposure (8). If other inhalation exposure conditions are used, justification for the modification should be documented. The duration of exposure by inhalation should be defined; a typical exposure is 4-6 h.

*LITERATURE:*

- (1) OECD (2009). Preliminary Review of OECD Test Guidelines for their Applicability to Manufactured Nanomaterials, Series on the Safety of Manufactured Nanomaterials No 15, ENV/JM/MONO(2009)21, OECD, Paris.
- (2) OECD (2000). Guidance Document on Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation; Environmental Health and Safety Publications, Series on Testing and Assessment No 19, ENV/JM/MONO(2000), OECD, Paris.
- (3) Solon E G, Kraus L (2002). Quantitative whole-body autoradiography in the pharmaceutical industry; Survey results on study design, methods, and regulatory compliance, *J Pharm and Tox Methods* 46: 73-81.
- (4) Stumpf WE (2005). Drug localization and targeting with receptor microscopic autoradiography. *J. Pharmacological and Toxicological Methods* 51: 25-40.
- (5) Loizou G, Spendiff M, Barton HA, Bessems J, Bois FY, d'Yvoire MB, Buist H, Clewell HJ 3rd, Meek B, Gundert-Remy U, Goerlitz G, Schmitt W. (2008). Development of good modelling practice for physiologically based pharmacokinetic models for use in risk assessment: The first steps. *Regulatory Toxicology and Pharmacology* 50: 400 – 411.
- (6) Chapter B.45 of this Annex, Skin Absorption: In Vitro Method.
- (7) IPCS (2010). Characterization and application of Physiologically-Based Pharmacokinetic Models in Risk Assessment. IPCS Harmonization Project Document No 9. Geneva, World Health Organization, International Programme on Chemical Safety.
- (8) OECD (2009). Guidance Document on Acute Inhalation Toxicity Testing, Series on Testing and Assessment No 39, ENV/JM/MONO(2009)28, OECD, Paris.
- (9) Chapter B.44 of this Annex, Skin Absorption: In Vivo Method.
- (10) Barton HA, *et al.* (2006). The Acquisition and Application of Absorption, Distribution, Metabolism, and Excretion (ADME) Data in Agricultural Chemical Safety Assessments, *Critical Reviews in Toxicology* 36: 9-35.
- (11) Gibaldi M and Perrier D, (1982), *Pharmacokinetics*, 2nd edition, Marcel Dekker, Inc., New York.

▼ **M4***Appendix*

## DEFINITIONS

**Absorption:** Process(es) of uptake of chemicals into or across tissues. Absorption refers to parent compound and all its metabolites. Not to be confused with 'bioavailability'.

**Accumulation (Bioaccumulation):** Increase of the amount of a test chemical over time within tissues (usually fatty tissues, following repeated exposure); if the input of a test chemical into the body is greater than the rate at which it is eliminated, the organism accumulates the test chemical and toxic concentrations of a test chemical might be achieved.

**ADME:** Acronym for 'Absorption, Distribution, Metabolism, and Excretion'.

**AUC:** (Area under the plasma concentration-time curve): Area under the curve in a plot of concentration of test chemical in plasma over time. It represents the total amount of test chemical absorbed by the body within a predetermined period of time. Under linear conditions, the AUC (from time zero to infinity) is proportional to the total amount of a test chemical absorbed by the body, irrespective of the rate of absorption.

**Autoradiography:** (Whole-body autoradiography): Used to determine qualitatively and/or quantitatively the tissue localisation of a radioactive test chemical, this technique uses X-ray film or more recently digital phosphor-imaging to visualize radioactively labelled molecules or fragments of molecules by recording the radiation emitted within the object under study. Quantitative whole-body autoradiography, compared to organ dissection, may have some advantages for the evaluation of test chemical distribution and the assessment of overall recovery and resolution of radioactive material in tissues. One significant advantage, for example, is it can be used in a pigmented animal model to assess possible association of the test chemical with melanin, which can bind certain molecules. However, while it may provide convenient whole body overviews of the high-capacity-low-affinity binding sites, this technique might be limited in recognising specific target sites such as receptor-binding sites where relatively high-resolution and high-sensitivity are needed for detection. When autoradiography is used, experiments intended to determine mass balance of administered compound should be conducted as a separate group or in a separate study from the tissue distribution experiment, where all excreta (which may also include expired air) and whole carcasses are homogenised and assayed by liquid scintillation counting.

**Biliary excretion:** Excretion via the bile ducts.

**Bioaccumulation:** See 'Accumulation'.

**Bioavailability:** Fraction of an administered dose that reaches the systemic circulation or is made available at the site of physiological activity. Usually, bioavailability of a test chemical refers to the parent compound, but it could refer to its metabolite. It considers only one chemical form. *Nota Bene:* bioavailability and absorption are not the same. The difference between e.g. oral absorption (i.e. presence in gut wall and portal circulation) and bioavailability (i.e. presence in systemic blood and in tissues) can arise from chemical degradation due to gut wall metabolism or efflux transport back to the intestinal lumen or presystemic metabolism in the liver, among other factors (10). Bioavailability of the toxic component (parent compound or a metabolite) is a critical parameter in human risk assessment (high-to-low dose extrapolation, route-to-route extrapolation) for derivation of an internal value from the external NOAEL or BMD (applied dose). For liver effects upon oral administration, it is the oral absorption that suffices. However, for every effect other than at the portal of entry, it is the bioavailability that is in general a more reliable parameter for further use in risk assessment, not the absorption.

**▼ M4**

**Biopersistence:** See 'Persistence'.

**Biotransformation:** (Usually enzymatic) chemical conversion of a test chemical of interest into a different chemical within the body. Synonymous with 'metabolism'.

**C<sub>max</sub>:** Either maximal (peak) concentration in blood (plasma/serum) after administration or maximal (peak) excretion (in urine or faeces) after administration.

**Clearance rate:** Quantitative measure of the rate at which a test chemical is removed from the blood, plasma or a certain tissue per unit time.

**Compartment:** Structural or biochemical portion (or unit) of a body, tissue or cell, that is separate from the rest.

**Detoxification pathways:** Series of steps leading to the elimination of toxic chemicals from the body, either by metabolic change or excretion.

**Distribution:** Dispersal of a test chemical and its derivatives throughout an organism.

**Enzymes/Isozymes:** Proteins that catalyse chemical reactions. Isozymes are enzymes that catalyse similar chemical reactions but differ in their amino acid sequence.

**Enzymatic Parameters:** K<sub>m</sub>: Michaelis constant and V<sub>max</sub>: maximum velocity.

**Excretion:** Process(es) by which an administered test chemical and/or its metabolites are removed from the body.

**Exogenously:** Introduced from or produced outside the organism or system.

**Extrapolation:** Inference of one or more unknown values on the basis of that which is known or has been observed.

**Half-life (t<sub>1/2</sub>):** The time taken for the concentration of the test chemical to decrease by one-half in a compartment. It typically refers to plasma concentration or the amount of the test chemical in the whole body.

**Induction/Enzyme induction:** Enzyme synthesis in response to an environmental stimulus or inducer molecule.

**Linearity/linear kinetics:** A process is linear in terms of kinetics when all transfer rates between compartments are proportional to the amounts or concentrations present, i.e. first order. Consequently, clearance and distribution volumes are constant, as well as half-lives. The concentrations achieved are proportional to the dosing rate (exposure), and accumulation is more easily predictable. Linearity/Non-linearity can be assessed by comparing the relevant parameters, e.g. AUC, after different doses or after single and repeated exposure. Lack of dose dependency may be indicative of saturation of enzymes involved in the metabolism of the compound, an increase of AUC after repeated exposure as compared to single exposure may be an indication for inhibition of metabolism and a decrease in AUC may be an indication for induction of metabolism [see also (11)].

**Mass balance:** Accounting of test chemical entering and leaving the system.

**▼ M4**

**Material balance:** See ‘mass balance’.

**Mechanism (Mode) of toxicity/Mechanism (Mode) of action:** Mechanism of action refers to specific biochemical interactions through which a test chemical produces its effect. Mode of action refers to more general pathways leading to the toxicity of a test chemical.

**Metabolism:** Synonymous with ‘biotransformation’.

**Metabolites:** Products of metabolism or metabolic processes.

**Oral Absorption:** The percentage of the dose of test chemical absorbed from the site of administration (i.e. GI tract). This critical parameter can be used to understand the fraction of the administered test chemical that reaches the portal vein, and subsequently the liver.

**Partition coefficient:** Also known as the distribution coefficient, it is a measure of the differential solubility of a chemical in two solvents.

**Peak blood (plasma/serum) levels:** Maximal (peak) blood (plasma/serum) concentration after administration (see also ‘C<sub>max</sub>’).

**Persistence (biopersistence):** Long-term presence of a chemical (in a biological system) due to resistance to degradation/elimination.

**Read-across:** The endpoint information for one or more chemicals is used to make a prediction of the endpoint for the target chemical.

**Receptor Microscopic Autoradiography** (or Receptor Microautoradiography): This technique may be used to probe xenobiotic interaction with specific tissue sites or cell populations as for instance in receptor binding or specific mode of action studies that may require high-resolution and high sensitivity which may not be feasible with other techniques such as whole-body autoradiography.

**Route of administration** (oral, IV, dermal, inhalation, etc.): Refers to the means by which chemicals are administered to the body (e.g. orally by gavage, orally by diet, dermal, by inhalation, intravenously, etc.).

**Saturation:** State whereby one or more of the kinetic (e.g. absorption, metabolism or clearance) process(es) are at a maximum (read ‘saturated’).

**Sensitivity:** Capability of a method or instrument to discriminate between measurement responses representing different levels of a variable of interest.

**Steady-state blood (plasma) levels:** Non-equilibrium state of an open system in which all forces acting on the system are exactly counter-balanced by opposing forces, in such a manner that all its components are stationary in concentration although matter is flowing through the system.

**Systems Modelling** (Physiologically-based Toxicokinetic, Pharmacokinetic-based, Physiologically-based Pharmacokinetic, Biologically-based, etc.): Abstract model that uses mathematical language to describe the behaviour of a system.

**Target tissue:** Tissue in which a principal adverse effect of a toxicant is manifested.

**▼ M4**

**Test chemical:** Any chemical or mixture tested using this Test Method.

**Tissue distribution:** Reversible movement of a test chemical from one location in the body to another. Tissue distribution can be studied by organ dissection, homogenisation, combustion and liquid scintillation counting or by qualitative and/or quantitative whole body autoradiography. The former is useful to obtain concentration and percent of recovery from tissues and remaining carcass of the same animals, but may lack resolution for all tissues and may have less than ideal overall recovery (< 90 %). See definition for the latter above.

**T<sub>max</sub>:** Time to reach C<sub>max</sub>.

**Toxicokinetics (Pharmacokinetics):** Study of the absorption, distribution, metabolism, and excretion of chemicals over time.

**Validation of models:** Process of assessing the adequacy of a model to consistently describe the available toxicokinetic data. Models may be evaluated via statistical and visual comparison of model predictions with experimental values against a common independent variable (e.g. time). The extent of evaluation should be justified in relation to the intended use of the model.

**▼B****B.37. DELAYED NEUROTOXICITY OF ORGANOPHOSPHORUS SUBSTANCES FOLLOWING ACUTE EXPOSURE****1. METHOD****1.1. INTRODUCTION**

In the assessment and evaluation of the toxic effects of substances, it is important to consider the potential of certain classes of substances to cause specific types of neurotoxicity that might not be detected in other toxicity studies. Certain organophosphorus substances have been observed to cause delayed neurotoxicity and should be considered as candidates for evaluation.

*In vitro* screening tests could be employed to identify those substances which may cause delayed polyneuropathy; however, negative findings from *in vitro* studies do not provide evidence that the test substance is not a neurotoxicant.

See General introduction Part B.

**1.2. DEFINITIONS**

Organophosphorus substances include uncharged organophosphorus esters, thioesters or anhydrides of organophosphoric, organophosphonic or organophosphoramidic acids or of related phosphorothioic, phosphonothioic or phosphorothioamidic acids, or other substances that may cause the delayed neurotoxicity sometimes seen in this class of substances.

Delayed neurotoxicity is a syndrome associated with prolonged delayed onset of ataxia, distal axonopathies in spinal cord and peripheral nerve, and inhibition and aging of neuropathy target esterase (NTE) in neural tissue.

**1.3. REFERENCE SUBSTANCES**

A reference substance may be tested with a positive control group as a means of demonstrating that under the laboratory test conditions, the response of the tested species has not changed significantly.

An example of a widely used neurotoxicant is tri-*o*-tolyl phosphate (CAS 78-30-8, Eines 201-103-5, CAS nomenclature: phosphoric acid, tris(2-methylphenyl)ester), also known as tris-*o*-cresylphosphate.

**1.4. PRINCIPLE OF THE TEST METHOD**

The test substance is administered orally in a single dose to domestic hens which have been protected from acute cholinergic effects, when appropriate. The animals are observed for 21 days for behavioural abnormalities, ataxia, and paralysis. Biochemical measurements, in particular neuropathy target esterase inhibition (NTE), are undertaken on hens randomly selected from each group, normally 24 and 48 hours after dosing. Twenty-one days after exposure, the remainder of the hens are killed and histopathological examination of selected neural tissues is undertaken.

**▼B**

## 1.5. DESCRIPTION OF THE TEST METHOD

1.5.1. **Preparations**

Healthy young adult hens free from interfering viral diseases and medication and without abnormalities of gait should be randomised and assigned to treatment and control groups and acclimatised to the laboratory conditions for at least five days prior to the start of the study.

Cages or enclosures which are large enough to permit free mobility of the hens, and easy observation of gait should be used.

Dosing with the test substance should normally be by the oral route using gavage, gelatine capsules, or a comparable method. Liquids may be given undiluted or dissolved in an appropriate vehicle such as corn oil; solids should be dissolved if possible since large doses of solids in gelatine capsules may not be absorbed efficiently. For non-aqueous vehicles the toxic characteristics of the vehicle should be known, and if not known should be determined before the test.

1.5.2. **Test conditions**1.5.2.1. *Test animals*

The young adult domestic laying hen (*Gallus gallus domesticus*), aged eight to 12 months, is recommended. Standard size breeds and strains should be employed and the hens normally should have been reared under conditions which permitted free mobility.

1.5.2.2. *Number and sex*

In addition to the treatment group, both a vehicle control group and a positive control group should be used. The vehicle control group should be treated in a manner identical to the treatment group, except that administration of the test substance is omitted.

Sufficient number of hens should be utilised in each group of birds so that at least six birds can be killed for biochemical determination (three at each of two time points) and six can survive the 21-day observation period for pathology.

The positive control group may be run concurrently or be a recent historical control group. It should contain at least six hens, treated with a known delayed neurotoxicant, three hens for biochemistry and three hens for pathology. Periodic updating of historical data is recommended. New positive control data should be developed when some essential element (e.g. strain, feed, housing conditions) of the conduct of the test has been changed by the performing laboratory.



**▼B**1.5.2.3. *Dose levels*

A preliminary study using an appropriate number of hens and dose levels groups should be performed to establish the level to be used in the main study. Some lethality is typically necessary in this preliminary study to define an adequate main study dose. However, to prevent death due to acute cholinergic effects, atropine or another protective agent, known to not interfere with delayed neurotoxic responses, may be used. A variety of test methods may be used to estimate the maximum non-lethal dose of test substances (See method B.1*bis*). Historical data in the hen or other toxicological information may also be helpful in dose selection.

The dose level of the test substance in the main study should be as high as possible taking into account the results of the preliminary dose selection study and the upper limit dose of 2 000 mg/kg body weight. Any mortality which might occur should not interfere with the survival of sufficient animals for biochemistry (six) and histology (six) at 21 days. Atropine or another protective agent, known to not interfere with delayed neurotoxic responses, should be used to prevent death due to acute cholinergic effects.

1.5.2.4. *Limit test*

If a test at a dose level of at least 2 000 mg/kg body weight/day, using the procedures described for this study, produces no observable toxic effects and if toxicity would not be expected based upon data from structurally related substances, then a study using a higher dose may not be considered necessary. The limit test applies except when human exposure indicates the need for a higher dose level to be used.

1.5.3. **Observation period**

Observation period should be 21 days.

1.5.4. **Procedure**

After administration of a protective agent to prevent death due to acute cholinergic effect, the test substance is administered in a single dose.

**General observation**

Observations should start immediately after exposure. All hens should be carefully observed several times during the first two days and thereafter at least once daily for a period of 21 days or until scheduled kill. All signs of toxicity should be recorded, including the time of onset, type, severity and duration of behavioural abnormalities. Ataxia should be measured on an ordinal grading scale consisting of at least four levels, and paralysis should be noted. At least twice a week the hens selected for pathology should be taken outside the cages and subjected to a period of forced motor activity, such as ladder climbing, in order to facilitate the observation of minimal toxic effects. Moribund animals and animals in severe distress or pain should be removed when noticed, humanely killed and necropsied.

**▼ B****Body weight**

All hens should be weighed just prior to administration of the test substance and at least once a week thereafter.

**Biochemistry**

Six hens randomly selected from each of the treatment and vehicle control groups, and three hens from the positive control group (when this group is run concurrently), should be killed within a few days after dosing, and the brain and lumbar spinal cord prepared and assayed for neuropathy target esterase inhibition activity. In addition, it may also be useful to prepare and assay sciatic nerve tissue for neuropathy target esterase inhibition activity. Normally, three birds of the control and each treatment group are killed after 24 hours and three at 48 hours, whereas the three hens of the positive controls should be killed at 24 hours. If observation of clinical signs of intoxication (this can often be assessed by observation of the time of onset of cholinergic signs) indicates that the toxic agent may be disposed of very slowly then it may be preferable to sample tissue from three birds at each of two times between 24 and as late as 72 hours after dosing.

Analyses of acetylcholinesterase (AChE) may also be performed on these samples, if deemed appropriate. However, spontaneous reactivation of AChE may occur *in vivo*, and so lead to underestimation of the potency of the substance as an AChE inhibitor.

**Gross necropsy**

Gross necropsy of all animals (scheduled killed and killed when moribund) should include observation of the appearance of the brain and spinal cord.

**Histopathological examination**

Neural tissue from animals surviving the observation period and not used for biochemical studies should be subjected to microscopic examination. Tissues should be fixed *in situ*, using perfusion techniques. Sections should include cerebellum (mid-longitudinal level), medulla oblongata, spinal cord, and peripheral nerves. The spinal cord sections should be taken from the upper cervical segment, the mid-thoracic and the lumbo-sacral regions. Sections of the distal region of the tibial nerve and its branches to the gastrocnemial muscle and of the sciatic nerve should be taken. Sections should be stained with appropriate myelin and axon-specific stains.

**2. DATA**

Negative results on the endpoints selected in this method (biochemistry, histopathology and behavioural observation) would not normally require further testing for delayed neurotoxicity. Equivocal or inconclusive results for these endpoints may require further evaluation.

Individual data should be provided. Additionally, all data should be summarised in tabular form, showing for each test group the number of animals at the start of the test, the number of animals showing lesions, behavioural or biochemical effects, the types and severity of these lesions or effects, and the percentage of animals displaying each type and severity of lesion or effect.

**▼B**

The findings of this study should be evaluated in terms of the incidence, severity, and correlation of behavioural, biochemical and histopathological effects and any other observed effects in the treated and control groups.

Numerical results should be evaluated by appropriate and generally acceptable statistical methods. The statistical methods used should be selected during the design of the study.

**3. REPORTING****TEST REPORT**

The test report shall, if possible, include the following information:

**3.1. Test animals:**

- strain used,
- number and age of animals,
- source, housing conditions, etc.,
- individual weights of animals at the start of the test.

**3.2. Test conditions:**

- details of test substance preparation, stability and homogeneity, where appropriate,
- justification for choice of vehicle,
- details of the administration of the test substance,
- details of food and water quality,
- rationale for dose selection,
- specification of doses administered, including details of the vehicle, volume and physical form of the material administered,
- identity and details of the administration of any protective agent.

**3.3. Results:**

- body weight data,
- toxic response data by group, including mortality,
- nature, severity and duration of clinic observations (whether reversible or not),
- a detailed description of biochemical methods and findings,
- necropsy findings,
- a detailed description of all histopathological findings,
- statistical treatment of results, where appropriate.

Discussion of results.

Conclusions.

**4. REFERENCES**

This method is analogous to OECD TG 418.

**▼B****B.38. DELAYED NEUROTOXICITY OF ORGANOPHOSPHORUS SUBSTANCES 28-DAY REPEATED DOSE STUDY****1. METHOD****1.1. INTRODUCTION**

In the assessment and evaluation of the toxic effects of substances, it is important to consider the potential of certain classes of substances to cause specific types of neurotoxicity that might not be detected in other toxicity studies. Certain organophosphorus substances have been observed to cause delayed neurotoxicity and should be considered as candidates for evaluation.

*In vitro* screening tests could be employed to identify those substances which may cause delayed polyneuropathy; however, negative findings from *in vitro* studies do not provide evidence that the test substance is not a neurotoxicant.

This 28-day delayed neurotoxicity test provides information on possible health hazards likely to arise from repeated exposures over a limited period of time. It will provide information on dose response and can provide an estimate of a no-observed-adverse effect level, which can be of use for establishing safety criteria for exposure.

See also General introduction Part B.

**1.2. DEFINITIONS**

Organophosphorus substances include uncharged organophosphorus esters, thioesters or anhydrides of organophosphoric, organophosphonic or organophosphoramidic acids or of related phosphorothioic, phosphonothioic or phosphorothioamidic acids or other substances that may cause the delayed neurotoxicity sometimes seen in this class of substances.

Delayed neurotoxicity is a syndrome associated with prolonged delayed onset of ataxia, distal axonopathies in spinal cord and peripheral nerve, and inhibition and ageing of neuropathy target esterase (NTE) in neural tissue.

**1.3. PRINCIPLE OF THE TEST METHOD**

Daily doses of the test substance are administered orally to domestic hens for 28 days. The animals are observed at least daily for behavioural abnormalities, ataxia and paralysis until 14 days after the last dose. Biochemical measurements, in particular neuropathy target esterase inhibition (NTE), are undertaken, on hens randomly selected from each group, normally 24 and 48 hours after the last dose. Two weeks after the last dose, the remainder of the hens are killed and histopathological examination of selected neural tissues is undertaken.

**▼ B**

## 1.4. DESCRIPTION OF THE TEST METHOD

1.4.1. **Preparations**

Healthy young adult hens free from interfering viral diseases and medication, and without abnormalities of gait should be randomised and assigned to treatment and control groups and acclimatised to the laboratory conditions for at least five days prior to the start of the study.

Cages or enclosures which are large enough to permit free mobility of the hens and easy observation of gait should be used.

Oral dosing each day, seven days per week, should be carried out, preferably by gavage or administration of gelatine capsules. Liquids may be given undiluted or dissolved in an appropriate vehicle such as corn oil; solids should be dissolved if possible since large doses of solids in gelatine capsules may not be absorbed efficiently. For non-aqueous vehicles the toxic characteristics of the vehicle should be known, and if not known should be determined before the test.

1.4.2. **Test conditions**1.4.2.1. *Test animals*

The young adult domestic laying hen (*Gallus gallus domesticus*), aged eight to 12 months, is recommended. Standard size, breeds and strains should be employed and the hens normally should have been reared under conditions which permitted free mobility.

1.4.2.2. *Number and sex*

Generally at least three treatment groups and a vehicle control group should be used. The vehicle control group should be treated in a manner identical to the treatment group, except that administration of the test substance is omitted.

Sufficient number of hens should be utilised in each group of birds so that at least six birds can be killed for biochemical determinations (three at each of two timepoints) and six birds can survive the 14-day post-treatment observation period for pathology.

1.4.2.3. *Dose levels*

Dose levels should be selected taking into account the results from an acute test on delayed neurotoxicity and any other existing toxicity or kinetic data available for the test compound. The highest dose level should be chosen with the aim of inducing toxic effects, preferably delayed neurotoxicity, but not death nor obvious suffering. Thereafter, a descending sequence of dose levels should be selected with a view to demonstrate any dose-related response and no-observed-adverse effects at the lowest dose level.

**▼ B**1.4.2.4. *Limit test*

If a test at a dose level of at least 1 000 mg/kg body weight/day, using the procedures described for this study, produces no observable toxic effects and if toxicity would not be expected based upon data from structurally related substances, then a study using a higher dose may not be considered necessary. The limit test applies except when expected human exposure indicates the need for a higher dose level to be used.

1.4.2.5. *Observation period*

All the animals should be observed at least daily during the exposure period and 14 days after, unless scheduled necropsy.

1.4.3. **Procedure**

Animals are dosed with the test substance on seven days per week for a period of 28 days.

*General observations*

Observations should start immediately after treatment begins. All hens should be carefully observed at least once daily on each of the 28 days of treatment, and for 14 days after dosing or until scheduled kill. All signs of toxicity should be recorded including their time of onset, type, severity and duration. Observations should include, but not be limited to, behavioural abnormalities. Ataxia should be measured on an ordinal grading scale consisting of at least four levels, and paralysis should be noted. At least twice a week the hens should be taken outside the cages and subjected to a period of forced motor activity, such as ladder climbing, in order to facilitate the observation of minimal toxic effects. Moribund animals in severe distress or pain should be removed when noticed, humanely killed and necropsied.

*Body weight*

All hens should be weighed just prior to the first administration of the test substance and at least once a week thereafter.

*Biochemistry*

Six hens randomly selected from each of the treatment and vehicle control groups should be killed within a few days after the last dose, and the brain and lumbar spinal cord prepared and assayed for neuropathy target esterase (NTE) inhibition activity. In addition, it may also be useful to prepare and assay sciatic nerve tissue for neuropathy target esterase (NTE) inhibition activity. Normally, three birds of the control and each treatment group are killed after 24 hours and three at 48 hours after the last dose. If data from the acute study or other studies (e.g. toxicokinetics) indicate that other times of killing after final dosing are preferable then these times should be used and the rationale documented.

Analyses of acetylcholinesterase (AChE) may also be performed on these samples, if deemed appropriate. However, spontaneous reactivation of AChE may occur *in vivo*, and so lead to underestimation of the potency of the substance as an AChE inhibitor.

**▼ B***Gross necropsy*

Gross necropsy of all animals (scheduled killed and killed when moribund) should include observation of the appearance of the brain and spinal cord.

*Histopathological examination*

Neural tissue from animals surviving the observation period and not used for biochemical studies should be subjected to microscopic examination. Tissues should be fixed *in situ*, using perfusion techniques. Sections should include cerebellum (mid longitudinal level), medulla oblongata, spinal cord and peripheral nerves. The spinal cord sections should be taken from the upper cervical segment, the mid-thoracic and the lumbo-sacral regions. Sections of the distal region of the tibial nerve and its branches to the gastrocnemial muscle and of the sciatic nerve should be taken. Sections should be stained with appropriate myelin and axon-specific stains. Initially, microscopic examination should be carried out on the preserved tissues of all animals in the control and high dose group. When there is evidence of effects in the high dose group, microscopic examination should also be carried out in hens from the intermediate and low dose groups.

**2. DATA**

Negative results on the endpoints selected in this method (biochemistry, histopathology and behavioural observation) would not normally require further testing for delayed neurotoxicity. Equivocal or inconclusive results for these endpoints may require further evaluation.

Individual data should be provided. Additionally, all data should be summarised in tabular form, showing for each test group the number of animals at the start of the test, the number of animals showing lesions, behavioural or biochemical effects, the types and severity of these lesions or effects, and the percentage of animals displaying each type and severity of lesion or effect.

The findings of this study should be evaluated in terms of the incidence, severity, and correlation of behavioural, biochemical and histopathological effects and any other observed effects in each of the treated and control groups.

Numerical results should be evaluated by appropriate and generally acceptable statistical methods. The statistical methods should be selected during the design of the study.

**3. REPORTING****TEST REPORT**

The test report shall, if possible, include the following information:

**3.1. Test animals:**

- strain used,
- number and age of animals,
- source, housing conditions, etc.,
- individual weights of animals at the start of the test.

**▼ B**

- 3.2. Test conditions:
- details of test substance preparation, stability and homogeneity, where appropriate,
  - justification for choice of vehicle,
  - details of the administration of the test substance,
  - details of food and water quality,
  - rationale for dose selection,
  - specification of doses administered, including details of the vehicle, volume and physical form of the material administered,
  - rationale for choosing other times for biochemical determination, if other than 24 and 48 h.

- 3.3. Results:
- body weight data,
  - toxic response data by dose level, including mortality,
  - no-observed adverse effect level,
  - nature, severity and duration of clinic observations (whether reversible or not),
  - a detailed description of biochemical methods and findings,
  - necropsy findings,
  - a detailed description of all histopathological findings,
  - statistical treatment of results, where appropriate.

Discussion of results.

Conclusions.

4. **REFERENCES**

This method is analogous to OECD TG 419.



**▼B****B.39. UNSCHEDULED DNA SYNTHESIS (UDS) TEST WITH MAMMALIAN LIVER CELLS IN VIVO****1. METHOD**

This method is a replicate of the OECD TG 486, Unscheduled DNA Synthesis (UDS) Test with Mammalian Liver Cells *In Vivo* (1997).

**1.1. INTRODUCTION**

The purpose of the unscheduled DNA Synthesis (UDS) test with mammalian liver cells *in vivo* is to identify test substances that induce DNA repair in liver cells of treated animals (see 1,2,3,4).

This *in vivo* test provides a method for investigating genotoxic effects of chemicals in the liver. The end-point measured is indicative of DNA damage and subsequent repair in liver cells. The liver is usually the major site of metabolism of absorbed compounds. It is thus an appropriate site to measure DNA damage *in vivo*.

If there is evidence that the test substance will not reach the target tissue, it is not appropriate to use this test.

The end-point of unscheduled DNA synthesis (UDS) is measured by determining the uptake of labelled nucleosides in cells that are not undergoing scheduled (S-phase) DNA synthesis. The most widely used technique is the determination of the uptake of tritium-labelled thymidine (<sup>3</sup>H-TdR) by autoradiography. Rat livers are preferably used for *in vivo* UDS tests. Tissues other than the livers may be used, but are not the subject of this method.

The detection of a UDS response is dependent on the number of DNA bases excised and replaced at the site of the damage. Therefore, the UDS test is particularly valuable to detect substance-induced 'longpatch repair' (20-30 bases). In contrast, 'shortpatch repair' (1-3 bases) is detected with much lower sensitivity. Furthermore, mutagenic events may result because of non-repair, misrepair or misreplication of DNA lesions. The extent of the UDS response gives no indication of the fidelity of the repair process. In addition, it is possible that a mutagen reacts with DNA but the DNA damage is not repaired via an excision repair process. The lack of specific information on mutagenic activity provided by the UDS test is compensated for by the potential sensitivity of this endpoint because it is measured in the whole genome.

See also General introduction Part B.

**1.2. DEFINITIONS**

**Cells in repair:** a net nuclear grain (NNG) higher than a preset value, to be justified at the laboratory conducting the test.

**Net nuclear grains (NNG):** quantitative measure for UDS activity of cells in autoradiographic UDS tests, calculated by subtracting the average number of cytoplasmic grains in nucleus-equivalent cytoplasmic areas (CG) from the number of nuclear grains (NG):  $NNG = NG - CG$ . NNG counts are calculated for individual cells and then pooled for cells in a culture, in parallel cultures, etc.

**▼ B**

**Unscheduled DNA Synthesis (UDS):** DNA repair synthesis after excision and removal of a stretch of DNA containing a region of damage induced by chemical substances or physical agents.

**1.3. PRINCIPLE OF THE TEST METHOD**

The UDS test with mammalian liver cells *in vivo* indicates DNA repair synthesis after excision and removal of a stretch of DNA containing a region of damage induced by chemical substances or physical agents. The test is usually based on the incorporation of <sup>3</sup>H-TdR into the DNA of liver cells which have a low frequency of cells in the S-phase of the cell cycle. The uptake of <sup>3</sup>H-TdR is usually determined by autoradiography, since this technique is not as susceptible to interference from S-phase cells as, for example, liquid scintillation counting.

**1.4. DESCRIPTION OF THE METHOD****1.4.1. Preparations****1.4.1.1. Selection of animal species**

Rats are commonly used, although any appropriate mammalian species may be used. Commonly used laboratory strains of young healthy adult animals should be employed. At the commencement of the study the weight variation of animals should be minimal and not exceed  $\pm 20\%$  of the mean weight for each sex.

**1.4.1.2. Housing and feeding conditions**

General conditions referred in the General introduction to Part B are applied although the aim for humidity should be 50-60 %.

**1.4.1.3. Preparation of the animals**

Healthy young adult animals are randomly assigned to the control and treatment groups. Cages should be arranged in such a way that possible effects due to cage placement are minimised. The animals are identified uniquely and kept in their cages for at least five days prior to the start of the study to allow for acclimatisation to the laboratory conditions.

**1.4.1.4. Test substance/Preparation**

Solid test substances should be dissolved or suspended in appropriate solvents or vehicles and diluted, if appropriate, prior to dosing of the animals. Liquid test substances may be dosed directly or diluted prior to dosing. Fresh preparations of the test substance should be employed unless stability data demonstrate the acceptability of storage.

**1.4.2. Test conditions****1.4.2.1. Solvent/Vehicle**

The solvent/vehicle should not produce toxic effects at the dose levels used, and should not be suspected of chemical reaction with the test substance. If other than well-known solvents/vehicles are used, their inclusion should be supported with data indicating their compatibility. It is recommended that wherever possible, the use of an aqueous solvent/vehicle should be considered first.

**▼B**1.4.2.2. *Controls*

Concurrent positive and negative controls (solvent/vehicle) should be included in each independently performed part of the experiment. Except for treatment with the test substance, animals in the control group should be handled in an identical manner to the animals in the treated groups.

Positive controls should be substances known to produce UDS when administered at exposure levels expected to give a detectable increase over background. Positive controls needing metabolic activation should be used at doses eliciting a moderate response (4). The doses may be chosen so that the effects are clear but do not immediately reveal the identity of the coded slides to the reader. Examples of positive control substances include:

Sampling Times	Substance	CAS No	EINECS No
Early sampling times (2-4 hours)	N-Nitrosodimethylamine	62-75-9	200-249-8
Late sampling times (12-16 hours)	N-2-Fluorenylacetamide (2-AAF)	53-96-3	200-188-6

Other appropriate positive control substances may be used. It is acceptable that the positive control should be administered by a route different from the test substance.

## 1.5. PROCEDURE

1.5.1. **Number and sex of animals**

An adequate number of animals should be used, to take account of natural biological variation in test response. The number of animals should be at least three analysable animals per group. Where a significant historical database has been accumulated, only one or two animals are required for the concurrent negative and positive control groups.

If at the time of the study there are data available from studies in the same species and using the same route of exposure that demonstrate that there are no substantial differences in toxicity between sexes, then testing in a single sex, preferably males, will be sufficient. Where human exposure to chemicals may be sex-specific, as for example with some pharmaceutical agents, the test should be performed with animals of the appropriate sex.

1.5.2. **Treatment schedule**

Test substances are generally administered as a single treatment.

1.5.3. **Dose levels**

Normally, at least two dose levels are used. The highest dose is defined as the dose producing signs of toxicity such that higher dose levels, based on the same dosing regimen, would be expected to produce lethality. In general, the lower dose should be 50 % to 25 % of the high dose.

**▼B**

Substances with specific biological activities at low non-toxic doses (such as hormones and mitogens) may be exceptions to the dose-setting criteria and should be evaluated on a case-by-case basis. If a range finding study is performed because there are no suitable data available, it should be performed in the same laboratory, using the same species, strain, sex, and treatment regimen to be used in the main study.

The highest dose may also be defined as a dose that produces some indication of toxicity in the liver (e.g. pyknotic nuclei).

**1.5.4. Limit test**

If a test at one dose level of at least 2 000 mg/kg body weight, applied in a single treatment, or in two treatments on the same day, produces no observable toxic effects, and if genotoxicity would not be expected, based upon data from structurally related substances, then a full study may not be necessary. Expected human exposure may indicate the need for a higher dose level to be used in the limit test.

**1.5.5. Administration of doses**

The test substance is usually administered by gavage using a stomach tube or a suitable intubation cannula. Other routes of exposure may be acceptable where they can be justified. However, the intraperitoneal route is not recommended as it could expose the liver directly to the test substance rather than via the circulatory system. The maximum volume of liquid that can be administered by gavage or injection at one time depends on the size of the test animal. The volume should not exceed 2 ml/100 g body weight. The use of volumes higher than these must be justified. Except for irritating or corrosive substances, which will normally reveal exacerbated effects with higher concentrations, variability in test volume should be minimised by adjusting the concentration to ensure a constant volume at all dose levels.

**1.5.6. Preparation of liver cells**

Liver cells are prepared from treated animals normally 12-16 hours after dosing. An additional earlier sampling time (normally two to four hours post-treatment) is generally necessary unless there is a clear positive response at 12-16 hours. However, alternative sampling times may be used when justified on the basis of toxicokinetic data.

Short-term cultures of mammalian liver cells are usually established by perfusing the liver *in situ* with collagenase and allowing freshly dissociated liver cells to attach themselves to a suitable surface. Liver cells from negative control animals should have a viability (5) of at least 50 %.

**1.5.7. Determination of UDS**

Freshly isolated mammalian liver cells are incubated usually with medium containing <sup>3</sup>H-TdR for an appropriate length of time, e.g. 3-8 hours. At the end of the incubation period, medium should be removed from the cells, which may then be incubated with medium containing excess unlabelled thymidine to diminish unincorporated radioactivity ('cold chase'). The cells are then rinsed, fixed and dried. For more prolonged incubation times, cold chase may not be necessary. Slides are dipped in autoradiographic emulsion, exposed in the dark (e.g. refrigerated for 7-14 days), developed, stained, and exposed silver grains are counted. Two to three slides are prepared from each animal.

**▼B**1.5.8. **Analysis**

The slide preparations should contain sufficient cells of normal morphology to permit a meaningful assessment of UDS. Preparations are examined microscopically for signs of overt cytotoxicity (e.g. pyknosis, reduced levels of radiolabelling).

Slides should be coded before grain counting. Normally 100 cells are scored from each animal from at least two slides; the scoring of less than 100 cells/animal should be justified. Grain counts are not scored for S-phase nuclei, but the proportion of S-phase cells may be recorded.

The amount of  $^3\text{H}$ -TdR incorporation in the nuclei and the cytoplasm of morphologically normal cells, as evidenced by the deposition of silver grains, should be determined by suitable methods.

Grain counts are determined over the nuclei (nuclear grains, NG) and nucleus equivalent areas over the cytoplasm (cytoplasmic grains, CG). CG counts are measured by either taking the most heavily labelled area of cytoplasm, or by taking an average of two to three random cytoplasmic grain counts adjacent to the nucleus. Other counting methods (e.g. whole cell counting) may be used if they can be justified (6).

2. **DATA**

## 2.1. TREATMENT OF RESULTS

Individual slide and animal data should be provided. Additionally, all data should be summarised in tabular form. Net nuclear grain (NNG) counts should be calculated for each cell, for each animal and for each dose and time by subtracting CG counts from NG counts. If 'cells in repair' are counted, the criteria for defining 'cells in repair' should be justified and based on historical or concurrent negative control data. Numerical results may be evaluated by statistical methods. If used, statistical tests should be selected and justified prior to conducting the study.

## 2.2. EVALUATION AND INTERPRETATION OF RESULTS

Examples of criteria for positive/negative responses include:

- |          |  |
|----------|--|
| positive | (i) NNG values above a pre-set threshold which is justified on the basis of laboratory historical data; or |
|          | (ii) NNG values significantly greater than concurrent control;   |
| negative | (i) NNG values within/below historical control threshold; or   |
|          | (ii) NNG values not significantly greater than concurrent control.   |

**▼B**

The biological relevance of data should be considered: i.e. parameters such as inter-animal variation, dose-response relationship and cytotoxicity should be taken into account. Statistical methods may be used as an aid in evaluating the test results. However, statistical significance should not be the only determining factor for a positive response.

Although most experiments will give clearly positive or negative results, in rare cases the data set will preclude making a definite judgement about the activity of the test substance. Results may remain equivocal or questionable regardless of the number of times the experiment is repeated.

A positive result from the UDS test with mammalian liver cells *in vivo* indicate that a test substance induces DNA damage in mammalian liver cells *in vivo* that can be repaired by unscheduled DNA synthesis *in vitro*. A negative result indicates that, under the test conditions, the test substance does not induce DNA damage that is detectable by this test.

The likelihood that the test substance reaches the general circulation or specifically the target tissue (e.g. systemic toxicity) should be discussed.

### 3. **REPORTING**

#### TEST REPORT

The test report must include the following information:

Solvent/Vehicle:

- justification for choice of vehicle,
- solubility and stability of the test substance in solvent/vehicle, if known.

Test animals:

- species/strain used,
- number, age and sex of animals,
- source, housing conditions, diet, etc.,
- individual weight of the animals at the start of the test, including body weight range, mean and standard deviation for each group,

Test conditions:

- positive and negative vehicle/solvent controls,
- data from range-finding study, if conducted,
- rationale for dose level selection,
- details of test substance preparation,
- details of the administration of the test substance,
- rationale for route of administration,
- methods for verifying that test agent reached the general circulation or target tissue, if applicable,

**▼B**

- conversion from diet/drinking water test substance concentration (ppm) to the actual dose (mg/kg body weight/day), if applicable,
- details of food and water quality,
- detailed description of treatment and sampling schedules,
- methods for measurement of toxicity,
- method of liver cell preparation and culture,
- autoradiographic technique used,
- number of slides prepared and numbers of cells scored,
- evaluation criteria,
- criteria for considering studies as positive, negative or equivocal,

## Results:

- individual slide, animal and group mean values for nuclear grains, cytoplasmic grains, and net nuclear grains,
- dose-response relationship, if available,
- statistical evaluation if any,
- signs of toxicity,
- concurrent negative (solvent/vehicle) and positive control data,
- historical negative (solvent/vehicle) and positive control data with range, means and standard deviations,
- number of 'cells in repair' if determined,
- number of S-phase cells if determined,
- viability of the cells.

Discussion of results.

Conclusions.

#### 4. REFERENCES

- (1) Ashby, J., Lefevre, P.A., Burlinson, B. and Penman, M.G., (1985) An Assessment of the *In Vivo* Rat Hepatocyte DNA Repair Assay. *Mutation Res.*, 156, p. 1-18.
- (2) Butterworth, B.E., Ashby, J., Bermudez, E., Casciano, D., Mirsalis, J., Probst, G. and Williams, G., (1987) A Protocol and Guide for the *In Vivo* Rat Hepatocyte DNA-Repair Assay. *Mutation Res.*, 189, p. 123-133.
- (3) Kennelly, J.C., Waters, R., Ashby, J., Lefevre, P.A., Burlinson, B., Benford, D.J., Dean, S.W. and Mitchell, I. de G., (1993) *In Vivo* Rat Liver UDS Assay. In: Kirkland D.J. and Fox M., (Eds) *Supplementary Mutagenicity Tests: UKEM Recommended Procedures*. UKEMS Subcommittee on Guidelines for Mutagenicity Testing. Report. Part II revised. Cambridge University Press, Cambridge, New York, Port Chester, Melbourne, Sydney, p. 52-77.

**▼B**

- (4) Madle, S., Dean, S.W., Andrae, U., Brambilla, G., Burlinson, B., Doolittle, D.J., Furihata, C., Hertner, T., McQueen, C.A. and Mori, H., (1993) Recommendations for the Performance of UDS Tests *In Vitro* and *In Vivo*. *Mutation Res.*, 312, p. 263-285.
- (5) Fautz, R., Hussain, B., Efstathiou, E. and Hechenberger-Freudl, C., (1993) Assessment of the Relation Between the Initial Viability and the Attachment of Freshly Isolated Rat Hepatocytes Used for the *In Vivo/In Vitro* DNA Repair Assay (UDS). *Mutation Res.*, 291, p. 21-27.
- (6) Mirsalis, J.C., Tyson, C.K. and Butterworth, B.E., (1982) Detection of Genotoxic Carcinogens in the *In Vivo/In Vitro* Hepatocyte DNA Repair Assay. *Environ. Mutagen*, 4, p. 553-562.



**▼B****B.40. IN VITRO SKIN CORROSION: TRANSCUTANEOUS ELECTRICAL RESISTANCE TEST (TER)****1. METHOD**

This testing method is equivalent to the OECD TG 430 (2004).

**1.1. INTRODUCTION**

Skin corrosion refers to the production of irreversible tissue damage in the skin following the application of a test material (as defined by the Globally Harmonised System for the Classification and Labelling of Chemical Substances and Mixtures (GHS)) (1). This method provides a procedure by which the assessment of corrosivity is not carried out in live animals.

The assessment of skin corrosivity has typically involved the use of laboratory animals (2). Concern for the pain and suffering of animals involved with this procedure has been addressed in the revision of testing method B.4 that allows for the determination of skin corrosion by using alternative, *in vitro*, methods, avoiding pain and suffering.

A first step towards defining alternative tests that could be used for skin corrosivity testing for regulatory purposes was the conduct of prevalidation studies (3). Following this, a formal validation study of *in vitro* methods for assessing skin corrosion (4)(5) was conducted (6)(7)(8). The outcome of these studies and other published literature led to the recommendation that the following tests could be used for the assessment of *in vivo* skin corrosivity (9)(10)(11): the human skin model test (see testing method B.40*bis*) and the transcutaneous electrical resistance test (this method).

A validation study and other published studies have reported that the rat skin transcutaneous electrical resistance (TER) assay (12)(13) is able to reliably discriminate between known skin corrosives and non-corrosives (5)(9).

The test described in this method allows the identification of corrosive chemical substances and mixtures. It further enables the identification of non-corrosive substances and mixtures when supported by a weight of evidence determination using other existing information (e.g. pH, structure-activity-relationships, human and/or animal data) (1)(2)(11)(14). It does not provide information on skin irritation, nor does it allow the sub-categorisation of corrosive substances as permitted in the Globally Harmonised Classification System (GHS) (1).

For a full evaluation of local skin effects after a single dermal exposure, it is recommended to follow the sequential testing strategy as appended to testing method B.4 (2) and provided in the Globally Harmonised System (1). This testing strategy includes the conduct of *in vitro* tests for skin corrosion (as described in this method) and skin irritation before considering testing in live animals.

**▼B**

## 1.2. DEFINITIONS

**Skin corrosion *in vivo*:** is the production of irreversible damage of the skin: namely, visible necrosis through the epidermis and into the dermis, following the application of a test substance for up to four hours. Corrosive reactions are typified by ulcers, bleeding, bloody scabs, and, by the end of the observation at 14 days, by discolouration due to blanching of the skin, complete areas of alopecia, and scars. Histopathology should be considered to evaluate questionable lesions.

**Transcutaneous Electrical Resistance (TER):** is a measure of the electrical impedance of the skin, as a resistance value in kilo Ohms. A simple and robust method of assessing barrier function by recording the passage of ions through the skin using a Wheatstone bridge apparatus.

## 1.3. REFERENCE SUBSTANCES

Table 1

## Reference chemicals

Name	EINECS No	CAS No	
1,2-Diaminopropane	201-155-9	78-90-0	Severely corrosive
Acrylic Acid	201-177-9	79-10-7	Severely Corrosive
2-tert. Butylphenol	201-807-2	88-18-6	Corrosive
Potassium hydroxide (10 %)	215-181-3	1310-58-3	Corrosive
Sulfuric acid (10 %)	231-639-5	7664-93-9	Corrosive
Octanoic acid (caprylic acid)	204-677-5	124-07-02	Corrosive
4-Amino-1,2,4-triazole	209-533-5	584-13-4	Not corrosive
Eugenol	202-589-1	97-53-0	Not corrosive
Phenethyl bromide	203-130-8	103-63-9	Not corrosive
Tetrachloroethylene	204-825-9	27-18-4	Not corrosive
Isostearic acid	250-178-0	30399-84-9	Not corrosive
4-(Methylthio)-benzaldehyde	222-365-7	3446-89-7	Not corrosive

Most of the chemicals listed are taken from the list of chemicals selected for the ECVAM international validation study (4). Their selection is based on the following criteria:

- (i) equal number of corrosive and non-corrosive substances;

**▼ B**

- (ii) commercially available substances covering most of the relevant chemical classes;
  
- (iii) inclusion of severely corrosive as well as less corrosive substances in order to enable discrimination based on corrosive potency;
  
- (iv) choice of chemicals that can be handled in a laboratory without posing other serious hazards than corrosivity.

#### 1.4. PRINCIPLE OF THE TEST METHOD

The test material is applied for up to 24 hours to the epidermal surfaces of skin discs in a two-compartment test system in which the skin discs function as the separation between the compartments. The skin discs are taken from humanely killed rats aged 28-30 days. Corrosive materials are identified by their ability to produce a loss of normal stratum corneum integrity and barrier function, which is measured as a reduction in the TER below a threshold level (12). For rat TER, a cut-off value of 5 k $\Omega$  has been selected based on extensive data for a wide range of chemicals where the vast majority of values were either clearly well above (often > 10 k $\Omega$ ), or well below (often < 3 k $\Omega$ ) this value (12). Generally, materials which are non-corrosive in animals but are irritating or non-irritating do not reduce the TER below this cut-off value. Furthermore, use of other skin preparations or other equipment may alter the cut-off value, necessitating further validation.

A dye-binding step is incorporated into the test procedure for confirmation testing of positive results in the TER including values around 5 k $\Omega$ . The dye-binding step determines if the increase in ionic permeability is due to physical destruction of the stratum corneum. The TER method utilising rat skin has shown to be predictive of *in vivo* corrosivity in the rabbit assessed under Testing Method B.4 (2). It should be noted that the *in vivo* rabbit test is highly conservative with respect to skin corrosivity and skin irritation when compared with the human skin patch test (15).

#### 1.5. DESCRIPTION OF THE TEST METHOD

##### 1.5.1. **Animals**

Rats are the species of choice because the sensitivity of their skin to chemicals in this test has been previously demonstrated (10). The age (when the skin is collected) and strain of the rat is particularly important to ensure that the hair follicles are in the dormant phase before adult hair growth begins.

**▼B**

The dorsal and flank hair from young, approximately 22 day-old, male or female rats (Wistar-derived or a comparable strain), is carefully removed with small clippers. Then, the animals are washed by careful wiping, whilst submerging the clipped area in antibiotic solution (containing, for example, streptomycin, penicillin, chloramphenicol, and amphotericin, at concentrations effective in inhibiting bacterial growth). Animals are washed with antibiotics again on the third or fourth day after the first wash and are used within three days of the second wash, when the stratum corneum has recovered from the hair removal.

**1.5.2. Preparation of the skin discs**

Animals are humanely killed when 28-30 days old; this age is critical. The dorso-lateral skin of each animal is then removed and stripped of excess subcutaneous fat by carefully peeling it away from the skin. Skin discs, with a diameter of approximately 20 mm each, are removed. The skin may be stored before disks are used where it is shown that positive and negative control data are equivalent to that obtained with fresh skin.

Each skin disc is placed over one of the ends of a PTFE (polytetrafluoroethylene) tube, ensuring that the epidermal surface is in contact with the tube. A rubber 'O' ring is press-fitted over the end of the tube to hold the skin in place and excess tissue is trimmed away. Tube and 'O' ring dimensions are shown in Figure 2. The rubber 'O' ring is then carefully sealed to the end of the PTFE tube with petroleum jelly. The tube is supported by a spring clip inside a receptor chamber containing MgSO<sub>4</sub> solution (154 mM) (Figure 1). The skin disc should be fully submerged in the MgSO<sub>4</sub> solution. As many as 10-15 skin discs can be obtained from a single rat skin.

Before testing begins, the electrical resistance of two skin discs is measured as a quality control procedure for each animal skin. Both discs should give resistance values greater than 10 kΩ for the remainder of the discs to be used for the test. If the resistance value is less than 10 kΩ, the remaining discs from that skin should be discarded.

**1.5.3. Application of the test and control substances**

Concurrent positive and negative controls should be used for each study to ensure adequate performance of the experimental model. Skin discs from a single animal should be used. The suggested positive and negative control substances are 10 M hydrochloric acid and distilled water, respectively.

Liquid test substances (150 µL) are applied uniformly to the epidermal surface inside the tube. When testing solid materials, a sufficient amount of the solid is applied evenly to the disc to ensure that the whole surface of the epidermis is covered. Deionised water (150 µL) is added on top of the solid and the tube is gently agitated. In order to achieve maximum contact with the skin, solids may need to be warmed to 30 °C to melt or soften the test substance, or ground to produce a granular material or powder.

**▼ B**

Three skin discs are used for each test and control substance. Test substances are applied for 24 hours at 20-23 °C. The test substance is removed by washing with a jet of tap water at up to 30 °C until no further material can be removed.

**1.5.4. TER measurements**

The skin impedance is measured as TER is measured by using a low-voltage, alternating current Wheatstone databridge (13). General specifications of the bridge are 1-3 Volt operating voltage, a sinus or rectangular shaped alternating current of 50 – 1 000 Hz, and a measuring range of at least 0,1 – 30 k $\Omega$ . The databridge used in the validation study measures inductance, capacitance and resistance up to values of 2 000 H, 2 000  $\mu$ F, and 2 M $\Omega$ , respectively at frequencies of 100 Hz or 1 kHz, using series or parallel values. For the purposes of the TER corrosivity assay measurements are recorded in resistance, at a frequency of 100 Hz and using series values. Prior to measuring the electrical resistance, the surface tension of the skin is reduced by adding a sufficient volume of 70 % ethanol to cover the epidermis. After a few seconds, the ethanol is removed from the tube and the tissue is then hydrated by the addition of 3 mL MgSO<sub>4</sub> solution (154 mM). The databridge electrodes are placed on either side of the skin disc to measure the resistance in k $\Omega$ /skin disc (Figure 1). Electrode dimensions and the length of the electrode exposed below the crocodile clips are shown in Figure 2. The clip attached to the inner electrode is rested on the top of the PTFE tube during resistance measurement to ensure that a consistent length of electrode is submerged in the MgSO<sub>4</sub> solution. The outer electrode is positioned inside the receptor chamber so that it rests on the bottom of the chamber. The distance between the spring clip and the bottom of the PTFE tube is maintained as a constant (Figure 2), because this distance affects the resistance value obtained. Consequently, the distance between the inner electrode and the skin disc should be constant and minimal (1-2 mm).

If the measured resistance value is greater than 20 k $\Omega$ , this may be due to the remains of the test substance coating the epidermal surface of the skin disc. Further removal of this coating can be attempted, for example, by sealing the PTFE tube with a gloved thumb and shaking it for approximately 10 seconds; the MgSO<sub>4</sub> solution is discarded and the resistance measurement is repeated with fresh MgSO<sub>4</sub>.

The properties and dimensions of the test apparatus and the experimental procedure used may influence the TER values obtained. The 5 k $\Omega$  corrosive threshold was developed from data obtained with the specific apparatus and procedure described in this method. Different threshold and control values may apply if the test conditions are altered or a different apparatus is used. Therefore, it is necessary to calibrate the methodology and resistance threshold values by testing a series of reference standards chosen from the chemicals used in the validation study (4)(5), or from similar chemical classes to the chemicals being investigated. A set of suitable reference chemicals is shown in Table 1.

**▼ B****1.5.5. Dye binding methods**

Exposure of certain non-corrosive materials can result in a reduction of resistance below the cut-off of 5 k $\Omega$  allowing the passage of ions through the stratum corneum, thereby reducing the electrical resistance (5). For example, neutral organics and chemicals that have surface-active properties (including detergents, emulsifiers and other surfactants) can remove skin lipids making the barrier more permeable to ions. Thus, if the TER values of test substances are less than or around 5 k $\Omega$  in the absence of visual damage, an assessment of dye penetration should be carried out on the control and treated tissues to determine if the TER values obtained were the result of increased skin permeability, or skin corrosion (3)(5). In case of the latter where the stratum corneum is disrupted, the dye sulforhodamine B, when applied to the skin surface rapidly penetrates and stains the underlying tissue. This particular dye is stable to a wide range of chemicals and is not affected by the extraction procedure described below.

**1.5.5.1. Sulforhodamine B dye application and removal**

Following TER assessment, the magnesium sulfate is discarded from the tube and the skin is carefully examined for obvious damage. If there is no obvious major damage, Sulforhodamine B dye (Acid Red 52; C.I. 45100; EINECS Number 222-529-8; CAS number 3520-42-1), 150  $\mu$ L of a 10 % (w/v) dilution in distilled water, is applied to the epidermal surface of each skin disc for two hours. These skin discs are then washed with tap water at up to room temperature for approximately 10 seconds to remove any excess/unbound dye. Each skin disc is carefully removed from the PTFE tube and placed in a vial (e.g. a 20 mL glass scintillation vial) containing deionised water (8 mL). The vials are agitated gently for five minutes to remove any additional unbound dye. This rinsing procedure is then repeated, after which the skin discs are removed and placed into vials containing 5 ml of 30 % (w/v) sodium dodecyl sulphate (SDS) in distilled water and are incubated overnight at 60 °C.

After incubation, each skin disc is removed and discarded and the remaining solution is centrifuged for eight minutes at 21 °C (relative centrifugal force  $\sim 175 \times g$ ). A 1 ml sample of the supernatant is diluted 1 in 5 (v/v) [i.e. 1 mL + 4 mL] with 30 % (w/v) SDS in distilled water. The optical density (OD) of the solution is measured at 565 nm.

**1.5.5.2. Calculation of dye content**

The sulforhodamine B dye content per disc is calculated from the OD values (5) (sulforhodamine B dye molar extinction coefficient at 565 nm =  $8,7 \times 10^4$ ; molecular weight = 580). The dye content is determined for each skin disc by the use of an appropriate calibration curve and a mean dye content is then calculated for the replicates.

**2. DATA**

Resistance values (k $\Omega$ ) and mean dye content values ( $\mu$ g/disc), where appropriate, for the test material, as well as for positive and negative controls should be reported in tabular form (individual trial data and means  $\pm$  S.D.), including data for replicates/repeat experiments, mean and individual values.

**▼B**

## 2.1. INTERPRETATION OF RESULTS

The mean TER results are accepted if the concurrent positive and negative control values fall within the acceptable ranges for the method in the testing laboratory. The acceptable resistance ranges for the methodology and apparatus described above are given in the following table:

Control	Substance	Resistance range (k $\Omega$ )
Positive	10M Hydrochloric acid	0,5-1,0
Negative	Distilled water	10-25

The mean dye binding results are accepted on condition that concurrent control values fall within the acceptable ranges for the method. Suggested acceptable dye content ranges for the control substances for the methodology and apparatus described above are given below:

Control	Substance	Dye content range ( $\mu$ g/disc)
Positive	10M Hydrochloric acid	40-100
Negative	Distilled water	15-35

The test substance is considered to be non-corrosive to skin:

- (i) if the mean TER value obtained for the test substance is greater than 5 k $\Omega$ ; or
- (ii) the mean TER value is less than or equal to 5 k $\Omega$ ; and
  - the skin disc is showing no obvious damage, and
  - the mean disc dye content is well below the mean disc dye content of the 10M HCl positive control obtained concurrently.

The test substance is considered to be corrosive to skin:

- (i) if the mean TER value is less than or equal to 5 k $\Omega$  and the skin disk is obviously damaged; or
- (ii) the mean TER value is less than or equal to 5 k $\Omega$ ; and
  - the skin disc is showing no obvious damage, but
  - the mean disc dye content is greater than or equal to the mean disc dye content of the 10M HCl positive control obtained concurrently.

**▼B****3. REPORTING****3.1. TEST REPORT**

The test report must include the following information:

Test and control substances:

- chemical name(s) such as IUPAC or CAS name and CAS number, if known,
- purity and composition of the substance or preparation (in percentage(s) by weight) and physical nature,
- physico-chemical properties such as physical state, pH, stability, water solubility, relevant to the conduct of the study,
- treatment of the test/control substances prior to testing, if applicable (e.g. warming, grinding),
- stability, if known.

Test animals:

- strain and sex used,
- age of the animals when used as donor animals,
- source, housing condition, diet, etc,
- details of the skin preparation.

Test conditions:

- calibration curves for test apparatus,
- calibration curves for dye binding test performance,
- details of the test procedure used for TER measurements,
- details of the test procedure used for the dye binding assessment; if appropriate,
- description of any modification of the test procedures,
- description of evaluation criteria used.

Results:

- tabulation of data from the TER and dye binding assay (if appropriate) for individual animals and individual skin samples;
- description of any effects observed.

Discussion of the results.

Conclusions.

**4. REFERENCES**

- (1) OECD, (2001) Harmonised Integrated Classification System for Human Health and Environmental Hazards of Chemical Substances and Mixtures. OECD Series on Testing and Assessment Number 33. ENV/JM/MONO(2001)6, Paris. [http://www.olis.oecd.org/olis/2001/doc.nsf/LinkTo/env-jm-mono\(2001\)6](http://www.olis.oecd.org/olis/2001/doc.nsf/LinkTo/env-jm-mono(2001)6).



**▼B**

- (2) Testing Method B.4. Acute Toxicity: Dermal Irritation/Corrosion.
- (3) Botham, P.A., Chamberlain, M., Barratt, M.D., Curren, R.D., Esdaile, D.J., Gardner, J.R., Gordon, V.C., Hildebrand, B., Lewis, R.W., Liebsch, M., Logemann, P., Osborne, R., Ponc, M., Regnier, J.F., Steiling, W., Walker, A.P., and Balls, M., (1995) A prevalidation study on *in vitro* skin corrosivity testing. The report and recommendations of ECVAM Workshop 6. ATLA 23, p. 219-255.
- (4) Barratt, M.D., Brantom, P.G., Fentem, J.H., Gerner, I., Walker, A.P., and Worth, A.P. (1998). The ECVAM international validation study on *in vitro* tests for skin corrosivity. 1. Selection and distribution of the test chemicals. Toxicology. *In Vitro* 12, p. 471-482.
- (5) Fentem, J.H., Archer, G.E.B., Balls, M., Botham, P.A., Curren, R.D., Earl, L.K., Esdaile, D.J., Holzhütter, H.-G., and Liebsch, M., (1998) The ECVAM international validation study on *in vitro* tests for skin corrosivity. 2. Results and evaluation by the Management Team. Toxicology. *In Vitro* 12, p. 483- 524.
- (6) OECD, (1996) Final Report of the OECD Workshop on Harmonisation of Validation and Acceptance Criteria for Alternative Toxicological Test Methods, p. 62.
- (7) Balls, M., Blaauboer, B.J., Fentem, J.H., Bruner, L., Combes, R.D., Ekwall, B., Fielder, R.J., Guillouzo, A., Lewis, R.W., Lovell, D.P., Reinhardt, C.A., Repetto, G., Sladowski, D., Spielmann, H., and Zucco, F., (1995) Practical aspects of the validation of toxicity test procedures. The report and recommendations of ECVAM workshops. ATLA 23, p. 129-147.
- (8) ICCVAM, (Interagency Coordinating Committee on the Validation of Alternative Methods), (1997) Validation and Regulatory Acceptance of Toxicological Test Methods. NIH Publication No 97-3981. National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA. <http://iccvam.niehs.nih.gov/docs/guidelines/validate.pdf>.
- (9) ECVAM, (1998) ECVAM News & Views. ATLA 26, 275-280.
- (10) ICCVAM (Interagency Coordinating Committee on the Validation of Alternative Methods), (2002) ICCVAM evaluation of EpiDerm™, EPISKIN™ (EPI-200), and the Rat Skin Transcutaneous Electrical Resistance (TER) assay: In Vitro test methods for assessing dermal corrosivity potential of chemicals. NIH Publication No 02-4502. National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA. [http://iccvam.niehs.nih.gov/methods/epiddocs/epis\\_brd.pdf](http://iccvam.niehs.nih.gov/methods/epiddocs/epis_brd.pdf).
- (11) OECD (2002) Extended Expert Consultation Meeting on The In Vitro Skin Corrosion Test Guideline Proposal, Berlin, 1st–2nd November 2001, Secretariat's Final Summary Report, 27<sup>th</sup> March 2002, OECD ENV/EHS, available upon request from the Secretariat.
- (12) Oliver, G.J.A., Pemberton, M.A., and Rhodes, C., (1986) An *in vitro* skin corrosivity test -modifications and validation. *Fd. Chem. Toxicol.* 24, 507-512.

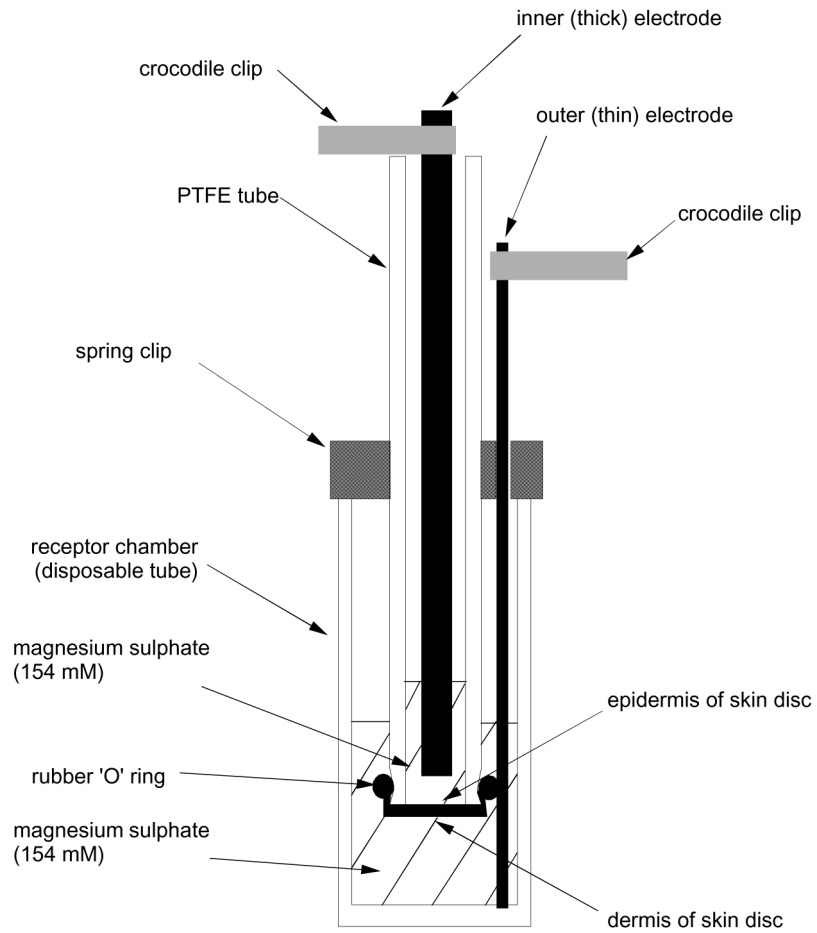
**▼B**

- (13) Botham, P.A., Hall, T.J., Dennett, R., McCall, J.C., Basketter, D.A., Whittle, E., Cheeseman, M., Esdaile, D.J., and Gardner, J., (1992) The skin corrosivity test *in vitro*: results of an inter-laboratory trial. *Toxic. In Vitro* 6, p. 191-194.
- (14) Worth A.P., Fentem J.H., Balls, M., Botham, P.A., Curren, R.D., Earl, L.K., Esdaile D.J., Liebsch, M., (1998) An Evaluation of the Proposed OECD Testing Strategy for Skin Corrosion. *ATLA* 26, p. 709-720
- (15) Basketter, D.A., Chamberlain, M., Griffiths, H.A., Rowson, M., Whittle, E., York, M., (1997) The classification of skin irritants by human patch test. *Fd. Chem. Toxicol.* 35, p. 845-852.
- (16) Oliver G.J.A, Pemberton M.A and Rhodes C., (1988) An In Vitro model for identifying skin-corrosive chemicals. I. Initial Validation. *Toxicology In Vitro.* 2, p. 7-17.

▼ B

Figure 1

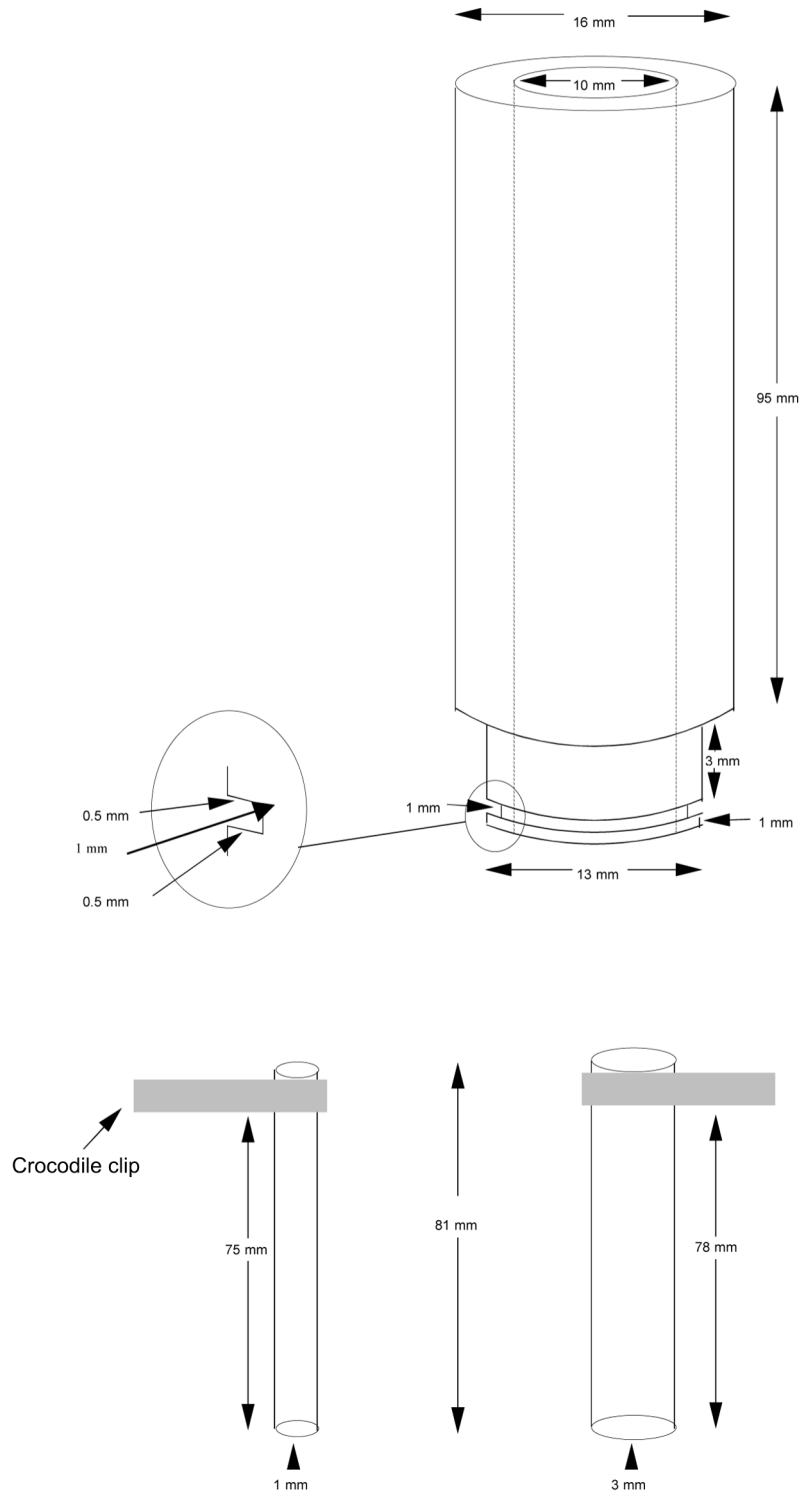
## Apparatus for the rat skin TER assay



▼B

Figure 2

Dimensions of the polytetrafluoroethylene (PFTE) and receptor tubes and electrodes used



**▼B****Critical factors of the apparatus shown above:**

- the inner diameter of the PTFE tube,
- the length of the electrodes relative to the PTFE tube and receptor tube, such that the skin disc is not touched by the electrodes and that a standard length of electrode is in contact with the  $\text{MgSO}_4$  solution,
- the amount of  $\text{MgSO}_4$  solution in the receptor tube should give a depth of liquid, relative to the level in the PTFE tube, as shown in Figure 1,
- the skin disk should be fixed well enough to the PTFE tube, such that the electrical resistance is a true measure of the skin properties.

**▼B****B.40 BIS. IN VITRO SKIN CORROSION: HUMAN SKIN MODEL TEST****1. METHOD**

This testing method is equivalent to the OECD TG 431 (2004).

**1.1. INTRODUCTION**

Skin corrosion refers to the production of irreversible tissue damage in the skin following the application of a test material [as defined by the Globally Harmonised System for the Classification and Labelling of Chemical Substances and Mixtures (GHS)] (1). This Testing Method does not require the use of live animals or animal tissue for the assessment of skin Corrosivity.

The assessment of skin corrosivity has typically involved the use of laboratory animals (2). Concern for the pain and suffering involved with this procedure has been addressed in the revision of testing method B.4 that allows for the determination of skin corrosion by using alternative, *in vitro*, methods, avoiding pain and suffering of animals.

A first step towards defining alternative tests that could be used for skin corrosivity testing for regulatory purposes was the conduct of prevalidation studies (3). Following this, a formal, validation study of *in vitro* methods for assessing skin corrosion (4)(5) was conducted (6)(7)(8). The outcome of these studies and other published literature (9) led to the recommendation that the following tests could be used for the assessment of the *in vivo* skin corrosivity (10)(11)(12)(13): the human skin model test (this method) and the transcutaneous electrical resistance test (see testing method B.40).

Validation studies have reported that tests employing human skin models (3)(4)(5)(9) are able to reliably discriminate between known skin corrosives and non-corrosives. The test protocol may also provide an indication of the distinction between severe and less severe skin corrosives.

The test described in this method allows the identification of corrosive chemical substances and mixtures. It further allows the identification of non-corrosive substances and mixtures when supported by a weight of evidence determination using other existing information (e.g. pH, structure-activity relationships, human and/or animal data) (1)(2)(13)(14). It does not normally provide adequate information on skin irritation, nor does it allow the subcategorisation of corrosive substances as permitted in the Globally Harmonised Classification System (GHS) (1).

For a full evaluation of local skin effects after single dermal exposure, it is recommended to follow the sequential testing strategy as appended to testing method B.4 (2) and provided in the Globally Harmonised System (GHS) (1). This testing strategy includes the conduct of *in vitro* tests for skin corrosion (as described in this method) and skin irritation before considering testing in live animals.

**▼B**

## 1.2. DEFINITIONS

**Skin corrosion *in vivo*:** is the production of irreversible damage of the skin: namely, visible necrosis through the epidermis and into the dermis, following the application of a test substance for up to four hours. Corrosive reactions are typified by ulcers, bleeding, bloody scabs, and, by the end of the observation at 14 days, by discolouration due to blanching of the skin, complete areas of alopecia, and scars. Histopathology should be considered to evaluate questionable lesions.

**Cell viability:** parameter measuring total activity of a cell population (e.g. ability of cellular mitochondrial dehydrogenases to reduce the vital dye MTT), which, depending on the end point measured and the test design used, correlates with the total number and/or vitality of the cells.

## 1.3. REFERENCE SUBSTANCES

Table 1

## Reference chemicals

Name	EINECS No	CAS No	
1,2-Diaminopropane	201-155-9	78-90-0	Severely corrosive
Acrylic Acid	201-177-9	79-10-7	Severely corrosive
2-tert. Butylphenol	201-807-2	88-18-6	Corrosive
Potassium hydroxide (10 %)	215-181-3	1310-58-3	Corrosive
Sulfuric acid (10 %)	231-639-5	7664-93-9	Corrosive
Octanoic acid (caprylic acid)	204-677-5	124-07-02	Corrosive
4-Amino-1,2,4-triazole	209-533-5	584-13-4	Not corrosive
Eugenol	202-589-1	97-53-0	Not corrosive
Phenethyl bromide	203-130-8	103-63-9	Not corrosive
Tetrachloroethylene	204-825-9	27-18-4	Not Corrosive
Isostearic acid	250-178-0	30399-84-9	Not corrosive
4-(Methylthio)-benzaldehyde	222-365-7	3446-89-7	Not corrosive

Most of the chemicals listed are taken from the list of chemicals selected for the ECVAM international validation study (4). Their selection is based on the following criteria:

- (i) equal number of corrosive and non-corrosive substances;
- (ii) commercially available substances covering most of the relevant chemical classes;
- (iii) inclusion of severely corrosive as well as less corrosive substances in order to enable discrimination based on corrosive potency;

**▼ B**

- (iv) choice of chemicals that can be handled in a laboratory without posing other serious hazards than corrosivity.

#### 1.4. PRINCIPLE OF THE TEST METHOD

The test material is applied topically to a three-dimensional human skin model, comprising at least a reconstructed epidermis with a functional stratum corneum. Corrosive materials are identified by their ability to produce a decrease in cell viability (as determined, for example, by using the MTT reduction assay (15)) below defined threshold levels at specified exposure periods. The principle of the human skin model assay is based on the hypothesis that corrosive chemicals are able to penetrate the stratum corneum by diffusion or erosion, and are cytotoxic to the underlying cell layers.

##### 1.4.1. Procedure

###### 1.4.1.1. Human skin models

Human skin models can be constructed or obtained commercially (e.g. the EpiDerm™ and EPISKIN™ models) (16)(17)(18)(19) or be developed or constructed in the testing laboratory (20)(21). It is recognised that the use of human skin is subject to national and international ethical considerations and conditions. Any new model should be validated (at least to the extent described under 1.4.1.1.2). Human skin models used for this test must comply with the following:

###### 1.4.1.1.1. General model conditions:

Human keratinocytes should be used to construct the epithelium. Multiple layers of viable epithelial cells should be present under a functional stratum corneum. The skin model may also have a stromal component layer. Stratum corneum should be multi-layered with the necessary lipid profile to produce a functional barrier with robustness to resist rapid penetration of cytotoxic markers. The containment properties of the model should prevent passage of material around the stratum corneum to the viable tissue. Passage of test chemicals around the stratum corneum will lead to poor modeling of the exposure to skin. The skin model should be free of contamination with bacteria (including mycoplasma) or fungi.

###### 1.4.1.1.2. Functional model conditions:

The magnitude of viability is usually quantified by using MTT or the other metabolically converted vital dyes. In these cases the optical density (OD) of the extracted (solubilised) dye from the negative control tissue should be at least 20 fold greater than the OD of the extraction solvent alone (for an overview, see (22)). The negative control tissue should be stable in culture (provide similar viability measurements) for the duration of the test exposure period. The stratum corneum should be sufficiently robust to resist the rapid penetration of certain cytotoxic marker chemicals (e.g. 1 % Triton X-100). This property can be estimated by the exposure time required to reduce cell viability by 50 % (ET<sub>50</sub>) (e.g. for the EpiDerm™ and EPISKIN™ models this is > 2 hours). The tissue should demonstrate reproductivity over time and preferably between laboratories. Moreover it should be capable of predicting the corrosive potential of the reference chemicals (see Table 1) when used in the testing protocol selected.



**▼ B**1.4.1.2. *Application of the test and control substances*

Two tissue replicates are used for each treatment (exposure time), including controls. For liquid materials, sufficient test substance must be applied to uniformly cover the skin surface: a minimum of 25  $\mu\text{L}/\text{cm}^2$  should be used. For solid materials, sufficient test substance must be applied evenly to cover the skin, and it should be moistened with deionised or distilled water to ensure good contact with the skin. Where appropriate, solids should be ground to a powder before application. The application method should be appropriate for the test substance (see e.g. reference 5). At the end of the exposure period, the test material must be carefully washed from the skin surface with an appropriate buffer, or 0,9 % NaCl.

Concurrent positive and negative controls should be used for each study to ensure adequate performance of the experimental model. The suggested positive control substances are glacial acetic acid or 8N KOH. The suggested negative controls are 0,9 % NaCl or water.

1.4.1.3. *Cell viability measurements*

Only quantitative, validated, methods can be used to measure cell viability. Furthermore, the measure of viability must be compatible with use in a three-dimensional tissue construct. Non-specific dye binding must not interfere with the viability measurement. Protein binding dyes and those which do not undergo metabolic conversion (e.g. neutral red) are therefore not appropriate. The most frequently used assay is MTT (3-(4,5-Dimethylthiazol-2yl)-2,5-diphenyltetrazolium bromide, Thiazolyl blue: EINECS number 206-069-5, CAS number 298-93-1) reduction, which has been shown to give accurate and reproducible results (5) but others may be used. The skin sample is placed in an MTT solution of appropriate concentration (e.g. 0,3-1 mg/mL) at appropriate incubation temperature for three hours. The precipitated blue formazan product is then extracted using a solvent (isopropanol), and the concentration of the formazan is measured by determining the OD at wavelength between 540 and 595 nm.

Chemical action by the test material on the vital dye may mimic that of cellular metabolism leading to a false estimate of viability. This has been shown to happen when such a test material is not completely removed from the skin by rinsing (9). If the test material directly acts on the vital dye, additional controls should be used to detect and correct for the test substances interference with the viability measurement (9)(23).

**2. DATA**

For each tissue, OD values and calculated percentage cell viability data for the test material, positive and negative controls, should be reported in tabular form, including data from replicate repeat experiments as appropriate, mean and individual values.

**▼B**

## 2.1. INTERPRETATION OF RESULTS

The OD values obtained for each test sample can be used to calculate a percentage viability relative to the negative control, which is arbitrarily set at 100 %. The cut-off percentage cell viability value distinguishing corrosive from non-corrosive test materials (or discriminating between different, corrosive classes), or the statistical procedure(s) used to evaluate the results and identify corrosive materials, must be clearly defined and documented, and be shown to be appropriate. In general, these cut-off values are established during test optimisation, tested during a prevalidation phase, and confirmed in a validation study. As an example, the prediction of Corrosivity associated with the EpiDerm<sup>TM</sup> model is (9):

The test substance is considered to be corrosive to skin:

- (i) if the viability after three minutes exposure is less than 50 %; or
- (ii) if the viability after three minutes exposure is greater than or equal to 50 % and the viability after 1 hour exposure is less than 15 %.

The test substance is considered to be non-corrosive to skin:

- (i) if the viability after three minutes exposure is greater than or equal to 50 % and the viability after 1 hour exposure is greater than or equal to 15 %.

3. **REPORTING**

## 3.1. TEST REPORT

The test report must include the following information:

Test and control substance:

- chemical names(s) such as IUPAC or CAS name and CAS number, if known,
- purity and composition of the substance or preparation (in percentage(s) by weight),
- physico-chemical properties such as physical state, pH, stability, water solubility relevant to the conduct of the study,
- treatment of the test/control substances prior to testing, if applicable (e.g. warming, grinding),
- stability, if known.

Justification of the skin model and protocol used.

Test conditions:

- cell system used,
- calibration information for measuring device used for measuring cell viability (e.g. Spectrophotometer),

**▼ B**

- complete supporting information for the specific skin model used including its validity,
- details of the test procedure used,
- test doses used,
- description of any modifications of the test procedure,
- reference to historical data of the model,
- description of evaluation criteria used.

## Results:

- tabulation of data from individual test samples,
- description of other effects observed.

## Discussion of the results.

## Conclusion.

## 4. REFERENCES

- (1) OECD (2001) Harmonised Integrated Classification System for Human Health and Environmental Hazards of Chemical Substances and Mixtures. OECD Series on Testing and Assessment Number 33. ENV/JM/MONO(2001)6, Paris. [http://www.olis.oecd.org/olis/2001doc.nsf/LinkTo/env-jm-mono\(2001\)6](http://www.olis.oecd.org/olis/2001doc.nsf/LinkTo/env-jm-mono(2001)6).
- (2) Testing Method B.4. Acute Toxicity: Dermal Irritation/Corrosion.
- (3) Botham, P.A., Chamberlain, M., Barratt, M.D., Curren, R.D., Esdaile, D.J., Gardner, J.R., Gordon, V.C., Hildebrand, B., Lewis, R.W., Liebsch, M., Logemann, P., Osborne, R., Poniec, M., Regnier, J.F., Steiling, W., Walker, A.P., and Balls, M., (1995) A prevalidation study on *in vitro* skin corrosivity testing. The report recommendations of ECVAM Workshop 6 ATLA 23, p. 219-255.
- (4) Barratt, M.D., Brantom, P.G., Fentem, J.H., Gerner, I., Walker, A.P., and Worth, A.P. (1998). The ECVAM international validation study on *in vitro* tests for skin corrosivity. 1. Selection and distribution of the test chemicals. Toxicology In Vitro 12, p. 471-482.
- (5) Fentem, J.H., Archer, G.E.B., Balls, M., Botham, P.A., Curren, R.D., Earl, L.K., Esdaile, D.J., Holzhutter, H.G. and Liebsch, M., (1998) The ECVAM international validation study on *in vitro* tests for skin corrosivity. 2. Results and evaluation by the Management Team. Toxicology In Vitro 12, p. 483-524.
- (6) OECD, (1996) Final Report of the OECD Workshop on Harmonisation of Validation and Acceptance Criteria for Alternative Toxicological Test Methods, p. 62.
- (7) Balls, M., Blaauboer, B.J., Fentem, J.H., Bruner, L., Combes, R.D., Ekwall, B., Fielder, R.J., Guillouzo, A., Lewis, R.W., Lovell, D.P., Reinhardt, C.A., Repetto, G., Sladowski, D., Spielmann, H., and Zucco, F., (1995) Practical aspects of the validation of toxicity test procedures. Test report and recommendations of ECVAM workshops. ATLA 23, p. 129-147.
- (8) ICCVAM (Interagency Coordinating Committee on the Validation of Alternative Methods), (1997) Validation and Regulatory Acceptance of Toxicological Test Methods. NIH Publication No 97-3981. National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA. <http://iccvam.niehs.nih.gov/docs/guidelines/validate.pdf>.

**▼B**

- (9) Liebsch, M., Traue, D., Barrabas, C., Spielmann, H., Uphill, P., Wilkins, S., McPherson, J.P., Wiemann, C., Kaufmann, T., Remmele, M. and Holzhutter, H.G., (2000) The ECVAM prevalidation study on the use of EpiDerm for skin Corrosivity testing. ATLA 28, p. 371-401.
- (10) ECVAM, (1998) ECVAM News & Views. ATLA 26, p. 275-280.
- (11) ECVAM, (2000) ECVAM News & Views. ATLA 28, p. 365-67.
- (12) ICCVAM (Interagency Coordinating Committee on the Validation of Alternative Methods), (2002) ICCVAM evaluation of EpiDerm™, EPISKIN™ (EPI-200), and the Rat Skin Transcutaneous Electrical Resistance (TER) assay: In Vitro test methods for assessing dermal corrosivity potential of chemicals. NIH Publication No 02-4502. National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA. [http://iccvam.niehs.nih.gov/methods/epiddocs/epis\\_brd.pdf](http://iccvam.niehs.nih.gov/methods/epiddocs/epis_brd.pdf).
- (13) OECD, (2002) Extended Expert Consultation Meeting on The In Vitro Skin Corrosion Test Guideline Proposal, Berlin, 1st–2nd November 2001, Secretariat's Final Summary Report, 27<sup>th</sup> March 2002, OECD ENV/EHS, available upon request from the Secretariat
- (14) Worth, A.P., Fentem, J.H., Balls, M., Botham, P.A., Curren, R.D., Earl, L.K., Esdaile, D.J., Liebsch, M., (1998) An Evaluation of the Proposed OECD Testing Strategy for Skin Corrosion. ATLA 26, p. 709-720.
- (15) Mosmann, T., (1983) Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J.Immunol. Meth.* 65, p. 55-63.
- (16) Cannon, C.L., Neal, P.J., Southee, J.A., Kubilus, J., and Klausner, M., 1994. New epidermal model for dermal irritancy testing. *Toxicology In Vitro* 8, p. 889-891.
- (17) Ponec, M., Boelsma, E., Weerheim, A., Mulder, A., Boutwstra, J., and Mommaas, M., 2000. Lipid and ultrastructural characterisation of reconstructed skin models. *International Journal of Pharmaceutics.* 203, p. 211-225.
- (18) Tinois, E., Gaetani, Q., Gayraud, B., Dupont, D., Rougier, A., Pouradier, D.X., (1994) The Episkin model: Successful reconstruction of human epidermis in vitro. In *In vitro Skin Toxicology*. Edited by A Rougier, AM Goldberg and HI Maibach, p. 133-140
- (19) Tinois E, Tiollier J, Gaucherand M, Dumas H, Tardy M, Thivolet J (1991). *In vitro* and post-transplantation differentiation of human keratinocytes grown on the human type IV collagen film of a bilayered dermal substitute. *Experimental Cell Research* 193, p. 310-319
- (20) Parentau, N.L., Bilbo, P., Molte, C.J., Mason, V.S., and Rosenberg, H., (1992) The organotypic culture of human skin keratinocytes and fibroblasts to achieve form and function. *Cytotechnology* 9, p. 163-171.
- (21) Wilkins, L.M., Watson, S.R., Prosky, S.J., Meunier, S.F., Parentau, N.L., (1994) Development of a bilayered living skin construct for clinical applications. *Biotechnology and Bioengineering* 43/8, p. 747-756.

**▼B**

- (22) Marshall, N.J., Goodwin, C.J., Holt, S.J., (1995) A critical assessment of the use of microculture tetrazolium assays to measure cell growth and function. *Growth Regulation* 5, p. 69-84.
- (23) Fentem, J.H., Briggs, D., Chesne', C., Elliot, G.R., Harbell, J.W., Heylings, J.R., Portes, P., Rouget, R., and van de Sandt, J.J.M., and Botham, P.A., (2001) A prevalidation study on *in vitro* tests for acute skin irritation: results and evaluation by the Management Team. *Toxicology In Vitro* 15, p. 57-93.

**▼B****B.41. IN VITRO 3T3 NRU PHOTOTOXICITY TEST****1. METHOD**

This method is equivalent to OECD TG 432 (2004).

**1.1. INTRODUCTION**

Phototoxicity is defined as a toxic response from a substance applied to the body which is either elicited or increased (apparent at lower dose levels) after subsequent exposure to light, or that is induced by skin irradiation after systemic administration of a substance.

The *in vitro* 3T3 NRU phototoxicity test is used to identify the phototoxic potential of a test substance induced by the excited chemical after exposure to light. The test evaluates photo-cytotoxicity by the relative reduction in viability of cells exposed to the chemical in the presence versus absence of light. Substances identified by this test are likely to be phototoxic *in vivo* following systemic application and distribution to the skin, or after topical application.

Many types of chemicals have been reported to induce phototoxic effects (1)(2)(3)(4). Their common feature is their ability to absorb light energy within the sunlight range. According to the first law of photochemistry (Grothaus-Draper Law), photoreaction requires sufficient absorption of light quanta. Thus, before biological testing is considered, a UV/vis absorption spectrum of the test chemical must be determined according to OECD Test Guideline 101. It has been suggested that if the molar extinction/absorption coefficient is less than  $10 \text{ litre} \times \text{mol}^{-1} \times \text{cm}^{-1}$  the chemical is unlikely to be photoreactive. Such chemical may not need to be tested in the *in vitro* 3T3 NRU phototoxicity test or any other biological test for adverse photochemical effects (1)(5). See also Appendix 1.

The reliability and relevance of the *in vitro* 3T3 NRU phototoxicity test was recently evaluated (6)(7)(8) (9). The *in vitro* 3T3 NRU phototoxicity test was shown to be predictive of acute phototoxicity effects in animals and humans *in vivo*. The test is not designed to predict other adverse effects that may arise from combined action of a chemical and light, e.g. it does not address photogenotoxicity, photoallergy, or photocarcinogenicity, nor does it allow an assessment of phototoxic potency. In addition, the test has not been designed to address indirect mechanisms of phototoxicity, effects of metabolites of the test substance, or effects of mixtures.

Whereas the use of metabolising systems is a general requirement for all *in vitro* tests for the prediction of genotoxic and carcinogenic potential, up to now, in the case of phototoxicology, there are only rare examples where metabolic transformation is needed for the chemical to act as a phototoxin *in vivo* or *in vitro*. Thus, it is neither considered necessary nor scientifically justified for the present test to be performed with a metabolic activation system.

**▼ B**

## 1.2. DEFINITIONS

**Irradiance:** the intensity of ultraviolet (UV) or visible light incident on a surface, measured in  $\text{W}/\text{m}^2$  or  $\text{mW}/\text{cm}^2$ .

**Dose of light:** the quantity (= intensity  $\times$  time) of ultraviolet (UV) or visible radiation incident on a surface, expressed in Joules (=  $\text{W} \times \text{s}$ ) per surface area, e.g.  $\text{J}/\text{m}^2$  or  $\text{J}/\text{cm}^2$ .

**UV light wavebands:** the designations recommended by the CIE (Commission Internationale de L'Eclairage) are: UVA (315-400 nm), UVB (280-315 nm) and UVC (100-280 nm). Other designations are also used; the division between UVB and UVA is often placed at 320 nm, and the UVA may be divided into UV-A1 and UV-A2 with a division made at about 340 nm.

**Cell viability:** parameter measuring total activity of a cell population (e.g. uptake of the vital dye Neutral Red into cellular lysosomes), which, depending on the endpoint measured and the test design used, correlates with the total number and/or vitality of the cells.

**Relative cell viability:** cell viability expressed in relation of solvent (negative) controls which have been taken through the whole test procedure (either +Irr or -Irr) but not treated with test chemical.

**PIF (Photo-Irritation-Factor):** factor generated by comparing two equally effective cytotoxic concentrations ( $\text{IC}_{50}$ ) of the test chemical obtained in the absence (-Irr) and in the presence (+Irr) of a non-cytotoxic irradiation with UVA/vis light.

**$\text{IC}_{50}$ :** the concentration of the test chemical by which the cell viability is reduced by 50 %.

**MPE (Mean-Photo-Effect):** measurement derived from mathematical analysis of the concentration response curves obtained in the absence (-Irr) and in the presence (+Irr) of a non-cytotoxic irradiation with UVA/vis light.

**Phototoxicity:** acute toxic response that is elicited after the first exposure of skin to certain chemicals and subsequent exposure to light, or that is induced similarly by skin irradiation after systemic administration of a chemical.

## 1.3. PRINCIPLE OF THE TEST METHOD

The *in vitro* 3T3 NRU phototoxicity test is based on a comparison of the cytotoxicity of a chemical when tested in the presence and in the absence of exposure to a non-cytotoxic dose of simulated solar light. Cytotoxicity in this test is expressed as a concentration-dependent reduction of the uptake of the vital dye Neutral Red when measured 24 hours after treatment with the test chemical and irradiation (10). NR is a weak cationic dye that readily penetrates cell membranes by non-diffusion, accumulating intracellularly in lysosomes. Alterations of the surface of the sensitive lysosomal membrane lead to lysosomal fragility and other changes that gradually become irreversible. Such changes brought about by the action of xenobiotics result in a decreased uptake and binding of NR. It is thus possible to distinguish between viable, damaged or dead cells, which is the basis of this test.

**▼ B**

Balb/c 3T3 cells are maintained in culture for 24 h for formation of monolayers. Two 96-well plates per test chemical are pre-incubated with eight different concentrations of the test substance for 1 h. Thereafter one of the two plates is exposed to the highest non-cytotoxic irradiation dose whereas the other plate is kept in the dark. In both plates the treatment medium is then replaced by culture medium and after another 24 h of incubation cell viability is determined by Neutral Red uptake. Cell viability is expressed as percentage of untreated solvent controls and is calculated for each test concentration. To predict the phototoxic potential, the concentration responses obtained in the presence and in the absence of irradiation are compared, usually at the IC<sub>50</sub> level, i.e., the concentration reducing cell viability to 50 % compared to the untreated controls.

#### 1.4. DESCRIPTION OF THE TEST METHOD

##### 1.4.1. Preparations

###### 1.4.1.1. Cells

A permanent mouse fibroblast cell line, Balb/c 3T3, clone 31, either from the American Type Culture Collection (ATCC), Manassas, VA, USA, or from the European Collection of Cell Cultures (ECACC), Salisbury, Wiltshire, UK, was used in the validation study, and therefore is recommended to obtain from a well qualified cell depository. Other cells or cell lines may be used with the same test procedure if culture conditions are adapted to the specific needs of the cells, but equivalency must be demonstrated.

Cells should be checked regularly for the absence of mycoplasma contamination and only used if none is found (11).

It is important that UV sensitivity of the cells is checked regularly according to the quality control procedure described in this method. Because the UVA sensitivity of cells may increase with the number of passages, Balb/c 3T3 cells of the lowest obtainable passage number, preferably less than 100, should be used. (See Section 1.4.2.2.2 and Appendix 2).

###### 1.4.1.2. Media and culture conditions

Appropriate culture media and incubation conditions should be used for routine cell passage and during the test procedure, e.g. for Balb/c 3T3 cells these are DMEM (Dulbecco's Modified Eagle's Medium) supplemented with 10 % new-born calf serum, 4 mM glutamine, penicillin (100 IU), and streptomycin (100 µg/mL), and humidified incubation at 37 °C, 5-7,5 % CO<sub>2</sub> depending on the buffer (See Section 1.4.1.4, second paragraph.). It is particularly important that cell culture conditions assure a cell cycle time within the normal historical range of the cells or cell line used.

###### 1.4.1.3. Preparation of cultures

Cells from frozen stock cultures are seeded in culture medium at an appropriate density and subcultured at least once before they are used in the *in vitro* 3T3 NRU phototoxicity test.



**▼ B**

Cells used for the phototoxicity test are seeded in culture medium at the appropriate density so that cultures will not reach confluence by the end of the test, i.e., when cell viability is determined 48 h after seeding of the cells. For Balb/c 3T3 cells grown in 96-well plates, the recommended cell seeding density is  $1 \times 10^4$  cells per well.

For each test chemical cells are seeded identically in two separate 96-well plates, which are then taken concurrently through the entire test procedure under identical culture conditions except for the time period where one of the plates is irradiated (+Irr) and the other one is kept in the dark (-Irr).

#### 1.4.1.4. *Preparation of test substance*

Test substances must be prepared fresh, immediately prior to use unless data demonstrate their stability in storage. It is recommended that all chemical handling and the initial treatment of cells be performed under light conditions that would avoid photoactivation or degradation of the test substance prior to irradiation.

Test chemicals shall be dissolved in buffered salt solutions, e.g. Earle's Balanced Salt Solution (EBSS), or other physiologically balanced buffer solutions, which must be free from protein components, light absorbing components (e.g. pH-indicator colours and vitamins) to avoid interference during irradiation. Since during irradiation cells are kept for about 50 minutes outside of the CO<sub>2</sub> incubator, care has to be taken to avoid alkalinisation. If weak buffers like EBSS are used this can be achieved by incubating the cells at 7,5 % CO<sub>2</sub>. If the cells are incubated at 5 % CO<sub>2</sub> only, a stronger buffer should be selected.

Test chemicals of limited solubility in water should be dissolved in an appropriate solvent. If a solvent is used it must be present at a constant volume in all cultures, i.e. in the negative (solvent) controls as well as in all concentrations of the test chemical, and be noncytotoxic at that concentration. Test chemical concentrations should be selected so as to avoid precipitate or cloudy solutions.

Dimethylsulphoxide (DMSO) and ethanol (ETOH) are the recommended solvents. Other solvents of low cytotoxicity may be appropriate. Prior to use, all solvents should be assessed for specific properties, e.g. reaction with the test chemical, quenching of the phototoxic effect, radical scavenging properties and/or chemical stability in the solvent.

Vortex mixing and/or sonication and/or warming to appropriate temperatures may be used to aid solubilisation unless this would affect the stability of the test chemical.

**▼ B**1.4.1.5. *Irradiation conditions*

## 1.4.1.5.1. Light source

The choice of an appropriate light source and filters is a crucial factor in phototoxicity testing. Light of the UVA and visible regions is usually associated with phototoxic reactions *in vivo* (3)(12), whereas generally UVB is of less relevance but is highly cytotoxic; the cytotoxicity increases 1 000-fold as the wavelength goes from 313 to 280 nm (13). Criteria for the choice of an appropriate light source must include the requirement that the light source emits wavelengths absorbed by the test chemical (absorption spectrum) and that the dose of light (achievable in a reasonable exposure time) should be sufficient for the detection of known photocytotoxic chemicals. Furthermore, the wavelengths and doses employed should not be unduly deleterious to the test system, e.g. the emission of heat (infrared region).

Simulation of sunlight with solar simulators is considered the optimal artificial light source. The irradiation power distribution of the filtered solar simulator should be close to that of outdoor daylight given in (14). Both, Xenon arcs and (doped) mercury-metal halide arcs are used as solar simulators (15). The latter has the advantage of emitting less heat and being cheaper, but the match to sunlight is less perfect compared to that of xenon arcs. Because all solar simulators emit significant quantities of UVB they should be suitably filtered to attenuate the highly cytotoxic UVB wavelengths. Because cell culture plastic materials contain UV stabilisers the spectrum should be measured through the same type of 96-well plate lid as will be used in the assay. Irrespective of measures taken to attenuate parts of the spectrum by filtering or by unavoidable filter effects of the equipment the spectrum recorded below these filters should not deviate from standardised outdoor daylight (14). An example of the spectral irradiance distribution of the filtered solar simulator used in the validation study of the *in vitro* 3T3 NRU phototoxicity test is given in (8)(16). See also Appendix 2 Figure 1.

## 1.4.1.5.2. Dosimetry

The intensity of light (irradiance) should be regularly checked before each phototoxicity test using a suitable broadband UV-meter. The intensity should be measured through the same type of 96-well plate lid as will be used in the assay. The UV-meter must have been calibrated to the source. The performance of the UV-meter should be checked, and for this purpose the use of a second, reference UV-meter of the same type and identical calibration is recommended. Ideally, at greater intervals, a spectroradiometer should be used to measure the spectral irradiance of the filtered light source and to check the calibration of the broadband UV-meter.

A dose of 5 J/cm<sup>2</sup> (as measured in the UVA range) was determined to be non-cytotoxic to Balb/c 3T3 cells and sufficiently potent to excite chemicals to elicit phototoxic reactions, (6) (17) e.g. to achieve 5 J/cm<sup>2</sup> within a time period of 50 min, irradiance was adjusted to 1,7 mW/cm<sup>2</sup>. See Appendix 2 Figure 2. If another cell line or a different light source are used, the irradiation dose may have to be calibrated so that a dose regimen can be selected that is not deleterious to the cells but sufficient to excite standard photo-toxins. The time of light exposure is calculated in the following way:

**▼B**

$$t(\text{min}) = \frac{\text{irradiation dose } (J/cm^2) \times 1000}{\text{irradiance } (mW/cm^2) \times 60} \quad (1 \text{ J} = 1 \text{ Wsec})$$

**1.4.2. Test conditions****1.4.2.1. Test substance concentrations**

The ranges of concentrations of a chemical tested in the presence (+Irr) and in the absence (-Irr) of light should be adequately determined in dose range-finding experiments. It may be useful to assess solubility initially and at 60 min (or whatever treatment time is to be used), as solubility can change during time or during the course of exposure. To avoid toxicity induced by improper culture conditions or by highly acidic or alkaline chemicals, the pH of the cell cultures with added test chemical should be in the range 6,5 - 7,8.

The highest concentration of the test substance should be within physiological test conditions, e.g. osmotic and pH stress should be avoided. Depending on the test chemical, it may be necessary to consider other physico-chemical properties as factors limiting the highest test concentration. For relatively insoluble substances that are not toxic at concentrations up to the saturation point the highest achievable concentration should be tested. In general, precipitation of the test chemical at any of the test concentrations should be avoided. The maximum concentration of a test substance should not exceed 1 000 µg/mL; osmolarity should not exceed 10 mmolar. A geometric dilution series of eight test substance concentrations with a constant dilution factor should be used (See Section 2.1, second paragraph).

If there is information (from a range finding experiment) that the test chemical is not cytotoxic up to the limit concentration in the dark experiment (-Irr), but is highly cytotoxic when irradiated (+Irr), the concentration ranges to be selected for the (+Irr) experiment may differ from those selected for the (-Irr) experiment to fulfill the requirement of adequate data quality.

**1.4.2.2. Controls****1.4.2.2.1. Radiation sensitivity of the cells, establishing of historical data:**

Cells should be checked regularly (about every fifth passage) for sensitivity to the light source by assessing their viability following exposure to increasing doses of irradiation. Several doses of irradiation, including levels substantially greater than those used for the 3T3 NRU Phototoxicity test should be used in this assessment. These doses are easiest quantitated by measurements of UV parts of the light source. Cells are seeded at the density used in the *in vitro* 3T3 NRU phototoxicity test and irradiated the next day. Cell viability is then determined one day later using Neutral Red uptake. It should be demonstrated that the resulting highest non-cytotoxic dose (e.g. in the validation study: 5 J/cm<sup>2</sup> [UVA]) was sufficient to classify the reference chemicals (Table 1) correctly.

**▼B**

## 1.4.2.2.2. Radiation sensitivity, check of current test:

The test meets the quality criteria if the irradiated negative/solvent controls show a viability of more than 80 % when compared with non-irradiated negative/solvent.

## 1.4.2.2.3. Viability of solvent controls:

The absolute optical density (OD<sub>540 NRU</sub>) of the Neutral Red extracted from the solvent controls indicates whether the  $1 \times 10^4$  cells seeded per well have grown with a normal doubling time during the two days of the assay. A test meets the acceptance criteria if the mean OD<sub>540 NRU</sub> of the untreated controls is  $\geq 0,4$  (i.e. approximately 20 times the background solvent absorbance).

## 1.4.2.2.4. Positive control:

A known phototoxic chemical shall be tested concurrently with each *in vitro* 3T3 NRU phototoxicity test. Chlorpromazine (CPZ) is recommended. For CPZ tested with the standard protocol in the *in vitro* 3T3 NRU phototoxicity test, the following test acceptance criteria were defined: CPZ irradiated (+Irr): IC<sub>50</sub> = 0,1 to 2,0 µg/ml, CPZ non-irradiated (-Irr): IC<sub>50</sub> = 7,0 to 90,0 µg/mL. The Photo Irritation Factor (PIF), should be  $> 6$ . The historical performance of the positive control should be monitored.

Other phototoxic chemicals, suitable for the chemical class or solubility characteristics of the chemical being evaluated, may be used as the concurrent positive controls in place of chlorpromazine.

## 1.4.3. Test procedure (6)(7)(8)(16)(17):

## 1.4.3.1. 1st day:

Dispense 100 µL culture medium into the peripheral wells of a 96-well tissue culture microtiter plate (= blanks). In the remaining wells, dispense 100 µL of a cell suspension of  $1 \times 10^5$  cells/mL in culture medium (=  $1 \times 10^4$  cells/well). Two plates should be prepared for each series of individual test substance concentrations, and for the solvent and positive controls.

Incubate cells for 24 h (See Section 1.4.1.2) until they form a half confluent monolayer. This incubation period allows for cell recovery, adherence, and exponential growth.

## 1.4.3.2. 2nd day:

After incubation, decant culture medium from the cells and wash carefully with 150 µL of the buffered solution used for incubation. Add 100 µL of the buffer containing the appropriate concentration of test chemical or solvent (solvent control). Apply eight different concentrations of the test chemical. Incubate cells with the test substance in the dark for 60 minutes (See Section 1.4.1.2 and 1.4.1.4 second paragraph).

From the two plates prepared for each series of test substance concentrations and the controls, one is selected, generally at random, for the determination of cytotoxicity (-Irr) (i.e., the control plate), and one (the treatment plate) for the determination of photocytotoxicity (+Irr).

**▼B**

To perform the +Irr exposure, irradiate the cells at room temperature for 50 minutes through the lid of the 96-well plate with the highest dose of radiation that is non-cytotoxic (see also Appendix 2). Keep non-irradiated plates (-Irr) at room temperature in a dark box for 50 min (= light exposure time).

Decant test solution and carefully wash twice with 150 µL of the buffered solution used for incubation, but not containing the test material. Replace the buffer with culture medium and incubate (See Section 1.4.1.2.) overnight (18-22 h).

1.4.3.3. *3rd day:*

1.4.3.3.1. **Microscopic evaluation**

Cells should be examined for growth, morphology, and integrity of the monolayer using a phase contrast microscope. Changes in cell morphology and effects on cell growth should be recorded.

1.4.3.3.2. **Neutral Red uptake test**

Wash the cells with 150 µL of the pre-warmed buffer. Remove the washing solution by gentle tapping. Add 100 µL of a 50 µg/mL Neutral Red (NR) (3-amino-7-dimethylamino-2-methylphenazine hydrochloride, EINECS number 209-035-8; CAS number 553-24-2; C.I. 50040) in medium without serum (16) and incubate as described in paragraph 1.4.1.2., for 3 h. After incubation, remove the NR medium, and wash cells with 150 µL of the buffer. Decant and remove excess buffer by blotting or centrifugation.

Add exactly 150 µL NR desorb solution (freshly prepared 49 parts water + 50 parts ethanol + 1 part acetic acid).

Shake the microtiter plate gently on a microtiter plate shaker for 10 min until NR has been extracted from the cells and has formed a homogeneous solution.

Measure the optical density of the NR extract at 540 nm in a spectrophotometer, using blanks as a reference. Save data in an appropriate electronic file format for subsequent analysis.

2. **DATA**

2.1. **QUALITY AND QUANTITY OF DATA**

The test data should allow a meaningful analysis of the concentration-response obtained in the presence and in the absence of irradiation, and if possible the concentration of test chemical by which cell viability is reduced to 50 % (IC<sub>50</sub>). If cytotoxicity is found, both the concentration range and the intercept of individual concentrations shall be set in a way to allow the fit of a curve to the experimental data.

For both clearly positive and clearly negative results (See Section 2.3, first paragraph), the primary experiment, supported by one or more preliminary dose range-finding experiment(s), may be sufficient.

**▼ B**

Equivocal, borderline, or unclear results should be clarified by further testing (see also section 2.4, second paragraph). In such cases, modification of experimental conditions should be considered. Experimental conditions that might be modified include the concentration range or spacing, the pre-incubation time, and the irradiation-exposure time. A shorter exposure time may be appropriate for water-unstable chemicals.

## 2.2. EVALUATION OF RESULTS

To enable evaluation of the data, a Photo-Irritation-Factor (PIF) or Mean Photo Effect (MPE) may be calculated.

For the calculation of the measures of photocytotoxicity (see below) the set of discrete concentration-response values has to be approximated by an appropriate continuous concentration-response curve (model). Fitting of the curve to the data is commonly performed by a non-linear regression method (18). To assess the influence of data variability on the fitted curve a bootstrap procedure is recommended.

A Photo-Irritation-Factor (PIF) is calculated using the following formula:

$$\text{PIF} = \frac{\text{IC}_{50}(-\text{Irr})}{\text{IC}_{50}(+\text{Irr})}$$

If an  $\text{IC}_{50}$  in the presence or absence of light cannot be calculated, a PIF cannot be determined for the test material. The mean photo effect (MPE) is based on comparison of the complete concentration-response curves (19). It is defined as the weighted average across a representative set of photo effect values

$$\text{MPE} = \frac{\sum_{i=1}^n w_i \text{PE}_{C_i}}{\sum_{i=1}^n w_i}$$

The photo effect  $\text{PE}_c$  at any concentration  $C$  is defined as the product of the response effect  $\text{RE}_c$  and the dose effect  $\text{DE}_c$  i.e.  $\text{PE}_c = \text{RE}_c \times \text{DE}_c$ . The response effect  $\text{RE}_c$  is the difference between the responses observed in the absence and presence of light, i.e.  $\text{RE}_c = R_c(-\text{Irr}) - R_c(+\text{Irr})$ . The dose-effect is given by

$$\text{DE}_c = \left| \frac{C/C^* - 1}{C/C^* + 1} \right|$$

where  $C^*$  represents the equivalence concentration, i.e. the concentration at which the  $+\text{Irr}$  response equals the  $-\text{Irr}$  response at concentration  $C$ . If  $C^*$  cannot be determined because the response values of the  $+\text{Irr}$  curve are systematically higher or lower than  $R_c(-\text{Irr})$  the dose effect is set to 1. The weighting factors  $w_i$  are given by the highest response value, i.e.  $w_i = \text{MAX} \{R_i(+\text{Irr}), R_i(-\text{Irr})\}$ . The concentration grid  $C_i$  is chosen such that the same number of points falls into each of the concentration intervals defined by the concentration values used in the experiment. The calculation of MPE is restricted to the maximum concentration value at which at least one of the two curves still exhibits a response value of at least 10 %. If this maximum concentration is higher than the highest concentration used in the  $+\text{Irr}$  experiment the residual part of the  $+\text{Irr}$  curve is set to the response value '0'. Depending on whether the MPE value is larger than a properly chosen cut-off value ( $\text{MPE}_c = 0,15$ ) or not, the chemical is classified as phototoxic.

**▼B**

A software package for the calculation of the PIF and MPE is available from (20).

### 2.3. INTERPRETATION OF RESULTS

Based on the validation study (8), a test substance with a PIF < 2 or an MPE < 0,1 predicts: 'no phototoxicity'. A PIF > 2 and < 5 or an MPE > 0,1 and < 0,15 predicts: 'probable phototoxicity'; and a PIF > 5 or an MPE > 0,15 predicts: 'phototoxicity'.

For any laboratory initially establishing this assay, the reference materials listed in Table 1 should be tested prior to the testing of test substances for phototoxic assessment. PIF or MPE values should be close to the values mentioned in Table 1.

**Table 1**

Chemical name	EINECS No	CAS No	PIF	MPE	Absorption peak	Solvent <sup>(1)</sup>
Amiodarone HCL	243-293-2	[19774-82-4]	> 3,25	0,2-0,54	242 nm 300 nm (shoulder)	ethanol
Chloropromazine HCL	200-701-3	[69-09-0]	> 14,4	0,33-0,63	309 nm	ethanol
Norfloxacin	274-614-4	[70458-96-7]	> 71,6	0,34-0,90	316 nm	acetonitrile
Anthracene	204-371-1	[120-12-7]	> 18,5	0,19-0,81	356 nm	acetonitrile
Protoporphyrin IX, Disodium	256-815-9	[50865-01-5]	> 45,3	0,54-0,74	402 nm	ethanol
L-Histidine		[7006-35-1]	no PIF	0,05-0,10	211 nm	water
Hexachlorophene	200-733-8	[70-30-4]	1,1-1,7	0,00-0,05	299 nm 317 nm (shoulder)	ethanol
Sodium lauryl sulphate	205-788-1	[151-21-3]	1,0-1,9	0,00-0,05	no absorption	water

<sup>(1)</sup> Solvent used for measuring absorption.

### 2.4. INTERPRETATION OF DATA

If phototoxic effects are observed only at the highest test concentration, (especially for water soluble test chemicals) additional considerations may be necessary for assessment of hazard. These may include data on skin absorption, and accumulation of the chemical in the skin and/or data from other tests, e.g. testing of the chemical in *in vitro* animal or human skin, or skin models.

**▼B**

If no toxicity is demonstrated (+Irr and -Irr), and if poor solubility limited the concentrations that could be tested, then the compatibility of the test substance with the assay may be questioned and confirmatory testing should be considered using, e.g. another model.

**3. REPORTING****TEST REPORT**

The test report must include at least the following information:

Test substance:

- identification data, common generic names and IUPAC and CAS number, if known,
- physical nature and purity,
- physicochemical properties relevant to conduct of the study,
- UV/vis absorption spectrum,
- stability and photostability, if known.

Solvent:

- justification for choice of solvent,
- solubility of the test chemical in solvent,
- percentage of solvent present in treatment medium.

Cells:

- type and source of cells,
- absence of mycoplasma,
- cell passage number, if known,
- Radiation sensitivity of cells, determined with the irradiation equipment used in the *in vitro* 3T3 NRU phototoxicity test.

Test conditions (1); *incubation before and after treatment*:

- type and composition of culture medium,
- incubation conditions (CO<sub>2</sub> concentration; temperature; humidity),
- duration of incubation (pre-treatment; post-treatment).

Test conditions (2); *treatment with the chemical*:

- rationale for selection of concentrations of the test chemical used in the presence and in the absence of irradiation,
- in case of limited solubility of the test chemical and absence of cytotoxicity: rationale for the highest concentration tested,
- type and composition of treatment medium (buffered salt solution),
- duration of the chemical treatment.

Test conditions (3); *irradiation*:

- rationale for selection of the light source used,



**▼ B**

- manufacturer and type of light source and radiometer,
- spectral irradiance characteristics of the light source,
- transmission and absorption characteristics of the filter(s) used,
- characteristics of the radiometer and details on its calibration,
- distance of the light source from the test system,
- UVA irradiance at this distance, expressed in  $\text{mW}/\text{cm}^2$ ,
- duration of the UV/vis light exposure,
- UVA dose (irradiance  $\times$  time), expressed in  $\text{J}/\text{cm}^2$ ,
- temperature of cell cultures during irradiation and cell cultures concurrently kept in the dark.

Test conditions (4); *Neutral Red viability test*:

- composition of Neutral Red treatment medium,
- duration of Neutral Red incubation,
- incubation conditions ( $\text{CO}_2$  concentration; temperature; humidity),
- Neutral Red extraction conditions (extractant; duration),
- wavelength used for spectrophotometric reading of Neutral Red optical density,
- second wavelength (reference), if used,
- content of spectrophotometer blank, if used.

Results:

- cell viability obtained at each concentration of the test chemical, expressed in percent viability of mean, concurrent solvent controls,
- concentration response curves (test chemical concentration vs. relative cell viability) obtained in concurrent +Irr and -Irr experiments,
- analysis of the concentration-response curves: if possible, computation/calculation of  $\text{IC}_{50}$  (+Irr) and  $\text{IC}_{50}$  (-Irr),
- comparison of the two concentration response curves obtained in the presence and in the absence of irradiation, either by calculation of the Photo-Irritation-Factor (PIF), or by calculation of the Mean-Photo-Effect (MPE),
- test acceptance criteria; concurrent solvent control:
- absolute viability (optical density of Neutral Red extract) of irradiated and non-irradiated cells,
- historic negative and solvent control data; means and standard deviations,
- test acceptance criteria; concurrent positive control,

**▼B**

- IC<sub>50</sub>(+Irr) and IC<sub>50</sub>(-Irr) and PIF/MPE of positive control chemical,
- historic positive control chemical data: IC<sub>50</sub>(+Irr) and IC<sub>50</sub>(-Irr) and PIF/MPE; means and standard deviations.

Discussion of the results.

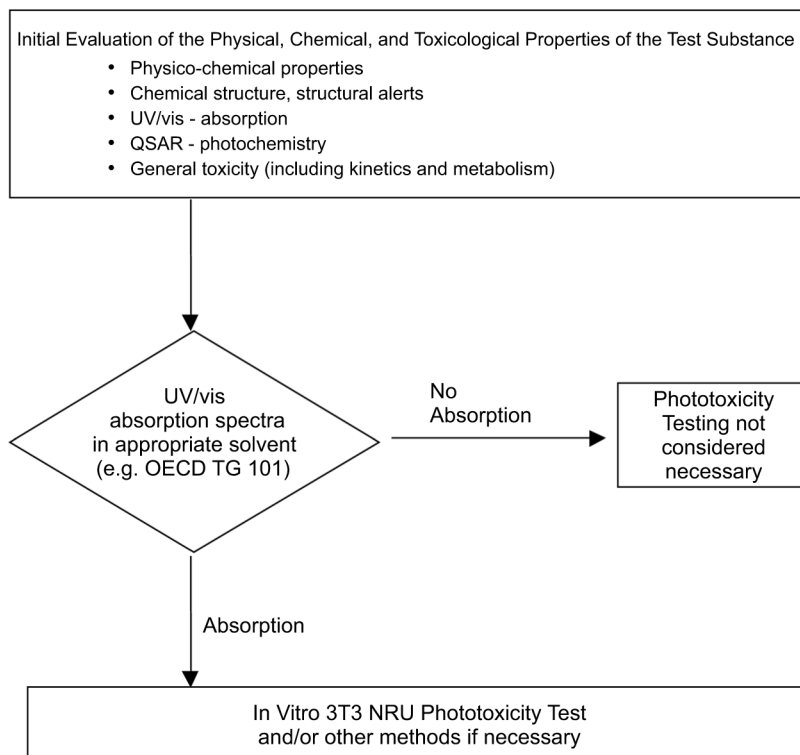
Conclusions.

#### 4. REFERENCES

- (1) Lovell W.W., (1993) A scheme for *in vitro* screening of substances for photoallergenic potential. *Toxicology In Vitro* 7, p. 95-102.
- (2) Santamaria, L. and Prino, G., (1972) List of the photodynamic substances. In 'Research Progress in Organic, Biological and Medicinal Chemistry' Vol. 3 part 1. North Holland Publishing Co. Amsterdam. p. XI-XXXV.
- (3) Spielmann, H., Lovell, W.W., Hölzle, E., Johnson, B.E., Maurer, T., Miranda, M.A., Pape, W.J.W., Sapora, O., and Sladowski, D., (1994) *In vitro* phototoxicity testing: The report and recommendations of ECVAM Workshop 2. *ATLA*, 22, p. 314-348.
- (4) Spikes, J.D., (1989) Photosensitisation. In 'The science of Photobiology' Edited by K.C. Smith. Plenum Press, New York. 2nd edition, p. 79-110.
- (5) OECD, (1997) Environmental Health and Safety Publications, Series on Testing and Assessment No 7 'Guidance Document On Direct Phototransformation Of Chemicals In Water' Environment Directorate, OECD, Paris.
- (6) Spielmann, H., Balls, M., Döring, B., Holzhütter, H.G., Kalweit, S., Klecak, G., L'Eplattenier, H., Liebsch, M., Lovell, W.W., Maurer, T., Moldenhauer, F. Moore, L., Pape, W., Pfannbecker, U., Potthast, J., De Silva, O., Steiling, W., and Willshaw, A., (1994) EEC/COLIPA project on *in vitro* phototoxicity testing: First results obtained with a Balb/c 3T3 cell phototoxicity assay. *Toxic. In Vitro* 8, p. 793-796.
- (7) Anon, (1998) Statement on the scientific validity of the 3T3 NRU PT test (an *in vitro* test for phototoxicity), European Commission, Joint Research Centre: ECVAM and DGXI/E/2, 3 November 1997, *ATLA*, 26, p. 7-8.
- (8) Spielmann, H., Balls, M., Dupuis, J., Pape, W.J.W., Pechovitch, G. De Silva, O., Holzhütter, H.G., Clothier, R., Desolle, P., Gerberick, F., Liebsch, M., Lovell, W.W., Maurer, T., Pfannenbecker, U., Potthast, J. M., Csato, M., Sladowski, D., Steiling, W., and Brantom, P., (1998) The international EU/COLIPA *In vitro* phototoxicity validation study: results of phase II (blind trial), part 1: the 3T3 NRU phototoxicity test. *Toxicology In Vitro* 12, p. 305-327.
- (9) OECD, (2002) Extended Expert Consultation Meeting on The *In Vitro* 3T3 NRU Phototoxicity Test Guideline Proposal, Berlin, 30<sup>th</sup>-31<sup>th</sup> October 2001, Secretariat's Final Summary Report, 15<sup>th</sup> March 2002, OECD ENV/EHS, available upon request from the Secretariat.
- (10) Borenfreund, E., and Puerner, J.A., (1985) Toxicity determination *in vitro* by morphological alterations and neutral red absorption. *Toxicology Lett.*, 24, p. 119-124.
- (11) Hay, R.J., (1988) The seed stock concept and quality control for cell lines. *Analytical Biochemistry* 171, p. 225-237.

**▼ B**

- (12) Lambert L.A., Warner W.G., and Kornhauser A., (1996) Animal models for phototoxicity testing. In 'Dermatotoxicology', edited by F.N. Marzulli and H.I. Maibach. Taylor & Francis, Washington DC. 5th Edition, p. 515-530.
- (13) Tyrrell R.M., Pidoux M., (1987) Action spectra for human skin cells: estimates of the relative cytotoxicity of the middle ultraviolet, near ultraviolet and violet regions of sunlight on epidermal keratinocytes. *Cancer Res.*, 47, p. 1825-1829.
- (14) ISO 10977., (1993) Photography — Processed photographic colour films and paper prints — Methods for measuring image stability.
- (15) Sunscreen Testing (UV.B) TECHNICALREPORT, CIE, International Commission on Illumination, Publication No 90, Vienna, 1993, ISBN 3 900 734 275
- (16) ZEBET/ECVAM/COLIPA — Standard Operating Procedure: *In Vitro* 3T3 NRU Phototoxicity Test. Final Version, 7 September, 1998. p. 18.
- (17) Spielmann, H., Balls, M., Dupuis, J., Pape, W.J.W., De Silva, O., Holzhütter, H.G., Gerberick, F., Liebsch, M., Lovell, W.W., and Pfannenbecker, U., (1998) A study on UV filter chemicals from Annex VII of the European Union *Directive 76/768/EEC*, in the *in vitro* 3T3 NRU phototoxicity test. *ATLA* 26, p. 679-708.
- (18) Holzhütter, H.G., and Quedenau, J., (1995) Mathematical modeling of cellular responses to external signals. *J. Biol. Systems* 3, p. 127-138.
- (19) Holzhütter, H.G., (1997) A general measure of *in vitro* phototoxicity derived from pairs of dose-response curves and its use for predicting the *in vivo* phototoxicity of chemicals. *ATLA*, 25, p. 445-462.
- (20) [http://www.oecd.org/document/55/0,2340,en\\_2649\\_34377\\_2349687\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/55/0,2340,en_2649_34377_2349687_1_1_1_1,00.html)

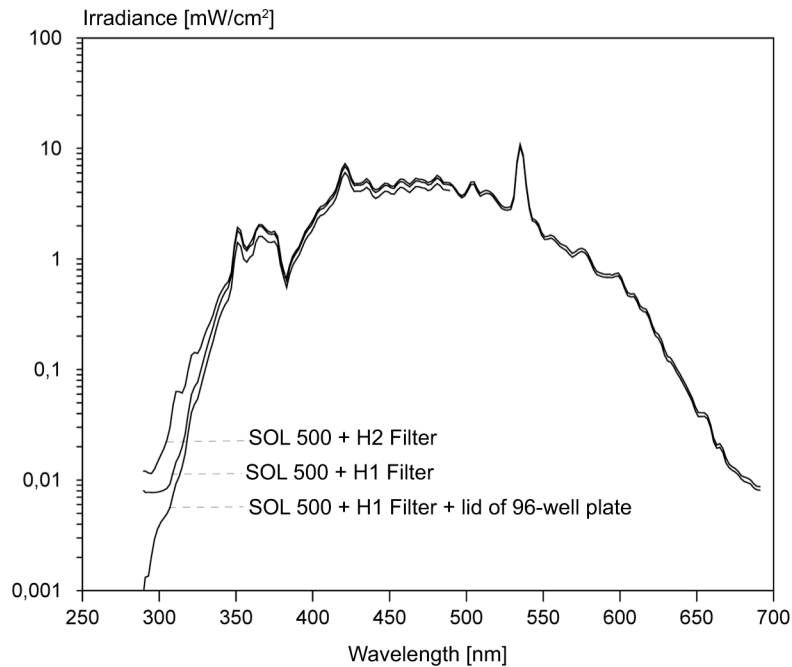
**▼B***Appendix 1***Role of the 3T3 NRU PT in a sequential approach to the phototoxicity testing of chemicals**

▼ B

## Appendix 2

Figure 1

## Spectral power distribution of a filtered solar simulator



(see Section 1.4.1.5, second paragraph)

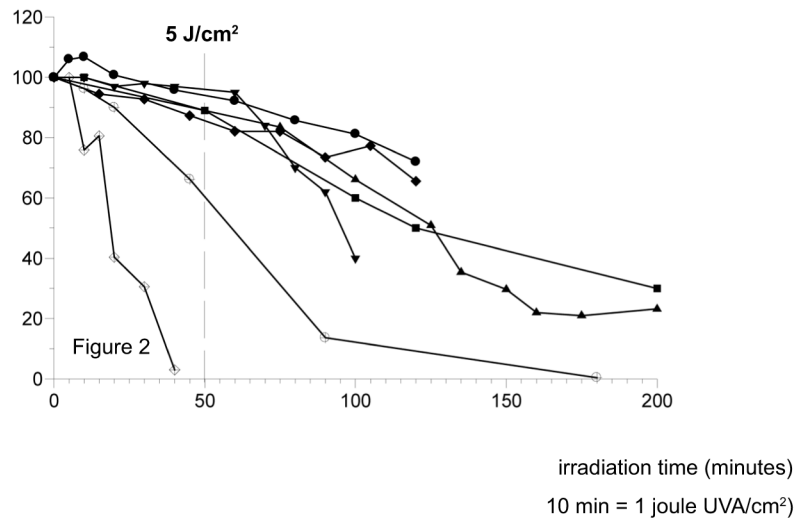
Figure 1 gives an example of an acceptable spectral irradiance distribution of a filtered solar simulator. It is from the doped metal halide source used in the validation trial of the 3T3 NRU PT (6)(8)(17). The effect of two different filters and the additional filtering effect of the lid of a 96-well cell culture plate are shown. The H2 filter was only used with test systems that can tolerate a higher amount of UVB (skin model test and red blood cell photo-haemolysis test). In the 3T3 NRU-PT the H1 filter was used. The figure shows that additional filtering effect of the plate lid is mainly observed in the UVB range, still leaving enough UVB in the irradiation spectrum to excite chemicals typically absorbing in the UVB range, like Amiodarone (see Table 1).

▼B

Figure 2

**Irradiation sensitivity of Balb/c 3T3 cells (as measured in the UVA range)**

Cell viability (% Neutral Red uptake of dark controls)



(see Sections 1.4.1.5.2 second paragraph; 1.4.2.2.1, 1.4.2.2.2)

Sensitivity of Balb/c 3T3 cells to irradiation with the solar simulator used in the validation trial of the 3T3NRU-phototoxicity test, as measured in the UVA range. Figure shows the results obtained in seven different laboratories in the pre-validation study (1). While the two curves with open symbols were obtained with aged cells (high number of passages), that had to be replaced by new cell stocks the curves with bold symbols show cells with acceptable irradiation tolerance.

From these data the highest non-cytotoxic irradiation dose of 5 J/cm<sup>2</sup> was derived (vertical dashed line). The horizontal dashed line shows in addition the maximum acceptable irradiation effect given in paragraph 1.4.2.2.

**▼M3****B.42. SKIN SENSITISATION: LOCAL LYMPH NODE ASSAY****INTRODUCTION**

1. OECD Guidelines for the Testing of Chemicals and EU Test Methods based on them are periodically reviewed in light of scientific progress, changing regulatory needs, and animal welfare considerations. The original Test Method (TM) for the determination of skin sensitisation in the mouse, the Local Lymph Node Assay (LLNA; OECD Test Guideline 429; Chapter B.42 of this Annex) was adopted previously (1). The details of the validation of the LLNA and a review of the associated work have been published (2) (3) (4) (5) (6) (7) (8) (9) (10) (11). The updated LLNA is based on the evaluation of experience and scientific data (12). This is the second TM to be designed for assessing skin sensitisation potential of chemicals (substances and mixtures) in animals. The other TM (*i.e.* OECD Test Guideline 406; Chapter B.6 of this Annex) utilises guinea pig tests, notably the guinea pig maximisation test and the Buehler test (13). The LLNA provides advantages over B.6 and OECD Test Guideline 406 (13) with regard to animal welfare. This updated LLNA TM includes a set of Performance Standards (PS) (Appendix 1) that can be used to evaluate the validation status of new and/or modified test methods that are functionally and mechanistically similar to the LLNA, in accordance with the principles of OECD Guidance Document No 34 (14).
2. The LLNA studies the induction phase of skin sensitisation and provides quantitative data suitable for dose-response assessment. It should be noted that the mild/moderate sensitisers which are recommended as suitable positive control chemicals (PC) for guinea pig test methods (*i.e.* B.6; OECD Test Guideline 406) (13) are also appropriate for use with the LLNA (6) (8) (15). A reduced LLNA (rLLNA) approach, which could use up to 40 % fewer animals is also described as an option in this TM (16) (17) (18). The rLLNA may be used when there is a regulatory need to confirm a negative prediction of skin sensitising potential, provided there is adherence to all other LLNA protocol specifications, as described in this TM. Prediction of a negative outcome should be made based on all available information as described in paragraph 4. Before applying the rLLNA approach, clear justifications and scientific rationale for its use should be provided. If, against expectations, a positive or equivocal result is obtained in the rLLNA, additional testing may be needed in order to interpret or clarify the finding. The rLLNA should not be used for the hazard identification of skin sensitising test substances when dose-response information is needed such as sub-categorisation for Regulation (EC) No 1272/2008 on classification, labelling and packaging of substances and mixtures and UN Globally Harmonised System of Classification and Labelling of Chemicals.

**DEFINITIONS**

3. Definitions used are provided in Appendix 2.

**INITIAL CONSIDERATIONS AND LIMITATIONS**

4. The LLNA provides an alternative method for identifying potential skin sensitising chemicals. This does not necessarily imply that in all instances the LLNA should be used in place of guinea pig tests (*i.e.* B.6; OECD Test Guideline 406) (13), but rather that the assay is of equal merit and may be employed as an alternative in which positive and negative results generally no longer require further confirmation. The testing laboratory should consider all available information on the test substance prior to conducting the study. Such information will include the identity and chemical structure

**▼ M3**

of the test substance; its physicochemical properties; the results of any other *in vitro* or *in vivo* toxicity tests on the test substance; and toxicological data on structurally related chemicals. This information should be considered in order to determine whether the LLNA is appropriate for the substance (given the incompatibility of limited types of chemicals with the LLNA — see paragraph 5) and to aid in dose selection.

5. The LLNA is an *in vivo* method and, as a consequence, will not eliminate the use of animals in the assessment of allergic contact sensitising activity. It has, however, the potential to reduce the number of animals required for this purpose. Moreover, the LLNA offers a substantial refinement (less pain and distress) of the way in which animals are used for allergic contact sensitisation testing. The LLNA is based upon consideration of immunological events stimulated by chemicals during the induction phase of sensitisation. Unlike guinea pig tests (*i.e.* B.6; OECD Test Guideline 406) (13) the LLNA does not require that challenge-induced dermal hypersensitivity reactions be elicited. Furthermore, the LLNA does not require the use of an adjuvant, as is the case for the guinea pig maximisation test (13). Thus, the LLNA reduces animal pain and distress. Despite the advantages of the LLNA over B.6 and OECD Test Guideline 406, it should be recognised that there are certain limitations that may necessitate the use of B.6 or OECD Test Guideline 406 (13) (*e.g.* false negative findings in the LLNA with certain metals, false positive findings with certain skin irritants (such as some surfactant type chemicals) (19) (20), or solubility of the test substance). In addition, chemical classes or substances containing functional groups shown to act as potential confounders (21) may necessitate the use of guinea pig tests (*i.e.* B.6; OECD Test Guideline 406) (13). Further, based on the limited validation database, which consisted primarily of pesticide formulations, the LLNA is more likely than the guinea pig test to yield a positive result for these types of test substances (22). However, when testing formulations, one could consider including similar substances with known results as benchmark substances to demonstrate that the LLNA is functioning properly (see paragraph 16). Other than such identified limitations, the LLNA should be applicable for testing any substances unless there are properties associated with these substances that may interfere with the accuracy of the LLNA.

**PRINCIPLE OF THE TEST**

6. The basic principle underlying the LLNA is that sensitisers induce proliferation of lymphocytes in the lymph nodes draining the site of test substance application. This proliferation is proportional to the dose and to the potency of the applied allergen and provides a simple means of obtaining a quantitative measurement of sensitisation. Proliferation is measured by comparing the mean proliferation in each test group to the mean proliferation in the vehicle treated control (VC) group. The ratio of the mean proliferation in each treated group to that in the concurrent VC group, termed the Stimulation Index (SI), is determined, and should be  $\geq 3$  before classification of the test substance as a potential skin sensitiser is warranted. The procedures described here are based on the use of *in vivo* radioactive labelling to measure an increased number of proliferating cells in the draining auricular lymph nodes. However, other endpoints for assessment of the number of proliferating cells may be employed provided the PS requirements are fully met (Appendix 1).



**▼ M3****DESCRIPTION OF THE ASSAY****Selection of animal species**

7. The mouse is the species of choice for this test. Young adult female mice of CBA/Ca or CBA/J strain, which are nulliparous and non-pregnant, are used. At the start of the study, animals should be between 8-12 weeks old, and the weight variation of the animals should be minimal and not exceed 20 % of the mean weight. Alternatively, other strains and males may be used when sufficient data are generated to demonstrate that significant strain and/or gender-specific differences in the LLNA response do not exist.

**Housing and feeding conditions**

8. Mice should be group-housed (23), unless adequate scientific rationale for housing mice individually is provided. The temperature of the experimental animal room should be  $22 \pm 3$  °C. Although the relative humidity should be at least 30 % and preferably not exceed 70 %, other than during room cleaning, the aim should be 50-60 %. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water.

**Preparation of animals**

9. The animals are randomly selected, marked to permit individual identification (but not by any form of ear marking), and kept in their cages for at least five days prior to the start of dosing to allow for acclimatisation to the laboratory conditions. Prior to the start of treatment all animals are examined to ensure that they have no observable skin lesions.

**Preparation of dosing solutions**

10. Solid chemicals should be dissolved or suspended in solvents/vehicles and diluted, if appropriate, prior to application to an ear of the mice. Liquid chemicals may be applied neat or diluted prior to dosing. Insoluble chemicals, such as those generally seen in medical devices, should be subjected to an exaggerated extraction in an appropriate solvent to reveal all extractable constituents for testing prior to application to an ear of the mice. Test substances should be prepared daily unless stability data demonstrate the acceptability of storage.

**Reliability check**

11. Positive control chemicals (PC) are used to demonstrate appropriate performance of the assay by responding with adequate and reproducible sensitivity as a sensitising test substance for which the magnitude of the response is well characterised. Inclusion of a concurrent PC is recommended because it demonstrates competency of the laboratory to successfully conduct each assay and allows for an assessment of intra- and inter-laboratory reproducibility and comparability. A PC for each study is also required by some regulatory authorities and therefore users are encouraged to consult the relevant authorities prior to conducting the LLNA. Accordingly, the routine use of a concurrent PC is encouraged to avoid the need for additional animal testing to meet such requirements that might arise from the use of a periodic PC (see paragraph 12). The PC should produce a positive LLNA response at an exposure level expected to give an increase in the SI  $> 3$  over the negative control (NC) group. The PC dose should be chosen such that it does not cause excessive skin

**▼ M3**

irritation or systemic toxicity and the induction is reproducible but not excessive (*i.e.* a SI > 20 would be excessive). Preferred PC are 25 % hexyl cinnamic aldehyde (Chemical Abstracts Service (CAS) No 101-86-0) in acetone: olive oil (4:1, v/v) and 5 % mercaptobenzothiazole (CAS No 149-30-4) in *N,N*-dimethylformamide (see Appendix 1, Table 1). There may be circumstances in which, given adequate justification, other PC, meeting the above criteria, may be used.

12. While inclusion of a concurrent PC group is recommended, there may be situations in which periodic testing (*i.e.* at intervals  $\leq$  6 months) of the PC may be adequate for laboratories that conduct the LLNA regularly (*i.e.* conduct the LLNA at a frequency of no less than once per month) and have an established historical PC database that demonstrates the laboratory's ability to obtain reproducible and accurate results with PCs. Adequate proficiency with the LLNA can be successfully demonstrated by generating consistent positive results with the PC in at least 10 independent tests conducted within a reasonable period of time (*i.e.* less than one year).
13. A concurrent PC group should always be included when there is a procedural change to the LLNA (*e.g.* change in trained personnel, change in test method materials and/or reagents, change in test method equipment, change in source of test animals), and such changes should be documented in laboratory reports. Consideration should be given to the impact of these changes on the adequacy of the previously established historical database in determining the necessity for establishing a new historical database to document consistency in the PC results.
14. Investigators should be aware that the decision to conduct a PC study on a periodic basis instead of concurrently has ramifications on the adequacy and acceptability of negative study results generated without a concurrent PC during the interval between each periodic PC study. For example, if a false negative result is obtained in the periodic PC study, negative test substance results obtained in the interval between the last acceptable periodic PC study and the unacceptable periodic PC study may be questioned. Implications of these outcomes should be carefully considered when determining whether to include concurrent PCs or to only conduct periodic PCs. Consideration should also be given to using fewer animals in the concurrent PC group when this is scientifically justified and if the laboratory demonstrates, based on laboratory-specific historical data, that fewer mice can be used (12).
15. Although the PC should be tested in the vehicle that is known to elicit a consistent response (*e.g.* acetone: olive oil; 4:1, v/v), there may be certain regulatory situations in which testing in a non-standard vehicle (clinically/chemically relevant formulation) will also be necessary (24). If the concurrent PC is tested in a different vehicle than the test substance, then a separate VC for the concurrent PC should be included.
16. In instances where test substances of a specific chemical class or range of responses are being evaluated, benchmark substances may also be useful to demonstrate that the test method is functioning properly for detecting the skin sensitisation potential of these types of test substances. Appropriate benchmark substances should have the following properties:
  - structural and functional similarity to the class of the test substance being tested;
  - known physical/chemical characteristics;
  - supporting data from the LLNA;
  - supporting data from other animal models and/or from humans.

**▼ M3****TEST PROCEDURE****Number of animals and dose levels**

17. A minimum of four animals is used per dose group, with a minimum of three concentrations of the test substance, plus a concurrent NC group treated only with the vehicle for the test substance, and a PC (concurrent or recent, based on laboratory policy in considering paragraphs 11-15). Testing multiple doses of the PC should be considered, especially when testing the PC on an intermittent basis. Except for absence of treatment with the test substance, animals in the control groups should be handled and treated in a manner identical to that of animals in the treatment groups.
  
18. Dose and vehicle selection should be based on the recommendations given in references (3) and (5). Consecutive doses are normally selected from an appropriate concentration series such as 100 %, 50 %, 25 %, 10 %, 5 %, 2,5 %, 1 %, 0,5 %, etc. Adequate scientific rationale should accompany the selection of the concentration series used. All existing toxicological information (*e.g.* acute toxicity and dermal irritation) and structural and physicochemical information on the test substance of interest (and/or structurally related substances) should be considered where available, in selecting the three consecutive concentrations so that the highest concentration maximises exposure while avoiding systemic toxicity and/or excessive local skin irritation (3) (25). In the absence of such information, an initial pre-screen test may be necessary (see paragraphs 21-24).
  
19. The vehicle should not interfere with or bias the test result and should be selected on the basis of maximising the solubility in order to obtain the highest concentration achievable while producing a solution/suspension suitable for application of the test substance. Recommended vehicles are acetone: olive oil (4:1, v/v), *N,N*-dimethylformamide, methyl ethyl ketone, propylene glycol, and dimethyl sulphoxide (19) but others may be used if sufficient scientific rationale is provided. In certain situations it may be necessary to use a clinically relevant solvent or the commercial formulation in which the test substance is marketed as an additional control. Particular care should be taken to ensure that hydrophilic test substances are incorporated into a vehicle system, which wets the skin and does not immediately run off, by incorporation of appropriate solubilisers (*e.g.* 1 % Pluronic® L92). Thus, wholly aqueous vehicles are to be avoided.
  
20. The processing of lymph nodes from individual mice allows for the assessment of inter-animal variability and a statistical comparison of the difference between test substance and VC group measurements (see paragraph 35). In addition, evaluating the possibility of reducing the number of mice in the PC group is feasible when individual animal data are collected (12). Further, some regulatory authorities require the collection of individual animal data. Nonetheless, pooled animal data may be considered acceptable by some regulatory authorities and in such situations, users may have the option of collecting either individual or pooled animal data.

**Pre-screen test**

21. In the absence of information to determine the highest dose to be tested (see paragraph 18), a pre-screen test should be performed in order to define the appropriate dose level to test in the LLNA. The purpose of the pre-screen test is to provide guidance for selecting the maximum dose level to use in the main LLNA study, where information on the concentration that induces systemic toxicity (see paragraph 24) and/or excessive local skin irritation (see paragraph 23) is not available. The maximum dose level tested should be 100 % of the test substance for liquids or the maximum possible concentration for solids or suspensions.

▼ **M3**

22. The pre-screen test is conducted under conditions identical to the main LLNA study, except there is no assessment of lymph node proliferation and fewer animals per dose group can be used. One or two animals per dose group are suggested. All mice will be observed daily for any clinical signs of systemic toxicity or local irritation at the application site. Body weights are recorded pre-test and prior to termination (Day 6). Both ears of each mouse are observed for erythema and scored using Table 1 (25). Ear thickness measurements are taken using a thickness gauge (*e.g.* digital micrometer or Peacock Dial thickness gauge) on Day 1 (pre-dose), Day 3 (approximately 48 hours after the first dose), and Day 6. Additionally, on Day 6, ear thickness could be determined by ear punch weight determinations, which should be performed after the animals are humanely killed. Excessive local skin irritation is indicated by an erythema score  $\geq 3$  and/or an increase in ear thickness of  $\geq 25\%$  on any day of measurement (26) (27). The highest dose selected for the main LLNA study will be the next lower dose in the pre-screen concentration series (see paragraph 18) that does not induce systemic toxicity and/or excessive local skin irritation.

*Table 1*  
**Erythema Scores**

Observation	Score
No erythema	0
Very slight erythema (barely perceptible)	1
Well-defined erythema	2
Moderate to severe erythema	3
Severe erythema (beet redness) to eschar formation preventing grading of erythema	4

23. In addition to a 25 % increase in ear thickness (26) (27), a statistically significant increase in ear thickness in the treated mice compared to control mice has also been used to identify irritants in the LLNA (28) (29) (30) (31) (32) (33) (34). However, while statistically significant increases can occur when ear thickness is less than 25 % they have not been associated specifically with excessive irritation (30) (32) (33) (34).
24. The following clinical observations may indicate systemic toxicity (35) (36) when used as part of an integrated assessment and therefore may indicate the maximum dose level to use in the main LLNA: changes in nervous system function (*e.g.* pilo-erection, ataxia, tremors, and convulsions); changes in behaviour (*e.g.* aggressiveness, change in grooming activity, marked change in activity level); changes in respiratory patterns (*i.e.* changes in frequency and intensity of breathing such as dyspnea, gasping, and rales), and changes in food and water consumption. In addition, signs of lethargy and/or unresponsiveness and any clinical signs of more than slight or momentary pain and distress, or a  $> 5\%$  reduction in body weight from Day 1 to Day 6, and mortality should be considered in the evaluation. Moribund animals or animals obviously in pain or showing signs of severe and enduring distress should be humanely killed (37).

**▼ M3****Main study experimental schedule**

25. The experimental schedule of the assay is as follows:
- *Day 1*: Individually identify and record the weight of each animal and any clinical observation. Apply 25 µL of the appropriate dilution of the test substance, the vehicle alone, or the PC (concurrent or recent, based on laboratory policy in considering paragraphs 11-15), to the dorsum of each ear.
  
  - *Days 2 and 3*: Repeat the application procedure carried out on Day 1.
  
  - *Days 4 and 5*: No treatment.
  
  - *Day 6*: Record the weight of each animal. Inject 250 µL of sterile phosphate-buffered saline (PBS) containing 20 µCi ( $7,4 \times 10^5$  Bq) of tritiated ( $^3\text{H}$ )-methyl thymidine into all test and control mice via the tail vein. Alternatively, inject 250 µL sterile PBS containing 2 µCi ( $7,4 \times 10^4$  Bq) of  $^{125}\text{I}$ -iododeoxyuridine and  $10^{-5}\text{M}$  fluorodeoxyuridine into all mice via the tail vein. Five hours (5 h) later, humanely kill the animals. Excise the draining auricular lymph nodes from each mouse ear and process together in PBS for each animal (individual animal approach); alternatively excise and pool the lymph nodes from each ear in PBS for each treatment group (pooled treatment group approach). Details and diagrams of the lymph node identification and dissection can be found in reference (12). To further monitor the local skin response in the main study, additional parameters such as scoring of ear erythema or ear thickness measurements (obtained either by using a thickness gauge, or ear punch weight determinations at necropsy) may be included in the study protocol.

**Preparation of cell suspensions**

26. A single-cell suspension of lymph node cells (LNC) excised bilaterally using the individual animal approach or alternatively, the pooled treatment group approach is prepared by gentle mechanical disaggregation through 200 micron-mesh stainless steel gauze or another acceptable technique for generating a single-cell suspension. The LNC are washed twice with an excess of PBS and the DNA is precipitated with 5 % trichloroacetic acid (TCA) at 4 °C for 18h (3). Pellets are either resuspended in 1 mL TCA and transferred to scintillation vials containing 10 mL of scintillation fluid for  $^3\text{H}$ -counting, or transferred directly to gamma counting tubes for  $^{125}\text{I}$ -counting.

**Determination of cellular proliferation (incorporated radioactivity)**

27. Incorporation of  $^3\text{H}$ -methyl thymidine is measured by  $\beta$ -scintillation counting as disintegrations per minute (DPM). Incorporation of  $^{125}\text{I}$ -iododeoxyuridine is measured by  $^{125}\text{I}$ -counting and also is expressed as DPM. Depending on the approach used, the incorporation is expressed as DPM/mouse (individual animal approach) or DPM/treatment group (pooled treatment group approach).

**Reduced LLNA**

28. In certain situations, when there is a regulatory need to confirm a negative prediction of skin sensitising potential, an optional rLLNA protocol (16) (17) (18) using fewer animals may be used, provided there is adherence to all other LLNA protocol specifications in this TM. Before applying the rLLNA approach, clear justifications and scientific rationale for its use should be provided. If a positive or equivocal result is obtained, additional testing may be needed in order to interpret or clarify the finding.

**▼ M3**

29. The reduction in number of dose groups is the only difference between the LLNA and the rLLNA test method protocols and for this reason the rLLNA does not provide dose-response information. Therefore, the rLLNA should not be used when dose-response information is needed. Like the multi-dose LLNA, the test substance concentration evaluated in the rLLNA should be the maximum concentration that does not induce overt systemic toxicity and/or excessive local skin irritation in the mouse (see paragraph 18).

**OBSERVATIONS****Clinical observations**

30. Each mouse should be carefully observed at least once daily for any clinical signs, either of local irritation at the application site or of systemic toxicity. All observations are systematically recorded with records being maintained for each mouse. Monitoring plans should include criteria to promptly identify those mice exhibiting systemic toxicity, excessive local skin irritation, or corrosion of skin for euthanasia (37).

**Body weights**

31. As stated in paragraph 25, individual animal body weights should be measured at the start of the test and at the scheduled humane kill.

**CALCULATION OF RESULTS**

32. Results for each treatment group are expressed as the SI. When using the individual animal approach, the SI is derived by dividing the mean DPM/mouse within each test substance group, and the PC group, by the mean DPM/mouse for the solvent/VC group. The average SI for the VCs is then one. When using the pooled treatment group approach, the SI is obtained by dividing the pooled radioactive incorporation for each treatment group by the incorporation of the pooled VC group; this yields a mean SI.
33. The decision process regards a result as positive when  $SI \geq 3$ . However, the strength of the dose-response, the statistical significance and the consistency of the solvent/vehicle and PC responses may also be used when determining whether a borderline result is declared positive (4) (5) (6).
34. If it is necessary to clarify the results obtained, consideration should be given to various properties of the test substance, including whether it has a structural relationship to known skin sensitizers, whether it causes excessive local skin irritation in the mouse, and the nature of the dose-response relationship seen. These and other considerations are discussed in detail elsewhere (7).
35. Collecting radioactivity data at the level of the individual mouse will enable a statistical analysis for presence and degree of dose-response relationship in the data. Any statistical assessment could include an evaluation of the dose-response relationship as well as suitably adjusted comparisons of test groups (e.g. pair-wise dosed group versus concurrent VC comparisons). Statistical analyses may include, e.g. linear regression or William's test to assess dose-response trends, and Dunnett's test for pair-wise comparisons. In choosing an appropriate method of statistical analysis, the investigator should maintain an awareness of possible inequalities of variances and other related problems that may necessitate a data transformation or a non-parametric statistical analysis. In any case the investigator may need to carry out SI calculations and statistical analyses with and without certain data points (sometimes called 'outliers').

**▼ M3**

## DATA AND REPORTING

**Data**

36. Data should be summarised in tabular form. When using the individual animal approach, show the individual animal DPM values, the group mean DPM/animal, its associated error term (*e.g.* SD, SEM), and the mean SI for each dose group compared against the concurrent VC group. When using the pooled treatment group approach, show the mean/median DPM and the mean SI for each dose group compared against the concurrent VC group.

**Test report**

37. The test report should contain the following information:

Test and control substances:

- identification data (*e.g.* CAS and EC numbers, if available; source; purity; known impurities; lot number);
- physical nature and physicochemical properties (*e.g.* volatility, stability, solubility);
- if mixture, composition and relative percentages of components;

Solvent/vehicle:

- identification data (purity; concentration, where appropriate; volume used);
- justification for choice of vehicle;

Test animals:

- source of CBA mice;
- microbiological status of the animals, when known;
- number and age of animals;
- source of animals, housing conditions, diet, etc.;

Test conditions:

- details of test substance preparation and application;
- justification for dose selection (including results from pre-screen test, if conducted);
- vehicle and test substance concentrations used, and total amount of test substance applied;
- details of food and water quality (including diet type/source, water source);
- details of treatment and sampling schedules;
- methods for measurement of toxicity;
- criteria for considering studies as positive or negative;
- details of any protocol deviations and an explanation on how the deviation affects the study design and results;

Reliability check:

- summary of results of latest reliability check, including information on test substance, concentration and vehicle used;
- concurrent and/or historical PC and concurrent NC data for testing laboratory;

▼ **M3**

- if a concurrent PC was not included, the date and laboratory report for the most recent periodic PC and a report detailing the historical PC data for the laboratory justifying the basis for not conducting a concurrent PC;

## Results:

- individual weights of mice at start of dosing and at scheduled kill; as well as mean and associated error term (*e.g.* SD, SEM) for each treatment group;
- time course of onset and signs of toxicity, including dermal irritation at site of administration, if any, for each animal;
- a table of individual mouse (individual animal approach) or mean/median (pooled treatment group approach) DPM values and SI values for each treatment group;
- mean and associated error term (*e.g.* SD, SEM) for DPM/mouse for each treatment group and the results of outlier analysis for each treatment group when using the individual animal approach;
- calculated SI and an appropriate measure of variability that takes into account the inter-animal variability in both the test substance and control groups when using the individual animal approach;
- dose-response relationship;
- statistical analyses, where appropriate;

## Discussion of results:

- a brief commentary on the results, the dose-response analysis, and statistical analyses, where appropriate, with a conclusion as to whether the test substance should be considered a skin sensitiser.

## LITERATURE

- (1) OECD (2002), Skin Sensitisation: Local Lymph Node Assay. OECD Guideline for the Testing of Chemicals No 429, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (2) Kimber, I. and Basketter, D.A. (1992), The murine local lymph node assay; collaborative studies and new directions: A commentary, *Food Chem. Toxicol.*, 30, 165-169.
- (3) Kimber, I., Dearman, R.J., Scholes, E.W. and Basketter, D.A. (1994), The local lymph node assay: developments and applications, *Toxicol.*, 93, 13-31.
- (4) Kimber, I., Hilton, J., Dearman, R.J., Gerberick, G.F., Ryan, C.A., Basketter, D.A., Lea, L., House, R.V., Ladies, G.S., Loveless, S.E. and Hastings, K.L. (1998), Assessment of the skin sensitisation potential of topical medicaments using the local lymph node assay: An interlaboratory exercise, *J. Toxicol. Environ. Health*, 53, 563-79.
- (5) Chamberlain, M. and Basketter, D.A. (1996), The local lymph node assay: status of validation, *Food Chem. Toxicol.*, 34, 999-1002.
- (6) Basketter, D.A., Gerberick, G.F., Kimber, I. and Loveless, S.E. (1996), The local lymph node assay: A viable alternative to currently accepted skin sensitisation tests, *Food Chem. Toxicol.*, 34, 985-997.
- (7) Basketter, D.A., Gerberick, G.F. and Kimber, I. (1998), Strategies for identifying false positive responses in predictive sensitisation tests, *Food Chem. Toxicol.*, 36, 327-33.



▼ M3

- (8) Van Och, F.M.M., Slob, W., De Jong, W.H., Vandebriel, R.J. and Van Loveren, H. (2000), A quantitative method for assessing the sensitising potency of low molecular weight chemicals using a local lymph node assay: employment of a regression method that includes determination of uncertainty margins, *Toxicol.*, 146, 49-59.
- (9) Dean, J.H., Twerdok, L.E., Tice, R.R., Sailstad, D.M., Hattan, D.G., Stokes, W.S. (2001), ICCVAM evaluation of the murine local lymph node assay: II. Conclusions and recommendations of an independent scientific peer review panel, *Reg. Toxicol. Pharmacol.*, 34: 258-273.
- (10) Haneke, K.E., Tice, R.R., Carson, B.L., Margolin, B.H., Stokes, W.S. (2001), ICCVAM evaluation of the murine local lymph node assay: III. Data analyses completed by the national toxicology program interagency center for the evaluation of alternative toxicological methods, *Reg. Toxicol. Pharmacol.*, 34, 274-286.
- (11) Sailstad, D.M., Hattan, D., Hill, R.N., Stokes, W.S. (2001), ICCVAM evaluation of the murine local lymph node assay: I. The ICCVAM review process, *Reg. Toxicol. Pharmacol.*, 34: 249-257.
- (12) ICCVAM (2009), Recommended Performance Standards: Murine Local Lymph Node Assay, NIH Publication Number 09-7357, Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: [[http://iccvam.niehs.nih.gov/docs/immunotox\\_docs/llnaps/LLNAPerfStds.pdf](http://iccvam.niehs.nih.gov/docs/immunotox_docs/llnaps/LLNAPerfStds.pdf)]
- (13) OECD (1992), *Skin Sensitisation*. OECD Guideline for Testing of Chemicals No 406, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (14) OECD (2005), *Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment*, Environment, Health and Safety Monograph, Series on Testing and Assessment No 34, ENV/JM/MONO(2005)14, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (15) Dearman, R.J., Hilton, J., Evans, P., Harvey, P., Basketter, D.A. and Kimber, I. (1998), Temporal stability of local lymph node assay responses to hexyl cinnamic aldehyde, *J. Appl. Toxicol.*, 18, 281-284.
- (16) Kimber, I., Dearman, R.J., Betts, C.J., Gerberick, G.F., Ryan, C.A., Kern, P.S., Patlewicz, G.Y. and Basketter, D.A. (2006), The local lymph node assay and skin sensitisation: a cut-down screen to reduce animal requirements? *Contact Dermatitis*, 54, 181-185.
- (17) ESAC (2007), Statement on the Reduced Local Lymph Node Assay (rLLNA), European Commission Directorate-General, Joint Research Centre, Institute for Health and Consumer Protection, European Centre for the Validation of Alternative Methods, April 2007. Available at: [[http://ecvam.jrc.it/ft\\_doc/ESAC26\\_statement\\_rLLNA\\_20070525-1.pdf](http://ecvam.jrc.it/ft_doc/ESAC26_statement_rLLNA_20070525-1.pdf)]
- (18) ICCVAM (2009), The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) Test Method Evaluation Report. The Reduced Murine Local Lymph Node Assay: An Alternative Test Method Using Fewer Animals to Assess the Allergic Contact Dermatitis Potential of Chemicals and Products, NIH Publication Number 09-6439, Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: [<http://iccvam.niehs.nih.gov/>]
- (19) ICCVAM (1999), The Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals/Compounds, The Results of an Independent Peer Review Evaluation Coordinated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), NIH Publication No 99-4494, Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: [[http://iccvam.niehs.nih.gov/docs/immunotox\\_docs/llna/lnarep.pdf](http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/lnarep.pdf)]

## ▼M3

- (20) Kreiling, R., Hollnagel, H.M., Hareng, L., Eigler, L., Lee, M.S., Griem, P., Dreesen, B., Kleber, M., Albrecht, A., Garcia, C. and Wendel, A. (2008), Comparison of the skin sensitising potential of unsaturated compounds as assessed by the murine local lymph node assay (LLNA) and the guinea pig maximization test (GPMT), *Food Chem. Toxicol.*, 46, 1896-1904.
- (21) Basketter, D., Ball, N., Cagen, S., Carrilo, J.C., Certa, H., Eigler, D., Garcia, C., Esch, H., Graham, C., Haux, C., Kreiling, R. and Mehling, A. (2009), Application of a weight of evidence approach to assessing discordant sensitisation datasets: implications for REACH, *Reg. Toxicol. Pharmacol.*, 55, 90-96.
- (22) ICCVAM (2009), ICCVAM Test Method Evaluation Report. Assessment of the Validity of the LLNA for Testing Pesticide Formulations and Other Products, Metals, and Substances in Aqueous Solutions, NIH Publication Number 10-7512, Research Triangle Park, NC: National Institute of Environmental Health Sciences, Available at: [<http://iccvam.niehs.nih.gov/>]
- (23) ILAR (1996), Institute of Laboratory Animal Research (ILAR) Guide for the Care and Use of Laboratory Animals, 7th ed. Washington, DC: National Academies Press.
- (24) McGarry, H.F. (2007), The murine local lymph node assay: regulatory and potency considerations under REACH, *Toxicol.*, 238, 71-89.
- (25) OECD (2002), Acute Dermal Irritation/Corrosion. OECD Guideline for Testing of Chemicals No 404, Paris, France. Available at: [[http://www.oecd.org/document/40/0,3343,en\\_2649\\_34377\\_37051368\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/40/0,3343,en_2649_34377_37051368_1_1_1_1,00.html)]
- (26) Reeder, M.K., Broomhead, Y.L., DiDonato, L. and DeGeorge, G.L. (2007), Use of an enhanced local lymph node assay to correctly classify irritants and false positive substances, *Toxicologist*, 96, 235.
- (27) ICCVAM (2009), Non-radioactive Murine Local Lymph Node Assay: Flow Cytometry Test Method Protocol (LLNA: BrdU-FC) Revised Draft Background Review Document, Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: [<http://iccvam.niehs.nih.gov/methods/immunotox/fcLLNA/BRDcomplete.pdf>]
- (28) Hayes, B.B., Gerber, P.C., Griffey, S.S. and Meade, B.J. (1998), Contact hypersensitivity to dicyclohexylcarbodiimide and diisopropylcarbodiimide in female B6C3F1 mice, *Drug. Chem. Toxicol.*, 21, 195-206.
- (29) Homey, B., von Schilling, C., Blumel, J., Schuppe, H.C., Ruzicka, T., Ahr, H.J., Lehmann, P. and Vohr, V.W. (1998), An integrated model for the differentiation of chemical-induced allergic and irritant skin reactions, *Toxicol. Appl. Pharmacol.*, 153, 83-94.
- (30) Woolhiser, M.R., Hayes, B.B. and Meade, B.J. (1998), A combined murine local lymph node and irritancy assay to predict sensitisation and irritancy potential of chemicals, *Toxicol. Meth.*, 8, 245-256.
- (31) Hayes, B.B. and Meade, B.J. (1999), Contact sensitivity to selected acrylate compounds in B6C3F1 mice: relative potency, cross reactivity, and comparison of test methods, *Drug. Chem. Toxicol.*, 22, 491-506.
- (32) Ehling, G., Hecht, M., Heusener, A., Huesler, J., Gamer, A.O., van Loveren, H., Maurer, T., Riecke, K., Ullmann, L., Ulrich, P., Vandebriel, R. and Vohr, H.W. (2005), A European inter-laboratory validation of alternative endpoints of the murine local lymph node assay: first round. *Toxicol.*, 212, 60-68.
- (33) Vohr, H.W. and Ahr, H.J. (2005), The local lymph node assay being too sensitive? *Arch. Toxicol.*, 79, 721-728.
- (34) Patterson, R.M., Noga, E. and Germolec, D. (2007), Lack of evidence for contact sensitisation by Pfisteria extract, *Environ. Health Perspect.*, 115, 1023-1028.
- (35) OECD (1987), *Acute Dermal Toxicity*, OECD Guideline for Testing of Chemicals No 402, Paris, France. Available at: [<http://www.oecd.org/env/testguidelines>]

**▼ M3**

- (36) ICCVAM (2009), Report on the ICCVAM-NICEATM/ECVAM/JaCVAM Scientific Workshop on Acute Chemical Safety Testing: Advancing In Vitro Approaches and Humane Endpoints for Systemic Toxicity Evaluations. Research Triangle Park, NC: National Institute of Environmental Health Sciences, Available at: [[http://iccvam.niehs.nih.gov/methods/acutetox/Tox\\_workshop.htm](http://iccvam.niehs.nih.gov/methods/acutetox/Tox_workshop.htm)]
- (37) OECD (2000), *Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation*, Environmental Health and Safety Monograph, Series on Testing and Assessment No 19, ENV/JM/MONO(2000)7, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]

**▼ M3***Appendix 1***Performance standards for assessment of proposed similar or modified llna test methods for skin sensitisation**

## INTRODUCTION

1. The purpose of Performance Standards (PS) is to communicate the basis by which new test methods, both proprietary (*i.e.* copyrighted, trademarked, registered) and non-proprietary can be determined to have sufficient accuracy and reliability for specific testing purposes. These PS, based on validated and accepted test methods, can be used to evaluate the reliability and accuracy of other similar methods (colloquially referred to as 'me-too' tests) that are based on similar scientific principles and measure or predict the same biological or toxic effect (14).
2. Prior to adoption of modified methods (*i.e.* proposed potential improvements to an approved test method), there should be an evaluation to determine the effect of the proposed changes on the test's performance and the extent to which such changes affect the information available for the other components of the validation process. Depending on the number and nature of the proposed changes, the generated data and supporting documentation for those changes, they should either be subjected to the same validation process as described for a new test, or, if appropriate, to a limited assessment of reliability and relevance using established PS (14).
3. Similar or modified methods proposed for use under this TM should be evaluated to determine their reliability and accuracy using chemicals representing the full range of the LLNA scores. To avoid unwarranted animal use, it is strongly recommended that model developers consult the appropriate authorities before starting validation studies in accordance with the PS and guidance provided in this TM.
4. These PS are based on the US-ICCVAM, EC-ECVAM and Japanese-JaCVAM harmonised PS (12), for evaluating the validity of similar or modified versions of the LLNA. The PS consists of essential test method components, recommended reference chemicals and standards for accuracy and reliability that the proposed method should meet or exceed.

**I. Essential test method components**

5. To ensure that a similar or modified LLNA method is functionally and mechanistically analogous to the LLNA and measures the same biological effect, the following components should be included in the test method protocol:

— The test substance should be applied topically to both ears of the mouse;

— Lymphocyte proliferation should be measured in the lymph nodes draining from the site of test substance application;

— Lymphocyte proliferation should be measured during the induction phase of skin sensitisation;

**▼ M3**

- For test substances, the highest dose selected should be the maximum concentration that does not induce systemic toxicity and/or excessive local skin irritation in the mouse. For positive reference chemicals, the highest dose should be at least as high as the LLNA EC3 values of the corresponding reference chemicals (see Table 1) without producing systemic toxicity and/or excessive local skin irritation in the mouse;
- A concurrent VC should be included in each study and, where appropriate, a concurrent PC should also be used;
- A minimum of four animals per dose group should be used;
- Either individual or pooled animal data may be collected.

If any of these criteria are not met, then these PS cannot be used for validation of the similar or modified method.

**II. Minimum list of reference chemicals**

6. The US-ICCVAM, EC-ECVAM and Japanese-JaCVAM harmonised PS (12) identified 18 minimum reference chemicals that should be used and four optional reference chemicals (*i.e.* substances that produced either false positive or false negative results in the LLNA, when compared to human and guinea pig results (B.6, or OECD Test Guideline 406) (13), and therefore provide the opportunity to demonstrate equal to or better performance than the LLNA) that are included in the LLNA PS. The selection criteria for identifying these chemicals were:
  - The list of reference chemicals represented the types of substances typically tested for skin sensitisation potential and the range of responses that the LLNA is capable of measuring or predicting;
  - The substances had well-defined chemical structures;
  - LLNA data from guinea pig tests (*i.e.* B.6; OECD Test Guideline 406) (13) and (where possible) data from humans were available for each substance; and
  - The substances were readily available from a commercial source.

The recommended reference chemicals are listed in Table 1. Studies using the proposed reference chemicals should be evaluated in the vehicle with which they are listed in Table 1. In situations where a listed substance may not be available, other substances that meet the selection criteria mentioned may be used, with adequate justification.

Table 1

## Recommended Reference chemicals for the LLNA PS.

Number	Chemicals <sup>(1)</sup>	CAS No	Form	Veh <sup>(2)</sup>	EC3 % <sup>(3)</sup>	N <sup>(4)</sup>	0,5x-2,0x EC3	Actual EC3 Range	LLNA vs GP	LLNA vs Human
1	5-Chloro-2-methyl-4-isothiazolin-3-one (CMI)/2-methyl-4-isothiazolin-3-one (MI) <sup>(5)</sup>	26172-55-4/ 2682-20-4	Liq	DMF	0,009	1	0,0045-0,018	NC	+/+	+/+
2	DNCB	97-00-7	Sol	AOO	0,049	15	0,025-0,099	0,02-0,094	+/+	+/+
3	4-Phenylenediamine	106-50-3	Sol	AOO	0,11	6	0,055-0,22	0,07-0,16	+/+	+/+
4	Cobalt chloride	7646-79-9	Sol	DMSO	0,6	2	0,3-1,2	0,4-0,8	+/+	+/+
5	Isoeugenol	97-54-1	Liq	AOO	1,5	47	0,77-3,1	0,5-3,3	+/+	+/+
6	2-Mercaptobenzothiazole	149-30-4	Sol	DMF	1,7	1	0,85-3,4	NC	+/+	+/+
7	Citral	5392-40-5	Liq	AOO	9,2	6	4,6-18,3	5,1-13	+/+	+/+
8	HCA	101-86-0	Liq	AOO	9,7	21	4,8-19,5	4,4-14,7	+/+	+/+
9	Eugenol	97-53-0	Liq	AOO	10,1	11	5,05-20,2	4,9-15	+/+	+/+
10	Phenyl benzoate	93-99-2	Sol	AOO	13,6	3	6,8-27,2	1,2-20	+/+	+/+
11	Cinnamic alcohol	104-54-1	Sol	AOO	21	1	10,5-42	NC	+/+	+/+
12	Imidazolidinyl urea	39236-46-9	Sol	DMF	24	1	12-48	NC	+/+	+/+
13	Methyl methacrylate	80-62-6	Liq	AOO	90	1	45-100	NC	+/+	+/+
14	Chlorobenzene	108-90-7	Liq	AOO	25	1	NA	NA	-/-	-/ (*)
15	Isopropanol	67-63-0	Liq	AOO	50	1	NA	NA	-/-	-/+

▼ M3

Number	Chemicals <sup>(1)</sup>	CAS No	Form	Veh <sup>(2)</sup>	EC3 % <sup>(3)</sup>	N <sup>(4)</sup>	0,5x-2,0x EC3	Actual EC3 Range	LLNA vs GP	LLNA vs Human
16	Lactic acid	50-21-5	Liq	DMSO	25	1	NA	NA	-/-	-/ (*)
17	Methyl salicylate	119-36-8	Liq	AOO	20	9	NA	NA	-/-	-/-
18	Salicylic acid	69-72-7	Sol	AOO	25	1	NA	NA	-/-	-/-

Optional Substances to Demonstrate Improved Performance Relative to the LLNA

19	Sodium lauryl sulphate	151-21-3	Sol	DMF	8,1	5	4,05-16,2	1,5-17,1	+/-	+/-
20	Ethylene glycol dimethacrylate	97-90-5	Liq	MEK	28	1	14-56	NC	+/-	+/+
21	Xylene	1330-20-7	Liq	AOO	95,8	1	47,9-100	NC	+/(*)	+/-
22	Nickel chloride	7718-54-9	Sol	DMSO	5	2	NA	NA	-/+	-/+

Abbreviations: AOO = acetone: olive oil (4:1, v/v); CAS No = Chemical Abstracts Service Number; DMF = *N,N*-dimethylformamide; DMSO = dimethyl sulfoxide; DNCB = 2,4-dinitrochlorobenzene; EC3 = estimated concentration needed to produce a stimulation index of 3; GP = guinea pig test result (*i.e.* B.6 or OECD Test Guideline 406) (13); HCA = hexyl cinnamic aldehyde; Liq = liquid; LLNA = murine local lymph node assay result (*i.e.* B.42 or OECD Test Guideline 429) (1); MEK = methyl ethyl ketone; NA = not applicable since stimulation index < 3; NC = not calculated since data was obtained from a single study; Sol = solid; Veh = test vehicle.

(\*) Presumed to be a non-sensitiser in humans based on the fact that no clinical patch test results were located, it is not included as a patch test kit allergen, and no case reports of human sensitisation were located.

(\*\*) GP data not available.

<sup>(1)</sup> Chemicals should be prepared daily unless stability data demonstrate the acceptability of storage.

<sup>(2)</sup> Because of the potential impact of different vehicles on the performance of the LLNA, the recommended vehicle for each reference chemical should be used (24) (32).

<sup>(3)</sup> Mean value where more than one EC3 value was available. For negative substances (*i.e.* with stimulation index < 3, the highest concentration tested is provided).

<sup>(4)</sup> Number of LLNA studies from which data were obtained.

<sup>(5)</sup> Commercially available as Kathon CG (CAS No 55965-84-9), which is a 3:1 mixture of CMI and MI. The relative concentrations of each component range from 1,1 % to 1,25 % (CMI) and 0,3 % to 0,45 % (MI). The inactive components are magnesium salts (21,5 % to 24 %) and copper nitrate (0,15 % to 0,17 %), with the remaining formulation 74 % to 77 % water. Kathon CG is readily available through Sigma-Aldrich and Rohm and Haas (now Dow Chemical Corporation).

**▼ M3****III. Defined reliability and accuracy standards**

7. The accuracy of a similar or modified LLNA method should meet or exceed that of the LLNA PS when it is evaluated using the 18 minimum reference chemicals that should be used. The new or modified method should result in the correct classification based on a 'yes/no' decision. However, the new or modified method might not correctly classify all of the minimum reference chemicals that should be used. If, for example, one of the weak sensitisers were misclassified, a rationale for the misclassification and appropriate additional data (e.g. test results that provide correct classifications for other substances with physical, chemical, and sensitising properties similar to those of the misclassified reference chemical) could be considered to demonstrate equivalent performance. Under such circumstances, the validation status of the new or modified LLNA test method would be evaluated on a case-by-case basis.

*Intra-laboratory reproducibility*

8. To determine intra-laboratory reproducibility, a new or modified LLNA method should be assessed using a sensitising substance that is well characterised in the LLNA. Therefore, the LLNA PS are based on the variability of results from repeated tests of hexyl cinnamic aldehyde (HCA). To assess intra-laboratory reliability, threshold estimated concentration (ECt) values for HCA should be derived on four separate occasions with at least one week between tests. Acceptable intra-laboratory reproducibility is indicated by a laboratory's ability to obtain, in each HCA test, ECt values between 5 % and 20 %, which represents the range of 0,5-2,0 times the mean EC3 specified for HCA (10 %) in the LLNA (see Table 1).

*Inter-laboratory reproducibility*

9. Inter-laboratory reproducibility of a new or modified LLNA method should be assessed using two sensitising substances that are well characterised in the LLNA. The LLNA PS are based on the variability of results from tests of HCA and 2,4-dinitrochlorobenzene (DNCB) in different laboratories. ECt values should be derived independently from a single study conducted in at least three separate laboratories. To demonstrate acceptable inter-laboratory reproducibility, each laboratory should obtain ECt values of 5 % to 20 % for HCA and 0,025 % to 0,1 % for DNCB, which represents the range of 0,5-2,0 times the mean EC3 concentrations specified for HCA (10 %) and DNCB (0,05 %), respectively, in the LLNA (see Table 1).



▼ **M3***Appendix 2***Definitions**

*Accuracy:* The closeness of agreement between test method results and accepted reference values. It is a measure of test method performance and one aspect of relevance. The term is often used interchangeably with ‘concordance’ to mean the proportion of correct outcomes of a test method (14).

*Benchmark substance:* A sensitising or non-sensitising substance used as a standard for comparison to a test substance. A benchmark substance should have the following properties: (i) consistent and reliable source(s); (ii) structural and functional similarity to the class of substances being tested; (iii) known physicochemical characteristics; (iv) supporting data on known effects; and (v) known potency in the range of the desired response.

*Estimated concentration threshold (EC<sub>t</sub>):* Estimated concentration of a test substance needed to produce a stimulation index that is indicative of a positive response.

*Estimated concentration three (EC<sub>3</sub>):* Estimated concentration of a test substance needed to produce a stimulation index of three.

*False negative:* A test substance incorrectly identified as negative or non-active by a test method, when in fact it is positive or active.

*False positive:* A test substance incorrectly identified as positive or active by a test, when in fact it is negative or non-active.

*Hazard:* The potential for an adverse health or ecological effect. The adverse effect is manifested only if there is an exposure of sufficient level.

*Inter-laboratory reproducibility:* A measure of the extent to which different qualified laboratories, using the same protocol and testing the same test substances, can produce qualitatively and quantitatively similar results. Inter-laboratory reproducibility is determined during the pre-validation and validation processes, and indicates the extent to which a test can be successfully transferred between laboratories, also referred to as between-laboratory reproducibility (14).

*Intra-laboratory reproducibility:* A determination of the extent that qualified people within the same laboratory can successfully replicate results using a specific protocol at different times. Also referred to as within-laboratory reproducibility (14).

*Me-too test:* A colloquial expression for a test method that is structurally and functionally similar to a validated and accepted reference test method. Such a test method would be a candidate for catch-up validation. Interchangeably used with similar test method (14).

*Outlier:* An outlier is an observation that is markedly different from other values in a random sample from a population.

*Performance standards (PS):* Standards, based on a validated test method, that provide a basis for evaluating the comparability of a proposed test method that is functionally and mechanistically similar. Included are: (i) essential test method components; (ii) a minimum list of Reference Chemicals selected from among the chemicals used to demonstrate the acceptable performance of the validated test method; and (iii) the similar levels of accuracy and reliability, based on what was obtained for the validated test method, that the proposed test method should demonstrate when evaluated using the minimum list of Reference Chemicals (14).

*Proprietary test method:* A test method for which manufacture and distribution is restricted by patents, copyrights, trademarks, etc.

**▼ M3**

*Quality assurance:* A management process by which adherence to laboratory testing standards, requirements, and record keeping procedures, and the accuracy of data transfer, are assessed by individuals who are independent from those performing the testing.

*Reference chemicals:* Chemicals selected for use in the validation process, for which responses in the *in vitro* or *in vivo* reference test system or the species of interest are already known. These chemicals should be representative of the classes of chemicals for which the test method is expected to be used, and should represent the full range of responses that may be expected from the chemicals for which it may be used, from strong, to weak, to negative. Different sets of reference chemicals may be required for the different stages of the validation process, and for different test methods and test uses (14).

*Relevance:* Description of relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method (14).

*Reliability:* Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility (14).

*Skin sensitisation:* An immunological process that results when a susceptible individual is exposed topically to an inducing chemical allergen, which provokes a cutaneous immune response that can lead to the development of contact sensitisation.

*Stimulation Index (SI):* A value calculated to assess the skin sensitisation potential of a test substance that is the ratio of the proliferation in treated groups to that in the concurrent vehicle control group.

*Test substance (also referred to as test chemical):* Any substance or mixture tested using this TM.

*Validated test method:* A test method for which validation studies have been completed to determine the relevance (including accuracy) and reliability for a specific purpose. It is important to note that a validated test method may not have sufficient performance in terms of accuracy and reliability to be found acceptable for the proposed purpose (14).

**▼B****B.43. NEUROTOXICITY STUDY IN RODENTS****1. METHOD**

This method is equivalent of OECD TG 424 (1997).

This test method has been designed to obtain the information necessary to confirm or to further characterise the potential neurotoxicity of chemicals in adult animals. It can either be combined with existing test methods for repeated dose toxicity studies or to be carried out as a separate study. It is recommended that the OECD Guidance Document on Neurotoxicity Testing Strategies and Methods (1) be consulted to assist in the design of studies based on this test method. This is particularly important when modifications of the observations and test procedures as recommended for routine use of this method are considered. The guidance document has been prepared to facilitate the selection of other test procedures for use in specific circumstances.

The assessment of developmental neurotoxicity is not the subject of this method.

**1.1. INTRODUCTION**

In the assessment and evaluation of the toxic characteristics of chemicals, it is important to consider the potential for neurotoxic effects. Already the test method for repeated dose systemic toxicity includes observations that screen for potential neurotoxicity. This test method can be used to design a study to obtain further information on, or to confirm, the neurotoxic effects observed in the repeated dose systemic toxicity studies. However, consideration of the potential neurotoxicity of certain classes of chemicals may suggest that they may be more appropriately evaluated using this Method without prior indications of the potential neurotoxicity from repeated dose systemic toxicity studies. Such considerations include, for example:

- observation of neurological signs or neuropathological lesions in toxicity studies other than repeated dose systemic toxicity studies, or
- structural relationship or other information linking them to known neurotoxicants.

In addition there may be other instances when use of this test method is appropriate; for further details see (1).

This method has been developed so that it can be tailored to meet particular needs to confirm the specific histopathological and behavioural neurotoxicity of a chemical as well as provide a characterisation and quantification of the neurotoxic responses.

**▼B**

In the past, neurotoxicity was equated with neuropathy involving neuropathological lesions or neurological dysfunctions, such as seizure, paralysis or tremor. Although neuropathy is an important manifestation of neurotoxicity, it is now clear that there are many other signs of nervous system toxicity (e.g. loss of motor co-ordination, sensory deficits, learning and memory dysfunctions) that may not be reflected in neuropathy or other types of studies.

This neurotoxicity test method is designed to detect major neurobehavioural and neuropathological effects in adult rodents. While behavioural effects, even in the absence of morphological changes, can reflect an adverse impact on the organism, not all behavioural changes are specific to the nervous system. Therefore, any changes observed should be evaluated in conjunction with correlative histopathological, haematological or biochemical data as well as data on other types of systemic toxicity. The testing called for in this method to provide a characterisation and quantification of the neurotoxic responses includes specific histopathological and behavioural procedures that may be further supported by electrophysiological and/or biochemical investigations (1)(2)(3)(4).

Neurotoxicants may act on a number of targets within the nervous system and by a variety of mechanisms. Since no single array of tests is capable of thoroughly assessing the neurotoxic potential of all substances, it may be necessary to utilise other *in vivo* or *in vitro* tests specific to the type of neurotoxicity observed or anticipated.

This test method can also be used, in conjunction with the guidance set out in the OECD Guidance Document on Neurotoxicity Testing Strategies and Methods (1) to design studies intended to further characterise or increase the sensitivity of the dose-response quantification in order to better estimate a no-observed-adverse effect level or to substantiate known or suspected hazards of the chemical. For example, studies may be designed to identify and evaluate the neurotoxic mechanism(s) or supplement the data already available from the use of basic neurobehavioural and neuropathological observation procedures. Such studies need not replicate data that would be generated from the use of the standard procedures recommended in this Method, if such data are already available and are not considered necessary for the interpretation of the results of the study.

This neurotoxicity study, when used alone or in combination, provides information that can:

- identify whether the nervous system is permanently or reversibly affected by the chemical tested;
  
- contribute to the characterisation of the nervous system alterations associated with exposure to the chemical, and to understanding the underlying mechanism.

**▼ B**

- determine dose-and time-response relationships in order to estimate a no-observed-adverse-effect level (which can be used to establish safety criteria for the chemical).

This test method uses oral administration of the test substance. Other routes of administration (e.g. dermal or inhalation) may be more appropriate, and may require modification of the procedures recommended. Considerations of the choice of the route of administration depend on the human exposure profile and available toxicological or kinetic information.

## 1.2. DEFINITIONS

**Adverse effect:** is any treatment-related alteration from baseline that diminishes an organism's ability to survive, reproduce or adapt to the environment.

**Dose:** is the amount of test substance administered. Dose is expressed as weight (g, mg) or as weight of test substance per unit weight of the test animal (e.g. mg/Kg), or as constant dietary concentrations (ppm).

**Dosage:** is a general term comprising of dose, its frequency and the duration of dosing.

**Neurotoxicity:** is an adverse change in the structure or function of the nervous system that results from exposure to a chemical, biological or physical agent.

**Neurotoxicant:** is any chemical, biological or physical agent having the potential to cause neurotoxicity.

**NOAEL:** is the abbreviation for no-observed-adverse effect level and is the highest dose level where no adverse treatment-related findings are observed.

## 1.3. PRINCIPLE OF THE TEST METHOD

The test chemical is administered by the oral route across a range of doses to several groups of laboratory rodents. Repeated doses are normally required, and the dosing regimen may be 28 days, subchronic (90 days) or chronic (1 year or longer). The procedures set out in this test method may also be used for an acute neurotoxicity study. The animals are tested to allow the detection or the characterisation of behavioural and/or neurological abnormalities. A range of behaviours that could be affected by neurotoxicants is assessed during each observation period. At the end of the test, a subset of animals of each sex from each group are perfused *in situ* and sections of the brain, spinal cord, and peripheral nerves are prepared and examined.

When the study is conducted as a stand-alone study to screen for neurotoxicity or to characterise neurotoxic effects, the animals in each group not used for perfusion and subsequent histopathology (see Table 1) can be used for specific neurobehavioural, neuropathological, neurochemical or electrophysiological procedures that may supplement the data obtained from the standard examinations required by this method (1). These supplemental procedures can be particularly useful when empirical observations or anticipated effects indicate a specific type or target of a chemical's neurotoxicity. Alternatively, the remaining animals can be used for evaluations such as those called for in test methods for repeated dose toxicity studies in rodents.

**▼B**

When the procedures of this test method are combined with those of other test methods, a sufficient number of animals is needed to satisfy the requirements for the observations of both studies.

**1.4. DESCRIPTION OF THE TEST METHOD****1.4.1. Selection of animal species**

The preferred rodent species is the rat, although other rodent species, with justification, may be used. Commonly used laboratory strains of young adult healthy animals should be employed. The females should be nulliparous and non-pregnant. Dosing should normally begin as soon as possible after weaning, preferably not later than when animals are six weeks, and, in any case, before the animals are nine weeks age. However, when this study is combined with other studies this age requirement may need adjustment. At the commencement of the study the weight variation of animals used should not exceed  $\pm 20\%$  of the mean weight of each sex. Where a repeated dose study of short duration is conducted as a preliminary to a long term study, animals from the same strain and source should be used in both studies.

**1.4.2. Housing and feeding conditions**

The temperature in the experimental animal room should be 22 °C ( $\pm 3$  °C). Although the relative humidity should be at least 30 % and preferably not exceed 70 % other than during room cleaning, the aim should be 50-60 %. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. Loud intermittent noise should be kept to a minimum. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water. The choice of diet may be influenced by the need to ensure a suitable admixture of a test substance when administered by this method. Animals may be housed individually, or be caged in small groups of the same sex.

**1.4.3. Preparation of animals**

Healthy young animals are randomly assigned to the treatment and control groups. Cages should be arranged in such a way that possible effects due to cage placement are minimised. The animals are identified uniquely and kept in their cages for at least (5) five days prior the start of the study to allow for acclimatisation to the laboratory conditions.

**1.4.4. Route of administration and preparation of doses**

This test method specifically addresses the oral administration of the test substance. Oral administration may be by gavage, in the diet, in drinking water or by capsules. Other routes of administration (e.g. dermal or inhalation) can be used but may require modification of the procedures recommended. Considerations of the choice of the route of administration depend on the human exposure profile and available toxicological or kinetic information. The rationale for choosing the route of administration as well as resulting modifications to the procedures of this test method should be indicated.

**▼B**

Where necessary, the test substance may be dissolved or suspended in a suitable vehicle. It is recommended that the use of an aqueous solution/suspension be considered first, followed by consideration of a solution/suspension in oil (e.g. corn oil) and then by possible solution/suspension in other vehicle. The toxic characteristics of the vehicle must be known. In addition, consideration should be given to the following characteristics of the vehicle: effects of the vehicle on absorption, distribution, metabolism, or retention of the test substance which may alter its toxic characteristics; and effects on the food or water consumption or the nutritional status of the animals.

**1.5. PROCEDURES****1.5.1. Number and sex animals**

When the study is conducted as a separate study, at least 20 animals (10 females and 10 males) should be used in each dose and control group for the evaluation of detailed clinical and functional observations. At least five males and five females, selected from these 10 males and 10 females, should be perfused *in situ* and used for detailed neurohistopathology at the end of the study. In cases where only a limited number of animals in a given dose group are observed for signs of neurotoxic effects, consideration should be given to the inclusion of these animals in those selected for perfusion. When the study is conducted in combination with a repeated dose toxicity study, adequate numbers of animals should be used to meet the objectives of both studies. The minimum numbers of animals per group for various combinations of studies are given in Table 1. If interim kills or recovery groups for observation of reversibility, persistence or delayed occurrence of toxic effects post treatment are planned or when supplemental observations are considered, then the number of animals should be increased to ensure that the number of animals required for observation and histopathology are available.

**1.5.2. Treatment and control group**

At least three dose groups and a control group should generally be used, but if from the assessment of other data, no effects would be expected at a repeated dose of 1 000 mg/kg body weight/day, a limit test may be performed. If there are no suitable data available, a range finding study may be performed to aid in the determination of the doses to be used. Except for treatment with the test substance, animals in the control group should be handled in an identical manner to the test group subjects. If a vehicle is used in administering the test substance, the control group should receive the vehicle at the highest volume used.

**1.5.3. Reliability check**

The laboratory performing the study should present data demonstrating its capability to carry out the study and the sensitivity of the procedures used. Such data should provide evidence of the ability to detect and quantify, as appropriate, changes in the different end points recommended for observation, such as autonomic signs, sensory reactivity, limb grip strength and motor activity. Information on chemicals that cause different types of neurotoxic responses and could be used as positive control substances can be found in references 2 to 9. Historical data may be used if the essential aspects of the experimental procedures remain the same. Periodic updating of historical data is recommended. New data that demonstrate the continuing sensitivity of the procedures should be developed when some essential element of the conduct of the test or procedures has been changed by the performing laboratory.

**▼B****1.5.4. Dose selection**

Dose levels should be selected by taking into account any previously observed toxicity and kinetic data available for the test compound or related materials. The highest dose level should be chosen with the aim of inducing neurotoxic effects or clear systemic toxic effects. Thereafter, a descending sequence of dose levels should be selected with a view to demonstrating any dose-related response and no-observed-adverse effect (NOAEL) at the lowest dose level. In principle, dose levels should be set so that primary toxic effects on the nervous system can be distinguished from effects related to systemic toxicity. Two to three intervals are frequently optimum and addition of a fourth test group is often preferable to using very large intervals (e.g. more than a factor of 10) between dosages. Where there is a reasonable estimation of human exposure this should also be taken into account.

**1.5.5. Limit test**

If a study at one dose level of at least 1 000 mg/kg body weight/day, using the procedures described, produces no observable neurotoxic effects and if toxicity would not be expected based upon data from structurally related compounds, then a full study using three dose levels may not be considered necessary. Expected human exposure may indicate the need for a higher oral dose level to be used in the limit test. For other types of administration, such as inhalation or dermal application, the physical chemical properties of the test substance often may dictate the maximum attainable level of exposure. For the conduct of an oral acute study, the dose for a limit test should be at least 2 000 mg/kg.

**1.5.6. Administration of doses**

The animals are dosed with the test substance daily, seven days each week, for a period at least 28 days; use of a five-day dosing regime or a shorter exposure period needs to be justified. When the test substance is administered by gavage, this should be done in a single dose using a stomach tube or a suitable intubation cannula. The maximum volume of a liquid that can be administered at one time depends on the size of the test animals. The volume should not exceed 1 ml/100 g body weight. However in the case of aqueous solutions, the use of up to 2 ml/100 g body weight can be considered. Except for irritating or corrosive substances, which will normally reveal exacerbated effects with higher concentrations, variability in test volume should be minimised by adjusting the concentration to ensure a constant volume at all dose levels.

For substances administered via the diet or drinking water, it is important to ensure that the quantities of the test substance involved do not interfere with normal nutrition or water balance. When the test substance is administered in the diet either a constant dietary concentration (ppm) or a constant dose level in terms of the animals' body weight may be used; the alternative used must be specified. For a substance administered by gavage, the dose should be given at similar times each day, and adjusted as necessary to maintain a constant dose level in terms of animal body weight. Where a repeat dose study is used as a preliminary to a long term study, a similar diet should be used in both studies. For acute studies, if a single dose is not possible, the dose may be given in smaller fractions over a period not exceeding 24 hours.



**▼B**

## 1.6. OBSERVATION

1.6.1. **Frequency of observations and tests**

In repeated dose studies, the observation period should cover the dosage period. In acute studies, 14-day post-treatment period should be observed. For animals in satellite groups which are kept without exposure during a post-treatment period, observations should cover this period as well.

Observations should be made with sufficient frequency to maximise the probability of detection of any behavioural and/or neurological abnormalities. Observations should be made preferably at the same times each day with consideration given to the peak period of anticipated effects after dosing. The frequency of clinical observations and functional tests is summarised in Table 2. If kinetic or other data generated from previous studies indicates the need to use different time points for observations, tests or post-observation periods, an alternative schedule should be adopted in order to achieve maximum information. The rationale for changes to the schedule should be provided.

1.6.1.1. *Observations of general health condition and mortality/morbidity*

All animals should be carefully observed at least once daily with respect to their health condition as well as at least twice daily for morbidity and mortality.

1.6.1.2. *Detailed clinical observations*

Detailed clinical observations should be made on all animals selected for this purpose (see Table 1) once before the first exposure (to allow for within-subject comparisons) and at different intervals thereafter, dependant on the duration of the study (see Table 2). Detailed clinical observations on satellite recovery groups should be made at the end of the recovery period. Detailed clinical observations should be made outside the home cage in a standard arena. They should be carefully recorded using scoring systems that include criteria or scoring scales for each measurement in the observations. The criteria or scales used should be explicitly defined by the testing laboratory. Effort should be made to ensure that variations in the test conditions are minimal (not systematically related to treatment) and that observations are conducted by trained observers unaware of the actual treatment.

It is recommended that the observations be carried out in a structured fashion in which well-defined criteria (including the definition of the normal 'range') are systematically applied to each animal at each observation time. The 'normal range' should be adequately documented. All observed signs should be recorded. Whenever feasible, the magnitude of the observed signs should also be recorded. Clinical observations should include, but not be limited to, changes in skin, fur, eyes, mucous membranes, occurrence of secretions and excretions and autonomic activity (e.g. lacrimation, piloerection, pupil size, unusual respiratory pattern and/or mouth breathing, any unusual signs of urination or defecation, and discoloured urine).

**▼B**

Any unusual responses with respect to body position, activity level (e.g. decreased or increased exploration of the standard arena) and co-ordination of movement should also be noted. Changes in gait (e.g. waddling, ataxia), posture (e.g. hunched-back) and reactivity to handling, placing or other environmental stimuli, as well as the presence of clonic or tonic movements, convulsions or tremors, stereotypes (e.g. excessive grooming, unusual head movements, repetitive circling) or bizarre behaviour (e.g. biting or excessive licking, self mutilation, walking backwards, vocalisation) or aggression should be recorded.

**1.6.1.3. Functional tests**

Similar to the detailed clinical observations, functional tests should also be conducted once prior to exposure and frequently thereafter in all animals selected for this purpose (see Table 1). The frequency of functional testing is also dependent on the study duration (see Table 2). In addition to the observation periods as set out in Table 2, functional observations on satellite recovery groups should also be made as close as possible to the terminal kill. Functional tests should include sensory reactivity to stimuli of different modalities (e.g. auditory, visual and proprioceptive stimuli (5)(6)(7)), assessment of limb grip strength (8) and assessment of motor activity (9). Motor activity should be measured with an automated device capable of detecting both decreases and increases in activity. If another defined system is used it should be quantitative and its sensitivity and reliability should be demonstrated. Each device should be tested to ensure reliability across time and consistency between devices. Further details of the procedures that can be followed are given in the respective references. If there are no data (e.g. structure-activity, epidemiological data, other toxicology studies) to indicate the potential neurotoxic effects, the inclusion of more specialised tests of sensory and motor function or learning and memory to examine these possible effects in greater details should be considered. More information on more specialised tests and their use is provided in (1).

Exceptionally, animals that reveal signs of toxicity to an extent that would significantly interfere with the functional test may be omitted from that test. Justification for the elimination of animals from a functional test should be provided.

**1.6.2. Body weight and food/water consumption**

For studies up to 90 days duration, all animals should be weighed at least once a week and measurements should be made of food consumption (water consumption, when the test substance is administered by that medium) at least weekly. For long term studies, all animals should be weighed at least once at week for the first 13 weeks and at least once every four weeks thereafter. Measurements should be made of food consumption (water consumption, when the test substance is administered by that medium) at least weekly for the first 13 weeks and then at approximately three-month intervals unless the health status or body weight changes dictate otherwise.

**▼ B****1.6.3. Ophthalmology**

For studies longer than 28 days duration, ophthalmologic examination, using an ophthalmoscope or an equivalent suitable instrument, should be made prior to the administration of the test substance and at the termination of the study, preferably on all animals, but at least on animals in the high dose and control groups. If changes in the eyes are detected or, if clinical signs indicate the need, all animals should be examined. For long term studies, an ophthalmologic examination should also be carried out at 13 weeks. Ophthalmologic examinations need not to be conducted if this data is already available from others studies of similar duration and at similar dose levels.

**1.6.4. Haematology and clinical biochemistry**

When the neurotoxicity study is carried out in combination with a repeated dose systemic toxicity study, haematological examinations and clinical biochemistry determinations should be carried out as set out in the respective method of the systemic toxicity study. Collection of samples should be carried out in such a way that any potential effects on neurobehaviour are minimised.

**1.6.5. Histopathology**

The neuropathological examination should be designed to complement and extend the observations made during the *in vivo* phase of the study. Tissues from at least five animals/sex/group (see Table 1 and next paragraph) should be fixed *in situ*, using generally recognised perfusion and fixation techniques (see reference 3, chapter 5 and reference 4, chapter 50). Any observable gross changes should be recorded. When the study is conducted as a stand-alone study screen for neurotoxicity or to characterise neurotoxic effects, the remainder of the animals may be used either for specific neurobehavioural (10)(11), neuropathological (10)(11)(12)(13), neurochemical (10)(11)(14)(15) or electrophysiological (10)(11)(16)(17) procedures that may supplement the procedures and examinations described here, or to increase the number of subjects examined for histopathology. These supplementary procedures are of particular use when empirical observations or anticipated effects indicate a specific type or target of neurotoxicity (2)(3). Alternatively, the remainder of the animals can also be used for routine pathological evaluations as described in Method for repeated dose studies.

A general staining procedure, such as haematoxylin and eosin (H&E), should be performed on all tissue specimens embedded in paraffin and microscopic examination should be carried out. If signs of peripheral neuropathy are observed or suspected, plastic-embedded samples of peripheral nerve tissue should be examined. Clinical signs may also suggest additional sites for examination or the use of special staining procedures. Guidance on additional sites to be examined can be found in (3)(4). Appropriate special stains to demonstrate specific types of pathological change may also be helpful (18).

**▼B**

Representative sections of the central and peripheral nervous system should be examined histologically (see reference 3, chapter 5 and reference 4, chapter 50). The areas examined should normally include: the forebrain, the centre of the cerebrum, including a section through the hippocampus, the midbrain, the cerebellum, the pons, the medulla oblongata, the eye with optic nerve and retina, the spinal cord at the cervical and lumbar swellings, the dorsal root ganglia, the dorsal and ventral root fibres, the proximal sciatic nerve, the proximal tibial nerve (at the knee) and the tibial nerve calf muscle branches. The spinal cord and peripheral nerve sections should include both cross or transverse and longitudinal sections. Attention should be given to the vasculature of the nervous system. A sample of skeletal muscle, particularly calf muscle, should also be examined. Special attention should be paid to sites with cellular and fibre structure and pattern in the CNS and PNS known to be particularly affected by neurotoxicants.

Guidance on neuropathological alterations that typically result from toxicant exposure can be found in the references (3)(4). A stepwise examination of tissue samples is recommended in which sections from the high dose group are first compared with those of the control group. If no neuropathological alterations are observed in the samples from these groups, subsequent analysis is not required. If neuropathological alterations are observed in the high dose group, sample from each of the potentially affected tissues from the intermediate and low dose groups should then be coded and examined sequentially.

If any evidence of neuropathological alterations is found in the qualitative examination, then a second examination should be performed on all regions of the nervous system showing these alterations. Sections from all dose groups from each of the potentially affected regions should be coded and examined at random without knowledge of the code. The frequency and severity of each lesion should be recorded. After all regions from all dose groups have been rated, the code can be broken and statistical analysis performed to evaluate dose-response relationships. Examples of different degrees of severity of each lesion should be described.

The neuropathological findings should be evaluated in the context of behavioural observations and measurements, as well as other data from preceding and concurrent systemic toxicity studies of the test substance.

## 2. DATA

### 2.1. TREATMENT OF RESULTS

Individual data should be provided. Additionally, all data should be summarised in tabular form showing for each test or control group the number of animals at the start of the test, the number of animals found dead during the test or killed for humane reasons and the time of any death or humane kill, the number showing signs of toxicity, a description of the signs of toxicity observed, including time of onset, duration, type and severity of any toxic effects, the number of animals showing lesions, including the type and severity of the lesion(s).

**▼B**

## 2.2. EVALUATION AND INTERPRETATION OF RESULTS

The findings of the study should be evaluated in terms of the incidence, severity and correlation of neurobehavioural and neuropathological effects (neurochemical or electrophysiological effects as well if supplementary examinations are included) and any other adverse effects observed. When possible, numerical results should be evaluated by an appropriate and generally acceptable statistical method. The statistical methods should be selected during the design of the study.

## 3. REPORTING

## 3.1. TEST REPORT

the test report must include the following information:

Test substance:

- physical nature (including isomerism, purity and physicochemical properties),
- identification data.

Vehicle (if appropriate):

- justification for choice of vehicle.

Test animals:

- species/strain used,
- number, age and sex of animals,
- source, housing conditions, acclimatisation, diet, etc,
- individual weights of animals at the start of the test.

Test conditions:

- details of test substance formulation/diet preparation, achieved concentration, stability and homogeneity of the preparation,
- specification of the doses administered, including details of the vehicle, volume and physical form of the material administered,
- details of the administration of the test substance,
- rationale for dose levels selected,
- rationale for the route and duration of the exposure,
- conversion from diet/drinking water test substance concentration (ppm) to the actual dose (mg/kg body weight/day), if applicable,
- details of the food and water quality.

Observation and test procedures:

- details of the assignment of animals in each group to the perfusion subgroups,
- details of scoring systems, including criteria and scoring scales for each measurement in the detailed clinical observations,

**▼B**

- details on the functional tests for sensory reactivity to stimuli of different modalities (e.g. auditory, visual and proprioceptive), for assessment of limb grip strength, for motor activity assessment (including details of automated devices for detecting activity), and other procedures used,
- details of ophthalmologic examinations and, if appropriate, haematological examinations and clinical biochemistry tests with relevant base-line values,
- details for specific neurobehavioural, neuropathological, neurochemical or electrophysiological procedures.

## Results:

- body weight/body weight changes including body weight at kill,
- food consumption and water consumption, as appropriate,
- toxic response data by sex and dose level, including signs of toxicity or mortality,
- nature, severity and duration (time of onset and subsequent course) of the detailed clinical observations (whether reversible or not),
- a detailed description of all functional test results,
- necropsy findings,
- a detailed description of all neurobehavioural, neuropathological, and neurochemical or electrophysiological findings, if available,
- absorption and metabolism data, if available,
- statistical treatment of results, where appropriate.

## Discussion of results:

- dose response information;
- relationship of any other toxic effects to a conclusion about the neurotoxic potential of the test chemical;
- no-observed-adverse effect level.

## Conclusions:

- a specific statement of the overall neurotoxicity of the test chemical is encouraged.

4. **REFERENCES**

- (1) OECD Guidance Document on Neurotoxicity Testing Strategies and Test Methods. OECD, Paris, In Preparation.
- (2) Test Guideline for a Developmental Neurotoxicity Study, OECD Guidelines for the Testing of Chemicals. In preparation.
- (3) World Health Organisation (WHO) (1986) Environmental Health Criteria document 60: Principles and Methods for the Assessment of Neurotoxicity associated with Exposure to Chemicals.
- (4) Spencer, P.S. and Schaumburg, H.H. (1980) Experimental and Clinical Neurotoxicology. Eds. Spencer, P.S. and Schaumburg, H.H. eds. Williams and Wilkins, Baltimore/London.

**▼B**

- (5) Tupper, D.E. and Wallace, R.B., (1980) Utility of the Neurological Examination in Rats. *Acta Neurobiol. Exp.*, 40, p. 999-1003.
- (6) Gad, S.C., (1982) A Neuromuscular Screen for Use in Industrial Toxicology. *J. Toxicol. Environ. Health*, 9, p. 691-704.
- (7) Moser, V.C., McDaniel, K.M. and Phillips, P.M., (1991) Rat Strain and Stock Comparisons Using a Functional Observational Battery: Baseline Values and Effects of amitraz. *Toxic. Appl. Pharmacol.*, 108, p. 267-283.
- (8) Meyer, O.A., Tilson, H.A., Byrd, W.C. and Riley, M.T., (1979) A Method for the Routine Assessment of Fore- and Hind- limb Grip Strength of Rats and Mice. *Neurobehav. Toxicol.*, 1, p. 233-236.
- (9) Crofton, K.M., Haward, J.L., Moser, V.C., Gill, M.W., Reirer, L.W., Tilson, H.A. and MacPhail, R.C. (1991) Interlaboratory Comparison of Motor Activity Experiments: Implication for Neurotoxicological Assessments. *Neurotoxicol. Teratol.*, 13, p. 599-609.
- (10) Tilson, H.A., and Mitchell, C.L. eds., (1992) *Neurotoxicology Target Organ Toxicology Series*. Raven Press, New York.
- (11) Chang, L.W., ed., (1995) *Principles of Neurotoxicology*. Marcel Dekker, New York.
- (12) Broxup, B., (1991) Neuropathology as a screen for Neurotoxicity Assessment. *J. Amer. Coll. Toxicol.*, 10, p. 689-695.
- (13) Moser, V.C., Anthony, D.C., Sette, W.F. and MacPhail, R.C., (1992) Comparison of Subchronic Neurotoxicity of 2-Hydroxyethyl Acrylate and Acrylamide in Rats. *Fund. Appl. Toxicol.*, 18, p. 343-352.
- (14) O'Callaghan, J.P. (1988). Neurotypic and Gliotypic Proteins as Biochemical Markers of Neurotoxicity. *Eurotoxicol. Teratol.*, 10, p. 445-452.
- (15) O'Callaghan J.P. and Miller, D.B., (1988) Acute Exposure of the Neonatal Rat to Triethyltin Results in Persistent Changes in Neurotypic and Gliotypic Proteins. *J. Pharmacol. Exp. Ther.*, 244, p. 368-378.
- (16) Fox, D.A., Lowndes, H.E. and Birkamper, G.G., (1982) Electrophysiological Techniques in Neurotoxicology. In: *Nervous System Toxicology*. Mitchell, C.L. ed. Raven Press, New York, p. 299-335.
- (17) Johnson, B.L., (1980) Electrophysiological Methods in neurotoxicity Testing. In: *Experimental and Clinical Neurotoxicology*. Spencer, P.S. and Schaumburg, H.H. eds., Williams and Wilkins Co., Baltimore/London, p. 726-742.
- (18) Bancroft, J.D. and Steven A., (1990) Theory and Practice of Histological Techniques. Chapter 17, *Neuropathological Techniques*. Lowe, James and Cox, Gordon eds. Churchill Livingstone.

▼B

Table 1

**Minimum numbers of animals needed per group when the neurotoxicity study is conducted separately or in combination with studies**

	NEUROTOXICITY STUDY CONDUCTED AS:			
	Separate study	Combined study with the 28-day study	Combined study with the 90-day study	Combined study with the chronic toxicity study
Total number of animals per group	10 males and 10 females	10 males and 10 females	15 males and 15 females	25 males and 25 females
Number of animals selected for functional testing including detailed clinical observations	10 males and 10 females	10 males and 10 females	10 males and 10 females	10 males and 10 females
Number of animals selected per perfusion <i>in situ</i> and neurohistopathology	5 males and 5 females	5 males and 5 females	5 males and 5 females	5 males and 5 females
Number of animals selected for repeated dose/sub-chronic/chronic toxicity observations, haematology, clinical biochemistry, histopathology, etc. as indicate in the respective <i>Guidelines</i>		5 males and 5 females	10 males † and 10 females †	20 males † and 20 females †
Supplemental observations, as appropriate	5 males and 5 females			

† Includes five animals selected for functional testing and detailed clinical observations as part of the neurotoxicity study.





Table 2

## Frequency of clinical observation and functional tests

Type of observations		Study duration			
		Acute	28-day	90-day	Chronic
In all animals	General health condition	daily	daily	daily	daily
	Mortality/morbidity	Twice daily	Twice daily	Twice daily	Twice daily
In animals selected for functional observations	Detailed clinical observations	<ul style="list-style-type: none"> <li>— prior to first exposure</li> <li>— within 8 hours of dosing at estimate time of peak effect</li> <li>— at day 7 and 14 after dosing</li> </ul>	<ul style="list-style-type: none"> <li>— prior to first exposure</li> <li>— once weekly thereafter</li> </ul>	<ul style="list-style-type: none"> <li>— prior to first exposure</li> <li>— once during the first or second week of exposure</li> <li>— monthly thereafter</li> </ul>	<ul style="list-style-type: none"> <li>— prior to first exposure</li> <li>— once at the end of the first month of exposure</li> <li>— every three months thereafter</li> </ul>
	Functional tests	<ul style="list-style-type: none"> <li>— prior to first exposure</li> <li>— within 8 hours of dosing at estimate time of peak effect</li> <li>— at day 7 and 14 after dosing</li> </ul>	<ul style="list-style-type: none"> <li>— prior to first exposure</li> <li>— during the fourth week of treatment as close as possible to the end of the exposure period</li> </ul>	<ul style="list-style-type: none"> <li>— prior to first exposure</li> <li>— once during the first or second week of exposure</li> <li>— monthly thereafter</li> </ul>	<ul style="list-style-type: none"> <li>— prior to first exposure</li> <li>— once at the end of the first month of exposure</li> <li>— every three months thereafter</li> </ul>

**▼B****B.44. SKIN ABSORPTION: *IN VIVO* METHOD****1. METHOD**

This testing method is equivalent to the OECD TG 427 (2004).

**1.1. INTRODUCTION**

Exposure to many chemicals occurs mainly *via* the skin whilst the majority of toxicological studies performed in laboratory animals use the oral route of administration. The *in vivo* percutaneous absorption study set out in this guideline provides the linkage necessary to extrapolate from oral studies when making safety assessments following dermal exposure.

A substance must cross a large number of cell layers of the skin before it can reach the circulation. The rate-determining layer for most substances is the *stratum corneum* consisting of dead cells. Permeability through the skin depends both on the lipophilicity of the chemical and the thickness of the outer layer of epidermis, as well on factors such as molecular weight and concentration of the substance. In general, the skin of rats and rabbits is more permeable than that of humans, whereas the skin permeability of guinea pigs and monkeys is more similar to that of humans.

The methods for measuring percutaneous absorption can be divided into two categories; *in vivo* and *in vitro*. The *in vivo* method is capable of providing good information, in various laboratory species, on skin absorption. More recently *in vitro* methods have been developed. These utilise transport across full or partial thickness animal or human skin to a fluid reservoir. The *in vitro* method is described in a separate testing method (1). It is recommended that the OECD Guidance Document for the Conduct of Skin Absorption Studies (2) be consulted to assist in the selection of the most appropriate method in the given situation, as it provides more details on the suitability of both *in vivo* and *in vitro* methods.

The *in vivo* method, described in this method, allows the determination of the penetration of the test substance through the skin into the systemic compartment. The technique has been widely used for many years (3)(4)(5)(6)(7). Although *in vitro* percutaneous absorption studies may in many cases be appropriate there may be situations in which only an *in vivo* study can provide the necessary data.

Advantages of the *in vivo* method are that it uses a physiologically and metabolically intact system, uses a species common to many toxicity studies and can be modified for use with other species. The disadvantages are the use of live animals, the need for radio-labelled material to facilitate reliable results, difficulties in determining the early absorption phase and the differences in permeability of the preferred species (rat) and human skin. Animal skin is generally more permeable and therefore may overestimate human percutaneous absorption (6)(8)(9). Caustic/corrosive substances should not be tested in live animals.

**▼ B**

## 1.2. DEFINITIONS

**Unabsorbed dose:** represents that washed from the skin surface after exposure and any present on the non-occlusive cover, including any dose shown to volatilise from the skin during exposure.

**Absorbed dose (*in vivo*):** comprises that present in urine, cage wash, faeces, expired air (if measured), blood, tissues (if collected) and the remaining carcass, following removal of application site skin.

**Absorbable dose:** represents that present on or in the skin following washing.

## 1.3. PRINCIPLE OF THE TEST METHOD

The test substance, preferably radiolabelled, is applied to the clipped skin of animals at one or more appropriate dose levels in the form of a representative in-use preparation. The test preparation is allowed to remain in contact with the skin for a fixed period of time under a suitable cover (non-occlusive, semi-occlusive, or occlusive) to prevent ingestion of the test preparation. At the end of the exposure time the cover is removed and the skin is cleaned with an appropriate cleansing agent, the cover and the cleansing materials are retained for analysis and a fresh cover applied. The animals are housed prior to, during and after the exposure period in individual metabolism cages and the excreta and expired air over these periods are collected for analysis. The collection of expired air can be omitted when there is sufficient information that little or no volatile radioactive metabolite is formed. Each study will normally involve several groups of animals that will be exposed to the test preparation. One group will be killed at the end of the exposure period. Other groups will be killed at scheduled time intervals thereafter (2). At the end of the sampling time the remaining animals are killed, blood is collected for analysis, the application site removed for analysis and the carcass is analysed for any unexcreted material. The samples are assayed by appropriate means and the degree of percutaneous absorption is estimated (6)(8)(9).

## 1.4. DESCRIPTION OF THE METHOD

## 1.4.1. Selection of animal species

The rat is the most commonly used species, but hairless strains and species having skin absorption rates more similar to those of human, can also be used (3)(6)(7)(8)(9). Young adult healthy animals of a single sex (with males as the default sex) of commonly used laboratory strains should be employed. At the commencement of the study, the weight variation of animals used should not exceed  $\pm 20\%$  of the mean weight. As an example, male rats of 200 g – 250 g are suitable, particularly in the upper half of this range.

**▼B****1.4.2. Number and sex of animals**

A group of at least four animals of one sex should be used for each test preparation and each scheduled termination time. Each group of animals will be killed after different time intervals, for example at the end of the exposure period (typically 6 or 24 hours) and subsequent occasions (e.g. 48 and 72 hours). If there are data available that demonstrate substantial differences in dermal toxicity between males and females, the more sensitive sex should be chosen. If there are no such data, then either gender can be used.

**1.4.3. Housing and feeding conditions**

The temperature in the experimental animal room should be 22 °C ( $\pm 3$  °C). Although the relative humidity should be at least 30 % and preferably not exceed 70 % other than during room cleaning, the aim should be 50-60 %. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used and should be freely available together with an unlimited supply of drinking water. During the study, and preferably also during the acclimatisation, the animals are individually housed in metabolism cages. Since food and water spillage would compromise the results, the probability of such events should be minimised.

**1.4.4. Preparation of animals**

The animals are marked to permit individual identification and kept in their cages for at least five days prior to the start of the study to allow for acclimatisation to the laboratory conditions.

Following the acclimatisation period, and approximately 24 hours prior to dosing, each animal will have an area of skin in the region of the shoulders and the back clipped. The permeation properties of damaged skin are different from intact skin and care should be taken to avoid abrading the skin. Following the clipping and approximately 24 hours before the test substance is applied to the skin, (See Section 1.4.7) the skin surface should be wiped with acetone to remove sebum. An additional soap and water wash is not recommended because any soap residue might promote test substance absorption. The area must be large enough to allow reliable calculation of the absorbed amount of test chemical per cm<sup>2</sup> skin, preferably at least 10 cm<sup>2</sup>. This area is practicable with rats of 200-250 g bodyweight. After preparation, the animals are returned to metabolism cages.

**1.4.5. Test substance**

The test substance is the entity whose penetration characteristics are to be studied. Ideally, the test substance should be radiolabelled.

**1.4.6. Test preparation**

The test substance preparation (e.g. neat, diluted, or formulated material containing the test chemical which is applied to the skin) should be the same (or realistic surrogate) as that to which humans or other potential target species may be exposed. Any variations from the 'in-use' preparation must be justified. Where necessary, the test substance is dissolved or suspended in a suitable vehicle. For vehicles other than water the absorption characteristics and potential interaction with the test substance should be known.

**▼B****1.4.7. Application to the skin**

An application site of a specific surface area is defined on the skin surface. A known amount of the test preparation is then evenly applied to the site. This amount should normally mimic potential human exposure, typically 1-5 mg/cm<sup>2</sup> for a solid or up to 10 µl/cm<sup>2</sup> for liquids. Any other quantities should be justified by the expected use conditions, the study objectives or physical characteristics of the test preparation. Following application, the treated site must be protected from grooming. An example of a typical device is shown in Figure 1. Normally, the application site will be protected by a non-occlusive cover (e.g. a permeable nylon gauze cover). However, for infinite applications the application site should be occluded. In case of evaporation of semivolatile test substances reduces the recovery rate of the test substance to an unacceptable extend (see also section 1.4.10, first paragraph), it is necessary to trap the evaporated substance in a charcoal filter covering the application device (see Figure 1). It is important that any device does not damage the skin, nor absorb or react with the test preparation. The animals are returned to individual metabolism cages in order to collect excreta.

**1.4.8. Duration of exposure and sampling**

The duration of exposure is the time interval between application and removal of test preparation by skin washing. A relevant exposure period (typically 6 or 24 hours) should be used, based on the expected human exposure duration. Following the exposure period, the animals are maintained in the metabolism cages until the scheduled termination. The animals should be observed for signs of toxicity/abnormal reactions at regular intervals for the entire duration of the study. At the end of the exposure period the treated skin should be observed for visible signs of irritation.

The metabolism cages should permit separate collection of urine and faeces throughout the study. They should also allow collection of <sup>14</sup>C-carbon dioxide and volatile <sup>14</sup>C-carbon compounds, which should be analysed when produced in quantity (> 5 %). The urine, faeces and trap fluids (e.g. <sup>14</sup>C-carbon dioxide and volatile <sup>14</sup>C-compounds) should be individually collected from each group at each sampling time. If there is sufficient information that little or no volatile radioactive metabolite is formed, open cages can be used.

Excreta are collected during the exposure period, up to 24 hours after the initial skin contact and then daily until the end of the experiment. Whilst three excreta collection intervals will normally be sufficient, the envisaged purpose of the test preparation or existing kinetic data may suggest more appropriate or additional time points for study.

At the end of the exposure period the protective device is removed from each animal and retained separately for analysis. The treated skin of all animals should be washed at least three times with cleansing agent using suitable swabs. Care must be taken to avoid contaminating other parts of the body. The cleansing agent should be representative of normal hygiene practice, e.g. aqueous soap solution. Finally, the skin should be dried. All swabs and washings must be retained for analysis. A fresh cover should be applied to protect the treated site of those animals forming later time point groups prior to their return to individual cages.

**▼B****1.4.9. Terminal procedures**

For each group, the individual animals should be killed at the scheduled time and blood collected for analysis. The protective device or cover should be removed for analysis. The skin from the application site and a similar area of non-dosed, clipped skin should be removed from each animal for separate analysis. The application site may be fractionated to separate the *stratum corneum* from the underlying epidermis to provide more information on the test chemical disposition. The determination of this disposition over a time course after the exposure period should provide some indication of the fate of any test chemical in the *stratum corneum*. To facilitate skin fractionation (following the final skin wash and killing the animal) each protective cover is removed. The application site skin, with annular ring of surrounding skin, is excised from the rat and pinned on a board. A strip of adhesive tape is applied to the skin surface using gentle pressure and the tape removed together with part of the *stratum corneum*. Successive strips of tape are applied until the tape no longer adheres to the skin surface, when all of the *stratum corneum* has been removed. For each animal, all the tape strips may be combined in a single container to which a tissue digestant is added to solubilise the *stratum corneum*. Any potential target tissues may be removed for separate measurement before the residual carcass is analysed for absorbed carcass dose. The carcasses of the individual animals should be retained for analysis. Usually analysis of the total content will be sufficient. Target organs may be removed for separate analysis (if indicated by other studies). Urine present in the bladder at scheduled kill should be added to the previous urine collection. After collection of the excreta from metabolism cages at the time scheduled kill, the cages and their traps should be washed with an appropriate solvent. Other potentially contaminated equipment should likewise be analysed.

**1.4.10. Analysis**

In all studies adequate recovery (i.e. mean of  $100 \pm 10$  % of the radioactivity) should be achieved. Recoveries outside this range must be justified. The amount of the administered dose in each sample should be analysed by suitably validated procedures.

Statistical considerations should include a measure of variance for the replicates for each application.

**2. DATA**

The following measurements should be made for each animal, at each sampling time for the test chemical and/or metabolites. In addition to individual data, data grouped according to sampling times should be reported as means.

— quantity associated with the protective appliances,

— quantity that can be dislodged from the skin,

— quantity in/on skin that cannot be washed from the skin,

**▼B**

- quantity in the sampled blood,
- quantity in the excreta and expired air (if appropriate),
- quantity remaining in the carcass and any organs removed for separate analysis.

The quantity of test substance and/or metabolites in the excreta, expired air, blood and in the carcass will allow determination of the total amount absorbed at each time point. A calculation of the amount of test chemical absorbed per cm<sup>2</sup> of skin exposed to the test substance over the exposure period can also be obtained.

**3. REPORTING****3.1. TEST REPORT**

The test report must include the requirements stipulated in the protocol, including a justification for the test system used and should comprise the following:

test substance:

- identification data (e.g. CAS number, if available, source, purity (radiochemical purity), known impurities, lot number),
- physical nature, physicochemical properties (e.g. pH, volatility, solubility, stability, molecular weight and log P<sub>ow</sub>).

Test preparation:

- formulation and justification of use,
- details of the test preparation, amount applied, achieved concentration, vehicle, stability and homogeneity.

Test animal:

- species/strain used,
- number, age and sex of animals,
- source of animals, housing conditions, diets, etc.,
- individual animal weights at start of test.

Test conditions:

- details of the administration of the test preparation (site of application, assay methods, occlusion/non-occlusion, volume, extraction, detection),
- details of food and water quality.

Results:

- any signs of toxicity,
- tabulated absorption data (expressed as rate, amount or percentage),

**▼B**

- overall recoveries of the experiment,
- interpretation of the results, comparison with any available data on percutaneous absorption of the test compound.

Discussion of the results.

Conclusions.

**4. REFERENCES**

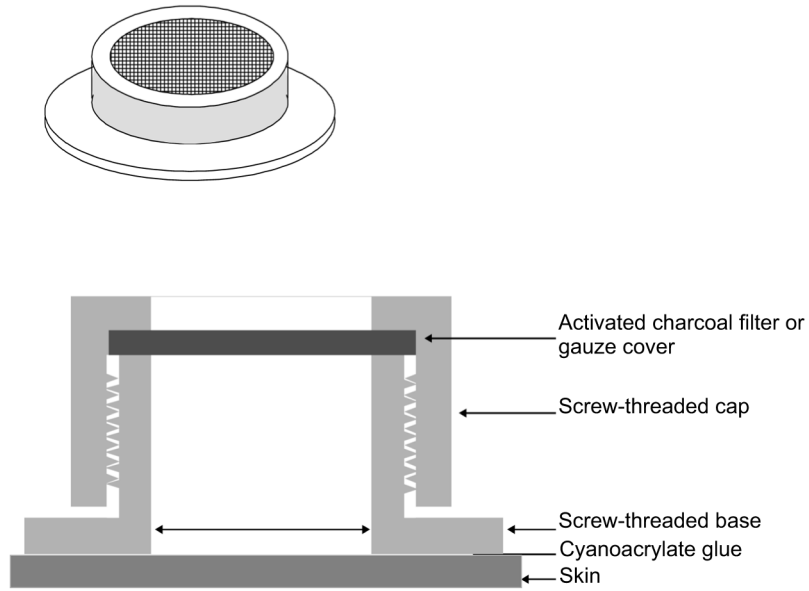
- (1) Testing Method B.45. Skin Absorption: *In vitro* Method.
- (2) OECD (2002). Guidance Document for the Conduct of Skin Absorption Studies. OECD, Paris.
- (3) ECETOC, (1993) Percutaneous Absorption. European Centre for Ecotoxicology and Toxicology of Chemicals, Monograph No 20.
- (4) Zendzian R.P. (1989) Skin Penetration Method suggested for Environmental Protection Agency Requirements. *J. Am. Coll. Toxicol.* 8(5), p. 829-835.
- (5) Kemppainen, B.W., Reifenrath WG (1990) Methods for skin absorption. CRC Press Boca Raton, FL, USA.
- (6) EPA, (1992) Dermal Exposure Assessment: Principles and Applications. Exposure Assessment Group, Office of Health and Environmental Assessment.
- (7) EPA, (1998) Health Effects Test Guidelines, OPPTS 870-7600, Dermal Penetration. Office of Prevention, Pesticides and Toxic Substances.
- (8) Bronaugh, R.L., Wester, R.C., Bucks, D., Maibach H.I. and Sarason, R., (1990) In vivo percutaneous absorption of fragrance ingredients in rhesus monkeys and humans. *Fd. Chem. Toxic.* 28, p. 369-373.
- (9) Feldman, R.J. and Maibach, H.I., (1970) Absorption of some organic compounds through the skin in man. *J. Invest Dermatol.* 54, p. 399-404.



▼ B

Figure 1

An example of a design of a typical device used to define and protect dermal application site during *in vivo* percutaneous absorption studies



**▼B****B.45. SKIN ABSORPTION: *IN VITRO* METHOD****1. METHOD**

This testing method is equivalent to the OECD TG 428 (2004).

**1.1. INTRODUCTION**

This method has been designed to provide information on absorption of a test substance applied to excised skin. It can either be combined with the method for skin absorption: *in vivo* method (1), or be conducted separately. It is recommended that the OECD Guidance Document for the Conduct of Skin Absorption Studies (2) be consulted to assist in the design of studies based on this method. The Guidance Document has been prepared to facilitate the selection of appropriate *in vitro* procedures for use in specific circumstances, to ensure the reliability of results obtained by this method.

The methods for measuring skin absorption and dermal delivery can be divided into two categories: *in vivo* and *in vitro*. *In vivo* methods on skin absorption are well established and provide pharmacokinetic information in a range of animal species. An *in vivo* method is separately described in another testing method (1). *In vitro* methods have also been used for many years to measure skin absorption. Although formal validation studies of the *in vitro* methods covered by this testing method have not been performed, OECD experts agreed in 1999 that there was sufficient data evaluated to support the *in vitro* method (3). Further details that substantiate this support, including a significant number of direct comparisons of *in vitro* and *in vivo* methods, are provided with the Guidance Document (2). There are a number of monographs that review this topic and provide detailed background on the use of an *in vitro* method (4)(5)(6)(7)(8)(9)(10)(11)(12). *In vitro* methods measure the diffusion of chemicals into and across skin to a fluid reservoir and can utilise non-viable skin to measure diffusion only, or fresh, metabolically active skin to simultaneously measure diffusion and skin metabolism. Such methods have found particular use as a screen for comparing delivery of chemicals into and through skin from different formulations and can also provide useful models for the assessment of percutaneous absorption in humans.

The *in vitro* method may not be applicable for all situations and classes of chemicals. It may be possible to use the *in vitro* test method for an initial qualitative evaluation of skin penetration. In certain cases, it may be necessary to follow this up with *in vivo* data. The guidance document (2) should be consulted for further elaboration of situations where the *in vitro* method would be suitable. Additional detailed information to support the decision is provided in (3).

This method presents general principles for measuring dermal absorption and delivery of a test substance using excised skin. Skin from many mammalian species, including humans, can be used. The permeability properties of skin are maintained after excision from the body because the principal diffusion barrier is the non-viable *stratum corneum*; active transport of chemicals through the skin has not been identified. The skin has been shown to have the capability to metabolise some chemicals during percutaneous absorption (6), but this process is not rate limiting in terms of actual absorbed dose, although it may affect the nature of the material entering the bloodstream.

**▼B**

## 1.2. DEFINITIONS

**Unabsorbed dose:** represents that washed from the skin surface after exposure and any present on the non-occlusive cover, including any dose shown to volatilise from the skin during exposure.

**Absorbed dose (*in vitro*):** mass of test substance reaching the receptor fluid or systemic circulation within a specified period of time.

**Absorbable dose (*in vitro*):** represents that present on or in the skin following washing.

## 1.3. PRINCIPLE OF THE TEST METHOD

The test substance, which may be radiolabelled, is applied to the surface of a skin sample separating the two chambers of a diffusion cell. The chemical remains on the skin for a specified time under specified conditions, before removal by an appropriate cleansing procedure. The receptor fluid is sampled at time points throughout the experiment and analysed for the test chemical and/or metabolites.

When metabolically active systems are used, metabolites of the test chemical may be analysed by appropriate methods. At the end of the experiment the distribution of the test chemical and its metabolites are quantified, when appropriate.

Using appropriate conditions, which are described in this method and the guidance document (2), absorption of a test substance during a given time period is measured by analysis of the receptor fluid and the treated skin. The test substance remaining in the skin should be considered as absorbed unless it can be demonstrated that absorption can be determined from receptor fluid values alone. Analysis of the other components (material washed off the skin and remaining within the skin layers) allows for further data evaluation, including total test substance disposition and percentage recovery.

To demonstrate the performance and reliability of the test system in the performing laboratory, the results for relevant reference chemicals should be available and in agreement with published literature for the method used. This requirement could be met by testing an appropriate reference substance (preferably of a lipophilicity close to the test substance) concurrently with the test substance or by providing adequate historical data for a number of reference substances of different lipophilicity (e.g. caffeine, benzoic acid, and testosterone).

## 1.4. DESCRIPTION OF THE TEST METHOD

## 1.4.1. Diffusion cell

A diffusion cell consists of a donor chamber and a receptor chamber between which the skin is positioned (an example of a typical design is provided in Figure 1). The cell should provide a good seal around the skin, enable easy sampling and good mixing of the receptor solution in contact with the underside of the skin, and good temperature control of the cell and its contents. Static and flow-through diffusion cells are both acceptable. Normally, donor chambers are left unoccluded during exposure to a finite dose of a test preparation. However, for infinite applications and certain scenarios for finite doses, the donor chambers may be occluded.

**▼B****1.4.2. Receptor fluid**

The use of a physiologically conducive receptor fluid is preferred although others may also be used provided that they are justified. The precise composition of the receptor fluid should be provided. Adequate solubility of the test chemical in the receptor fluid should be demonstrated so that it does not act as a barrier to absorption. In addition, the receptor fluid should not affect skin preparation integrity. In a flow-through system, the rate of flow must not hinder diffusion of a test substance into the receptor fluid. In a static cell system, the fluid should be continuously stirred and sampled regularly. If metabolism is being studied, the receptor fluid must support skin viability throughout the experiment.

**1.4.3. Skin preparations**

Skin from human or animal sources can be used. It is recognised that the use of human skin is subject to national and international ethical considerations and conditions. Although viable skin is preferred, non-viable skin can also be used provided that the integrity of the skin can be demonstrated. Either epidermal membranes (enzymically, heat or chemically separated) or split thickness skin (typically 200-400 µm thick) prepared with a dermatome, are acceptable. Full thickness skin may be used but excessive thickness (approximately > 1 mm) should be avoided unless specifically required for determination of the test chemical in layers of the skin. The selection of species, anatomical site and preparative technique must be justified. Acceptable data from a minimum of four replicates per test preparation are required.

**1.4.4. Skin preparation integrity**

It is essential that the skin is properly prepared. Inappropriate handling may result in damage to the *stratum corneum*, hence the integrity of the prepared skin must be checked. When skin metabolism is being investigated, freshly excised skin should be used as soon as possible, and under conditions known to support metabolic activity. As a general guidance, freshly excised skin should be used within 24 hours, but the acceptable storage period may vary depending on the enzyme system involved in metabolism and storage temperatures (13). When skin preparations have been stored prior to use, evidence should be presented to show that barrier function is maintained.

**1.4.5. Test substance**

The test substance is the entity whose penetration characteristics are to be studied. Ideally, the test substance should be radiolabelled.

**1.4.6. Test preparation**

The test substance preparation (e.g. neat, diluted or formulated material containing the test substance which is applied to the skin) should be the same (or a realistic surrogate) as that to which humans or other potential target species may be exposed. Any variation from the 'in-use' preparation must be justified.

**▼ B****1.4.7. Test substances concentrations and formulations**

Normally more than one concentration of the test substance is used spanning the upper of potential human exposures. Likewise, testing a range of typical formulations should be considered.

**1.4.8. Application to the skin**

Under normal conditions of human exposure to chemicals, finite doses are usually encountered. Therefore, an application that mimics human exposure, normally 1-5 mg/cm<sup>2</sup> of skin for a solid and up to 10 µl/cm<sup>2</sup> for liquids, should be used. The quantity should be justified by the expected use conditions, the study objectives or physical characteristics of the test preparation. For example, applications to the skin surface may be infinite, where large volumes per unit area are applied.

**1.4.9. Temperature**

The passive diffusion of chemicals (and therefore their skin absorption) is affected by temperature. The diffusion chamber and skin should be maintained at a constant temperature close to normal skin temperature of 32 ± 1 °C. Different cell designs will require different water bath or heated block temperatures to ensure that the receptor/skin is at its physiological norm. Humidity should preferably be between 30 and 70 %.

**1.4.10. Duration of exposure and sampling**

Skin exposure to the test preparation may be for the entire duration of the experiment or for shorter times (i.e., to mimic a specific type of human exposure). The skin should be washed of excess test preparation with a relevant cleansing agent, and the rinses collected for analysis. The removal procedure of the test preparation will depend on the expected use condition, and should be justified. A period of sampling of 24 hours is normally required to allow for adequate characterisation of the absorption profile. Since skin integrity may start to deteriorate beyond 24 hours, sampling times should not normally exceed 24 hours. For test substances that penetrate the skin rapidly this may not be necessary but, for test substances that penetrate slowly, longer times may be required. Sampling frequency of the receptor fluid should allow the absorption profile of the test substance to be presented graphically.

**1.4.11. Terminal procedures**

All components of the test system should be analysed and recovery is to be determined. This includes the donor chamber, the skin surface rinsing, the skin preparation and the receptor fluid/chamber. In some cases, the skin may be fractionated into the exposed area of skin and area of skin under the cell flange, and into *stratum corneum*, epidermis and dermis fractions, for separate analysis.

**1.4.12. Analysis**

In all studies adequate recovery should be achieved (the aim should be a mean of 100 ± 10 % of the radioactivity and any deviation should be justified). The amount of test substance in the receptor fluid, skin preparation, skin surface washings and apparatus rinse should be analysed, using a suitable technique.

**▼ B****2. DATA**

The analysis of receptor fluid, the distribution of the test substance chemical in the test system and the absorption profile with time, should be presented. When finite dose conditions of exposure are used, the quantity washed from the skin, the quantity associated with the skin (and in the different skin layers if analysed) and the amount present in the receptor fluid (rate, and amount or percentage of applied dose) should be calculated. Skin absorption may sometimes be expressed using receptor fluid data alone. However, when the test substance remains in the skin at the end of the study, it may need to be included in the total amount absorbed (see paragraph 66 in reference (3)). When infinite dose conditions of exposure are used the data may permit the calculation of a permeability constant ( $K_p$ ). Under the latter conditions, the percentage absorbed is not relevant.

**3. REPORTING****3.1. TEST REPORT**

The test report must include the requirements stipulated in the protocol, including a justification for the test system used and should, comprise the following:

test substance:

- physical nature, physicochemical properties (at least molecular weight and  $\log P_{ow}$ ), purity (radiochemical purity),
- identification information (e.g. batch number),
- solubility in receptor fluid.

Test preparation:

- formulation and justification of use,
- homogeneity.

Test conditions:

- sources and site of skin, method of preparation, storage conditions prior to use, any pre-treatment (cleaning, antibiotic treatments, etc.), skin integrity measurements, metabolic status, justification of use,
- cell design, receptor fluid composition, receptor fluid flow rate or sampling times and procedures,
- details of application of test preparation and quantification of dose applied,
- duration of exposure,
- details of removal of test preparation from the skin, e.g. skin rinsing,
- details of analysis of skin and any fractionation techniques employed to demonstrate skin distribution,

**▼B**

- cell and equipment washing procedures,
- assay methods, extraction techniques, limits of detection and analytical method validation.

## Results:

- overall recoveries of the experiment (Applied dose  $\equiv$  Skin washings + Skin + Receptor fluid + Cell washings),
- tabulation of individual cell recoveries in each compartment,
- absorption profile,
- tabulated absorption data (expressed as rate, amount or percentage).

## Discussion of results.

## Conclusions.

4. **REFERENCES**

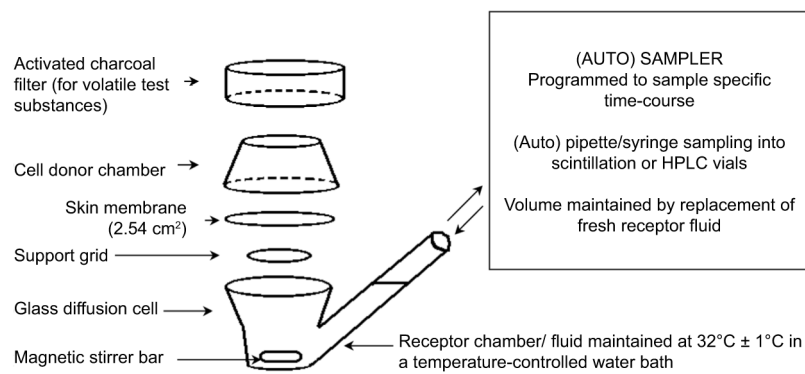
- (1) Testing Method B.44. Skin Absorption: *In Vivo* Method.
- (2) OECD, (2002) Guidance Document for the Conduct of Skin Absorption Studies. OECD, Paris.
- (3) OECD, (2000) Report of the Meeting of the OECD Extended Steering Committee for Percutaneous Absorption Testing, Annex 1 to ENV/JM/TG(2000)5. OECD, Paris.
- (4) Kemppainen B.W. and Reifenrath W.G., (1990) Methods for skin absorption. CRC Press, Boca Raton.
- (5) Bronaugh R.L. and Collier, S.W., (1991) Protocol for In vitro Percutaneous Absorption Studies, in *In vitro Percutaneous Absorption: Principles, Fundamentals and Applications*, RL Bronaugh and HI Maibach, Eds., CRC Press, Boca Raton, p. 237-241.
- (6) Bronaugh R.L. and Maibach H.I., (1991) *In vitro Percutaneous Absorption: Principles, Fundamentals and Applications*. CRC Press, Boca Raton.
- (7) European Centre for Ecotoxicology and Toxicology of Chemicals, (1993) Monograph No 20, Percutaneous Absorption, ECETOC, Brussels.
- (8) Diembeck W, Beck H, Benech-Kieffer F, Courtellemont P, Dupuis J, Lovell W, Paye M, Spengler J, Steiling W., (1999) Test Guidelines for *In Vitro* Assessment of Dermal Absorption and Percutaneous Penetration of Cosmetic Ingredients, *Fd Chem Tox*, 37, p. 191-205.
- (9) Recommended Protocol for *In vitro* Percutaneous Absorption Rate Studies (1996). US Federal Register, Vol. 61, No 65.
- (10) Howes, D., Guy, R., Hadgraft, J., Heylings, J.R. *et al.* (1996). Methods for assessing Percutaneous absorption. ECVAM Workshop Report ATLA 24, 81 R10.

▼ B

- (11) Schaefer, H. and Redelmeier, T.E., (1996). Skin barrier: principles of percutaneous absorption. Karger, Basel.
- (12) Roberts, M.S. and Walters, K.A., (1998). Dermal absorption and toxicity assessment. Marcel Dekker, New York.
- (13) Jewell, C., Heylings, J.R., Clowes, H.M. and Williams, F.M. (2000). Percutaneous absorption and metabolism of dinitrochlorobenzene in vitro. Arch Toxicol 74, p. 356-365.

Figure 1

**An example of a typical design of a static diffusion cell for *in vitro* percutaneous absorption studies**





**▼ M3****B.46. *IN VITRO* SKIN IRRITATION: RECONSTRUCTED HUMAN EPIDERMIS TEST METHOD**

## INTRODUCTION

1. Skin irritation refers to the production of reversible damage to the skin following the application of a test chemical for up to 4 hours (as defined by the United Nations (UN) Globally Harmonised System of Classification and Labelling of Chemicals (GHS) and Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures (1) (3)). This Test Method (TM) provides an *in vitro* procedure that may be used for the hazard identification of irritant chemicals (substances and mixtures) in accordance with UN GHS and EU CLP Category 2 (1) (2) (3). In the EU and other regions, that have not adopted the optional UN GHS Category 3 (mild irritants), this TM can also be used to identify non-classified chemicals, i.e. UN GHS and EU CLP 'No Category' (1) (3). This TM may be used to determine the skin irritancy of chemicals as a stand-alone replacement test for *in vivo* skin irritation testing within a tiered testing strategy (4 and Chapter B.4 in this Annex).
2. The assessment of skin irritation has typically involved the use of laboratory animals (OECD Test Guideline 404; Chapter B.4 in this Annex) (4). In relation to animal welfare concerns, B.4 was revised in 2004 allowing for the determination of skin corrosion/irritation by applying a tiered testing strategy, using validated *in vitro* or *ex vivo* test methods, thus avoiding pain and suffering of animals. Three validated *in vitro* test methods have been adopted as OECD Test Guidelines 430, 431 and 435 (5) (6) (7) and two of them as Chapters B.40 and B.40a of this Annex, to be used for the corrosivity part of the tiered testing strategy of B.4 or OECD Test Guideline 404 (4).
3. This TM addresses the human health endpoint skin irritation. It is based on reconstructed human epidermis (RhE), which in its overall design (the use of human derived non-transformed epidermis keratinocytes as cell source and use of representative tissue and cytoarchitecture) closely mimics the biochemical and physiological properties of the upper parts of the human skin, i.e. the epidermis. This TM also includes a set of Performance Standards (PS) (Appendix 2) for the assessment of similar and modified RhE-based methods developed by EC-ECVAM (8), in accordance with the principles of OECD Guidance Document No 34 (9).
4. There are three validated methods that adhere to this TM. Prevalidation, optimisation and validation studies have been completed for an *in vitro* method (10) (11) (12) (13) (14) (15) (16) (17) (18) (19) (20), using a RhE model, commercially available as EpiSkin™ (designated the Validated Reference Method — VRM). Two other commercially available *in vitro* skin irritation RhE methods have shown similar results to the VRM according to PS-based validation (21), and these are the EpiDerm™ SIT (EPI-200) and the SkinEthic™ RHE methods (22).

**▼ M3**

5. Before a proposed similar or modified *in vitro* RhE method other than the VRM, EpiDerm™ SIT (EPI-200) or SkinEthic™ RHE methods can be used for regulatory purposes, its reliability, relevance (accuracy), and limitations for its proposed use should be determined in order to ensure that it can be regarded as similar to that of the VRM, in accordance with the requirements of the PS set out in this TM (Appendix 2). Moreover, it is recommended to consult the OECD Explanatory Background Document on *in vitro* skin irritation testing before developing and validating a similar or modified *in vitro* RhE method and submitting it for regulatory adoption (23).

## DEFINITIONS

6. Definitions used are provided in Appendix 1.

## INITIAL CONSIDERATIONS AND LIMITATIONS

7. A limitation of the TM, as demonstrated by the validation study (16), is that it does not allow the classification of chemicals to the optional UN GHS Category 3 (mild irritants) (1). When used as a partial replacement test, follow-up *in vivo* testing may be required to fully characterise skin irritation potential (4 and Chapter B.4 of this Annex). It is recognised that the use of human skin is subject to national and international ethical considerations and conditions.
8. This TM addresses the *in vitro* skin irritation component of the tiered testing strategy of B.4 (OECD Test Guideline 404) on dermal corrosion/irritation (4). While this TM does not provide adequate information on skin corrosion, it should be noted that B.40a (OECD Test Guideline 431) on skin corrosion is based on the same RhE test system, though using another protocol (Chapter B.40a). This method is based on RhE-models using human keratinocytes, which therefore represent *in vitro* the target organ of the species of interest. It moreover directly covers the initial step of the inflammatory cascade/mechanism of action (cell damage and tissue damage resulting in localised trauma) that occurs during irritation *in vivo*. A wide range of chemicals has been tested in the validation underlying this TM and the empirical database of the validation study amounted to 58 chemicals in total (16) (18) (23). This is applicable to solids, liquids, semi-solids and waxes. The liquids may be aqueous or non-aqueous; solids may be soluble or insoluble in water. Whenever possible, solids should be ground to a fine powder before application; no other pre-treatment of the sample is required. Gases and aerosols have not been assessed yet in a validation study (24). While it is conceivable that these can be tested using RhE technology, the current TM does not allow testing of gases and aerosols. It should also be noted that highly coloured chemicals may interfere with the cell viability measurements and need the use of adapted controls for corrections (see paragraphs 24-26).
9. A single testing run composed of three replicate tissues should be sufficient for a test chemical when the classification is unequivocal. However, in cases of borderline results, such as non-concordant replicate measurements and/or mean percent viability equal to  $50 \pm 5\%$ , a second run should be considered, as well as a third one in case of discordant results between the first two runs.

▼ **M3**

## PRINCIPLE OF THE TEST

10. The test chemical is applied topically to a three-dimensional RhE model, comprised of non-transformed human-derived epidermal keratinocytes, which have been cultured to form a multilayered, highly differentiated model of the human *epidermis*. It consists of organised basal, spinous and granular layers, and a multilayered *stratum corneum* containing inter-cellular lamellar lipid layers representing main lipid classes analogous to those found *in vivo*.
11. Chemical-induced skin irritation, manifested by erythema and oedema, is the result of a cascade of events beginning with penetration of the *stratum corneum* and damage to the underlying layers of keratinocytes. The dying keratinocytes release mediators that begin the inflammatory cascade which acts on the cells in the *dermis*, particularly the stromal and endothelial cells. It is the dilatation and increased permeability of the endothelial cells that produce the observed erythema and oedema (24). The RhE-based methods measure the initiating events in the cascade.
12. Cell viability in RhE models is measured by enzymatic conversion of the vital dye MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, Thiazolyl blue); CAS number 298-93-1 into a blue formazan salt that is quantitatively measured after extraction from tissues (25). Irritant chemicals are identified by their ability to decrease cell viability below defined threshold levels (*i.e.*  $\leq 50\%$ , for UN GHS/EU CLP Category 2). Depending on the regulatory framework in which the results of this TM are used, chemicals that produce cell viabilities above the defined threshold level, may be considered non-irritants (*i.e.*  $> 50\%$ , No Category).

## DEMONSTRATION OF PROFICIENCY

13. Prior to routine use of any of the three validated methods that adhere to this TM, laboratories should demonstrate technical proficiency, using the ten Reference Chemicals listed in Table 1. For similar methods developed under this TM or for modifications of any of the three validated methods, the PS requirements described in Appendix 2 of this TM should be met prior to using the method for regulatory testing.
14. As part of the proficiency exercise, it is recommended that the user verifies the barrier properties of the tissues after receipt as specified by the RhE model producer. This is particularly important if tissues are shipped over long distance/time periods. Once a method has been successfully established and proficiency in its use has been demonstrated, such verification will not be necessary on a routine basis. However, when using a method routinely, it is recommended to continue to assess the barrier properties in regular intervals.

Table 1

Reference Chemicals <sup>(1)</sup>

Chemical	CAS NR	<i>In vivo</i> score <sup>(2)</sup>	Physical state	UN GHS/EU CLP Category
naphthalene acetic acid	86-87-3	0	Solid	No Cat.
isopropanol	67-63-0	0,3	Liquid	No Cat.

## ▼ M3

Chemical	CAS NR	<i>In vivo</i> score <sup>(2)</sup>	Physical state	UN GHS/EU CLP Category
methyl stearate	112-61-8	1	Solid	No Cat.
heptyl butyrate	5870-93-9	1,7	Liquid	No Cat. (Optional Cat. 3) <sup>(3)</sup> , <sup>(4)</sup>
hexyl salicylate	6259-76-3	2	Liquid	No Cat. (Optional Cat. 3) <sup>(3)</sup> , <sup>(4)</sup>
cyclamen aldehyde	103-95-7	2,3	Liquid	Cat. 2
1-bromohexane	111-25-1	2,7	Liquid	Cat. 2
potassium hydroxide (5 % aq.)	1310-58-3	3	Liquid	Cat. 2
1-methyl-3-phenyl-1-piperazine	5271-27-2	3,3	Solid	Cat. 2
Heptanal	111-71-7	3,4	Liquid	Cat. 2

<sup>(1)</sup> These Reference Chemicals are a subset of the Reference Chemicals used in the validation study.

<sup>(2)</sup> *In vivo* score in accordance with B.4 and OECD Test Guideline 404 <sup>(4)</sup>.

<sup>(3)</sup> Under this Test Method, the UN GHS optional Category 3 (mild irritants) (1) is considered as No Category.

<sup>(4)</sup> The UN GHS optional Category 3 is not applicable under the EU CLP.

## PROCEDURE

15. The following is a description of the components and procedures of a RhE method for skin irritation assessment. A RhE model should be reconstructed, and can be in-house-prepared or obtained commercially. Standard Operating Procedures (SOPs) for the EpiSkin<sup>TM</sup>, EpiDerm<sup>TM</sup> SIT (EPI-200) and SkinEthic<sup>TM</sup> RHE are available (26) (27) (28). Testing should be performed according to the following:

**RhE Test Method Components***General conditions*

16. Non-transformed human keratinocytes should be used to reconstruct the epithelium. Multiple layers of viable epithelial cells (basal layer, *stratum spinosum*, *stratum granulosum*) should be present under a functional *stratum corneum*. *Stratum corneum* should be multilayered containing the essential lipid profile to produce a functional barrier with robustness to resist rapid penetration of cytotoxic marker chemicals, e.g. sodium dodecyl sulphate (SDS) or Triton X-100. The barrier function should be demonstrated and may be assessed either by determination of the concentration at which a marker chemical reduces the viability of the tissues by 50 % (IC<sub>50</sub>) after a fixed exposure time, or by determination of the exposure time required to reduce cell viability by 50 % (ET<sub>50</sub>) upon application of the marker chemical at a specified, fixed concentration. The containment properties of the RhE model should prevent the passage of material around the *stratum corneum* to the viable tissue, which would lead to poor modelling of skin exposure. The RhE model should be free of contamination by bacteria, viruses, mycoplasma, or fungi.

▼ **M3***Functional conditions*

## Viability

17. The assay used for determining the magnitude of viability is the MTT-assay (25). The RhE model users should ensure that each batch of the RhE model used meets defined criteria for the negative control (NC). The optical density (OD) of the extraction solvent alone should be sufficiently small, *i.e.*  $OD < 0,1$ . An acceptability range (upper and lower limit) for the negative control OD values (in the Skin Irritation Test Method conditions) are established by the RhE model developer/supplier, and the acceptability ranges for the 3 validated methods are given in Table 2. It should be documented that the tissues treated with NC are stable in culture (provide similar viability measurements) for the duration of the test exposure period.

Table 2

**Acceptability ranges for negative control OD values**

	Lower acceptance limit	Upper acceptance limit
EpiSkin™ (SM)	$\geq 0,6$	$\leq 1,5$
EpiDerm™ SIT (EPI-200)	$\geq 1,0$	$\leq 2,5$
SkinEthic™ RHE	$\geq 1,2$	$\leq 2,5$

## Barrier function

18. The *stratum corneum* and its lipid composition should be sufficient to resist the rapid penetration of cytotoxic marker chemicals, *e.g.* SDS or Triton X-100, as estimated by  $IC_{50}$  or  $ET_{50}$  (Table 3).

## Morphology

19. Histological examination of the RhE model should be performed demonstrating human *epidermis*-like structure (including multilayered *stratum corneum*).

## Reproducibility

20. The results of the positive control chemical (PC) and negative controls (NC) of the test method should demonstrate reproducibility over time.

## Quality control (QC)

21. The RhE model developer/supplier should ensure and demonstrate that each batch of the RhE model used meets defined production release criteria, among which those for *viability* (paragraph 17), *barrier function* (paragraph 18) and *morphology* (paragraph 19) are the most relevant. These data should be provided to the method users, so that they are able to include this information in the test report. An acceptability range (upper and lower limit) for the  $IC_{50}$  or the  $ET_{50}$  should be established by the RhE model developer/supplier (or investigator when using an in-house model). Only results produced with qualified tissues can be accepted for reliable prediction of irritation classification. As an example, the acceptability ranges for the three validated methods are given in Table 3.

▼ **M3**

Table 3

**Examples of QC batch release criteria**

	Lower acceptance limit	Upper acceptance limit
EpiSkin™ (SM) (18 hours treatment with SDS) (26)	IC <sub>50</sub> = 1,0 mg/ml	IC <sub>50</sub> = 3,0 mg/ml
EpiDerm™ SIT (EPI-200) (1 % Triton X-100) (27)	ET <sub>50</sub> = 4,8 hr	ET <sub>50</sub> = 8,7 hr
SkinEthic™ RHE (1 % Triton X-100) (28)	ET <sub>50</sub> = 4,0 hr	ET <sub>50</sub> = 9,0 hr

*Application of the Test and Control Chemicals*

22. At least three replicates should be used for each test chemical and for the controls in each run. For liquids as well as solids, sufficient amount of test chemical should be applied to uniformly cover the *epidermis* surface while avoiding an infinite dose, *i.e.* a minimum of 25 µL/cm<sup>2</sup> or 25 mg/cm<sup>2</sup> should be used. For solids, the *epidermis* surface should be moistened with deionised or distilled water before application, to improve contact between the test chemical and the *epidermis* surface. Whenever possible, solids should be tested as a fine powder. At the end of the exposure period, the test chemical should be carefully washed from the *epidermis* surface with aqueous buffer, or 0,9 % NaCl. Depending on which of the three validated RhE methods is used, the exposure period varies between 15 and 60 minutes, and the incubation temperature between 20 and 37 °C. These exposure periods and temperatures are optimised for each RhE method and represent the different intrinsic properties of the methods, for details, see the Standard Operating Procedures (SOPs) for the methods (26) (27) (28).
23. Concurrent NC and PC should be used in each run to demonstrate that viability (with the NC), barrier function and resulting tissue sensitivity (with the PC) of the tissues are within a defined historical acceptance range. The suggested PC is 5 % aqueous SDS. The suggested NC chemicals are water or phosphate buffered saline (PBS).

*Cell Viability Measurements*

24. The most important element of the test procedure is that viability measurements are not performed immediately after the exposure to the test chemicals, but after a sufficiently long post-treatment incubation period of the rinsed tissues in fresh medium. This period allows both for recovery from weak cytotoxic effects and for appearance of clear cytotoxic effects. The test optimisation phase (11) (12) (13) (14) (15) demonstrated that a 42 hours post-treatment incubation period was optimal.
25. The MTT assay is a validated quantitative method which should be used to measure cell viability under this TM. It is compatible with use in a three-dimensional tissue construct. The tissue sample is placed in MTT solution of appropriate concentration (*e.g.* 0,3-1 mg/mL) for 3 hours. The precipitated blue formazan product is then extracted from the tissue using a solvent (*e.g.* isopropanol, acidic isopropanol), and the concentration of formazan is measured by determining the OD at 570 nm using a filter band pass of maximum ± 30 nm.

**▼ M3**

26. Optical properties of the test chemical or its chemical action on the MTT may interfere with the assay leading to a false estimate of viability (because the test chemical may prevent or reverse the colour generation as well as cause it). This may occur when a specific test chemical is not completely removed from the tissue by rinsing or when it penetrates the *epidermis*. If a test chemical acts directly on the MTT (MTT-reducer), is naturally coloured, or becomes coloured during tissue treatment, additional controls should be used to detect and correct for test chemical interference with the viability measurement technique. Detailed description of how to correct direct MTT reduction and interferences by colouring agents is available in the SOPs for the three validated methods (26) (27) (28).

*Acceptability Criteria*

27. For each method using valid RhE model batches (see paragraph 21), tissues treated with the NC should exhibit OD reflecting the quality of the tissues that followed shipment, receipt steps and all protocol processes. Control OD values should not be below historically established boundaries. Similarly, tissues treated with the PC, *i.e.* 5 % aqueous SDS, should reflect their ability to respond to an irritant chemical under the conditions of the TM(26) (27) (28). Associated and appropriate measures of variability between tissue replicates should be defined (*e.g.* if standard deviations (SD) are used they should be within the 1-sided 95 % tolerance interval calculated from historical data; for the VRM SD < 18 %).

*Interpretation of Results and Prediction Model*

28. The OD values obtained with each test chemical can be used to calculate the percentage of viability normalised to NC, which is set to 100 %. The cut-off value of percentage cell viability distinguishing irritant from non-classified test chemicals and the statistical procedure(s) used to evaluate the results and identify irritant chemicals should be clearly defined, documented, and proven to be appropriate. The cut-off values for the prediction of irritation are given below:

- the test chemical is considered to be irritant to skin in accordance with UN GHS/EU CLP Category 2 if the tissue viability after exposure and post-treatment incubation is less than or equal ( $\leq$ ) to 50 %.
  
- depending on the regulatory framework in which the results of this TM are used, the test chemical may be considered to be non-irritant to skin in accordance with UN GHS/EU CLP No Category if the tissue viability after exposure and post-treatment incubation is more than ( $>$ ) 50 %.

**DATA AND REPORTING***Data*

29. For each run, data from individual replicate tissues (*e.g.* OD values and calculated percentage cell viability data for each test chemical, including classification) should be reported in tabular form, including data from repeat experiments as appropriate. In addition means  $\pm$  SD for each run should be reported. Observed interactions with MTT reagent and coloured test chemicals should be reported for each tested chemical.

**▼ M3***Test Report*

30. The test report should include the following information:

Test and Control Chemicals:

- Chemical name(s) such as CAS name and number, EC name and number, if known;
- Purity and composition of the chemical (in percentage(s) by weight);
- Physical/chemical properties relevant to the conduct of the study (*e.g.* physical state, stability, volatility, pH and water solubility if known);
- Treatment of the test/control chemicals prior to testing, if applicable (*e.g.* warming, grinding);
- Storage conditions;

Justification of the RhE model and protocol used

Test Conditions:

- Cell system used;
- Complete supporting information for the specific RhE model used including its performance. This should include, but is not limited to:
  - i) viability
  - ii) barrier function
  - iii) morphology
  - iv) reproducibility and predictivity
  - v) Quality controls (QC) of the model
- Details of the test procedure used;
- Test doses used, duration of exposure and post treatment incubation period;
- Description of any modifications of the test procedure;
- Reference to historical data of the model. This should include, but is not limited to:
  - i) acceptability of the QC data with reference to historical batch data
  - ii) acceptability of the positive and negative control values with reference to positive and negative control means and ranges
- Description of evaluation criteria used including the justification for the selection of the cut-off point(s) for the prediction model;
- Reference to historical control data;



▼ **M3**

Results:

- Tabulation of data from individual test chemicals for each run and each replicate measurement;
- Indication of controls used for direct MTT-reducers and/or colouring test chemicals;
- Description of other effects observed;

Discussion of the results

Conclusion

## LITERATURE

- (1) UN (2009), United Nations Globally Harmonised System of Classification and Labelling of Chemicals (GHS), Third revised edition, UN New York and Geneva. Available at: [[http://www.unece.org/trans/danger/publi/ghs/ghs\\_rev03/03files\\_e.html](http://www.unece.org/trans/danger/publi/ghs/ghs_rev03/03files_e.html)]
- (2) EC-ECVAM (2009), Statement on the 'Performance under UN GHS of three *in vitro* assays for skin irritation testing and the adaptation of the Reference Chemicals and Defined Accuracy Values of the ECVAM skin irritation Performance Standards', issued by the ECVAM Scientific Advisory Committee (ESAC30), 9 April 2009. Available at: [<http://ecvam.jrc.ec.europa.eu>]
- (3) Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. OJ L 353, 31.12.2008, p. 1.
- (4) OECD (2004), Acute Dermal Irritation/Corrosion, OECD Guideline for the Testing of Chemicals No 404, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (5) OECD (2004), *In Vitro* Skin Corrosion: Transcutaneous Electrical Resistance (TER), OECD Guideline for the Testing of Chemicals No 430, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (6) OECD (2004), *In Vitro* Skin Corrosion: Human Skin Model Test, OECD Guideline for the Testing of Chemicals No 431, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (7) OECD (2006), *In Vitro* Membrane Barrier Test Method for Skin Corrosion, OECD Guideline for the Testing of Chemicals No 435, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (8) EC-ECVAM (2009), Performance Standards for *in vitro* skin irritation test methods based on Reconstructed human Epidermis (RhE)? Available at: [<http://ecvam.jrc.ec.europa.eu>]
- (9) OECD (2005), Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment, OECD Series on Testing and Assessment No 34, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (10) Fentem, J.H., Briggs, D., Chesné, C., Elliot, G.R., Harbell, J.W., Heylings, J.R., Portes, P., Roguet, R., van de Sandt, J.J. M. and Botham, P. (2001), A prevalidation study on *in vitro* tests for acute skin irritation, Results and evaluation by the Management Team, *Toxicol. in Vitro* 15, 57-93.
- (11) Portes, P., Grandidier, M.-H., Cohen, C. and Roguet, R. (2002), Refinement of the EPISKIN protocol for the assessment of acute skin irritation of chemicals: follow-up to the ECVAM prevalidation study, *Toxicol. in Vitro* 16, 765-770.

▼ **M3**

- (12) Kandárová, H., Liebsch, M., Genschow, E., Gerner, I., Traue, D., Slawik, B. and Spielmann, H. (2004), Optimisation of the EpiDerm test protocol for the upcoming ECVAM validation study on *in vitro* skin irritation tests, ALTEX 21, 107-114.
- (13) Kandárová, H., Liebsch, M., Gerner, I., Schmidt, E., Genschow, E., Traue, D. and Spielmann, H. (2005), The EpiDerm test protocol for the upcoming ECVAM validation study on *in vitro* skin irritation tests — An assessment of the performance of the optimised test, ATLA 33, 351-367.
- (14) Cotovio, J., Grandidier, M.-H., Portes, P., Roguet, R. and Rubinsteen, G. (2005), The *in vitro* acute skin irritation of chemicals: optimisation of the EPISKIN prediction model within the framework of the ECVAM validation process, ATLA 33, 329-349.
- (15) Zuang, V., Balls, M., Botham, P.A., Coquette, A., Corsini, E., Curren, R.D., Elliot, G.R., Fentem, J.H., Heylings, J.R., Liebsch, M., Medina, J., Roguet, R., van De Sandt, J.J.M., Wiemann, C. and Worth, A. (2002), Follow-up to the ECVAM prevalidation study on *in vitro* tests for acute skin irritation, The European Centre for the Validation of Alternative Methods Skin Irritation Task Force report 2, ATLA 30, 109-129.
- (16) Spielmann, H., mailto:Hoffmann, S., Liebsch, M., Botham, P., Fentem, J., Eskes, C., Roguet, R., Cotovio, J., Cole, T., Worth, A., Heylings, J., Jones, P., Robles, C., Kandárová, H., mailto:Gamer, A., Remmele, M., Curren, R., Raabe, H., Cockshott, A., Gerner, I. and Zuang, V. (2007), The ECVAM international validation study on *in vitro* tests for acute skin irritation: Report on the validity of the EPISKIN and EpiDerm assays and on the skin integrity function test, ATLA 35, 559-601.
- (17) Hoffmann, S. (2006), ECVAM skin irritation validation study phase II: Analysis of the primary endpoint MTT and the secondary endpoint IL1- $\alpha$ . Available at: [<http://ecvam.jrc.ec.europa.eu>]
- (18) Eskes, C., Cole, T., Hoffmann, S., Worth, A., Cockshott, A., Gerner, I. and Zuang, V. (2007), The ECVAM international validation study on *in vitro* tests for acute skin irritation: selection of test chemicals, ATLA 35, 603-619.
- (19) Cotovio, J., Grandidier, M.-H., Lelièvre, D., Roguet, R., Tinois-Tessonnaud, E. and Leclaire, J. (2007), *In vitro* acute skin irritancy of chemicals using the validated EPISKIN model in a tiered strategy — Results and performances with 184 cosmetic ingredients, AATEX, 14, 351-358.
- (20) EC-ECVAM (2007), Statement on the validity of *in vitro* tests for skin irritation, issued by the ECVAM Scientific Advisory Committee (ESAC26), 27 April 2007. Available at: [<http://ecvam.jrc.ec.europa.eu>]
- (21) EC-ECVAM (2007), Performance Standards for applying human skin models to *in vitro* skin irritation testing. Available at: [<http://ecvam.jrc.ec.europa.eu>]
- (22) EC-ECVAM (2008), Statement on the scientific validity of *in vitro* tests for skin irritation testing, issued by the ECVAM Scientific Advisory Committee (ESAC29), 5 November 2008. Available at: [<http://ecvam.jrc.ec.europa.eu>]
- (23) OECD (2010), Explanatory background document to the OECD draft Test Guideline on *in vitro* skin irritation testing. OECD Series on Testing and Assessment, No 137, OECD, Paris. Available at: [[http://www.oecd.org/document/24/0,3746,en\\_2649\\_34377\\_47858904\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/24/0,3746,en_2649_34377_47858904_1_1_1_1,00.html)]
- (24) Welss, T., Basketter, D.A. and Schröder, K.R. (2004), *In vitro* skin irritation: fact and future. State of the art review of mechanisms and models, Toxicol. in Vitro 18, 231-243.

▼ **M3**

- (25) Mosmann, T. (1983), Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays, *J. Immunol. Methods* 65, 55-63.
- (26) EpiSkin™ SOP, Version 1.8 (February 2009), ECVAM Skin Irritation Validation Study: Validation of the EpiSkin™ test method 15 min-42 hours for the prediction of acute skin irritation of chemicals. Available at: [<http://ecvam.jrc.ec.europa.eu>]
- (27) EpiDerm™ SOP, Version 7.0 (Revised March 2009), Protocol for: *In vitro* EpiDerm™ skin irritation test (EPI-200-SIT), For use with MatTek Corporation's reconstructed human epidermal model EpiDerm (EPI-200). Available at: [<http://ecvam.jrc.ec.europa.eu>]
- (28) SkinEthic™ RHE SOP, Version 2.0 (February 2009), SkinEthic skin irritation test-42a test method for the prediction of acute skin irritation of chemicals: 42 minutes application + 42 hours post-incubation. Available at: [<http://ecvam.jrc.ec.europa.eu>]
- (29) Harvell, J.D., Lamminstausta, K., and Maibach, H.I. (1995), Irritant contact dermatitis, In: *Practical Contact Dermatitis*, pp 7-18, (Ed. Guin J. D.). Mc Graw-Hill, New York.
- (30) Commission Directive 2001/59/EC of 6 August 2001 adapting to technical progress for the 28th time Council Directive 67/548/EEC on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances, *OJ L* 225, 21.8.2001, p. 1.
- (31) Basketter, D.A., York, M., McFadden, J.P. and Robinson, M.K. (2004), Determination of skin irritation potential in the human 4-h patch test. *Contact Dermatitis* 51, 1-4.
- (32) Jirova, D., Liebsch, M., Basketter, D., Spiller, E., Kejlova, K., Bendova, H., Marriott, M. and Kandarova, H. (2007), Comparison of human skin irritation and photo-irritation patch test data with cellular *in vitro* assays and animal *in vivo* data, *ALTEX*, 14, 359-365.
- (33) Jirová, D., Basketter, D., Liebsch, M., Bendová, H., Kejlová, K., Marriott, M. and Kandárová, H. (2010), Comparison of human skin irritation patch test data with *in vitro* skin irritation assays and animal data, *Contact Dermatitis*, 62, 109-116.

▼ **M3***Appendix 1***Definitions**

*Accuracy*: The closeness of agreement between test method results and accepted reference values. It is a measure of test method performance and one aspect of relevance. The term is often used interchangeably with ‘concordance’ to mean the proportion of correct outcomes of a test method (9).

*Cell viability*: Parameter measuring total activity of a cell population *e.g.* as ability of cellular mitochondrial dehydrogenases to reduce the vital dye MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, Thiazolyl blue), which depending on the endpoint measured and the test design used, correlates with the total number and/or vitality of living cells.

*Concordance*: This is a measure of test method performance for test methods that give a categorical result, and is one aspect of relevance. The term is used interchangeably with accuracy, and is defined as the proportion of all chemicals tested that are correctly classified as positive or negative. (9).

*ET<sub>50</sub>*: Can be estimated by determination of the exposure time required to reduce cell viability by 50 % upon application of the marker chemical at a specified, fixed concentration, see also IC<sub>50</sub>.

*EU CLP (Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of Substances and Mixtures)*: Implements in the European Union (EU) the UN GHS system for the classification and labelling of chemicals (substances and mixtures) (3).

*GHS (Globally Harmonised System of Classification and Labelling of Chemicals by the United Nations (UN))*: A system proposing the classification of chemicals (substances and mixtures) according to standardised types and levels of physical, health and environmental hazards, and addressing corresponding communication elements, such as pictograms, signal words, hazard statements, precautionary statements and safety data sheets, so that to convey information on their adverse effects with a view to protect people (including employers, workers, transporters, consumers and emergency responders) and the environment (1).

*IC<sub>50</sub>*: Can be estimated by determination of the concentration at which a marker chemical reduces the viability of the tissues by 50 % (IC<sub>50</sub>) after a fixed exposure time, see also ET<sub>50</sub>.

*Infinite dose*: Amount of test chemical applied to the *epidermis* exceeding the amount required to completely and uniformly cover the *epidermis* surface.

*Me-too test*: A colloquial expression for a test method that is structurally and functionally similar to a validated and accepted reference test method. Such a test method would be a candidate for catch-up validation. Interchangeably used with similar test method (9).

*Performance standards (PS)*: Standards, based on a validated test method, that provide a basis for evaluating the comparability of a proposed test method that is mechanistically and functionally similar. Included are; (i) essential test method components; (ii) a minimum list of Reference Chemicals selected from among the chemicals used to demonstrate the acceptable performance of the validated test method; and (iii) the comparable levels of accuracy and reliability, based on what was obtained for the validated test method, that the proposed test method should demonstrate when evaluated using the minimum list of Reference Chemicals (9).

**▼ M3**

*Reference chemicals:* Chemicals selected for use in the validation process, for which responses in the *in vitro* or *in vivo* reference test system or the species of interest are already known. These chemicals should be representative of the classes of chemicals for which the test method is expected to be used, and should represent the full range of responses that may be expected from the chemicals for which it may be used, from strong, to weak, to negative. Different sets of reference chemicals may be required for the different stages of the validation process, and for different test methods and test uses (9).

*Relevance:* Description of relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method (9).

*Reliability:* Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility (9).

*Replacement test:* A test which is designed to substitute for a test that is in routine use and accepted for hazard identification and/or risk assessment, and which has been determined to provide equivalent or improved protection of human or animal health or the environment, as applicable, compared to the accepted test, for all possible testing situations and chemicals (9).

*Sensitivity:* The proportion of all positive/active test chemicals that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results, and is an important consideration in assessing the relevance of a test method (9).

*Skin irritation:* The production of reversible damage to the skin following the application of a test chemical for up to 4 hours. Skin irritation is a locally arising, non-immunogenic reaction, which appears shortly after stimulation (29). Its main characteristic is its reversible nature involving inflammatory reactions and most of the clinical characteristic signs of irritation (erythema, oedema, itching and pain) related to an inflammatory process.

*Specificity:* The proportion of all negative/inactive test chemicals that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results and is an important consideration in assessing the relevance of a test method (9).

*Tiered testing strategy:* Testing which uses test methods in a sequential manner; the test methods selected in each succeeding level are decided based on the results in the previous level of testing (9).

*Test chemical (also referred to as test substance):* Any substance or mixture tested using this TM.

▼ **M3***Appendix 2***Performance Standards for assessment of proposed similar or modified *in vitro* reconstructed human epidermis (RhE) methods for skin irritation**

## INTRODUCTION

1. The purpose of Performance Standards (PS) is to communicate the basis by which new methods, both proprietary (*i.e.* copyrighted, trademarked, registered) and non-proprietary can be determined to have sufficient accuracy and reliability for specific testing purposes. These PS, based on validated and accepted methods, can be used to evaluate the reliability and accuracy of other analogous methods (colloquially referred to as ‘me-too’ tests) that are based on similar scientific principles and measure or predict the same biological or toxic effect (9).
2. Prior to adoption of modified methods, *i.e.* proposed potential improvements to an approved method, there should be an evaluation to determine the effect of the proposed changes on the test’s performance and the extent to which such changes affect the information available for the other components of the validation process. Depending on the number and nature of the proposed changes, the generated data and supporting documentation for those changes, they should either be subjected to the same validation process as described for a new test, or, if appropriate, to a limited assessment of reliability and relevance using established PS (9).
3. Similar (me-too) or modified methods of any of the three validated methods (EpiSkin™ (Validated Reference Method — VRM), EpiDerm™ SIT (EPI-200) and SkinEthic™ RHE) proposed for use under this TM should be evaluated to determine their reliability and accuracy using chemicals representing the full range of the Draize irritancy scores. When evaluated using the 20 recommended Reference Chemicals of the PS (Table 1), the proposed similar or modified methods should have reliability and accuracy values which are comparable or better than those derived from the VRM (Table 2) (2) (16). The reliability and accuracy values that should be achieved are provided in paragraphs 8 to 12 of this Appendix. Non-classified (UN GHS/EU CLP No Category) and classified (UN GHS/EU CLP Category 2) (1) chemicals, representing different chemical classes are included, so that the reliability and accuracy (sensitivity, specificity and overall accuracy) of the proposed method can be compared to that of the VRM. The reliability of the method, as well as its ability to correctly identify UN GHS/EU CLP Category 2 irritant chemicals and, depending on the regulatory framework for which data are produced, also its ability to correctly identify UN GHS/EU CLP No Category chemicals (should be determined prior to its use for testing new test chemicals).
4. These PS are based on the EC-ECVAM PS (8), updated according to the UN GHS and EU CLP systems on classification and labelling (1) (3). The original PS were defined after the completion of the validation study (21) and were based on the EU classification system as laid down in Commission Directive 2001/59/EC of 6 August 2001 adapting to technical progress for the 28th time Council Directive 67/548/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances (1). Due to the adoption of the UN GHS system for classification

(1) OJ L 225, 21.8.2001, p. 1.

**▼ M3**

and labelling in EU (EU CLP) (3), which took place between the finalisation of the validation study and the completion of this TM, the PS have been updated (8). This update concerns mainly changes (i), in the set of the PS Reference Chemicals; and (ii), the defined reliability and accuracy values (2) (23).

PERFORMANCE STANDARDS FOR *IN VITRO* RhE TEST METHODS FOR SKIN IRRITATION

5. The PS comprises the following three elements (9):

- I) Essential Test Method Components
- II) Minimum List of Reference Chemicals
- III) Defined Reliability and Accuracy Values

**I) Essential Test Method Components**

6. These consist of essential structural, functional, and procedural elements of a validated method that should be included in the protocol of a proposed, mechanistically and functionally similar or modified method. These components include unique characteristics of the method, critical procedural details, and quality control measures. Adherence to essential test method components will help to assure that a similar or modified proposed method is based on the same concepts as the corresponding VRM (9). The essential test method components are described in detail in paragraphs 16 to 21 of the TM and testing should be performed according to the following:

- The general conditions (paragraph 16)
- The functional conditions, which include:
  - viability (paragraph 17);
  - barrier function (paragraph 18);
  - morphology (paragraph 19);
  - reproducibility (paragraph 20); and,
  - quality control (paragraph 21)

**II) Minimum List of Reference Chemicals**

7. Reference Chemicals are used to determine if the reliability and accuracy of a proposed similar or modified method, proven to be structurally and functionally sufficiently similar to the VRM, or representing a minor modification of one of the three validated methods, are comparable or better than those of the VRM (2) (8) (16) (23). The 20 recommended Reference Chemicals listed in Table 1 include chemicals representing different chemical classes (*i.e.* chemical categories based on functional groups), and are representative of the full range of Draize irritancy scores (from non-irritant to strong irritant). The chemicals included in this list comprise 10 UN GHS/EU CLP Category 2 chemicals and 10 non-categorised chemicals, of which 3 are optional UN GHS Category 3 chemicals. Under this Test Method, the optional Category 3 is considered as No Category. The chemicals listed in Table 1 are selected from the chemicals used in the optimisation phase that followed prevalidation and in the validation study of the VRM, with regard to chemical functionality and physical state (14) (18). These Reference Chemicals represent the minimum number of chemicals that should be used to evaluate the accuracy and reliability of a proposed similar or modified method, but

## ▼ M3

should not be used for the development of new methods. In situations where a listed chemical is unavailable, other chemicals for which adequate *in vivo* reference data are available could be used, primarily from the chemicals used in the optimisation phase following prevalidation or the validation study of the VRM. If desired, additional chemicals representing other chemical classes and for which adequate *in vivo* reference data are available may be added to the minimum list of Reference Chemicals to further evaluate the accuracy of the proposed method.

Table 1

**Minimum List of Reference Chemicals for Determination of Accuracy and Reliability Values for Similar or Modified RhE Skin Irritation Methods <sup>(1)</sup>**

Chemical	CAS Number	Physical state	<i>In vivo</i> score	VRM <i>in vitro</i> Cat.	UN GHS/EU CLP <i>in vivo</i> Cat.
1-bromo-4-chlorobutane	6940-78-9	Liquid	0	Cat. 2	No Cat.
diethyl phthalate	84-66-2	Liquid	0	No Cat.	No Cat.
naphthalene acetic acid	86-87-3	Solid	0	No Cat.	No Cat.
allyl phenoxy-acetate	7493-74-5	Liquid	0,3	No Cat.	No Cat.
isopropanol	67-63-0	Liquid	0,3	No Cat.	No Cat.
4-methyl-thio-benzaldehyde	3446-89-7	Liquid	1	Cat. 2	No Cat.
methyl stearate	112-61-8	Solid	1	No Cat.	No Cat.
heptyl butyrate	5870-93-9	Liquid	1,7	No Cat.	No Cat.
hexyl salicylate	6259-76-3	Liquid	2	No Cat.	No Cat.
Cinnamaldehyde	104-55-2	Liquid	2	Cat. 2	No Cat. (Optional Cat. 3) <sup>(2)</sup>
<i>1-decanol</i> <sup>(2)</sup>	<i>112-30-1</i>	<i>Liquid</i>	<i>2,3</i>	<i>Cat. 2</i>	<i>Cat. 2</i>
cyclamen aldehyde	103-95-7	Liquid	2,3	Cat. 2	Cat. 2
1-bromohexane	111-25-1	Liquid	2,7	Cat. 2	Cat. 2
2-chloromethyl-3,5-dimethyl-4-methoxypyridine HCl	86604-75-3	Solid	2,7	Cat. 2	Cat. 2
<i>di-n-propyl disulphide</i> <sup>(2)</sup>	<i>629-19-6</i>	<i>Liquid</i>	<i>3</i>	<i>No Cat.</i>	<i>Cat. 2</i>
potassium hydroxide (5 % aq.)	1310-58-3	Liquid	3	Cat. 2	Cat. 2
benzenethiol, 5-(1,1-dimethylethyl)-2-methyl	7340-90-1	Liquid	3,3	Cat. 2	Cat. 2
1-methyl-3-phenyl-1-piperazine	5271-27-2	Solid	3,3	Cat. 2	Cat. 2



## ▼ M3

Chemical	CAS Number	Physical state	<i>In vivo</i> score	VRM <i>in vitro</i> Cat.	UN GHS/EU CLP <i>in vivo</i> Cat.
Heptanal	111-71-7	Liquid	3,4	Cat. 2	Cat. 2
Tetrachloroethylene	127-18-4	Liquid	4	Cat. 2	Cat. 2

(<sup>1</sup>) The chemical selection is based on the following criteria: (i) the chemicals are commercially available; (ii) they are representative of the full range of Draize irritancy scores (from non-irritant to strong irritant); (iii) they have a well-defined chemical structure; (iv) they are representative of the chemical functionality used in the validation process; and (v) they are not associated with an extremely toxic profile (e.g. carcinogenic or toxic to the reproductive system) and they are not associated with prohibitive disposal costs.

(<sup>2</sup>) Chemicals that are irritant in the rabbit but for which there is reliable evidence that they are non-irritant in humans (31) (32) (33).

(<sup>3</sup>) Under the UN GHS, not in the EU CLP

### III) Defined Reliability and Accuracy Values

8. For purposes of establishing the reliability and relevance of proposed similar or modified methods to be transferred between laboratories, all 20 Reference Chemicals in Table 1 should be tested in at least three laboratories. However, if the proposed method is to be used in a single laboratory only, multi-laboratory testing will not be required for validation. It is however essential that such validation studies are independently assessed by internationally recognised validation bodies, in agreement with international guidelines (9). In each laboratory, all 20 Reference Chemicals should be tested in three independent runs performed with different tissue batches and at sufficiently spaced time points. Each run should consist of a minimum of three concurrently tested tissue replicates for each included test chemical, NC and PC.

9. The calculation of the reliability and accuracy values of the proposed method should be done considering all four criteria below together, ensuring that the values for reliability and relevance are calculated in a predefined and consistent manner:

1. Only the data of runs from complete run sequences qualify for the calculation of the method within, and between-laboratory variability and predictive capacity (accuracy).
2. The final classification for each Reference Chemicals in each participating laboratory should be obtained by using the mean value of viability over the different runs of a complete run sequence.
3. Only the data obtained for chemicals that have complete run sequences in all participating laboratories qualify for the calculation of the method between-laboratory variability.
4. The calculation of the accuracy values should be done on the basis of the individual laboratory predictions obtained for the 20 Reference Chemicals by the different participating laboratories.

In this context, a **run sequence** consists of three independent runs from one laboratory for one test chemical. A **complete run sequence** is a run sequence from one laboratory for one test chemical where all three runs are valid. This means that any single invalid run invalidates an entire run sequence of three runs.

#### *Within-laboratory reproducibility*

10. An assessment of within-laboratory reproducibility should show a concordance of classifications (UN GHS/EU CLP Category 2 and No Category) obtained in different, independent test runs of the 20 Reference Chemicals within one single laboratory equal or higher ( $\geq$ ) than 90 %.

▼ **M3***Between-laboratory reproducibility*

11. An assessment of between-laboratory reproducibility is not essential if the proposed method is to be used in a single laboratory only. For methods to be transferred between laboratories, the concordance of classifications (UN GHS/EU CLP Category 2 and No Category) obtained in different, independent test runs of the 20 Reference Chemicals between preferentially a minimum of three laboratories should be equal or higher ( $\geq$ ) than 80 %.

*Predictive capacity (accuracy)*

12. The accuracy (sensitivity, specificity and overall accuracy) of the proposed similar or modified method should be comparable or better to that of the VRM, taking into consideration additional information relating to relevance in the species of interest (Table 2). The sensitivity should be equal or higher ( $\geq$ ) than 80 % (2) (8) (23). However, a further specific restriction applies to the sensitivity of the proposed *in vitro* method inasmuch as only two *in vivo* Category 2 chemicals, *1-decanol* and *di-n-propyl disulphide*, may be misclassified as No Category by more than one participating laboratory. The specificity should be equal or higher ( $\geq$ ) than 70 % (2) (8) (23). There is no further restriction with regard to the specificity of the proposed *in vitro* method, *i.e.* any participating laboratory may misclassify any *in vivo* No Category chemical as long as the final specificity of the test method is within the acceptable range. The overall accuracy should be equal or higher ( $\geq$ ) than 75 % (2) (8) (23). Although the sensitivity of the VRM calculated for the 20 Reference Chemicals listed in Table 1 is equal to 90 %, the defined minimum sensitivity value required for any similar or modified method to be considered valid is set at 80 % since both *1-decanol* (a borderline chemical) and *di-n-propyl disulphide* (a false negative of the VRM) are known to be non-irritant in humans (31) (32) (33), although being identified as irritants in the rabbit test. Since RhE models are based on cells of human origin, they may predict these chemicals as non-irritant (UN GHS/EU CLP No Category).

Table 2

**Required predictive values for sensitivity, specificity and overall accuracy for any similar or modified method to be considered valid**

Sensitivity	Specificity	Overall Accuracy
$\geq 80 \%$	$\geq 70 \%$	$\geq 75 \%$

*Study Acceptance Criteria*

13. It is possible that one or several tests pertaining to one or more test chemicals does/do not meet the test acceptance criteria for the test and control chemicals or is/are not acceptable for other reasons. To complement missing data, for each test chemical a maximum number of two additional tests is admissible ('retesting'). More precisely, since in case of retesting also PC and NC have to be concurrently tested, a maximum number of two additional runs may be conducted for each test chemical.
14. It is conceivable that even after retesting, the minimum number of three valid runs required for each tested chemical is not obtained for every Reference Chemical in every participating laboratory, leading to an incomplete data matrix. In such cases the following three criteria should all be met in order to consider the datasets acceptable:

1. All 20 Reference Chemicals should have at least one complete run sequence.

**▼ M3**

2. In each of at least three participating laboratories, a minimum of 85 % of the run sequences need to be complete (for 20 chemicals; *i.e.* 3 invalid run sequences are allowed in a single laboratory).
3. A minimum of 90 % of all possible run sequences from at least three laboratories need to be complete (for 20 chemicals tested in 3 laboratories; *i.e.* 6 invalid run sequences are allowed in total).

**▼ M2****B. 47. BOVINE CORNEAL OPACITY AND PERMEABILITY TEST METHOD FOR IDENTIFYING OCULAR CORROSIVES AND SEVERE IRRITANTS**

## INTRODUCTION

1. The Bovine Corneal Opacity and Permeability (BCOP) test method is an *in vitro* test method that can be used, under certain circumstances and with specific limitations, to classify substances and mixtures as ‘ocular corrosives and severe irritants’ (1) (2) (3). For the purpose of this test method, severe irritants are defined as those that induce ocular lesions that persist in the rabbit for at least 21 days after administration. While it is not considered valid as a complete replacement for the *in vivo* rabbit eye test, the BCOP is recommended for use as part of a tiered-testing strategy for regulatory classification and labelling within a specific applicability domain (4) (5). Test substances and mixtures (6) can be classified as ocular corrosives or severe irritants without further testing in rabbits. A substance that tests negative would need to be tested in rabbits using a sequential testing strategy, as outlined in OECD Test Guideline 405 (7) (chapter B. 5 of this Annex).
  
2. The purpose of this test method is to describe the procedures used to evaluate the potential ocular corrosivity or severe irritancy of a test substance as measured by its ability to induce opacity and increased permeability in an isolated bovine cornea. Toxic effects to the cornea are measured by: (i) decreased light transmission (opacity), and (ii) increased passage of sodium fluorescein dye (permeability). The opacity and permeability assessments of the cornea following exposure to a test substance are combined to derive an *In Vitro* Irritancy Score (IVIS), which is used to classify the irritancy level of the test substance.
  
3. Ocular irritants that induce lesions that resolve in less than 21 days and non-irritants have also been tested using the BCOP test method. However, the accuracy and reliability of the BCOP test method for substances in these categories have not been formally evaluated.
  
4. Definitions are provided in Appendix 1.

## INITIAL CONSIDERATIONS AND LIMITATIONS

5. This test method is based on the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) BCOP test method protocol (8), which was developed following an international validation study (4)(5)(9), with contributions from the European Centre for the Validation of Alternative Methods (ECVAM) and the Japanese Center for the Validation of Alternative Methods (JaCVAM). The protocol is based on information obtained from the Institute for In Vitro Sciences (IIVS) and INVITTOX Protocol 124 (10), which represents the protocol used for the European Community-sponsored prevalidation study of the BCOP assay conducted in 1997-1998. Both of these protocols are based on the BCOP assay methodology first reported by Gautheron *et al.* (11).

**▼ M2**

6. The identified limitations for this test method are based on the high false positive rates for alcohols and ketones and the high false negative rate for solids observed in the validation database (see paragraph 44) (5). When substances within these chemical and physical classes are excluded from the database, the accuracy of BCOP across the EU, EPA, and GHS classification systems is substantially improved (5). Based on the purpose of this assay (*i.e.*, to identify ocular corrosives/severe irritants only), false negative rates are not critical since such substances would be subsequently tested in rabbits or with other adequately validated *in vitro* tests, depending on regulatory requirements, using a sequential testing strategy in a weight of evidence approach. Furthermore, the current validation database did not allow for an adequate evaluation of some chemical or product classes (*e.g.*, mixtures). However, investigators could consider using this test method for all types of test material (including mixtures), whereby a positive result could be accepted as indicative of an ocular corrosive or severe irritant response. However, positive results obtained with alcohols or ketones should be interpreted cautiously due to risk of over-prediction.
7. All procedures with bovine eyes and bovine corneas should follow the testing facility's applicable regulations and procedures for handling animal-derived materials, which include, but are not limited to, tissues and tissue fluids. Universal laboratory precautions are recommended (12).
8. A limitation of the test method is that, although it takes into account some of the ocular effects evaluated in the rabbit ocular irritancy test method and to some degree their severity, it does not consider conjunctival and iridal injuries. Also, although the reversibility of corneal lesions cannot be evaluated *per se* in the BCOP assay, it has been proposed, based on rabbit eye studies, that an assessment of the initial depth of corneal injury can be used to distinguish between irreversible and reversible effects (13). Finally, the BCOP does not allow for an assessment of the potential for systemic toxicity associated with ocular exposure.
9. Efforts are ongoing to further characterize the usefulness and limitations of the BCOP assay for identifying non-severe irritants and non-irritants (see also paragraph 45). Users are also encouraged to provide specimens and/or data to validation organizations for a formal evaluation of possible future uses of the BCOP test method, including for the identification of non-severe irritants and non-irritants.
10. For any laboratory initially establishing this assay, the proficiency chemicals provided in Appendix 2 should be used. A laboratory can use these chemicals to demonstrate their technical competence in performing the BCOP test method prior to submitting BCOP assay data for regulatory hazard classification purposes.

**PRINCIPLE OF THE TEST**

11. The BCOP test method is an organotypic model that provides short-term maintenance of normal physiological and biochemical function of the bovine cornea *in vitro*. In this test method, damage by the test substance is assessed by quantitative measurements of changes in corneal opacity and permeability with an opacitometer and a visible light spectrophotometer, respectively. Both measurements are used to calculate an IVIS, which is used to assign an *in vitro* irritancy hazard classification category for prediction of the *in vivo* ocular irritation potential of a test substance (see Decision criteria).

▼ **M2**

12. The BCOP test method uses isolated corneas from the eyes of freshly slaughtered cattle. Corneal opacity is measured quantitatively as the amount of light transmission through the cornea. Permeability is measured quantitatively as the amount of sodium fluorescein dye that passes across the full thickness of the cornea, as detected in the medium in the posterior chamber. Test substances are applied to the epithelial surface of the cornea by addition to the anterior chamber of the corneal holder. Appendix 3 provides a description and a diagram of a corneal holder used in the BCOP. Corneal holders can be obtained commercially from different sources or can be constructed.

#### Source and Age of Bovine Eyes and Selection of Animal Species

13. Cattle sent to slaughterhouses are typically killed either for human consumption or for other commercial uses. Only healthy animals considered suitable for entry into the human food chain are used as a source of corneas for use in the BCOP. Because cattle have a wide range of weights, depending on breed, age, and sex, there is no recommended weight for the animal at the time of slaughter.
  
14. Variations in corneal dimensions can result when using eyes from animals of different ages. Corneas with a horizontal diameter  $> 30,5$  mm and central corneal thickness (CCT) values  $\geq 1\ 100$   $\mu\text{m}$  are generally obtained from cattle older than eight years, while those with a horizontal diameter  $< 28,5$  mm and CCT  $< 900$   $\mu\text{m}$  are generally obtained from cattle less than five years old (14). For this reason, eyes from cattle greater than 60 months old are not typically used. Eyes from cattle less than 12 months of age have not traditionally been used since the eyes are still developing and the corneal thickness and corneal diameter are considerably smaller than that reported for eyes from adult cattle. However, the use of corneas from young animals (*i.e.*, 6 to 12 months old) is permissible since there are some advantages, such as increased availability, a narrow age range, and decreased hazards related to potential worker exposure to Bovine Spongiform Encephalopathy (15). As further evaluation of the effect of corneal size or thickness on responsiveness to corrosive and irritant substances would be useful, users are encouraged to report the estimated age and/or weight of the animals providing the corneas used in a study.

#### Collection and Transport of Eyes to the Laboratory

15. Eyes are collected by slaughterhouse employees. To minimize mechanical and other types of damage to the eyes, the eyes should be enucleated as soon as possible after death. To prevent exposure of the eyes to potentially irritant substances, the slaughterhouse employees should not use detergent when rinsing the head of the animal.
  
16. Eyes should be immersed completely in Hanks' Balanced Salt Solution (HBSS) in a suitably sized container, and transported to the laboratory in such a manner as to minimize deterioration and/or bacterial contamination. Because the eyes are collected during the slaughter process, they might be exposed to blood and other biological substances, including bacteria and other microorganisms. Therefore, it is important to ensure that the risk of contamination is minimized (*e.g.*, by keeping the container containing the eyes on wet ice, by adding antibiotics to the HBSS used to store the eyes during transport [*e.g.*, penicillin at 100 IU/mL and streptomycin at 100  $\mu\text{g/mL}$ ]).

**▼ M2**

17. The time interval between collection of the eyes and use of corneas in the BCOP should be minimized (typically collected and used on the same day) and should be demonstrated to not compromise the assay results. These results are based on the selection criteria for the eyes, as well as the positive and negative control responses. All eyes used in the assay should be from the same group of eyes collected on a specific day.

**Selection Criteria for Eyes Used in the BCOP**

18. The eyes, once they arrive at the laboratory, are carefully examined for defects including increased opacity, scratches, and neovascularisation. Only corneas from eyes free of such defects are to be used.
  
19. The quality of each cornea is also evaluated at later steps in the assay. Corneas that have an opacity greater than seven opacity units (NOTE: the opacitometer should be calibrated with opacity standards that are used to establish the opacity units, see Appendix 3) after an initial one hour equilibration period are to be discarded.
  
20. Each treatment group (test substance, concurrent negative and positive controls) consists of a minimum of three eyes. Three corneas should be used for the negative control corneas in the BCOP assay. Since all corneas are excised from the whole globe, and mounted in the corneal chambers, there is the potential for artefacts from handling upon individual corneal opacity and permeability values (including negative control). Furthermore, the opacity and permeability values from the negative control corneas are used to correct the test article and positive control-treated corneal opacity and permeability values in the IVIS calculations.

**PROCEDURE****Preparation of the Eyes**

21. Corneas free of defects are dissected with a 2 to 3 mm rim of sclera remaining to assist in subsequent handling, with care taken to avoid damage to the corneal epithelium and endothelium. Isolated corneas are mounted in specially designed corneal holders that consist of anterior and posterior compartments, which interface with the epithelial and endothelial sides of the cornea, respectively. Both chambers are filled to excess with pre-warmed Eagle's Minimum Essential Medium (EMEM) (posterior chamber first), ensuring that no bubbles are formed. The device is then equilibrated at  $32 \pm 1$  °C for at least one hour to allow the corneas to equilibrate with the medium and to achieve normal metabolic activity, to the extent possible (the approximate temperature of the corneal surface *in vivo* is 32 °C).
  
22. Following the equilibration period, fresh pre-warmed EMEM is added to both chambers and baseline opacity readings are taken for each cornea. Any corneas that show macroscopic tissue damage (*e.g.*, scratches, pigmentation, neovascularisation) or an opacity > 7 opacity units are discarded. The mean opacity of all equilibrated corneas is calculated. A minimum of three corneas with opacity values close to the median value for all corneas are selected as negative (or solvent) control corneas. The remaining corneas are then distributed into treatment and positive control groups.

**▼ M2**

23. Because the heat capacity of water is higher than that of air, water provides more stable temperature conditions for incubation. Therefore, the use of a water bath for maintaining the corneal holder and its contents at  $32 \pm 1$  °C is recommended. However, air incubators might also be used, assuming precautions to maintain temperature stability (*e.g.*, by pre-warming of holders and media).

**Application of the Test Substance**

24. Two different treatment protocols are used, one for liquids and surfactants (solids or liquids), and one for non-surfactant solids.
25. Liquids are tested undiluted, while surfactants are tested at a concentration of 10 % w/v in a 0,9 % sodium chloride solution, distilled water, or other solvent that has been demonstrated to have no adverse effects on the test system. Semi-solids, creams, and waxes are typically tested as liquids. Appropriate justification should be provided for alternative dilution concentrations. Corneas are exposed to liquids and surfactants for 10 minutes. Use of other exposure times should be accompanied by adequate scientific rationale.
26. Non-surfactant solids are typically tested as solutions or suspensions at 20 % concentration in a 0,9 % sodium chloride solution, distilled water, or other solvent that has been demonstrated to have no adverse effects on the test system. In certain circumstances and with proper scientific justification, solids may also be tested neat by direct application onto the corneal surface using the open chamber method (see paragraph 29). Corneas are exposed to solids for four hours, but as with liquids and surfactants, alternative exposure times may be used with appropriate scientific rationale.
27. Different treatment methods can be used, depending on the physical nature and chemical characteristics (*e.g.*, solids, liquids, viscous vs. non-viscous liquids) of the test substance. The critical factor is ensuring that the test substance adequately covers the epithelial surface and that it is adequately removed during the rinsing steps. A closed-chamber method is typically used for non-viscous to slightly viscous liquid test substances, while an open-chamber method is typically used for semi-viscous and viscous liquid test substances and for neat solids.
28. In the closed-chamber method, sufficient test substance (750 µL) to cover the epithelial side of the cornea is introduced into the anterior chamber through the dosing holes on the top surface of the chamber, and the holes are subsequently sealed with the chamber plugs during the exposure. It is important to ensure that each cornea is exposed to a test substance for the appropriate time interval.
29. In the open-chamber method, the window-locking ring and glass window from the anterior chamber are removed prior to treatment. The control or test substance (750 µL, or enough test substance to completely cover the cornea) is applied directly to the epithelial surface of the cornea using a micropipette. If a test substance is difficult to pipette, the test substance can be pressure-loaded into a positive displacement pipette to aid in dosing. The pipette tip of the positive displacement pipette is inserted into the dispensing tip of the syringe so that the material can be loaded into the displacement tip under pressure. Simultaneously, the syringe plunger is depressed as the pipette piston is drawn upwards. If air bubbles appear in the pipette tip, the test article is removed (expelled) and the process repeated until the tip is filled without air bubbles. If necessary, a normal syringe (without a needle) can be used since it permits measuring an accurate volume of test substance and an easier application to the epithelial surface of the cornea. After dosing, the glass window is replaced on the anterior chamber to recreate a closed system.



**▼ M2****Post-Exposure Incubation**

30. After the exposure period, the test substance, the negative control, or the positive control substance is removed from the anterior chamber and the epithelium washed at least three times (or until no visual evidence of test substance can be observed) with EMEM (containing phenol red). Phenol red-containing medium is used for rinsing since a colour change in the phenol red may be monitored to determine the effectiveness of rinsing acidic or alkaline materials. The corneas are washed more than three times if the phenol red is still discoloured (yellow or purple), or the test substance is still visible. Once the medium is free of test substance, the corneas are given a final rinse with EMEM (without phenol red). The EMEM (without phenol red) is used as a final rinse to ensure removal of the phenol red from the anterior chamber prior to the opacity measurement. The anterior chamber is then refilled with fresh EMEM without phenol red.
31. For liquids or surfactants, after rinsing, the corneas are incubated for an additional two hours at  $32 \pm 1$  °C. Longer post-exposure time may be useful in certain circumstances and could be considered on a case-by-case basis. Corneas treated with solids are rinsed thoroughly at the end of the four-hour exposure period, but do not require further incubation.
32. At the end of the post-exposure incubation period for liquids and surfactants and at the end of the four-hour exposure period for non-surfactant solids, the opacity and permeability of each cornea are recorded. Also, each cornea is observed visually and pertinent observations recorded (e.g., tissue peeling, residual test substance, non-uniform opacity patterns). These observations could be important as they may be reflected by variations in the opacimeter readings.

**Control Substances**

33. Concurrent negative or solvent/vehicle controls and positive controls are included in each experiment.
34. When testing a liquid substance at 100 %, a concurrent negative control (e.g., 0,9 % sodium chloride solution or distilled water) is included in the BCOP test method so that nonspecific changes in the test system can be detected and to provide a baseline for the assay endpoints. It also ensures that the assay conditions do not inappropriately result in an irritant response.
35. When testing a diluted liquid, surfactant, or solid, a concurrent solvent/vehicle control group is included in the BCOP test method so that nonspecific changes in the test system can be detected and to provide a baseline for the assay endpoints. Only a solvent/vehicle that has been demonstrated to have no adverse effects on the test system can be used.
36. A known ocular irritant is included as a concurrent positive control in each experiment to verify that an appropriate response is induced. As the BCOP assay is being used in this test method to identify corrosive or severe irritants, ideally the positive control should be a reference substance that induces a severe response in this test method. However, to ensure that variability in the positive control response across time can be assessed, the magnitude of irritant response should not be excessive.
37. Examples of positive controls for liquid test substances are dimethylformamide or 1 % sodium hydroxide. An example of a positive control for solid test substances is 20 % (weight to volume) imidazole in 0,9 % sodium chloride solution.

**▼ M2**

38. Benchmark substances are useful for evaluating the ocular irritancy potential of unknown chemicals of a specific chemical or product class, or for evaluating the relative irritancy potential of an ocular irritant within a specific range of irritant responses.

**Endpoints Measured**

39. Opacity is determined by the amount of light transmission through the cornea. Corneal opacity is measured quantitatively with the aid of an opacimeter, resulting in opacity values measured on a continuous scale.
40. Permeability is determined by the amount of sodium fluorescein dye that penetrates all corneal cell layers (*i.e.*, the epithelium on the outer cornea surface through the endothelium on the inner cornea surface). 1 mL sodium fluorescein solution (4 or 5 mg/mL when testing liquids and surfactants or non-surfactant solids, respectively) is added to the anterior chamber of the corneal holder, which interfaces with the epithelial side of the cornea, while the posterior chamber, which interfaces with the endothelial side of the cornea, is filled with fresh EMEM. The holder is then incubated in a horizontal position for  $90 \pm 5$  min at  $32 \pm 1$  °C. The amount of sodium fluorescein that crosses into the posterior chamber is quantitatively measured with the aid of UV/VIS spectrophotometry. Spectrophotometric measurements evaluated at 490 nm are recorded as optical density (OD<sub>490</sub>) or absorbance values, which are measured on a continuous scale. The fluorescein permeability values are determined using OD<sub>490</sub> values based upon a visible light spectrophotometer using a standard 1 cm path length.
41. Alternatively, a 96-well microtiter plate reader may be used provided that; (i) the linear range of the plate reader for determining fluorescein OD<sub>490</sub> values can be established; and (ii), the correct volume of fluorescein samples are used in the 96-well plate to result in OD<sub>490</sub> values equivalent to the standard 1 cm path length (this could require a completely full well [usually 360 µL]).

**DATA AND REPORTING****Data Evaluation**

42. Once the opacity and mean permeability (OD<sub>490</sub>) values have been corrected for background opacity and the negative control permeability OD<sub>490</sub> values, the mean opacity and permeability OD<sub>490</sub> values for each treatment group should be combined in an empirically-derived formula to calculate an *in vitro* irritancy score (IVIS) for each treatment group as follows:

$$\text{IVIS} = \text{mean opacity value} + (15 \times \text{mean permeability OD}_{490} \text{ value})$$

Sina *et al.* (16) reported that this formula was derived during in-house and inter-laboratory studies. The data generated for a series of 36 compounds in a multi-laboratory study were subjected to a multivariate analysis to determine the equation of best fit between *in vivo* and *in vitro* data. This analysis was performed by scientists at two separate companies, who derived nearly identical equations.

43. The opacity and permeability values should also be evaluated independently to determine whether a test substance induced corrosivity or severe irritation through only one of the two endpoints (see Decision Criteria).

**▼ M2****Decision Criteria**

44. A substance that induces an IVIS  $\geq 55,1$  is defined as a corrosive or severe irritant. As stated in paragraph 1, if the test substance is not identified as an ocular corrosive or severe irritant, additional testing should be conducted for classification and labelling purposes. The BCOP test method has an overall accuracy of 79 % (113/143) to 81 % (119/147), a false positive rate of 19 % (20/103) to 21 % (22/103), and a false negative rate of 16 % (7/43) to 25 % (10/40), when compared to *in vivo* rabbit eye test method data classified according to the EPA (1), EU (2), or GHS (3) classification systems. When substances within certain chemical (*i.e.*, alcohols, ketones) or physical (*i.e.*, solids) classes are excluded from the database, the accuracy of BCOP across the EU, EPA, and GHS classification systems ranges from 87 % (72/83) to 92 % (78/85), the false positive rates range from 12 % (7/58) to 16 % (9/56), and the false negative rates range from 0 % (0/27) to 12 % (3/26).
45. Even if an ocular corrosive or severe irritant classification is not obtained for a test substance, BCOP data can be useful, in conjunction with test data from the *in vivo* rabbit eye test or from an adequately validated *in vitro* test, to further evaluate the usefulness and limitations of the BCOP test method for identifying non-severe irritants and non-irritants (a Guidance Document on the use of *in vitro* ocular toxicity test methods is under development).

**Study Acceptance Criteria**

46. A test is considered acceptable if the positive control gives an IVIS that falls within two standard deviations of the current historical mean, which is to be updated at least every three months, or each time an acceptable test is conducted in laboratories where tests are conducted infrequently (*i.e.*, less than once a month). The negative or solvent/vehicle control responses should result in opacity and permeability values that are less than the established upper limits for background opacity and permeability values for bovine corneas treated with the respective negative or solvent/vehicle control.

**Test Report**

47. The test report should include the following information, if relevant to the conduct of the study:

*Test and Control Substances*

Chemical name(s) such as the structural name used by the Chemical Abstracts Service (CAS), followed by other names, if known;

The CAS Registry Number (RN), if known;

Purity and composition of the substance or mixture (in percentage(s) by weight), to the extent this information is available;

Physicochemical properties such as physical state, volatility, pH, stability, chemical class, water solubility relevant to the conduct of the study;

Treatment of the test/control substances prior to testing, if applicable (*e.g.*, warming, grinding);

Stability, if known.

**▼ M2***Information Concerning the Sponsor and the Test Facility*

Name and address of the sponsor, test facility and study director;

Identification of the source of the eyes (*i.e.*, the facility from which they were collected);

Storage and transport conditions of eyes (*e.g.*, date and time of eye collection, time interval prior to initiating testing, transport media and temperature conditions, any antibiotics used);

If available, specific characteristics of the animals from which the eyes were collected (*e.g.*, age, sex, weight of the donor animal).

*Justification of the Test Method and Protocol Used**Test Method Integrity*

The procedure used to ensure the integrity (*i.e.*, accuracy and reliability) of the test method over time (*e.g.*, periodic testing of proficiency substances, use of historical negative and positive control data).

*Criteria for an Acceptable Test*

Acceptable concurrent positive and negative control ranges based on historical data;

If applicable, acceptable concurrent benchmark control ranges based on historical data.

*Test Conditions*

Description of test system used;

Type of corneal holder used;

Calibration information for devices used for measuring opacity and permeability (*e.g.*, opacitometer and spectrophotometer);

Information on the bovine corneas used, including statements regarding their quality;

Details of test procedure used;

Test substance concentration(s) used;

Description of any modifications of the test procedure;

Reference to historical data of the model (*e.g.*, negative and positive controls, proficiency substances, benchmark substances);

Description of evaluation criteria used.

*Results*

Tabulation of data from individual test samples (*e.g.*, opacity and OD<sub>490</sub> values and calculated IVIS for the test substance and the positive, negative, and benchmark controls [if included], reported in tabular form, including data from replicate repeat experiments as appropriate, and means  $\pm$  the standard deviation for each experiment);

Description of other effects observed.

*Discussion of the Results**Conclusion*

## LITERATURE

- (1) U.S. EPA (1996). Label Review Manual: 2nd Edition. EPA737-B-96-001. Washington, DC: U.S. Environmental Protection Agency.

▼ M2

- (2) Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directive 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. OJ L 353, 31.12.2008, p. 1.
- (3) UN (2007). Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Second revised edition, New York & Geneva: United Nations Publications, 2007. Available:  
  
[[http://www.unece.org/trans/danger/publi/ghs/ghs\\_rev02/02files\\_e.html](http://www.unece.org/trans/danger/publi/ghs/ghs_rev02/02files_e.html)]
- (4) ESAC (2007). Statement on the conclusion of the ICCVAM retrospective study on organotypic *in vitro* assays as screening tests to identify potential ocular corrosives and severe eye irritants. Available:  
  
[<http://ecvam.jrc.it/index.htm>]
- (5) ICCVAM (2007). Test Method Evaluation Report - *In Vitro* Ocular Toxicity Test Methods for Identifying Ocular Severe Irritants and Corrosives. Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). NIH Publication No: 07-4517. Available:  
  
[[http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu\\_tmer.htm](http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu_tmer.htm)]
- (6) Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. OJ L 396, 30.12.2006, p. 1.
- (7) OECD (2002). Test Guideline 405. OECD Guideline for Testing of Chemicals. Acute eye irritation/corrosion. Available:  
  
[[http://www.oecd.org/document/40/0,2340,en\\_2649\\_34377\\_37051368\\_1\\_1\\_1\\_1,00.htm](http://www.oecd.org/document/40/0,2340,en_2649_34377_37051368_1_1_1_1,00.htm)]
- (8) ICCVAM (2007). ICCVAM Recommended BCOP Test Method Protocol. In: ICCVAM Test Method Evaluation Report - *In Vitro* Ocular Toxicity Test Methods for Identifying Ocular Severe Irritants and Corrosives. Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). NIH Publication No: 07-4517. Available:  
  
[[http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu\\_tmer.htm](http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu_tmer.htm)]
- (9) ICCVAM. (2006). Current Status of *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants: Bovine Corneal Opacity and Permeability Test Method. NIH Publication No: 06-4512. Research Triangle Park: National Toxicology Program. Available:  
  
[[http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu\\_brd\\_ice.htm](http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu_brd_ice.htm)]
- (10) INVITTOX (1999). Protocol 124: Bovine Corneal Opacity and Permeability Assay – SOP of Microbiological Associates Ltd. Ispra, Italy: European Centre for the Validation of Alternative Methods (ECVAM).

▼ M2

- (11) Gautheron, P., Dukic, M., Alix, D. and Sina, J.F. (1992). Bovine corneal opacity and permeability test: An *in vitro* assay of ocular irritancy. *Fundam. Appl. Toxicol.* 18:442-449.
- (12) Siegel, J.D., Rhinehart, E., Jackson, M., Chiarello, L., and the Healthcare Infection Control Practices Advisory Committee (2007). Guideline for Isolation Precautions: Preventing Transmission of Infectious Agents in Healthcare Settings. Available:  
  
[<http://www.cdc.gov/ncidod/dhqp/pdf>].
- (13) Maurer, J.K., Parker, R.D. and Jester, J.V. (2002). Extent of corneal injury as the mechanistic basis for ocular irritation: key findings and recommendations for the development of alternative assays. *Reg. Tox. Pharmacol.* 36:106-117.
- (14) Doughty, M.J., Petrou, S. and Macmillan, H. (1995). Anatomy and morphology of the cornea of bovine eyes from a slaughterhouse. *Can. J. Zool.* 73:2159-2165.
- (15) Collee, J. and Bradley, R. (1997). BSE: A decade on - Part I. *The Lancet* 349: 636-641.
- (16) Sina, J.F., Galer, D.M., Sussman, R.S., Gautheron, P.D., Sargent, E.V., Leong, B., Shah, P.V., Curren, R.D., and Miller, K. (1995). A collaborative evaluation of seven alternatives to the Draize eye irritation test using pharmaceutical intermediates. *Fundam Appl Toxicol* 26:20-31.
- (17) ICCVAM (2006). Background review document, Current Status of In Vitro Test Methods for Identifying Ocular Corrosives and Severe Irritants: Bovine Corneal Opacity and Permeability (BCOP) Test Method. Available:  
  
[[http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu\\_brd\\_bcop.htm](http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu_brd_bcop.htm)]
- (18) ICCVAM (2006). Background review document, Current Status of In Vitro Test Methods for Identifying Ocular Corrosives and Severe Irritants: Isolated Chicken Eye (ICE) Test Method. Available:  
  
[[http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu\\_brd\\_bcop.htm](http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu_brd_bcop.htm)]

**▼ M2***Appendix 1*

## DEFINITIONS

**Accuracy:** The closeness of agreement between test method results and accepted reference values. It is a measure of test method performance and one aspect of 'relevance'. The term is often used interchangeably with 'concordance', to mean the proportion of correct outcomes of a test method.

**Benchmark substance:** A substance used as a standard for comparison to a test substance. A benchmark substance should have the following properties; (i) a consistent and reliable source(s); (ii) structural and functional similarity to the class of substances being tested; (iii) known physical/chemical characteristics; (iv) supporting data on known effects, and (v) known potency in the range of the desired response.

**Cornea:** The transparent part of the front of the eyeball that covers the iris and pupil and admits light to the interior.

**Corneal opacity:** Measurement of the extent of opaqueness of the cornea following exposure to a test substance. Increased corneal opacity is indicative of damage to the cornea. Opacity can be evaluated subjectively as done in the Draize rabbit eye test, or objectively with an instrument such as an 'opacimeter'.

**Corneal permeability:** Quantitative measurement of damage to the corneal epithelium by a determination of the amount of sodium fluorescein dye that passes through all corneal cell layers.

**EPA Category 1:** Corrosive (irreversible destruction of ocular tissue) or corneal involvement or irritation persisting for more than 21 days (1).

**EU Category R41:** Production of tissue damage in the eye, or serious physical decay of vision, following application of a test substance to the anterior surface of the eye, which is not fully reversible within 21 days of application (2).

**False negative rate:** The proportion of all positive substances falsely identified by a test method as negative. It is one indicator of test method performance.

**False positive rate:** The proportion of all negative substances that are falsely identified by a test method as positive. It is one indicator of test method performance.

**GHS (Globally Harmonized System of Classification and Labelling of Chemicals):** A system proposing the classification of chemicals (substances and mixtures) according to standardized types and levels of physical, health and environmental hazards, and addressing corresponding communication elements, such as pictograms, signal words, hazard statements, precautionary statements and safety data sheets, so that to convey information on their adverse effects with a view to protect people (including employers, workers, transporters, consumers and emergency responders) and the environment (3).

**▼ M2**

**GHS Category 1:** Production of tissue damage in the eye, or serious physical decay of vision, following application of a test substance to the anterior surface of the eye, which is not fully reversible within 21 days of application (3).

**Hazard:** Inherent property of an agent or situation having the potential to cause adverse effects when an organism, system or (sub) population is exposed to that agent.

**In Vitro Irritancy Score (IVIS):** An empirically-derived formula used in the BCOP assay whereby the mean opacity and mean permeability values for each treatment group are combined into a single *in vitro* score for each treatment group. The *IVIS* = mean opacity value + (15 × mean permeability value).

**Negative control:** An untreated replicate containing all components of a test system. This sample is processed with test substance-treated samples and other control samples to determine whether the solvent interacts with the test system.

**Non-irritant:** Substances that are not classified as EPA Category I, II, or III; EU Category R41 or R36; or GHS Category 1, 2A, or 2B ocular irritants.

**Ocular corrosive:** (a) A substance that causes irreversible tissue damage to the eye; (b) Substances that are classified as GHS Category 1, EPA Category I, or EU Category R41 ocular irritants (1) (2) (3).

**Ocular irritant:** (a) A substance that produces a reversible change in the eye following application to the anterior surface of the eye; (b) Substances that are classified as EPA Category II or III, EU Category R36, or GHS Category 2A or 2B ocular irritants (1) (2) (3).

**Ocular severe irritant:** (a) A substance that causes tissue damage in the eye following application to the anterior surface of the eye that does not resolve within 21 days of application or causes serious physical decay of vision; (b) Substances that are classified as GHS Category 1, EPA Category I, or EU Category R41 ocular irritants (1) (2) (3).

**Opacitometer:** An instrument used to measure 'corneal opacity' by quantitatively evaluating light transmission through the cornea. The typical instrument has two compartments, each with its own light source and photocell. One compartment is used for the treated cornea, while the other is used to calibrate and zero the instrument. Light from a halogen lamp is sent through a control compartment (empty chamber without windows or liquid) to a photocell and compared to the light sent through the experimental compartment, which houses the chamber containing the cornea, to a photocell. The difference in light transmission from the photocells is compared and a numeric opacity value is presented on a digital display.

**Positive control:** A replicate containing all components of a test system and treated with a substance known to induce a positive response. To ensure that variability in the positive control response across time can be assessed, the magnitude of the severe response should not be excessive.



**▼ M2**

**Reliability:** Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility and intra-laboratory repeatability.

**Solvent/vehicle control:** An untreated sample containing all components of a test system, including the solvent or vehicle that is processed with the test substance-treated and other control samples to establish the baseline response for the samples treated with the test substance dissolved in the same solvent or vehicle. When tested with a concurrent negative control, this sample also demonstrates whether the solvent or vehicle interacts with the test system.

**Tiered testing:** A stepwise testing strategy where all existing information on a test substance is reviewed, in a specified order, using a weight of evidence process at each tier to determine if sufficient information is available for a hazard classification decision, prior to progression to the next tier. If the irritancy potential of a test substance can be assigned based on the existing information, no additional testing is required. If the irritancy potential of a test substance cannot be assigned based on the existing information, a step-wise sequential animal testing procedure is performed until an unequivocal classification can be made.

**Validated test method:** A test method for which validation studies have been completed to determine the relevance (including accuracy) and reliability for a specific purpose. It is important to note that a validated test method may not have sufficient performance in terms of accuracy and reliability to be found acceptable for the proposed purpose.

**Weight-of-evidence:** The process of considering the strengths and weaknesses of various pieces of information in reaching and supporting a conclusion concerning the hazard potential of a substance.

## ▼ M2

## Appendix 2

## Proficiency substances for the BCOP test method

Prior to routine use of a test method that adheres to this test method, laboratories may wish to demonstrate technical proficiency by correctly identifying the ocular corrosivity classification of the 10 substances recommended in Table 1. These substances were selected to represent the range of responses for local eye irritation/corrosion, which is based on results in the *in vivo* rabbit eye test (TG 405) (*i.e.*, Categories 1, 2A, 2B, or Not Classified and Labelled according to the UN GHS (3) (7). However, considering the validated usefulness of these assays (*i.e.*, to identify ocular corrosives/severe irritants only), there are only two test outcomes for classification purposes (corrosive/severe irritant or non-corrosive/non-severe irritant) to demonstrate proficiency. Other selection criteria were that substances are commercially available, there are high quality *in vivo* reference data available, and there are high quality data from the two *in vitro* methods for which Test Guidelines are being developed. For this reason, irritant substances were selected from the ICCVAM recommended list of 122 reference substances for the validation of *in vitro* ocular toxicity test methods (see Appendix H: ICCVAM Recommended Reference Substances) (5). Reference data are available in the ICCVAM Background Review Documents for BCOP and Isolated Chicken Eye (ICE) test method (17) (18).

Table 1

Recommended substances for demonstrating technical proficiency with BCOP

Substance	CASRN	Chemical Class (1)	Physical Form	<i>In Vivo</i> Classification (2)	<i>In Vitro</i> Classification (3)
Benzalkonium chloride (5 %)	8001-54-5	Onium compound	Liquid	Category 1	Corrosive/Severe Irritant
Chlorhexidine	55-56-1	Amine, Amidine	Solid	Category 1	Corrosive/Severe Irritant
Dibenzoyl-L-tartaric acid	2743-38-6	Carboxylic acid, Ester	Solid	Category 1	Corrosive/Severe Irritant
Imidazole	288-32-4	Heterocyclic	Solid	Category 1	Corrosive/Severe Irritant
Trichloroacetic acid (30 %)	76-03-9	Carboxylic Acid	Liquid	Category 1	Corrosive/Severe Irritant
2,6-Dichlorobenzoyl chloride	4659-45-4	Acyl halide	Liquid	Category 2A	Non corrosive/Non severe irritant
Ethyl-2-methyl-aceto-acetate	609-14-3	Ketone, Ester	Liquid	Category 2B	Non corrosive/Non severe irritant
Ammonium nitrate	6484-52-2	Inorganic salt	Solid	Category 2A	Non corrosive/Non severe irritant

▼ **M2**

Substance	CASRN	Chemical Class <sup>(1)</sup>	Physical Form	<i>In Vivo</i> Classification <sup>(2)</sup>	<i>In Vitro</i> Classification <sup>(3)</sup>
Glycerol	56-81-5	Alcohol	Liquid	Not Labelled	Non corrosive/Non severe irritant
n-Hexane	110-54-3	Hydrocarbon (acyclic)	Liquid	Not Labeled	Noncorrosive/Non severe irritant

Abbreviations: CASRN = Chemical Abstracts Service Registry Number

<sup>(1)</sup> Chemical classes were assigned to each test substance using a standard classification scheme, based on the National Library of Medicine Medical Subject Headings (MeSH) classification system (available at <http://www.nlm.nih.gov/mesh>).

<sup>(2)</sup> Based on results from the *in vivo* rabbit eye test (OECD TG 405) and using the UN GHS (3)(7).

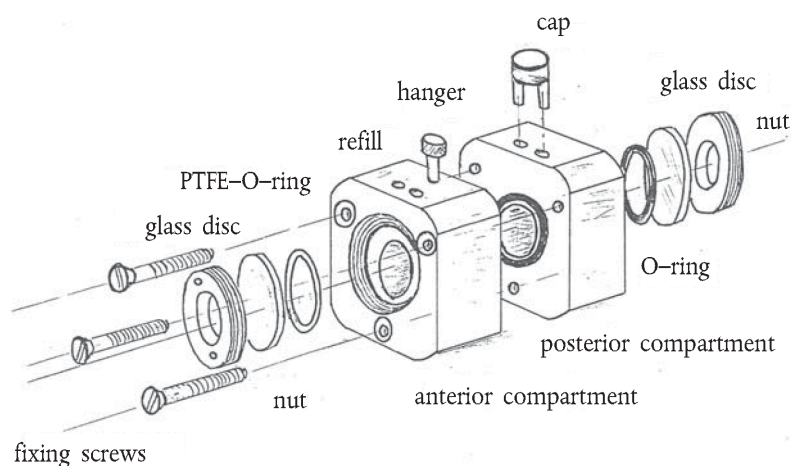
<sup>(3)</sup> Based on results in BCOP and ICE.

## ▼ M2

## Appendix 3

## THE BCOP CORNEAL HOLDER

1. The BCOP corneal holders are made of an inert material (*e.g.*, polypropylene). The holders are comprised of two halves (an anterior and posterior chamber), and have two similar cylindrical internal chambers. Each chamber holds a volume of 5 mL and terminates in a glass window, through which opacity measurements are recorded. Each of the inner chambers is 1,7 cm in diameter and 2,2 cm in depth <sup>(1)</sup>. An o-ring located on the posterior chamber is used to prevent leaks. The corneas are placed endothelial side down on the o-ring of the posterior chambers and the anterior chambers are placed on the epithelial side of the corneas. The chambers are maintained in place by three stainless screws located on the outer edges of the chamber. The end of each chamber houses a glass window which can be removed for easy access to the cornea. An o-ring is also located between the glass window and the chamber to prevent leaks. Two holes on the top of each chamber permit introduction and removal of medium and test compounds. They are closed with rubber caps during the treatment and incubation periods.

*Glossary**Glass disc:**PTFE-O-ring:**Refill:**Hanger:**Cap:**Nut:**O-ring:**Posterior compartment:**Anterior compartment:**Fixing screws:*

<sup>(1)</sup> The dimensions provided are based on a corneal holder that is used for cows ranging in age from 12 to 60 months old. In the event that animals 6 to 12 months are being used, the holder would instead need to be designed such that each chamber holds a volume of 4 mL, and each of the inner chambers is 1,5 cm in diameter and 2,2 cm in depth. With any newly designed corneal holder, it is very important that the ratio of exposed corneal surface area to posterior chamber volume should be the same as the ratio in the traditional corneal holder. This is necessary to assure that permeability values are correctly determined for the calculation of the IVIS by the proposed formula.

**▼ M2**

## THE OPACITOMETER

2. The opacitometer is a light transmission measuring device. Light from a halogen lamp is sent through a control compartment (empty chamber without windows or liquid) to a photocell and compared to the light sent through the experimental compartment, which houses the chamber containing the cornea, to a photocell. The difference in light transmission from the photocells is compared and a numeric opacity value is presented on a digital display. The opacity units are established.
3. The opacitometer should provide a linear response through a range of opacity readings covering the cut-offs used for the different classifications described by the Prediction Model (*i.e.*, up to the cut-off determining corrosiveness/severe irritancy). To ensure linear and accurate readings up to 75-80 opacity units, it is necessary to calibrate the opacitometer using a series of calibrators. Calibrators (opaque sheets of polyester) are placed into the calibration chamber (a corneal chamber designed to hold the calibrators) and read on the opacitometer. The calibration chamber is designed to hold the calibrators at approximately the same distance between the light and photocell that the corneas would be placed during the opacity measurements. The opacitometer is first calibrated to 0 opacity units using the calibration chamber without a calibrator. Three different calibrators are then placed into the calibration chamber one by one and the opacities are measured. Calibrators 1, 2 and 3 should result in opacity readings equal to their set values of 75, 150, and 225 opacity units, respectively,  $\pm 5\%$ .

**▼ M2****B. 48. ISOLATED CHICKEN EYE TEST METHOD FOR IDENTIFYING OCULAR CORROSIVES AND SEVERE IRRITANTS**

## INTRODUCTION

1. The Isolated Chicken Eye (ICE) test method is an *in vitro* test method that can be used, under certain circumstances and with specific limitations, to classify substances and mixtures as ocular corrosives and severe irritants, (1) (2) (3). For the purpose of this test method, severe irritants are defined as those that induce ocular lesions that persist in the rabbit for at least 21 days after administration. While it is not considered valid as a complete replacement for the *in vivo* rabbit eye test, the ICE is recommended for use as part of a tiered testing strategy for regulatory classification and labelling within a specific applicability domain (4) (5). Test substances and mixtures (6) that are positive in this assay can be classified as ocular corrosives or severe irritants without further testing in rabbits. A substance that tests negative would need to be tested in rabbits using a sequential testing strategy, as outlined in OECD Test Guideline 405 (7) (chapter B. 5 of this Annex).
  
2. The purpose of this test method is to describe the procedures used to evaluate the potential ocular corrosivity or severe irritancy of a test substance as measured by its ability to induce toxicity in an enucleated chicken eye. Toxic effects to the cornea are measured by (i) a qualitative assessment of opacity, (ii) a qualitative assessment of damage to epithelium based on application of fluorescein to the eye (fluorescein retention), (iii) a quantitative measurement of increased thickness (swelling), and (iv) a qualitative evaluation of macroscopic morphological damage to the surface. The corneal opacity, swelling, and damage assessments following exposure to a test substance are assessed individually and then combined to derive an Eye Irritancy Classification.
  
3. Ocular irritants that induce lesions that resolve in less than 21 days and non-irritants have also been tested using the ICE test method. However, the accuracy and reliability of the ICE test method for substances in these categories have not been formally evaluated.
  
4. Definitions are provided in Appendix 1.

## INITIAL CONSIDERATIONS AND LIMITATIONS

5. This test method is based on the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) ICE test method protocol (8), which was developed following an international validation study (4) (5) (9), with contributions from the European Centre for the Validation of Alternative Methods, the Japanese Center for the Validation of Alternative Methods, and TNO Quality of Life Department of Toxicology and Applied Pharmacology (Netherlands). The protocol is based on information obtained from published protocols, as well as the current protocol used by TNO (10) (11) (12) (13) (14).

**▼ M2**

6. The identified limitations for this method are based upon the false positive rate for alcohols and the false negative rates for solids and surfactants (see paragraph 47) (4). When substances within these chemical and physical classes are excluded from the database, the accuracy of ICE across the EU, EPA, and GHS classification systems is substantially improved (4). Based on the purpose of this assay (*i.e.*, to identify ocular corrosives/severe irritants only), false negative rates are not critical since such substances would be subsequently tested in rabbits or with other adequately validated *in vitro* tests, depending on regulatory requirements, using a sequential testing strategy in a weight of evidence approach. Furthermore, the current validation database did not allow for an adequate evaluation of some chemical or product classes (*e.g.*, mixtures). However, investigators could consider using this test method for testing all types of material (including mixtures), whereby a positive result could be accepted as indicative of an ocular corrosive or severe irritant response. However, positive results obtained with alcohols should be interpreted cautiously due to risk of over-prediction.
7. All procedures with chicken eyes should follow the test facility's applicable regulations and procedures for handling of human or animal-derived materials, which include, but are not limited to, tissues and tissue fluids. Universal laboratory precautions are recommended (15).
8. A limitation of the test method is that, although it takes into account some of the ocular effects evaluated in the rabbit ocular irritancy test method and to some degree their severity, it does not consider conjunctival and iridal injuries. Also, although the reversibility of corneal lesions cannot be evaluated *per se* in the ICE test method, it has been proposed, based on rabbit eye studies, that an assessment of the initial depth of corneal injury can be used to distinguish between irreversible and reversible effects (16). Finally, the ICE test method does not allow for an assessment of the potential for systemic toxicity associated with ocular exposure.
9. Efforts are ongoing to further characterize the usefulness and limitations of the ICE test method for identifying non-severe irritants and non-irritants (see also paragraph 48). Users are also encouraged to provide specimens and/or data to validation organizations for a formal evaluation of possible future uses of the ICE test method, including for the identification of non-severe ocular irritants and non-irritants.
10. For any laboratory initially establishing this assay, the proficiency chemicals provided in Appendix 2 should be used. A laboratory can use these chemicals to demonstrate their technical competence in performing the ICE test method prior to submitting ICE data for regulatory hazard classification purposes.

**PRINCIPLE OF THE TEST**

11. The ICE test method is an organotypic model that provides short-term maintenance of the chicken eye *in vitro*. In this test method, damage by the test substance is assessed by determination of corneal swelling, opacity, and fluorescein retention. While the latter two parameters involve a qualitative assessment, analysis of corneal swelling provides for a quantitative assessment. Each measurement is either converted into a quantitative score used to calculate an overall Irritation Index, or assigned a qualitative categorization that is used to assign an *in vitro* ocular corrosivity and severe irritancy classification. Either of these outcomes can then be used to predict the *in vivo* ocular corrosivity and severe irritation potential of a test substance (see Decision Criteria).

**▼ M2****Source and Age of Chicken Eyes**

12. Historically, eyes collected from chickens obtained from a slaughterhouse where they are killed for human consumption have been used for this assay, eliminating the need for laboratory animals. Only the eyes of healthy animals considered suitable for entry into the human food chain are used.
13. Although a controlled study to evaluate the optimum chicken age has not been conducted, the age and weight of the chickens used historically in this test method are that of spring chickens traditionally processed by a poultry slaughterhouse (*i.e.*, approximately 7 weeks old, 1,5-2,5 kg).

**Collection and Transport of Eyes to the Laboratory**

14. Heads should be removed immediately after sedation of the chickens, usually by electric shock, and incision of the neck for bleeding. A local source of chickens close to the laboratory should be located so that their heads can be transferred from the slaughterhouse to the laboratory quickly enough to minimize deterioration and/or bacterial contamination. The time interval between collection of the chicken heads and use of eyes in the ICE test method should be minimized (typically within two hours) and should be demonstrated to not compromise the assay results. These results are based on the selection criteria for the eyes, as well as the positive and negative control responses. All eyes used in the assay should be from the same group of eyes collected on a specific day.
15. Because eyes are dissected in the laboratory, the intact heads are transported from the slaughterhouse at ambient temperature in plastic boxes humidified with towels moistened with isotonic saline.

**Selection Criteria for Eyes Used in the ICE**

16. Eyes that have high baseline fluorescein staining (*i.e.*, > 0,5) or corneal opacity score (*i.e.*, > 0,5) after they are enucleated are rejected.
17. Each treatment group and concurrent positive control consists of at least three eyes. The negative control group or the solvent control (if using a solvent other than saline) consists of at least one eye.

**PROCEDURE****Preparation of the Eyes**

18. The eyelids are carefully excised, taking care not to damage the cornea. Corneal integrity is quickly assessed with a drop of 2 % (w/v) sodium fluorescein applied to the corneal surface for a few seconds, and then rinsed with isotonic saline. Fluorescein-treated eyes are then examined with a slit-lamp microscope to ensure that the cornea is undamaged (*i.e.*, fluorescein retention and corneal opacity scores  $\leq 0,5$ ).
19. If undamaged, the eye is further dissected from the skull, taking care not to damage the cornea. The eyeball is pulled from the orbit by holding the nictitating membrane firmly with surgical forceps, and the eye muscles are cut with a bent, blunt-tipped scissor. It is important to avoid causing corneal damage due to excessive pressure (*i.e.*, compression artifacts).
20. When the eye is removed from the orbit, a visible portion of the optic nerve should be left attached. Once removed from the orbit, the eye is placed on an absorbent pad and the nictitating membrane and other connective tissue are cut away.



**▼ M2**

21. The enucleated eye is mounted in a stainless steel clamp with the cornea positioned vertically. The clamp is then transferred to a chamber of the superfusion apparatus (16). The clamps should be positioned in the superfusion apparatus such that the entire cornea is supplied with the isotonic saline drip. The chambers of the superfusion apparatus should be temperature controlled at  $32 \pm 1,5$  °C. Appendix 3 provides a diagram of a typical superfusion apparatus and the eye clamps, which can be obtained commercially or constructed. The apparatus can be modified to meet the needs of an individual laboratory (*e.g.*, to accommodate a different number of eyes).
22. After being placed in the superfusion apparatus, the eyes are again examined with a slit-lamp microscope to ensure that they have not been damaged during the dissection procedure. Corneal thickness should also be measured at this time at the corneal apex using the depth measuring device on the slit-lamp microscope. Eyes with; (i), a fluorescein retention score of  $> 0,5$ ; (ii) corneal opacity  $> 0,5$ ; or, (iii), any additional signs of damage should be replaced. For eyes that are not rejected based on any of these criteria, individual eyes with a corneal thickness deviating more than 10 % from the mean value for all eyes are to be rejected. Users should be aware that slit-lamp microscopes could yield different corneal thickness measurements if the slit-width setting is different. The slit-width should be set at 0,095 mm.
23. Once all eyes have been examined and approved, the eyes are incubated for approximately 45 to 60 minutes to equilibrate them to the test system prior to dosing. Following the equilibration period, a zero reference measurement is recorded for corneal thickness and opacity to serve as a baseline (*i.e.*, time = 0). The fluorescein score determined at dissection is used as the baseline measurement for that endpoint.

**Application of the Test Substance**

24. Immediately following the zero reference measurements, the eye (in its holder) is removed from the superfusion apparatus, placed in a horizontal position, and the test substance is applied to the cornea.
25. Liquid test substances are typically tested undiluted, but may be diluted if deemed necessary (*e.g.*, as part of the study design). The preferred solvent for diluted substances is physiological saline. However, alternative solvents may also be used under controlled conditions, but the appropriateness of solvents other than physiological saline should be demonstrated.
26. Liquid test substances are applied to the cornea such that the entire surface of the cornea is evenly covered with the test substance; the standard volume is 0,03 mL.
27. If possible, solid substances should be ground as finely as possible in a mortar and pestle, or comparable grinding tool. The powder is applied to the cornea such that the surface is uniformly covered with the test substance; the standard amount is 0,03 g.
28. The test substance (liquid or solid) is applied for 10 seconds and then rinsed from the eye with isotonic saline (approximately 20 mL) at ambient temperature. The eye (in its holder) is subsequently returned to the superfusion apparatus in the original upright position.

**Control Substances**

29. Concurrent negative or solvent/vehicle controls and positive controls should be included in each experiment.

**▼ M2**

30. When testing liquids at 100 % or solids, physiological saline is used as the concurrent negative control in the ICE test method to detect non-specific changes in the test system, and to ensure that the assay conditions do not inappropriately result in an irritant response.
31. When testing diluted liquids, a concurrent solvent/vehicle control group is included in the test method to detect non-specific changes in the test system, and to ensure that the assay conditions do not inappropriately result in an irritant response. As stated in paragraph 25, only a solvent/vehicle that has been demonstrated to have no adverse effects on the test system can be used.
32. A known ocular irritant is included as a concurrent positive control in each experiment to verify that an appropriate response is induced. As the ICE assay is being used in this test method to identify corrosive or severe irritants, the positive control should be a reference substance that induces a severe response in this test method. However, to ensure that variability in the positive control response across time can be assessed, the magnitude of the severe response should not be excessive. Sufficient *in vitro* data for the positive control should be generated such that a statistically defined acceptable range for the positive control can be calculated. If adequate historical ICE test method data are not available for a particular positive control, studies may need to be conducted to provide this information.
33. Examples of positive controls for liquid test substances are 10 % acetic acid or 5 % benzalkonium chloride, while examples of positive controls for solid test substances are sodium hydroxide or imidazole.
34. Benchmark substances are useful for evaluating the ocular irritancy potential of unknown chemicals of a specific chemical or product class, or for evaluating the relative irritancy potential of an ocular irritant within a specific range of irritant responses.

**Endpoints Measured**

35. Treated corneas are evaluated pretreatment and starting at 30, 75, 120, 180, and 240 minutes ( $\pm$  5 minutes) after the post-treatment rinse. These time points provide an adequate number of measurements over the four-hour treatment period, while leaving sufficient time between measurements for the requisite observations to be made for all eyes.
36. The endpoints evaluated are corneal opacity, swelling, fluorescein retention, and morphological effects (*e.g.*, pitting or loosening of the epithelium). All of the endpoints, with the exception of fluorescein retention (which is determined only at pretreatment and 30 minutes after test substance exposure) are determined at each of the above time points.
37. Photographs are advisable to document corneal opacity, fluorescein retention, morphological effects and, if conducted, histopathology
38. After the final examination at four hours, users are encouraged to preserve eyes in an appropriate fixative (*e.g.*, neutral buffered formalin) for possible histopathological examination.

▼ **M2**

39. Corneal swelling is determined from corneal thickness measurements made with an optical pachymeter on a slit-lamp microscope. It is expressed as a percentage and is calculated from corneal thickness measurements according to the following formula:

$$\left[ \frac{\text{corneal thickness at time } t - \text{corneal thickness at time } = 0}{\text{corneal thickness at time } = 0} \right] \times 100$$

40. The mean percentage of corneal swelling for all test eyes is calculated for all observation time points. Based on the highest mean score for corneal swelling, as observed at any time point, an overall category score is then given for each test substance.
41. Corneal opacity is calculated by using the area of the cornea that is most densely opacified for scoring. The mean corneal opacity value for all test eyes is calculated for all observation time points. Based on the highest mean score for corneal opacity, as observed at any time point, an overall category score is then given for each test substance (Table 1).

Table 1

Corneal opacity scores

Score	Observation
0	No opacity
0,5	Very faint opacity
1	Scattered or diffuse areas; details of the iris are clearly visible
2	Easily discernible translucent area; details of the iris are slightly obscured
3	Severe corneal opacity; no specific details of the iris are visible; size of the pupil is barely discernible
4	Complete corneal opacity; iris invisible

42. The mean fluorescein retention value for all test eyes is calculated for the 30-minute observation time point only, which is used for the overall category score given for each test substance (Table 2).

Table 2

Fluorescein retention scores

Score	Observation
0	No fluorescein retention
0,5	Very minor single cell staining
1	Single cell staining scattered throughout the treated area of the cornea
2	Focal or confluent dense single cell staining
3	Confluent large areas of the cornea retaining fluorescein

▼ **M2**

43. Morphological effects include ‘pitting’ of corneal epithelial cells, ‘loosening’ of epithelium, ‘roughening’ of the corneal surface and ‘sticking’ of the test substance to the cornea. These findings can vary in severity and may occur simultaneously. The classification of these findings is subjective according to the interpretation of the investigator.

**DATA AND REPORTING****Data Evaluation**

44. Results from corneal opacity, swelling, and fluorescein retention should be evaluated separately to generate an ICE class for each endpoint. The ICE classes for each endpoint are then combined to generate an Irritancy Classification for each test substance.

**Decision Criteria**

45. Once each endpoint has been evaluated, ICE classes can be assigned based on a predetermined range. Interpretation of corneal thickness (Table 3), opacity (Table 4), and fluorescein retention (Table 5) using four ICE classes is done according to the following scales:

*Table 3*

ICE classification criteria for corneal thickness

Mean Corneal Swelling (%) (*)	ICE Class
0 to 5	I
> 5 to 12	II
> 12 to 18 (> 75 min after treatment)	II
> 12 to 18 (≤ 75 min after treatment)	III
> 18 to 26	III
> 26 to 32 (> 75 min after treatment)	III
> 26 to 32 (≤ 75 min after treatment)	IV
> 32	IV

(\*) Corneal swelling scores only applicable if thickness is measured with a Haag-Streit BP900 slit-lamp microscope with depth-measuring device No I and slit-width setting at 9½, equaling 0,095 mm. Users should be aware that slit-lamp microscopes could yield different corneal thickness measurements if the slit-width setting is different.

*Table 4*

ICE classification criteria for opacity

Mean Maximum Opacity Score (*)	ICE Class
0,0-0,5	I
0,6-1,5	II
1,6-2,5	III
2,6-4,0	IV

(\*) See Table 1.

## ▼ M2

Table 5

ICE classification criteria for mean fluorescein retention

Mean Fluorescein Retention Score at 30 minutes post-treatment (*)	ICE Class
0,0-0,5	I
0,6-1,5	II
1,6-2,5	III
2,6-3,0	IV

(\*) See Table 2.

46. The overall *in vitro* irritancy classification for a test substance is assessed by reading the irritancy classification that corresponds to the combination of categories obtained for corneal swelling, corneal opacity, and fluorescein retention and applying the scheme presented in Table 6.

Table 6

Overall *in vitro* irritancy classifications

Classification	Combinations of the 3 Endpoints
Corrosive/Severe Irritant	3 × IV 2 × IV, 1 × III 2 × IV, 1 × II (*) 2 × IV, 1 × I (*) Corneal opacity ≥ 3 at 30 min (in at least 2 eyes) Corneal opacity = 4 at any time point (in at least 2 eyes) Severe loosening of the epithelium (in at least 1 eye)

(\*) Combinations less likely to occur.

47. As stated in paragraph 1, if the test substance is not identified as an ocular corrosive or severe irritant, additional testing should be conducted for classification and labelling purposes. The ICE test method has an overall accuracy of 83 % (120/144) to 87 % (134/154), a false positive rate of 6 % (7/122) to 8 % (9/116), and a false negative rate of 41 % (13/32) to 50 % (15/30) for the identification of ocular corrosives and severe irritants, when compared to *in vivo* rabbit eye test method data classified according to the EPA (1), EU (2), or GHS (3) classification systems. When substances within certain chemical (*i.e.*, alcohols and surfactants) and physical (*i.e.*, solids) classes are excluded from the database, the accuracy of ICE across the EU, EPA, and GHS classification systems ranges from 91 % (75/82) to 92 % (69/75), the false positive rates range from 5 % (4/73) to 6 % (4/70), and the false negative rates range from 29 % (2/7) to 33 % (3/9) (4).

48. Even if an ocular corrosive or severe irritant classification is not obtained for a test substance, ICE data can be useful in conjunction with test data from the *in vivo* rabbit eye test or from an adequately validated *in vitro* test to further evaluate the usefulness and limitations of the ICE test method for identifying non-severe irritants and non-irritants (a Guidance Document on the use of *in vitro* ocular toxicity test methods is under development).

**▼ M2****Study Acceptance Criteria**

49. A test is considered acceptable if the concurrent negative or vehicle/solvent controls and the concurrent positive controls give an Irritancy Classification that falls within nonirritant and severe irritant/corrosive classes, respectively.

**Test Report**

50. The test report should include the following information, if relevant to the conduct of the study:

*Test and Control Substances*

Chemical name(s) such as the structural name used by the Chemical Abstracts Service (CAS), followed by other names, if known;

The CAS Registry Number (RN), if known;

Purity and composition of the substance or mixture (in percentage(s) by weight), to the extent this information is available;

Physicochemical properties such as physical state, volatility, pH, stability, chemical class water solubility relevant to the conduct of the study;

Treatment of the test/control substances prior to testing, if applicable (*e.g.*, warming, grinding);

Stability, if known;

*Information Concerning the Sponsor and the Test Facility*

Name and address of the sponsor, test facility and study director;

Identification on the source of the eyes (*e.g.*, the facility from which they were collected);

Storage and transport conditions of eyes (*e.g.*, date and time of eye collection, time interval prior to initiating testing);

If available, specific characteristics of the animals from which the eyes were collected (*e.g.*, age, sex, weight of the donor animal);

*Justification of the Test Method and Protocol Used**Test Method Integrity*

The procedure used to ensure the integrity (*i.e.*, accuracy and reliability) of the test method over time (*e.g.*, periodic testing of proficiency substances, use of historical negative and positive control data).

*Criteria for an Acceptable Test*

If applicable, acceptable concurrent benchmark control ranges based on historical data;

*Test Conditions*

Description of test system used;

Slit-lamp microscope used (*e.g.*, model);

Instrument settings for the slit-lamp microscope used;

Information for the chicken eyes used, including statements regarding their quality;

▼ M2

Details of test procedure used;

Test substance concentration(s) used;

Description of any modifications of the test procedure;

Reference to historical data of the model (*e.g.*, negative and positive controls, proficiency substances, benchmark substances);

Description of evaluation criteria used;

#### *Results*

Description of other effects observed;

If appropriate, photographs of the eye;

#### *Discussion of the Results*

#### *Conclusion*

#### LITERATURE

(1) U.S. EPA (1996). Label Review Manual: 2nd Edition. EPA737-B-96-001. Washington, DC: U.S. Environmental Protection Agency.

(2) Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directive 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. OJ L 353, 31.12.2008, p. 1.

(3) United nations (UN) (2007). Globally Harmonized System of Classification and Labelling of Chemicals (GHS), Second revised edition, UN New York and Geneva, 2007. Available at:

[[http://www.unece.org/trans/danger/publi/ghs/ghs\\_rev02/02files\\_e.html](http://www.unece.org/trans/danger/publi/ghs/ghs_rev02/02files_e.html)]

(4) ICCVAM (2007). Test Method Evaluation Report - *In Vitro Ocular Toxicity* Test Methods for Identifying Ocular Severe Irritants and Corrosives. Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). NIH Publication No: 07-4517. Available:

[[http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu\\_tmer.htm](http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu_tmer.htm)]

(5) ESAC (2007). Statement on the conclusion of the ICCVAM retrospective study on organotypic *in vitro* assays as screening tests to identify potential ocular corrosives and severe eye irritants. Available:

[<http://ecvam.jrc.it/index.htm>].

(6) Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. OJ L 396, 30.12.2006, p. 1.

(7) OECD (2002). Test Guideline 405. OECD Guideline for Testing of Chemicals. Acute eye irritation/corrosion. Available:

[[http://www.oecd.org/document/40/0,2340,en\\_2649\\_34377\\_37051368\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/40/0,2340,en_2649_34377_37051368_1_1_1_1,00.html)]

▼ M2

- (8) ICCVAM (2007). ICCVAM Recommended ICE Test Method Protocol. In: ICCVAM Test Method Evaluation Report - *In Vitro* Ocular Toxicity Test Methods for Identifying Ocular Severe Irritants and Corrosives. Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). NIH Publication No: 07-4517. Available:

[[http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu\\_tmter.htm](http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu_tmter.htm)]

- (9) ICCVAM. (2006). Current Status of *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants: Isolated Chicken Eye Test Method. NIH Publication No: 06-4513. Research Triangle Park: National Toxicology Program. Available:

[[http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu\\_brd\\_ice.htm](http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu_brd_ice.htm)]

- (10) Prinsen, M.K. and Koëter, B.W.M. (1993). Justification of the enucleated eye test with eyes of slaughterhouse animals as an alternative to the Draize eye irritation test with rabbits. *Fd. Chem. Toxicol.* 31:69-76.

- (11) INVITTOX (1994). Protocol 80: Chicken enucleated eye test (CEET). Available:

[<http://ecvam.jrc.it/index.htm>]

- (12) Balls, M., Botham, P.A., Bruner, L.H. and Spielmann H. (1995). The EC/HO international validation study on alternatives to the Draize eye irritation test. *Toxicol. In Vitro* 9:871-929.

- (13) Prinsen, M.K. (1996). The chicken enucleated eye test (CEET): A practical (pre)screen for the assessment of eye irritation/corrosion potential of test materials. *Food Chem. Toxicol.* 34:291-296.

- (14) Chamberlain, M., Gad, S.C., Gautheron, P. and Prinsen, M.K. (1997). IRAG Working Group I: Organotypic models for the assessment/prediction of ocular irritation. *Food Chem. Toxicol.* 35:23-37.

- (15) Siegel, J.D., Rhinehart, E., Jackson, M., Chiarello, L., and the Healthcare Infection Control Practices Advisory Committee (2007). Guideline for Isolation Precautions: Preventing Transmission of Infectious Agents in Healthcare Settings. Available:

[<http://www.cdc.gov/ncidod/dhqp/pdf/isolation2007.pdf>].

- (16) Maurer, J.K., Parker, R.D. and Jester J.V. (2002). Extent of corneal injury as the mechanistic basis for ocular irritation: key findings and recommendations for the development of alternative assays. *Reg. Tox. Pharmacol.* 36:106-117.

- (17) Burton, A.B.G., M. York and R.S. Lawrence (1981). The *in vitro* assessment of severe irritants. *Fd. Cosmet.- Toxicol.* 19, 471-480.



**▼ M2**

- (18) ICCVAM (2006). Background review document, Current Status of *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants: Bovine Corneal Opacity and Permeability (BCOP) Test Method. Available:

[[http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu\\_brd\\_bcop.htm](http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu_brd_bcop.htm)]

- (19) ICCVAM (2006). Background review document, Current Status of *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants Isolated Chicken Eye (ICE) Test Method. Available:

[[http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu\\_brd\\_bcop.htm](http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu_brd_bcop.htm)]

▼ M2*Appendix 1*

## DEFINITIONS

**Accuracy:** The closeness of agreement between test method results and accepted reference values. It is a measure of test method performance and one aspect of 'relevance'. The term is often used interchangeably with 'concordance', to mean the proportion of correct outcomes of a test method.

**Benchmark substance:** A substance used as a standard for comparison to a test substance. A benchmark substance should have the following properties; (i), a consistent and reliable source(s); (ii), structural and functional similarity to the class of substances being tested; (iii), known physical/chemical characteristics; (iv), supporting data on known effects; and (v), known potency in the range of the desired response

**Cornea:** The transparent part of the front of the eyeball that covers the iris and pupil and admits light to the interior.

**Corneal opacity:** Measurement of the extent of opaqueness of the cornea following exposure to a test substance. Increased corneal opacity is indicative of damage to the cornea.

**Corneal swelling:** An objective measurement in the ICE test of the extent of distention of the cornea following exposure to a test substance. It is expressed as a percentage and is calculated from baseline (pre-dose) corneal thickness measurements and the thickness recorded at regular intervals after exposure to the test material in the ICE test. The degree of corneal swelling is indicative of damage to the cornea.

**EPA Category 1:** Corrosive (irreversible destruction of ocular tissue) or corneal involvement or irritation persisting for more than 21 days (1).

**EU Category R41:** Production of tissue damage in the eye, or serious physical decay of vision, following application of a test substance to the anterior surface of the eye, which is not fully reversible within 21 days of application (2).

**False negative rate:** The proportion of all positive substances falsely identified by a test method as negative. It is one indicator of test method performance.

**False positive rate:** The proportion of all negative substances that are falsely identified by a test method as positive. It is one indicator of test method performance.

**Fluorescein retention:** A subjective measurement in the ICE test of the extent of fluorescein sodium that is retained by epithelial cells in the cornea following exposure to a test substance. The degree of fluorescein retention is indicative of damage to the corneal epithelium.

**GHS (Globally Harmonized System of Classification and Labelling of Chemicals):** A system proposing the classification of chemicals (substances and mixtures) according to standardized types and levels of physical, health and environmental hazards, and addressing corresponding communication elements, such as pictograms, signal words, hazard statements, precautionary statements and safety data sheets, so that to convey information on their adverse effects with a view to protect people (including employers, workers, transporters, consumers and emergency responders) and the environment (3).

**▼ M2**

**GHS Category 1:** Production of tissue damage in the eye, or serious physical decay of vision, following application of a test substance to the anterior surface of the eye, which is not fully reversible within 21 days of application (3).

**Hazard:** Inherent property of an agent or situation having the potential to cause adverse effects when an organism, system or (sub) population is exposed to that agent.

**Negative control:** An untreated replicate containing all components of a test system. This sample is processed with test substance-treated samples and other control samples to determine whether the solvent interacts with the test system.

**Non-irritant:** Substances that are not classified as EPA Category I, II, or III; EU Category R41 or R36; or GHS Category 1, 2A, or 2B ocular irritants (1)(2)(3).

**Ocular corrosive:** (a) A substance that causes irreversible tissue damage to the eye. (b) Substances that are classified as GHS Category 1, EPA Category I, or EU Category R41 ocular irritants (1)(2)(3).

**Ocular irritant:** (a) A substance that produces a reversible change in the eye following application to the anterior surface of the eye; (b) Substances that are classified as EPA Category II or III; EU Category R36; or GHS Category 2A, or 2B ocular irritants (1)(2)(3).

**Ocular severe irritant:** (a) A substance that causes tissue damage in the eye following application to the anterior surface of the eye that is not reversible within 21 days of application or causes serious physical decay of vision. (b) Substances that are classified as GHS Category 1, EPA Category I, or EU Category R41 ocular irritants (1)(2)(3).

**Positive control:** A replicate containing all components of a test system and treated with a substance known to induce a positive response. To ensure that variability in the positive control response across time can be assessed, the magnitude of the severe response should not be excessive.

**Reliability:** Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility and intra-laboratory repeatability.

**Slit-lamp microscope:** An instrument used to directly examine the eye under the magnification of a binocular microscope by creating a stereoscopic, erect image. In the ICE test method, this instrument is used to view the anterior structures of the chicken eye as well as to objectively measure corneal thickness with a depth-measuring device attachment.

**Solvent/vehicle control:** An untreated sample containing all components of a test system, including the solvent or vehicle that is processed with the test substance-treated and other control samples to establish the baseline response for the samples treated with the test substance dissolved in the same solvent or vehicle. When tested with a concurrent negative control, this sample also demonstrates whether the solvent or vehicle interacts with the test system.

**▼ M2**

**Tiered testing:** A stepwise testing strategy where all existing information on a test substance is reviewed, in a specified order, using a weight of evidence process at each tier to determine if sufficient information is available for a hazard classification decision, prior to progression to the next tier. If the irritancy potential of a test substance can be assigned based on the existing information, no additional testing is required. If the irritancy potential of a test substance cannot be assigned based on the existing information, a step-wise sequential animal testing procedure is performed until an unequivocal classification can be made.

**Validated test method:** A test method for which validation studies have been completed to determine the relevance (including accuracy) and reliability for a specific purpose. It is important to note that a validated test method may not have sufficient performance in terms of accuracy and reliability to be found acceptable for the proposed purpose.

**Weight-of-evidence:** The process of considering the strengths and weaknesses of various pieces of information in reaching and supporting a conclusion concerning the hazard potential of a substance.

## ▼M2

## Appendix 2

## PROFICIENCY CHEMICALS FOR THE ICE TEST METHOD

Prior to routine use of a test method that adheres to this test method, laboratories may wish to demonstrate technical proficiency by correctly identifying the ocular corrosivity classification of the 10 substances recommended in Table 1. These substances were selected to represent the range of responses for local eye irritation/corrosion, which is based on results in the *in vivo* rabbit eye test (TG 405) (*i.e.*, Categories 1, 2A, 2B, or Not Classified or Labeled according to the UN GHS)(3)(7). However, considering the validated usefulness of these assays (*i.e.*, to identify ocular corrosives/severe irritants only), there are only two test outcomes for classification purposes (corrosive/severe irritant or non-corrosive/non-severe irritant) to demonstrate proficiency. Other selection criteria were that substances are commercially available, there are high quality *in vivo* reference data available, and there are high quality data from the two *in vitro* methods for which Test Guidelines are being developed. For this reason, irritant substances were selected from the ICCVAM recommended list of 122 reference substances for the validation of *in vitro* ocular toxicity test methods (see Appendix H, ICCVAM Recommended Reference Substances List)(4). Reference data are available in the ICCVAM Background Review Documents for the Bovine Corneal Opacity and Permeability (BCOP) and the ICE test methods (18) (19).

Table 1

## Recommended substances for demonstrating technical proficiency with ICE

Chemical	CASRN	Chemical Class (1)	Physical Form	<i>In Vivo</i> Classification (2)	<i>In Vitro</i> classification (3)
Benzalkonium chloride (5 %)	8001-54-5	Onium compound	Liquid	Category 1	Corrosive/Severe Irritant
Chlorhexidine	55-56-1	Amine, Amidine	Solid	Category 1	Corrosive/Severe Irritant
Dibenzoyl-L-tartaric acid	2743-38-6	Carboxylic acid, Ester	Solid	Category 1	Corrosive/Severe Irritant
Imidazole	288-32-4	Heterocyclic	Solid	Category 1	Corrosive/Severe Irritant
Trichloroacetic acid (30 %)	76-03-9	Carboxylic Acid	Liquid	Category 1	Corrosive/Severe Irritant
2,6-Dichlorobenzoyl chloride	4659-45-4	Acyl halide	Liquid	Category 2A	Non-corrosive/Non-severe irritant
Ethyl-2-methylacetoacetate	609-14-3	Ketone, Ester	Liquid	Category 2B	Non-corrosive/Non-severe irritant
Ammonium nitrate	6484-52-2	Inorganic salt	Solid	Category 2A	Non-corrosive/Non-severe irritant

▼ **M2**

Chemical	CASRN	Chemical Class <sup>(1)</sup>	Physical Form	<i>In Vivo</i> Classification <sup>(2)</sup>	<i>In Vitro</i> lassification <sup>(3)</sup>
Glycerol	56-81-5	Alcohol	Liquid	Not Labeled	Non-corrosive/Non-severe irritant
n-Hexane	110-54-3	Hydrocarbon (acyclic)	Liquid	Not Labeled	Non-corrosive/Non-severe irritant

Abbreviations: CASRN = Chemical Abstracts Service Registry Number

<sup>(1)</sup> Chemical classes were assigned to each test substance using a standard classification scheme, based on the National Library of Medicine Medical Subject Headings (MeSH) classification system (available at <http://www.nlm.nih.gov/mesh>).

<sup>(2)</sup> Based on results from the *in vivo* rabbit eye test (OECD TG 405) and using the UN GHS (3)(7).

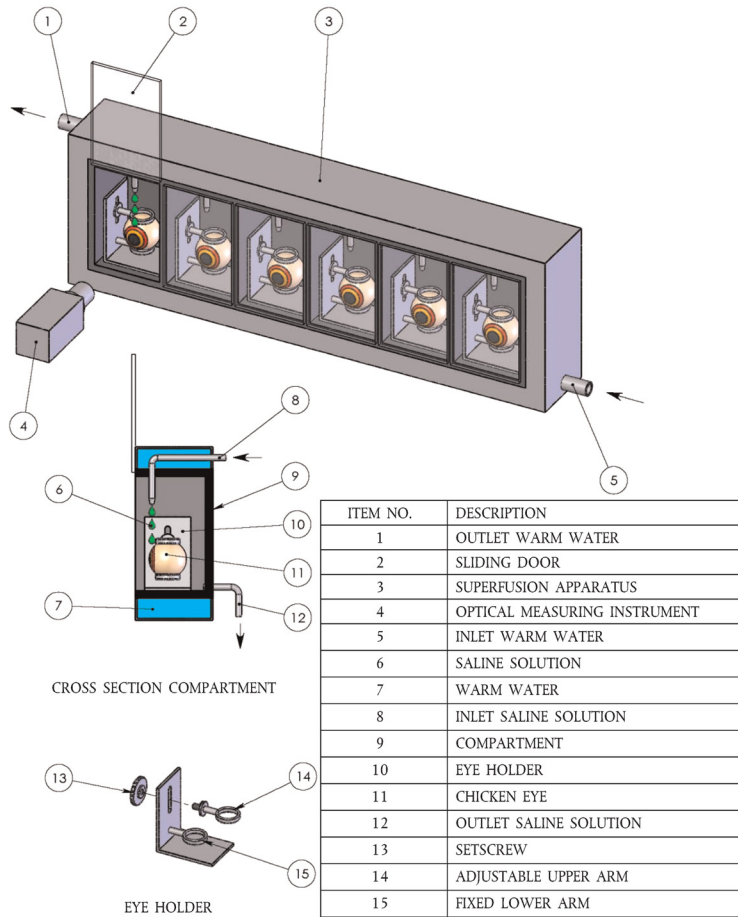
<sup>(3)</sup> Based on results in BCOP and ICE.

▼ **M2**

Appendix 3

**Diagrams of the ICE superfusion apparatus and eye clamps**

(See Burton et al. (17) for additional generic descriptions of the superfusion apparatus and eye clamp)



▼ **M3****B.49. IN VITRO MAMMALIAN CELL MICRONUCLEUS TEST**

## INTRODUCTION

1. The *in vitro* micronucleus (MNvit) assay is a genotoxicity test for the detection of micronuclei (MN) in the cytoplasm of interphase cells. Micronuclei may originate from acentric chromosome fragments (*i.e.* lacking a centromere), or whole chromosomes that are unable to migrate to the poles during the anaphase stage of cell division. The assay detects the activity of clastogenic and aneugenic chemicals (substances and mixtures) (1) (2) in cells that have undergone cell division during or after exposure to the test substance. This Test Method (TM) allows the use of protocols with and without the actin polymerisation inhibitor cytochalasin B (cytoB). The addition of cytoB prior to the targeted mitosis allows for the identification and selective analysis of micronucleus frequency in cells that have completed one mitosis because such cells are binucleate (3) (4). This TM also allows the use of protocols without cytokinesis block, provided there is evidence that the cell population analysed has undergone mitosis.
2. In addition to using the MNvit assay to identify chemicals (substances and mixtures) that induce micronuclei, the use of a cytokinesis block, immunochemical labelling of kinetochores, or hybridisation with centromeric/te-lomeric probes (fluorescence *in situ* hybridisation (FISH)), also can provide information on the mechanisms of chromosome damage and micro-nucleus formation (5) (6) (7) (8) (9) (10) (11) (12) (13) (14) (15) (16). The labelling and hybridisation procedures can be used when there is an increase in micronucleus formation and the investigator wishes to determine if the increase was the result of clastogenic and/or aneugenic events.
3. Micronuclei represent damage that has been transmitted to daughter cells, whereas chromosome aberrations scored in metaphase cells may not be transmitted. Because micronuclei in interphase cells can be assessed relatively objectively, laboratory personnel need only determine whether or not the cells have undergone division and how many cells contain a micronucleus. As a result, the preparations can be scored relatively quickly and analysis can be automated. This makes it practical to score thousands instead of hundreds of cells per treatment, increasing the power of the assay. Finally, as micronuclei may arise from lagging chromosomes, there is the potential to detect aneuploidy-inducing agents that are difficult to study in conventional chromosomal aberration tests, *e.g.* OECD Test Guideline 473 (Chapter B.10 of this Annex) (17). However, the MNvit assay does not allow for the differentiation of chemicals inducing poly-ploidy from those inducing clastogenicity without special techniques such as FISH described under paragraph 2.
4. The MNvit assay is an *in vitro* method that typically uses cultured human or rodent cells. It provides a comprehensive basis for investigating chro-mosome damaging potential *in vitro* because both aneugens and clastogens can be detected.
5. The MNvit assay is robust and effective in a variety of cell types, and in the presence or absence of cytoB. There are extensive data to support the validity of the MNvit assay using various rodent cell lines (CHO, V79, CHL/IU, and L5178Y) and human lymphocytes (18) (19) (20) (21) (22) (23) (24) (25) (26) (27) (28) (29) (30) (31). These include, in particular, the international validation studies coordinated by the Société Française de Toxicologie Génétique (SFTG) (18) (19) (20) (21) (22) and the reports of the International Workshop on Genotoxicity Testing (4) (16). The available data have also been re-evaluated in a weight-of-evidence retrospective



**▼ M3**

validation study by the European Centre for the Validation of Alternative Methods (ECVAM) of the European Commission, and the test method has been endorsed as scientifically valid by the ECVAM Scientific Advisory Committee (ESAC) (32) (33) (34). The use of the human TK6 lymphoblastoid cell line (35), HepG2 cells (36) (37) and primary Syrian Hamster Embryo cells (38) has been described, although they have not been used in validation studies.

## DEFINITIONS

6. Definitions used are provided in Appendix 1.

## INITIAL CONSIDERATIONS

7. Tests conducted *in vitro* generally require the use of an exogenous source of metabolic activation unless the cells are metabolically competent with respect to the substances being tested. The exogenous metabolic activation system does not entirely mimic *in vivo* conditions. Care should also be taken to avoid conditions that would lead to artefactual positive results which do not reflect intrinsic mutagenicity, and may arise from such factors as marked changes in pH or osmolality, or by high levels of cytotoxicity (39) (40) (41). If the test chemical causes a change in the pH of the medium at the time of addition, the pH should be adjusted, preferably by buffering the stock solution so that all the volumes at all test concentrations, and for all controls, remain the same.
8. To analyse the induction of micronuclei, it is essential that mitosis has occurred in both treated and untreated cultures. The most informative stage for scoring micronuclei is in cells that have completed one mitosis during or after treatment with the test substance.

## PRINCIPLE OF THE TEST

9. Cell cultures of human or mammalian origin are exposed to the test substance both with and without an exogenous source of metabolic activation unless cells with an adequate metabolising capability are used. Concurrent solvent/vehicle (VC) and positive control chemicals (PC) are included in all tests.
10. During or after exposure to the test substance, the cells are grown for a period sufficient to allow chromosome or spindle damage to lead to the formation of micronuclei in interphase cells. For induction of aneuploidy, the test substance should ordinarily be present during mitosis. Harvested and stained interphase cells are analysed for the presence of micronuclei. Ideally, micronuclei should only be scored in those cells that have completed mitosis during exposure to the test substance or during the post-exposure period, if one is used. In cultures that have been treated with a cytokinesis blocker, this is achieved by scoring only binucleate cells. In the absence of a cytokinesis blocker, it is important to demonstrate that the cells analysed are likely to have undergone cell division during or after exposure to the test substance. For all protocols, it is important to demonstrate that cell proliferation has occurred in both the control and treated cultures, and the extent of test substance-induced cytotoxicity or cytostasis should be assessed in the cultures (or in parallel cultures) that are scored for micronuclei.

**▼ M3**

## DESCRIPTION OF THE ASSAY

*Preparations*

11. Cultured primary human peripheral blood lymphocytes (5) (19) (42) (43) and a number of rodent cell lines such as CHO, V79, CHL/IU, and L5178Y cells may be used (18) (19) (20) (21) (22) (25) (26) (27) (28) (30). The use of other cell lines and types should be justified based on their demonstrated performance in the assay, as described in the Acceptability Criteria section. Because the background frequency of micronuclei will influence the sensitivity of the assay, it is recommended that cell types with a low, stable background frequency of micronucleus formation be used.
  
12. Human peripheral blood lymphocytes should be obtained from young (approximately 18-35 years of age), healthy, non-smoking individuals with no known recent exposures to genotoxic chemicals or radiation. If cells from more than one donor are pooled for use, the number of donors should be specified. The micronucleus frequency increases with age and this trend is more marked in females than in males (44) and this should be taken into account in the selection of donor cells for pooling.

*Media and culture conditions*

13. Appropriate culture medium and incubation conditions (culture vessels, CO<sub>2</sub> concentration, temperature, and humidity) should be used for maintaining cultures. Established cell lines and strains should be checked routinely for the stability of the modal chromosome number and the absence of mycoplasma contamination, and should not be used if contaminated or if the modal chromosome number has changed. The normal cell cycle time for the culture conditions used in the testing laboratory should be known. If the cytokinesis-block method is used then the concentration of the cytokinesis inhibitor should be optimised for the particular cell type and should be shown to produce a good yield of binucleate cells for scoring.

*Preparation of cultures*

14. Established cell lines and strains: cells are propagated from stock cultures, seeded in culture medium at a density such that the cultures will not reach confluency in monolayers, and suspension cultures will not reach excessive density before the time of harvest, and incubated at 37 °C.
  
15. Lymphocytes: whole blood treated with an anti-coagulant (*e.g.* heparin), or separated lymphocytes, are cultured in the presence of a mitogen *e.g.* phytohaemagglutinin (PHA) prior to exposure to the test substance and cytoB.

*Metabolic activation*

16. Exogenous metabolising systems should be used when using cells with inadequate endogenous metabolic capacity. The most commonly used system is a co-factor-supplemented post-mitochondrial fraction (S9) prepared from the livers of rodents treated with enzyme-inducing agents such as Aroclor 1254 (45) (46) or a combination of phenobarbitone and  $\beta$ -naphthoflavone (46) (47) (48) (49). The latter combination does not conflict with the Stockholm Convention on Persistent Organic Pollutants (50) and Regulation (EC) No 850/2004 on Persistent Organic Pollutants (66) and has been shown to be as effective as Aroclor 1254 for inducing mixed-function oxidases (46) (47) (48) (49). The S9 fraction typically is used at concentrations ranging from 1-10 % (v/v) in the final test medium. The condition of a metabolic activation system may depend upon the class of chemical being tested and in some cases it may be appropriate to utilise more than one S9 concentration.

**▼ M3**

17. Genetically engineered cell lines expressing specific human or rodent activating enzymes may eliminate the need for an exogenous metabolic activation system, and may be used as the test cells. In such cases the choice of the cell lines used should be scientifically justified, *e.g.* by relevance of the mixed function oxidases for the metabolism of the test substance (51), and their responsiveness to known clastogens and aneugens (see separate section on Acceptability Criteria). It should be recognised that the substance being tested may not be metabolised by the expressed mixed function oxidase(s); in this case, the negative results would not indicate that the test substance cannot induce micronuclei.

## Test substance preparation

18. Solid chemicals should be dissolved in appropriate solvents or vehicles and diluted, if appropriate, prior to treatment of the cells. Liquid chemicals may be added directly to the test systems and/or diluted prior to treatment. Gases or volatile chemicals should be tested by appropriate modifications to the standard protocols, such as treatment in sealed vessels (52) (53). Fresh preparations of the test substance should be used unless stability data demonstrate the acceptability of storage.

*Test Conditions*

## Solvents/vehicles

19. The solvent/vehicle should not react with the test substance, or be incompatible with the survival of the cells or with the maintenance of S9 activity at the concentration used. If other than well established solvent/vehicles (*e.g.* water, cell culture medium, dimethyl sulfoxide) are used, their use should be supported by data indicating their compatibility with the test substance and their lack of genetic toxicity. It is recommended that, wherever possible, the use of an aqueous solvent/vehicle should be considered first.

## Use of cytoB as a cytokinesis blocker

20. One of the most important considerations in the performance of the MNvit assay is ensuring that the cells being scored have completed mitosis during the treatment or the post-treatment incubation period, if one is used. CytoB is the agent that has been most widely used to block cytokinesis because it inhibits actin assembly, and thus prevents separation of daughter cells after mitosis, leading to the formation of binucleated cells (5) (54) (55). Micro-nucleus scoring, therefore, can be limited to cells that have gone through mitosis during or after treatment. The effect of the test substance on cell proliferation kinetics can be measured simultaneously. CytoB should be used as a cytokinesis blocker when human lymphocytes are used because cell cycle times will be variable within cultures and among donors and because not all lymphocytes will respond to PHA. Other methods have been used when testing cell lines to determine if the cells being scored have divided; these are addressed below (see Paragraph 26).
21. The appropriate concentration of cytoB should be determined by the laboratory for each cell type to achieve the optimal frequency of binucleated cells in the solvent/vehicle control cultures. The appropriate concentration of cytoB is usually between 3 and 6 µg/ml.

**▼ M3**

Measuring cell proliferation and cytotoxicity and choosing exposure concentrations

22. When determining the highest test substance concentration to be tested, concentrations that have the capability of producing artefactual positive responses, such as those producing excessive cytotoxicity, precipitation in the culture medium, and marked changes in pH or osmolality (39) (40) (41), should be avoided.
23. Measurements of cell proliferation are made to ensure that the treated cells have undergone mitosis during the assay and that the treatments are conducted at appropriate levels of cytotoxicity (see Paragraph 29). Cytotoxicity should be determined with and without metabolic activation in cells that require metabolic activation using the relative increase in cell counts (RICC) or relative population doubling (RPD) (see Appendix 2 for formulas) unless cytoB is used. When cytoB is used, cytotoxicity can be determined using the replication index (RI) (see Appendix 2 for formula).
24. Treatment of cultures with cytoB, and measurement of the relative frequencies of mononucleate, binucleate, and multi-nucleate cells in the culture, provides an accurate method of quantifying the effect on cell proliferation and the cytotoxic or cytostatic activity of a treatment (5), and ensures that only cells that divided during or after treatment are scored.
25. In studies with cytoB, cytostasis/cytotoxicity can be quantified from the cytokinesis-block proliferation index (CBPI) (5) (26) (56) or may be derived from the RI from at least 500 cells per culture (see Appendix 2 for formulas). When cytoB is used to assess cell proliferation, a CBPI or RI should be determined from at least 500 cells per culture. These measurements among others can be used to estimate cytotoxicity by comparing values in the treated and control cultures. Assessment of other markers of cytotoxicity (e.g. confluency, cell number, apoptosis, necrosis, metaphase counting) can provide useful information.
26. In studies without cytoB, it is necessary to demonstrate that the cells scored in the culture have undergone division during or following treatment with the test substance, otherwise false negative responses may be produced. Methods that have been used for ensuring that divided cells are being scored include incorporation and subsequent detection of bromodeoxyuridine (BrdU) to identify cells that have replicated (57), the formation of clones when cells from permanent cell lines are treated and scored *in situ* on a microscope slide (Proliferation Index (PI)) (25) (26) (27) (28), or the measurement of Relative Population Doubling (RPD) or Relative Increase in Cell Count (RICC) or other proven methods (16) (56) (58) (59) (see Appendix 2 for formulas). Assessment of other markers for cytotoxicity or cytostasis (e.g. confluency, cell number, apoptosis, necrosis, metaphase counting) can provide useful information.
27. At least three analysable test concentrations should be evaluated. In order to achieve this, it may be necessary to perform the experiment using a larger number of closely spaced concentrations and analyse micronucleus formation in those concentrations providing the appropriate range of cytotoxicities. An alternative strategy is to perform a preliminary cytotoxicity test to narrow the range for the definitive test.
28. The highest concentration should aim to produce  $55 \pm 5$  % cytotoxicity. Higher levels may induce chromosome damage as a secondary effect of cytotoxicity (60). Where cytotoxicity occurs, the test concentrations selected should cover a range from that producing  $55 \pm 5$  % cytotoxicity, to little or no cytotoxicity.

**▼ M3**

29. If no cytotoxicity or precipitate is observed, the highest test concentration should correspond to 0,01 M, 5 mg/mL or 5 µL/mL, whichever is the lowest. The concentrations selected for analysis should, in general, be separated by a spacing of no more than 10. For test substances that exhibit a steep concentration-response curve, it may be necessary to more closely space the test substance concentrations so that cultures in the moderate and low toxicity ranges also will be scored.
  
30. When solubility is a limiting factor, the maximum concentration, if not limited by cytotoxicity, should be the lowest concentration at which minimal precipitate is visible in cultures, provided there is no interference with scoring. Evaluation of precipitation should be done by methods such as light microscopy, noting precipitate that persists, or appears during culture (by the end of treatment).

**Controls**

31. Concurrent positive and solvent/vehicle controls both with and without metabolic activation should be included in each experiment.
  
32. PC are needed to demonstrate the ability of the cells used, and the test protocol, to identify clastogens and aneugens, and to affirm the metabolic capability of the S9 preparation. The PC should employ known inducers of micronucleus formation at concentrations expected to give small, but reproducible increases over background, and demonstrate the sensitivity of the test system. PC concentrations should be chosen so that the effects are clear but do not immediately reveal the identity of the coded slides to the reader.
  
33. A clastogen that requires metabolic activation (*e.g.* cyclophosphamide; benzo(a)pyrene) should be used to demonstrate both the metabolic competence and the ability of the test system to detect clastogens. Other PC may be used if justified. Because some PC that need metabolic activation may be active without exogenous metabolic activation under certain treatment conditions or in certain cell lines, the need for metabolic activation, and the activity of the S9 preparation, should be tested in the selected cell line and at the selected concentrations.
  
34. At the present time, no aneugens are known that require metabolic activation for their genotoxic activity (16). Currently accepted PC for aneugenic activity are, for example, colchicine and vinblastine. Other chemicals may be used if they induce micronuclei solely, or primarily, through aneugenic activity. To avoid the need for two PC (for clastogenicity and aneugenicity) without metabolic activation, the aneugenicity control can serve as the PC without S9, and the clastogenicity control can be used to test the adequacy of the metabolic activation system used. PC for both clastogenicity and aneugenicity should be used in cells that do not require S9. Suggested PC are included in Appendix 3.
  
35. The use of chemical class-related PC may be considered, when suitable chemicals are available. All PC used should be appropriate for the cell type and activation conditions.

▼ **M3**

36. Solvent/vehicle controls should be included for every harvest time. In addition, untreated NC (lacking solvent/vehicle) should also be used unless there are published or laboratory historical control data demonstrating that no genotoxic or other deleterious effects are induced by the chosen solvent at the concentrations used.

## TEST PROCEDURE

*Treatment Schedule*

37. In order to maximise the probability of detecting an aneugen or clastogen acting at a specific stage in the cell cycle, it is important that sufficient numbers of cells are treated with the test substance during all stages of their cell cycles. The treatment schedule for cell lines and primary cell cultures may, therefore, differ somewhat from that for lymphocytes which require mitogenic stimulation to begin their cell cycle and these are considered in Paragraphs 41-43 (16).
38. Theoretical considerations, together with published data (18) indicate that most aneugens and clastogens will be detected by a short term treatment period of 3 to 6 hrs in the presence and absence of S9, followed by removal of the test substance and a growth period of 1,5-2,0 cell cycles (6). Cells are sampled at a time equivalent to about 1,5-2,0 times the normal (*i.e.* untreated) cell cycle length either after the beginning or at the end of treatment (See Table 1). Sampling or recovery times may be extended if it is known or suspected that the test substance affects the cell cycling time (*e.g.* when testing nucleoside analogues).
39. Because of the potential cytotoxicity of S9 preparations for cultured mammalian cells, an extended exposure treatment of 1,5-2,0 normal cell cycles is used only in the absence of S9. In the extended treatment, options are offered to allow treatment of the cells with the test chemical in the absence or presence of cytoB. These options address situations where there may be concern regarding possible interactions between the test substance and cytoB.
40. The suggested cell treatment schedules are presented in Table 1. These general treatment schedules may be modified depending on the stability or reactivity of the test substance or the particular growth characteristics of the cells being used. All treatments should commence and end while the cells are growing exponentially. These schedules are presented in more details in paragraphs 41-47 following.

Table 1

**Cell treatment and harvest times for the MNvit assay**

Lymphocytes, primary cells and cell lines treated with cytoB	+ S9	Treat for 3-6 hrs in the presence of S9; remove the S9 and treatment medium; add fresh medium and cytoB; harvest 1,5-2,0 normal cell cycles later.
	- S9 Short exposure	Treat for 3-6 hrs; remove the treatment medium; add fresh medium and cytoB; harvest 1,5-2,0 normal cell cycles later.

▼ M3

	– S9 Extended exposure	<p><i>Option A:</i> Treat for 1,5-2 normal cell cycles in the presence of cytoB; harvest at the end of the exposure period.</p> <p><i>Option B:</i> Treat for 1,5-2,0 normal cell cycles; remove the test substance; add fresh medium and cytoB; harvest 1,5-2,0 normal cell cycles later.</p>
--	---------------------------	--

Cell lines treated without cytoB

(Identical to the treatment schedules outlined above with the exception that no cytoB is added)

*Lymphocytes, primary cells, and cell lines with cytoB*

41. For lymphocytes, the most efficient approach is to start the exposure to the test substance at 44-48 hrs after PHA stimulation, when cycle synchronisation will have disappeared (5). In the initial assay, cells are treated for 3 to 6 hrs with the test substance in the absence and presence of S9. The treatment medium is removed and replaced with fresh medium containing cytoB, and the cells are harvested 1,5-2,0 normal cell cycles later.
42. If both initial tests of the short (3-6 hrs) treatment are negative or equivocal, a subsequent, extended exposure treatment without S9 is used. Two treatment options are available and are equally acceptable. However, it might be more appropriate to follow Option A for stimulated lymphocytes where exponential growth may be declining at 96 hrs following stimulation. Also, cultures of cells should not have reached confluence by the final sampling time in Option B.
- Option A: The cells are treated with the test substance for 1,5-2,0 normal cell cycles, and harvested at the end of the treatment time.
  - Option B: The cells are treated with the test substance for 1,5-2,0 normal cell cycles. The treatment medium is removed and replaced with fresh medium, and the cells are harvested after additional 1,5-2,0 normal cell cycles.
43. Primary cells and cell lines should be treated in a similar manner to lymphocytes except that it is not necessary to stimulate them with PHA for 44-48 hrs. Cells other than lymphocytes should be exposed such that at the time of study termination, the cells are still in log-phase growth.

*Cell lines without cytoB*

44. Cells should be treated for 3-6 hrs in the presence and absence of S9. The treatment medium is removed and replaced with fresh medium, and the cells are harvested 1,5-2,0 normal cell cycles later.

**▼ M3**

45. If both initial tests of the short (3-6 hrs) treatment are negative or equivocal, a subsequent, extended exposure treatment (without S9) is used. Two treatment options are available, both of which are equally acceptable:
- Option A: The cells are treated with the test substance for 1,5-2,0 normal cell cycles, and harvested at the end of the treatment time.
  - Option B: The cells are treated with the test substance for 1,5-2,0 normal cell cycles. The treatment medium is removed and replaced with fresh medium, and the cells are harvested after additional 1,5-2,0 normal cell cycles.
46. In monolayers, mitotic cells (identifiable as being round and detaching from the surface) may be present at the end of the 3-6 hr treatment. Because these mitotic cells are easily detached, they can be lost when the medium containing the test substance is removed. Care should be taken to collect these when cultures are washed, and to return them to the cultures, to avoid losing cells that are in mitosis, and at risk for micronuclei, at the time of harvest.

*Number of cultures*

47. Duplicate cultures should be used for each test substance concentration and for the vehicle/solvent and NC cultures. Where minimal variation between duplicate cultures can be demonstrated from historical laboratory data, it may be acceptable for single cultures to be used. If single cultures are used, it is recommended that an increased number of concentrations be analysed.

*Cell harvest and slide preparation*

48. Each culture is harvested and processed separately. Cell preparation may involve hypotonic treatment, but this step is not necessary if adequate cell spreading is otherwise achieved. Different techniques can be used in slide preparation provided that high-quality cell preparations for scoring are obtained. Cell cytoplasm should be retained to allow the detection of micronuclei and (in the cytokinesis-block method) reliable identification of binucleate cells.
49. The slides can be stained using various methods, such as Giemsa or fluorescent DNA specific dyes (59). The use of a DNA specific stain (*e.g.* acridine orange (61) or Hoechst 33258 plus pyronin-Y (62)) can eliminate some of the artefacts associated with using a non-DNA specific stain. Anti-kinetochore antibodies, FISH with pancentromeric DNA probes, or primed *in situ* labelling with pancentromere-specific primers, together with appropriate DNA counterstaining, can be used to identify the contents (chromosome/chromosomal fragment) of micronuclei if mechanistic information of their formation is of interest (15) (16). Other methods for differentiation between clastogens and aneugens may be used if they have been shown to be effective.

*Analysis*

50. All slides, including those of the solvent/vehicle and the controls, should be independently coded before the microscopic analysis. Alternatively, coded samples can be analysed using a validated, automated flow cytometric or image analysis system.



**▼ M3**

51. In cytoB-treated cultures, micronucleus frequencies should be analysed in at least 2 000 binucleated cells per concentration (at least 1 000 binucleated cells per culture; two cultures per concentration). If single cultures are used, at least 2 000 binucleated cells per concentration should be scored from that culture. If substantially fewer than 1 000 binucleate cells per culture, or 2 000 if a single culture is used, are available for scoring at each concentration, and if a significant increase in micronuclei is not detected, the test should be repeated using more cells, or at less toxic concentrations, whichever is appropriate. Care should be taken not to score binucleate cells with irregular shapes or where the two nuclei differ greatly in size; neither should binucleate cells be confused with poorly spread multi-nucleate cells. Cells containing more than two main nuclei should not be analysed for micronuclei, as the baseline micronucleus frequency may be higher in these cells (63) (64) Scoring of mononucleate cells is acceptable if the test substance is shown to interfere with cytoB activity.
52. In cell lines assayed without cytoB treatment, micronuclei should be scored in at least 2 000 cells per concentration (at least 1 000 cells per culture; two cultures per concentration). Where only one culture per concentration is used, at least 2 000 cells should be scored from that culture.
53. When cytoB is used, a CBPI or an RI should be determined to assess cell proliferation (see Appendix 2) using at least 500 cells per culture. When treatments are performed in the absence of cytoB, it is essential to provide evidence that the cells being scored have proliferated, as discussed in Paragraphs 24-27.

*Acceptability criteria*

54. A laboratory proposing to use the MNvit assay described in this TM should demonstrate its ability to reliably and accurately detect chemicals of known aneugenic and clastogenic activity, with and without metabolic activation, as well as known negative chemicals, using the reference chemicals in Appendix 3. As evidence of its ability to perform this TM correctly, the laboratory should provide evidence that the cells being scored for micronucleus formation have completed one nuclear division if the test is performed without the use of cytoB.
55. The chemicals in Appendix 3 are recommended for use as reference chemicals. Substitute or additional chemicals can be included if their activity is known and if they induce micronuclei by the same mechanisms of action, and if they are shown to be relevant to the chemicals that will be tested using the MNvit procedure. Justification could include a validation study employing a broad variety of substances or focused on a narrower spectrum based on the chemical class of the test substance or the mechanism of damage being studied.
56. Solvent/vehicle control and untreated cultures should give reproducibly low and consistent micronuclei frequencies (typically 5-25 micronuclei/1 000 cells for the cell types identified in paragraph 11). Other cell types may have different ranges of responses which should be determined when validating them for use in the MNvit assay. Data from negative, solvent, and PC should be used to establish historical control ranges. These values should be used in deciding the adequacy of the concurrent NC/PC for an experiment.

**▼ M3**

57. If minor changes to the protocol (*e.g.* use of automated instead of manual scoring techniques; use of a new cell type) are proposed for the assay, then the effectiveness of the change should be demonstrated before the modified protocol can be considered acceptable for use. Demonstration of effectiveness includes demonstration that the major mechanisms of chromosome breakage and gain or loss can be detected, and that appropriate positive and negative results can be achieved for the class of the individual substance, or the broad range of substances, to be tested.

**DATA AND REPORTING***Treatment of results*

58. If the cytokinesis-block technique is used, only the frequencies of binucleate cells with micronuclei (independent of the number of micronuclei per cell) are used in the evaluation of micronucleus induction. Scoring of the numbers of cells with one, two, or more micronuclei could provide useful information, but is not mandatory.
59. Concurrent measures of cytotoxicity and/or cytostasis for all treated and solvent/vehicle control cultures should be determined (58). The CBPI or the RI should be calculated for all treated and control cultures as measurements of cell cycle delay when the cytokinesis-block method is used. In the absence of cytoB, the RPD or the RICC or PI should be used (see Appendix 2).
60. Individual culture data should be provided. Additionally, all data should be summarised in tabular form.
61. Chemicals that induce micronuclei in the MNvit assay may do so because they induce chromosome breakage, chromosome loss, or a combination of the two. Further analysis using anti-kinetochore antibodies, centromere specific *in situ* probes, or other methods may be used to determine whether the mechanism of micronucleus induction is due to clastogenic and/or aneugenic activity.

*Evaluation and interpretation of results*

62. There is no requirement for verification by additional testing of a clear positive or negative response. Equivocal results may be clarified by analysis of another 1 000 cells from all the cultures to avoid loss of blinding. If this approach does not resolve the result, further testing should be performed. Modification of study parameters over an extended or narrowed range of conditions, as appropriate, should be considered in follow-up experiments. Study parameters that might be modified include the test concentration spacing, the timing of treatment and cell harvest, and/or the metabolic activation conditions.
63. There are several criteria for determining a positive result, such as a concentration-related increase or a statistically significant increase in the number of cells containing micronuclei. The biological relevance of the results should be considered first. Consideration of whether the observed values are within or outside of the historical control range can provide guidance when evaluating the biological significance of the response. Appropriate statistical methods may be used as an aid in evaluating the test results (65). However, the results of statistical testing should be assessed with respect to dose-response relationship. Reproducibility and historical data should also be taken into consideration.

**▼ M3**

64. Although most experiments will give clearly positive or negative results, in some cases the data set will preclude making a definite judgement about the activity of the test substance. These equivocal or questionable responses may occur regardless of the number of times the experiment is repeated.
65. Positive results from the MNvit assay indicate that the test substance induces chromosome breakage or chromosome loss, in cultured mammalian cells. Negative results indicate that, under the test conditions used, the test substance does not induce chromosome breaks and/or gain or loss in cultured mammalian cells.

*Test Report*

66. The test report should include at least the following information, if relevant to the conduct of the study:

*Test chemical:*

- identification data and Chemical Abstract Services Registry Number and EC Number;
- physical nature and purity;
- physico-chemical properties relevant to the conduct of the study;
- reactivity of the test chemical with the solvent/vehicle or cell culture media;

*Solvent/Vehicle:*

- justification for choice of solvent/vehicle;
- solubility and stability of the test substance in solvent/vehicle;

*Cells:*

- type and source of cells used;
- suitability of the cell type used;
- absence of mycoplasma, if applicable;
- information on cell cycle length, doubling time or proliferation index;
- where lymphocytes are used, sex, age and number of blood donors, if applicable;
- where lymphocytes are used, whether whole blood or separated lymphocytes are exposed;
- number of passages, if applicable;
- methods for maintenance of cell cultures, if applicable;
- modal number of chromosomes;
- normal (negative control) cell cycle time;

*Test Conditions:*

- identity of cytokinesis blocking substance (e.g. cytoB), if used, and its concentration and duration of cell exposure;
- rationale for selection of concentrations and number of cultures, including cytotoxicity data and solubility limitations, if available;
- composition of media, CO<sub>2</sub> concentration, if applicable;
- concentrations of test substance;

**▼ M3**

- concentration (and/or volume) of vehicle and test substance added;
- incubation temperature and time;
- duration of treatment;
- harvest time after treatment;
- cell density at seeding, if applicable;
- type and composition of metabolic activation system, including acceptability criteria;
- positive control chemicals and negative controls;
- methods of slide preparation and staining technique used;
- criteria for micronucleus identification;
- numbers of cells analysed;
- methods for the measurements of cytotoxicity;
- any supplementary information relevant to cytotoxicity;
- criteria for considering studies as positive, negative, or equivocal;
- method(s) of statistical analysis used;
- methods, such as use of kinetochore antibody, to characterise whether micronuclei contain whole or fragmented chromosomes, if applicable;

*Results:*

- measurement of cytotoxicity used, *e.g.* CBPI or RI in the case of cytokinesis-block method; RICC, RPD or PI when cytokinesis-block methods are not used; other observations when applicable, *e.g.* cell confluency, apoptosis, necrosis, metaphase counting, frequency of binucleated cells;
- signs of precipitation;
- data on pH and osmolality of the treatment medium, if determined;
- definition of acceptable cells for analysis;
- distribution of mono-, bi-, and multi-nucleated cells if a cytokinesis block method is used;
- number of cells with micronuclei given separately for each treated and control culture, and defining whether from binucleate or mononucleate cells, where appropriate;
- concentration-response relationship, where possible;
- concurrent negative (solvent/vehicle) and positive control chemical data (concentrations and solvents);
- historical negative (solvent/vehicle) and positive control chemical data, with ranges, means and standard deviation and confidence interval (*e.g.* 95 %);
- statistical analysis; p-values if any;

*Discussion of the results**Conclusions*

▼ **M3**

## LITERATURE

- (1) Kirsch-Volders, M. (1997), Towards a validation of the micronucleus test. *Mutation Res.*, 392, 1-4.
- (2) Parry, J.M. and Sors, A. (1993), The detection and assessment of the aneuploidy potential of environmental chemicals: the European Community aneuploidy project. *Mutation Res.*, 287, 3-15.
- (3) Fenech, M. and Morley, A.A. (1985), Solutions to the kinetic problem in the micronucleus assay. *Cytobios.*, 43, 233-246.
- (4) Kirsch-Volders, M., Sofuni, T., Aardema, M., Albertini, S., Eastmond, D., Fenech, M., Ishidate, M. Jr, Lorge, E., Norppa, H., Surrallés, J., von der Hude, W. and Wakata, A. (2000), Report from the *In Vitro* Micronucleus Assay Working Group. *Environ. Mol. Mutagen.*, 35, 167-172.
- (5) Fenech, M. (2007), Cytokinesis-block micronucleus cytome assay. *Nature Protocols*, 2 (5), 1084-1104.
- (6) Fenech, M. and Morley, A.A. (1986), Cytokinesis-block micronucleus method in human lymphocytes: effect of *in-vivo* ageing and low dose X-irradiation. *Mutation Res.*, 161, 193-198.
- (7) Eastmond, D.A. and Tucker, J.D. (1989), Identification of aneuploidy-inducing agents using cytokinesis-blocked human lymphocytes and an antikinetochore antibody. *Environ. Mol. Mutagen.*, 13, 34-43.
- (8) Eastmond, D.A. and Pinkel, D. (1990), Detection of aneuploidy and aneuploidy-inducing agents in human lymphocytes using fluorescence *in-situ* hybridisation with chromosome-specific DNA probes. *Mutation Res.*, 234, 9-20.
- (9) Miller, B.M., Zitzelsberger, H.F., Weier, H.U. and Adler, I.D. (1991), Classification of micronuclei in murine erythrocytes: immunofluorescent staining using CREST antibodies compared to *in situ* hybridization with biotinylated gamma satellite DNA. *Mutagenesis*, 6, 297-302.
- (10) Farooqi, Z., Darroudi, F. and Natarajan, A.T. (1993), The use of fluorescence *in-situ* hybridisation for the detection of aneuploidy in cytokinesis-blocked mouse splenocytes. *Mutagenesis*, 8, 329-334.
- (11) Migliore, L., Bocciardi, R., Macri, C. and Lo Jacono, F. (1993), Cytogenetic damage induced in human lymphocytes by four vanadium compounds and micronucleus analysis by fluorescence *in situ* hybridization with a centromeric probe. *Mutation Res.*, 319, 205-213.
- (12) Norppa, H., Renzi, L. and Lindholm, C. (1993), Detection of whole chromosomes in micronuclei of cytokinesis-blocked human lymphocytes by antikinetochore staining and *in situ* hybridization. *Mutagenesis*, 8, 519-525.
- (13) Eastmond, D.A., Rupa, D.S. and Hasegawa, L.S. (1994), Detection of hyperdiploidy and chromosome breakage in interphase human lymphocytes following exposure to the benzene metabolite hydroquinone using multicolor fluorescence *in situ* hybridization with DNA probes. *Mutation Res.*, 322, 9-20.
- (14) Marshall, R.R., Murphy, M., Kirkland, D.J. and Bentley, K.S. (1996), Fluorescence *in situ* hybridisation (FISH) with chromosome-specific centromeric probes: a sensitive method to detect aneuploidy. *Mutation Res.*, 372, 233-245.
- (15) Zijno, P., Leopardi, F., Marcon, R. and Crebelli, R. (1996), Analysis of chromosome segregation by means of fluorescence *in situ* hybridization: application to cytokinesis-blocked human lymphocytes. *Mutation Res.*, 372, 211-219.
- (16) Kirsch-Volders, M., Sofuni, T., Aardema, M., Albertini, S., Eastmond, D., Fenech, M., Ishidate Jr., M., Lorge, E., Norppa, H., Surrallés, J., von der Hude, W. and Wakata, A. (2003), Report from the *in vitro* micronucleus assay working group. *Mutation Res.*, 540, 153-163.

## ▼M3

- (17) OECD (1997), *In Vitro Mammalian Chromosome Aberration Test*, Test Guideline No 473, OECD Guidelines for Testing of Chemicals, OECD, Paris. Available at: [[www.oecd.org/env/testguidelines](http://www.oecd.org/env/testguidelines)]
- (18) Lorge, E., Thybaud, V., Aardema, M.J., Oliver, J., Wakata, A., Lorenzon G. and Marzin, D. (2006), SFTG International collaborative Study on *in vitro* micronucleus test. I. General conditions and overall conclusions of the study, *Mutation Res.*, 607, 13-36.
- (19) Clare, G., Lorenzon, G., Akhurst, L.C., Marzin, D., van Delft, J., Montero, R., Botta, A., Bertens, A., Cinelli, S., Thybaud, V. and Lorge, E. (2006), SFTG International collaborative study on the *in vitro* micronucleus test. II. Using human lymphocytes, *Mutation Res.*, 607, 37-60.
- (20) Aardema, M.J., Snyder, R.D., Spicer, C., Divi, K., Morita, T., Mauthe, R.J., Gibson, D.P., Soelter, S., Curry, P.T., Thybaud, V., Lorenzon, G., Marzin, D. and Lorge, E. (2006), SFTG International collaborative study on the *in vitro* micronucleus test, III. Using CHO cells, *Mutation Res.*, 607, 61-87.
- (21) Wakata, A., Matsuoka, A., Yamakage, K., Yoshida, J., Kubo, K., Kobayashi, K., Senjyu, N., Itoh, S., Miyajima, H., Hamada, S., Nishida, S., Araki, H., Yamamura, E., Matsui, A., Thybaud, V., Lorenzon, G., Marzin, D. and Lorge, E. (2006), SFTG International collaborative study on the *in vitro* micronucleus test, IV. Using CHO/IU cells, *Mutation Res.*, 607, 88-124.
- (22) Oliver, J., Meunier, J.-R., Awogi, T., Elhajouji, A., Ouldeldhkim, M.-C., Bichet, N., Thybaud, V., Lorenzon, G., Marzin, D. and Lorge, E. (2006), SFTG International collaborative study on the *in vitro* micronucleus test, V. Using L5178Y cells, *Mutation Res.*, 607, 125-152.
- (23) Albertini, S., Miller, B., Chetelat, A.A. and Locher, F. (1997), Detailed data on *in vitro* MNT and *in vitro* CA: industrial experience, *Mutation Res.*, 392, 187-208.
- (24) Miller, B., Albertini, S., Locher, F., Thybaud, V. and Lorge, E. (1997), Comparative evaluation of the *in vitro* micronucleus test and the *in vitro* chromosome aberration test: industrial experience, *Mutation Res.*, 392, 45-59.
- (25) Miller, B., Potter-Locher, F., Seelbach, A., Stopper, H., Utesch, D. and Madle, S. (1998), Evaluation of the *in vitro* micronucleus test as an alternative to the *in vitro* chromosomal aberration assay: position of the GUM Working Group on the *in vitro* micronucleus test. Gesellschaft für Umwelt-Mutations-forschung, *Mutation Res.*, 410, 81-116.
- (26) Kalweit, S., Utesch, U., von der Hude, W. and Madle, S. (1999), Chemically induced micronucleus formation in V79 cells — comparison of three different test approaches, *Mutation Res.* 439, 183-190.
- (27) Kersten, B., Zhang, J., Brendler Schwaab, S.Y., Kasper, P. and Müller, L. (1999), The application of the micronucleus test in Chinese hamster V79 cells to detect drug-induced photogenotoxicity, *Mutation Res.* 445, 55-71.
- (28) von der Hude, W., Kalweit, S., Engelhardt, G., McKiernan, S., Kasper, P., Slacik-Erben, R., Miltenburger, H.G., Honarvar, N., Fahrig, R., Gorlitz, B., Albertini, S., Kirchner, S., Utesch, D., Potter-Locher, F., Stopper, H. and Madle, S. (2000), *In vitro* micronucleus assay with Chinese hamster V79 cells — results of a collaborative study with *in situ* exposure to 26 chemical substances, *Mutation Res.*, 468, 137-163.
- (29) Garriott, M.L., Phelps, J.B. and Hoffman, W.P. (2002), A protocol for the *in vitro* micronucleus test, I. Contributions to the development of a protocol suitable for regulatory submissions from an examination of 16 chemicals with different mechanisms of action and different levels of activity, *Mutation Res.*, 517, 123-134.
- (30) Matsushima, T., Hayashi, M., Matsuoka, A., Ishidate, M. Jr., Miura, K.F., Shimizu, H., Suzuki, Y., Morimoto, K., Ogura, H., Mure, K., Koshi, K. and Sofuni, T. (1999), Validation study of the *in vitro* micronucleus test in a Chinese hamster lung cell line (CHL/IU), *Mutagenesis*, 14, 569-580.

## ▼ M3

- (31) Elhajouji, A., and Lorge, E. (2006), Special Issue: SFTG International collaborative study on *in vitro* micronucleus test, *Mutation Res.*, 607, 1-152.
- (32) ECVAM (2006), Statement by the European Centre for the Validation of Alternative Methods (ECVAM) Scientific Advisory Committee (ESAC) on the scientific validity of the *in vitro* micronucleus test as an alternative to the *in vitro* chromosome aberration assay for genotoxicity testing. ESAC 25th meeting, 16-17 November, 2006, Available at: [<http://ecvam.jrc.it/index.htm>]
- (33) ESAC (2006), ECVAM Scientific Advisory Committee (ESAC) Peer Review, Retrospective Validation of the *In Vitro* Micronucleus Test, Summary and Conclusions of the Peer Review Panel, Available at: [<http://ecvam.jrc.it/index.htm>]
- (34) Corvi, R., Albertini, S., Hartung, T., Hoffmann, S., Maurici, D., Pfuhrer, S., van Benthem, J., Vanparys P. (2008), ECVAM Retrospective Validation of *in vitro* Micronucleus Test (MNT), *Mutagenesis*, 23, 271-283.
- (35) Zhang, L.S., Honma, M., Hayashi, M., Suzuki, T., Matsuoka, A. and Sofuni, T. (1995), A comparative study of TK6 human lymphoblastoid and L5178Y mouse lymphoma cell lines in the *in vitro* micronucleus test, *Mutation Res.*, 347, 105-115.
- (36) Ehrlich, V., Darroudi, F., Uhl, M., Steinkellner, S., Zsivkovits, M. and Knasmeuller, S. (2002), Fumonisin B<sub>1</sub> is genotoxic in human derived hepatoma (HepG2) cells, *Mutagenesis*, 17, 257-260.
- (37) Knasmüller, S., Mersch-Sundermann, V., Kevekordes, S., Darroudi, F., Huber, W.W., Hoelzl, C., Bichler, J. and Majer, B.J. (2004), Use of human-derived liver cell lines for the detection of environmental and dietary genotoxicants; current state of knowledge, *Toxicol.*, 198, 315-328.
- (38) Gibson, D.P., Brauning, R., Shaffi, H.S., Kerckaert, G.A., LeBoeuf, R.A., Isfort, R.J. and Aardema, M.J. (1997), Induction of micronuclei in Syrian hamster embryo cells: comparison to results in the SHE cell transformation assay for National Toxicology Program test chemicals, *Mutation Res.*, 392, 61-70.
- (39) Scott, D., Galloway, S.M., Marshall, R.R., Ishidate, M. Jr., Brusick, D., Ashby, J. and Myhr, B.C. (1991), International Commission for Protection Against Environmental Mutagens and Carcinogens, Genotoxicity under extreme culture conditions. A report from ICPEMC Task Group 9, *Mutation Res.*, 257, 147-205.
- (40) Morita, T., Nagaki, T., Fukuda, I. and Okumura, K. (1992), Clastogenicity of low pH to various cultured mammalian cells, *Mutation Res.*, 268, 297-305.
- (41) Brusick, D. (1986), Genotoxic effects in cultured mammalian cells produced by low pH treatment conditions and increased ion concentrations, *Environ. Mutagen.*, 8, 789-886.
- (42) Fenech, M. and Morley, A.A. (1985), Measurement of micronuclei in lymphocytes, *Mutation Res.*, 147, 29-36.
- (43) Fenech, M. (1997), The advantages and disadvantages of cytokinesis-blood micronucleus method, *Mutation Res.*, 392, 11-18.
- (44) Bonassi, S., Fenech, M., Lando, C., Lin, Y.P., Ceppi, M., Chang, W.P., Holland, N., Kirsch-Volders, M., Zeiger, E., Ban, S., Barale, R., Bigatti, M.P., Bolognesi, C., Jia, C., Di Giorgio, M., Ferguson, L.R., Fucic, A., Lima, O.G., Hrelia, P., Krishnaja, A.P., Lee, T.K., Migliore, L., Mikhalevich, L., Mirkova, E., Mosesso, P., Muller, W.U., Odagiri, Y., Scarffi, M.R., Szabova, E., Vorobtsova, I., Vral, A. and Zijno, A. (2001), HUMAN MicroNucleus Project: international database comparison for results with the cytokinesis-block micronucleus assay in human lymphocytes, I. Effect of laboratory protocol, scoring criteria and host factors on the frequency of micronuclei, *Environ. Mol. Mutagen.* 37, 31-45.

## ▼ M3

- (45) Maron, D.M. and Ames, B.N. (1983), Revised methods for the Salmonella mutagenicity test, *Mutation Res.*, 113, 173-215.
- (46) Ong, T.-m., Mukhtar, M., Wolf, C.R. and Zeiger, E. (1980), Differential effects of cytochrome P450-inducers on promutagen activation capabilities and enzymatic activities of S-9 from rat liver, *J. Environ. Pathol. Toxicol.*, 4, 55-65.
- (47) Elliott, B.M., Combes, R.D., Elcombe, C.R., Gatehouse, D.G., Gibson, G.G., Mackay, J.M. and Wolf, R.C. (1992), Alternatives to Aroclor 1254-induced S9 in *in-vitro* genotoxicity assays. *Mutagenesis*, 7, 175-177.
- (48) Matsushima, T., Sawamura, M., Hara, K. and Sugimura, T. (1976), A safe substitute for Polychlorinated Biphenyls as an Inducer of Metabolic Activation Systems, *In: de Serres, F.J., Fouts, J. R., Bend, J.R. and Philpot, R.M. (eds), In Vitro Metabolic Activation in Mutagenesis Testing*, Elsevier, North-Holland, pp. 85-88.
- (49) Johnson, T.E., Umbenhauer, D.R. and Galloway, S.M. (1996), Human liver S-9 metabolic activation: proficiency in cytogenetic assays and comparison with phenobarbital/beta-naphthoflavone or Aroclor 1254 induced rat S-9, *Environ. Mol. Mutagen.*, 28, 51-59.
- (50) UNEP (2001), Stockholm Convention on Persistent Organic Pollutants, United Nations Environment Programme (UNEP). Available at: [<http://www.pops.int/>]
- (51) Doherty, A.T., Ellard, S., Parry, E.M. and Parry, J.M. (1996), An investigation into the activation and deactivation of chlorinated hydrocarbons to genotoxins in metabolically competent human cells, *Mutagenesis*, 11, 247-274.
- (52) Krahn, D.F., Barsky, F.C. and McCooey, K.T. (1982), CHO/HGPRT Mutation Assay: Evaluation of Gases and Volatile Liquids, *In: Tice, R.R., Costa, D.L. and Schaich, K.M. (eds), Genotoxic Effects of Airborne Agents*. New York, Plenum, pp. 91-103.
- (53) Zamora, P.O., Benson, J.M., Li, A.P. and Brooks, A.L. (1983), Evaluation of an exposure system using cells grown on collagen gels for detecting highly volatile mutagens in the CHO/HGPRT mutation assay, *Environ. Mutagenesis* 5, 795-801.
- (54) Fenech, M. (1993), The cytokinesis-block micronucleus technique: a detailed description of the method and its application to genotoxicity studies in human populations, *Mutation Res.*, 285, 35-44.
- (55) Phelps, J.B., Garriott, M.L., and Hoffman, W.P. (2002), A protocol for the *in vitro* micronucleus test. II. Contributions to the validation of a protocol suitable for regulatory submissions from an examination of 10 chemicals with different mechanisms of action and different levels of activity, *Mutation Res.*, 521, 103-112.
- (56) Kirsch-Volders, M., Sofuni, T., Aardema, M., Albertini, S., Eastmond, D., Fenech, M., Ishidate, M. Jr., Kirchner, S., Lorge, E., Morita, T., Norppa, H., Surralles, J., Vanhauwaert, A. and Wakata, A. (2004), Corrigendum to 'Report from the *in vitro* micronucleus assay working group', *Mutation Res.*, 564, 97-100.
- (57) Pincu, M., Bass, D. and Norman, A. (1984), An improved micronuclear assay in lymphocytes, *Mutation Res.*, 139, 61-65.
- (58) Lorge, E., Hayashi, M., Albertini, S. and Kirkland, D. (2008), Comparison of different methods for an accurate assessment of cytotoxicity in the *in vitro* micronucleus test. I. Theoretical aspects, *Mutation Res.*, 655, 1-3.
- (59) Surralles, J., Xamena, N., Creus, A., Catalan, J., Norppa, H. and Marcos, R. (1995), Induction of micronuclei by five pyrethroid insecticides in whole-blood and isolated human lymphocyte cultures, *Mutation Res.*, 341, 169-184.
- (60) Galloway, S. (2000), Cytotoxicity and chromosome aberrations *in vitro*: Experience in industry and the case for an upper limit on toxicity in the aberration assay, *Environ. Molec. Mutagenesis* 35, 191-201.



**▼ M3**

- (61) Hayashi, M., Sofuni, T., and Ishidate, M. Jr. (1983), An Application of Acridine Orange Fluorescent Staining to the Micronucleus Test, *Mutation Res.*, 120, 241-247.
- (62) MacGregor, J. T., Wehr, C. M., and Langlois, R. G. (1983), A Simple Fluorescent Staining Procedure for Micronuclei and RNA in Erythrocytes Using Hoechst 33258 and Pyronin Y, *Mutation Res.*, 120, 269-275.
- (63) Hayashi, M., Sofuni, T. and Ishidate, M. Jr. (1983), An application of acridine orange fluorescent staining to the micronucleus test, *Mutation Res.*, 120, 241-247.
- (64) Fenech, M., Chang, W.P., Kirsch-Volders, M., Holland, N., Bonassi, S. and Zeiger, E. (2003), HUMN project: detailed description of the scoring criteria for the cytokinesis-block micronucleus assay using isolated human lymphocyte cultures, *Mutation Res.*, 534, 65-75.
- (65) Hoffman, W.P., Garriott, M.L. and Lee, C. (2003), *In vitro* micronucleus test, *In: Encyclopedia of Biopharmaceutical Statistics*, Second edition. S. Chow (ed.), Marcel Dekker, Inc. New York, NY, pp. 463-467.
- (66) Regulation (EC) No 850/2004 of the European Parliament and of the Council of 29 April 2004 on persistent organic pollutants and amending Directive 79/117/EEC, OJ L 229, 30.4.2004, p. 5.

▼ **M3***Appendix 1***Definitions**

*Aneugen*: any substance or process that, by interacting with the components of the mitotic and meiotic cell division cycle, leads to aneuploidy in cells or organisms.

*Aneuploidy*: any deviation from the normal diploid (or haploid) number of chromosomes by a single chromosome or more than one, but not by entire set(s) of chromosomes (polyploidy).

*Apoptosis*: programmed cell death characterised by a series of steps leading to a disintegration of cells into membrane-bound particles that are then eliminated by phagocytosis or by shedding.

*Cell proliferation*: increase in cell number as a result of mitotic cell division.

*Centromere*: DNA region of a chromosome where both chromatids are held together and on which both kinetochores are attached side-to-side.

*Clastogen*: any substance or process which causes structural chromosomal aberrations in populations of cells or organisms.

*Cytokinesis*: the process of cell division immediately following mitosis to form two daughter cells, each containing a single nucleus.

*Cytokinesis-Block Proliferation index (CBPI)*: the proportion of second-division cells in the treated population relative to the untreated control (see Appendix 2 for formula).

*Cytostasis*: inhibition of cell growth (see Appendix 2 for formula).

*Cytotoxicity*: harmful effects to cell structure or function ultimately causing cell death.

*Genotoxic*: a general term encompassing all types of DNA or chromosome damage, including breaks, adducts rearrangements, mutations, chromosome aberrations, and aneuploidy. Not all types of genotoxic effects result in mutations or stable chromosome damage.

*Interphase cells*: cells not in the mitotic stage.

*Kinetochores*: a protein-containing structure that assembles at the centromere of a chromosome to which spindle fibres associate during cell division, allowing orderly movement of daughter chromosomes to the poles of the daughter cells.

*Micronuclei*: small nuclei, separate from and additional to the main nuclei of cells, produced during telophase of mitosis or meiosis by lagging chromosome fragments or whole chromosomes.

*Mitosis*: division of the cell nucleus usually divided into prophase, prometaphase, metaphase, anaphase and telophase.

*Mitotic index*: the ratio of cells in metaphase divided by the total number of cells observed in a population of cells; an indication of the degree of cell proliferation of that population.

*Mutagenic*: produces a heritable change of DNA base-pair sequences(s) in genes or of the structure of chromosomes (chromosome aberrations).

**▼ M3**

*Non-disjunction*: failure of paired chromatids to disjoin and properly segregate to the developing daughter cells, resulting in daughter cells with abnormal numbers of chromosomes.

*Polyploidy*: numerical chromosome aberrations in cells or organisms involving entire set(s) of chromosomes, as opposed to an individual chromosome or chromosomes (aneuploidy).

*Proliferation Index (PI)*: method for cytotoxicity measurement when cytoB is not used (see Appendix 2 for formula).

*Relative Increase in Cell Count (RICC)*: method for cytotoxicity measurement when cytoB is not used (see Appendix 2 for formula).

*Relative Population Doubling (RPD)*: method for cytotoxicity measurement when cytoB is not used (see Appendix 2 for formula).

*Replication Index (RI)*: the proportion of cell division cycles completed in a treated culture, relative to the untreated control, during the exposure period and recovery (see Appendix 2 for formula).

*Test chemical (also referred to as test substance)*: Any substance or mixture tested using this TM.

▼ M3

## Appendix 2

## Formulas for Cytotoxicity Assessment

1. *When cytoB is used*, evaluation of cytotoxicity should be based on the Cytokinesis-Block Proliferation Index (CBPI) or Replicative Index (RI) (16) (58). The CBPI indicates the average number of cell cycles per cell during the period of exposure to cytoB, and may be used to calculate cell proliferation. The RI indicates the relative number of nuclei in treated cultures compared to control cultures and can be used to calculate the % cytostasis:

$$\% \text{ Cytostasis} = 100 - 100\{(\text{CBPI}_T - 1) \div (\text{CBPI}_C - 1)\}$$

And:

T = test chemical treatment culture

C = vehicle control culture

Where:

$$\text{CBPI} = \frac{((\text{No. mononucleate cells}) + (2 \times \text{No. binucleate cells}) + (3 \times \text{No. multinucleate cells}))}{(\text{Total number of cells})}$$

Thus, a CBPI of 1 (all cells are mononucleate) is equivalent to 100 % cytostasis.

$$\text{Cytostasis} = 100 - \text{RI}$$

$$\text{RI} = \frac{((\text{No. binucleated cells}) + (2 \times \text{No. multinucleate cells})) \div (\text{Total number of cells})_T}{((\text{No. binucleated cells}) + (2 \times \text{No. multinucleate cells})) \div (\text{Total number of cells})_C} \times 100$$

T = treated cultures

C = control cultures

2. Thus, an RI of 53 % means that, compared to the numbers of cells that have divided to form binucleate and multinucleate cells in the control culture, only 53 % of this number divided in the treated culture, *i.e.* 47 % cytostasis.

3. *When cytoB is not used*, evaluation of cytotoxicity based on Relative Increase in Cell Counts (RICC) or on Relative Population Doubling (RPD) is recommended (58), as both take into account the proportion of the cell population which has divided.

$$\text{RICC} = \frac{(\text{Increase in number of cells in treated cultures (final - starting)})}{(\text{Increase in number of cells in control cultures (final - starting)})} \times 100$$

$$\text{RPD} = \frac{(\text{No. of Population doublings in treated cultures})}{(\text{No. of Population doublings in control cultures})} \times 100$$

**▼ M3**

where:

Population Doubling =  $[\log (\text{Post-treatment cell number} \div \text{Initial cell number})] \div \log 2$

4. Thus, a RICC, or a RPD of 53 % indicates 47 % cytotoxicity/cytostasis.
5. By using a Proliferation Index (PI), cytotoxicity may be assessed via counting the number of clones consisting of 1 cell (c1), 2 cells (c2), 3 to 4 cells (c4) and 5 to 8 cells (c8)

$$\text{PI} = \frac{((1 \times \text{c1}) + (2 \times \text{c2}) + (3 \times \text{c4}) + (4 \times \text{c8}))}{(\text{c1} + \text{c2} + \text{c4} + \text{c8})}$$

6. The PI has been used as a valuable and reliable cytotoxicity parameter also for cell lines cultured *in situ* in the absence of cytoB (25) (26) (27) (28).

▼ **M3***Appendix 3***Reference chemicals recommended for assessing performance <sup>(1)</sup>**

Category	Chemical	CAS No	EC No
1. Clastogens active without metabolic activation			
	Cytosine arabinoside	147-94-4	205-705-9
	Mitomycin C	50-07-7	200-008-6
2. Clastogens requiring metabolic activation			
	Benzo(a)pyrene	50-32-8	200-028-5
	Cyclophosphamide	50-18-0	200-015-4
3. Aneugens			
	Colchicine	64-86-8	200-598-5
	Vinblastine	143-67-9	205-606-0
4. Negative substances			
	Di(2-ethylhexyl)phthalate	117-81-7	204-211-0
	Nalidixic acid	389-08-2	206-864-7
	Pyrene	129-00-0	204-927-3
	Sodium chloride	7647-14-5	231-598-3

<sup>(1)</sup> The reference chemicals are the recommended chemicals for use. Substitution or adding of chemicals to the list of reference chemicals can be done if their activity is known and if they induce micronuclei by the same mechanisms of action, and if they are shown to be relevant to the chemicals that will be tested using the MNvit procedure. Depending on the purpose, justification could also include a validation study employing a broad variety of substances or focused on a narrower spectrum based on the chemical class of the test substance or the mechanism of damage being studied.

**▼ M3****B.50. SKIN SENSITISATION: LOCAL LYMPH NODE ASSAY: DA**

## INTRODUCTION

1. OECD Guidelines for the Testing of Chemicals and EU Test Methods are periodically reviewed in light of scientific progress, changing regulatory needs, and animal welfare considerations. The first Test Method (TM) (B.42) for the determination of skin sensitisation in the mouse, the Local Lymph Node Assay (LLNA; OECD Test Guideline 429) has been revised (1) The details of the validation of the LLNA and a review of the associated work have been published (2) (3) (4) (5) (6) (7) (8) (9). In the LLNA, radioisotopic thymidine or iodine is used to measure lymphocyte proliferation and therefore the assay has limited use where the acquisition, use, or disposal of radioactivity is problematic. The LLNA: DA (developed by Daicel Chemical Industries, Ltd) is a non-radioactive modification to the LLNA, which quantifies adenosine triphosphate (ATP) content via bio-luminescence as an indicator of lymphocyte proliferation. The LLNA: DA test method has been validated and reviewed and recommended by an international peer review panel as considered useful for identifying skin sensitising and non-sensitising chemicals, with certain limitations (10) (11) (12) (13). This TM is designed for assessing skin sensitisation potential of chemicals (substances and mixtures) in animals. Chapter B.6 of this Annex and OECD Test Guideline 406 utilise guinea pig tests, notably the guinea pig maximisation test and the Buehler test (14) The LLNA (Chapter B.42 of this Annex; OECD Test Guideline 429) and the two non-radioactive modifications, LLNA: DA (Chapter B.50 of this Annex; OECD Test Guideline 442 A) and LLNA: BrdU-ELISA (Chapter B.51 of this Annex; OECD Test Guideline 442 B), all provide an advantage over the guinea pig tests in B.6 and OECD Test Guideline 406 (14) in terms of reduction and refinement of animal use.
2. Similar to the LLNA, the LLNA: DA studies the induction phase of skin sensitisation and provides quantitative data suitable for dose-response assessment. Furthermore, an ability to detect skin sensitisers without the necessity for using a radiolabel for DNA eliminates the potential for occupational exposure to radioactivity and waste disposal issues. This in turn may allow for the increased use of mice to detect skin sensitisers, which could further reduce the use of guinea pigs to test for skin sensitisation potential (*i.e.* B.6; OECD Test Guideline 406) (14).

## DEFINITIONS

3. Definitions used are provided in Appendix 1.

## INITIAL CONSIDERATIONS AND LIMITATIONS

4. The LLNA: DA is a modified LLNA method for identifying potential skin sensitising chemicals, with specific limitations. This does not necessarily imply that in all instances the LLNA: DA should be used in place of the LLNA or guinea pig tests (*i.e.* B.6; OECD Test Guideline 406) (14), but rather that the assay is of equal merit and may be employed as an alternative in which positive and negative results generally no longer require further confirmation (10) (11). The testing laboratory should consider all available information on the test substance prior to conducting the study. Such information will include the identity and chemical structure of the test

**▼ M3**

substance; its physicochemical properties; the results of any other *in vitro* or *in vivo* toxicity tests on the test substance; and toxicological data on structurally related chemicals. This information should be considered in order to determine whether the LLNA: DA is appropriate for the test substance (given the incompatibility of limited types of chemicals with the LLNA: DA (see paragraph 5) and to aid in dose selection.

5. The LLNA: DA is an *in vivo* method and, as a consequence, will not eliminate the use of animals in the assessment of allergic contact sensitising activity. It has, however, the potential to reduce animal use for this purpose when compared to the guinea pig tests (B.6; OECD Test Guideline 406) (14). Moreover, the LLNA: DA offers a substantial refinement (less pain and distress) of the way in which animals are used for allergic contact sensitisation testing, since unlike the B.6 and OECD Test Guideline 406, the LLNA: DA does not require that challenge-induced dermal hypersensitivity reactions be elicited. Despite the advantages of the LLNA: DA over B.6 and OECD Test Guideline 406 (14), there are certain limitations that may necessitate the use of B.6 or OECD Test Guideline 406 (e.g. the testing of certain metals, false positive findings with certain skin irritants (such as some surfactant-type substances) (6) (1 and Chapter B.42 in this Annex), solubility of the test substance). In addition, chemical classes or substances containing functional groups shown to act as potential confounders (16) may necessitate the use of guinea pig tests (i.e. B.6; OECD Test Guideline 406 (14)). Limitations that have been identified for the LLNA (1 and Chapter B.42 in this Annex) have been recommended to apply also to the LLNA: DA (10). Additionally, the use of the LLNA: DA might not be appropriate for testing substances that affect ATP levels (e.g. substances that function as ATP inhibitors) or those that affect the accurate measurement of intracellular ATP (e.g. presence of ATP degrading enzymes, presence of extracellular ATP in the lymph node). Other than such identified limitations, the LLNA: DA should be applicable for testing any substances unless there are properties associated with these substances that may interfere with the accuracy of the LLNA: DA. In addition, consideration should be given to the possibility of borderline positive results when Stimulation Index (SI) values between 1,8 and 2,5 are obtained (see paragraphs 31-32). This is based on the validation database of 44 substances using an  $SI \geq 1,8$  (see paragraph 6) for which the LLNA: DA correctly identified all 32 LLNA sensitisers, but incorrectly identified three of 12 LLNA non-sensitisers with SI values between 1,8 and 2,5 (i.e. borderline positive) (10). However, as the same dataset was used for setting the SI-values and calculating the predictive properties of the test, the stated results may be an over-estimation of the real predictive properties.

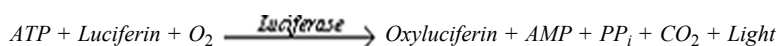
**PRINCIPLE OF THE TEST METHOD**

6. The basic principle underlying the LLNA: DA is that sensitisers induce proliferation of lymphocytes in the lymph nodes draining the site of test substance application. This proliferation is proportional to the dose and to the potency of the applied allergen and provides a simple means of obtaining a quantitative measurement of sensitisation. Proliferation is measured by comparing the mean proliferation in each test group to the mean proliferation in the vehicle treated control (VC) group. The ratio of the mean proliferation in each treated group to that in the concurrent VC group, termed the SI, is determined, and should be  $\geq 1,8$  before further



**▼ M3**

evaluation of the test substance as a potential skin sensitiser is warranted. The procedures described here are based on the use of measuring ATP content by bioluminescence (known to correlate with living cell number) (17) to indicate an increased number of proliferating cells in the draining auricular lymph nodes (18) (19). The bioluminescent method utilises the luciferase enzyme to catalyse the formation of light from ATP and luciferin according to the following reaction:



The emitted light intensity is linearly related to the ATP concentration and is measured using a luminometer. The luciferin-luciferase assay is a sensitive method for ATP quantitation used in a wide variety of applications (20).

**DESCRIPTION OF THE ASSAY****Selection of animal species**

- The mouse is the species of choice for this test. Validation studies for the LLNA: DA were conducted exclusively with the CBA/J strain, which is therefore considered the preferred strain (12) (13). Young adult female mice, which are nulliparous and non-pregnant, are used. At the start of the study, animals should be between 8-12 weeks old, and the weight variation of the animals should be minimal and not exceed 20 % of the mean weight. Alternatively, other strains and males may be used when sufficient data are generated to demonstrate that significant strain and/or gender-specific differences in the LLNA: DA response do not exist.

**Housing and feeding conditions**

- Mice should be group-housed (21), unless adequate scientific rationale for housing mice individually is provided. The temperature of the experimental animal room should be  $22 \pm 3$  °C. Although the relative humidity should be at least 30 % and preferably not exceed 70 %, other than during room cleaning, the aim should be 50-60 %. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water.

**Preparation of animals**

- The animals are randomly selected, marked to permit individual identification (but not by any form of ear marking), and kept in their cages for at least five days prior to the start of dosing to allow for acclimatisation to the laboratory conditions. Prior to the start of treatment all animals are examined to ensure that they have no observable skin lesions.

**Preparation of dosing solutions**

- Solid chemicals should be dissolved or suspended in solvents/vehicles and diluted, if appropriate, prior to application to an ear of the mice. Liquid chemicals may be applied neat or diluted prior to dosing. Insoluble chemicals, such as those generally seen in medical devices, should be subjected to an exaggerated extraction in an appropriate solvent to reveal all extractable constituents for testing prior to application to an ear of the mice. Test substances should be prepared daily unless stability data demonstrate the acceptability of storage.

**▼ M3****Reliability check**

11. Positive control chemicals (PC) are used to demonstrate appropriate performance of the assay by responding with adequate and reproducible sensitivity to a sensitising test substance for which the magnitude of the response is well characterised. Inclusion of a concurrent PC is recommended because it demonstrates competency of the laboratory to successfully conduct each assay and allows for an assessment of intra- and inter-laboratory reproducibility and comparability. Some regulatory authorities also require a PC for each study and therefore users are encouraged to consult the relevant authorities prior to conducting the LLNA: DA. Accordingly, the routine use of a concurrent PC is encouraged to avoid the need for additional animal testing to meet such requirements that might arise from the use of a periodic PC (see paragraph 12). The PC should produce a positive LLNA: DA response at an exposure level expected to give an increase in the SI  $\geq 1,8$  over the negative control (NC) group. The PC dose should be chosen such that it does not cause excessive skin irritation or systemic toxicity and the induction is reproducible but not excessive (e.g. SI  $> 10$  would be considered excessive). Preferred PC are 25 % hexyl cinnamic aldehyde (Chemical Abstracts Service (CAS) number 101-86-0) and 25 % eugenol (CAS number 97-53-0) in acetone: olive oil (4:1, v/v). There may be circumstances in which, given adequate justification, other PC, meeting the above criteria, may be used.
  
12. While inclusion of a concurrent PC group is recommended, there may be situations in which periodic testing (i.e. at intervals  $\leq 6$  months) of the PC may be adequate for laboratories that conduct the LLNA: DA regularly (i.e. conduct the LLNA: DA at a frequency of no less than once per month) and have an established historical PC database that demonstrates the laboratory's ability to obtain reproducible and accurate results with PCs. Adequate proficiency with the LLNA: DA can be successfully demonstrated by generating consistent positive results with the PC in at least 10 independent tests conducted within a reasonable period of time (i.e. less than one year).
  
13. A concurrent PC group should always be included when there is a procedural change to the LLNA: DA (e.g. change in trained personnel, change in test method materials and/or reagents, change in test method equipment, change in source of test animals), and such changes should be documented in laboratory reports. Consideration should be given to the impact of these changes on the adequacy of the previously established historical database in determining the necessity for establishing a new historical database to document consistency in the PC results.
  
14. Investigators should be aware that the decision to conduct a PC study on a periodic basis instead of concurrently has ramifications on the adequacy and acceptability of negative study results generated without a concurrent PC during the interval between each periodic PC study. For example, if a false negative result is obtained in the periodic PC study, negative test substance results obtained in the interval between the last acceptable periodic PC study and the unacceptable periodic PC study may be questioned. Implications of these outcomes should be carefully considered when determining whether to include concurrent PCs or to only conduct periodic PCs. Consideration should also be given to using fewer animals in the concurrent PC group when this is scientifically justified and if the laboratory demonstrates, based on laboratory-specific historical data, that fewer mice can be used (22).

**▼ M3**

15. Although the PC should be tested in the vehicle that is known to elicit a consistent response (e.g. acetone: olive oil; 4:1, v/v), there may be certain regulatory situations in which testing in a non-standard vehicle (clinically/chemically relevant formulation) will also be necessary (23). If the concurrent PC is tested in a different vehicle than the test substance, then a separate VC for the concurrent PC should be included.
  
16. In instances where substances of a specific chemical class or range of responses are being evaluated, benchmark substances may also be useful to demonstrate that the test method is functioning properly for detecting the skin sensitisation potential of these types of substances. Appropriate benchmark substances should have the following properties:
  - structural and functional similarity to the class of the test substance being tested;
  
  - known physical chemical characteristics;
  
  - supporting data from the LLNA: DA;
  
  - supporting data from other animal models and/or from humans.

**TEST PROCEDURE****Number of animals and dose levels**

17. A minimum of four animals is used per dose group, with a minimum of three concentrations of the test substance, plus a concurrent NC group treated only with the vehicle for the test substance, and a PC (concurrent or recent, based on laboratory policy in considering paragraphs 11-15). Testing multiple doses of the PC should be considered, especially when testing the PC on an intermittent basis. Except for absence of treatment with the test substance, animals in the control groups should be handled and treated in a manner identical to that of animals in the treatment groups.
  
18. Dose and vehicle selection should be based on the recommendations given in references (2) and (24). Consecutive doses are normally selected from an appropriate concentration series such as 100 %, 50 %, 25 %, 10 %, 5 %, 2,5 %, 1 %, 0,5 %, etc. Adequate scientific rationale should accompany the selection of the concentration series used. All existing toxicological information (e.g. acute toxicity and dermal irritation) and structural and physicochemical information on the test substance of interest (and/or structurally related substances) should be considered, where available, in selecting the three consecutive concentrations so that the highest concentration maximises exposure while avoiding systemic toxicity and/or excessive local skin irritation (24) (25). In the absence of such information, an initial pre-screen test may be necessary (see paragraphs 21-24).
  
19. The vehicle should not interfere with or bias the test result and should be selected on the basis of maximising the solubility in order to obtain the highest concentration achievable while producing a solution/suspension suitable for application of the test substance. Recommended vehicles are acetone: olive oil (4:1 v/v), N,N-dimethylformamide, methyl ethyl ketone, propylene glycol, and dimethyl sulphoxide (6) but others may be used if sufficient scientific rationale is provided. In certain situations it may be necessary to use a clinically relevant solvent or the commercial formulation in which the test substance is marketed as an additional control. Particular care should be taken to ensure that hydrophilic substances are incorporated into a vehicle system, which wets the skin and does not immediately run off, by incorporation of appropriate solubilisers (e.g. 1 % Pluronic® L92). Thus, wholly aqueous vehicles are to be avoided.

**▼ M3**

20. The processing of lymph nodes from individual mice allows for the assessment of inter-animal variability and a statistical comparison of the difference between test substance and VC group measurements (see paragraph 33). In addition, evaluating the possibility of reducing the number of mice in the PC group is only feasible when individual animal data are collected (22). Further, some regulatory authorities require the collection of individual animal data. Regular collection of individual animal data provides an animal welfare advantage by avoiding duplicate testing that would be necessary if the test substance results originally collected in one manner (e.g. via pooled animal data) were to be considered later by regulatory authorities with other requirements (e.g. individual animal data).

**Pre-screen test**

21. In the absence of information to determine the highest dose to be tested (see paragraph 18), a pre-screen test should be performed in order to define the appropriate dose level to test in the LLNA: DA. The purpose of the pre-screen test is to provide guidance for selecting the maximum dose level to use in the main LLNA: DA study, where information on the concentration that induces systemic toxicity (see paragraph 24) and/or excessive local skin irritation (see paragraph 23) is not available. The maximum dose level tested should be 100 % of the test substance for liquids or the maximum possible concentration for solids or suspensions.
22. The pre-screen test is conducted under conditions identical to the main LLNA: DA study, except there is no assessment of lymph node proliferation and fewer animals per dose group can be used. One or two animals per dose group are suggested. All mice will be observed daily for any clinical signs of systemic toxicity or local irritation at the application site. Body weights are recorded pre-test and prior to termination (Day 8). Both ears of each mouse are observed for erythema and scored using Table 1 (25). Ear thickness measurements are taken using a thickness gauge (e.g. digital micrometer or Peacock Dial thickness gauge) on Day 1 (pre-dose), Day 3 (approximately 48 hours after the first dose), Day 7 (24 hours prior to termination) and Day 8. Additionally on Day 8, ear thickness could be determined by ear punch weight determinations, which should be performed after the animals are humanely killed. Excessive local irritation is indicated by an erythema score  $\geq 3$  and/or ear thickness of  $\geq 25$  % on any day of measurement (26) (27). The highest dose selected for the main LLNA: DA study will be the next lower dose in the pre-screen concentration series (see paragraph 18) that does not induce systemic toxicity and/or excessive local skin irritation

Table 1

**Erythema Scores**

Observation	Score
No erythema	0
Very slight erythema (barely perceptible)	1
Well-defined erythema	2

▼ **M3**

Observation	Score
Moderate to severe erythema	3
Severe erythema (beet redness) to eschar formation preventing grading of erythema	4

23. In addition to a 25 % increase in ear thickness (26) (27), a statistically significant increase in ear thickness in the treated mice compared to control mice has also been used to identify irritants in the LLNA (28) (29) (30) (31) (32) (33) (34). However, while statistically significant increases can occur when ear thickness is less than 25 % they have not been associated specifically with excessive irritation (30) (31) (32) (33) (34).
24. The following clinical observations may indicate systemic toxicity (35) when used as part of an integrated assessment and therefore may indicate the maximum dose level to use in the main LLNA: DA: changes in nervous system function (*e.g.* pilo-erection, ataxia, tremors, and convulsions); changes in behaviour (*e.g.* aggressiveness, change in grooming activity, marked change in activity level); changes in respiratory patterns (*i.e.* changes in frequency and intensity of breathing such as dyspnea, gasping, and rales), and changes in food and water consumption. In addition, signs of lethargy and/or unresponsiveness and any clinical signs of more than slight or momentary pain and distress, or a > 5 % reduction in body weight from Day 1 to Day 8 and mortality, should be considered in the evaluation. Moribund animals or animals showing signs of severe pain and distress should be humanely killed (36).

**Main study experimental schedule**

25. The experimental schedule of the assay is as follows:
- *Day 1:* Individually identify and record the weight of each animal and any clinical observation. Apply 1 % sodium lauryl sulfate (SLS) aqueous solution to the dorsum of each ear by using a brush dipped in the SLS solution to cover the entire dorsum of each ear with four to five strokes. One hour after the SLS treatment, apply 25 µL of the appropriate dilution of the test substance, the vehicle alone, or the PC (concurrent or recent, based on laboratory policy in considering paragraphs 11-15), to the dorsum of each ear.
  - *Days 2, 3 and 7:* Repeat the 1 % SLS aqueous solution pre-treatment and test substance application procedure carried out on Day 1.
  - *Days 4, 5, and 6:* No treatment.
  - *Day 8:* Record the weight of each animal and any clinical observation. Approximately 24 to 30 hours after the start of application on Day 7, humanely kill the animals. Excise the draining auricular lymph nodes from each mouse ear and process separately in phosphate buffered saline (PBS) for each animal. Details and diagrams of the lymph node identification and dissection can be found in reference (22). To further monitor the local skin response in the main study, additional parameters such as scoring of ear erythema or ear thickness measurements (obtained either by using a thickness gauge, or ear punch weight determinations at necropsy) may be included in the study protocol.

**▼ M3****Preparation of cell suspensions**

26. From each mouse, a single-cell suspension of lymph node cells (LNC) excised bilaterally is prepared by sandwiching the lymph nodes between two glass slides and applying light pressure to crush the nodes. After confirming that the tissue has spread out thinly pull the two slides apart. Suspend the tissue on both slides in PBS by holding each slide at an angle over the Petri dish and rinsing with PBS while concurrently scraping the tissue off of the slide with a cell scraper. Further, the lymph nodes in NC animals are small, so careful operation is important to avoid any artificial effects on SI values. A total volume of 1 mL PBS should be used for rinsing both slides. The LNC suspension in the Petri dish should be homogenised lightly with the cell scraper. A 20 µL aliquot of the LNC suspension is then collected with a micropipette, taking care not to take up the membrane that is visible to the eye, and subsequently mixed with 1,98 mL of PBS to yield a 2 mL sample. A second 2 mL sample is then prepared using the same procedure so that two samples are prepared for each animal.

**Determination of cellular proliferation (measurement of ATP content of lymphocytes)**

27. Increases in ATP content in the lymph nodes are measured by the luciferin/luciferase method using an ATP measurement kit, which measures bioluminescence in Relative Luminescence Units (RLU). The assay time from time of animal sacrifice to measurement of ATP content for each individual animal should be kept uniform, within approximately 30 minutes, because the ATP content is considered to gradually decrease with time after animal sacrifice (12). Thus, the series of procedures from excision of auricular lymph nodes to ATP measurement should be completed within 20 minutes by the pre-determined time schedule that is the same for each animal. ATP luminescence should be measured in each 2 mL sample so that a total of two ATP measurements are collected for each animal. The mean ATP luminescence is then determined and used in subsequent calculations (see paragraph 30).

**OBSERVATIONS****Clinical observations**

28. Each mouse should be carefully observed at least once daily for any clinical signs, either of local irritation at the application site or of systemic toxicity. All observations are systematically recorded with records being maintained for each mouse. Monitoring plans should include criteria to promptly identify those mice exhibiting systemic toxicity, excessive local skin irritation, or corrosion of skin for euthanasia (36).

**Body weights**

29. As stated in paragraph 25, individual animal body weights should be measured at the start of the test and at the scheduled humane kill.

**CALCULATION OF RESULTS**

30. Results for each treatment group are expressed as the mean SI. The SI is derived by dividing the mean RLU/mouse within each test substance group and the PC group by the mean RLU/mouse for the solvent/VC group. The average SI for the VCs is then one.

**▼ M3**

31. The decision process regards a result as positive when  $SI \geq 1,8$  (10). However, the strength of the dose-response relationship, the statistical significance and the consistency of the solvent/vehicle and PC responses may also be used when determining whether a borderline result (*i.e.* SI value between 1,8 and 2,5) is declared positive (2) (3) (37).
32. For a borderline positive response between an SI of 1,8 and 2,5, users may want to consider additional information such as dose-response relationship, evidence of systemic toxicity or excessive irritation, and where appropriate, statistical significance together with SI values to confirm that such results are positives (10). Consideration should also be given to various properties of the test substance, including whether it has a structural relationship to known skin sensitisers, whether it causes excessive skin irritation in the mouse, and the nature of the dose-response relationship observed. These and other considerations are discussed in detail elsewhere (4).
33. Collecting data at the level of the individual mouse will enable a statistical analysis for presence and degree of dose-response relationship in the data. Any statistical assessment could include an evaluation of the dose-response relationship as well as suitably adjusted comparisons of test groups (*e.g.* pair-wise dosed group versus concurrent solvent/vehicle control comparisons). Statistical analyses may include, *e.g.* linear regression or William's test to assess dose-response trends, and Dunnett's test for pair-wise comparisons. In choosing an appropriate method of statistical analysis, the investigator should maintain an awareness of possible inequalities of variances and other related problems that may necessitate a data transformation or a non-parametric statistical analysis. In any case, the investigator may need to carry out SI calculations and statistical analyses with and without certain data points (sometimes called 'outliers').

**DATA AND REPORTING****Data**

34. Data should be summarised in tabular form showing the individual animal RLU values, the group mean RLU/animal, its associated error term (*e.g.* SD, SEM), and the mean SI for each dose group compared against the concurrent solvent/vehicle control group.

**Test report**

35. The test report should contain the following information:

Test and control chemicals:

- identification data (*e.g.* CAS number and EC number, if available; source; purity; known impurities; lot number);
- physical nature and physicochemical properties (*e.g.* volatility, stability, solubility);
- if mixture, composition and relative percentages of components;

Solvent/vehicle:

- identification data (purity; concentration, where appropriate; volume used);
- justification for choice of vehicle;

**▼ M3**

## Test animals:

- source of CBA mice;
- microbiological status of the animals, when known;
- number and age of animals;
- source of animals, housing conditions, diet, etc.;

## Test conditions:

- the source, lot number and manufacturer's quality assurance/quality control data for the ATP kit;
- details of test substance preparation and application;
- justification for dose selection (including results from pre-screen test, if conducted);
- vehicle and test substance concentrations used, and total amount of test substance applied;
- details of food and water quality (including diet type/source, water source);
- details of treatment and sampling schedules;
- methods for measurement of toxicity;
- criteria for considering studies as positive or negative;
- details of any protocol deviations and an explanation on how the deviation affects the study design and results;

## Reliability check:

- a summary of results of latest reliability check, including information on test substance, concentration and vehicle used;
- concurrent and/or historical PC and concurrent negative (solvent/vehicle) control data for testing laboratory;
- if a concurrent PC was not included, the date and laboratory report for the most recent periodic PC and a report detailing the historical PC data for the laboratory justifying the basis for not conducting a concurrent PC;

## Results:

- individual weights of mice at start of dosing and at scheduled kill; as well as mean and associated error term (*e.g.* SD, SEM) for each treatment group;
- time course of onset and signs of toxicity, including dermal irritation at site of administration, if any, for each animal;
- time of animal termination and time of ATP measurement for each animal;
- a table of individual mouse RLU values and SI values for each dose treatment group;
- mean and associated error term (*e.g.* SD, SEM) for RLU/mouse for each treatment group and the results of outlier analysis for each treatment group;



▼ **M3**

- calculated SI and an appropriate measure of variability that takes into account the inter-animal variability in both the test substance and control groups;
- dose response relationship;
- statistical analyses, where appropriate;

Discussion of results:

- a brief commentary on the results, the dose-response analysis, and statistical analyses, where appropriate, with a conclusion as to whether the test substance should be considered a skin sensitiser.

## LITERATURE

- (1) OECD (2010), *Skin Sensitisation: Local Lymph Node Assay*, Test Guideline No 429, Guidelines for the Testing of Chemicals, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (2) Chamberlain, M. and Basketter, D.A. (1996), The local lymph node assay: status of validation. *Food Chem, Toxicol.*, 34, 999-1002.
- (3) Basketter, D.A., Gerberick, G.F., Kimber, I. and Loveless, S.E. (1996), The local lymph node assay: A viable alternative to currently accepted skin sensitisation tests. *Food Chem, Toxicol.*, 34, 985-997.
- (4) Basketter, D.A., Gerberick, G.F. and Kimber, I. (1998), Strategies for identifying false positive responses in predictive sensitisation tests. *Food Chem. Toxicol.*, 36, 327-333.
- (5) Van Och, F.M.M., Slob, W., De Jong, W.H., Vandebriel, R.J. and Van Loveren, H. (2000), A quantitative method for assessing the sensitising potency of low molecular weight chemicals using a local lymph node assay: employment of a regression method that includes determination of uncertainty margins. *Toxicol.*, 146, 49-59.
- (6) ICCVAM (1999), The murine local lymph node Assay: A test method for assessing the allergic contact dermatitis potential of chemicals/compounds: The results of an independent peer review evaluation coordinated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program Center for the Evaluation of Alternative Toxicological Methods (NICETAM). NIH Publication No: 99-4494. Research Triangle Park, N.C. Available at: [[http://iccvam.niehs.nih.gov/docs/immunotox\\_docs/llna/llnarep.pdf](http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/llnarep.pdf)]
- (7) Dean, J.H., Twerdok, L.E., Tice, R.R., Sailstad, D.M., Hattan, D.G., Stokes, W.S. (2001), ICCVAM evaluation of the murine local lymph node assay: II. Conclusions and recommendations of an independent scientific peer review panel. *Reg. Toxicol. Pharmacol.* 34, 258-273.
- (8) Haneke, K.E., Tice, R.R., Carson, B.L., Margolin, B.H., Stokes, W.S. (2001), ICCVAM evaluation of the murine local lymph node assay: III. Data analyses completed by the national toxicology program interagency center for the evaluation of alternative toxicological methods. *Reg. Toxicol. Pharmacol.* 34, 274-286.
- (9) Sailstad, D.M., Hattan, D., Hill, R.N., Stokes, W.S. (2001), ICCVAM evaluation of the murine local lymph node assay: I. The ICCVAM review process. *Reg. Toxicol. Pharmacol.* 34, 249-257.
- (10) ICCVAM (2010), ICCVAM Test Method Evaluation Report. Nonradioactive local lymph node assay: modified by Daicel Chemical Industries, Ltd, based on ATP content test method protocol (LLNA: DA). NIH Publication No 10-7551A/B. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: [<http://iccvam.niehs.nih.gov/methods/immunotox/llna-DA/TMER.htm>]

▼ **M3**

- (11) ICCVAM (2009), Independent Scientific Peer Review Panel Report: Updated validation status of new versions and applications of the murine local lymph node assay: a test method for assessing the allergic contact dermatitis potential of chemicals and products. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: [[http://iccvam.niehs.nih.gov/docs/immunotox\\_docs/LLNAPRPRpt2009.pdf](http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRpt2009.pdf)].
- (12) Idehara, K., Yamagishi, G., Yamashita, K. and Ito, M. (2008), Characterization and evaluation of a modified local lymph node assay using ATP content as a non-radio isotopic endpoint. *J. Pharmacol. Toxicol. Meth.*, 58, 1-10.
- (13) Omori, T., Idehara, K., Kojima, H., Sozu, T., Arima, K., Goto, H., Hanada, T., Ikarashi, Y., Inoda, T., Kanazawa, Y., Kosaka, T., Maki, E., Morimoto, T., Shinoda, S., Shinoda, N., Takeyoshi, M., Tanaka, M., Uratani, M., Usami, M., Yamanaka, A., Yoneda, T., Yoshimura, I. and Yuasa, A. (2008), Interlaboratory validation of the modified murine local lymph node assay based on adenosine triphosphate measurement. *J. Pharmacol. Toxicol. Meth.*, 58, 11-26.
- (14) OECD (1992), *Skin Sensitisation*, Test Guideline No 406, Guidelines for Testing of Chemicals, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (15) Kreiling, R., Hollnagel, H.M., Hareng, L., Eigler, L., Lee, M.S., Griem, P., Dreesen, B., Kleber, M., Albrecht, A., Garcia, C. and Wendel, A. (2008), Comparison of the skin sensitising potential of unsaturated compounds as assessed by the murine local lymph node assay (LLNA) and the guinea pig maximization test (GPMT). *Food Chem. Toxicol.*, 46, 1896-1904.
- (16) Basketter, D., Ball, N., Cagen, S., Carrillo, J.C., Certa, H., Eigler, D., Garcia, C., Esch, H., Graham, C., Haux, C., Kreiling, R. and Mehling, A. (2009), Application of a weight of evidence approach to assessing discordant sensitisation datasets: implications for REACH. *Reg. Toxicol. Pharmacol.*, 55, 90-96.
- (17) Crouch, S.P., Kozlowski, R., Slater, K.J. and Fletcher J. (1993), The use of ATP bioluminescence as a measure of cell proliferation and cytotoxicity. *J. Immunol. Meth.*, 160, 81-88.
- (18) Ishizaka, A., Tono-oka, T. and Matsumoto, S. (1984), Evaluation of the proliferative response of lymphocytes by measurement of intracellular ATP. *J. Immunol. Meth.*, 72, 127-132.
- (19) Dexter, S.J., Cámara, M., Davies, M. and Shakesheff, K.M. (2003), Development of a bioluminescent ATP assay to quantify mammalian and bacterial cell number from a mixed population. *Biomat.*, 24, 27-34.
- (20) Lundin A. (2000), Use of firefly luciferase in ATP-related assays of biomass, enzymes, and metabolites. *Meth. Enzymol.*, 305, 346-370.
- (21) ILAR (1996), Institute of Laboratory Animal Research (ILAR) Guide for the Care and Use of Laboratory Animals. 7th ed. Washington, DC: National Academies Press.
- (22) ICCVAM (2009), Recommended Performance Standards: Murine Local Lymph Node Assay, NIH Publication Number 09-7357, Research Triangle Park, NC: National Institute of Environmental Health Science. Available at: [[http://iccvam.niehs.nih.gov/docs/immunotox\\_docs/llnaps/LLNAPerfStds.pdf](http://iccvam.niehs.nih.gov/docs/immunotox_docs/llnaps/LLNAPerfStds.pdf)]
- (23) McGarry, H.F. (2007), The murine local lymph node assay: regulatory and potency considerations under REACH. *Toxicol.*, 238, 71-89.
- (24) Kimber, I., Dearman, R.J., Scholes E.W. and Basketter, D.A. (1994), The local lymph node assay: developments and applications. *Toxicol.*, 93, 13-31.
- (25) OECD (2002), *Acute Dermal Irritation/Corrosion*, Test Guideline No 404, Guidelines for Testing of Chemicals, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]

## ▼ M3

- (26) Reeder, M.K., Broomhead, Y.L., DiDonato, L. and DeGeorge, G.L. (2007), Use of an enhanced local lymph node assay to correctly classify irritants and false positive substances. *Toxicologist*, 96, 235.
- (27) ICCVAM (2009), Nonradioactive Murine Local Lymph Node Assay: Flow Cytometry Test Method Protocol (LLNA: BrdU-FC) Revised Draft Background Review Document. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: [<http://iccvam.niehs.nih.gov/methods/immunotox/fcLLNA/BRDcomplete.pdf>].
- (28) Hayes, B.B., Gerber, P.C., Griffey, S.S. and Meade, B.J. (1998), Contact hypersensitivity to dicyclohexylcarbodiimide and diisopropylcarbodiimide in female B6C3F1 mice. *Drug. Chem. Toxicol.*, 21, 195-206.
- (29) Homey, B., von Schilling, C., Blumel, J., Schuppe, H.C., Ruzicka, T., Ahr, H.J., Lehmann, P. and Vohr, V.W. (1998), An integrated model for the differentiation of chemical-induced allergic and irritant skin reactions. *Toxicol. Appl. Pharmacol.*, 153, 83-94.
- (30) Woolhiser, M.R., Hayes, B.B. and Meade, B.J. (1998), A combined murine local lymph node and irritancy assay to predict sensitisation and irritancy potential of chemicals. *Toxicol. Meth.*, 8, 245-256.
- (31) Hayes, B.B. and Meade, B.J. (1999), Contact sensitivity to selected acrylate compounds in B6C3F1 mice: relative potency, cross reactivity, and comparison of test methods. *Drug Chem. Toxicol.*, 22, 491-506.
- (32) Ehling, G., Hecht, M., Heusener, A., Huesler, J., Gamer, A.O., van Loveren, H., Maurer, T., Riecke, K., Ullmann, L., Ulrich, P., Vandebriel, R. and Vohr, H.W. (2005), A European inter-laboratory validation of alternative endpoints of the murine local lymph node assay: first round. *Toxicol.*, 212, 60-68.
- (33) Vohr, H.W. and Ahr, H.J. (2005), The local lymph node assay being too sensitive? *Arch. Toxicol.*, 79, 721-728.
- (34) Patterson, R.M., Noga, E. and Germolec, D. (2007), Lack of evidence for contact sensitisation by Pfiesteria extract. *Environ. Health Perspect.*, 115, 1023-1028.
- (35) ICCVAM (2009), Report on the ICCVAM-NICEATM/ECVAM/JaCVAM Scientific Workshop on Acute Chemical Safety Testing: Advancing *In Vitro* Approaches and Humane Endpoints for Systemic Toxicity Evaluations. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: [[http://iccvam.niehs.nih.gov/methods/acetox/Tox\\_workshop.htm](http://iccvam.niehs.nih.gov/methods/acetox/Tox_workshop.htm)]
- (36) OECD (2000), *Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation*, Environmental Health and Safety Monograph Series on Testing and Assessment No 19, ENV/JM/MONO(2000)7, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (37) Kimber, I., Hilton, J., Dearman, R.J., Gerberick, G.F., Ryan, C.A., Basketter, D.A., Lea, L., House, R.V., Ladies, G.S., Loveless, S.E. and Hastings, K.L. (1998), Assessment of the skin sensitisation potential of topical medicaments using the local lymph node assay: An interlaboratory exercise. *J. Toxicol. Environ. Health*, 53 563-79.
- (38) OECD (2005), *Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment*, Environment, Health and Safety Monograph Series on Testing and Assessment No 34, ENV/JM/MONO (2005)14, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]

▼ **M3***Appendix 1*

## DEFINITIONS

*Accuracy:* The closeness of agreement between test method results and accepted reference values. It is a measure of test method performance and one aspect of relevance. The term is often used interchangeably with 'concordance' to mean the proportion of correct outcomes of a test method (38).

*Benchmark substance:* A sensitising or non-sensitising substance used as a standard for comparison to a test substance. A benchmark substance should have the following properties; (i) a consistent and reliable source(s); (ii) structural and functional similarity to the class of substances being tested; (iii) known physicochemical characteristics; (iv) supporting data on known effects, and (v) known potency in the range of the desired response.

*False negative:* A substance incorrectly identified as negative or non-active by a test method, when in fact it is positive or active.

*False positive:* A substance incorrectly identified as positive or active by a test, when in fact it is negative or non-active.

*Hazard:* The potential for an adverse health or ecological effect. The adverse effect is manifested only if there is an exposure of sufficient level.

*Inter-laboratory reproducibility:* A measure of the extent to which different qualified laboratories, using the same protocol and testing the same test substances, can produce qualitatively and quantitatively similar results. Inter-laboratory reproducibility is determined during the pre-validation and validation processes, and indicates the extent to which a test can be successfully transferred between laboratories, also referred to as between-laboratory reproducibility (38).

*Intra-laboratory reproducibility:* A determination of the extent that qualified people within the same laboratory can successfully replicate results using a specific protocol at different times. Also referred to as within-laboratory reproducibility (38).

*Outlier:* An outlier is an observation that is markedly different from other values in a random sample from a population.

*Quality assurance:* A management process by which adherence to laboratory testing standards, requirements, and record keeping procedures, and the accuracy of data transfer, are assessed by individuals who are independent from those performing the testing.

*Reliability:* Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility (38).

*Skin sensitisation:* An immunological process that results when a susceptible individual is exposed topically to an inducing chemical allergen, which provokes a cutaneous immune response that can lead to the development of contact sensitisation.

*Stimulation Index (SI):* A value calculated to assess the skin sensitisation potential of a test substance that is the ratio of the proliferation in treated groups to that in the concurrent vehicle control group.

*Test substance (also referred to as test chemical):* Any substance or mixture tested using this TM.

**▼ M3****B.51. SKIN SENSITISATION: LOCAL LYMPH NODE ASSAY:  
BrdU-ELISA**

## INTRODUCTION

1. OECD Guidelines for the Testing of Chemicals and EU Test Methods are periodically reviewed in light of scientific progress, changing regulatory needs, and animal welfare considerations. The first Test Method (TM) (B.42) for the determination of skin sensitisation in the mouse, the Local Lymph Node Assay (LLNA; OECD Test Guideline 429) has been revised (1 and Chapter B.42 in this Annex). The details of the validation of the LLNA and a review of the associated work have been published (2) (3) (4) (5) (6) (7) (8) (9). In the LLNA, radioisotopic thymidine or iodine is used to measure lymphocyte proliferation and therefore the assay has limited use where the acquisition, use, or disposal of radioactivity is problematic. The LLNA: BrdU-ELISA (Enzyme-Linked Immunosorbent Assay) is a non-radioactive modification to the LLNA TM, which utilises non-radiolabelled 5-bromo-2-deoxyuridine (BrdU) (Chemical Abstracts Service (CAS) No 59-14-3) in an ELISA-based test system to measure lymphocyte proliferation. The LLNA: BrdU-ELISA has been validated and reviewed and recommended by an international independent scientific peer review panel as considered useful for identifying skin sensitising and non-sensitising chemicals with certain limitations (10) (11) (12). This TM is designed for assessing skin sensitisation potential of chemicals (substances and mixtures) in animals. Chapter B.6 of this Annex and OECD Test Guideline 406 utilise guinea pig tests, notably the guinea pig maximisation test and the Buehler test (13). The LLNA (Chapter B.42 of this Annex; OECD Test Guideline 429) and the two non-radioactive modifications, LLNA: BrdU-ELISA (Chapter B.51 of this Annex; OECD Test Guideline 442 B) and LLNA: DA (Chapter B.50 of this Annex; OECD Test Guideline 442 A), all provide an advantage over the guinea pig tests in B.6 and OECD Test Guideline 406 (13) in terms of reduction and refinement of animal use.
2. Similar to the LLNA, the LLNA: BrdU-ELISA studies the induction phase of skin sensitisation and provides quantitative data suitable for dose-response assessment. Furthermore, an ability to detect skin sensitisers without the necessity for using a radiolabel for DNA eliminates the potential for occupational exposure to radioactivity and waste disposal issues. This in turn may allow for the increased use of mice to detect skin sensitisers, which could further reduce the use of guinea pigs to test for skin sensitisation potential (*i.e.* B.6; OECD Test Guideline 406) (13).

## DEFINITIONS

3. Definitions used are provided in Appendix 1.

## INITIAL CONSIDERATIONS AND LIMITATIONS

4. The LLNA: BrdU-ELISA is a modified LLNA method for identifying potential skin sensitising chemicals, with specific limitations. This does not necessarily imply that in all instances the LLNA: BrdU-ELISA should be used in place of the LLNA or guinea pig tests (*i.e.* B.6; OECD Test Guideline 406) (13), but rather that the assay is of equal merit and may be employed as an alternative in which positive and negative results generally no longer require further confirmation (10) (11). The testing laboratory should consider all available information on the test substance prior to conducting the study. Such information will include the

**▼ M3**

identity and chemical structure of the test substance; its physicochemical properties; the results of any other *in vitro* or *in vivo* toxicity tests on the test substance; and toxicological data on structurally related chemicals. This information should be considered in order to determine whether the LLNA: BrdU-ELISA is appropriate for the test substance (given the incompatibility of limited types of chemicals with the LLNA: BrdU-ELISA (see paragraph 5)) and to aid in dose selection.

5. The LLNA: BrdU-ELISA is an *in vivo* method and, as a consequence, will not eliminate the use of animals in the assessment of allergic contact sensitising activity. It has, however, the potential to reduce the animal use for this purpose when compared to the guinea pig tests (B.6; OECD Test Guideline 406) (13). Moreover, the LLNA: BrdU-ELISA offers a substantial refinement of the way in which animals are used for allergic contact sensitisation testing, since unlike the B.6 and OECD Test Guideline 406, the LLNA: BrdU-ELISA does not require that challenge-induced dermal hypersensitivity reactions be elicited. Furthermore, the LLNA: BrdU-ELISA does not require the use of an adjuvant, as is the case for the guinea pig maximisation test (Chapter B.6 of this Annex, 13). Thus, the LLNA: BrdU-ELISA reduces animal distress. Despite the advantages of the LLNA: BrdU-ELISA over B.6 and OECD Test Guideline 406 (13), there are certain limitations that may necessitate the use of B.6 or OECD Test Guideline 406 (*e.g.* the testing of certain metals, false positive findings with certain skin irritants (such as some surfactant-type substances) (6) (1 and Chapter B.42 in this Annex), solubility of the test substance). In addition, chemical classes or substances containing functional groups shown to act as potential confounders (15) may necessitate the use of guinea pig tests (*i.e.* B.6; OECD Test Guideline 406 (13)). Limitations that have been identified for the LLNA (1 and Chapter B.42 in this Annex) have been recommended to apply also to the LLNA: BrdU-ELISA (10). Other than such identified limitations, the LLNA: BrdU-ELISA should be applicable for testing any chemicals unless there are properties associated with these chemicals that may interfere with the accuracy of the LLNA: BrdU-ELISA. In addition, consideration should be given to the possibility of borderline positive results when Stimulation Index (SI) values between 1,6 and 1,9 are obtained (see paragraphs 31-32). This is based on the validation database of 43 substances using an SI  $\geq 1,6$  (see paragraph 6) for which the LLNA: BrdU-ELISA correctly identified all 32 LLNA sensitisers, but incorrectly identified two of 11 LLNA non-sensitisers with SI values between 1,6 and 1,9 (*i.e.* borderline positive) (10). However, as the same dataset was used for setting the SI-values and calculating the predictive properties of the test, the stated results may be an over-estimation of the real predictive properties.

**PRINCIPLE OF THE TEST METHOD**

6. The basic principle underlying the LLNA: BrdU-ELISA is that sensitisers induce proliferation of lymphocytes in the lymph nodes draining the site of test substance application. This proliferation is proportional to the dose and to the potency of the applied allergen and provides a simple means of obtaining a quantitative measurement of sensitisation. Proliferation is measured by comparing the mean proliferation in each test group to the mean proliferation in the vehicle treated control group (VC). The ratio of the mean proliferation in each treated group to that in the concurrent VC group, termed the SI, is determined, and should be  $\geq 1,6$  before further evaluation of the test substance as a potential skin sensitiser is warranted.

**▼ M3**

The procedures described here are based on the use of measuring BrdU content to indicate an increased number of proliferating cells in the draining auricular lymph nodes. BrdU is an analogue of thymidine and is similarly incorporated into the DNA of proliferating cells. The incorporation of BrdU is measured by ELISA, which utilises an antibody specific for BrdU that is also labelled with peroxidase. When the substrate is added, the peroxidase reacts with the substrate to produce a coloured product that is quantified at a specific absorbance using a microtitre plate reader.

**DESCRIPTION OF THE ASSAY****Selection of animal species**

7. The mouse is the species of choice for this test. Validation studies for the LLNA: BrdU-ELISA were conducted exclusively with the CBA/JN strain, which is therefore considered the preferred strain (10) (12). Young adult female mice, which are nulliparous and non-pregnant, are used. At the start of the study, animals should be between 8-12 weeks old, and the weight variation of the animals should be minimal and not exceed 20 % of the mean weight. Alternatively, other strains and males may be used when sufficient data are generated to demonstrate that significant strain and/or gender-specific differences in the LLNA: BrdU-ELISA response do not exist.

**Housing and feeding conditions**

8. Mice should be group-housed (16), unless adequate scientific rationale for housing mice individually is provided. The temperature of the experimental animal room should be  $22 \pm 3$  °C. Although the relative humidity should be at least 30 % and preferably not exceed 70 %, other than during room cleaning, the aim should be 50-60 %. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water.

**Preparation of animals**

9. The animals are randomly selected, marked to permit individual identification (but not by any form of ear marking), and kept in their cages for at least five days prior to the start of dosing to allow for acclimatisation to the laboratory conditions. Prior to the start of treatment all animals are examined to ensure that they have no observable skin lesions.

**Preparation of dosing solutions**

10. Solid chemicals should be dissolved or suspended in solvents/vehicles and diluted, if appropriate, prior to application to an ear of the mice. Liquid chemicals may be applied neat or diluted prior to dosing. Insoluble chemicals, such as those generally seen in medical devices, should be subjected to an exaggerated extraction in an appropriate solvent to reveal all extractable constituents for testing prior to application to an ear of the mice. Test substances should be prepared daily unless stability data demonstrate the acceptability of storage.

**Reliability check**

11. Positive control chemicals (PC) are used to demonstrate appropriate performance of the assay by responding with adequate and reproducible sensitivity as a sensitising test substance for which the magnitude of the response is well characterised. Inclusion of a concurrent PC is recommended because it demonstrates competency of the laboratory to successfully conduct each assay and allows for an assessment of intra- and inter-laboratory reproducibility and comparability. Some regulatory authorities also require a PC for each study and therefore users are

**▼ M3**

encouraged to consult the relevant authorities prior to conducting the LLNA: BrdU-ELISA. Accordingly, the routine use of a concurrent PC is encouraged to avoid the need for additional animal testing to meet such requirements that might arise from the use of a periodic PC (see paragraph 12). The PC should produce a positive LLNA: BrdU-ELISA response at an exposure level expected to give an increase in the SI  $\geq 1,6$  over the negative control (NC) group. The PC dose should be chosen such that it does not cause excessive skin irritation or systemic toxicity and the induction is reproducible but not excessive (*e.g.* SI  $> 14$  would be considered excessive). Preferred PC are 25 % hexyl cinnamic aldehyde (CAS No 101-86-0) and 25 % eugenol (CAS No 97-53-0) in acetone: olive oil (4:1, v/v). There may be circumstances in which, given adequate justification, other PC, meeting the above criteria, may be used.

12. While inclusion of a concurrent PC group is recommended, there may be situations in which periodic testing (*i.e.* at intervals  $\leq 6$  months) of the PC may be adequate for laboratories that conduct the LLNA: BrdU-ELISA regularly (*i.e.* conduct the LLNA: BrdU-ELISA at a frequency of no less than once per month) and have an established historical PC database that demonstrates the laboratory's ability to obtain reproducible and accurate results with PCs. Adequate proficiency with the LLNA: BrdU-ELISA can be successfully demonstrated by generating consistent positive results with the PC in at least 10 independent tests conducted within a reasonable period of time (*i.e.* less than one year).
13. A concurrent PC group should always be included when there is a procedural change to the LLNA: BrdU-ELISA (*e.g.* change in trained personnel, change in test method materials and/or reagents, change in test method equipment, change in source of test animals), and such changes should be documented in laboratory reports. Consideration should be given to the impact of these changes on the adequacy of the previously established historical database in determining the necessity for establishing a new historical database to document consistency in the PC results.
14. Investigators should be aware that the decision to conduct a PC study on a periodic basis instead of concurrently has ramifications on the adequacy and acceptability of negative study results generated without a concurrent PC during the interval between each periodic PC study. For example, if a false negative result is obtained in the periodic PC study, negative test substance results obtained in the interval between the last acceptable periodic PC study and the unacceptable periodic PC study may be questioned. Implications of these outcomes should be carefully considered when determining whether to include concurrent PCs or to only conduct periodic PCs. Consideration should also be given to using fewer animals in the concurrent PC group when this is scientifically justified and if the laboratory demonstrates, based on laboratory-specific historical data, that fewer mice can be used (17).
15. Although the PC should be tested in the vehicle that is known to elicit a consistent response (*e.g.* acetone: olive oil; 4:1, v/v), there may be certain regulatory situations in which testing in a non-standard vehicle (clinically/chemically relevant formulation) will also be necessary (18). If the concurrent PC is tested in a different vehicle than the test substance, then a separate VC for the concurrent PC should be included.



**▼ M3**

16. In instances where test substances of a specific chemical class or range of responses are being evaluated, benchmark substances may also be useful to demonstrate that the test method is functioning properly for detecting the skin sensitisation potential of these types of test substances. Appropriate benchmark substances should have the following properties:
- structural and functional similarity to the class of the test substance being tested;
  - known physical chemical characteristics;
  - supporting data from the LLNA: BrdU-ELISA;
  - supporting data from other animal models and/or from humans.

**TEST PROCEDURE****Number of animals and dose levels**

17. A minimum of four animals is used per dose group, with a minimum of three concentrations of the test substance, plus a concurrent NC group treated only with the vehicle for the test substance, and a PC group (concurrent or recent, based on laboratory policy in considering paragraphs 11-15). Testing multiple doses of the PC should be considered especially when testing the PC on an intermittent basis. Except for absence of treatment with the test substance, animals in the control groups should be handled and treated in a manner identical to that of animals in the treatment groups.
18. Dose and vehicle selection should be based on the recommendations given in the references 2 and 19. Consecutive doses are normally selected from an appropriate concentration series such as 100 %, 50 %, 25 %, 10 %, 5 %, 2,5 %, 1 %, 0,5 %, etc. Adequate scientific rationale should accompany the selection of the concentration series used. All existing toxicological information (*e.g.* acute toxicity and dermal irritation) and structural and physicochemical information on the test substance of interest (and/or structurally related substances) should be considered, where available, in selecting the three consecutive concentrations so that the highest concentration maximises exposure while avoiding systemic toxicity and/or excessive local skin irritation (19) (20 and Chapter B.4 of this Annex). In the absence of such information, an initial pre-screen test may be necessary (see paragraphs 21-24).
19. The vehicle should not interfere with or bias the test result and should be selected on the basis of maximising the solubility in order to obtain the highest concentration achievable while producing a solution/suspension suitable for application of the test substance. Recommended vehicles are acetone: olive oil (4:1 v/v), *N,N*-dimethylformamide, methyl ethyl ketone, propylene glycol, and dimethyl sulphoxide (6) but others may be used if sufficient scientific rationale is provided. In certain situations it may be necessary to use a clinically relevant solvent or the commercial formulation in which the test substance is marketed as an additional control. Particular care should be taken to ensure that hydrophilic test substances are incorporated into a vehicle system, which wets the skin and does not immediately run off, by incorporation of appropriate solubilisers (*e.g.* 1 % Pluronic® L92). Thus, wholly aqueous vehicles are to be avoided.

**▼ M3**

20. The processing of lymph nodes from individual mice allows for the assessment of inter-animal variability and a statistical comparison of the difference between test substance and VC group measurements (see paragraph 33). In addition, evaluating the possibility of reducing the number of mice in the PC group is only feasible when individual animal data are collected (17). Further, some regulatory authorities require the collection of individual animal data. Regular collection of individual animal data provides an animal welfare advantage by avoiding duplicate testing that would be necessary if the test substance results originally collected in one manner (*e.g.* via pooled animal data) were to be considered later by regulatory authorities with other requirements (*e.g.* individual animal data).

**Pre-screen test**

21. In the absence of information to determine the highest dose to be tested (see paragraph 18), a pre-screen test should be performed in order to define the appropriate dose level to test in the LLNA: BrdU-ELISA. The purpose of the pre-screen test is to provide guidance for selecting the maximum dose level to use in the main LLNA: BrdU-ELISA study, where information on the concentration that induces systemic toxicity (see paragraph 24) and/or excessive local skin irritation (see paragraph 23) is not available. The maximum dose level tested should be a concentration of 100 % of the test substance for liquids or the maximum possible concentration for solids or suspensions.
22. The pre-screen test is conducted under conditions identical to the main LLNA: BrdU-ELISA study, except there is no assessment of lymph node proliferation and fewer animals per dose group can be used. One or two animals per dose group are suggested. All mice will be observed daily for any clinical signs of systemic toxicity or local irritation at the application site. Body weights are recorded pre-test and prior to termination (Day 6). Both ears of each mouse are observed for erythema and scored using Table 1 (20 and Chapter B.4 of this Annex). Ear thickness measurements are taken using a thickness gauge (*e.g.* digital micrometer or Peacock Dial thickness gauge) on Day 1 (pre-dose), Day 3 (approximately 48 hours after the first dose), and Day 6. Additionally, on Day 6, ear thickness could be determined by ear punch weight determinations, which should be performed after the animals are humanely killed. Excessive local irritation is indicated by an erythema score  $\geq 3$  and/or ear thickness of  $\geq 25$  % on any day of measurement (21) (22). The highest dose selected for the main LLNA: BrdU-ELISA study will be the next lower dose in the pre-screen concentration series (see paragraph 18) that does not induce systemic toxicity and/or excessive local skin irritation.

Table 1

**Erythema Scores**

Observation	Score
No erythema	0
Very slight erythema (barely perceptible)	1
Well-defined erythema	2

▼ **M3**

Observation	Score
Moderate to severe erythema	3
Severe erythema (beet redness) to eschar formation preventing grading of erythema	4

23. In addition to a 25 % increase in ear thickness (21) (22), a statistically significant increase in ear thickness in the treated mice compared to control mice has also been used to identify irritants in the LLNA (22) (23) (24) (25) (26) (27) (28). However, while statistically significant increases can occur when ear thickness is less than 25 % they have not been associated specifically with excessive irritation (25) (26) (27) (28) (29).

24. The following clinical observations may indicate systemic toxicity (30) when used as part of an integrated assessment and therefore may indicate the maximum dose level to use in the main LLNA: BrdU-ELISA: changes in nervous system function (*e.g.* pilo-erection, ataxia, tremors, and convulsions); changes in behaviour (*e.g.* aggressiveness, change in grooming activity, marked change in activity level); changes in respiratory patterns (*i.e.* changes in frequency and intensity of breathing such as dyspnea, gasping, and rales), and changes in food and water consumption. In addition, signs of lethargy and/or unresponsiveness and any clinical signs of more than slight or momentary pain and distress, or a > 5 % reduction in body weight from Day 1 to Day 6 and mortality should be considered in the evaluation. Moribund animals or animals showing signs of severe pain and distress should be humanely killed (31).

**Main study experimental schedule**

25. The experimental schedule of the assay is as follows:

- *Day 1:* Individually identify and record the weight of each animal and any clinical observation. Apply 25 µL of the appropriate dilution of the test substance, the vehicle alone, or the PC (concurrent or recent, based on laboratory policy in considering paragraphs 11-15), to the dorsum of each ear.
- *Days 2 and 3:* Repeat the application procedure carried out on Day 1.
- *Day 4:* No treatment.
- *Day 5:* Inject 0,5 mL (5 mg/mouse) of BrdU (10 mg/mL) solution intra-peritoneally.
- *Day 6:* Record the weight of each animal and any clinical observation. Approximately 24 hours (24 h) after BrdU injection, humanely kill the animals. Excise the draining auricular lymph nodes from each mouse ear and process separately in phosphate buffered saline (PBS) for each animal. Details and diagrams of the lymph node identification and dissection can be found in reference (17). To further monitor the local skin response in the main study, additional parameters such as scoring of ear erythema or ear thickness measurements (obtained either by using a thickness gauge, or ear punch weight determinations at necropsy) may be included into the study protocol.

**▼ M3****Preparation of cell suspensions**

26. From each mouse, a single-cell suspension of lymph node cells (LNC) excised bilaterally is prepared by gentle mechanical disaggregation through 200 micron-mesh stainless steel gauze or another acceptable technique for generating a single-cell suspension (*e.g.* use of a disposable plastic pestle to crush the lymph nodes followed by passage through a #70 nylon mesh). The procedure for preparing the LNC suspension is critical in this assay and therefore every operator should establish the skill in advance. Further, the lymph nodes in NC animals are small, so careful operation is important to avoid any artificial effects on SI values. In each case, the target volume of the LNC suspension should be adjusted to a determined optimised volume (approximately 15 mL). The optimised volume is based on achieving a mean absorbance of the NC group within 0,1-0,2.

**Determination of cellular proliferation (measurement of BrdU content in DNA of lymphocytes)**

27. BrdU is measured by ELISA using a commercial kit (*e.g.* Roche Applied Science, Mannheim, Germany, Catalogue Number 11 647 229 001). Briefly, 100 µL of the LNC suspension is added to the wells of a flat-bottom microplate in triplicate. After fixation and denaturation of the LNC, anti-BrdU antibody is added to each well and allowed to react. Subsequently the anti-BrdU antibody is removed by washing and the substrate solution is then added and allowed to produce chromogen. Absorbance at 370 nm with a reference wavelength of 492 nm is then measured. In all cases, assay test conditions should be optimised (see paragraph 26).

**OBSERVATIONS****Clinical observations**

28. Each mouse should be carefully observed at least once daily for any clinical signs, either of local irritation at the application site or of systemic toxicity. All observations are systematically recorded with records being maintained for each mouse. Monitoring plans should include criteria to promptly identify those mice exhibiting systemic toxicity, excessive local skin irritation, or corrosion of skin for euthanasia (31).

**Body weights**

29. As stated in paragraph 25, individual animal body weights should be measured at the start of the test and at the scheduled humane kill.

**CALCULATION OF RESULTS**

30. Results for each treatment group are expressed as the mean SI. The SI is derived by dividing the mean BrdU labelling index/mouse within each test substance group and the PC group by the mean BrdU labelling index for the solvent/VC group. The average SI for the VCs is then one.

The BrdU labelling index is defined as:

$$\text{BrdU labelling index} = (\text{ABS}_{\text{em}} - \text{ABS blank}_{\text{em}}) - (\text{ABS}_{\text{ref}} - \text{ABS blank}_{\text{ref}})$$

Where: em = emission wavelength; and ref = reference wavelength.

**▼ M3**

31. The decision process regards a result as positive when  $SI \geq 1,6$  (10). However, the strength of the dose-response relationship, the statistical significance and the consistency of the solvent/vehicle and PC responses may also be used when determining whether a borderline result (*i.e.* SI value between 1,6 and 1,9) is declared positive (3) (6) (32).
32. For a borderline positive response between an SI of 1,6 and 1,9, users may want to consider additional information such as dose-response relationship, evidence of systemic toxicity or excessive irritation, and where appropriate, statistical significance together with SI values to confirm that such results are positives (10). Consideration should also be given to various properties of the test substance, including whether it has a structural relationship to known skin sensitisers, whether it causes excessive skin irritation in the mouse, and the nature of the dose-response observed. These and other considerations are discussed in detail elsewhere (4).
33. Collecting data at the level of the individual mouse will enable a statistical analysis for presence and degree of dose-response relationship in the data. Any statistical assessment could include an evaluation of the dose-response relationship as well as suitably adjusted comparisons of test groups (*e.g.* pair-wise dosed group versus concurrent solvent/vehicle control comparisons). Statistical analyses may include, *e.g.* linear regression or William's test to assess dose-response trends, and Dunnett's test for pair-wise comparisons. In choosing an appropriate method of statistical analysis, the investigator should maintain an awareness of possible inequalities of variances and other related problems that may necessitate a data transformation or a non-parametric statistical analysis. In any case, the investigator may need to carry out SI calculations and statistical analyses with and without certain data points (sometimes called 'outliers').

**DATA AND REPORTING****Data**

34. Data should be summarised in tabular form showing the individual animal BrdU labelling index values, the group mean BrdU labelling index/animal, its associated error term (*e.g.* SD, SEM), and the mean SI for each dose group compared against the concurrent solvent/vehicle control group.

**Test report**

35. The test report should contain the following information:

Test and control chemicals:

- identification data (*e.g.* CAS number and EC number, if available; source; purity; known impurities; lot number);
- physical nature and physicochemical properties (*e.g.* volatility, stability, solubility);
- if mixture, composition and relative percentages of components;

Solvent/vehicle:

- identification data (purity; concentration, where appropriate; volume used);
- justification for choice of vehicle;

**▼ M3**

## Test animals:

- source of CBA mice;
- microbiological status of the animals, when known;
- number and age of animals;
- source of animals, housing conditions, diet, etc.;

## Test conditions:

- source, lot number, and manufacturer's quality assurance/quality control data (antibody sensitivity and specificity and the limit of detection) for the ELISA kit;
- details of test substance preparation and application;
- justification for dose selection (including results from pre-screen test, if conducted);
- vehicle and test substance concentrations used, and total amount of test substance applied;
- details of food and water quality (including diet type/source, water source);
- details of treatment and sampling schedules;
- methods for measurement of toxicity;
- criteria for considering studies as positive or negative;
- details of any protocol deviations and an explanation on how the deviation affects the study design and results;

## Reliability check:

- a summary of results of latest reliability check, including information on test substance, concentration and vehicle used;
- concurrent and/or historical PC and concurrent negative (solvent/vehicle) control data for testing laboratory;
- if a concurrent PC was not included, the date and laboratory report for the most recent periodic PC and a report detailing the historical PC data for the laboratory justifying the basis for not conducting a concurrent PC;

## Results:

- individual weights of mice at start of dosing and at scheduled humane kill; as well as mean and associated error term (*e.g.* SD, SEM) for each treatment group;
- time course of onset and signs of toxicity, including dermal irritation at site of administration, if any, for each animal;
- a table of individual mouse BrdU labelling indices and SI values for each treatment group;
- mean and associated error term (*e.g.* SD, SEM) for BrdU labelling index/mouse for each treatment group and the results of outlier analysis for each treatment group;

▼ **M3**

- calculated SI and an appropriate measure of variability that takes into account the inter-animal variability in both the test substance and control groups;
- dose-response relationship;
- statistical analyses, where appropriate;

## Discussion of results:

- a brief commentary on the results, the dose-response analysis, and statistical analyses, where appropriate, with a conclusion as to whether the test substance should be considered a skin sensitiser.

## LITERATURE

- (1) OECD (2010), Skin Sensitisation: Local Lymph Node Assay, Test Guideline No 429, Guidelines for the Testing of Chemicals, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (2) Chamberlain, M. and Basketter, D.A. (1996), The local lymph node assay: status of validation. *Food Chem. Toxicol.*, 34, 999-1002.
- (3) Basketter, D.A., Gerberick, G.F., Kimber, I. and Loveless, S.E. (1996), The local lymph node assay: A viable alternative to currently accepted skin sensitisation tests. *Food Chem. Toxicol.*, 34, 985-997.
- (4) Basketter, D.A., Gerberick, G.F. and Kimber, I. (1998), Strategies for identifying false positive responses in predictive sensitisation tests. *Food Chem. Toxicol.*, 36, 327-33.
- (5) Van Och, F.M.M., Slob, W., De Jong, W.H., Vandebriel, R.J. and Van Loveren, H. (2000), A quantitative method for assessing the sensitising potency of low molecular weight chemicals using a local lymph node assay: employment of a regression method that includes determination of uncertainty margins. *Toxicol.*, 146, 49-59.
- (6) ICCVAM (1999), The murine local lymph node Assay: A test method for assessing the allergic contact dermatitis potential of chemicals/compounds: The results of an independent peer review evaluation coordinated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program Center for the Evaluation of Alternative Toxicological Methods (NICETAM). NIH Publication No: 99-4494. Research Triangle Park, N.C. Available at: [[http://iccvam.niehs.nih.gov/docs/immunotox\\_docs/llna/llnarep.pdf](http://iccvam.niehs.nih.gov/docs/immunotox_docs/llna/llnarep.pdf)]
- (7) Dean, J.H., Twerdok, L.E., Tice, R.R., Sailstad, D.M., Hattan, D.G., Stokes, W.S. (2001), ICCVAM evaluation of the murine local lymph node assay: II. Conclusions and recommendations of an independent scientific peer review panel. *Reg. Toxicol. Pharmacol.*, 34(3), 258-273.
- (8) Haneke, K.E., Tice, R.R., Carson, B.L., Margolin, B.H., Stokes, W.S. (2001), ICCVAM evaluation of the murine local lymph node assay: III. Data analyses completed by the national toxicology program interagency center for the evaluation of alternative toxicological methods. *Reg. Toxicol. Pharmacol.*, 34(3), 274-286.
- (9) Sailstad, D.M., Hattan, D., Hill, R.N., Stokes, W.S. (2001), ICCVAM evaluation of the murine local lymph node assay: I. The ICCVAM review process. *Reg. Toxicol. Pharmacol.*, 34(3), 249-257.
- (10) ICCVAM (2010), ICCVAM Test Method Evaluation Report. Nonradioactive local lymph node assay: BrdU-ELISA Test Method Protocol (LLNA: BrdU-ELISA). NIH Publication No 10-7552A/B. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: [<http://iccvam.niehs.nih.gov/methods/immunotox/llna-ELISA/TMER.htm>]

▼ **M3**

- (11) ICCVAM (2009), Independent Scientific Peer Review Panel Report: Updated validation status of new versions and applications of the murine local lymph node assay: a test method for assessing the allergic contact dermatitis potential of chemicals and products. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: [[http://iccvam.niehs.nih.gov/docs/immunotox\\_docs/LLNAPRPRpt2009.pdf](http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAPRPRpt2009.pdf)]
- (12) Takeyoshi, M., Iida, K., Shiraishi, K. and Hoshuyama, S. (2005), Novel approach for classifying chemicals according to skin sensitising potency by non-radioisotopic modification of the local lymph node assay. *J. Appl. Toxicol.*, 25, 129-134.
- (13) OECD (1992), Skin Sensitisation, Test Guideline No 406, Guidelines for Testing of Chemicals, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (14) Kreiling, R., Hollnagel, H.M., Hareng, L., Eigler, L., Lee, M.S., Griem, P., Dreesen, B., Kleber, M., Albrecht, A., Garcia, C. and Wendel, A. (2008), Comparison of the skin sensitising potential of unsaturated compounds as assessed by the murine local lymph node assay (LLNA) and the guinea pig maximization test (GPMT). *Food Chem. Toxicol.*, 46, 1896-1904.
- (15) Basketter, D., Ball, N., Cagen, S., Carrilo, J.C., Certa, H., Eigler, D., Garcia, C., Esch, H., Graham, C., Haux, C., Kreiling, R. and Mehling, A. (2009), Application of a weight of evidence approach to assessing discordant sensitisation datasets: implications for REACH. *Toxicol. Pharmacol.*, 55, 90-96.
- (16) ILAR (1996), Institute of Laboratory Animal Research (ILAR) Guide for the Care and Use of Laboratory Animals. 7th ed. Washington, DC: National Academies Press.
- (17) ICCVAM (2009), Recommended Performance Standards: Murine Local Lymph Node Assay. NIH Publication Number 09-7357. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: [[http://iccvam.niehs.nih.gov/docs/immunotox\\_docs/llnaps/LLNAPerfStds.pdf](http://iccvam.niehs.nih.gov/docs/immunotox_docs/llnaps/LLNAPerfStds.pdf)]
- (18) McGarry, H.F. (2007), The murine local lymph node assay: regulatory and potency considerations under REACH. *Toxicol.*, 238, 71-89.
- (19) Kimber, I., Dearman, R.J., Scholes E.W. and Basketter, D.A. (1994), The local lymph node assay: developments and applications. *Toxicol.*, 93, 13-31.
- (20) OECD (2002), Acute Dermal Irritation/Corrosion, Test Guideline No 404, Guidelines for Testing of Chemicals, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (21) Reeder, M.K., Broomhead, Y.L., DiDonato, L. and DeGeorge, G.L. (2007), Use of an enhanced local lymph node assay to correctly classify irritants and false positive substances. *Toxicologist*, 96, 235.
- (22) ICCVAM (2009), Nonradioactive Murine Local Lymph Node Assay: Flow Cytometry Test Method Protocol (LLNA: BrdU-FC) Revised Draft Background Review Document. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: [<http://iccvam.niehs.nih.gov/methods/immunotox/fcLLNA/BRDcomplete.pdf>].
- (23) Hayes, B.B., Gerber, P.C., Griffey, S.S. and Meade, B.J. (1998), Contact hypersensitivity to dicyclohexylcarbodiimide and diisopropylcarbodiimide in female B6C3F1 mice. *Drug Chem. Toxicol.*, 21, 195-206.
- (24) Homey, B., von Schilling, C., Blumel, J., Schuppe, H.C., Ruzicka, T., Ahr, H.J., Lehmann, P. and Vohr, V.W. (1998), An integrated model for the differentiation of chemical-induced allergic and irritant skin reactions. *Toxicol. Appl. Pharmacol.*, 153, 83-94.
- (25) Woolhiser, M.R., Hayes, B.B. and Meade, B.J. (1998), A combined murine local lymph node and irritancy assay to predict sensitisation and irritancy potential of chemicals. *Toxicol. Meth.*, 8, 245-256.



▼ **M3**

- (26) Hayes, B.B. and Meade, B.J. (1999), Contact sensitivity to selected acrylate compounds in B6C3F1 mice: relative potency, cross reactivity, and comparison of test methods. *Drug. Chem. Toxicol.*, 22, 491-506.
- (27) Ehling, G., Hecht, M., Heusener, A., Huesler, J., Gamer, A.O., van Loveren, H., Maurer, T., Riecke, K., Ullmann, L., Ulrich, P., Vandebriel, R. and Vohr, H.W. (2005), A European inter-laboratory validation of alternative endpoints of the murine local lymph node assay: first round. *Toxicol.*, 212, 60-68.
- (28) Vohr, H.W. and Ahr, H.J. (2005), The local lymph node assay being too sensitive? *Arch. Toxicol.*, 79, 721-728.
- (29) Patterson, R.M., Noga, E. and Germolec D. (2007), Lack of evidence for contact sensitisation by Pfiesteria extract. *Environ. Health Perspect.*, 115, 1023-1028.
- (30) ICCVAM (2009), Report on the ICCVAM-NICEATM/ECVAM/JaCVAM Scientific Workshop on Acute Chemical Safety Testing: Advancing *In Vitro* Approaches and Humane Endpoints for Systemic Toxicity Evaluations. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: [[http://iccvam.niehs.nih.gov/methods/acute/tox\\_workshop.htm](http://iccvam.niehs.nih.gov/methods/acute/tox_workshop.htm)].
- (31) OECD (2000), Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation, Environmental Health and Safety Monograph Series on Testing and Assessment No 19, ENV/JM/MONO(2000)7, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (32) Kimber, I., Hilton, J., Dearman, R.J., Gerberick, G.F., Ryan, C.A., Basketter, D.A., Lea, L., House, R.V., Ladies, G.S., Loveless, S.E. and Hastings, K.L. (1998), Assessment of the skin sensitisation potential of topical medicaments using the local lymph node assay: An interlaboratory exercise. *J. Toxicol. Environ. Health*, 53, 563-79.
- (33) OECD (2005), Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment, Environment, Health and Safety Monograph Series on Testing and Assessment No 34, ENV/JM/MONO(2005)14, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]

▼ **M3***Appendix 1*

## DEFINITIONS

*Accuracy:* The closeness of agreement between test method results and accepted reference values. It is a measure of test method performance and one aspect of relevance. The term is often used interchangeably with 'concordance' to mean the proportion of correct outcomes of a test method (33).

*Benchmark substance:* A sensitising or non-sensitising substance used as a standard for comparison to a test substance. A benchmark substance should have the following properties: (i) a consistent and reliable source(s); (ii) structural and functional similarity to the class of substances being tested; (iii) known physical/chemical characteristics; (iv) supporting data on known effects; and (v) known potency in the range of the desired response.

*False negative:* A test substance incorrectly identified as negative or non-active by a test method, when in fact it is positive or active (33).

*False positive:* A test substance incorrectly identified as positive or active by a test, when in fact it is negative or non-active (33).

*Hazard:* The potential for an adverse health or ecological effect. The adverse effect is manifested only if there is an exposure of sufficient level.

*Inter-laboratory reproducibility:* A measure of the extent to which different qualified laboratories, using the same protocol and testing the same test substance, can produce qualitatively and quantitatively similar results. Inter-laboratory reproducibility is determined during the pre-validation and validation processes, and indicates the extent to which a test can be successfully transferred between laboratories, also referred to as between-laboratory reproducibility (33).

*Intra-laboratory reproducibility:* A determination of the extent that qualified people within the same laboratory can successfully replicate results using a specific protocol at different times. Also referred to as within-laboratory reproducibility (33).

*Outlier:* An outlier is an observation that is markedly different from other values in a random sample from a population.

*Quality assurance:* A management process by which adherence to laboratory testing standards, requirements, and record keeping procedures, and the accuracy of data transfer, are assessed by individuals who are independent from those performing the testing.

*Reliability:* Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility (33).

*Skin sensitisation:* An immunological process that results when a susceptible individual is exposed topically to an inducing chemical allergen, which provokes a cutaneous immune response that can lead to the development of contact sensitisation.

*Stimulation Index (SI):* A value calculated to assess the skin sensitisation potential of a test substance that is the ratio of the proliferation in treated groups to that in the concurrent vehicle control group.

*Test substance (also referred to as test chemical):* Any substance or mixture tested using this TM.

**▼M4****B.52. ACUTE INHALATION TOXICITY — ACUTE TOXIC CLASS METHOD**

## INTRODUCTION

1. This Test Method is equivalent to OECD Test Guideline (TG) 436 (2009). The first acute inhalation TG 403 was adopted in 1981, and has since been revised (see chapter B.2 of this Annex (1)). Development of an Inhalation Acute Toxic Class (ATC) method (2) (3) (4) was considered appropriate following the adoption of the revised oral ATC method (chapter B.1 tris of this Annex) (5). A retrospective performance assessment of the ATC test method for acute inhalation toxicity showed that the method is suitable for being used for Classification and Labelling purposes (6). The inhalation ATC Test Method will allow the use of serial steps of fixed target concentrations to provide a ranking of test chemical toxicity. Lethality is used as key endpoint, however, animals in severe pain or distress, suffering or impending death should be humanely killed to minimise suffering. Guidance on humane endpoints is available in the OECD Guidance Document No 19 (7).
2. Guidance on the conduct and interpretation of this Test Method can be found in the Guidance Document No 39 on Acute Inhalation Toxicity Testing (GD 39) (8).
3. Definitions used in the context of this Test Method are provided in Appendix 1 and in GD 39 (8).
4. The Test Method provides information on the hazardous properties and allows the test chemical to be ranked and classified according to the Regulation (EC) No 1272/2008 for the classification of chemicals that cause acute toxicity (9). In case point estimates of LC<sub>50</sub>-values or concentration-response analyses are required, chapter B.2 of this Annex (1) is the appropriate Test Method to use. Further guidance on Test Method selection can be found in GD 39 (8). This Test Method is not specially intended for the testing of specialized materials, such as poorly soluble isometric or fibrous materials or manufactured nanomaterials.

## INITIAL CONSIDERATIONS

5. Before considering testing in accordance with this Test Method, all available information on the test chemical, including existing studies whose data would support not doing additional testing should be considered by the testing laboratory in order to minimize animal usage. Information that may assist in the selection of the most appropriate species, strain, sex, mode of exposure and appropriate test concentrations include the identity, chemical structure, and physico-chemical properties of the test chemical; results of any *in vitro* or *in vivo* toxicity tests; anticipated use(s) and potential for human exposure; available (Q)SAR data and toxicological data on structurally related chemicals. Concentrations that are expected to cause severe pain and distress, due to corrosive<sup>(1)</sup> or severely irritant actions, should not be tested with this Test Method [see GD 39 (8)].

<sup>(1)</sup> The corrosivity evaluation could be based on expert judgment using such evidence as: human and animal experience, existing (in vitro) data, e.g. chapter B.40 (10), B.40 bis (11) of this annex or OECD TG 435 (12), pH values, information from similar chemicals or any other pertinent data.

**▼M4**

## PRINCIPLE OF THE TEST

6. It is the principle of the test that based on a stepwise procedure, sufficient information is obtained on the acute inhalation toxicity of the test chemical during an exposure period of 4 hours to enable its classification. Other durations of exposure may apply to serve specific regulatory purposes. At any of the defined concentration steps, 3 animals of each sex are tested. Depending on the mortality and/or the moribund status of the animals, 2 steps may be sufficient to allow judgement on the acute toxicity of the test chemical. If evidence is provided that one sex is more susceptible than the other, then the test may be continued with the more susceptible sex only. The outcome of the previous step will determine the following step such that:
  - a) No further testing is needed,
  - b) Testing of three animals per sex, or
  - c) Testing with 6 animals of the more susceptible sex only *i.e.* the lower boundary estimates of the toxic class should be based on 6 animals per test concentration group, regardless of sex.
7. Moribund animals or animals obviously in pain or showing signs of severe and enduring distress should be humanely killed, and are considered in the interpretation of the test results in the same way as animals that died on test. Criteria for making the decision to kill moribund or severely suffering animals, and guidance on the recognition of predictable or impending death, are the subject of Guidance Document No 19 on Humane Endpoints (7).

## DESCRIPTION OF THE METHOD

**Selection of animal species**

8. Healthy young adult animals of commonly used laboratory strains should be used. The preferred species is the rat and justifications should be provided if other species are used.

**Preparation of animals**

9. Females should be nulliparous and non-pregnant. On the exposure day, animals should be young adults 8 to 12 weeks of age, and body weights should be within  $\pm 20$  % of the mean weight for each sex of any previously exposed animals at the same age. The animals are randomly selected, marked for individual identification. The animals are kept in their cages for at least 5 days prior to the start of the test to allow for acclimatisation to laboratory conditions. Animals should also be acclimatised to the test apparatus for a short period prior to testing, as this will lessen the stress caused by introduction to the new environment.

**Animal husbandry**

10. The temperature of the experimental animal maintenance room should be  $22 \pm 3$  °C. The relative humidity should ideally be maintained in the range of 30 to 70 %, though this may not be possible when using water as a vehicle. Before and after exposures, animals generally should be caged in groups by sex and concentration, but the number of animals per cage should not interfere with clear observation of each animal and should minimise losses due to cannibalism and fighting. When animals are to be exposed nose-only, it may be necessary for them to be acclimated to the restraining tubes. The restraining tubes should not impose undue physical, thermal, or immobilisation stress on the animals. Restraint may affect physiological

**▼M4**

endpoints such as body temperature (hyperthermia) and/or respiratory minute volume. If generic data are available to show that no such changes occur to any appreciable extent, then pre-adaptation to the restraining tubes is not necessary. Animals exposed whole-body to an aerosol should be housed individually during exposure to prevent them from filtering the test aerosol through the fur of their cage mates. Conventional and certified laboratory diets may be used, except during exposure, accompanied with an unlimited supply of municipal drinking water. Lighting should be artificial, the sequence being 12 hours light/12 hours dark.

**Inhalation chambers**

11. The nature of the test chemical and the objective of the test should be considered when selecting an inhalation chamber. The preferred mode of exposure is nose-only (which term includes head-only, nose-only or snout-only). Nose-only exposure is generally preferred for studies of liquid or solid aerosols and for vapours that may condense to form aerosols. Special objectives of the study may be better achieved by using a whole-body mode of exposure, but this should be justified in the study report. To ensure atmosphere stability when using a whole-body chamber, the total volume of the test animals should not exceed 5 % of the chamber volume. Principles of the nose-only and whole body exposure techniques and their particular advantages and disadvantages are described in GD 39 (8).

**EXPOSURE CONDITIONS****Administrations of concentrations**

12. A fixed duration of exposure for four hours, excluding equilibration time, is recommended. Other durations may be needed to meet specific requirements, however, justification should be provided in the study report [see GD 39 (8)]. Animals exposed in whole-body chambers should be housed individually to prevent ingestion of test chemical due to grooming of cage mates. Feed should be withheld during the exposure period. Water may be provided throughout a whole-body exposure.
13. Animals are exposed to the test chemical as a gas, vapour, aerosol, or a mixture thereof. The physical state to be tested depends on the physico-chemical properties of the test chemical, the selected concentration, and/or the physical form most likely present during the handling and use of the test chemical. Hygroscopic and chemically reactive test chemicals should be tested under dry air conditions. Care should be taken to avoid generating explosive concentrations.

**Particle-size distribution**

14. Particle sizing should be performed for all aerosols and for vapours that may condense to form aerosols. To allow for exposure of all relevant regions of the respiratory tract, aerosols with mass median aerodynamic diameters (MMAD) ranging from 1 to 4  $\mu\text{m}$  with a geometric standard deviation ( $\sigma_g$ ) in the range of 1,5 to 3,0 are recommended (8) (13) (14). Although a reasonable effort should be made to meet this standard, expert judgment should be provided if it cannot be achieved. For example, metal fumes may be smaller than this standard, and charged particles, fibres, and hygroscopic materials (which increase in size in the moist environment of the respiratory tract) may exceed this standard.

**▼ M4****Test chemical preparation in a vehicle**

15. A vehicle may be used to generate an appropriate concentration and particle size of the test chemical in the atmosphere. As a rule, water should be given preference. Particulate material may be subjected to mechanical processes to achieve the required particle size distribution, however, care should be taken not to decompose or alter the test chemical. In cases where mechanical processes are believed to have altered test chemical composition (e.g. extreme temperature from excessive milling due to friction), the composition of the test chemical should be verified analytically. Adequate care should be taken to not contaminate the test chemical. It is not necessary to test non-friable granular materials which are purposefully formulated to be un-inhalable. An attrition test should be used to demonstrate that respirable particles are not produced when the granular material is handled. If an attrition test produces respirable particles, an inhalation toxicity test should be performed.

**Control animals**

16. A concurrent negative (air) control group is not necessary. When a vehicle other than water is used to assist in generating the test atmosphere, a vehicle control group should only be used when historical inhalation toxicity data are not available. If a toxicity study of a test chemical formulated in a vehicle reveals no toxicity, it follows that the vehicle is non-toxic at the concentration tested; thus, there is no need for a vehicle control.

**MONITORING OF EXPOSURE CONDITIONS****Chamber airflow**

17. The flow of air through the chamber should be carefully controlled, continuously monitored, and recorded at least hourly during each exposure. The monitoring of test atmosphere concentration (or stability) is an integral measurement of all dynamic parameters and provides an indirect means to control all relevant dynamic atmosphere generation parameters. Special consideration should be given to avoiding re-breathing in nose-only chambers in cases where airflow through the exposure system are inadequate to provide dynamic flow of test chemical atmosphere. There are prescribed methodologies that can be used to demonstrate that re-breathing does not occur under the selected operation conditions (8) (15). Oxygen concentration should be at least 19 % and carbon dioxide concentration should not exceed 1 %. If there is reason to believe that these standards cannot be met, oxygen and carbon dioxide concentrations should be measured.

**Chamber temperature and relative humidity**

18. Chamber temperature should be maintained at  $22 \pm 3$  °C. Relative humidity in the animals' breathing zone, for both nose-only and whole-body exposures, should be monitored and recorded at least three times for durations up to 4 hrs, and hourly for shorter durations. The relative humidity should ideally be maintained in the range of 30 to 70 %, but this may either be unattainable (e.g. when testing water based mixtures) or not measurable due to test chemical interference with the test method.

**▼ M4****Test chemical: nominal concentration**

19. Whenever feasible, the nominal exposure chamber concentration should be calculated and recorded. The nominal concentration is the mass of generated test chemical divided by the total volume of air passed through the chamber system. The nominal concentration is not used to characterise the animals' exposure, but a comparison of the nominal and the actual concentration gives an indication of the generation efficiency of the test system, and thus may be used to discover generation problems.

**Test chemical: actual concentration**

20. The actual concentration is the test chemical concentration at the animals' breathing zone in an inhalation chamber. Actual concentrations can be obtained either by specific methods (*e.g.* direct sampling, adsorptive or chemical reactive methods, and subsequent analytical characterisation) or by non-specific methods such as gravimetric filter analysis. The use of gravimetric analysis is acceptable only for single component powder aerosols or aerosols of low volatility liquids and should be supported by appropriate pre-study test chemical-specific characterisations. Multi-component powder aerosol concentration may also be determined by gravimetric analysis. However, this requires analytical data which demonstrate that the composition of airborne material is similar to the starting material. If this information is not available, a reanalysis of the test chemical (ideally in its airborne state) at regular intervals during the course of the study may be necessary. For aerosolised agents that may evaporate or sublime, it should be shown that all phases were collected by the method chosen. The target, nominal, and actual concentrations should be provided in the study report, but only actual concentrations are used in statistical analyses to calculate lethal concentration values.
21. One lot of the test chemical should be used, if possible, and the test sample should be stored under conditions that maintain its purity, homogeneity, and stability. Prior to the start of the study, there should be a characterisation of the test chemical, including its purity and, if technically feasible, the identity, and quantities of identified contaminants and impurities. This can be demonstrated by, but is not limited to, the following data: retention time and relative peak area, molecular weight from mass spectroscopy or gas chromatography analyses, or other estimates. Although the test sample's identity is not the responsibility of the test laboratory, it may be prudent for the test laboratory to confirm the sponsor's characterisation at least in a limited way (*e.g.* colour, physical nature, etc.).
22. The exposure atmosphere shall be held as constant as practicable and monitored continuously and/or intermittently depending on the method of analysis. When intermittent sampling is used, chamber atmosphere samples should be taken at least twice in a four hour study. If not feasible due to limited air flow rates or low concentrations, one sample may be collected over the entire exposure period. If marked sample-to-sample fluctuations occur, the next concentrations tested should use four samples per exposure. Individual chamber concentration samples should not deviate from the mean chamber concentration by more than  $\pm 10\%$  for gases and vapours, and by no more than  $\pm 20\%$  for liquid or solid aerosols. Time to chamber equilibration ( $t_{95}$ ) should be calculated and recorded. The duration of an exposure spans the time that the test chemical is generated and this takes into account the times required to attain  $t_{95}$ . Guidance for estimating  $t_{95}$  can be found in GD 39 (8).

**▼M4**

23. For very complex mixtures consisting of vapours/gases, and aerosols (*e.g.* combustion atmospheres and test chemicals propelled from purpose-driven end-use products/devices), each phase may behave differently in an inhalation chamber so at least one indicator substance (analyte), normally the principal active substance in the mixture, of each phase (vapour/gas and aerosol) should be selected. When the test chemical is a mixture, the analytical concentration should be reported for the total mixture and not just for the active ingredient or the component (analyte). Additional information regarding actual concentrations can be found in GD 39 (8).

**Test chemical: particle size distribution**

24. The particle size distribution of aerosols should be determined at least twice during each 4 hour exposure by using a cascade impactor or an alternative instrument such as an aerodynamic particle sizer. If equivalence of the results obtained by a cascade impactor or an alternative instrument can be shown, then the alternative instrument may be used throughout the study. A second device, such as a gravimetric filter or an impinger/gas bubbler, should be used in parallel to the primary instrument to confirm the collection efficiency of the primary instrument. The mass concentration obtained by particle size analysis should be within reasonable limits of the mass concentration obtained by filter analysis [see GD 39 (8)]. If equivalence can be demonstrated in the early phase of the study, then further confirmatory measurements may be omitted. For animal welfare reasons, measures should be taken to minimize inconclusive data which may lead to a need to repeat an exposure. Particle sizing should be performed for vapours if there is any possibility that vapour condensation may result in the formation of an aerosol, or if particles are detected in a vapour atmosphere with potential for mixed phases (see paragraph 14).

**PROCEDURE****Main test**

25. Three animals per sex, or six animals of the more susceptible sex, are used for each step. If rodent species other than rats are exposed nose-only, maximum exposure durations may be adjusted to minimise species-specific distress. The concentration level to be used as the starting dose is selected from one of four fixed levels and the starting concentration level should be that which is most likely to produce toxicity in some of the dosed animals. The testing schemes for gases, vapours and aerosols (included in Appendixes 2-4) represent the testing with the cut-off values of the CLP categories 1-4 (9) for gases (100, 500, 2 500, 20 000 ppm/4h) (Appendix 2), for vapours (0,5, 2, 10, 20 mg/l/4h) (Appendix 3) and for aerosols (0,05, 0,5, 1, 5 mg/l/4h) (Appendix 4). Category 5, which is not implemented in Regulation (EC) No 1272/2008 (9) relates to concentrations above the respective limit concentrations. For each starting concentration, the respective testing scheme applies. Depending on the number of humanely killed or dead animals, the test procedure follows the indicated arrows until a categorisation can be made.
26. The time interval between exposure groups is determined by the onset, duration, and severity of toxic signs. Exposure of animals at the next concentration level should be delayed until there is reasonable confidence in the survival of the previously tested animals. A period of three or four days between the exposures at each concentration level is recommended to allow for the observation of delayed toxicity. The time interval may be adjusted as appropriate, *e.g.* in case of inconclusive responses.



**▼M4****Limit test**

27. The limit test is used when the test chemical is known or expected to be virtually non-toxic, i.e. eliciting a toxic response only above the regulatory limit concentration. Information about the toxicity of the test chemical can be gained from knowledge about similar tested substances or similar mixtures, taking into consideration the identity and percentage of components known to be of toxicological significance. In those situations where there is little or no information about its toxicity, or the test chemical is expected to be toxic, the main test should be performed [further guidance can be found in GD 39 (8)].
28. Using the normal procedure, three animals per sex, or six animals of the more susceptible sex, are exposed at concentrations of 20 000 ppm for gases, 20 mg/l for vapours and 5 mg/l for dusts/mists, respectively (if achievable), which serves as the limit test for this Test Method. When testing aerosols, the primary goal should be to achieve a respirable particle size (*i.e.* an MMAD of 1-4 µm). This is possible with most test chemicals at a concentration of 2 mg/l Aerosol testing at greater than 2 mg/l should only be attempted if a respirable particle size can be achieved [see GD 39 (8)]. In accordance with GHS (16), testing in excess of a limit concentration is discouraged for animal welfare reasons. Testing in GHS Category 5 (16), which is not implemented in Regulation (EC) No 1272/2008 (9), should only be considered when there is a strong likelihood that results of such a test would have direct relevance for protecting human health, and justification provided in the study report. In the case of potentially explosive test chemicals, care should be taken to avoid conditions favourable for an explosion. To avoid an unnecessary use of animals, a test run without animals should be conducted prior to the limit test to ensure that the chamber conditions for a limit test can be achieved.

**OBSERVATIONS**

29. The animals should be clinically observed frequently during the exposure period. Following exposure, clinical observations should be made at least twice on the day of exposure, or more frequently when indicated by the response of the animals to treatment, and at least once daily thereafter for a total of 14 days. The length of the observation period is not fixed, but should be determined by the nature and time of onset of clinical signs and length of the recovery period. The times at which signs of toxicity appear and disappear are important, especially if there is a tendency for signs of toxicity to be delayed. All observations are systematically recorded with individual records being maintained for each animal. Animals found in a moribund condition and animals showing severe pain and/or enduring signs of severe distress should be humanely killed for animal welfare reasons. Care should be taken when conducting examinations for clinical signs of toxicity that initial poor appearance and transient respiratory changes, resulting from the exposure procedure, are not mistaken for treatment-related effects. The principles and criteria summarised in the Humane Endpoints Guidance Document should be taken into consideration (7). When animals are killed for humane reasons or found dead, the time of death should be recorded as precisely as possible.

**▼ M4**

30. Cage-side observations should include changes in the skin and fur, eyes and mucous membranes, and also respiratory, circulatory, autonomic and central nervous systems, and somato-motor activity and behaviour patterns. When possible, any differentiation between local and systemic effects should be noted. Attention should be directed to observations of tremors, convulsions, salivation, diarrhoea, lethargy, sleep and coma. The measurement of rectal temperatures may provide supportive evidence of reflex bradypnea or hypo/hyperthermia related to treatment or confinement.

**Body weights**

31. Individual animal weights should be recorded once during the acclimatisation period, on the day of exposure prior to exposure (day 0) and at least on days 1, 3 and 7 (and weekly thereafter), and at the time of death or euthanasia if exceeding day 1. Body weight is recognised as a critical indicator of toxicity and animals exhibiting a sustained decrement of  $\geq 20\%$ , compared to pre-study values, should be closely monitored. Surviving animals are weighed and humanely killed at the end of the post-exposure period.

**Pathology**

32. All test animals, including those which die during the test or are euthanised and removed from the study for animal welfare reasons, should be subjected to gross necropsy. If necropsy cannot be performed immediately after a dead animal is discovered, the animal should be refrigerated (not frozen) at temperatures low enough to minimize autolysis. Necropsies should be performed as soon as possible, normally within a day or two. All gross pathological changes should be recorded for each animal with particular attention to any changes in the respiratory tract.
33. Additional examinations included *a priori* by design may be considered to extend the interpretive value of the study, such as measuring lung weight of surviving rats and/or providing evidence of irritation by microscope examination of the respiratory tract. Examined organs may include those showing evidence of gross pathology in animals surviving 24 or more hours, and organs known or expected to be affected. Microscopic examination of the entire respiratory tract may provide useful information for test chemicals that are reactive with water, such as acids and hygroscopic test chemicals.

**DATA AND REPORTING****Data**

34. Individual animal data on body weights and necropsy findings should be provided. Clinical observation data should be summarised in tabular form, showing for each test group the number of animals used, the number of animals displaying specific signs of toxicity, the number of animals found dead during the test or killed for humane reasons, time of death of individual animals, a description and time course of toxic effects and reversibility, and necropsy findings.

**▼ M4****Test report**

35. The test report should include the following information, as appropriate:

*Test animals and husbandry*

- Description of caging conditions, including: number (or change in number) of animals per cage, bedding material, ambient temperature and relative humidity, photoperiod, and identification of diet;
- Species/strain used and justification for using a species other than the rat;
- Number, age, and sex of animals;
- Method of randomisation;
- Details of food and water quality (including diet type/source, water source);
- Description of any pre-test conditioning including diet, quarantine, and treatment for disease;

*Test chemical*

- Physical nature, purity, and, where relevant, physico-chemical properties (including isomerisation);
- Identification data and Chemical Abstract Services (CAS) Registry Number, if known;

*Vehicle*

- Justification for use of vehicle and justification for choice of vehicle (if other than water);
- Historical or concurrent data demonstrating that the vehicle does not interfere with the outcome of the study;

*Inhalation chamber*

- Description of the inhalation chamber including dimensions and volume;
- Source and description of equipment used for the exposure of animals as well as generation of atmosphere;
- Equipment for measuring temperature, humidity, particle-size, and actual concentration;
- Source of air, treatment of air supplied/extracted and system used for conditioning;
- Methods used for calibration of equipment to ensure a homogeneous test atmosphere;
- Pressure difference (positive or negative);
- Exposure ports per chamber (nose-only); location of animals in the system (whole-body);
- Temporal homogeneity/stability of test atmosphere;
- Location of temperature and humidity sensors and sampling of test atmosphere in the chamber;
- Air flow rates, air flow rate/exposure port (nose-only), or animal load/chamber (whole-body);
- Information about the equipment used to measure oxygen and carbon dioxide, if applicable;

**▼ M4**

- Time required to reach inhalation chamber equilibrium ( $t_{95}$ );
- Number of volume changes per hour;
- Metering devices (if applicable);

*Exposure data*

- Rationale for target concentration selection in the main study;
- Nominal concentrations (total mass of test chemical generated into the inhalation chamber divided by the volume of air passed through the chamber);
- Actual test chemical concentrations collected from the animals' breathing zone; for test mixtures that produce heterogeneous physical forms (gases, vapours, aerosols), each may be analysed separately;
- All air concentrations should be reported in units of mass (e.g. mg/l, mg/m<sup>3</sup>, etc.), units of volume (e.g. ppm, ppb) may also be reported parenthetically;
- Particle size distribution, mass median aerodynamic diameter (MMAD), and geometric standard deviation ( $\sigma_g$ ), including their methods of calculation. Individual particle size analyses should be reported;

*Test conditions*

- Details of test chemical preparation, including details of any procedures used to reduce the particle size of solid substances or to prepare solutions of the test chemical. In cases where mechanical processes may have altered test chemical composition, include the results of analyses to verify the composition of the test chemical;
- A description (preferably including a diagram) of the equipment used to generate the test atmosphere and to expose the animals to the test atmosphere;
- Details of the chemical analytical method used and method validation (including efficiency of recovery of test chemical from the sampling medium);
- The rationale for the selection of test concentrations;

*Results*

- Tabulation of chamber temperature, humidity, and airflow;
- Tabulation of chamber nominal and actual concentration data;
- Tabulation of particle size data including analytical sample collection data, particle size distribution, and calculations of the MMAD and  $\sigma_g$ ;
- Tabulation of response data and concentration level for each animal (i.e. animals showing signs of toxicity including mortality, nature, severity, and duration of effects);
- Individual body weights of animals collected on study days, date and time of death if prior to scheduled euthanasia; time course of onset of signs of toxicity, and whether these were reversible for each animal;

**▼M4**

- Necropsy findings and histopathological findings for each animal, if available;
- The CLP category classification and the LC<sub>50</sub> cut-off value;

*Discussion and interpretation of results*

- Particular emphasis should be made to the description of methods used to meet this Test Method's criteria, e.g. the limit concentration or the particle size;
- The respirability of particles in light of the overall findings should be addressed, especially if the particle-size criteria could not be met;
- The consistency of methods used to determine nominal and actual concentrations, and the relation of actual concentration to nominal concentration should be included in the overall assessment of the study;
- The likely cause of death and predominant mode of action (systemic versus local) should be addressed;
- An explanation should be provided if there was a need to humanely sacrifice animals in pain or showing signs of severe and enduring distress, based on the criteria in the OECD Guidance Document on Humane Endpoints (7).

*LITERATURE:*

- (1) Chapter B.2 of this Annex, Acute Toxicity (Inhalation).
- (2) Holzhütter H-G, Genschow E, Diener W, and Schlede E (2003). Dermal and Inhalation Acute Toxicity Class Methods: Test Procedures and Biometric Evaluations for the Globally Harmonized Classification System. *Arch. Toxicol.* 77: 243-254.
- (3) Diener W, Kayser D and Schlede E (1997). The Inhalation Acute-Toxic-Class Method; Test Procedures and Biometric Evaluations. *Arch. Toxicol.* 71: 537-549.
- (4) Diener W and Schlede E (1999). Acute Toxic Class Methods: Alternatives to LD/LC<sub>50</sub> Tests. *ALTEX* 1: 129-134.
- (5) Chapter B.1 tris of this Annex, Acute Oral Toxicity — Acute Toxic Class Method.
- (6) OECD (2009). Report on Biostatistical Performance Assessment of the Draft TG 436 Acute Toxic Class Testing Method for Acute Inhalation Toxicity. Environmental Health and Safety Monograph Series on Testing and Assessment No 105, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (7) OECD (2000). Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation. Environmental Health and Safety Monograph Series on Testing and Assessment No 19. Available at: [<http://www.oecd.org/env/testguidelines>]
- (8) OECD (2009). Guidance Document on Acute Inhalation Toxicity Testing. Environmental Health and Safety Monograph Series on Testing and Assessment No 39, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (9) Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006 (OJ L 353, 31.12.2008, p. 1).

**▼M4**

- (10) Chapter B.40 of this Annex, In Vitro Skin Corrosion: Transcutaneous Electrical Resistance Test (TER).
- (11) Chapter B.40 bis of this Annex, In Vitro Skin Corrosion: Human Skin Model Test.
- (12) OECD (2005). In Vitro Membrane Barrier Test Method for Skin Corrosion. OECD Guideline for testing of chemicals No 435, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
- (13) Phalen RF (2009). Inhalation Studies: Foundations and Techniques. (2<sup>nd</sup> Edition) Informa Healthcare, New York.
- (14) SOT (1992). Technical Committee of the Inhalation Specialty Section, Society of Toxicology (SOT). Recommendations for the Conduct of Acute Inhalation Limit Tests. *Fund. Appl. Toxicol.* 18: 321-327.
- (15) Pauluhn J and Thiel A (2007). A Simple Approach to Validation of Directed-Flow Nose-Only Inhalation Chambers. *J. Appl. Toxicol.* 27: 160-167
- (16) UN (2007), United Nations Globally Harmonized System of Classification and Labelling of Chemicals (GHS), ST/SG/AC.10/30, UN New York and Geneva. Available: [[http://www.unece.org/trans/danger/publi/ghs/ghs\\_welcome\\_e.html](http://www.unece.org/trans/danger/publi/ghs/ghs_welcome_e.html)]

▼ M4

*Appendix 1*

DEFINITION

**Test chemical:** Any substance or mixture tested using this Test Method.

**▼ M4***Appendix 2***Procedure to be followed by each of the starting concentrations for gases (ppm/4h)**General remarks <sup>(1)</sup>

For each starting concentration, the respective testing schemes as included in this Appendix outline the procedure to be followed.

Appendix 2a: Starting concentration is 100 ppm

Appendix 2b: Starting concentration is 500 ppm

Appendix 2c: Starting concentration is 2 500 ppm

Appendix 2d: Starting concentration is 20 000 ppm

Depending on the number of humanely killed or dead animals, the test procedure follows the indicated arrows.

---

<sup>(1)</sup> In the following tables reference is made to GHS (Globally Harmonised System of Classification and Labelling of Chemicals (GHS). The EU equivalent is Regulation (EC) No 1272/2008. In the case of Acute Inhalation Toxicity, the Regulation (EC) No 1272/2008 (9) does not implement Category 5.

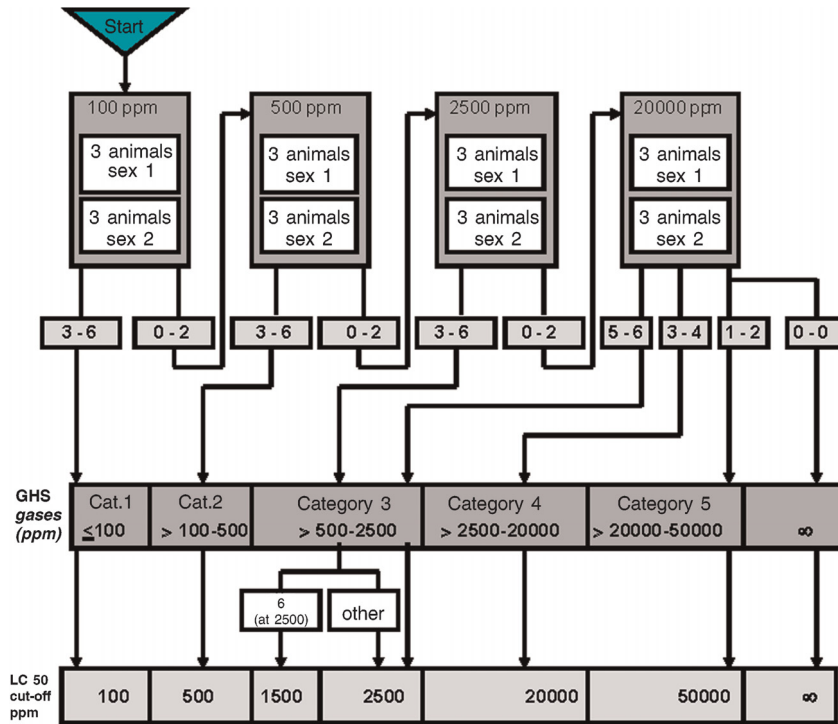


▼ **M4**

Appendix 2a

**Acute Inhalation Toxicity:**

**Test Procedure with a starting concentration of 100 ppm/4h for gases**



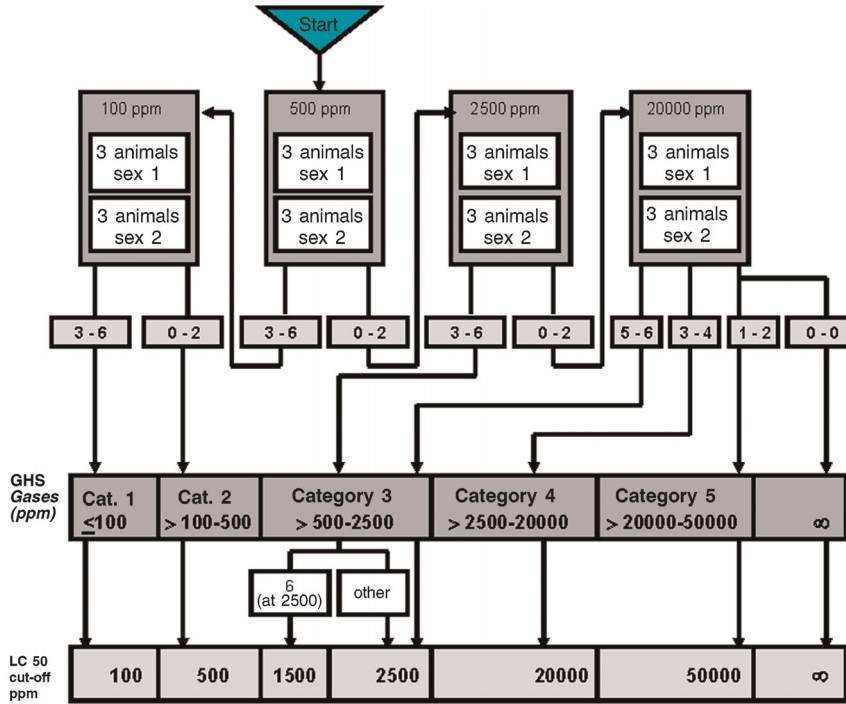
- 3 ♂ + 3 ♀, or 6 animals of the more susceptible sex are used per step
- 0-6: Number of moribund or dead animals/tested concentration
- GHS: Globally Harmonized Classification System
- ∞: unclassified
- Testing at ≥ 20000 ppm/4h: see Guidance Document 39 (8)

▼ **M4**

Appendix 2b

**Acute Inhalation Toxicity:**

**Test Procedure with a starting concentration of 500 ppm/4h for gases**



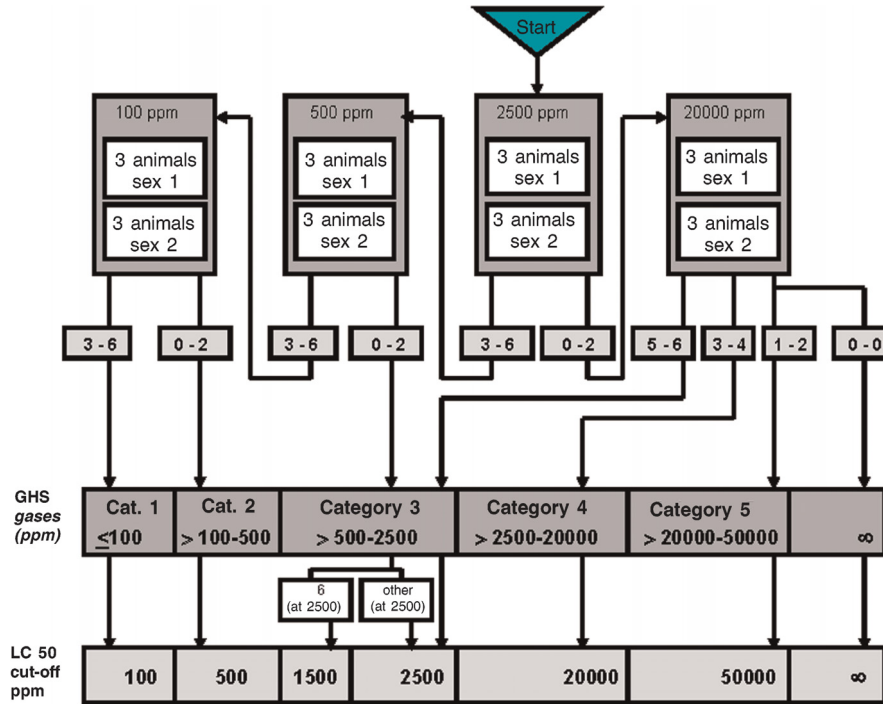
- 3  $\sigma$  + 3  $\phi$ , or 6 animals of the more susceptible sex are used per step
- 0-6: Number of moribund or dead animals/tested concentraion
- GHS: Globally Harmonized Classification System
- $\infty$ : unclassified
- Testing at  $\geq 20000$  ppm/4h: see Guidance Document 39 (8)

▼ **M4**

Appendix 2c

**Acute Inhalation Toxicity:**

**Test Procedure with a starting concentration of 2 500 ppm/4h for gases**



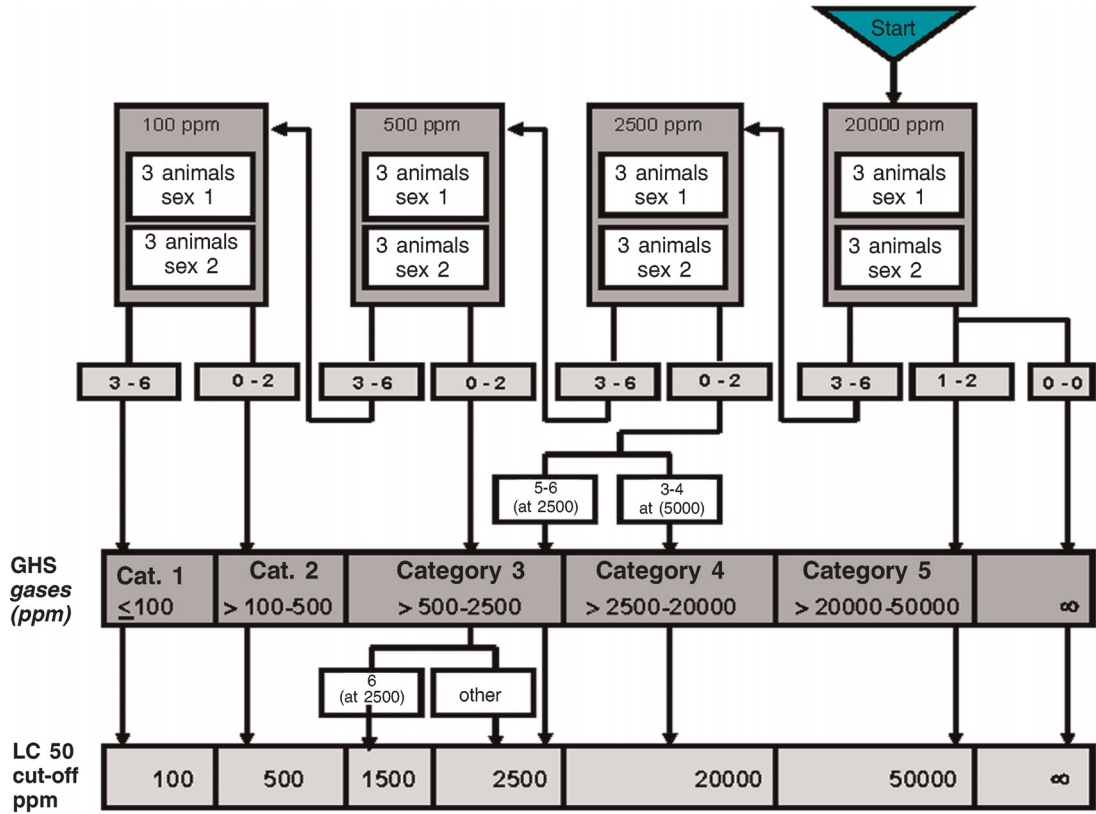
- 3 ♂ + 3 ♀, or 6 animals of the more susceptible sex are used per step
- 0-6: Number of moribund or dead animals/tested concentration
- GHS: Globally Harmonized Classification System
- ∞: unclassified
- Testing at ≥ 20000 ppm/4h: see Guidance Document 39 (8)

▼ **M4**

Appendix 2d

**Acute Inhalation Toxicity:**

**Test Procedure with a starting concentration of 20 000 ppm/4h for gases**



- 3 ♂ + 3 ♀, or 6 animals of the more susceptible sex are used per step
- 0-6: Number of moribund or dead animals/tested concentration
- GHS: Globally Harmonized Classification System
- ∞: unclassified
- Testing at ≥ 20 000 ppm/4h: see Guidance Document 39 (8)

**▼ M4***Appendix 3***Procedure to be followed by each of the starting concentrations for vapour (mg/l/4h)**General remarks <sup>(1)</sup>

For each starting concentration, the respective testing schemes as included in this Appendix outline the procedure to be followed.

Appendix 3a: Starting concentration is 0,5 mg/l

Appendix 3b: Starting concentration is 2,0 mg/l

Appendix 3c: Starting concentration is 10 mg/l

Appendix 3d: Starting concentration is 20 mg/l

Depending on the number of humanely killed or dead animals, the test procedure follows the indicated arrows.

---

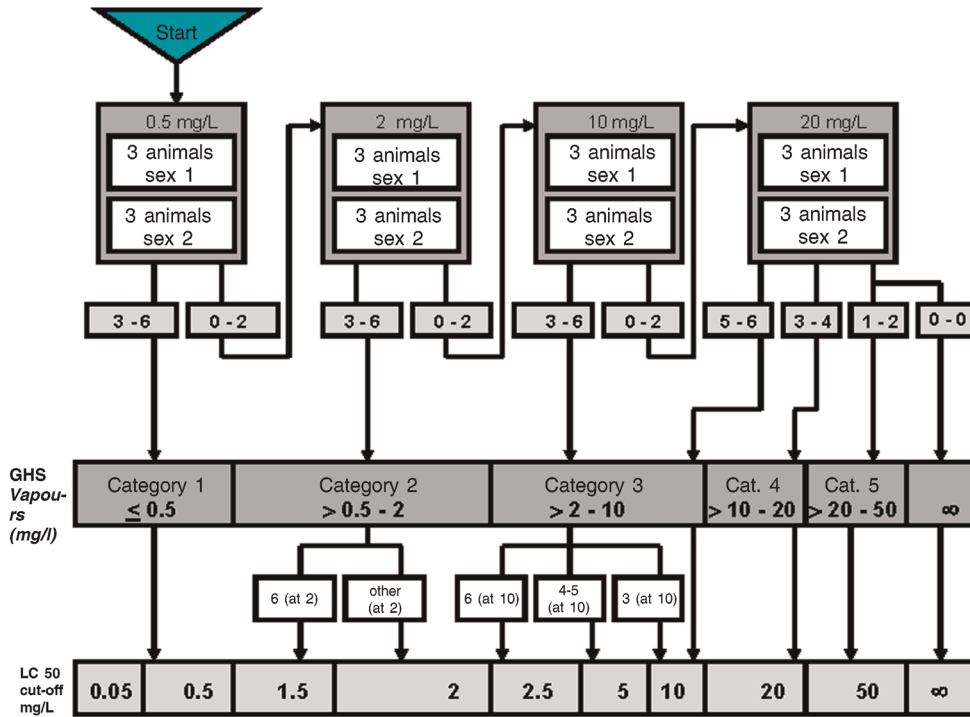
<sup>(1)</sup> In the following tables reference is made to GHS (Globally Harmonised System of Classification and Labelling of Chemicals (GHS). The EU equivalent is Regulation (EC) No 1272/2008. In the case of Acute Inhalation Toxicity, the Regulation (EC) No 1272/2008 (9) does not implement Category 5.

▼ **M4**

Appendix 3a

Acute Inhalation Toxicity:

Test procedure with a starting concentration of 0,5 mg/L/4h for vapours



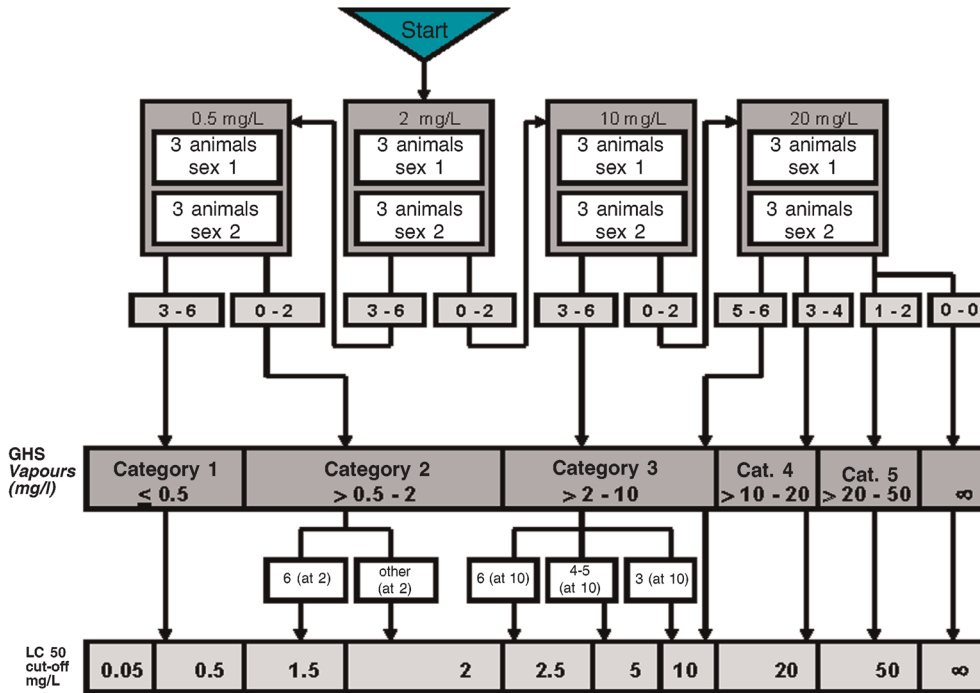
- 3 ♂ + 3 ♀, or 6 animals of the more susceptible sex are used per step
- 0-6: Number of moribund or dead animals/tested concentration
- GHS: Globally Harmonized Classification System
- ∞: unclassified
- Testing at 50 mg/L/4h: see Guidance Document 39 (8)

▼ **M4**

Appendix 3b

Acute Inhalation Toxicity:

Test procedure with a starting concentration of 2 mg/L/4h for vapours



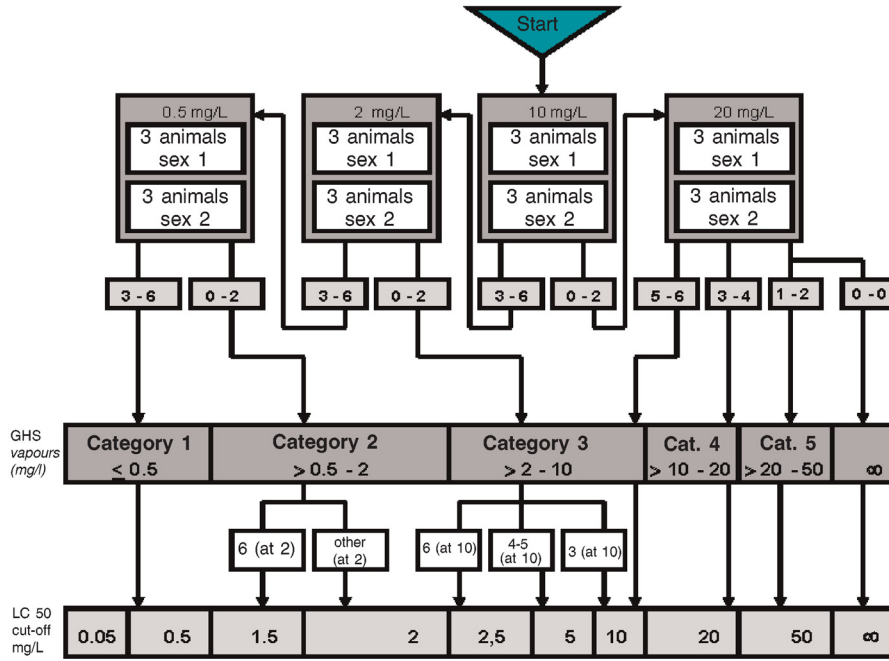
- 3 ♂ + 3 ♀, or 6 animals of the more susceptible sex are used per step
- 0-6: Number of moribund or dead animals/tested concentration
- GHS: Globally Harmonized Classification System
- ∞: unclassified
- Testing at 50 mg/L/4h: see Guidance Document 39 (8)

▼ **M4**

Appendix 3c

Acute Inhalation Toxicity:

Test procedure with a starting concentration of 10 mg/L/4h for vapours



- 3 ♂ + 3 ♀, or 6 animals of the more susceptible sex are used per step
- 0-6: Number of moribund or dead animals/tested concentration
- GHS: Globally Harmonized Classification System
- ∞: unclassified
- Testing at 50 mg/L/4h: see Guidance Document 39 (8)

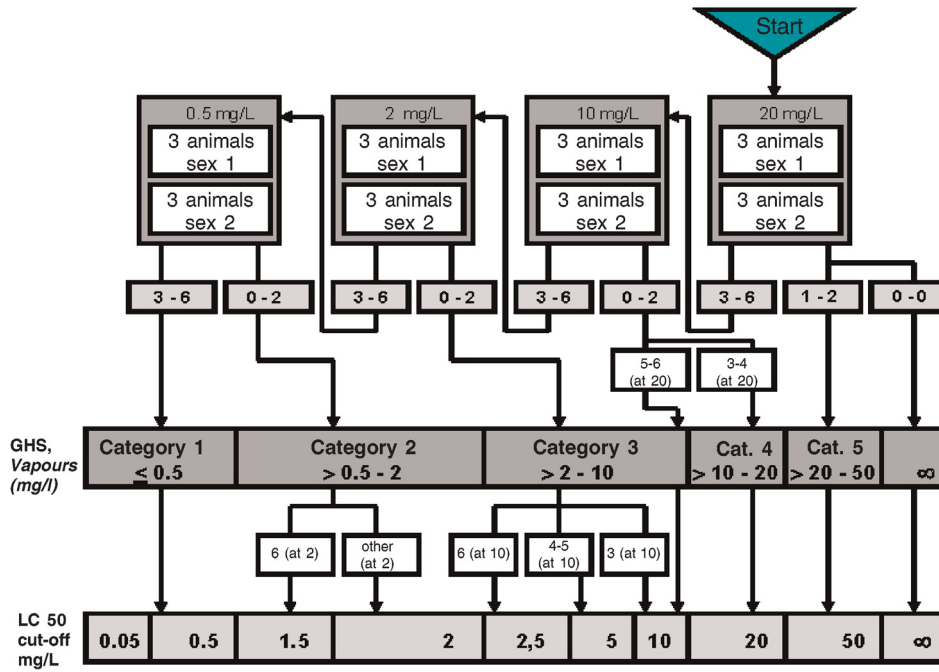


▼ **M4**

Appendix 3d

Acute Inhalation Toxicity:

Test procedure with a starting concentration of 20 mg/L/4h for vapours



- 3 ♂ + 3 ♀, or 6 animals of the more susceptible sex are used per step
- 0-6: Number of moribund or dead animals/tested concentration
- GHS: Globally Harmonized Classification System
- ∞: unclassified
- Testing at 50 mg/L/4h: see Guidance Document 39(8)

**▼ M4***Appendix 4***Procedure to be followed by each of the starting concentrations for aerosols (mg/l/4h)**General remarks <sup>(1)</sup>

For each starting concentration, the respective testing schemes as included in this Appendix outline the procedure to be followed.

Appendix 4a: Starting concentration is 0,05 mg/l

Appendix 4b: Starting concentration is 0,5 mg/l

Appendix 4c: Starting concentration is 1 mg/l

Appendix 4d: Starting concentration is 5 mg/l

Depending on the number of humanely killed or dead animals, the test procedure follows the indicated arrows.

---

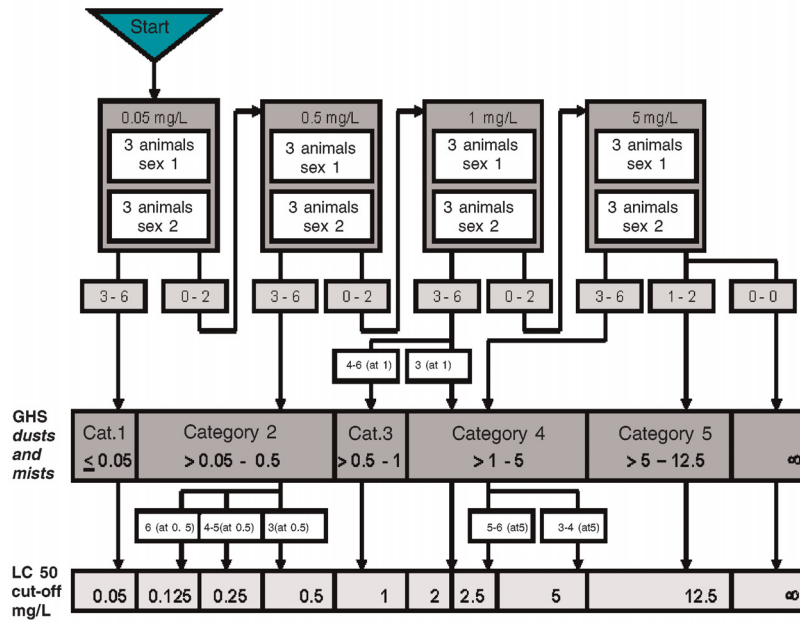
<sup>(1)</sup> In the following tables reference is made to GHS (Globally Harmonised System of Classification and Labelling of Chemicals (GHS). The EU equivalent is Regulation (EC) No 1272/2008. In the case of Acute Inhalation Toxicity the Regulation (EC) No 1272/2008 (9) does not implement Category 5.

▼ **M4**

Appendix 4a

**Acute Inhalation Toxicity:**

Test procedure with a starting concentration of 0,05 mg/L/4h for aerosols



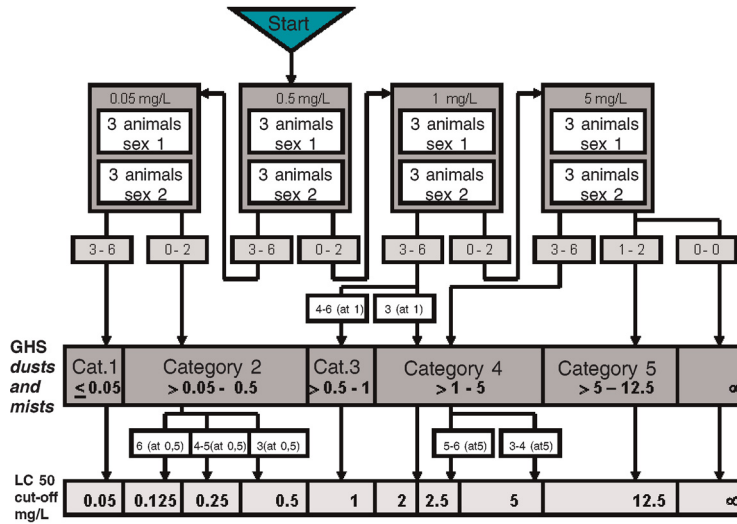
- 3 ♂ + 3 ♀, or 6 animals of the more susceptible sex are used per step
- 0-6: Number of moribund or dead animals/tested concentration
- GHS: Globally Harmonized Classification System
- ∞: unclassified
- Testing at 12.5 mg/L/4h: see Guidance Document 39 (8)

▼ **M4**

Appendix 4b

**Acute Inhalation Toxicity:**

Test procedure with a starting concentration of 0,5 mg/L/4h for aerosols



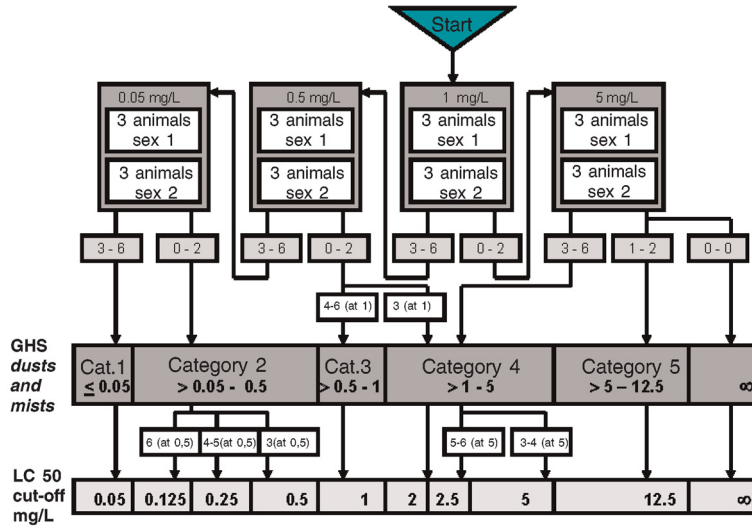
- 3 ♂ + 3 ♀, or 6 animals of the more susceptible sex are used per step
- 0-6: Number of moribund or dead animals/tested concentraion
- GHS: Globally Harmonized Classification System
- ∞: unclassified
- Testing at 12.5 mg/L/4h: see Guidance Document 39 (8)

▼ **M4**

Appendix 4c

**Acute Inhalation Toxicity:**

**Test procedure with a starting concentration of 1 mg/L/4h for aerosols**



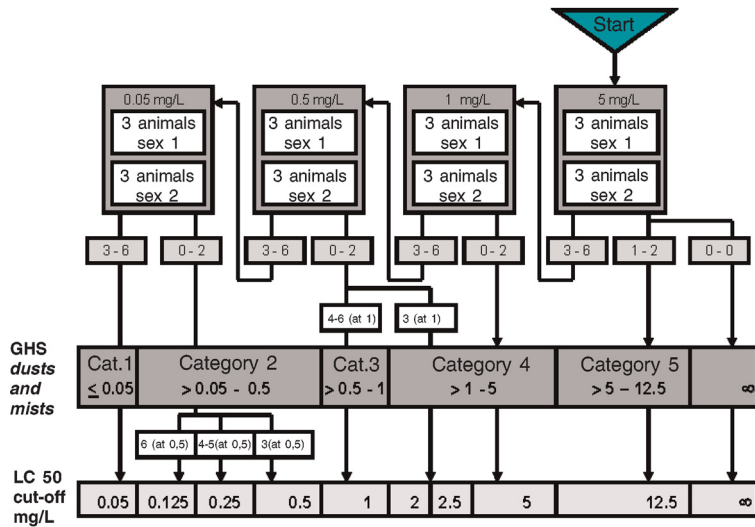
- 3 ♂ + 3 ♀, or 6 animals of the more susceptible sex are used per step
- 0-6: Number of moribund or dead animals/tested concentration
- GHS: Globally Harmonized Classification System
- ∞: unclassified
- Testing at 12.5 mg/L/4h: see Guidance Document 39 (8)

▼ **M4**

*Appendix 4d*

**Acute Inhalation Toxicity:**

**Test procedure with a starting concentration of 5 mg/L/4h for aerosols**



- 3 ♂ + 3 ♀, or 6 animals of the more susceptible sex are used per step
- 0-6 : Number of moribund or dead animals/tested concentraion
- GHS: Globally Harmonized Classification System
- ∞: unclassified
- Testing at 12.5 mg/L/4h: see Guidance Document 39 (8)

▼ M5**B.53 DEVELOPMENTAL NEUROTOXICITY STUDY**

## INTRODUCTION

1. This test method is equivalent to OECD Test Guideline (TG) 426 (2007). In Copenhagen in June 1995, an OECD Working Group on Reproduction and Developmental Toxicity discussed the need to update existing OECD test guidelines for reproduction and developmental toxicity, and the development of new guidelines for endpoints not yet covered (1). The working group recommended that a test guideline for developmental neurotoxicity should be written based on a US EPA guideline, which has since been revised (2). In June 1996, a second consultation meeting was held in Copenhagen to provide the Secretariat with guidance on the outline of a new test guideline on developmental neurotoxicity, including the major elements, e.g. details concerning choice of animal species, dosing period, testing period, endpoints to be assessed, and criteria for evaluating results. A US neurotoxicity risk assessment guideline was published in 1998 (3). An OECD Expert Consultation Meeting and an ILSI Risk Science Institute Workshop were held back-to-back in October 2000 and an expert consultation meeting was held in Tokyo 2005. These meetings were held to discuss the scientific and technical issues related to the current test guideline and the recommendations from the meetings (4)(5)(6)(7) were considered in the development of this test method. Additional information on the conduct, interpretation and terminology used for this test method can be found in OECD Guidance Documents No 43 on 'Reproductive Toxicity Testing and Assessment' (8) and No 20 on 'Neurotoxicity Testing' (9).

## INITIAL CONSIDERATIONS

2. A number of chemicals is known to produce developmental neurotoxic effects in humans and other species (10)(11)(12)(13). Determination of the potential for developmental neurotoxicity may be needed to assess and evaluate the toxic characteristics of a chemical. Developmental neurotoxicity studies are designed to provide data, including dose-response characterisations, on the potential functional and morphological effects on the developing nervous system of the offspring that may arise from exposure in utero and during early life.
3. A developmental neurotoxicity study can be conducted as a separate study, incorporated into a reproductive toxicity and/or adult neurotoxicity study (e.g. test methods B.34 (14), B.35 (15), B.43 (16)), or added onto a prenatal developmental toxicity study (e.g. test method B.31 (17)). When the developmental neurotoxicity study is incorporated within or attached to another study, it is imperative to preserve the integrity of both study types. All testing should comply with applicable legislation or government and institutional guidelines for the use of laboratory animals in research (e.g. 18).
4. The testing laboratory should consider all available information on the test chemical prior to conducting the study. Such information will include the identity and structure of the chemical; its physico-chemical properties; the results of any other *in vitro* or *in vivo* toxicity tests on the chemical; toxicological data on structurally related chemicals; and the anticipated use(s) of the chemical. This information is necessary to satisfy all concerned that the test is relevant for the protection of human health, and will help in the selection of an appropriate starting dose.

**▼ M5**

## PRINCIPLE OF THE TEST

5. The test chemical is administered to animals during gestation and lactation. Dams are tested to assess effects in pregnant and lactating females and may also provide comparative information (dams versus offspring). Offspring are randomly selected from within litters for neurotoxicity evaluation. The evaluation consists of observations to detect gross neurologic and behavioural abnormalities, including the assessment of physical development, behavioural ontogeny, motor activity, motor and sensory function, and learning and memory; and the evaluation of brain weights and neuropathology during postnatal development and adulthood.
  
6. When the test method is conducted as a separate study, additional available animals in each group could be used for specific neurobehavioral, neuropathological, neurochemical or electrophysiological procedures that may supplement the data obtained from the examinations recommended by this test method (16)(19)(20)(21). The supplemental procedures can be particularly useful when empirical observation, anticipated effects, or mechanism/mode-of-action indicate a specific type of neurotoxicity. These supplemental procedures may be used in the dams as well as in the pups. In addition, *ex vivo* or *in vitro* procedures may also be used, as long as these procedures do not alter the integrity of the *in vivo* procedures.

## PREPARATIONS FOR THE TEST

**Selection of animal species**

7. The preferred test species is the rat; other species can be used when appropriate. Note, however, the gestational and postnatal days specified in this test method are specific to commonly used strains of rats, and comparable days should be selected if a different species or unusual strain is used. The use of another species should be justified based on toxicological, pharmacokinetic, and/or other data. Justification should include availability of species-specific postnatal neurobehavioral and neuropathological assessments. If there was an earlier test that raised concerns, the species/strain that raised a concern should be considered. Because of the differing performance attributes of different rat strains, there should be evidence that the strain selected for use has adequate fecundity and responsiveness. The reliability and sensitivity of other species to detect developmental neurotoxicity should be documented.

**Housing and feeding conditions**

8. The temperature in the experimental animal room should be  $22 \pm 3$  °C. Although the relative humidity should be at least 30 % and preferably not exceed 70 % other than during room cleaning, the aim should be 50-60 %. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. It is also possible to reverse the light cycle prior to mating and for the duration of the study, in order to perform the assessments of functional and behavioural endpoints during the dark period (under red light), i.e. during the time the animals are normally active (22). Any changes in the light-dark cycle should include adequate acclimation time to allow animals to adapt to the new cycle. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water. The type of food and water should be reported and both should be analysed for contaminants.



**▼ M5**

9. Animals may be housed individually or be caged in small groups of the same sex. Mating procedures should be carried out in cages suitable for the purpose. After evidence of copulation or no later than day 15 of pregnancy, mated animals should be caged separately in delivery or maternity cages. Cages should be arranged in such a way that possible effects due to cage placement are minimised. Mated females should be provided with appropriate and defined nesting materials when parturition is near. It is well known that inappropriate handling or stress during pregnancy can result in adverse outcomes, including prenatal loss and altered foetal and postnatal development. To guard against foetal loss from factors which are not treatment-related, animals should be carefully handled during pregnancy, and stress from outside factors such as excessive outside noise should be avoided.

**Preparation of the animals**

10. Healthy animals should be used, which have been acclimated to laboratory conditions and have not been subjected to previous experimental procedures, unless the study is incorporated in another study (see paragraph 3). The test animals should be characterised as to species, strain, source, sex, weight and age. Each animal should be assigned and marked with a unique identification number. The animals of all test groups should, as nearly as practicable, be of uniform weight and age, and should be within the normal range of the species and strain under study. Young adult nulliparous female animals should be used at each dose level. Siblings should not be mated, and care should be taken to ensure this. Gestation Day (GD) 0 is the day on which a vaginal plug and/or sperm are observed. Adequate acclimation time (e.g. 2-3 days) should be allowed when purchasing time-pregnant animals from a supplier. Mated females should be assigned in an unbiased way to the control and treatment groups, and as far as possible, they should be evenly distributed among the groups (e.g. a stratified random procedure is recommended to provide even distribution among all groups, such as that based on body weight). Females inseminated by the same male should be equalised across groups.

**PROCEDURE****Number and sex of animals**

11. Each test and control group should contain a sufficient number of pregnant females to be exposed to the test chemical to ensure that an adequate number of offspring are produced for neurotoxicity evaluation. A total of 20 litters are recommended at each dose level. Replicate and staggered-group dosing designs are allowed if total numbers of litters per group are achieved, and appropriate statistical models are used to account for replicates.
12. On or before postnatal day (PND) 4 (day of delivery is PND 0), the size of each litter should be adjusted by eliminating extra pups by random selection to yield a uniform litter size for all litters (23). The litter size should not exceed the average litter size for the strain of rodents used (8-12). The litter should have, as nearly as possible, equal numbers of male and female pups. Selective elimination of pups, e.g. based upon body weight, is not appropriate. After standardisation of litters (culling) and prior to further testing of functional endpoints, individual pups that are scheduled for pre-weaning or post-weaning testing should be identified uniquely, using any suitable humane method for pup identification (e.g. 24).

▼ **M5****Assignment of animals for functional and behavioural tests, brain weights, and neuropathological evaluations**

13. The test method allows various approaches with respect to the assignment of animals exposed *in utero* and through lactation to functional and behavioural tests, sexual maturation, brain weight determination, and neuropathological evaluation (25). Other tests of neurobehavioral function (e.g. social behaviour), neurochemistry or neuropathology can be added on a case-by-case basis, as long as the integrity of the original required tests are not compromised.
14. Pups are selected from each dose group and assigned for endpoint assessments on or after PND 4. Selection of pups should be performed so that to the extent possible both sexes from each litter in each dose group are equally represented in all tests. For motor activity testing the same pair of male and female pups should be tested at all pre-weaning ages (see paragraph 35). For all other tests the same or separate pairs of male and female animals may be assigned to different behavioural tests. Different pups may need to be assigned to weanling versus adult tests of cognitive function in order to avoid confounding the effects of age and prior training on these measurements (26)(27). At weaning (PND 21), pups not selected for testing can be disposed of humanely. Any alterations in pup assignments should be reported. The statistical unit of measure should be the litter (or dam) and not the pup.
15. There are different ways to assign pups to the pre-weaning and post-weaning examinations, cognitive tests, pathological examinations, etc., (see Figure 1 for general design and Appendix 1 for examples of assignment). Recommended minimum numbers of animals in each dose group for pre-weaning and post-weaning examinations are as follows:

Clinical observations and bodyweight	All animals
Detailed clinical observations	20/sex (1/sex/litter)
Brain weight (post fixation) PND 11-22	10/sex (1/litter)
Brain weight (unfixed) ~ PND 70	10/sex (1/litter)
Neuropathology (immersion or perfusion fixation) PND 11-22	10/sex (1/litter)
Neuropathology (perfusion fixation) PND ~ 70	10/sex (1/litter)
Sexual maturation	20/sex (1/sex/litter)
Other developmental landmarks (optional)	All animals
Behavioural ontogeny	20/sex (1/sex/litter)
Motor activity	20/sex (1/sex/litter)
Motor and sensory function	20/sex (1/sex/litter)
Learning and memory	10/sex <sup>(a)</sup> (1/litter)

<sup>(a)</sup> Depending on the sensitivity of cognitive function tests, investigation of a large higher number of animals should be considered e.g. up to 1 male and 1 female per litter (for animal assignments see Appendix 1) (further guidance on sample size is provided in the OECD Guidance Document 43 (8)).

**▼ M5****Dosage**

16. At least three dose levels and a concurrent control should be used. The dose levels should be spaced to produce a gradation of toxic effects. Unless limited by the physico-chemical nature or biological properties of the chemical, the highest dose level should be chosen with the aim to induce some maternal toxicity (e.g. clinical signs, decreased body weight gain (not more than 10 %) and/or evidence of dose-limiting toxicity in a target organ). The high dose may be limited to 1 000 mg/kg/day body weight, with some exceptions. For example, expected human exposure may indicate the need for a higher dose level to be used. Alternatively, pilot studies or preliminary range-finding studies should be performed to determine the highest dosage to be used which should produce a minimal degree of maternal toxicity. If the test chemical has been shown to be developmentally toxic either in a standard developmental toxicity study or in a pilot study, the highest dose level should be the maximum dose which will not induce excessive offspring toxicity, or *in utero* or neonatal death or malformations, sufficient to preclude a meaningful evaluation of neurotoxicity. The lowest dose level should aim to not produce any evidence of either maternal or developmental toxicity including neurotoxicity. A descending sequence of dose levels should be selected with a view to demonstrating any dose-related response and a No-Observed-Adverse Effect Level (NOAEL), or doses near the limit of detection that would allow the determination of a benchmark dose. Two- to four-fold intervals are frequently optimal for setting the descending dose levels, and the addition of a fourth dose group is often preferable to using very large intervals (e.g. more than a factor of 10) between dosages.
  
17. Dose levels should be selected taking into account all existing toxicity data as well as additional information on metabolism and toxicokinetics of the test chemical or related materials. This information may also assist in demonstrating the adequacy of the dosing regimen. Direct dosing of pups should be considered based on exposure and pharmacokinetic information (28)(29). Careful consideration of benefits and disadvantages should be made prior to conducting direct dosing studies (30).
  
18. The concurrent control group should be a sham-treated control group or a vehicle-control group if a vehicle is used in administering the test chemical. All animals should normally be administered the same volume of either test chemical or vehicle on a body weight basis. If a vehicle or other additive is used to facilitate dosing, consideration should be given to the following characteristics: effects on the absorption, distribution, metabolism, or retention of the test chemical; effects on the chemical properties of the test chemical which may alter its toxic characteristics; and effects on the food or water consumption or the nutritional status of the animals. The vehicle should not cause effects that could interfere with the interpretation of the study neither should it be neurobehaviourally toxic nor have effects on reproduction or development. For novel vehicles, a sham-treated control group should be included in addition to a vehicle control group. Animals in the control group(s) should be handled in an identical manner to test group animals.

**▼ M5****Administration of doses**

19. The test chemical or vehicle should be administered by the route most relevant to potential human exposure, and based on available metabolism and distribution information in the test animals. The route of administration will generally be oral (e.g. gavage, dietary, via drinking water), but other routes (e.g. dermal, inhalation) may be used depending on the characteristics and anticipated or known human exposure routes (further guidance is provided in the Guidance Document 43(8)). Justification should be provided for the route of administration chosen. The test chemical should be administered at approximately the same time every day.
  
20. The dose administered to each animal should normally be based on the most recent individual body weight determination. However, caution should be exercised when adjusting the doses during the last third of pregnancy. If excess toxicity is noted in the treated dams, those animals should be humanely killed.
  
21. The test chemical or vehicle should, as a minimum, be administered daily to mated females from the time of implantation (GD 6) throughout lactation (PND 21), so that the pups are exposed to the test chemical during pre- and postnatal neurological development. The age at which dosing starts, and the duration and frequency of dosing, may be adjusted if evidence supports an experimental design more relevant to human exposures. Dosing durations should be adjusted for other species to ensure exposure during all early periods of brain development (i.e. equivalent to prenatal and early postnatal human brain growth). Dosing may begin from the initiation of pregnancy (GD 0) although consideration should be given to the potential of the test chemical to cause pre-implantation loss. Administration beginning at GD 6 would avoid this risk, but the developmental stages between GD 0 and 6 would not be treated. When a laboratory purchases time-mated animals, it is impractical to begin dosing at GD 0, and thus GD 6 would be a good starting day. The testing laboratory should set the dosing regimen according to relevant information about the effects of the test chemical, prior experience, and logistical considerations; this may include extension of dosing past weaning. Dosing should not occur on the day of parturition in those animals which have not completely delivered their offspring. In general, it is assumed that exposure of the pups will occur through the maternal milk; however, direct dosing of pups should be considered in those cases where there is a lack of evidence of continued exposure to offspring. Evidence of continuous exposure can be retrieved from e.g. pharmacokinetic information, offspring toxicity or changes in bio-markers (28).

**OBSERVATIONS****Observations on dams**

22. All dams should be carefully observed at least once daily with respect to their health condition, including morbidity and mortality.
  
23. During the treatment and observation periods, more detailed clinical observations should be conducted periodically (at least twice during the gestational dosing period and twice during the lactational dosing period) using at least 10 dams per dose level. The animals should be observed outside the home cage by trained technicians who are unaware of the animals' treatment, using standardised procedures to minimise animal stress and observer bias, and maximise inter-observer reliability. Where possible, it is advisable that the observations in a given study be made by the same technician.

**▼ M5**

24. The presence of observed signs should be recorded. Whenever feasible, the magnitude of the observed signs should also be recorded. Clinical observations should include, but not be limited to, changes in skin, fur, eyes, mucous membranes, occurrence of secretions, and autonomic activity (e.g. lacrimation, piloerection, pupil size, unusual respiratory pattern and/or mouth breathing, and any unusual signs of urination or defecation).
25. Any unusual responses with respect to body position, activity level (e.g. decreased or increased exploration of the standard area) and co-ordination of movement should also be noted. Changes in gait, (e.g. waddling, ataxia), posture (e.g. hunched-back) and reactivity to handling, placing or other environmental stimuli, as well as the presence of clonic or tonic movements, convulsions, tremors, stereotypies (e.g. excessive grooming, unusual head movements, repetitive circling), bizarre behaviour (e.g. biting or excessive licking, self-mutilation, walking backwards, vocalisation), or aggression should be recorded.
26. Signs of toxicity should be recorded, including the day of onset, time of day, degree, and duration.
27. Animals should be weighed at the time of dosing at least weekly throughout the study, on or near the day of delivery, and on PND 21 (weaning). For gavage studies dams should be weighed at least twice weekly. Doses should be adjusted at the time of each body weight determination, as appropriate. Food consumption should be measured weekly at a minimum during gestation and lactation. Water consumption should be measured at least weekly if exposure is via the water supply.

**Observations on offspring**

28. All offspring should be carefully observed at least daily for signs of toxicity and for morbidity and mortality.
29. During the treatment and observation periods, more detailed clinical observations of the offspring should be conducted. The offspring (at least one pup/sex/litter) should be observed by trained technicians who are unaware of the animals' treatment, using standardised procedures to minimise bias and maximise inter-observer reliability. Where possible, it is advisable that the observations are made by the same technician. At a minimum, the endpoints described in paragraphs 24 and 25 should be monitored as appropriate for the developmental stage being observed.
30. All signs of toxicity in the offspring should be recorded, including the day of onset, time of day, degree, and duration.

**Physical and developmental landmarks**

31. Changes in pre-weaning landmarks of development (e.g. pinna unfolding, eye opening, incisor eruption) are highly correlated with body weight (30)(31). Body weight may be the best indicator of physical development. Measurement of developmental landmarks is, therefore, recommended only when there is prior evidence that these endpoints will provide additional information. Timing for the assessment of these parameters is indicated in Table 1. Depending on the anticipated effects, and the results of the initial measurements, it may be advisable to add additional time points or to perform the measurements in other developmental stages.

## ▼ M5

32. It is advisable to use post-coital age instead of postnatal age when assessing physical development (33). If pups are tested on the day of weaning, it is recommended that this testing be carried out prior to actual weaning to avoid a confounding effect by the stress associated with weaning. In addition, any post-weaning testing of pups should not occur during the two days after weaning.

Table 1

**Timing of the assessment of physical and developmental landmarks, and functional/behavioural endpoints <sup>(a)</sup>**

Age Periods Endpoints	Pre-weaning <sup>(b)</sup>	Adolescence <sup>(b)</sup>	Young adults <sup>(b)</sup>
--------------------------	----------------------------	----------------------------	-----------------------------

**Physical and developmental landmarks**

Body weight and Clinical Observations	weekly <sup>(c)</sup>	at least every two weeks	at least every two weeks
Brain weight	PND 22 <sup>(d)</sup>		at termination
Neuropathology	PND 22 <sup>(d)</sup>		at termination
Sexual maturation	—	as appropriate	—
Other developmental landmarks <sup>(e)</sup>	as appropriate	—	—

**Functional/behavioural endpoints**

Behavioural ontogeny	At least two measures		
Motor activity (including habituation)	1–3 times <sup>(f)</sup>	—	once
Motor and sensory function	—	once	once
Learning and memory	—	once	once

<sup>(a)</sup> This table presents the minimum number of times when measurements should be performed. Depending on the anticipated effects, and the results of the initial measurements, it may be advisable to add additional time points (e.g. aged animals) or to perform the measurements in other developmental stages.

<sup>(b)</sup> It is recommended that pups not be tested during the two days after weaning (see paragraph 32). Recommended ages for adolescent testing are: learning and memory = PND 25 ± 2; motor and sensory function = PND 25 ± 2. Recommended ages for testing young adults is PND 60–70.

<sup>(c)</sup> Body weights should be measured at least twice weekly when directly dosing pups for adjustment of doses at a time of rapid body weight gain.

<sup>(d)</sup> Brain weights and neuropathology may be assessed at some earlier time (e.g. PND 11), if appropriate (see paragraph 39).

<sup>(e)</sup> Other developmental landmarks in addition to the body weight (e.g. eye opening) should be recorded when appropriate (see paragraph 31).

<sup>(f)</sup> See paragraph 35.

33. Live pups should be counted and sexed e.g. by visual inspection or measurement of anogenital distance (34)(35), and each pup within a litter should be weighed individually at birth or soon thereafter, at least weekly throughout lactation, and at least once every two weeks thereafter. When sexual maturation is evaluated, the age and body weight of the animal when vaginal patency (36) or preputial separation (37) occurs should be determined for at least one male and one female per litter.

**▼ M5****Behavioural ontogeny**

34. Ontogeny of selected behaviours should be measured in at least one pup/sex/litter during the appropriate age period, with the same pups being used on all test days for all behaviours assessed. The measurement days should be spaced evenly over that period to define either the normal or treatment-related change in ontogeny of that behaviour (38). The following are some examples of behaviours for which their ontogeny could be assessed: righting reflex, negative geotaxis and motor activity (38)(39)(40).

**Motor activity**

35. Motor activity should be monitored (41)(42)(43)(44)(45) during the pre-weaning and adult age periods. For testing at the time of weaning, see paragraph 32. The test session should be long enough to demonstrate intra-session habituation for non-treated controls. Use of motor activity to assess behavioural ontogeny is strongly recommended. If used as a test of behavioural ontogeny, then testing should utilise the same animals for all pre-weaning test sessions. Testing should be frequent enough to assess the ontogeny of intra-session habituation (44). This may require three or more time periods prior to, and including the day of weaning (e.g. PND 13, 17, 21). Testing of the same animals, or littermates, should also occur at an adult age close to study termination (e.g. PND 60-70). Testing on additional days may be done as necessary. Motor activity should be monitored by an automated activity recording apparatus which should be capable of detecting both increases and decreases in activity, (i.e. baseline activity as measured by the device should not be so low as to preclude detection of decreases, nor so high as to preclude detection of increases in activity). Each device should be tested by standard procedures to ensure, to the extent possible, reliability of operation across devices and across days. To the extent possible, treatment groups should be balanced across devices. Each animal should be tested individually. Treatment groups should be counter-balanced across test times to avoid confounding by circadian rhythms of activity. Efforts should be made to ensure that variations in the test conditions are minimal and are not systematically related to treatment. Among the variables that can affect many measures of behaviour, including motor activity, are sound level, size and shape of the test cage, temperature, relative humidity, light conditions, odours, use of home cage or novel test cage and environmental distractions.

**Motor and sensory function**

36. Motor and sensory function should be examined in detail at least once for the adolescent period and once during the young adult period (e.g. PND 60-70). For testing at the time of weaning, see paragraph 32. Sufficient testing should be conducted to ensure an adequate quantitative sampling of sensory modalities (e.g. somato-sensory, vestibular) and motor functions (e.g. strength, coordination). A few examples of tests for motor and sensory function are extensor thrust response (46), righting reflex (47)(48), auditory startle habituation (40)(49)(50)(51)(52)(53)(54), and evoked potentials (55).

**▼ M5****Learning and memory tests**

37. A test of associative learning and memory should be conducted post-weaning (e.g.  $25 \pm 2$  days) and for young adults (PND 60 and older). For testing at the time of weaning, see paragraph 32. The same or separate test(s) may be used at these two stages of development. Some flexibility is allowed in the choice of test(s) for learning and memory in weanling and adult rats. However, the test(s) should be designed so as to fulfil two criteria. First, learning should be assessed either as a change across several repeated learning trials or sessions, or, in tests involving a single trial, with reference to a condition that controls for non-associative effects of the training experience. Second, the test(s) should include some measure of memory (short-term or long-term) in addition to original learning (acquisition), but this measure of memory cannot be reported in the absence of a measure of acquisition obtained from the same test. If the test(s) of learning and memory reveal(s) an effect of the test chemical, additional tests to rule out alternative interpretations based on alterations in sensory, motivational, and/or motor capacities may be considered. In addition to the above two criteria, it is recommended that the test of learning and memory be chosen on the basis of its demonstrated sensitivity to the class of chemical under investigation, if such information is available in the literature. In the absence of such information, examples of tests that could be made to meet the above criteria include: passive avoidance (43)(56)(57), delayed-matching-to-position for the adult rat (58) and for the infant rat (59), olfactory conditioning (43)(60), Morris water maze (61)(62)(63), Biel or Cincinnati maze (64)(65), radial arm maze (66), T-maze (43), and acquisition and retention of schedule-controlled behaviour (26)(67)(68). Additional tests are described in the literature for weanling (26)(27) and adult rats (19)(20).

**Post-mortem examination**

38. Maternal animals can be euthanised after weaning of the offspring.
39. Neuropathological evaluation of the offspring will be conducted using tissues from animals humanely killed at PND 22 or at an earlier time point between PND 11 and PND 22, as well as at study termination. For offspring killed through PND 22, brain tissues should be evaluated; for animals killed at termination, both central nervous system (CNS) tissues and peripheral nervous system (PNS) tissues should be evaluated. Animals killed on PND 22 or earlier may be fixed either by immersion or perfusion. Animals killed at study termination should be fixed by perfusion. All aspects of the preparation of tissue samples, from the perfusion of animals, through the dissection of tissue samples, tissue processing, and staining of slides should employ a counterbalanced design such that each batch contains representative samples from each dose group. Additional guidance on neuropathology can be found in OECD Guidance Document No 20(9), see also (103).

**Processing of tissue samples**

40. All gross abnormalities apparent at the time of necropsy should be noted. Tissue samples taken should represent all major regions of the nervous system. The tissue samples should be retained in an appropriate fixative and processed according to standardised published histological protocols (69)(70)(71)(103). Paraffin embedding is acceptable for tissues of the CNS and PNS, but the use of osmium in post-fixation, together with



**▼ M5**

epoxy embedding, may be appropriate when a higher degree of resolution is required (e.g. for peripheral nerves when a peripheral neuropathy is suspected and/or for morphometric analysis of peripheral nerves). Brain tissue collected for morphometric analysis should be embedded in appropriate media at all dose levels at the same time in order to avoid shrinkage artefacts that may be associated with prolonged storage in fixative (6).

**Neuropathological examination**

41. The purposes of the qualitative examination are:
  - (i) to identify regions within the nervous system exhibiting evidence of neuropathological alterations;
  - (ii) to identify types of neuropathological alterations resulting from exposure to the test chemical; and
  - (iii) to determine the range of severity of the neuropathological alterations.

Representative histological sections from the tissue samples should be examined microscopically by an appropriately trained pathologist for evidence of neuropathological alterations. All neuropathologic alterations should be assigned a subjective grade indicating severity. A hematoxylin and eosin stain may be sufficient for evaluating brain sections from animals humanely killed at PND 22, or earlier. However, a myelin stain (e.g. luxol fast blue/cresyl violet) and a silver stain (e.g. Bielschowsky's or Bodian's stains) are recommended for sections of CNS and PNS tissues from animals killed at study termination. Subject to the professional judgement of the pathologist and the kind of alterations observed, other stains may be considered appropriate to identify and characterise particular types of alterations (e.g. glial fibrillary acidic protein (GFAP) or lectin histochemistry to assess glial and microglial alterations (72), fluoro-jade to detect necrosis (73)(74), or silver stains specific for neural degeneration (75)).

42. Morphometric (quantitative) evaluation should be performed as these data may assist in the detection of a treatment-related effect and are valuable in the interpretation of treatment-related differences in brain weight or morphology (76)(77). Nervous tissue should be sampled and prepared to enable morphometric evaluation. Morphometric evaluations may include e.g. linear or areal measurements of specific brain regions (78). Linear or areal measurements require the use of homologous sections carefully selected based on reliable microscopic landmarks (6). Stereology may be used to identify treatment-related effects on parameters such as volume or cell number for specific neuroanatomic regions (79)(80)(81)(82)(83)(84).
43. The brains should be examined for any evidence of treatment-related neuropathological alterations and adequate samples should be taken from all major brain regions (e.g. olfactory bulbs, cerebral cortex, hippocampus, basal ganglia, thalamus, hypothalamus, midbrain (tectum, tegmentum, and cerebral peduncles), pons, medulla oblongata, cerebellum) to ensure a thorough examination. It is important that sections for all animals are taken in the same plane. In adults humanely killed at study termination, representative sections of the spinal cord and the PNS should be sampled. The areas examined should include the eye with optic nerve and retina, the spinal cord at the cervical and lumbar swellings, the dorsal and ventral root fibres, the proximal sciatic nerve, the proximal tibial nerve (at the knee), and the tibial nerve calf muscle branches. The spinal cord and peripheral nerve sections should include both cross or transverse and longitudinal sections.

▼ **M5**

44. Neuropathological evaluation should include an examination for indications of developmental damage to the nervous system (6)(85)(86)(87)(88)(89), in addition to the cellular alterations (e.g. neuronal vacuolation, degeneration, necrosis) and tissue changes (e.g. gliosis, leukocytic infiltration, cystic formation). In this regard, it is important that treatment-related effects be distinguished from normal developmental events known to occur at a developmental stage corresponding to the time of sacrifice (90). Examples of significant alterations indicative of developmental insult include, but are not restricted to:

- alterations in the gross size or shape of the olfactory bulbs, cerebrum or cerebellum;
- alterations in the relative size of various brain regions, including decreases or increases in the size of regions resulting from the loss or persistence of normally transient populations of cells or axonal projections (e.g. external germinal layer of cerebellum, corpus callosum);
- alterations in proliferation, migration, and differentiation, as indicated by areas of excessive apoptosis or necrosis, clusters or dispersed populations of ectopic, disoriented or malformed neurons or alterations in the relative size of various layers of cortical structures;
- alterations in patterns of myelination, including an overall size reduction or altered staining of myelinated structures;
- evidence of hydrocephalus, in particular enlargement of the ventricles, stenosis of the cerebral aqueduct and thinning of the cerebral hemispheres.

**Analysis of the dose-response relationship of neuropathological alterations**

45. The following stepwise procedure is recommended for the qualitative and quantitative neuropathological analyses. First, sections from the high dose group are compared with those of the control group. If no evidence of neuropathological alterations is found in animals of the high dose group, no further analysis is required. If evidence of neuropathological alterations is found in the high dose group, then animals from the intermediate and low dose groups are examined. If the high dose group is terminated due to death or other confounding toxicity, the high and intermediate dose groups should be analysed for neuropathological alterations. If there is any indication of neurotoxicity in lower dose groups, neuropathological analysis should be performed in those groups. If any treatment-related neuropathological alterations are found in the qualitative or quantitative examination, the dose-dependence of the incidence, frequency and severity grade of the lesions or of the morphometric alterations should be determined, based on an evaluation of all animals from all dose groups. All regions of the brain that exhibit any evidence of neuropathologic alteration should be included in this evaluation. For each type of lesion, the characteristics used to define each severity grade should be described, indicating the features used to differentiate each grade. The frequency of each type of lesion and its severity grade should be recorded and a statistical analysis should be performed to evaluate the nature of a dose-response relationships. The use of coded slides is recommended (91).

**DATA AND REPORTING**

**Data**

46. Data should be reported individually and summarised in tabular form, showing for each test group the types of change and the number of dams, offspring by sex, and litters displaying each type of change. If direct postnatal exposure of the offspring has been performed, the route, duration and period of exposure should be reported.

**▼ M5****Evaluation and interpretation of results**

47. A developmental neurotoxicity study will provide information on the effects of repeated exposure to a chemical during *in utero* and early postnatal development. Since emphasis is placed on both general toxicity and developmental neurotoxicity endpoints, the results of the study will allow for the discrimination between neurodevelopmental effects occurring in the absence of general maternal toxicity, and those which are only expressed at levels that are also toxic to the maternal animal. Due to the complex interrelationships among study design, statistical analysis, and biological significance of the data, adequate interpretation of developmental neurotoxicity data will involve expert judgment (107)(109). The interpretation of test results should use a weight-of-evidence-approach (20)(92)(93)(94). Patterns of behavioural or morphological findings, if present, as well as evidence of dose-response should be discussed. Data from all studies relevant to the evaluation of developmental neurotoxicity, including human epidemiological studies or case reports, and experimental animal studies (e.g. toxicokinetic data, structure-activity information, data from other toxicity studies) should be included in this characterisation. This includes the relationship between the doses of the test chemical and the presence or absence, incidence, and extent of any neurotoxic effect for each sex (20)(95).
48. Evaluation of data should include a discussion of both the biological and statistical significance. Statistical analysis should be viewed as a tool that guides rather than determines the interpretation of data. Lack of statistical significance should not be the sole rationale for concluding a lack of treatment related effect, just as statistical significance should not be the sole justification for concluding a treatment-related effect. To guard against possible false-negative findings and the inherent difficulties in 'proving a negative,' available positive and historical control data should be discussed, especially when there are no treatment-related effects (102)(106). The probability of false positives should be discussed in light of the total statistical evaluation of the data (96). The evaluation should include the relationship, if any, between observed neuropathological and behavioural alterations.
49. All results should be analysed using statistical models appropriate to the experimental design (108). The choice of a parametric or a nonparametric analysis should be justified by considering factors such as the nature of the data (transformed or not) and their distribution, as well as the relative robustness of the statistical analysis selected. The purpose and design of the study should guide the choice of statistical analyses to minimise Type I (false positive) and Type II (false negative) errors (96)(97)(104)(105). Developmental studies using multiparous species where multiple pups per litter are tested should include the litter in the statistical model to guard against an inflated Type I error rates (98)(99)(100)(101). The statistical unit of measure should be the litter and not the pup. Experiments should be designed such that littermates are not treated as independent observations. Any endpoint repeatedly measured in the same subject should be analysed using statistical models that account for the non-independence of those measures.

**Test report**

50. The test report should include the following information:

*Test chemical:*

- physical nature and, where relevant, physiochemical properties;
  
- identification data, including source;

**▼ M5**

- purity of the preparation, and known and/or anticipated impurities.

*Vehicle (if appropriate):*

- justification for choice of vehicle, if other than water or physiological saline solution.

*Test animals:*

- species and strain used, and a justification if other than the rat;
- supplier of test animals;
- number, age at start, and sex of animals;
- source, housing conditions, diet, water, etc.;
- individual weights of animals at the start of the test.

*Test conditions:*

- rationale for dose level selection;
- rationale for dosing route and time period;
- specifications of the doses administered, including details of the vehicle, volume and physical form of the material administered;
- details of test chemical formulation/diet preparation, achieved concentration, stability and homogeneity of the preparation;
- method used for unique identification of dams and offspring;
- a detailed description of the randomisation procedure(s) used to assign dams to treatment groups, to select pups for culling, and to assign pups to test groups;
- details of the administration of the test chemical;
- conversion from diet/drinking water or inhalation test chemical concentration (ppm) to the actual dose (mg/kg body weight/day), if applicable;
- environmental conditions;
- details of food and water (e.g. tap, distilled) quality;
- dates of study start and end.

*Observations and test procedures:*

- a detailed description of the procedures used to standardise observations and procedures as well as operational definitions for scoring observations;
- a list of all test procedures used, and justification for their use;
- details of the behavioural/functional, pathological, neurochemical or electrophysiological procedures used, including information and details on automated devices;
- procedures for calibrating and ensuring the equivalence of devices and the balancing of treatment groups in testing procedures;
- a short justification explaining any decisions involving professional judgement.

**▼ M5**

*Results (individual and summary, including mean and variance when appropriate):*

- the number of animals at the start of the study and the number at the end of the study;
- the number of animals and litters used for each test method;
- identification number of each animal and the litter from which it came;
- litter size and mean weight at birth by sex;
- body weight and body weight change data, including terminal body weight for dams and offspring;
- food consumption data, and water consumption data if appropriate (e.g. if test chemical is administered via water);
- toxic response data by sex and dose level, including signs of toxicity or mortality, including time and cause of death, if appropriate;
- nature, severity, duration, day of onset, time of day, and subsequent course of the detailed clinical observations;
- score on each developmental landmark (weight, sexual maturation and behavioural ontogeny) at each observation time;
- a detailed description of all behavioural, functional, neuropathological, neurochemical, electrophysiological findings by sex, including both increases and decreases from controls;
- necropsy findings;
- brain weights;
- any diagnoses derived from neurological signs and lesions, including naturally-occurring diseases or conditions;
- images of exemplar findings;
- low-power images to assess homology of sections used for morphometry;
- absorption and metabolism data, including complementary data from a separate toxicokinetic study, if available;
- statistical treatment of results, including statistical models used to analyse the data, and the results, regardless of whether they were significant or not;
- list of study personnel, including professional training.

*Discussion of results:*

- dose response information, by sex and group;
- relationship of any other toxic effects to a conclusion about the neurotoxic potential of the test chemical, by sex and group;
- impact of any toxicokinetic information on the conclusions;
- similarities of effects to any known neurotoxicants;

▼ **M5**

- data supporting the reliability and sensitivity of the test method (i.e. positive and historical control data);
- relationships, if any, between neuropathological and functional effects;
- NOAEL or benchmark dose for dams and offspring, by sex and group.

*Conclusions:*

- a discussion of the overall interpretation of the data based on the results, including a conclusion of whether or not the test chemical caused developmental neurotoxicity and the NOAEL.

**LITERATURE**

- (1) OECD (1995). Draft Report of the OECD *Ad Hoc* Working Group on Reproduction and Developmental Toxicity. Copenhagen, Denmark, 13-14 June 1995.
- (2) US EPA (1998). U.S. Environmental Protection Agency Health Effects Test Guidelines. OPPTS 870.6300. Developmental Neurotoxicity Study. US EPA 712-C-98-239. Available: [[http://www.epa.gov/opptsfrs/OPPTS\\_Harmonized/870\\_Health\\_Effects\\_Test\\_Guidelines/Series/](http://www.epa.gov/opptsfrs/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Series/)].
- (3) US EPA (1998). Guidelines for Neurotoxicity Risk Assessment. US EPA 630/R-95/001F. Available: [<http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?PrintVersion=True&deid=12479>].
- (4) Cory-Slechta, D.A., Crofton, K.M., Foran, J.A., Ross, J.F., Sheets, L.P., Weiss, B., Mileson, B. (2001). Methods to identify and characterize developmental neurotoxicity for human health risk assessment: I. Behavioral effects. *Environ. Health Perspect.*, 109:79-91.
- (5) Dorman, D.C., Allen, S.L., Byczkowski, J.Z., Claudio, L., Fisher, J.E. Jr., Fisher, J.W., Harry, G.J., Li, A.A., Makris, S.L., Padilla, S., Sultatos, L.G., Mileson, B.E. (2001). Methods to identify and characterize developmental neurotoxicity for human health risk assessment: III. Pharmacokinetic and pharmacodynamic considerations. *Environ. Health Perspect.*, 109:101-111.
- (6) Garman, R.H., Fix, A.S., Jortner, B.S., Jensen, K.F., Hardisty, J.F., Claudio, L., Ferenc, S. (2001). Methods to identify and characterize developmental neurotoxicity for human health risk assessment: II. Neuropathology. *Environ. Health Perspect.*, 109:93-100.
- (7) OECD (2003). Report of the OECD Expert Consultation Meeting on Developmental Neurotoxicity Testing. Washington D.C., US, 23-25 October 2000.
- (8) OECD (2008). OECD Environment, Health and Safety Publications Series on Testing and Assessment No 43. Guidance Document on Mammalian Reproductive Toxicity Testing and Assessment. Environment Directorate, OECD, Paris. July 2008 Available: [[http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono\(2008\)16&doclanguage=en](http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono(2008)16&doclanguage=en)].
- (9) OECD (2003). OECD Environment, Health and Safety Publications Series on Testing and Assessment No 20. Guidance Document for Neurotoxicity Testing. Environment Directorate, OECD, Paris, September 2003. Available: [[http://www.oecd.org/document/22/0,2340,en\\_2649\\_34377\\_1916054\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/22/0,2340,en_2649_34377_1916054_1_1_1_1,00.html)].
- (10) Kimmel, C.A., Rees, D.C., Francis, E.Z. (1990) Qualitative and quantitative comparability of human and animal developmental neurotoxicity. *Neurotoxicol. Teratol.*, 12: 173-292.
- (11) Spencer, P.S., Schaumburg, H.H., Ludolph, A.C. (2000) *Experimental and Clinical Neurotoxicology, 2nd Edition*, ISBN 0195084772, Oxford University Press, New York.

▼ **M5**

- (12) Mendola, P., Selevan, S.G., Gutter, S., Rice, D. (2002) Environmental factors associated with a spectrum of neurodevelopmental deficits. *Ment. Retard. Dev. Disabil. Res. Rev.* 8:188-197.
- (13) Slikker, W.B., Chang, L.W. (1998) *Handbook of Developmental Neurotoxicology, 1<sup>st</sup> Edition*, ISBN 0126488606, Academic Press, New York.
- (14) Chapter B.34 of this Annex, One-generation reproduction toxicity study.
- (15) Chapter B.35 of this Annex, Two-generation reproduction toxicity study.
- (16) Chapter B.43 of this Annex, Neurotoxicity Study in Rodents.
- (17) Chapter B.31 of this Annex, Prenatal developmental toxicity study.
- (18) Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. OJ L 276, 20.10.2010, p. 33
- (19) WHO (1986) *Principles and Methods for the Assessment of Neurotoxicity Associated with Exposure to Chemicals*, (Environmental Health Criteria 60), Albany, New York: World Health Organization Publications Center, USA. Available: [<http://www.inchem.org/documents/ehc/ehc/ehc060.htm>].
- (20) WHO (2001) *Neurotoxicity Risk Assessment for Human Health: Principles and Approaches*, (Environmental Health Criteria 223), World Health Organization Publications, Geneva. Available: [<http://www.intox.org/databank/documents/supplem/supp/ehc223.htm>].
- (21) Chang, L.W., Slikker, W. (1995) *Neurotoxicology: Approaches and Methods, 1<sup>st</sup> Edition*, ISBN 012168055X, Academic Press, New York.
- (22) De Cabo, C., Viveros, M.P. (1997) Effects of neonatal naltrexone on neurological and somatic development in rats of both genders. *Neurotoxicol. Teratol.*, 19:499-509.
- (23) Agnish, N.D., Keller, K.A. (1997) The rationale for culling of rodent litters. *Fundam. Appl. Toxicol.*, 38:2-6.
- (24) Avery, D.L., Spyker, J.M. (1977) Foot tattoo of neonatal mice. *Lab. Animal Sci.*, 27:110-112.
- (25) Wier, P.J., Guerriero, F.J., Walker, R.F. (1989) Implementation of a primary screen for developmental neurotoxicity. *Fundam. Appl. Toxicol.*, 13:118-136.
- (26) Spear, N.E., Campbell, B.A. (1979) *Ontogeny of Learning and Memory*. ISBN 0470268492, Erlbaum Associates, New Jersey.
- (27) Krasnegor, N.A., Blass, E.M., Hofer, M.A., Smotherman, W. (1987) *Perinatal Development: A Psychobiological Perspective*. Academic Press, Orlando.
- (28) Zoetis, T., Walls, I. (2003) *Principles and Practices for Direct Dosing of Pre-Weaning Mammals in Toxicity Testing and Research*. ILSI Press, Washington, DC.
- (29) Moser, V., Walls, I., Zoetis, T. (2005) Direct dosing of preweaning rodents in toxicity testing and research: Deliberations of an ILSI RSI expert working group. *Int. J. Toxicol.*, 24:87-94.
- (30) Conolly, R.B., Beck, B.D., Goodman, J.I. (1999) Stimulating research to improve the scientific basis of risk assessment. *Toxicol. Sci.*, 49: 1-4.
- (31) ICH (1993) ICH Harmonised Tripartite Guideline: Detection of Toxicity to Reproduction for Medical Products (S5A). International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use.
- (32) Lochry, E.A. (1987) Concurrent use of behavioral/functional testing in existing reproductive and developmental toxicity screens: Practical considerations. *J. Am. Coll. Toxicol.*, 6:433-439.

## ▼ M5

- (33) Tachibana, T., Narita, H., Ogawa, T., Tanimura, T. (1998) Using postnatal age to determine test dates leads to misinterpretation when treatments alter gestation length, results from a collaborative behavioral teratology study in Japan. *Neurotoxicol. Teratol.*, 20:449-457.
- (34) Gallavan, R.H. Jr., Holson, J.F., Stump, D.G., Knapp, J.F., Reynolds, V.L. (1999) Interpreting the toxicologic significance of alterations in anogenital distance: potential for confounding effects of progeny body weights. *Reprod. Toxicol.*, 13:383-390.
- (35) Gray, L.E. Jr., Ostby, J., Furr, J., Price, M., Veeramachaneni, D.N., Parks, L. (2000) Perinatal exposure to the phthalates DEHP, BBP, and DINP, but not DEP, DMP, or DOTP, alters sexual differentiation of the male rat. *Toxicol. Sci.*, 58:350-365.
- (36) Adams, J., Buelke-Sam, J., Kimmel, C.A., Nelson, C.J., Reiter, L.W., Sobotka, T.J., Tilson, H.A., Nelson, B.K. (1985) Collaborative behavioral teratology study: Protocol design and testing procedure. *Neurobehav. Toxicol. Teratol.*, 7:579-586.
- (37) Korenbrot, C.C., Huhtaniemi, I.T., Weiner, R.W. (1977) Preputial separation as an external sign of pubertal development in the male rat. *Biol. Reprod.*, 17:298-303.
- (38) Spear, L.P. (1990) Neurobehavioral assessment during the early postnatal period. *Neurotoxicol. Teratol.*, 12:489-95.
- (39) Altman, J., Sudarshan, K. (1975) Postnatal development of locomotion in the laboratory rat. *Anim. Behav.*, 23:896-920.
- (40) Adams, J. (1986) Methods in Behavioral Teratology. In: *Handbook of Behavioral Teratology*. Riley, E.P., Vorhees, C.V. (eds.) Plenum Press, New York, pp. 67-100.
- (41) Reiter, L.W., MacPhail, R.C. (1979) Motor activity: A survey of methods with potential use in toxicity testing. *Neurobehav. Toxicol.*, 1:53-66.
- (42) Robbins, T.W. (1977) A critique of the methods available for the measurement of spontaneous motor activity, *Handbook of Psychopharmacology*, Vol. 7, Iverson, L.L., Iverson, D.S., Snyder, S.H., (eds.) Plenum Press, New York, pp. 37-82.
- (43) Crofton, K.M., Peele, D.B., Stanton, M.E. (1993) Developmental neurotoxicity following neonatal exposure to 3,3'-iminodipropionitrile in the rat. *Neurotoxicol. Teratol.*, 15:117-129.
- (44) Ruppert, P.H., Dean, K.F., Reiter, L.W. (1985) Development of locomotor activity of rat pups in figure-eight mazes. *Dev. Psychobiol.*, 18:247-260.
- (45) Crofton, K.M., Howard, J.L., Moser, V.C., Gill, M.W., Reiter, L.W., Tilson, H.A., MacPhail, R.C. (1991) Interlaboratory comparison of motor activity experiments: Implications for neurotoxicological assessments. *Neurotoxicol. Teratol.*, 13:599-609.
- (46) Ross, J. F., Handley, D. E., Fix, A. S., Lawhorn, G. T., Carr, G. J. (1997) Quantification of the hind-limb extensor thrust response in rats. *Neurotoxicol. Teratol.*, 19:1997. 405-411.
- (47) Handley, D.E., Ross, J.F., Carr, G.J. (1998) A force plate system for measuring low-magnitude reaction forces in small laboratory animals. *Physiol. Behav.*, 64:661-669.
- (48) Edwards, P.M., Parker, V.H. (1977) A simple, sensitive, and objective method for early assessment of acrylamide neuropathy in rats. *Toxicol. Appl. Pharmacol.*, 40:589-591.
- (49) Davis, M. (1984) The mammalian startle response. In: *Neural Mechanisms of Startle Behavior*, Eaton, R.C. (ed), Plenum Press, New York, pp. 287-351
- (50) Koch, M. (1999) The neurobiology of startle. *Prog. Neurobiol.*, 59:107-128.
- (51) Crofton, K.M. (1992) Reflex modification and the assessment of sensory dysfunction. In *Target Organ Toxicology Series: Neurotoxicology*, Tilson, H., Mitchell, C. (eds). Raven Press, New York, pp. 181-211.



▼ M5

- (52) Crofton, K.M., Sheets, L.P. (1989) Evaluation of sensory system function using reflex modification of the startle response. *J. Am. Coll. Toxicol.*, 8:199-211.
- (53) Crofton, K.M., Lassiter, T.L., Rebert, C.S. (1994) Solvent-induced ototoxicity in rats: An atypical selective mid-frequency hearing deficit. *Hear. Res.*, 80:25-30.
- (54) Ison, J.R. (1984) Reflex modification as an objective test for sensory processing following toxicant exposure. *Neurobehav. Toxicol. Teratol.*, 6:437-445.
- (55) Mattsson, J.L., Boyes, W.K., Ross, J.F. (1992) Incorporating evoked potentials into neurotoxicity test schemes. In: *Target Organ Toxicology Series: Neurotoxicity*, Tilson, H., Mitchell, C., (eds.), Raven Press, New York. pp. 125-145.
- (56) Peele, D.B., Allison, S.D., Crofton, K.M. (1990) Learning and memory deficits in rats following exposure to 3,3'-iminopropionitrile. *Toxicol. Appl. Pharmacol.*, 105:321-332.
- (57) Bammer, G. (1982) Pharmacological investigations of neurotransmitter involvement in passive avoidance responding: A review and some new results. *Neurosci. Behav. Rev.*, 6:247-296.
- (58) Bushnell, P.J. (1988) Effects of delay, intertrial interval, delay behavior and trimethyltin on spatial delayed response in rats. *Neurotoxicol. Teratol.*, 10:237-244.
- (59) Green, R.J., Stanton, M.E. (1989) Differential ontogeny of working memory and reference memory in the rat. *Behav. Neurosci.*, 103:98-105.
- (60) Kucharski, D., Spear, N.E. (1984) Conditioning of aversion to an odor paired with peripheral shock in the developing rat. *Develop. Psychobiol.*, 17:465-479.
- (61) Morris, R. (1984) Developments of a water-maze procedure for studying spatial learning in the rat. *J. Neurosci. Methods*, 11:47-60.
- (62) Brandeis, R., Brandys, Y., Yehuda, S. (1989) The use of the Morris water maze in the study of memory and learning. *Int. J. Neurosci.*, 48:29-69.
- (63) D'Hooge, R., De Deyn, P.P. (2001) Applications of the Morris water maze in the study of learning and memory. *Brain Res. Rev.*, 36:60-90.
- (64) Vorhees, C.V. (1987) Maze learning in rats: A comparison of performance in two water mazes in progeny prenatally exposed to different doses of phenytoin. *Neurotoxicol. Teratol.*, 9:235-241.
- (65) Vorhees, C.V. (1997) Methods for detecting long-term CNS dysfunction after prenatal exposure to neurotoxins. *Drug Chem. Toxicol.*, 20:387-399.
- (66) Akaike, M., Tanaka, K., Goto, M., Sakaguchi, T. (1988) Impaired Biel and Radial arm maze learning in rats with methyl-nitrosurea induced microcephaly. *Neurotoxicol. Teratol.*, 10:327-332.
- (67) Cory-Slechta, D.A., Weiss, B., Cox, C. (1983) Delayed behavioral toxicity of lead with increasing exposure concentration. *Toxicol. Appl. Pharmacol.*, 71:342-352.
- (68) Campbell, B.A., Haroutunian, V. (1981) Effects of age on long-term memory: Retention of fixed interval responding. *J. Gerontol.*, 36:338-341.
- (69) Fix, A.S., Garman, R.H. (2000) Practical aspects of neuropathology: A technical guide for working with the nervous system. *Toxicol. Pathol.*, 28: 122-131.
- (70) Prophet, E.B., Mills, B., Arrington, J.B., Sobin, L.H. (1994) *Laboratory Methods in Histotechnology*, American Registry of Pathology, Washington, DC, pp. 84-107.
- (71) Bancroft, J.D., Gamble, M. (2002) *Theory and Practice of Histological Techniques*, 5<sup>th</sup> edition, Churchill Livingstone, London.

## ▼ M5

- (72) Fix, A.S., Ross, J.F., Stitzel, S.R., Switzer, R.C. (1996) Integrated evaluation of central nervous system lesions: stains for neurons, astrocytes, and microglia reveal the spatial and temporal features of MK-801-induced neuronal necrosis in the rat cerebral cortex. *Toxicol. Pathol.*, 24: 291-304.
- (73) Schmued, L.C., Hopkins, K.J. (2000) Fluoro-Jade B: A high affinity tracer for the localization of neuronal degeneration. *Brain Res.*, 874:123-130.
- (74) Krinke, G.J., Classen, W., Vidotto, N., Suter, E., Wurmlin, C.H. (2001) Detecting necrotic neurons with fluoro-jade stain. *Exp. Toxic. Pathol.*, 53:365-372.
- (75) De Olmos, I.S., Beltramino, C.A., and de Olmos de Lorenzo, S. (1994) Use of an amino-cupric-silver technique for the detection of early and semiacute neuronal degeneration caused by neurotoxicants, hypoxia and physical trauma. *Neurotoxicol. Teratol.*, 16, 545-561.
- (76) De Groot, D.M.G., Bos-Kuijpers, M.H.M., Kaufmann, W.S.H., Lammers, J.H.C.M., O'Callaghan, J.P., Pakkenberg, B., Pelgrim, M.T.M., Waalkens-Berendsen, I.D.H., Waanders, M.M., Gundersen, H.J. (2005a) Regulatory developmental neurotoxicity testing: A model study focusing on conventional neuropathology endpoints and other perspectives. *Environ. Toxicol. Pharmacol.*, 19:745-755.
- (77) De Groot, D.M.G., Hartgring, S., van de Horst, L., Moerkens, M., Otto, M., Bos-Kuijpers, M.H.M., Kaufmann, W.S.H., Lammers, J.H.C.M., O'Callaghan, J.P., Waalkens-Berendsen, I.D.H., Pakkenberg, B., Gundersen, H.J. (2005b) 2D and 3D assessment of neuropathology in rat brain after prenatal exposure to methylazoxymethanol, a model for developmental neurotoxicity. *Reprod. Toxicol.*, 20:417-432.
- (78) Rodier, P.M., Gramann, W.J. (1979) Morphologic effects of interference with cell proliferation in the early fetal period. *Neurobehav. Toxicol.*, 1:129-135.
- (79) Howard, C.V., Reed, M.G. (1998) *Unbiased Stereology: Three-Dimensional Measurement in Microscopy*, Springer-Verlag, New York.
- (80) Hyman, B.T., Gomez-Isla, T., Irizarry, M.C. (1998) Stereology: A practical primer for neuropathology. *J. Neuropathol. Exp. Neurol.*, 57: 305-310.
- (81) Korbo, L., Andersen, B.B., Ladefoged, O., Møller, A. (1993) Total numbers of various cell types in rat cerebellar cortex estimated using an unbiased stereological method. *Brain Res.*, 609: 262-268.
- (82) Schmitz, C. (1997) Towards more readily comprehensible procedures in disector stereology. *J. Neurocytol.*, 26:707-710.
- (83) West, M.J. (1999) Stereological methods for estimating the total number of neurons and synapses: Issues of precision and bias. *Trends Neurosci.*, 22:51-61.
- (84) Schmitz, C., Hof, P.R. (2005) Design-based stereology in neuroscience. *Neuroscience*, 130: 813-831.
- (85) Gavin, C.E., Kates, B., Gerken, L.A., Rodier, P.M. (1994) Patterns of growth deficiency in rats exposed *in utero* to undernutrition, ethanol, or the neuroteratogen methylazoxymethanol (MAM). *Teratology*, 49:113-121.
- (86) Ohno, M., Aotani, H., Shimada, M. (1995) Glial responses to hypoxic/ischemic encephalopathy in neonatal rat cerebrum. *Develop. Brain Res.*, 84:294-298.
- (87) Jensen KF, Catalano SM. (1998) Brain morphogenesis and developmental neurotoxicology. In: *Handbook of Developmental Neurotoxicology*, Slikker, Jr. W., Chang, L.W. (eds) Academic Press, New York, pp. 3-41.
- (88) Ikonomidou, C., Bosch, F., Miksa, M., Bittigau, P., Vöckler, J., Dikranian, K., Tenkova, T.I., Stefovská, V., Turski, L., Olney, J.W. (1999) Blockade of NMDA receptors and apoptotic neurodegeneration in the developing brain. *Science*, 283:70-74.

## ▼ M5

- (89) Ikonomidou, C., Bittigau, P., Ishimaru, M.J., Wozniak, D.F., Koch, C., Genz, K., Price, M.T., Sefovska, V., Hörster, F., Tenkova, T., Dikranian, K., Olney, J.W. (2000) Ethanol-induced apoptotic degeneration and fetal alcohol syndrome. *Science*, 287:1056–1060.
- (90) Friede, R. L. (1989) *Developmental Neuropathology*. Second edition. Springer-Verlag, Berlin.
- (91) House, D.E., Berman, E., Seeley, J.C., Simmons, J.E. (1992) Comparison of open and blind histopathologic evaluation of hepatic lesions. *Toxicol. Let.*, 63:127-133.
- (92) Tilson, H.A., MacPhail, R.C., Crofton, K.M. (1996) Setting exposure standards: a decision process. *Environ. Health Perspect.*, 104:401-405.
- (93) US EPA (2005) Guidelines for Carcinogen Risk Assessment. US EPA NCEA-F-0644A.
- (94) US EPA (1996) Guidelines for Reproductive Toxicity Risk Assessment, Federal Register 61(212): 56274-56322.
- (95) Danish Environmental Protection Agency (1995) *Neurotoxicology*. Review of Definitions, Methodology, and Criteria. Miljøprojekt nr. 282. Ladefoged, O., Lam, H.R., Østergaard, G., Nielsen, E., Arlien-Søborg, P.
- (96) Muller, K.E., Barton, C.N., Benignus, V.A. (1984). Recommendations for appropriate statistical practice in toxicologic experiments. *Neurotoxicology*, 5:113-126.
- (97) Gad, S.C. (1989) Principles of screening in toxicology with special emphasis on applications to Neurotoxicology. *J. Am. Coll. Toxicol.*, 8:21-27.
- (98) Abby, H., Howard, E. (1973) Statistical procedures in developmental studies on a species with multiple offspring. *Dev. Psychobiol.*, 6:329-335.
- (99) Haseman, J.K., Hogan, M.D. (1975) Selection of the experimental unit in teratology studies. *Teratology*, 12:165-172.
- (100) Holson, R.R., Pearce, B. (1992) Principles and pitfalls in the analysis of prenatal treatment effects in multiparous species. *Neurotoxicol. Teratol.*, 14: 221-228.
- (101) Nelson, C.J., Felton, R.P., Kimmel, C.A., Buelke-Sam, J., Adams, J. (1985) Collaborative Behavioral Teratology Study: Statistical approach. *Neurobehav. Toxicol. Teratol.*, 7:587-90.
- (102) Crofton, K.M., Makris, S.L., Sette, W.F., Mendez, E., Raffaele, K.C. (2004) A qualitative retrospective analysis of positive control data in developmental neurotoxicity studies. *Neurotoxicol. Teratol.*, 26:345-352.
- (103) Bolon, B., Garman, R., Jensen, K., Krinke, G., Stuart, B., and an *ad hoc* working group of the STP Scientific and Regulatory Policy Committee. (2006) A ‘best practices’ approach to neuropathological assessment in developmental neurotoxicity testing — for today. *Toxicol. Pathol.* 34:296-313.
- (104) Tamura, R.N., Buelke-Sam, J. (1992) The use of repeated measures analysis in developmental toxicology studies. *Neurotoxicol. Teratol.*, 14(3):205-210.
- (105) Tukey, J.W., Ciminera, J.L., Heyse, J.F. (1985) Testing the statistical certainty of a response to increasing doses of a drug. *Biometrics*, 41:295-301.
- (106) Crofton, K.M., Foss, J.A., Haas, U., Jensen, K., Levin, E.D., and Parker, S.P. (2008) Undertaking positive control studies as part of developmental neurotoxicity testing: report from the ILSI Research Foundation/Risk Science Institute expert working group on neurodevelopmental endpoints. *Neurotoxicology and Teratology*, 30(4):266-287.
- (107) Raffaele, K.C., Fisher, E., Hancock, S., Hazelden, K., and Sobrian, S.K. (2008) Determining normal variability in a developmental neurotoxicity test: report from the ILSI Research Foundation/Risk Science Institute expert working group on neurodevelopmental endpoints. *Neurotoxicology and Teratology*, 30(4):288-325.

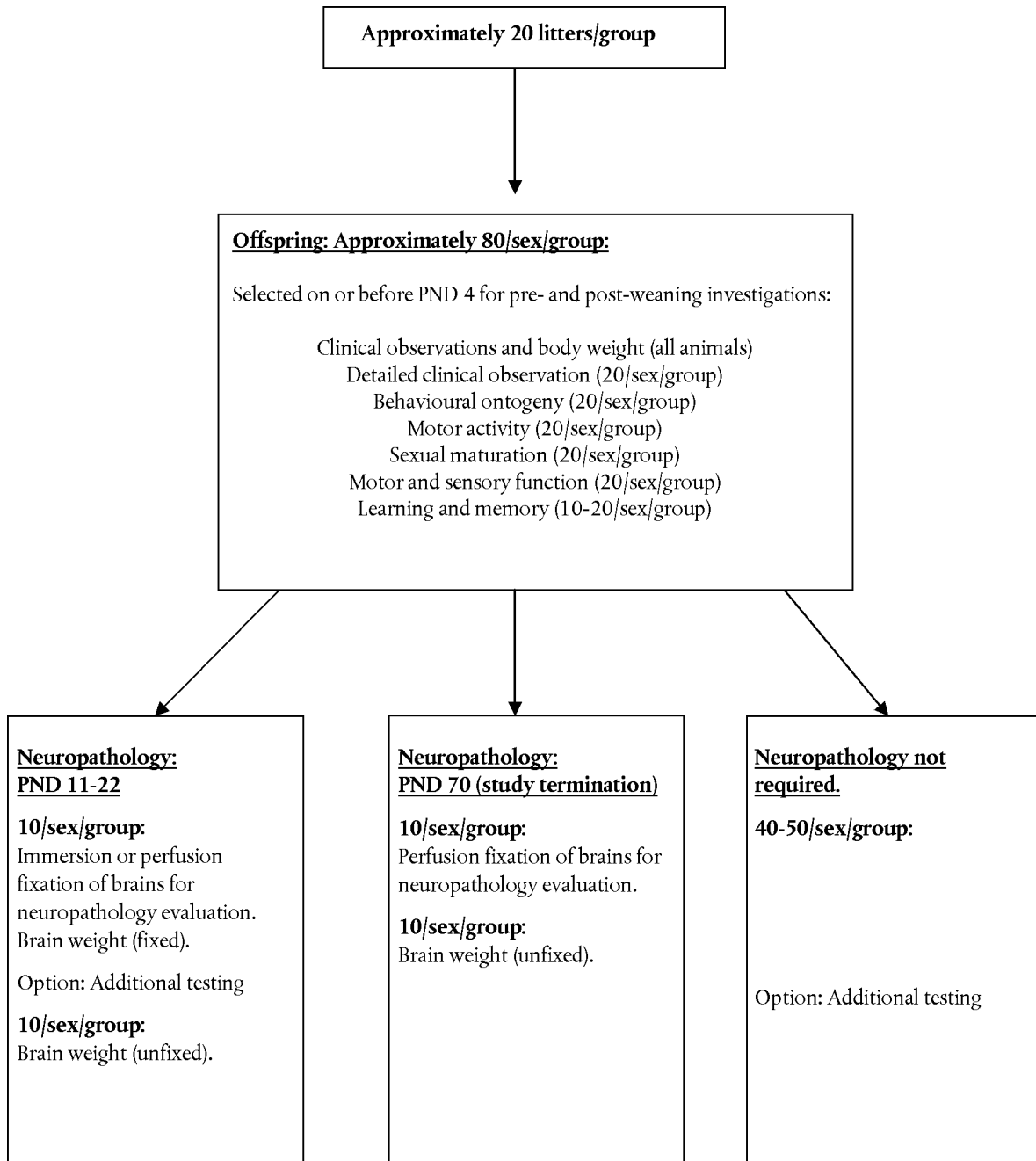
**▼ M5**

- (108) Holson, R.R., Freshwater, L., Maurissen, J.P.J., Moser, V.C., and Phang, W. (2008) Statistical issues and techniques appropriate for developmental neurotoxicity testing: a report from the ILSI Research Foundation/Risk Science Institute expert working group on neurodevelopmental endpoints. *Neurotoxicology and Teratology*, 30(4):326-348.
- (109) Tyl, R.W., Crofton, K.M., Moretto, A., Moser, V.C., Sheets, L.P., and Sobotka, T.J. (2008) Identification and interpretation of developmental neurotoxicity effects: a report from the ILSI Research Foundation/Risk Science Institute expert working group on neurodevelopmental endpoints. *Neurotoxicology and Teratology*, 30(4):349-381.

▼ M5

Figure 1

General testing scheme for functional/behavioural tests, neuropathology evaluation, and brain weights. This diagram is based on the description in paragraphs 13-15 (PND=postnatal day). Examples of animal assignment are given in Appendix 1.



▼ **M5***Appendix 1*

1. Examples of possible assignments are described and tabulated below. These examples are provided to illustrate that assignment of study animals to various testing paradigms can be accomplished in a number of different ways.

**Example 1**

2. One set of 20 pups/sex/dose level (i.e. 1 male and 1 female per litter) is used for pre-weaning testing of behavioural ontogeny. Out of these animals, 10 pups/sex/dose level (i.e. 1 male or 1 female per litter) are humanely killed at PND 22. The brains are removed, weighed and processed for histopathologic evaluation. In addition, brain weight data are collected using unfixated brains from the remaining 10 males and 10 females per dose level.
3. Another set of 20 animals/sex/dose level (i.e. 1 male and 1 female per litter) is used for post-weaning functional/behavioral tests (detailed clinical observations, motor activity, auditory startle and cognitive function testing in adolescents) and assessing age of sexual maturation. Of these animals, 10 animals/sex/dose level (i.e. 1 male or 1 female per litter), are anaesthetised and fixed via perfusion at study termination (approximately PND 70). After additional fixation *in situ*, the brain is removed and processed for neuropathological evaluation.
4. For cognitive function testing in young adults (e.g. PND 60-70), a third set of 20 pups/sex/dose level is used (i.e. 1 male and 1 female per litter). Of these animals, 10 animals/sex/group (1 male or 1 female per litter) are killed at study termination and the brain is removed and weighed.
5. The remaining 20 animals/sex/group are reserved for possible additional tests.

*Table 1*

Pup No <sup>(a)</sup>		No of pups assigned to test	Examination/Test
m	f		
1	5	20 m + 20 f	Behavioural ontogeny
		10 m + 10 f	PND 22 brain weight/neuropathology/morphometry
		10 m + 10 f	PND 22 brain weight
2	6	20 m + 20 f	Detailed clinical observations
		20 m + 20 f	Motor activity
		20 m + 20 f	Sexual maturation
		20 m + 20 f	Motor and sensory function
		20 m + 20 f	Learning and memory (PND 25)
		10 m + 10 f	Young adult brain weight/neuropathology/morphometry ~ PND 70

## ▼ M5

Pup No <sup>(a)</sup>		No of pups assigned to test	Examination/Test
m	f		
3	7	20 m + 20 f	Learning and memory (young adults)
		10 m + 10 f	Young adult brain weight ~ PND 70
4	8	—	Reserve animals for replacements or additional tests

<sup>(a)</sup> For this example, litters are culled to 4 males + 4 females; male pups are numbered 1 through 4, female pups 5 through 8.

**Example 2**

6. One set of 20 pups/sex/dose level (i.e. 1 male and 1 female per litter) is used for pre-weaning testing of behavioural ontogeny. Out of these animals, 10 pups/sex/dose level (1 male or 1 female per litter), are humanely killed at PND 11. The brains are removed, weighed and processed for histopathologic evaluation.
7. Another set of 20 animals/sex/dose level (1 male and 1 female per litter) is used for post-weaning examinations (detailed clinical observations, motor activity, assessing age of sexual maturation and motor and sensory function). Of these animals, 10 animals/sex/dose level (i.e. 1 male or 1 female per litter) are anaesthetised and fixed via perfusion at study termination (approximately PND 70). After additional fixation *in situ*, the brain is removed, weighed and processed for neuropathological evaluation.
8. For cognitive function testing in adolescents and young adults, 10 pups/sex/dose level are used (i.e. 1 male or 1 female per litter). Different animals are used for testing for cognitive function tests at PND 23 and young adults. At termination, the 10 animals/sex/group tested as adults are killed, the brain is removed and weighed.
9. The remaining 20 animals/sex/group not selected for testing are killed and discarded at weaning.

Table 2

Pup No <sup>(a)</sup>		No of pups assigned to test	Examination/Test
m	f		
1	5	20 m + 20 f	Behavioural ontogeny
		10 m + 10 f	PND 11 brain weight/neuropathology/morphometry
2	6	20 m + 20 f	Detailed clinical observations
		20 m + 20 f	Motor activity
		20 m + 20 f	Sexual maturation
		20 m + 20 f	Motor and sensory function
		10 m + 10 f	Young adult brain weight/neuropathology/morphometry ~ PND 70

## ▼ M5

Pup No <sup>(a)</sup>		No of pups assigned to test	Examination/Test
m	f		
3	7	10 m + 10 f <sup>(b)</sup>	Learning and memory (PND 23)
3	7	10 m + 10 f <sup>(b)</sup>	Learning and memory (young adults) Young adult brain weight
4	8	—	Animals killed and discarded PND 21.

<sup>(a)</sup> For this example, litters are culled to 4 males + 4 females; male pups are numbered 1 through 4, female pups 5 through 8.

<sup>(b)</sup> Different pups are used for cognitive tests at PND 23 and in young adults (e.g. even/odd litters from total of 20).

## Example 3

10. One set 20 pups/sex/dose level (i.e. 1 male and 1 female per litter) is used for brain weight and neuropathology assessment at PND 11. Out of these animals, 10 pups/sex/dose level (i.e. 1 male or 1 female per litter) are humanely killed at PND 11 and brains are removed, weighed and processed for histopathologic evaluation. In addition, brain weight data are collected using unfixed brains from the remaining 10 males and 10 females per dose level.
11. Another set of 20 animals/sex/dose level (i.e. 1 male and 1 female per litter) are used for behavioural ontogeny (motor activity), post-weaning examinations (motor activity and assessing age of sexual maturation), and cognitive function testing in adolescents.
12. Another set of 20 animals/sex/dose level (i.e. 1 male and 1 female per litter) is used for motor and sensory function tests (auditory startle) and detailed clinical observations. Of these animals, 10 animals/sex/dose level (i.e. 1 male or 1 female per litter) are anaesthetised and fixed via perfusion at study termination (approximately PND 70). After additional fixation *in situ*, the brain is removed, weighed and processed for neuropathological evaluation.
13. Another set of 20 pups/sex/dose level are used for cognitive function testing in young adults (i.e. 1 male and 1 female per litter). Of these, 10 animals/sex/group (i.e. 1 male or 1 female per litter) are killed at termination, the brain removed and weighed.

Table 3

Pup No <sup>(a)</sup>		No of pups assigned to test	Examination/Test
m	f		
1	5	10 m + 10 f	PND 11 brain weight/neuropathology/morphometry
		10 m + 10 f	PND 11 brain weight
2	6	20 m + 20 f	Behavioural ontogeny (motor activity)
		20 m + 20 f	Motor activity
		20 m + 20 f	Sexual maturation
		20 m + 20 f	Learning and memory (PND 27)



▼ **M5**

Pup No <sup>(a)</sup>		No of pups assigned to test	Examination/Test
m	f		
3	7	20 m + 20 f	Auditory startle (adolescents and young adults)
		20 m + 20 f	Detailed clinical observations
		10 m + 10 f	Young adult brain weight/neuropathology/morphometry ~ PND 70
4	8	20 m + 20 f	Learning and memory (young adults)
		10 m + 10 f	Young adult brain weight

<sup>(a)</sup> For this example, litters are culled to 4 males + 4 females; male pups are numbered 1 through 4, female pups 5 through 8.

▼ M5

*Appendix 2*

**Definitions**

Chemical: A substance or a mixture

Test chemical: Any substance or mixture tested using this test method

**▼ M5****B.54 UTEROTROPHIC BIOASSAY IN RODENTS: A SHORT-TERM SCREENING TEST FOR OESTROGENIC PROPERTIES**

## INTRODUCTION

1. This test method is equivalent to OECD Test Guideline (TG) 440 (2007). The OECD initiated a high-priority activity in 1998 to revise existing guidelines and to develop new guidelines for the screening and testing of potential endocrine disruptors (1). One element of the activity was to develop a test guideline for the rodent Uterotrophic Bioassay. The rodent Uterotrophic Bioassay then underwent an extensive validation programme including the compilation of a detailed background document (2)(3) and the conduct of extensive intra- and interlaboratory studies to show the relevance and reproducibility of the bioassay with a potent reference oestrogen, weak oestrogen receptor agonists, a strong oestrogen receptor antagonist, and a negative reference chemical (4)(5)(6)(7)(8)(9). This test method B.54 is the outcome of the experience gained during the validation test programme and the results obtained thereby with oestrogenic agonists.
2. The Uterotrophic Bioassay is a short-term screening test that originated in the 1930s (27)(28) and was first standardised for screening by an expert committee in 1962 (32)(35). It is based on the increase in uterine weight or uterotrophic response (for review, see 29). It evaluates the ability of a chemical to elicit biological activities consistent with agonists or antagonists of natural oestrogens (e.g. 17 $\beta$ -estradiol), however, its use for antagonist detection is much less common than for agonists. The uterus responds to oestrogens in two ways. An initial response is an increase in weight due to water imbibition. This response is followed by a weight gain due to tissue growth (30). The uterus responses in rats and mice qualitatively are comparable.
3. This bioassay serves as an *in vivo* screening assay and its application should be seen in the context of the 'OECD Conceptual Framework for the Testing and Assessment of Endocrine Disrupting Chemicals' (Appendix 2). In this Conceptual Framework the Uterotrophic Bioassay is contained in Level 3 as an *in vivo* assay providing data about a single endocrine mechanism, i.e. oestrogenicity.
4. The Uterotrophic Bioassay is intended to be included in a battery of *in vitro* and *in vivo* tests to identify chemicals with potential to interact with the endocrine system, ultimately leading to risk assessments for human health or the environment. The OECD validation programme used both strong and weak oestrogen agonists to evaluate the performance of the assay to identify oestrogenic chemicals (4)(5)(6)(7)(8). Thereby the sensitivity of the test procedure for oestrogen agonists was well demonstrated besides a good intra- and interlaboratory reproducibility.
5. With regard to negative chemicals, only one 'negative' reference chemical already reported negative by uterotrophic assay as well as *in vitro* receptor binding and receptor assays was included in the validation programme, but additional test data, not related to the OECD validation programme, have been evaluated, giving further support to the specificity of the Uterotrophic Bioassay for the screening of oestrogen agonists (16).

▼ M5

## INITIAL CONSIDERATIONS AND LIMITATIONS

6. Oestrogen agonists and antagonists act as ligands for oestrogen receptors  $\alpha$  and  $\beta$  and may activate or inhibit, respectively, the transcriptional action of the receptors. This may have the potential to lead to adverse health hazards, including reproductive and developmental effects. Therefore, the need exists to rapidly assess and evaluate a chemical as a possible oestrogen agonist or antagonist. While informative, the affinity of a ligand for an oestrogen receptor or transcriptional activation of reporter genes *in vitro* is only one of several determinants of possible hazard. Other determinants can include metabolic activation and deactivation upon entering the body, distribution to target tissues, and clearance from the body, depending at least in part on the route of administration and the chemical being tested. This leads to the need to screen the possible activity of a chemical *in vivo* under relevant conditions, unless the chemical's characteristics regarding Absorption — Distribution — Metabolism — Elimination (ADME) already provide appropriate information. Uterine tissues respond with rapid and vigorous growth to stimulation by oestrogens, particularly in laboratory rodents, where the oestrous cycle lasts approximately 4 days. Rodent species, particularly the rat, are also widely used in toxicity studies for hazard characterisation. Therefore, the rodent uterus is an appropriate target organ for the *in vivo* screening of oestrogen agonists and antagonists.
  
7. This test method is based on those protocols employed in the OECD validation study which have been shown to be reliable and repeatable in intra- and interlaboratory studies (5)(7). Currently two methods, namely the ovariectomised adult female method (ovx-adult method) and the immature non-ovariectomised method (immature method) are available. It was shown in the OECD validation test programme that both methods have comparable sensitivity and reproducibility. However, the immature, as it has an intact hypothalamic-pituitary-gonadal (HPG) axis, is somewhat less specific but covers a larger scope of investigation than the ovariectomised animal because it can respond to chemicals that interact with the HPG axis rather than just the oestrogen receptor. The HPG axis of the rat is functional at about 15 days of age. Prior to that, puberty cannot be accelerated with treatments like GnRH. As the females begin to reach puberty, prior to vaginal opening, the female will have several silent cycles that do not result in vaginal opening or ovulation, but there are some hormonal fluctuations. If a chemical stimulates the HPG axis directly or indirectly, precocious puberty, early ovulation and accelerated vaginal opening result. Not only chemicals that act on the HPG axis do this but some diets with higher metabolisable energy levels than others will stimulate growth and accelerate vaginal opening without being oestrogenic. Such chemicals would not induce an uterotrophic response in OVX adult animals as their HPG axis does not work.
  
8. For animal welfare reasons preference should be given to the method using immature rats, avoiding surgical pre-treatment of the animals and avoiding also a possible non-use of those animals which indicate any evidence entering oestrous (see paragraph 30).

**▼ M5**

9. The uterotrophic response is not entirely of oestrogenic origin, i.e. chemicals other than agonists or antagonists of oestrogens may also provide a response. For example, relatively high doses of progesterone, testosterone, or various synthetic progestins may all lead to a stimulative response (30). Any response may be analysed histologically for keratinisation and cornification of the vagina (30). Irrespective of the possible origin of the response, a positive outcome of an Uterotrophic Bioassay should normally initiate actions for further clarification. Additional evidence of oestrogenicity could come from *in vitro* assays, such as the ER binding assays and transcriptional activation assays, or from other *in vivo* assays such as the female pubertal assay.
  
10. Taking into account that the Uterotrophic Bioassay serves as an *in vivo* screening assay, the validation approach taken served both animal welfare considerations and a tiered testing strategy. To this end, effort was directed at rigorously validating reproducibility and sensitivity for oestrogenicity — the main concern for many chemicals-, while little effort was directed at the antioestrogenicity component of the assay. Only one antioestrogen with strong activity was tested since the number of chemicals with a clear antioestrogenic profile (not obscured by some oestrogenic activity) is very limited. Thus this test method is dedicated to the oestrogenic protocol, while the protocol describing the antagonist mode of the assay is included in a Guidance Document (37). The reproducibility and sensitivity of the assay for chemicals with purely anti-oestrogenic activity will be more clearly defined later on, after the test procedure has been in routine use for some time and more chemicals with this modality of action are identified.
  
11. It is acknowledged that all animal based procedures will conform to local standards of animal care; the descriptions of care and treatment set forth below are minimal performance standards, and will be superseded by local regulations such as Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes (38). Further guidance of the humane treatment of animals is given by the OECD (25).
  
12. As with all assays using live animals, it is essential to ensure that the data are truly necessary prior to the start of the assay. For example, two conditions where the data may be required are:
  - high exposure potential (Level 1 of the Conceptual Framework, Appendix 2) or indications for oestrogenicity (Level 2) to investigate whether such effects may occur *in vivo*;
  
  - effects indicating oestrogenicity in Level 4 or 5 *in vivo* tests to substantiate that the effects were related to an oestrogenic mechanism that cannot be elucidated using an *in vitro* test.
  
13. Definitions used in this test method are given in Appendix 1.

**▼ M5**

## PRINCIPLE OF THE TEST

14. The Uterotrophic Bioassay relies for its sensitivity on an animal test system in which the hypothalamic-pituitary-ovarian axis is not functional, leading to low endogenous levels of circulating oestrogen. This will ensure a low baseline uterine weight and a maximum range of response to administered oestrogens. Two oestrogen sensitive states in the female rodent meet this requirement:
  - (i) immature females after weaning and prior to puberty; and
  - (ii) young adult females after ovariectomy with adequate time for uterine tissues to regress.
  
15. The test chemical is administered daily by oral gavage or subcutaneous injection. Graduated test chemical doses are administered to a minimum of two treatment groups (see paragraph 33 for guidance) of experimental animals using one dose level per group and an administration period of three consecutive days for immature method and a minimum administration period of three consecutive days for ovx-adult method. The animals are necropsied approximately 24 hours after the last dose. For oestrogen agonists, the mean uterine weight of the treated animal groups relative to the vehicle group is assessed for a statistically significant increase. A statistically significant increase in the mean uterine weight of a test group indicates a positive response in this bioassay.

## DESCRIPTION OF THE METHOD

**Selection of animal species**

16. Commonly used laboratory rodent strains may be used. As an example, Sprague-Dawley and Wistar strains of rats were used during the validation. Strains with uteri known or suspected to be less responsive should not be used. The laboratory should demonstrate the sensitivity of the strain used as described in paragraphs 26 and 27.
  
17. The rat and mouse have been routinely used in the Uterotrophic Bioassay since the 1930s. The OECD validation studies were only performed with rats based on an understanding that both species are expected to be equivalent and therefore one species should be enough for the world-wide validation in order to save resources and animals. The rat is the species of choice in most reproductive and developmental toxicity studies. Taking into consideration that a vast historical database exists for mice and thus to broaden the scope of the Uterotrophic Bioassay test method in rodents to the use of mice as test species, a limited follow-up validation study was carried out in mice (16). A bridging approach with a limited number of test chemicals, participating laboratories and without coded sample testing has been selected in keeping with the original intent to save resources and animals. This bridging validation study shows for the Uterotrophic Bioassay in young adult ovariectomised mice that, qualitatively and quantitatively, the data obtained in rats and mice correspond well with each other. Where the Uterotrophic Bioassay result may be preliminary to a long-term study, this allows animals from the same strain and source to be used in both studies. The bridging approach was limited to the OVX mice and the report does not provide a robust data set to validate the immature model, thus the immature model for mice is not considered under the scope of the current test method.

**▼M5**

18. Thus, in some cases mice may be used instead of rats. A rationale should be given for this species, based on toxicological, pharmacokinetic, and/or other criteria. Modifications of the protocol may be necessary for mice. For example, the food consumption of mice on a body weight basis is higher than that of rats and therefore the phyto-oestrogen content in food should be lower for mice than for rats (9)(20)(22).

**Housing and feeding conditions**

19. All procedures should conform with local standards of laboratory animal care. These descriptions of care and treatment are minimum standards and will be superseded by local regulations such as Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes (38). The temperature in the experimental animal room should be 22 °C (with an approximate range  $\pm 3$  °C). The relative humidity should be a minimum of 30 % and preferably should not exceed a maximum 70 %, other than during room cleaning. The aim should be relative humidity of 50-60 %. Lighting should be artificial. The daily lighting sequence should be 12 hours light, 12 hours dark.
20. Laboratory diet and drinking water should be provided *ad libitum*. Young adult animals may be housed individually or be caged in groups of up to three animals. Due to the young age of the immature animals, social group housing is recommended.
21. High levels of phyto-oestrogens in laboratory diets have been known to increase uterine weights in rodents to a degree enough as to interfere with the Uterotrophic Bioassay (13)(14)(15). High levels of phyto-oestrogens and of metabolisable energy in laboratory diets may also result in early puberty, if immature animals are used. The presence of phyto-oestrogens results primarily from the inclusion of soy and alfalfa products in the laboratory diets and concentrations of phyto-oestrogens have been shown to vary from batch-to-batch of standard laboratory diets (23). Body weight is an important variable, as the quantity of food consumed is related to body weight. Therefore, the actual phyto-oestrogen dose consumed from the same diet may vary among species and by age (9). For immature female rats, food consumption on a body weight basis may be approximately double that of ovariectomised young adult females. For young adult mice, food consumption on a body weight basis may be approximately quadruple that of ovariectomised young adult female rats.
22. Uterotrophic Bioassay results (9)(17)(18)(19), however, show that limited quantities of dietary phyto-oestrogens are acceptable and do not reduce the sensitivity of the bioassay. As a guide, dietary levels of phyto-oestrogens should not exceed 350  $\mu\text{g}$  of genistein equivalents/gram of laboratory diet for immature female Sprague Dawley and Wistar rats (6)(9). Such diets should also be appropriate when testing in young adult ovariectomised rats because food consumption on a body weight basis is less in young adult as compared to immature animals. If adult ovariectomised mice or more phyto-oestrogen-sensitive rats are to be used, proportional reduction in dietary phyto-oestrogen levels must be considered (20). In addition, the differences in available metabolic energy from different diets may lead to time shifts for the onset of puberty (21)(22).

**▼ M5**

23. Prior to the study, careful selection is required of a diet without an elevated level of phyto-oestrogens (for guidance see (6)(9)) or metabolisable energy, that can confound the results (15)(17)(19)(22)(36). Ensuring the proper performance of the test system used by the laboratory as specified in paragraphs 26 and 27 is an important check on both of these factors. As a safeguard consistent with good laboratory practice (GLP) representative sampling of each batch of diet administered during the study should be conducted for possible analysis of phyto-oestrogen content (e.g. in the case of high uterine control weight relative to historic controls or an inadequate response to the reference oestrogen, 17 alpha ethinyl estradiol). Aliquots should be analysed as part of the study or frozen at – 20 °C or in such a way as to prevent the sample from decomposing prior to analysis.
  
24. Some bedding materials may contain naturally occurring oestrogenic or antioestrogenic chemicals (e.g. corn cob is known to affect the cyclicity of rats and appears to be antioestrogenic). The selected bedding material should contain a minimum level of phyto-oestrogens.

**Preparation of animals**

25. Experimental animals without evidence of any disease or physical abnormalities are randomly assigned to the control and treatment groups. Cages should be arranged in such a way that possible effects due to cage placement are minimised. The animals should be identified uniquely. Preferably, immature animals should be caged with dams or foster dams until weaning during acclimatisation. The acclimatisation period prior to the start of the study should be about 5 days for young adult animals and for the immature animals delivered with dams or foster dams. If immature animals are obtained as weanlings without dams a shorter duration of the acclimatisation period may become necessary as dosing should start immediately after weaning (see paragraph 29).

**PROCEDURE****Verification of Laboratory Proficiency**

26. Two different options can be used to verify laboratory proficiency:
  - Periodic verification, relying on an initial baseline positive control study (see paragraph 27). At least every 6 months and each time there is a change that may influence the performance of the assay (e.g. a new formulation of diet, change in personnel performing dissections, change in animal strain or supplier, etc.), the responsiveness of the test system (animal model) should be verified using an appropriate dose (based on the baseline positive control study described in paragraph 27) of a reference oestrogen: 17 $\alpha$ -ethinyl estradiol (CAS No 57-63-6) (EE).
  
  - Use of concurrent controls, by including a group administered with an appropriate dose of reference oestrogen in each assay.

If the system does not respond as expected, the experimental conditions should be examined and modified accordingly. It is recommended that the dose of reference oestrogen to be used in either approach be approximately the ED70 to 80.



**▼ M5**

27. **Baseline Positive Control Study** — Before a laboratory conducts a study under this test method for the first time, laboratory proficiency should be demonstrated by testing the responsiveness of the animal model, by establishing the dose response of a reference oestrogen: 17 $\alpha$ -ethinyl estradiol (CAS No 57-63-6) (EE) with a minimum of four doses. The uterine weight response will be compared to established historical data (see reference (5)). If this baseline positive control study does not yield the anticipated results the experimental conditions should be examined and modified.

**Number and condition of animals**

28. Each treated and control group should include at least 6 animals (for both immature and ovx-adult method protocols).

**Age of immature animals**

29. For the Uterotrophic Bioassay with immature animals the day of birth must be specified. Dosing should begin early enough to ensure that, at the end of test chemical administration, the physiological rise of endogenous oestrogens associated with puberty has not yet taken place. On the other hand, there is evidence that very young animals may be less sensitive. For defining the optimal age each laboratory should take its own background data on maturation into consideration.

As a general guide, dosing in rats may begin immediately after early weaning on postnatal day 18 (with the day of birth being postnatal day 0). Dosing in rats preferably should be completed on postnatal day 21 but in any case prior to postnatal day 25, because, after this age, the hypothalamic-pituitary-ovarian axis becomes functional and endogenous oestrogen levels may begin to rise with a concomitant increase in baseline uterine weight means and an increase in the group standard deviations (2)(3)(10)(11)(12).

**Procedure for ovariectomy**

30. For the ovariectomised female rat and mouse (treatment and control groups), ovariectomy should occur between 6 and 8 weeks of age. For rats, a minimum of 14 days should elapse between ovariectomy and the first day of administration in order to allow the uterus to regress to a minimum, stable baseline. For mice, at least 7 days should elapse between ovariectomy and the first day of administration. As small amounts of ovarian tissue are sufficient to produce significant circulating levels of oestrogens (3), the animals should be tested prior to use by observing epithelial cells swabbed from the vagina on at least five consecutive days (e.g. days 10-14 after ovariectomy for rats). If the animals indicate any evidence entering oestrous, the animals should not be used. Further, at necropsy, the ovarian stubs should be examined for any evidence that ovarian tissue is present. If so, the animal should not be used in the calculations (3).
31. The ovariectomy procedure begins with the animal in ventral recumbency after the animal has been properly anaesthetised. The incision opening the dorso-lateral abdominal wall should be approximately 1 cm lengthways at the mid-point between the costal inferior border and the iliac crest, and a few millimetres lateral to the lateral margin of the lumbar muscle. The ovary should be removed from the abdominal cavity onto an aseptic field. The ovary should be disconnected at the junction of the oviduct and the uterine body. After confirming that no massive bleeding is occurring, the abdominal wall should be closed by a suture and the skin closed by autoclips or appropriate suture. The ligation points are shown schematically in Figure 1. Appropriate post-operative analgesia should be used as recommended by a veterinarian experienced in rodent care.

**▼ M5****Body weight**

32. In the ovx-adult method, body weight and uterine weight are not correlated because uterine weight is affected by hormones like oestrogens but not by the growth factors that regulate body size. On the contrary, body weight is related to uterine weight in the immature model, while it is maturing (34). Thus, at the commencement of the study the weight variation of animals used, in the immature model, should be minimal and not exceed  $\pm 20\%$  of the mean weight. This means that the litter size should be standardised by the breeder, to ensure that offspring of different mother animals will be fed approximately the same. Animals should be assigned to groups (both control and treatment) by randomised weight distribution, so that mean body weight of each group is not statistically different from any other group. Consideration should be given to avoid assignment of littermates to the same treatment group as far as practicable without increasing the number of litters to be used for the investigation.

**Dosage**

33. In order to establish whether a test chemical can have oestrogenic action *in vivo*, two dose groups and a control are normally sufficient and this design is therefore preferred for animal welfare reasons. If the purpose is either to obtain a dose-response curve or to extrapolate to lower doses, at least 3 dose groups are needed. If information beyond identification of oestrogenic activity (such as an estimate of potency) is required, a different dosing regimen should be considered. Except for treatment with the test chemical, animals in the control group should be handled in an identical manner to the test group subjects. If a vehicle is used in administering the test chemical, the control group should receive the same amount of vehicle used with the treated groups (or highest volume used with the test groups if different among groups).
  
34. The objective in the case of the Uterotrophic Bioassay is to select doses that ensure animal survival and that are without significant toxicity or distress to the animals after three consecutive days of chemical administration up to a maximum dose of 1 000 mg/kg/d. All dose levels should be proposed and selected taking into account any existing toxicity and (toxico-) kinetic data available for the test chemical or related materials. The highest dose level should first take into consideration the LD50 and/or acute toxicity information in order to avoid death, severe suffering or distress in the animals (24)(25)(26). The highest dose should represent the maximum tolerated dose (MTD); a study conducted at a dose level that induced a positive uterotrophic response would be accepted too. As a screen, large intervals (e.g. one half log units corresponding to a dose progression of 3,2 or even up to one log units) between dosages are generally acceptable. If there are no suitable data available, a range finding study may be performed to aid the determination of the doses to be used.
  
35. Alternatively, if the oestrogenic potency of an agonist can be estimated by *in vitro* (or *in silico*) data, these may be taken into consideration for dose selection. For example, the amount of the test chemical that would produce uterotrophic responses equivalent to the reference agonist (ethinyl estradiol) is estimated by its relative *in vitro* potencies to ethinyl estradiol. The highest test dose would be given by multiplying this equivalent dose by an appropriate factor e.g. 10 or 100.

**▼ M5****Considerations for range finding**

36. If necessary, a preliminary range finding study can be carried out with few animals. In this respect, OECD Guidance Document No 19(25) may be used defining clinical signs indicative of toxicity or distress to the animals. If feasible within this range finding study after three days of administration, the uteri may be excised and weighed approximately 24-hours after the last dose. These data could then be used to assist the main study design (select an acceptable maximum and lower doses and recommend the number of dose groups).

**Administration of doses**

37. The test chemical is administered by oral gavage or subcutaneous injection. Animal welfare considerations as well as toxicological aspects like the relevance to the human route of exposure to the chemical (e.g. oral gavage to model ingestion, subcutaneous injection to model inhalation or dermal adsorption), the physical/chemical properties of the test material and especially existing toxicological information and data on metabolism and kinetics (e.g. need to avoid first pass metabolism, better efficiency via a particular route) have to be taken into account when choosing the route of administration.
38. It is recommended that, wherever possible, the use of an aqueous solution/suspension be considered first. But as most oestrogen ligands or their metabolic precursors tend to be hydrophobic, the most common approach is to use a solution/suspension in oil (e.g. corn, peanut, sesame or olive oil). However, these oils have different caloric and fat content, thus the vehicle might affect total metabolisable energy (ME) intake, thereby potentially altering measured endpoints such as the uterine weight especially in the immature method (33). Thus, prior to the study, any vehicle to be used should be tested against controls without vehicles. Test chemicals can be dissolved in a minimal amount of 95 % ethanol or other appropriate solvents and diluted to final working concentrations in the test vehicle. The toxic characteristics of the solvent must be known, and should be tested in a separate solvent-only control group. If the test chemical is considered stable, gentle heating and vigorous mechanical action can be used to assist in dissolving the test chemical. The stability of the test chemical in the vehicle should be determined. If the test chemical is stable for the duration of the study, then one starting aliquot of the test chemical may be prepared, and the specified dosage dilutions prepared daily.
39. Dosage timing will depend of the model used (refer to paragraph 29 for the immature model and to paragraph 30 for ovx-adult model). Immature female rats are dosed with the test chemical daily for three consecutive days. A three-day treatment is also recommended for ovariectomised female rats but longer exposures are acceptable and may improve the detection of weakly active chemicals. With ovariectomised female mice, an application duration of 3 days should be sufficient without a significant advantage by an extension of up to seven days for strong oestrogen agonists, however, this relation was not demonstrated for weak oestrogens in the validation study (16) thus dosage should be extended up to 7 consecutive days in ovx-adult mice. The dose should be given at similar times each day. They should be adjusted as necessary to maintain a constant dose level in terms of animal body weight (e.g. mg of test chemical per kg of body weight per day). Regarding the test volume, its variability, on a body weight basis, should be minimised by adjusting the concentration of the dosing solution to ensure a constant volume on a body weight basis at all dose levels and for any route of administration.

**▼ M5**

40. When the test chemical is administered by gavage, this should be done in a single daily dose to the animals using a stomach tube or a suitable intubation cannula. The maximum volume of liquid that can be administered at one time depends on the size of the test animal. Local animal care guidelines should be followed, but the volume should not exceed 5 ml/kg body weight, except in the case of aqueous solutions where 10 ml/kg body weight may be used.
  
41. When the test chemical is administered by subcutaneous injection, this should be done in a single daily dose. Doses should be administered to the dorsoscapular or lumbar regions via sterile needle (e.g. 23- or 25-gauge) and a tuberculin syringe. Shaving the injection site is optional. Any losses, leakage at the injection site or incomplete dosing should be recorded. The total volume injected per rat per day should not exceed 5 ml/kg body weight, divided into 2 injection sites, except in the case of aqueous solutions where 10 ml/kg body weight may be used.

**Observations***General and clinical observations*

42. General clinical observations should be made at least once a day and more frequently when signs of toxicity are observed. Observations should be carried out preferably at the same time(s) each day and considering the period of anticipated peak effects after dosing. All animals are to be observed for mortality, morbidity and general clinical signs such as changes in behaviour, skin, fur, eyes, mucous membranes, occurrence of secretions and excretions and autonomic activity (e.g. lacrimation, piloerection, pupil size, unusual respiratory pattern).

*Body weight and food consumption*

43. All animals should be weighed daily to the nearest 0,1 g, starting just prior to initiation of treatment i.e. when the animals are allocated into groups. As an optional measurement, the amount of food consumed during the treatment period may be measured per cage by weighing the feeders. The food consumption results should be expressed in grams per rat per day.

*Dissection and measurement of uterus weight*

44. Twenty-four hours after the last treatment, the rats will be humanely killed. Ideally, the necropsy order will be randomised across groups to avoid progression directly up or down dose groups that could subtly affect the data. The bioassay objective is to measure both the wet and blotted uterus weights. The wet weight includes the uterus and the luminal fluid contents. The blotted weight is measured after the luminal contents of the uterus have been expressed and removed.
  
45. Before dissection the vagina will be examined for opening status in immature animals. The dissection procedure begins by opening the abdominal wall starting at the pubic symphysis. Then, uterine horn and ovaries, if present, are detached from the dorsal abdominal wall. The urinary bladder and ureters are removed from the ventral and lateral side of uterus and vagina. Fibrous adhesion between the rectum and the vagina is detached until the junction of vaginal orifice and perineal skin can be identified. The uterus and vagina are detached from the body by incising the vaginal wall just above the junction between perineal skin as shown in

▼ **M5**

Figure 2. The uterus should be detached from the body wall by gently cutting the uterine mesentery at the point of its attachment along the full length of the dorsolateral aspect of each uterine horn. Once removed from the body, uterine handling should be sufficiently rapid to avoid desiccation of the tissues. Loss of weight due to desiccation becomes more important with small tissues such as the uterus (23). If ovaries are present, the ovaries are removed at the oviduct avoiding loss of luminal fluid from the uterine horn. If the animal has been ovariectomised, the stubs should be examined for the presence of any ovarian tissue. Excess fat and connective tissue should be trimmed away. The vagina is removed from the uterus just below the cervix so that the cervix remains with the uterine body as shown in Figure 2.

46. Each uterus should be transferred to a uniquely marked and weighed container (e.g. a petri-dish or plastic weight boat) with continuing care to avoid desiccation before weighing (e.g. filter paper slightly dampened with saline may be placed in the container). The uterus with luminal fluid will be weighed to the nearest 0,1 mg (wet uterine weight).
47. Each uterus will then be individually processed to remove the luminal fluid. Both uterine horns will be pierced or cut longitudinally. The uterus will be placed on lightly moistened filter paper (e.g. Whatman No 3) and gently pressed with a second piece of lightly moistened filter paper to completely remove the luminal fluid. The uterus without the luminal contents will be weighed to the nearest 0,1 mg (blotted uterine weight).
48. The uterus weight at termination can be used to ensure that the appropriate age in the immature intact rat was not exceeded, however, the historical data of the rat strain used by the laboratory are decisive in this respect (see paragraph 56 for interpretation of the results).

*Optional investigations*

49. After weighing, the uterus may be fixed in 10 % neutral buffered formalin to be examined histopathologically after Haematoxylin & Eosin (HE)-staining. The vagina may be investigated accordingly (see paragraph 9). In addition, morphometric measurement of endometrial epithelium may be done for quantitative comparison.

**DATA AND REPORTING****Data**

50. Study data should include:

- the number of animals at the start of the assay,
- the number and identity of animals found dead during the assay or killed for humane reasons and the date and time of any death or humane kill,
- the number and identity of animals showing signs of toxicity, and a description of the signs of toxicity observed, including time of onset, duration, and severity of any toxic effects, and
- the number and identity of animals showing any lesions and a description of the type of lesions.

**▼ M5**

51. Individual animal data should be recorded for the body weights, the wet uterine weight, and the blotted uterine weight. One-tailed statistical analyses for agonists should be used to determine whether the administration of a test chemical resulted in a statistically significant ( $p < 0,05$ ) increase in the uterine weight. Appropriate statistical analyses should be carried out to test for treatment related changes in blotted and wet uterine weight. For example, the data may be evaluated by an analysis of covariance (ANCOVA) approach with body weight at necropsy as the co-variable. A variance-stabilising logarithmic transformation may be carried out on the uterine data prior to the data analysis. Dunnett and Hsu's test are appropriate for making pair wise comparisons of each dosed group to vehicle controls and to calculate the confidence intervals. Studentised residual plots can be used to detect possible outliers and to assess homogeneity of variances. These procedures were applied in the OECD validation programme using the PROC GLM in the Statistical Analysis System (SAS Institute, Cary, NC), version 8 (6)(7).

52. A final report shall include:

*Testing facility:*

- Responsible personnel and their study responsibilities
- Data from the Baseline Positive Control Test and periodic positive control data (see paragraphs 26 and 27)

*Test chemical:*

- Characterisation of test chemicals
- Physical nature and where relevant physicochemical properties
- Method and frequency of preparation of dilutions
- Any data generated on stability
- Any analyses of dosing solutions

*Vehicle:*

- Characterisation of test vehicle (nature, supplier and lot)
- Justification of choice of vehicle (if other than water)

*Test animals:*

- Species and strain and justification for their choice
- Supplier and specific supplier facility
- Age on supply with birth date
- If immature animals, whether or not supplied with dam or foster dam and date of weaning
- Details of animal acclimatisation procedure
- Number of animals per cage
- Detail and method of individual animal and group identification

*Assay Conditions:*

- Details of randomisation process (i.e. method used)
- Rationale for dose selection

**▼ M5**

- Details of test chemical formulation, its achieved concentrations, stability and homogeneity
- Details of test chemical administration and rationale for the choice of exposure route
- Diet (name, type, supplier, content, and, if known, phyto-oestrogen levels)
- Water source (e.g. tap water or filtered water) and supply (by tubing from a large container, in bottles, etc.)
- Bedding (name, type, supplier, content)
- Record of caging conditions, lighting interval, room temperature and humidity, room cleaning
- Detailed description of necropsy and uterine weighing procedures
- Description of statistical procedures

*Results**For individual animals:*

- All daily individual body weights (from allocation into groups through necropsy) (to the nearest 0,1 g)
- Age of each animal (in days counting day of birth as day 0) when administration of test chemical begins
- Date and time of each dose administration
- Calculated volume and dosage administered and observations of any dosage losses during or after administration
- Daily record of status of animal, including relevant symptoms and observations
- Suspected cause of death (if found during study in moribund state or dead)
- Date and time of humane killing with time interval to last dosing
- Wet uterine weight (to the nearest 0,1 mg) and any observations of luminal fluid losses during dissection and preparation for weighing
- Blotted uterine weight (to the nearest 0,1 mg)

*For each group of animals:*

- Mean daily body weights (to the nearest 0,1 g) and standard deviations (from allocation into groups through necropsy)
- Mean wet uterine weights and mean blotted uterine weights (to the nearest 0,1 mg) and standard deviations
- If measured, daily food consumption (calculated as grams of food consumed per animal)

▼ **M5**

- The results of statistical analyses comparing both the wet and blotted uterine weights of treated groups relative to the same measures in the vehicle control groups.
  
- The results of statistical analysis comparing the total body weight and the body weight gain of treated groups relative to the same measures in the vehicle control groups.

## 53. Summary of the important guidance facts of the test method

	<b>Rat</b>	<b>Mice</b>
<b>Animals</b>		
Strain	Commonly used laboratory rodent strain	
Number of animals	A minimum of 6 animals per dose group	
Number of groups	A minimum of 2 test groups (see paragraph 33 for guidance) and a negative control group For guidance on positive control groups see paragraphs 26 and 27	
<b>Housing and feeding conditions</b>		
T° in animal room	22 °C ± 3 °C	
Relative humidity	50-60 % and not below 30 % or above 70 %	
Daily lighting sequence	12 hours light, 12 hours dark	
Diet and drinking water	Ad libitum	
Housing	Individually or in groups of up to three animals (social group housing is recommended for immature animals)	
Diet and bedding	Low level of phyto-oestrogens recommended in diet and bedding	
<b>Protocol</b>		
Method	Immature non-ovariectomised method (the preferred one). Ovariectomised adult female method	Ovariectomised adult female method
Age of dosing for immature animals	PND 18 at the earliest. Dosing should be completed prior to PND 25	Not relevant under the scope of the current test method.
Age of ovariectomy	Between 6 and 8 weeks of age.	
Age of dosing for ovariectomised animals	A minimum of 14 days should elapse between ovariectomy and the 1 <sup>st</sup> day of administration.	A minimum of 7 days should elapse between ovariectomy and the 1 <sup>st</sup> day of administration.
Body weight	Body weight variation should be minimal and not exceed ± 20 % of the mean weight.	



▼ **M5**

	<b>Rat</b>	<b>Mice</b>
<b>Dosing</b>		
Route of administration	Oral gavage or subcutaneous injection	
Frequency of administration	Single daily dose	
Volume amount for gavage and injection	≤ 5 ml/kg body weight (or up to 10 ml/kg body weight in case of aqueous solutions) (in 2 injection sites for subcutaneous route)	
Duration of administration	3 consecutive days for immature model Minimum of 3 consecutive days for the OVX model	7 consecutive days for the OVX model
Time of necropsy	Approximately 24 hours after the last dose	
<b>Results</b>		
Positive response	Statistically significant increase of the mean uterus weight (wet and/or blotted)	
Reference oestrogen	17 $\alpha$ -ethinyl estradiol	

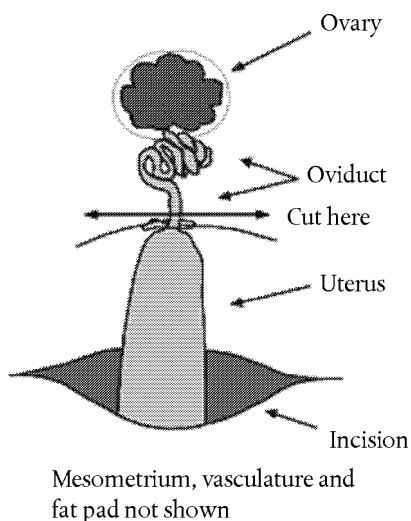
## GUIDANCE FOR THE INTERPRETATION AND ACCEPTANCE OF THE RESULTS

54. In general, a test for oestrogenicity should be considered positive if there is a statistically significant increase in uterine weight ( $p < 0,05$ ) at least at the high dose level as compared to the solvent control group. A positive result is further supported by the demonstration of a biologically plausible relationship between the dose and the magnitude of the response, bearing in mind that overlapping oestrogenic and antioestrogenic activities of the test chemical may affect the shape of the dose-response curve.
55. Care must be taken in order not to exceed the maximum tolerated dose to allow a meaningful interpretation of the data. Reduction of body weight, clinical signs and other findings should be thoroughly assessed in this respect.
56. An important consideration for the acceptance of the data from the Uterotrophic Bioassay is the uterine weights of the vehicle control group. High control values may compromise the responsiveness of the bioassay and the ability to detect very weak oestrogen agonists. Literature reviews and the data generated during the validation of the Uterotrophic Bioassay suggest that instances of high control means do occur spontaneously, particularly in immature animals (2)(3)(6)(9). As the uterine weight of immature rats depends on many variables like strain or body weight, no definitive upper limit for the uterine weight can be given. As a guide, if blotted uterine weights in immature control rats are comprised between 40 and 45 mg, results should be considered as suspicious and uterine weights above 45 mg may lead to rerun the test. However, this needs to be considered on a case by case basis (3)(6)(8). When testing in adult rats incomplete ovariectomy will leave ovarian tissue that can produce endogenous oestrogen and retard the regression of the uterine weight.

▼ **M5**

57. Blotted vehicle control uterine weights less than 0,09 % of body weight for immature female rats and less than 0,04 % for ovariectomised young adult females appear to yield acceptable results (see Table 31 (2)). If the control uterine weights are greater than these numbers, various factors should be scrutinised including the age of the animals, proper ovariectomy, dietary phyto-oestrogens, and so on, and a negative assay result (no indication for oestrogenic activity) should be used with caution.
58. Historical data for vehicle control groups should be maintained in the laboratory. Historical data for responses to positive reference oestrogens, such as 17 $\alpha$ -ethinyl estradiol, should also be maintained in the laboratory. Laboratories may also test the response to known weak oestrogen agonists. All these data can be compared to available data (2)(3)(4)(5)(6)(7)(8) to ensure that the laboratory's methods yield sufficient sensitivity.
59. The blotted uterine weights showed less variability in the course of the OECD validation study than the wet uterine weights (6)(7). However, a significant response in either measure would indicate that the test chemical is positive for oestrogenic activity.
60. The uterotrophic response is not entirely of oestrogenic origin, however, a positive result of the Uterotrophic Bioassay should generally be interpreted as evidence for oestrogenic potential *in vivo*, and should normally initiate actions for further clarification (see paragraph 9 and the 'OECD Conceptual Framework for the Testing and Assessment of Endocrine Disrupting Chemicals', Annex 2).

Figure 1

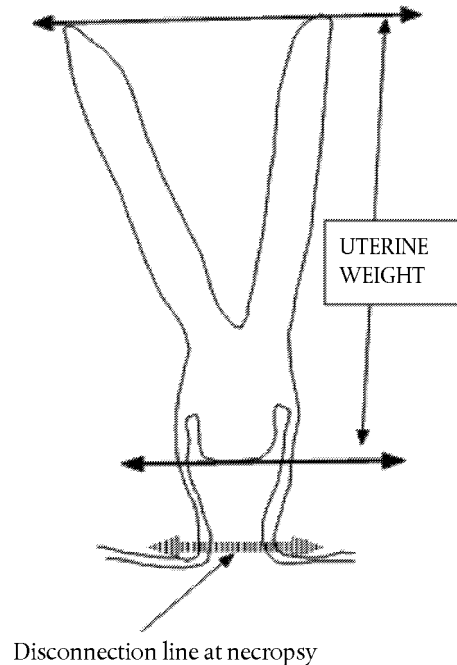
**Schematic diagram showing the surgical removal of the ovaries**

The procedure begins by opening dorso-lateral abdominal wall at the mid-point between the costal inferior border and the iliac crest, and a few millimetres lateral to the lateral margin of the lumbar muscle. Within the abdominal cavity, the ovaries should be located. On an aseptic field, the ovaries are then physically removed from the abdominal cavity, a ligature placed between the ovary and uterus to control bleeding, and the ovary detached by incision above the ligature at the junction of the oviduct and each uterine horn. After confirming that no significant bleeding persists, the abdominal wall should be closed by suture, and the skin closed, e.g. by autoclips or suture. The animals should be allowed to recover and the uterus weight to regress for a minimum of 14 days before use.

▼ M5

Figure 2

The removal and preparation of the uterine tissues for weight measurement.



The procedure begins by opening the abdominal wall at the pubic symphysis. Then, each ovary, if present and uterine horn is detached from the dorsal abdominal wall. Urinary bladder and ureters are removed from the ventral and lateral side of uterus and vagina. Fibrous adhesion between the rectum and the vagina are detached until the junction of vaginal orifice and perineal skin can be identified. The uterus and vagina are detached from the body by incising the vaginal wall just above the junction between perineal skin as shown in the figure. The uterus should be detached from the body wall by gently cutting the uterine mesentery at the point of its attachment along the full length of the dorsolateral aspect of each uterine horn. After removal from the body, the excess fat and connective tissue is trimmed away. If ovaries are present, the ovaries are removed at the oviduct avoiding loss of luminal fluid from the uterine horn. If the animal has been ovariectomised, the stubs should be examined for the presence of any ovarian tissue. The vagina is removed from the uterus just below the cervix so that the cervix remains with the uterine body as shown in the figure. The uterus can then be weighed.

▼ **M5***Appendix 1*

## DEFINITIONS:

**Antioestrogenicity** is the capability of a chemical to suppress the action of estradiol 17 $\beta$  in a mammalian organism.

**Chemical** means a substance or a mixture.

**Date of birth** is postnatal day 0.

**Dosage** is a general term comprising of dose, its frequency and the duration of dosing.

**Dose** is the amount of test chemical administered. For the Uterotrophic Bioassay, the dose is expressed as weight of test chemical per unit body weight of test animal per day (e.g. mg/kg body weight/day).

**Maximum Tolerable Dose (MTD)** is the highest amount of a chemical that, when introduced into the body does not kill test animals (denoted by LD<sub>0</sub>) (IUPAC, 1993)

**Oestrogenicity** is the capability of a chemical to act like estradiol 17 $\beta$  in a mammalian organism.

**Postnatal day X** is the Xth day of life after the day of birth.

**Sensitivity** is the proportion of all positive/active chemicals that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results, and is an important consideration in assessing the relevance of a test method.

**Specificity** is the proportion of all negative/inactive chemicals that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results and is an important consideration in assessing the relevance of a test method.

**Test chemical** means any substance or mixture tested using this test method.

**Uterotrophic** is a term used to describe a positive influence on the growth of uterine tissues.

**Validation** is a scientific process designed to characterise the operational requirements and limitations of a test method and to demonstrate its reliability and relevance for a particular purpose.

Appendix 2

**Note:** Document prepared by the Secretariat of the Test Guidelines Programme based on the agreement reached at the 6th Meeting of the EDTA Task Force

**OECD Conceptual Framework for the Testing and Assessment of Endocrine Disrupting Chemicals**

<p><b>Level 1</b> Sorting &amp; prioritization based upon existing information</p>	<ul style="list-style-type: none"> <li>— physical &amp; chemical properties, e.g., MW, reactivity, volatility, biodegradability,</li> <li>— human &amp; environmental exposure, e.g., production volume, release, use patterns</li> <li>— hazard, e.g., available toxicological data</li> </ul>	
<p><b>Level 2</b> <i>In vitro</i> assays providing mechanistic data</p>	<ul style="list-style-type: none"> <li>— ER, AR, TR receptor binding affinity</li> <li>— Transcriptional activation</li> <li>— Aromatase and steroidogenesis <i>in vitro</i></li> <li>— Aryl hydrocarbon receptor recognition/binding</li> <li>— QSARs</li> </ul>	<ul style="list-style-type: none"> <li>— High Through Put Prescreens</li> <li>— Thyroid function</li> <li>— Fish hepatocyte VTG assay</li> <li>— Others (as appropriate)</li> </ul>
<p><b>Level 3</b> <i>In vivo</i> assays providing data about single endocrine Mechanisms and effects</p>	<ul style="list-style-type: none"> <li>— Uterotrophic assay (estrogenic related)</li> <li>— Hershberger assay (androgenic related)</li> <li>— Non -receptor mediated hormone function</li> <li>— Others (e.g. thyroid)</li> </ul>	<ul style="list-style-type: none"> <li>— Fish VTG (vitellogenin) assay (estrogenic related)</li> </ul>
<p><b>Level 4</b> <i>In vivo</i> assays providing data about multiple endocrine Mechanisms and effects</p>	<ul style="list-style-type: none"> <li>— enhanced OECD 407 (endpoints based on endocrine mechanisms)</li> <li>— male and female pubertal assays</li> <li>— adult intact male assay</li> </ul>	<ul style="list-style-type: none"> <li>— Fish gonadal histopathology assay</li> <li>— Frog metamorphosis assay</li> </ul>
<p><b>Level 5</b> <i>In vivo</i> assays providing data on effects from endocrine &amp; other mechanisms</p>	<ul style="list-style-type: none"> <li>— 1-generation assay (TG415 enhanced)<sup>1</sup></li> <li>— 2-generation assay (TG416 enhanced)<sup>1</sup></li> <li>— reproductive screening test (TG421 enhanced)<sup>1</sup></li> <li>— combined 28 day/reproduction screening test (TG 422 enhanced)<sup>1</sup></li> </ul> <p><sup>1</sup> Potential enhancements will be considered by VMG mamm</p>	<ul style="list-style-type: none"> <li>— Partial and full life cycle assays in fish, birds, amphibians &amp; invertebrates (developmental and reproduction)</li> </ul>

VMG mamm: Validation Management Group on Mammalian Testing and Assessment

**▼ M5**

## NOTES TO THE FRAMEWORK:

- Note 1:* Entering at all levels and exiting at all levels is possible and depends upon the nature of existing information needs for hazard and risk assessment purposes
- Note 2:* In level 5, ecotoxicology should include endpoints that indicate mechanisms of adverse effects, and potential population damage
- Note 3:* When a multimodal model covers several of the single endpoint assays, that model would replace the use of those single endpoint assays
- Note 4:* The assessment of each chemical should be based on a case by case basis, taking into account all available information, bearing in mind the function of the framework levels.
- Note 5:* The framework should not be considered as all inclusive at the present time. At levels 3, 4 and 5 it includes assays that are either available or for which validation is under way. With respect to the latter, these are provisionally included. Once developed and validated, they will be formally added to the framework.
- Note 6:* Level 5 should not be considered as including definitive tests only. Tests included at that level are considered to contribute to general hazard and risk assessment.

▼ **M5****LITERATURE**

- (1) OECD (1998). Report of the First Meeting of the OECD Endocrine Disrupter Testing and Assessment (EDTA) Task Force, 10<sup>th</sup>-11<sup>th</sup> March 1998, ENV/MC/CHEM/RA(98)5.
- (2) OECD (2003). Detailed Background Review of the Uterotrophic Bioassay: Summary of the Available Literature in Support of the Project of the OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA) to Standardise and Validate the Uterotrophic Bioassay. OECD Environmental Health and Safety Publication Series on Testing and Assessment No 38. ENV/JM/MONO(2003)1.
- (3) Owens JW, Ashby J. (2002). Critical Review and Evaluation of the Uterotrophic Bioassay for the Identification of Possible Estrogen Agonists and Antagonists: In Support of the Validation of the OECD Uterotrophic Protocols for the Laboratory Rodent. Crit. Rev. Toxicol. 32:445-520.
- (4) OECD (2006). OECD Report of the Initial Work Towards the Validation of the Rodent Uterotrophic Assay — Phase 1. OECD Environmental Health and Safety Publication Series on Testing and Assessment No 65. ENV/JM/MONO(2006)33.
- (5) Kanno, J, Onyon L, Haseman J, Fenner-Crisp P, Ashby J, Owens W. (2001). The OECD program to validate the rat uterotrophic bioassay to screen compounds for *in vivo* estrogenic responses: Phase 1. Environ Health Perspect. 109:785-94.
- (6) OECD (2006). OECD Report of the Validation of the Rodent Uterotrophic Bioassay: Phase 2 — Testing of Potent and Weak Oestrogen Agonists by Multiple Laboratories. OECD Environmental Health and Safety Publication Series on Testing and Assessment No 66. ENV/JM/MONO(2006)34.
- (7) Kanno J, Onyon L, Peddada S, Ashby J, Jacob E, Owens W. (2003). The OECD program to validate the rat uterotrophic bioassay: Phase Two — Dose Response Studies. Environ. Health Persp.111:1530-1549
- (8) Kanno J, Onyon L, Peddada S, Ashby J, Jacob E, Owens W. (2003). The OECD program to validate the rat uterotrophic bioassay: Phase Two — Coded Single Dose Studies. Environ. Health Persp.111:1550-1558.
- (9) Owens W, Ashby J, Odum J, Onyon L. (2003). The OECD program to validate the rat uterotrophic bioassay: Phase Two — Dietary phytoestrogen analyses. Environ. Health Persp. 111:1559-1567.
- (10) Ogasawara Y, Okamoto S, Kitamura Y, Matsumoto K. (1983). Proliferative pattern of uterine cells from birth to adulthood in intact, neonatally castrated, and/or adrenalectomized mice assayed by incorporation of [<sup>125</sup>I]iododeoxyuridine. Endocrinology 113:582-587.
- (11) Branham WS, Sheehan DM, Zehr DR, Ridlon E, Nelson CJ. (1985). The postnatal ontogeny of rat uterine glands and age-related effects of 17 $\beta$ -estradiol. Endocrinology 117:2229-2237.
- (12) Schlumpf M, Berger L, Cotton B, Conscience-Egli M, Durrer S, Fleischmann I, Haller V, Maerker K, Lichtensteiger W. (2001). Estrogen active UV screens. SÖFW-J. 127:10-15.
- (13) Zarrow MX, Lazo-Wasem EA, Shoger RL. (1953). Estrogenic activity in a commercial animal ration. Science 118:650-651.
- (14) Drane HM, Patterson DSP, Roberts BA, Saba N. (1975). The chance discovery of oestrogenic activity in laboratory rat cake. Fd. Cosmet. Toxicol. 13:425-427.
- (15) Boettger-Tong H, Murphy L, Chiappetta C, Kirkland JL, Goodwin B, Adlercreutz H, Stancel GM, Makela S. (1998). A case of a laboratory animal feed with high estrogenic activity and its impact on *in vivo* responses to exogenously administered estrogens. Environ. Health Perspec.106:369-373.

▼ **M5**

- (16) OECD (2007). Additional data supporting the Test Guideline on the Uterotrophic Bioassay in rodents. OECD Environmental Health and Safety Publication Series on Testing and Assessment No 67.
- (17) Degen GH, Janning P, Diel P, Bolt HM. (2002). Estrogenic isoflavones in rodent diets. *Toxicol. Lett.* 128:145-157.
- (18) Wade MG, Lee A, McMahon A, Cooke G, Curran I. (2003). The influence of dietary isoflavone on the uterotrophic response in juvenile rats. *Food Chem. Toxicol.* 41:1517-1525.
- (19) Yamasaki K, Sawaki M, Noda S, Wada T, Hara T, Takatsuki M. (2002). Immature uterotrophic assay of estrogenic compounds in rats given different phytoestrogen content diets and the ovarian changes in the immature rat uterotrophic of estrogenic compounds with ICI 182,780 or antide. *Arch. Toxicol.* 76:613-620.
- (20) Thigpen JE, Haseman JK, Saunders HE, Setchell KDR, Grant MF, Forsythe D. (2003). Dietary phytoestrogens accelerate the time of vaginal opening in immature CD-1 mice. *Comp. Med.* 53:477-485.
- (21) Ashby J, Tinwell H, Odum J, Kimber I, Brooks AN, Pate I, Boyle CC. (2000). Diet and the aetiology of temporal advances in human and rodent sexual development. *J. Appl. Toxicol.* 20:343-347.
- (22) Thigpen JE, Lockear J, Haseman J, Saunders HE, Caviness G, Grant MF, Forsythe DB. (2002). Dietary factors affecting uterine weights of immature CD-1 mice used in uterotrophic bioassays. *Cancer Detect. Prev.* 26:381-393.
- (23) Thigpen JE, Li L-A, Richter CB, Lebetkin EH, Jameson CW. (1987). The mouse bioassay for the detection of estrogenic activity in rodent diets: I. A standardized method for conducting the mouse bioassay. *Lab. Anim. Sci.* 37:596-601.
- (24) OECD (2008). Acute oral toxicity — up-and-down procedure. OECD Guideline for the testing of chemicals No 425.
- (25) OECD (2000). Guidance document on the recognition, assessment and use of clinical signs as humane endpoints for experimental animals used in safety evaluation. Environmental Health and Safety Monograph Series on Testing and Assessment No 19. ENV/JM/MONO(2000)7.
- (26) OECD (2001). Guidance document on acute oral toxicity. Environmental Health and Safety Monograph Series on Testing and Assessment No 24. ENV/JM/MONO(2001)4.
- (27) Bulbring, E., and Burn, J.H. (1935). The estimation of oestrin and of male hormone in oily solution. *J. Physiol.* 85: 320 — 333.
- (28) Dorfman, R.I., Gallagher, T.F. and Koch, F.C (1936). The nature of the estrogenic substance in human male urine and bull testis. *Endocrinology* 19: 33 — 41.
- (29) Reel, J.R., Lamb IV, J.C. and Neal, B.H. (1996). Survey and assessment of mammalian estrogen biological assays for hazard characterization. *Fundam. Appl. Toxicol.* 34: 288 — 305.
- (30) Jones, R.C. and Edgren, R.A. (1973). The effects of various steroid on the vaginal histology in the rat. *Fertil. Steril.* 24: 284 — 291.
- (31) OECD (1982). Organization for Economic Co-operation and Development — Principles of Good Laboratory Practice, ISBN 92-64-12367-9, Paris.
- (32) Dorfman R.I. (1962). *Methods in Hormone Research*, Vol. II, Part IV: Standard Methods Adopted by Official Organization. New York, Academic Press.
- (33) Thigpen J. E. et al. (2004). Selecting the appropriate rodent diet for endocrine disruptor research and testing studies. *ILAR J* 45(4): 401-416.



**▼M5**

- (34) Gray L.E. and Ostby J. (1998). Effects of pesticides and toxic substances on behavioral and morphological reproductive development: endocrine versus non-endocrine mechanism. *Toxicol Ind Health*. 14 (1-2): 159-184.
- (35) Booth AN, Bickoff EM and Kohler GO. (1960). Estrogen-like activity in vegetable oils and mill by-products. *Science* 131:1807-1808.
- (36) Kato H, Iwata T, Katsu Y, Watanabe H, Ohta Y, Iguchi T (2004). Evaluation of estrogenic activity in diets for experimental animals using in vitro assay. *J. Agric Food Chem*. 52, 1410-1414.
- (37) OECD (2007). Guidance Document on the Uterotrophic Bioassay Procedure to Test for Antioestrogenicity. Series on Testing and Assessment. No 71.
- (38) Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes (OJ L 276, 20.10.2010, p. 33).

**▼ M5****B.55 HERSHBERGER BIOASSAY IN RATS: A SHORT-TERM SCREENING ASSAY FOR (ANTI)ANDROGENIC PROPERTIES**

## INTRODUCTION

1. This test method is equivalent to OECD Test Guideline (TG) 441 (2009). The OECD initiated a high-priority activity in 1998 to revise existing guidelines and to develop new guidelines for the screening and testing of potential endocrine disrupters (1). One element of the activity was to develop a test guideline for the rat Hershberger Bioassay. After several decades of use by the pharmaceutical industry, this assay was first standardised by an official expert committee in 1962 as a screening tool for androgenic chemicals (2). In 2001-2007, the rat Hershberger Bioassay has undergone an extensive validation programme including the generation of a Background Review Document (23), compilation of a detailed methods paper (3), development of a dissection guide (21) and the conduct of extensive intra- and interlaboratory studies to show the reliability and reproducibility of the bioassay. These validation studies were conducted with a potent reference androgen (testosterone propionate (TP)), two potent synthetic androgens (trenbolone acetate and methyl testosterone), a potent antiandrogenic pharmaceutical (flutamide), a potent inhibitor of the synthesis (finasteride) of the natural androgen (dihydrotestosterone-DHT), several weakly antiandrogenic pesticides (linuron, vinclozolin, procymidone, p,p' DDE), a potent 5 $\alpha$  reductase inhibitor (finasteride) and two known negative chemicals (dinitrophenol and nonylphenol) (4) (5) (6) (7) (8). This test method is the outcome of the long historical experience with the bioassay and the experience gained during the validation test programme and the results obtained therein.
2. The Hershberger Bioassay is a short-term *in vivo* screening test using accessory tissues of the male reproductive tract. The assay originated in the 1930s and was modified in the 1940s to include androgen-responsive muscles in the male reproductive tract (2) (9-15). In the 1960s, over 700 possible androgens were evaluated using a standardised version of the protocol (2) (14), and use of the assay for both androgens and anti-androgens was considered a standard method in the 1960s (2) (15). The current bioassay is based on the changes in weight of five androgen-dependent tissues in the castrate-peripubertal male rat. It evaluates the ability of a chemical to elicit biological activities consistent with androgen agonists, antagonists or 5 $\alpha$ -reductase inhibitors. The five target androgen-dependent tissues included in this test method are the ventral prostate (VP), seminal vesicle (SV) (plus fluids and coagulating glands), levator ani-bulbocavernosus (LABC) muscle, paired Cowper's glands (COW) and the glans penis (GP). In the castrate-peripubertal male rat, these five tissues all respond to androgens with an increase in absolute weight. When these same tissues are stimulated to increase in weight by administration of a potent reference androgen, these five tissues all respond to antiandrogens with a decrease in absolute weight. The primary model for the Hershberger bioassay has been the surgically castrated peripubertal male, which was validated in Phases 1, 2 and 3 of the Hershberger validation programme.
3. The Hershberger bioassay serves as a mechanistic *in vivo* screening assay for androgen agonists, androgen antagonists and 5 $\alpha$ -reductase inhibitors and its application should be seen in the context of the 'OECD Conceptual Framework for the Testing and Assessment of Endocrine Disrupting Chemicals' (Appendix 2). In this Conceptual Framework the Hershberger Bioassay is contained in Level 3 as an *in vivo* assay providing data

▼ M5

about a single endocrine mechanism, i.e. (anti)androgenicity. It is intended to be included in a battery of *in vitro* and *in vivo* tests to identify chemicals with potential to interact with the endocrine system, ultimately leading to hazard and risk assessments for human health or the environment.

4. Due to animal welfare concerns with the castration procedure, the intact (uncastrated) stimulated weanling male was sought as an alternative model for the Hershberger Bioassay to avoid the castration step. The stimulated weanling test method was validated (24); however, in the validation studies, the weanling version of the Hershberger Bioassay did not appear to be able to consistently detect effects on androgen-dependent organ weights from weak anti-androgens at the doses tested. Therefore, it was not included in this test method. However, recognising that its use may provide not only animal welfare benefits but also may provide information on other modes of action, it is available in OECD Guidance Document 115(25).

## INITIAL CONSIDERATIONS AND LIMITATIONS

5. Androgen agonists and antagonists act as ligands for the androgen receptor and may activate or inhibit, respectively, gene transcription controlled by the receptor. In addition, some chemicals inhibit the conversion of testosterone to the more potent natural androgen dihydrotestosterone in some androgen target tissues (5 $\alpha$ -reductase inhibitors). Such chemicals have the potential to lead to adverse health hazards, including reproductive and developmental effects. Therefore, the regulatory need exists to rapidly assess and evaluate a chemical as a possible androgen agonist or antagonist or 5 $\alpha$ -reductase inhibitor. While informative, the affinity of a ligand for an androgen receptor as measured by receptor binding or transcriptional activation of reporter genes *in vitro* is not the only determinant of possible hazard. Other determinants include metabolic activation and deactivation upon entering the body, chemical distribution to target tissues, and clearance from the body. This leads to the need to screen the possible activity of a chemical *in vivo* under relevant conditions and exposure. *In vivo* evaluation is less critical if the chemical's characteristics regarding Absorption — Distribution — Metabolism — Elimination (ADME) are known. Androgen-dependent tissues respond with rapid and vigorous growth to stimulation by androgens, particularly in castrate-peripubertal male rats. Rodent species, particularly the rat, are also widely used in toxicity studies for hazard characterisation. Therefore, the assay version, using the castrated peripubertal rat and the five target tissues in this assay, is appropriate for the *in vivo* screening of androgen agonists and antagonists and 5 $\alpha$ -reductase inhibitors.
6. This test method is based on those protocols employed in the OECD validation study which have been shown to be reliable and reproducible in intra- and inter-laboratory studies (4)(5)(6)(7)(8). Both androgen and anti-androgen procedures are presented in this test method.
7. Although there was some variation in the dose of TP used to detect anti-androgens in the OECD Hershberger Bioassay validation programme by the different laboratories (0,2 versus 0,4 mg/kg/d, subcutaneous injection) there was little difference between these two protocol variations in the ability to detect weak or strong antiandrogenic activity. However, it is clear that the dose of TP should not be too high to block the effects of weak androgen receptor (AR) antagonists or so low that the androgenic tissues display little growth response even without antiandrogen coadministration.

▼ **M5**

8. The growth response of the individual androgen-dependent tissues is not entirely of androgenic origin, i.e. chemicals other than androgen agonists can alter the weight of certain tissues. However, the growth response of several tissues concomitantly substantiates a more androgen-specific mechanism. For example, high doses of potent oestrogens can increase the weight of the seminal vesicles; however, the other androgen-dependent tissues in the assay do not respond in a similar manner. Antiandrogenic chemicals can act either as androgen receptor antagonists or 5 $\alpha$ -reductase inhibitors. 5 $\alpha$ -reductase inhibitors have a variable effect, because the conversion to more potent dihydrotestosterone varies by tissue. Antiandrogens that inhibit 5 $\alpha$ -reductase, like finasteride, have more pronounced effects in the ventral prostate than other tissues as compared to a potent AR antagonist, like flutamide. This difference in tissue response can be used to differentiate between AR mediated and 5 $\alpha$ -reductase mediated modes of action. In addition, the androgen receptor is evolutionarily related to that of other steroid hormones, and some other hormones, when administered at high, supraphysiological dosage levels, can bind and antagonise the growth-promoting effects of TP (13). Further, it also is plausible that enhanced steroid metabolism and a consequent lowering of serum testosterone could reduce androgen-dependent tissue growth. Therefore, any positive outcome in the Hershberger Bioassay should normally be evaluated using a weight of evidence approach, including *in vitro* assays, such as the AR and oestrogen receptor (ER) binding assays and corresponding transcriptional activation assays, or from other *in vivo* assays that examine similar androgen target tissues such as the male pubertal assay, 15-day intact adult male assay, or 28-day or 90-day repeat dose studies.
9. Experience indicates that xenobiotic androgens are rarer than xenobiotic antiandrogens. The expectation then is that the Hershberger bioassay will be used most often for the screening of antiandrogens. However, the procedure to test for androgens could, nevertheless, be recommended for steroidal or steroid-like chemicals or for chemicals for which an indication of possible androgenic effects was derived from methods contained in Level 1 or 2 of the conceptual framework (Appendix 2). Similarly, adverse effects associated with (anti)androgenic profiles may be observed in Level 5 assays, leading to the need to assess whether a chemical operates by an endocrine mode of action.
10. It is acknowledged that all animal-based procedures should conform to local standards of animal care; the descriptions of care and treatment set forth below are minimal performance standards, and will be superseded by local regulations such as Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes (26). Further guidance of the humane treatment of animals is given by the OECD (17).
11. As in any bioassay using experimental animals, careful considerations should be given to the necessity to carry out this study. Basically there may be two reasons for such a decision:
  - high exposure potential (Level 1 of the Conceptual Framework) or indications for (anti)androgenicity in *in vitro* assays (Level 2) supporting investigations whether such effects may occur *in vivo*;
  - effects consistent with (anti)androgenicity in Level 4 or 5 *in vivo* tests supporting investigations of the specific mode of action, e.g. to determine whether the effects were due to an (anti)androgenic mechanism.

**▼ M5**

12. Definitions used in this test method are given in Appendix 1.

## PRINCIPLE OF THE TEST

13. The Hershberger Bioassay achieves its sensitivity by using males with minimal endogenous androgen production. This is achieved through the use of castrated males provided an adequate time after castration for the target tissues to regress to a minimal and uniform baseline weight is allowed. Thus, when screening of potential androgenic activity, there are low endogenous levels of circulating androgens, the hypothalamic — pituitary — gonad axis is rendered unable to compensate via feedback mechanisms, the ability of the tissues to respond is maximised, and the starting tissue weight variability is minimised. When screening of potential anti-androgenic activity, a more consistent tissue weight gain can be achieved when the tissues are stimulated by a reference androgen. As a result, the Hershberger Bioassay requires only 6 animals per dose group whereas other assays with intact pubertal or adult males suggest using 15 males per dose group.
14. Castration of peripubertal male rats should be done in an appropriate manner using approved anaesthetics and aseptic technique. Analgesics should be administered on the first few days following surgery to eliminate post-surgical discomfort. Castration enhances the precision of the assay to detect weak androgens and antiandrogens by eliminating compensatory endocrine feed-back mechanisms present in the intact animal that can attenuate the effects of administered androgens and anti-androgens and by eliminating the large inter-individual variability in serum testosterone levels. Hence, castration reduces the numbers of animals required to screen for these endocrine activities.
15. When screening for potential androgenic activity, the test chemical is administered daily by oral gavage or subcutaneous (sc) injection for a period of 10 consecutive days. Test chemicals are administered to a minimum of two treatment groups of experimental animals using one dose level per group. The animals are necropsied approximately 24 hours after the last dose. A statistically significant increase in two or more target organ weights of the test chemical groups compared to the vehicle control group indicates that the test chemical is positive for potential androgenic activity (See paragraph 60). Androgens, like trenbolone that cannot be 5 $\alpha$ -reduced have more pronounced effects on the LABC and GP versus TP, but all tissues should display increased growth.
16. When screening for potential antiandrogenic activity, the test chemical is administered daily by oral gavage or subcutaneous injection for a period of 10 consecutive days in concert with daily TP doses (0,2 or 0,4 mg/kg/d) by sc injection. It was determined in the validation programme that either 0,2 or 0,4 mg/kg/d of TP could be used as both were effective in the detection of antiandrogens and, therefore, only one dose should be selected for use in the assay. Graduated test chemical doses are administered to a minimum of three treatment groups of experimental animals using one dose level per group. The animals are necropsied approximately 24 hours after the last dose. A statistically significant decrease in two or more target organ weights of the test chemical plus TP groups compared to the TP only control group indicates that the test chemical is positive for potential anti-androgenic activity (See paragraph 61).

**▼ M5**

## DESCRIPTION OF THE METHOD

**Selection of species and strain**

17. The rat has been routinely used in the Hershberger Bioassay since the 1930s. Although it is biologically plausible that both the rat and mouse would display similar responses, based upon 70 years of experience with the rat model, the rat is the species of choice for the Hershberger Bioassay. In addition, since Hershberger Bioassay data may be preliminary to a long-term multigenerational study, this allows animals from the same species, strain and source to be used in both studies.
  
18. This protocol allows laboratories to select the strain of rat to be used in the assay which should generally be that used historically by the participating laboratory. Commonly used laboratory rat strains may be used; however, strains that mature significantly later than 42 days of age should not be used since castration of these males at 42 days of age could preclude measurement of glans penis weights, which can only be done after the prepuce is separated from the penile shaft. Thus, strains derived from the Fisher 344 rat should not be used, except in rare cases. The Fisher 344 rat has a different timing of sexual development compared with other more commonly used strains such as Sprague Dawley or Wistar strains (16). If such a strain is to be used, the laboratory should castrate them at a slightly older age and be able to demonstrate the sensitivity of the strain used. The rationale for the choice of rat strain should be clearly stated by the laboratory. Where the screening assay may be preliminary to a repeated dose oral study, a reproductive and developmental study, or a long-term study, preferably animals from the same strain and source should be used in all studies.

**Housing and feeding conditions**

19. All procedures should conform to all local standards of laboratory animal care. These descriptions of care and treatment are minimum standards and will be superseded by more stringent local regulations, such as Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes (26). The temperature in the experimental animal room should be 22 °C (with an approximate range  $\pm 3$  °C). The relative humidity should be a minimum of 30 % and preferably should not exceed a maximum 70 %, other than during room cleaning. The aim should be relative humidity of 50-60 %. Lighting should be artificial. The daily lighting sequence should be 12 hours light, 12 hours dark.
  
20. Group housing is preferable to isolation because of the young age of the animals and the fact that rats are social animals. Housing of two or three animals per cage avoids crowding and associated stress that may interfere with the hormonal control of the development of the sex accessory tissue. Cages should be thoroughly cleaned to remove possible contaminants and arranged in such a way that possible effects due to cage placement are minimised. Cages of a proper size (~ 2 000 square centimetres) will prevent overcrowding.
  
21. Each animal should be identified individually (e.g. ear mark or tag) using a humane method. The method of identification should be recorded.

▼ **M5**

22. Laboratory diet and drinking water should be provided *ad libitum*. Laboratories executing the Hershberger Bioassay should use the laboratory diet normally used in their chemical testing work. In the validation studies of the Bioassay, no effects or variability were observed that were attributable to the diet. The diet used will be recorded and a sample of the laboratory diet should be retained for possible future analysis.

**Performance Criteria for androgen-dependent organ weights**

23. During the validation study, there was no evidence that a decrease in body weight affected increases or decreases in the growth of tissue weights for target tissues (i.e. that should be weighted in this study).
24. Among the different strains of rat used successfully in the validation programme, androgen-dependent organ weights are larger in the heavier rat strains than in the lighter strains. Therefore, the Hershberger Bioassay performance criteria do not include absolute expected organ weights for positive and negative controls.
25. Because the Coefficient of Variation (CV) for a tissue has an inverse relationship with statistical power, the Hershberger Bioassay performance criteria are based on maximum CV values for each tissue (Table 1). The CVs are derived from the OECD validation studies. In the case of negative outcomes, laboratories should examine the CVs from the control group and the high dose treatment group to determine if the maximum CV performance criteria have been exceeded.
26. The study should be repeated when: 1) three or more of the 10 possible individual CVs in the control and high dose treatment groups exceed the maximums designated for agonist and antagonist studies in Tables 1 and 2) at least two target tissues were marginally insignificant, i.e. *r* values between 0,05 and 0,10. <sup>(1)</sup>

*Table 1*

**Maximum allowable CVs Determined for the Target Sex Accessory Tissues for the castrate model in the OECD Validation Studies <sup>(1)</sup>.**

Tissue	Antiandrogenic effects	Androgenic effects
Seminal vesicles	40 %	40 %
Ventral prostate	40 %	45 %
LABC	20 %	30 %
Cowper's glands	35 %	55 %
Glans penis	17 %	22 %

<sup>(1)</sup> The threshold CV for a given tissue was identified from a graph of CV values — arranged from smallest sequentially to largest — for all means from all experiments in the validation exercise using a specific model (agonist or antagonist). The threshold CV was read from the point at which the increments between to the next highest CVs in the series are dramatically larger than the preceding few CVs- the 'breakpoint'. It should be noted that although this analysis identified relatively reliable 'breakpoints' for the antagonist model of the assay, CV curves for the agonist assay showed a more uniform increase making identification of a threshold CV by this method somewhat arbitrary.

**▼ M5****PROCEDURE****Regulatory compliance and laboratory verification**

27. Unlike the Uterotrophic assay (Chapter B.54 of this Annex), a demonstration of laboratory competence prior to the initiation of the study is not necessary for the Hershberger assay because concurrent positive (Testosterone Propionate and Flutamide) and negative controls are run as an integral part of the assay.

**Number and condition of animals**

28. Each treated and control group should include a minimum of 6 animals. This applies to both the androgenic and antiandrogenic protocols.

**Castration**

29. There should be an initial acclimatisation period of several days after receipt of the animals to ensure that the animals are healthy and thriving. Since animals castrated before 42 days of age or postnatal day (pnd) 42 may not display preputial separation, animals should be castrated on pnd 42 or thereafter, not before. The animals are castrated under anaesthesia by placing an incision in the scrotum and removing both testes and epididymides with ligation of blood vessels and seminal ducts. After confirming that no bleeding is occurring, the scrotum should be closed with suture or autoclips. Animals should be treated with analgesics for the first few days after surgery to alleviate any post-surgical discomfort. If castrated animals are purchased from an animal supplier, the age of animals and stage of sexual maturity should be assured by the supplier.

**Acclimatisation after castration**

30. The animals should continue acclimation to the laboratory conditions to allow for the regression in the target tissue weights for a minimum of 7 days following castration. Animals should be observed daily, and any animals with evidence of disease or physical abnormalities should be removed. Thus, treatment with initiation of dosing (on study) may commence as early as pnd 49 days of age, but not later than pnd 60. Age at necropsy should not be greater than pnd 70. This flexibility allows a laboratory to schedule the experimental work efficiently.

**Body weight and group randomisation**

31. Differences in individual body weights are a source of variability in tissue weights both within and among groups of animals. Increasing tissue weight variability results in an increased coefficient of variation (CV) and decreases the statistical power of the assay (sometimes referred to as assay sensitivity). Therefore, variations in body weight should be both experimentally and statistically controlled.
32. Experimental control involves producing small variations in body weight within and among the study groups. First, unusually small or large animals should be avoided and not placed in the study cohort. At study commencement the weight variation of animals used should not exceed  $\pm 20\%$  of the mean weight (e.g.  $175\text{ g} \pm 35\text{ g}$  for castrated peripubertal rats). Second, animals should be assigned to groups (both control and treatment) by randomised weight distribution, so that mean body weight of each group is not statistically different from any other group. The block randomisation procedure used should be recorded.



**▼ M5**

33. Because toxicity may decrease the body weight of treated groups relative to the control group, the body weight on the first day of test chemical administration could be used as the statistical covariate, not the body weight at necropsy.

**Dosage**

34. In order to establish whether a test chemical can have androgenic action *in vivo*, two dose groups of the test chemical plus positive and vehicle (negative) controls (See paragraph 43) are normally sufficient, and this design is therefore preferred for animal welfare reasons. If the purpose is either to obtain a dose-response curve or to extrapolate to lower doses, at least 3 dose groups are needed. If information beyond identification of androgenic activity (such as an estimate of potency) is required, a different dosing regimen should be considered. To test for antiandrogens, the test chemical is administered together with a reference androgen agonist. A minimum of 3 test groups with different doses of the test chemical and a positive and a negative control (See paragraph 44) should be used. Except for treatment with the test chemical, animals in the control group should be handled in an identical manner to the test group subjects. If a vehicle is used in administering the test chemical, the control group should receive the vehicle in the highest volume used with the test groups.

35. All dose levels should be proposed and selected taking into account any existing toxicity and (toxico-) kinetic data available for the test chemical or related materials. The highest dose level should first take into consideration the LD<sub>50</sub> and/or acute toxicity information in order to avoid death, severe suffering or distress in the animals (17)(18)(19)(20) and, second, take into consideration available information on the doses used in subchronic and chronic studies. In general, the highest dose should not cause a reduction in the final body weight of the animals greater than 10 % of control weight. The highest dose should be either 1) the highest dose that ensures animal survival and that is without significant toxicity or distress to the animals after 10 consecutive days of administration up to a maximal dose of 1 000 mg/kg/day (See paragraph 36) or 2) a dose inducing (anti)androgenic effects, whichever is lower. As a screen, large intervals, e.g. one half log units (corresponding to a dose progression of 3,2) or even one log units, between dosages are acceptable. If there are no suitable data available, a range finding study (See paragraph 37) may be performed to aid the determination of the doses to be used.

**Limit dose level**

36. If a test at the limit dose of 1 000 mg/kg body weight/day and a lower dose using the procedures described for this study fails to produce a statistically significant change in reproductive organ weights, then additional dose levels may be considered unnecessary. The limit dose applies except when human exposure data indicate the need for a higher dose level to be used.

**Considerations for range finding**

37. If necessary, a preliminary range finding study can be carried out with a few animals to select the appropriate dose groups [using methods for acute toxicity testing (Chapters B.1 bis, B.1 tris of this Annex (27), OECD TG 425 (19))]. The objective in the case of the Hershberger Bioassay is to select doses that ensure animal survival and that are without significant toxicity or distress to the animals after 10 consecutive days of chemical

**▼M5**

administration up to a limit dose of 1 000 mg/kg/d as noted in paragraphs 35 and 36. In this respect an OECD Guidance Document (17) may be used defining clinical signs indicative of toxicity or distress to the animals. If feasible within this range finding study after 10 days of administration, the target tissues may be excised and weighed approximately 24-hours after the last dose is administered. These data could then be used to assist the selection of the doses in the main study.

**Reference chemicals and vehicle**

38. The reference androgen agonist should be Testosterone Propionate (TP), CAS No 57-82-5. The reference TP dosage may be either 0,2 mg/kg-bw/d or 0,4 mg/kg-bw/d. The reference androgen antagonist should be Flutamide (FT), CAS No 1311-84-7. The reference FT dosage should be 3 mg/kg-bw/d, and the FT should be co-administered with the reference TP dosage.
  
39. It is recommended that, wherever possible, the use of an aqueous solution/suspension be considered first. However, since many androgen ligands or their metabolic precursors tend to be hydrophobic, the most common approach is to use a solution/suspension in oil (e.g. corn, peanut, sesame or olive oil). Test chemicals can be dissolved in a minimal amount of 95 % ethanol or other appropriate solvents and diluted to final working concentrations in the test vehicle. The toxic characteristics of the solvent should be known, and should be tested in a separate solvent-only control group. If the test chemical is considered stable, gentle heating and vigorous mechanical action can be used to assist in dissolving the test chemical. The stability of the test chemical in the vehicle should be determined. If the test chemical is stable for the duration of the study, then one starting aliquot of the test chemical may be prepared, and the specified dosage dilutions prepared daily using care to avoid contamination and spoilage of the samples.

**Administration of doses**

40. TP should be administered by subcutaneous injection, and FT by oral gavage.
  
41. The test chemical is administered by oral gavage or subcutaneous injection. Animal welfare considerations and the physical/chemical properties of the test chemical need to be taken into account when choosing the route of administration. In addition, toxicological aspects like the relevance to the human route of exposure to the chemical (e.g. oral gavage to model ingestion, subcutaneous injection to model inhalation or dermal adsorption) and existing toxicological information and data on metabolism and kinetics (e.g. need to avoid first pass metabolism, better efficiency via a particular route) should be taken into account before extensive, long-term testing is initiated if positive results are obtained by injection.
  
42. The animals should be dosed in the same manner and time sequence for 10 consecutive days at approximately 24 hour intervals. The dosage level should be adjusted daily based on the concurrent daily measures of body weight. The volume of dose and time that it is administered should be recorded on each day of exposure. Care should be taken in order not to exceed the maximum dose described in paragraph 35 to allow a meaningful interpretation of the data. Reduction of body weight, clinical signs, and other findings should be thoroughly assessed in this respect. For oral gavage, a stomach tube or a suitable intubation cannula should be used. The maximum volume of liquid that can be administered at one time

**▼ M5**

depends on the size of the test animal. Local animal care guidelines should be followed, but the volume should not exceed 5 ml/kg body weight, except in the case of aqueous solutions where 10 ml/kg body weight may be used. For subcutaneous injections, doses should be administered to the dorsoscapular and or lumbar regions via sterile needle (e.g. 23- or 25-gauge) and a tuberculin syringe. Shaving the injection site is optional. Any losses, leakage at the injection site or incomplete dosing should be recorded. The total volume injected per rat per day should not exceed 0,5 ml/kg body weight.

**Specific procedures for androgen agonists**

43. For the test for androgen agonists, the vehicle is the negative control, and the TP-treated group is the positive control. Biological activity consistent with androgen agonists is tested by administering a test chemical to treatment groups at the selected doses for 10 consecutive days. The weights of the five sex accessory tissues from the test chemical groups are compared to the vehicle group for statistically significant increases in weight.

**Specific procedures for androgen antagonists and 5 $\alpha$ -reductase inhibitors**

44. For the test for androgen antagonists and 5 $\alpha$ -reductase inhibitors, the TP-treated group is the negative control, and the group coadministered with reference doses of TP and FT is the positive control. Biological activity consistent with androgen antagonists and 5 $\alpha$ -reductase inhibitors is tested by administering a reference dose of TP and administering the test chemical for 10 consecutive days. The weights of the five sex accessory tissues from the TP plus test chemical groups are compared to the reference TP-only group for statistically significant decreases in weights.

**OBSERVATIONS****Clinical observations**

45. General clinical observations should be made at least once a day and more frequently when signs of toxicity are observed. Observations should be carried out preferably at the same time(s) each day and considering the period of anticipated peak effects after dosing. All animals should be observed for mortality, morbidity and general clinical signs such as changes in behaviour, skin, fur, eyes, mucous membranes, occurrence of secretions and excretions and autonomic activity (e.g. lacrimation, piloerection, pupil size, unusual respiratory pattern).
46. Any animal found dead should be removed and disposed of without further data analysis. Any mortality of animals prior to necropsy should be included in the study record together with any apparent reasons for mortality. Any moribund animals should be humanely terminated. Any moribund and subsequently euthanised animals should be included in the study record with apparent reasons for morbidity.

**▼ M5****Body weight and food consumption**

47. All animals should be weighed daily to the nearest 0,1 g, starting just prior to initiation of treatment, i.e. when the animals are allocated into groups. As an optional measurement, the amount of food consumed during the treatment period may be measured per cage by weighing the feeders. The food consumption results should be expressed in grams per rat per day.

**Dissection and measurement of tissue and organ weights**

48. Approximately 24 hours after the last administration of the test chemical, the rats should be euthanised and exsanguinated according to the normal procedures of the conducting laboratory, and necropsy carried out. The method of humane killing should be recorded in the laboratory report.
  
49. Ideally, the necropsy order should be randomised across groups to avoid progression directly up or down dose groups that could affect the data. Any finding at necropsy, i.e. pathological changes/visible lesions should be noted and reported.
  
50. The five androgen-dependent tissues (VP, SV, LABC, COW, GP) should be weighted. These tissues should be excised, carefully trimmed of excess adhering tissue and fat, and their fresh (unfixed) weights determined. Each tissue should be handled with particular care to avoid the loss of fluids and to avoid desiccation, which may introduce significant errors and variability by decreasing the recorded weights. Several of the tissues may be very small or difficult to dissect, and this will introduce variability. Therefore, it is important that persons carrying out the dissection of the sex accessory tissues are familiar with standard dissection procedures for these tissues. A standard operating procedure (SOP) manual for dissection is available from the OECD (21). Careful training according to the SOP guide will minimise a potential source of variation in the study. Ideally the same prosector should be responsible for the dissection of a given tissue to eliminate inter-individual differences in tissue processing. If this is not possible, the necropsy should be designed such that each prosector dissects a given tissue from all treatment groups as opposed to one individual dissecting all tissues from a control group, while someone else is responsible for the treated groups. Each sex accessory tissues should be weighed without blotting to the nearest 0,1 mg, and the weights recorded for each animal.
  
51. Several of the tissues may be very small or difficult to dissect, and this will introduce variability. Previous work has indicated a range of coefficient of variations (CVs) that appears to differ based upon the proficiency of the laboratory. In a few cases, large differences in the absolute weights of the tissues such as the VP and COWS have been observed within a particular laboratory.
  
52. Liver, paired kidney, and paired adrenal weights are optional measurements. Again, tissues should be trimmed free of any adhering fascia and fat. The liver should be weighed and recorded to the nearest 0,1 g and the paired kidneys and paired adrenals should be weighed and recorded to the nearest 0,1 mg. The liver, kidney and adrenals are not only influenced by androgens; they also provide useful indices of systemic toxicity.

▼ **M5**

53. Measurement of serum luteinising hormone (LH), follicular stimulating hormone (FSH) and testosterone (T) is optional. Serum T levels are useful to determine if the test chemical induces liver metabolism of testosterone, lowering serum levels. Without the T data, such an effect might appear to be via an antiandrogenic mechanism. LH levels provides information about the ability of an antiandrogen to not only reduce organ weights, but also to affect hypothalamic-pituitary function, which in long term studies can induce testis tumors. FSH is an important hormone for spermatogenesis. Serum T4 and T3 also are optional measures that would provide useful supplemental information about the ability to disrupt thyroid hormone homeostasis. If hormone measurements are to be made, the rats should be anesthetised prior to necropsy and blood taken by cardiac puncture, and the method of anaesthesia should be chosen with care so that it does not affect hormone measurement. The method of serum preparation, the source of radioimmunoassay or other measurement kits, the analytical procedures, and the results should be recorded. LH levels should be reported as ng per ml of serum, and T should also be reported as ng per ml of serum.
54. The dissection of the tissues is described as follows with a detailed dissection guide with photographs published as supplementary materials as part of the validation programme (21). A dissection video is also available from the Korea Food and Drug Administration web page (22).
- With the ventral surface of the animal upwards, determine if the prepuce of the penis has separated from the glans penis. If so, then retract the prepuce and remove the glans penis, weigh (nearest 0,1 mg), and record the weight;
  - Open the abdominal skin and wall, exposing the viscera. If the optional organs are weighed, remove and weigh liver to nearest 0,1 g, remove the stomach and intestines, remove and weigh the paired kidneys and paired adrenals to the nearest 0,1 mg. This dissection exposes the bladder and begins the dissection of the target male accessory tissues.
  - To dissect the VP, separate bladder from the ventral muscle layer by cutting connective tissue along the midline. Displace the bladder anteriorly towards the seminal vesicles (SV), revealing the left and right lobes of the ventral prostate (covered by a layer of fat). Carefully tease the fat from the right and left lobes of the VP. Gently displace the VP right lobe from the urethra and dissect the lobe from the urethra. While still holding the VP right lobe, gently displace the VP left lobe from the urethra and then dissect; weigh to nearest 0,1 mg and record the weight.
  - To dissect the SVCG, displace the bladder caudally, exposing the vas deferens and right and left lobes of the seminal vesicles plus coagulating glands (SVCG). Prevent leakage of fluid by clamping a haemostat at the base of the SVCGs, where the vas deferens joins the urethra. Carefully dissect the SVCGs, with the haemostat in place trim fat and adnexa away, place in a tared weigh-boat, remove the haemostat, and weigh to the nearest 0,1 mg and record the weight.

**▼ M5**

- To dissect the levator ani plus bulbocavernosus muscles (LABC), the muscles and the base of the penis are exposed. The LA muscles wrap around the colon, while the anterior LA and BC muscles are attached to the penile bulbs. The skin and adnexa from the perianal region extending from the base of the penis to the anterior end of the anus are removed. The BC muscles are gradually dissected from the penile bulb and tissues. The colon is cut in two and, the full LABC can be dissected and removed. The LABC should be trimmed of fat and adnexa, weighed to the nearest 0,1 mg, and record the weight.
  - After the LABC has been removed, the round Cowper's or bulbourethral glands (COW) are visible at the base of, and slightly dorsal to, the penile bulbs. Careful dissection is required to avoid nicking the thin capsule in order to prevent fluid leakage. Weigh the paired COW to the nearest 0,1 mg, and record the weight.
  - In addition, if fluid is lost from any gland during the necropsy and dissection, this should be recorded.
55. If the evaluation of each chemical requires necropsy of more animals than is reasonable for a single day, the study start may be staggered on two consecutive days, resulting in the staggering of the necropsy and the related work over two days. If staggered in this manner, one-half of the animals per treatment group should be used per day.
56. Carcasses should be disposed of in an appropriate manner following necropsy.

**REPORTING****Data**

57. Data should be reported individually (i.e. body weight, accessory sex tissue weights, optional measurements and other responses and observations) and for each group of animals (means and standard deviations of all measurement taken). The data should be summarised in tabular form. The data should show the number of animals at the start of the test, the number of animals found dead during the test or found showing signs of toxicity, a description of the signs of toxicity observed, including time of onset, duration and severity.
58. A final report should include:

*Testing facility*

- Name of facility, location
- Study director and other personnel and their study responsibilities
- Dates the study began and ended, i.e. first day of test chemical administration and last day of necropsy, respectively.

*Test chemical*

- Source, lot/batch number, identity, purity, full address of the supplier and characterisation of the test chemical(s)
- Physical nature and, where relevant, physicochemical properties;
- Storage conditions and the method and frequency of dilution preparation
- Any data generated on stability
- Any analyses of dosing solutions/suspensions.

**▼ M5***Vehicle*

- Characterisation of the vehicle (identity, supplier and lot #)
- Justification of the vehicle choice (if other than water)

*Test animals and animal husbandry procedures*

- Species/strain used and rationale for choice
- Source or supplier of animals, including full address
- Number and age of animals supplied
- Housing conditions (temperature, lighting, and so on)
- Diet (name, type, supplier, lot number, content and if known, phytoestrogens levels)
- Bedding (name, type, supplier, content)
- Caging conditions and number of animals per cage;

*Assay Conditions*

- Age at castration and duration of acclimatisation after castration;
- Individual weights of animals at the start of the study (to nearest 0,1 g);
- Randomisation process and a record of the assignment to vehicle, reference, test chemical groups, and cages
- Mean and standard deviation of the body weights for each group for each weigh day throughout the study;
- Rationale for dose selection
- Route of administration of test chemical and rationale for the choice of exposure route
- If an assay for antiandrogenicity, the TP treatment (dose and volume),
- Test chemical treatment (dose and volume),
- Time of dosing
- Necropsy procedures, including means of exsanguinations and any anaesthesia
- If serum analyses are performed, details of the method should be supplied. For example, if RIA is used, the RIA procedure, source of RIA kits, kit expiration dates, procedure for scintillation counting, and standardisation should be reported.

*Results*

- Daily observations for each animal during dosing, including:
  - Body weights (to the nearest 0,1 g),
  - Clinical signs (if any),
  - Any measurement or notes of food consumption.
- Necropsy observations for each animal, including:

**▼ M5**

- Date of necropsy,
- Animal treatment group,
- Animal ID,
- Prosector,
- Time of day necropsy and dissection are performed,
- Animal age,
- Final body weight at necropsy, noting any statistically significant increase or decrease,
- Order of animal exsanguination and dissection at necropsy,
- Weights of the five target androgen dependent tissues:
  - Ventral prostate (to the nearest 0,1 mg)
  - Seminal vesicles plus coagulating glands, including fluid (paired, to nearest 0,1 mg)
  - Levator ani plus bulbocavernosus muscle complex (to nearest 0,1 mg)
  - Cowper's glands (fresh weight — paired, to nearest 0,1 mg).
  - Glans penis (fresh weight to nearest 0,1 mg)
- Weights of optional tissues, if performed:
  - Liver (to nearest 0,1 g)
  - Kidney (paired, to nearest 0,1 mg)
  - Adrenal (paired, to nearest 0,1 mg)
- General remarks and comments
- Analyses of serum hormones, if performed.
  - Serum LH (optional — ng per ml of serum), and
  - Serum T (optional — ng per ml of serum)
- General remarks and comments

*Data summarisation*

Data should be summarised in tabular form containing the sample size for each group, the mean of the value, and the standard error of the mean or the standard deviation. Tables should include necropsy body weights, body weight changes from the beginning of dosing until necropsy, target accessory sex tissues weights, and any optional organ weights.

*Discussion of the results***Analysis of results**

59. Necropsy body and organ weights should be statistically analysed for characteristics such as homogeneity of variance with appropriate data transformations as needed. Treatment groups should be compared to a control group using techniques such as ANOVA followed by pairwise comparisons (e.g. Dunnett's one tailed test) and the criterion for statistical difference, for example,  $p \leq 0,05$ . Those groups attaining statistical significance should be identified. However, 'relative organ' weights should be avoided due to the invalid statistical assumptions underlying this data manipulation.



**▼ M5**

60. For androgen agonism, the control should be the vehicle-only test group. The mode of action characteristics of a test chemical can lead to different relative responses amongst the tissues, for example trenbolone, which cannot be 5 alpha-reduced, has more pronounced effects on the LABC and GP than does TP. A statistically significant increase ( $p \leq 0,05$ ) in any two or more of the five target androgen-dependent tissue weights (VP, LABC, GP, CG and SVCG) should be considered a positive androgen agonist result, and all the target tissues should display some degree of increased growth. Combined evaluation of all accessory sex organs (ASO) tissue responses could be achieved using appropriate multivariate data analysis. This could improve the analysis, especially in cases where only a single tissue gives a statistically significant response.
61. For androgen antagonism, the control should be the reference androgen (testosterone propionate only) test group. The mode of action characteristics of a test chemical can lead to different relative responses amongst the tissues, for example 5 alpha  $\alpha$ -reductase inhibitors, like finasteride, have more pronounced effects on the ventral prostate than other tissues as compared to potent AR antagonists, like flutamide. A statistically significant reduction ( $p \leq 0,05$ ) in any two or more of the five target androgen-dependent tissue weights (VP, LABC, GP, CG and SVCG) relative to TP treatment alone should be considered a positive androgen antagonist result and all the target tissues should display some degree of reduced growth. Combined evaluation of all ASO tissue responses could be achieved using appropriate multivariate data analysis. This could improve the analysis, especially in cases where only a single tissue gives a statistically significant response.
62. Data should be summarised in tabular form containing the mean, standard error of the mean (standard deviation would also be acceptable) and sample size for each group. Individual data tables should also be included. The individual values, mean, SE (SD) and CV values for the control data should be examined to determine if they meet acceptable criteria for consistency with expected historical values. CVs that exceed CV values listed in Table 1 (see paragraphs 25 and 26) for each organ weight should determine if there are errors in data recording or entry or if the laboratory has not yet mastered accurate dissection of the androgen-dependent tissues and further training/practice is warranted. Generally, CVs (the standard deviation divided by the mean organ weight) are reproducible from lab to lab and study to study. Data presented should include at least: ventral prostate, seminal vesicle, levator ani plus bulbocavernosus, Cowper's glands, glans penis, liver, and body weights and body weight change from the beginning of dosing until necropsy. Data also may be presented after covariance adjustment for body weight, but this should not replace presentation of the unadjusted data. In addition, if preputial separation (PPS) does not occur in any of the groups, the incidence of PPS should be recorded and statistically compared to the control group using Fisher Exact test.

**▼ M5**

63. When verifying the computer data entries with the original data sheets for accuracy, organ weight values that are not biologically plausible or vary by more than three standard deviations from that treatment group means should be carefully scrutinised and may need to be discarded, likely being recording errors.
64. Comparison of study results with OECD CV values (in Table 1) is often an important step in interpretation as to the validity of the study results. Historical data for vehicle control groups should be maintained in the laboratory. Historical data for responses to positive reference chemicals, such as TP and FT, should also be maintained in the laboratory. Laboratories may also periodically test the response to known weak androgen agonists and antagonists and maintain these data. These data can be compared to available OECD data to ensure that the laboratory's methods yield sufficient statistical precision and power.

**▼ M5***Appendix 1*

## DEFINITIONS:

**Androgenic** is a term used to describe a positive influence on the growth of androgen-dependent tissues

**Antiandrogenic** is the capability of a chemical to suppress the action of TP in a mammalian organism.

**Chemical** means a substance or a mixture.

**Date of birth** is postnatal day 0.

**Dose** is the amount of test chemical administered. For the Hershberger Bioassay, the dose is expressed as weight of test chemical per unit body weight of test animal per day (e.g. mg/kg body weight/day).

**Dosage** is a general term comprising of dose, its frequency and the duration of dosing.

**Moribund** is a term used to describe an animal in a dying state, i.e. near the point of death.

**Postnatal day X** is the Xth day of life after the day of birth.

**Sensitivity** is the capability of a test method to correctly identify chemicals having the property that is being tested for.

**Specificity** is the capability of a test method to correctly identify chemicals not having the property that is being tested for.

**Test chemical** means any substance or mixture tested using this test method.

**Validation** is a scientific process designed to characterise the operational requirements and limitations of a test method and to demonstrate its reliability and relevance for a particular purpose.

Appendix 2

**Note:** Document prepared by the Secretariat of the Test Guidelines Programme based on the agreement reached at the 6th Meeting of the EDTA Task Force

**OECD Conceptual Framework for the Testing and Assessment of Endocrine Disrupting Chemicals**

<p><b>Level 1</b> Sorting &amp; prioritization based upon existing information</p>	<ul style="list-style-type: none"> <li>— physical &amp; chemical properties, e.g., MW, reactivity, volatility, biodegradability,</li> <li>— human &amp; environmental exposure, e.g., production volume, release, use patterns</li> <li>— hazard, e.g., available toxicological data</li> </ul>	
<p><b>Level 2</b> <i>In vitro</i> assays providing mechanistic data</p>	<ul style="list-style-type: none"> <li>— ER, AR, TR receptor binding affinity</li> <li>— Transcriptional activation</li> <li>— Aromatase and steroidogenesis <i>in vitro</i></li> <li>— Aryl hydrocarbon receptor recognition/binding</li> <li>— QSARs</li> </ul>	<ul style="list-style-type: none"> <li>— High Through Put Prescreens</li> <li>— Thyroid function</li> <li>— Fish hepatocyte VTG assay</li> <li>— Others (as appropriate)</li> </ul>
<p><b>Level 3</b> <i>In vivo</i> assays providing data about single endocrine Mechanisms and effects</p>	<ul style="list-style-type: none"> <li>— Uterotrophic assay (estrogenic related)</li> <li>— Hershberger assay (androgenic related)</li> <li>— Non-receptor mediated hormone function</li> <li>— Others (e.g. thyroid)</li> </ul>	<ul style="list-style-type: none"> <li>— Fish VTG (vitellogenin) assay (estrogenic related)</li> </ul>
<p><b>Level 4</b> <i>In vivo</i> assays providing data about multiple endocrine Mechanisms and effects</p>	<ul style="list-style-type: none"> <li>— enhanced OECD 407 (endpoints based on endocrine mechanisms)</li> <li>— male and female pubertal assays</li> <li>— adult intact male assay</li> </ul>	<ul style="list-style-type: none"> <li>— Fish gonadal histopathology assay</li> <li>— Frog metamorphosis assay</li> </ul>
<p><b>Level 5</b> <i>In vivo</i> assays providing data on effects from endocrine &amp; other mechanisms</p>	<ul style="list-style-type: none"> <li>— 1-generation assay (TG415 enhanced)<sup>1</sup></li> <li>— 2-generation assay (TG416 enhanced)<sup>1</sup></li> <li>— reproductive screening test (TG421 enhanced)<sup>1</sup></li> <li>— combined 28 day/reproduction screening test (TG 422 enhanced)</li> </ul> <p><sup>1</sup> Potential enhancements will be considered by VMG mamm</p>	<ul style="list-style-type: none"> <li>— Partial and full life cycle assays in fish, birds, amphibians &amp; invertebrates (developmental and reproduction)</li> </ul>

VMG mamm: Validation Management Group on Mammalian Testing and Assessment

**▼ M5**

## NOTES TO THE FRAMEWORK:

- Note 1:* Entering at all levels and exiting at all levels is possible and depends upon the nature of existing information needs for hazard and risk assessment purposes
- Note 2:* In level 5, ecotoxicology should include endpoints that indicate mechanisms of adverse effects, and potential population damage
- Note 3:* When a multimodal model covers several of the single endpoint assays, that model would replace the use of those single endpoint assays
- Note 4:* The assessment of each chemical should be based on a case by case basis, taking into account all available information, bearing in mind the function of the framework levels.
- Note 5:* The framework should not be considered as all inclusive at the present time. At levels 3, 4 and 5 it includes assays that are either available or for which validation is under way. With respect to the latter, these are provisionally included. Once developed and validated, they will be formally added to the framework.
- Note 6:* Level 5 should not be considered as including definitive tests only. Tests included at that level are considered to contribute to general hazard and risk assessment.

## LITERATURE

- (1) OECD (1998). Report of the First Meeting of the OECD Endocrine Disrupter Testing and Assessment (EDTA) Task Force, 10th-11th March 1998, ENV/MC/CHEM/RA(98)5.
- (2) Dorfman RI (1962). Standard methods adopted by official organization. Academic Press, NY.
- (3) Gray LE Jr, Furr J and Ostby JS (2005). Hershberger assay to investigate the effects of endocrine disrupting compounds with androgenic and anti-androgenic activity in castrate-immature male rats. In: *Current Protocols in Toxicology* 16.9.1-16.9.15. J Wiley and Sons Inc.
- (4) OECD (2006). Final OECD report of the initial work towards the validation of the rat Hershberger assay. Phase 1. Androgenic response to testosterone propionate and anti-androgenic effects of flutamide. Environmental Health and Safety, Monograph Series on Testing and Assessment N° 62. ENV/JM/MONO(2006)30.
- (5) OECD (2008). Report of the OECD Validation of the Rat Hershberger Bioassay: Phase 2: Testing of Androgen Agonists, Androgen Antagonists and a 5 $\alpha$ -Reductase Inhibitor in Dose Response Studies by Multiple Laboratories. Environmental Health and Safety, Monograph Series on Testing and Assessment N° 86. ENV/JM/MONO(2008)3.
- (6) OECD (2007). Report of the Validation of the Rat Hershberger Assay: Phase 3: Coded Testing of Androgen Agonists, Androgen Antagonists and Negative Reference Chemicals by Multiple Laboratories. Surgical Castrate Model Protocol. Environmental Health and Safety, Monograph Series on Testing and Assessment N° 73. ENV/JM/MONO(2007)20.
- (7) Owens, W, Zeiger E, Walker M, Ashby J, Onyon L, Gray, Jr, LE (2006). The OECD programme to validate the rat Hershberger bioassay to screen compounds for in vivo androgen and antiandrogen responses. Phase 1: Use of a potent agonist and a potent antagonist to test the standardized protocol. *Env. Health Persp.* 114:1265-1269.

▼ **M5**

- (8) Owens W, Gray LE, Zeiger E, Walker M, Yamasaki K, Ashby J, Jacob E (2007). The OECD program to validate the rat Hershberger bioassay to screen compounds for in vivo androgen and antiandrogen responses: phase 2 dose-response studies. *Environ Health Perspect.* 115(5):671-8.
- (9) Korenchevsky V (1932). The assay of testicular hormone preparations. *Biochem J*26:413-422.
- (10) Korenchevsky V, Dennison M, Schalit R (1932). The response of castrated male rats to the injection of the testicular hormone. *Biochem J*26:1306-1314.
- (11) Eisenberg E, Gordan GS (1950). The levator ani muscle of the rat as an index of myotrophic activity of steroidal hormones. *J Pharmacol Exp Therap* 99:38-44.
- (12) Eisenberg E, Gordan GS, Elliott HW (1949). Testosterone and tissue respiration of the castrate male rat with a possible test for myotrophic activity. *Endocrinology* 45:113-119.
- (13) Hershberger L, Shipley E, Meyer R (1953). Myotrophic activity of 19-nortestosterone and other steroids determined by modified levator ani muscle method. *Proc Soc Exp Biol Med* 83:175-180.
- (14) Hilgar AG, Vollmer EP (1964). *Endocrine bioassay data: Androgenic and myogenic.* Washington DC: United States Public Health Service.
- (15) Dorfman RI (1969). Androgens and anabolic agents. In: *Methods in Hormone Research, volume IIA.* (Dorfman RI, ed.) New York:Academic Press, 151-220.
- (16) Massaro EJ (2002). *Handbook of Neurotoxicology, volume I.* New York: Humana Press, p 38.
- (17) OECD (2000). Guidance document on the recognition, assessment and use of clinical signs as humane endpoints for experimental animals used in safety evaluation. *Environmental Health and Safety Monograph Series on Testing and Assessment No 19.* ENV/JM/MONO(2000)7.
- (18) OECD (1982). *Organization for Economic Co-operation and Development — Principles of Good Laboratory Practice,* ISBN 92-64-12367-9, Paris.
- (19) OECD (2008). *Acute oral toxicity — up-and-down procedure.* OECD Guideline for the testing of chemicals No 425.
- (20) OECD (2001). *Guidance document on acute oral toxicity.* *Environmental Health and Safety Monograph Series on Testing and Assessment No 24.* ENV/JM/MONO(2001)4.
- (21) Supplemental materials for Owens et al. (2006). The OECD programme to validate the rat Hershberger bioassay to screen compounds for in vivo androgen and antiandrogen responses. Phase 1: Use of a potent agonist and a potent antagonist to test the standardized protocol. *Env. Health Persp.* 114:1265-1269. See, section II, The dissection guidance provided to the laboratories: <http://www.ehponline.org/docs/2006/8751/suppl.pdf>.
- (22) Korea Food and Drug Administration. *Visual reference guide on Hershberger assay procedure, including a dissection video.* [http://ndmoa.kfda.go.kr/endocrine/reference/education\\_fr.html](http://ndmoa.kfda.go.kr/endocrine/reference/education_fr.html)
- (23) OECD (2008). *Background Review Document on the Rodent Hershberger Bioassay.* *Environmental Health and Safety Monograph Series on Testing and Assessment No 90.* ENV/JM/MONO(2008)17.
- (24) OECD (2008). *Draft Validation report of the Intact, Stimulated, Weanling Male Rat Version of the Hershberger Bioassay.*
- (25) OECD (2009). *Guidance Document on the Weanling Hershberger Bioassay in rats: A shortterm screening assay for (anti)androgenic properties.* *Series on Testing and Assessment, Number 115.*

▼ **M5**

- (26) Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes (OJ L 276, 20.10.2010, p. 33).
- (27) The following chapters of this Annex:
  - B.1 bis, Acute oral toxicity — fixed dose procedure
  - B.1 tris, Acute oral toxicity — acute toxic class method

**▼ M5****B.56 EXTENDED ONE-GENERATION REPRODUCTIVE TOXICITY STUDY**

## INTRODUCTION

1. This test method is equivalent to OECD Test Guideline (TG) 443 (2012). It is based on the International Life Science Institute (ILSI)-Health and Environmental Sciences Institute (HESI), Agricultural Chemical Safety Assessment (ACSA) Technical Committee proposal for a life stage F<sub>1</sub> extended one generation reproductive study as published in Cooper et al., 2006 (1). Several improvements and clarifications have been made to the study design to provide flexibility and to stress the importance of starting with existing knowledge, while using in-life observations to guide and tailor the testing. This test method provides a detailed description of the operational conduct of an Extended One-Generation Reproductive Toxicity Study. The test method describes three cohorts of F<sub>1</sub> animals:

Cohort 1: assesses reproductive/developmental endpoints; this cohort may be extended to include an F<sub>2</sub> generation.

Cohort 2: assesses the potential impact of chemical exposure on the developing nervous system.

Cohort 3: assesses the potential impact of chemical exposure on the developing immune system.

2. Decisions on whether to assess the second generation and to omit the developmental neurotoxicity cohort and/or developmental immunotoxicity cohort should reflect existing knowledge for the chemical being evaluated, as well as the needs of various regulatory authorities. The purpose of the test method is to provide details on how the study can be conducted and to address how each cohort should be evaluated.
3. Procedure for the decision on the internal triggering for producing a second generation is described in OECD Guidance Document 117(39) for those regulatory authorities using internal triggers.

## INITIAL CONSIDERATIONS AND OBJECTIVES

4. The main objective of the Extended One-Generation Reproductive Toxicity Study is to evaluate specific life stages not covered by other types of toxicity studies and test for effects that may occur as a result of pre- and postnatal chemical exposure. For reproductive endpoints, it is envisaged that, as a first step and when available, information from repeat-dose studies (including screening reproductive toxicity studies, e.g. OECD TG 422 (32)), or short term endocrine disrupter screening assays, (e.g. Uterotrophic assay — test method B.54 (36); and Hershberger assay — test method B.55 (37)) is used to detect effects on reproductive organs for males and females. This might include spermatogenesis (testicular histopathology) for males and oestrous cycles, follicle counts/oocyte maturation and ovarian integrity (histopathology) for females. The Extended One-Generation Reproductive Toxicity Study then serves as a test for reproductive endpoints that require the interaction of males with females, females with conceptus, and females with offspring and the F<sub>1</sub> generation until after sexual maturity (see OECD Guidance Document 151 supporting this test method (40)).



**▼ M5**

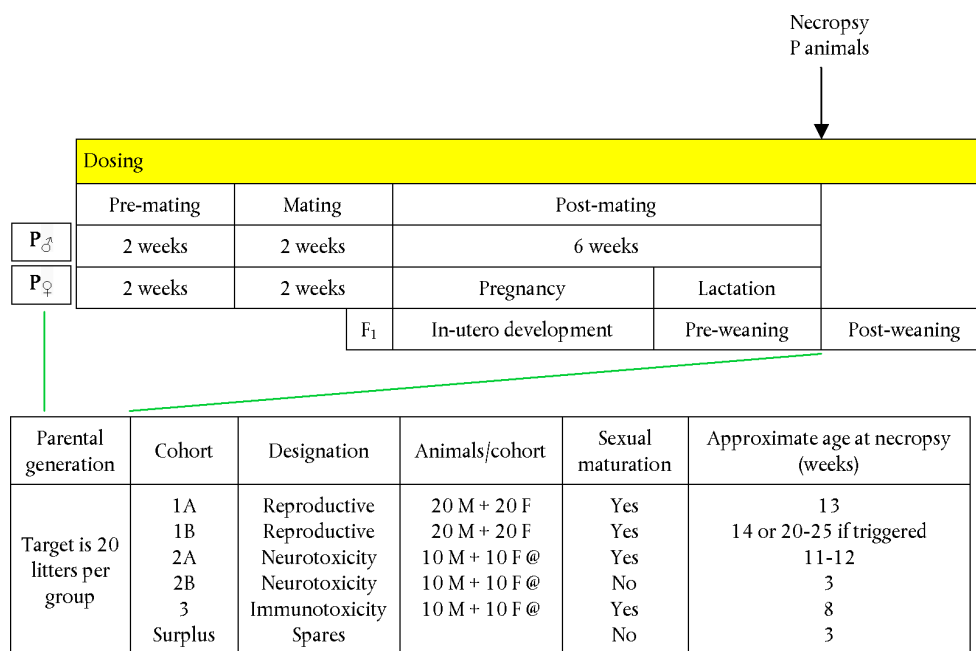
5. The test method is designed to provide an evaluation of the pre- and postnatal effects of chemicals on development as well as a thorough evaluation of systemic toxicity in pregnant and lactating females and young and adult offspring. Detailed examination of key developmental endpoints, such as offspring viability, neonatal health, developmental status at birth, and physical and functional development until adulthood, is expected to identify specific target organs in the offspring. In addition, the study will provide and/or confirm information about the effects of a test chemical on the integrity and performance of the adult male and female reproductive systems. Specifically, but not exclusively, the following parameters are considered: gonadal function, the oestrous cycle, epididymal sperm maturation, mating behaviour, conception, pregnancy, parturition, and lactation. Furthermore, the information obtained from the developmental neurotoxicity and developmental immunotoxicity assessments will characterise potential effects in those systems. The data derived from these tests should allow the determination of No-Observed Adverse Effect Levels (NOAELs), Lowest Observed Adverse Effect Levels (LOAELs) and/or benchmark doses for the various endpoints and/or be used to characterise effects detected in previous repeat-dose studies and/or serve as a guide for subsequent testing.
  
6. A schematic drawing of the protocol is presented in Figure 1. The test chemical is administered continuously in graduated doses to several groups of sexually mature males and females. This parental (P) generation is dosed for a defined pre-mating period (selected based on the available information for the test chemical; but for a minimum of two weeks) and a two-week mating period. P males are further treated at least until weaning of the F<sub>1</sub>. They should be treated for a minimum of 10 weeks. They may be treated for longer if there is a need to clarify effects on reproduction. Treatment of the P females is continued during pregnancy and lactation until termination after the weaning of their litters (i.e. 8-10 weeks of treatment). The F<sub>1</sub> offspring receive further treatment with the test chemical from weaning to adulthood. If a second generation is assessed (see OECD Guidance Document 117(39)), the F<sub>1</sub> offspring will be maintained on treatment until weaning of the F<sub>2</sub>, or until termination of the study.
  
7. Clinical observations and pathology examinations are performed on all animals for signs of toxicity, with special emphasis on the integrity and performance of the male and female reproductive systems and the health, growth, development and function of the offspring. At weaning, selected offspring are assigned to specific subgroups (cohorts 1-3, see paragraphs 33 and 34 and Figure 1) for further investigations, including sexual maturation, reproductive organ integrity and function, neurological and behavioural endpoints, and immune functions.
  
8. In conducting the study, the guiding principles and considerations outlined in the OECD Guidance Document No 19 on the recognition, assessment, and use of clinical signs as humane endpoints for experimental animals used in safety evaluations (34) should be followed.

## ▼ M5

9. When a sufficient number of studies are available to ascertain the impact of this new study design, the test method will be reviewed and if necessary revised in light of experience gained.

Figure 1

## Scheme of the Extended One-Generation Reproductive Toxicity Study



@ one per litter and representative of 20 litters in total where possible

## DESCRIPTION OF THE METHOD/PREPARATIONS FOR THE TEST

**Animals***Selection of animal species and strain*

10. The choice of species for the reproductive toxicity test should be carefully considered in light of all available information. However, because of the extent of background data and the comparability to general toxicity tests, the rat is normally the preferred species, and criteria and recommendations given in this test method refer to this species. If another species is used, justification should be given and appropriate modifications to the protocol will be necessary. Strains with low fecundity or a well-known high incidence of spontaneous developmental defects should not be used.

*Age, body weight and inclusion criteria*

11. Healthy parental animals, which have not been subjected to previous experimental procedures, should be used. Both males and females should be studied and the females should be nulliparous and non-pregnant. The P animals should be sexually mature, of similar weight (within sex) at initiation of dosing, similar age (approximately 90 days) at mating, and representative of the species and strain under study. Animals should be acclimated for at least 5 days after arrival. The animals are randomly assigned to the control and treatment groups, in a manner, which results in comparable mean body weight values among the groups (i.e.  $\pm 20\%$  of the mean).

**▼ M5***Housing and feeding conditions*

12. The temperature in the experimental animal room should be 22 °C ( $\pm 3$  °C). Relative humidity should be between 30-70 %, with an ideal range of 50-60 %. Artificial lighting should be set at 12 hours light, 12 hours dark. Conventional laboratory diets may be used with an unlimited supply of drinking water. Careful attention should be given to diet phytoestrogen content, as a high level of phytoestrogen in the diet might affect some reproductive endpoints. Standardised, open-formula diets in which estrogenic chemicals have been reduced are recommended (2)(30). The choice of diet may be influenced by the need to ensure a suitable admixture of a test chemical when administered by this method. Content, homogeneity and stability of the test chemical in the diets should be verified. The feed and drinking water should be regularly analysed for contaminants. Samples of each batch of the diet used during the study should be retained under appropriate conditions (e.g. frozen at  $-20$  °C), until finalisation of the report, in case the results necessitate a further analysis of diet ingredients.
  
13. Animals should be caged in small groups of the same sex and treatment group. They may be housed individually to avoid possible injuries (e.g. males after the mating period). Mating procedures should be carried out in suitable cages. After evidence of copulation, females that are presumed to be pregnant are housed separately in parturition or maternity cages where they are provided with appropriate and defined nesting materials. Litters are housed with their mothers until weaning. F<sub>1</sub> animals should be housed in small groups of the same sex and treatment group from weaning to termination. If scientifically justified, animals can be housed individually. The level of phytoestrogens contained in the selected bedding material should be minimal.

*Number and identification of animals*

14. Normally, each test and control group should contain a sufficient number of mating pairs to yield at least 20 pregnant females per dose group. The objective is to produce enough pregnancies to ensure a meaningful evaluation of the potential of the chemical to affect fertility, pregnancy and maternal behaviour of the P generation and growth and development of the F<sub>1</sub> offspring, from conception to maturity. Failure to achieve the desired number of pregnant animals does not necessarily invalidate the study and should be evaluated on a case-by-case basis, considering a possible causal relationship to the test chemical.
  
15. Each P animal is assigned a unique identification number before dosing starts. If laboratory historical data suggest that a significant proportion of females may not show regular (4 or 5-day) oestrous cycles, then an assessment of oestrous cycles before start of treatment is advised. Alternatively, the group size may be increased to ensure that at least 20 females in each group would have regular (4 or 5-day) oestrous cycles at start of treatment. All F<sub>1</sub> offspring are uniquely identified when neonates are first examined on postnatal day (PND) 0 or 1. Records indicating the litter of origin should be maintained for all F<sub>1</sub> animals, and F<sub>2</sub> animals where applicable, throughout the study.

**▼ M5****Test chemical***Available information on the test chemical*

16. The review of existing information is important for decisions on the route of administration, the choice of the vehicle, the selection of animal species, the selection of dosages and potential modifications of the dosing schedule. Therefore, all the relevant available information on the test chemical, i.e. physico-chemical, toxicokinetics (including species-specific metabolism), toxicodynamic properties, structure-activity relationships (SARs), *in vitro* metabolic processes, results of previous toxicity studies and relevant information on structural analogues should be taken into consideration in planning the Extended One-Generation Reproductive Toxicity Study. Preliminary information on absorption, distribution, metabolism and elimination (ADME) and bioaccumulation may be derived from chemical structure, physico-chemical data, extent of plasma protein binding or toxicokinetic (TK) studies, while results from toxicity studies give additional information, e.g. on NOAEL, metabolism or induction of metabolism.

*Consideration of toxicokinetic data*

17. Although not required, TK data from previously conducted dose range-finding or other studies are extremely useful in the planning of the study design, selection of dose levels and interpretation of results. Of particular utility are data which: 1) verify exposure of developing fetuses and pups to the test chemical (or relevant metabolites), 2) provide an estimate of internal dosimetry, and 3) evaluate for potential dose-dependent saturation of kinetic processes. Additional TK data, such as metabolite profiles, concentration-time courses, etc. should also be considered, if they are available. Supplemental TK data may also be collected during the main study, provided that it does not interfere with the collection and interpretation of the main study endpoints.

As a general guide, the following TK data set would be useful in planning the Extended One-Generation Reproductive Toxicity Study:

— Late pregnancy (e.g. Gestation Day 20) — maternal blood and foetal blood

— Mid-lactation (PND 10) — maternal blood, pup blood and/or milk

— Early post-weaning (e.g. PND 28) — weanling blood samples.

Flexibility should be employed in determining the specific analytes (e.g. parent chemical and/or metabolites) and sampling scheme. For example, the number and timing of sample collection on a given sampling day will be dependent upon route of exposure and prior knowledge of TK properties in non-pregnant animals. For dietary studies, sampling at a single consistent time on each of these days is sufficient, whereas gavage dosing may warrant additional sampling times to obtain a better estimate of the range of internal doses. However, it is not necessary to generate a full concentration time-course on any of the sampling days. If necessary, blood can be pooled by sex within litters for fetal and neonatal analyses.

**▼ M5***Route of administration*

18. Selection of the route should take into consideration the route(s) most relevant for human exposure. Although the protocol is designed for administration of the test chemical through the diet, it can be modified for administration by other routes (drinking water, gavage, inhalation, dermal), depending on the characteristics of the chemical and the information required.

*Choice of the vehicle*

19. Where necessary, the test chemical is dissolved or suspended in a suitable vehicle. It is recommended that, where possible, the use of an aqueous solution/suspension is considered first, followed by consideration of a solution/suspension in oil (e.g. corn oil). For vehicles other than water, the toxic characteristics of the vehicle should be known. Use of vehicles with potential intrinsic toxicity should be avoided (e.g. acetone, DMSO). The stability of the test chemical in the vehicle should be determined. Considerations should be given to the following characteristics if a vehicle or other additive is used to facilitate dosing: effects on the absorption, distribution, metabolism, or retention of the test chemical; effects on the chemical properties of the test chemical that may alter its toxic characteristics; and effects on the food or water consumption or the nutritional status of the animals.

*Dose selection*

20. Normally, the study should include at least three dose levels and a concurrent control. When selecting appropriate dose levels, the investigator should consider all available information, including the dosing information from previous studies, TK data from pregnant or non-pregnant animals, the extent of lactational transfer, and estimates of human exposure. If TK data are available which indicate dose-dependent saturation of TK processes, care should be taken to avoid high dose levels which clearly exhibit saturation, provided of course, that human exposures are expected to be well below the point of saturation. In such cases, the highest dose level should be at, or just slightly above the inflection point for transition to nonlinear TK behaviour.
21. In the absence of relevant TK data, the dose levels should be based on toxic effects, unless limited by the physical/chemical nature of the test chemical. If dose levels are based on toxicity, the highest dose should be chosen with the aim to induce some systemic toxicity, but not death or severe suffering of the animals.
22. A descending sequence of dose levels should be selected in order to demonstrate any dose-related effect and to establish NOAELs or doses near the limit of detection that would allow for derivation of a benchmark dose for the most sensitive endpoint(s). To avoid large dose spacing between NOAELs and LOAELs, two- or four-fold intervals are frequently optimal. The addition of a fourth test group is often preferable to using a very large interval (e.g. more than a factor of 10) between doses.
23. Except for treatment with the test chemical, animals in the control group are handled in an identical manner to the test group subjects. This group should be untreated or sham-treated or a vehicle-control group if a vehicle is used in administering the test chemical. If a vehicle is used, the control group should receive the vehicle in the highest volume used.

**▼ M5***Limit test*

24. If there is no evidence of toxicity at a dose of at least 1 000 mg/kg body weight/day in repeat-dose studies, or if toxicity would not be expected based upon data from structurally- and/or metabolically-related chemicals, indicating similarity in the *in vivo/in vitro* metabolic properties, a study using several dose levels may not be necessary. In such cases, the Extended One-Generation Reproductive Toxicity Study could be conducted using a control group and a single dose of at least 1 000 mg/kg body weight/day. However, should evidence for reproductive or developmental toxicity be found at this limit dose, further studies at lower dose levels will be required to identify a NOAEL. These limit test considerations apply only when human exposure does not indicate the need for a higher dose level.

## PROCEDURES

**Exposure of offspring**

25. Dietary exposure is the preferred method of administration. If gavage studies are performed, it should be noted that the pups will normally only receive test chemical indirectly through the milk, until direct dosing commences for them at weaning. In diet or drinking water studies, the pups will additionally receive test chemical directly when they commence eating for themselves during the last week of the lactation period. Modifications to the study design should be considered when excretion of the test chemical in milk is poor and where there is lack of evidence for a continuous exposure of the offspring. In these cases, direct dosing of pups during the lactation period should be considered based on available TK information, offspring toxicity or changes in bio-markers (3) (4). Careful consideration of benefits and disadvantages should be made prior to conducting direct-dosing studies on nursing pups (5).

**Dosing schedule and administration of doses**

26. Some information on oestrous cycles, male and female reproductive tract histopathology and testicular/epididymal sperm analysis may be available from previous repeat-dose toxicity studies of adequate duration. The duration of the pre-mating treatment in the Extended One-Generation Reproductive Toxicity Study is therefore aimed at the detection of effects on functional changes that may interfere with mating behaviour and fertilisation. The pre-mating treatment should be sufficiently long to achieve steady-state exposure conditions in P males and females. A 2-week pre-mating treatment for both sexes is considered adequate in most cases. For females, this covers 3-4 complete oestrous cycles and should be sufficient to detect any adverse effects on cyclicity. For males, this is equivalent to the time required for epididymal transit of maturing spermatozoa and should allow the detection of post-testicular effects on sperm (during the final stages of spermiation and epididymal sperm maturation) at mating. At the time of termination, when testicular and epididymal histopathology and analysis of sperm parameters are scheduled, the P and F<sub>1</sub> males, will have been exposed for at least one entire spermatogenic process ((6) (7) (8) (9) and OECD Guidance Document 151(40)).

**▼ M5**

27. Pre-mating exposure scenarios for males could be adapted if testicular toxicity (impairment of spermatogenesis) or effects on sperm integrity and function have been clearly identified in previous studies. Similarly, for females, known effects of the test chemical on the oestrous cycle and thus sexual receptivity, may justify different pre-mating exposure scenarios. In special cases it may be acceptable that treatment of the P females is initiated only after a sperm-positive smear has been obtained (see OECD Guidance Document 151(40)).
  
28. Once the pre-mating dosing period is established, the animals should be treated with the test chemical continuously on a 7-days/week basis until necropsy. All animals should be dosed by the same method. Dosing should continue during the 2-week mating period and, for P females, throughout gestation and lactation up to the day of termination after weaning. Males should be treated in the same manner until termination at the time when the F<sub>1</sub> animals are weaned. For necropsy, priority should be given to females which should be necropsied on the same/similar day of lactation. Necropsy of males can be spread over a larger number of days, depending on laboratory facilities. Unless already initiated during the lactation period, direct dosing of the selected F<sub>1</sub> males and females should begin at weaning and continue until scheduled necropsy, depending on cohort assignment.
  
29. For chemicals administered via the diet or drinking water, it is important to ensure that the quantities of the test chemical involved do not interfere with normal nutrition or water balance. When the test chemical is administered in the diet, either a constant dietary concentration (ppm) or a constant dose level in terms of the body weight of the animal may be employed; the option chosen should be specified.
  
30. When the test chemical is administered by gavage, the volume of liquid administered at one time should not normally exceed 1 ml/100 g body weight (0,4 ml/100 g body weight is the maximum for oil, e.g. corn oil). Except for irritant or corrosive chemicals, which will normally reveal exacerbated effects with higher concentrations, variability in test volume should be minimised by adjusting the concentration to ensure a constant volume at all dose levels. The treatment should be given at similar times each day. The dose to each animal should normally be based on the most recent individual bodyweight determination and adjusted at least weekly in adult males and adult non-pregnant females, and every two days in pregnant females and F<sub>1</sub> animals when administered prior to weaning and during the 2 weeks following weaning. If TK data indicate a low placental transfer of the test chemical, the gavage dose during the last week of pregnancy may have to be adjusted to prevent administration of an excessively toxic dose to the dam. Females should not be treated by gavage, or any other route of treatment where the animal needs to be handled, on the day of parturition; omission of test chemical administration on that day is preferable to a disturbance of the birth process.

**▼ M5****Mating**

31. Each P female should be placed with a single, randomly selected, unrelated male from the same dose group (1:1 pairing) until evidence of copulation is observed or 2 weeks have elapsed. If there are insufficient males, for example due to male death before pairing, then male(s) which have already mated may be paired (1:1) with a second female(s) such that all females are paired. Day 0 of pregnancy is defined as the day on which mating evidence is confirmed (a vaginal plug or sperm are found). Animals should be separated as soon as possible after evidence of copulation is observed. If mating has not occurred after 2 weeks, the animals should be separated without further opportunity for mating. Mating pairs should be clearly identified in the data.

**Litter size**

32. On day 4 after birth, the size of each litter may be adjusted by eliminating extra pups by random selection to yield, as nearly as possible, five males and five females per litter. Selective elimination of pups, e.g. based upon body weight, is not appropriate. Whenever the number of male or female pups prevents having five of each sex per litter, partial adjustment (for example, six males and four females) is acceptable.

**Selection of pups for post-weaning studies (see Figure 1)**

33. At weaning (around PND 21) pups from all available litters up to 20 per dose and control group are selected for further examinations and maintained until sexual maturation (unless earlier testing is required). Pups are selected randomly, with the exception that obvious runts (animals with a body weight more than two standard deviations below the mean pup weight of the respective litter) should not be included, as they are unlikely to be representative of the treatment group.

On PND 21, the selected F<sub>1</sub> pups are randomly assigned to one of three cohorts of animals, as follows:

Cohort 1 (1A and 1B)= Reproductive/developmental toxicity testing

Cohort 2 (2A and 2B)= Developmental neurotoxicity testing

Cohort 3= Developmental immunotoxicity testing

Cohort 1A: One male and one female/litter/group (20/sex/group): priority selection for primary assessment of effects upon reproductive systems and of general toxicity.

Cohort 1B: One male and one female/litter/group (20/sex/group): priority selection for follow-up assessment of reproductive performance by mating F1 animals, when assessed (see OECD Guidance Document 117(39)), and for obtaining additional histopathology data in cases of suspected reproductive or endocrine toxicants, or when results from cohort 1A are equivocal.

Cohort 2A: Total of 20 pups per group (10 males and 10 females per group; one male or one female per litter) assigned for neurobehavioral testing followed by neurohistopathology assessment as adults.

Cohort 2B: Total of 20 pups per group (10 males and 10 females per group; one male or one female per litter) assigned for neurohistopathology assessment at weaning (PND 21 or PND 22). If there are insufficient numbers of animals, preference should be given to assign animals to Cohort 2A.



**▼ M5**

Cohort 3: Total of 20 pups per group (10 males and 10 females per group; one per litter, where possible). Additional pups may be required from the control group to act as positive control animals in the T-cell dependant antibody response assay (TDAR) at PND  $56 \pm 3$ .

34. Should there be an insufficient number of pups in a litter to serve all cohorts, the cohort 1 takes precedence, as it can be extended to produce an F<sub>2</sub> generation. Additional pups may be assigned to any of the cohorts in case of specific concern, e.g. if a chemical is suspected to be a neurotoxicant, immunotoxicant or reproductive toxicant. These pups may be used for examinations at different timepoints or for the evaluation of supplementary endpoints. Pups not assigned to cohorts will be submitted to clinical biochemistry (paragraph 55) and gross necropsy (paragraph 68).

**Second mating of the P animals**

35. A second mating is not normally recommended for the P animals, as it comes at the expense of losing important information on the number of implantation sites (and thus post-implantation and peri-natal loss data, indicators of a possible teratogenic potential) for the first litter. The need to verify or elucidate an effect in exposed females would be served better by extending the study to include a mating of the F<sub>1</sub> generation. However, a second mating of the P males with untreated females is always an option to clarify equivocal findings or for further characterisation of effects on fertility observed in the first mating.

**IN-LIFE OBSERVATIONS****Clinical observations**

36. For the P and the selected F<sub>1</sub> animals, a general clinical observation is made once a day. In the case of gavage dosing, the timing of clinical observations should be prior to and post dosing (for possible signs of toxicity associated with peak plasma concentration). Pertinent behavioural changes, signs of difficult or prolonged parturition and all signs of toxicity are recorded. Twice daily, during the weekend once daily, all animals are observed for severe toxicity, morbidity and mortality.
  
37. In addition, a more detailed examination of all P and F<sub>1</sub> animals (after weaning) is conducted on a weekly basis and could conveniently be performed on an occasion when the animal is weighed, which would minimise handling stress. Observations should be carefully conducted and recorded using scoring systems that have been defined by the testing laboratory. Efforts should be made to ensure that variations in the test conditions are minimal. Signs noted should include, but not be limited to, changes in skin, fur, eyes, mucous membranes, occurrence of secretions and excretions and autonomic activity (e.g. lacrimation, piloerection, pupil size, unusual respiratory pattern). Changes in gait, posture, response to handling, as well as the presence of clonic or tonic movements, stereotypy (e.g. excessive grooming, repetitive circling) or bizarre behaviour (e.g. self-mutilation, walking backwards) should also be recorded.

**▼ M5****Body weight and food/water consumption**

38. P animals are weighed on the first day of dosing and at least weekly thereafter. In addition, P females are weighed during lactation on the same days as the weighing of the pups in their litters (see paragraph 44). All F<sub>1</sub> animals are weighed individually at weaning (PND 21) and at least weekly thereafter. Body weight is also recorded on the day when they attain puberty (completion of preputial separation or vaginal patency). All animals are weighed at sacrifice.
39. During the study, food and water consumption (in the case of test chemical administration in the drinking water) are recorded at least weekly on the same days as animal body weights (except during cohabitation). The food consumption of each cage of F<sub>1</sub> animals is recorded weekly commencing with selection to a respective cohort.

**Oestrous cycles**

40. Preliminary information of test chemical-related effects on the oestrous cycle may already be available from previous repeat-dose toxicity studies, and may be used in designing a test chemical-specific protocol for the Extended One-Generation Reproductive Toxicity Study. Normally the assessment of oestrous cyclicity (by vaginal cytology) will start at the beginning of the treatment period and continue until confirmation of mating or the end of the 2-week mating period. If females have been screened for normal oestrous cycles before treatment, then it is useful to continue smearing as treatment starts, but if there is concern about non-specific effects at the start of treatment (such as an initial marked reduction in food consumption) then animals may be allowed to adapt to treatment for up to two weeks before the start of the 2-week smearing period leading into pairing. If the female treatment period is extended in this way (i.e. to a 4-week pre-mating treatment) then consideration should be made to purchasing animals younger and to extending the period of male treatment before pairing. When obtaining vaginal/cervical cells, care should be taken to avoid disturbance of mucosa and subsequently, the induction of pseudopregnancy (10) (11).
41. Vaginal smears should be examined daily for all F<sub>1</sub> females in cohort 1A, after the onset of vaginal patency, until the first cornified smear is recorded, in order to determine the time interval between these two events. Oestrous cycles for all F<sub>1</sub> females in cohort 1A should also be monitored for a period of two weeks, commencing around PND 75. In addition, should mating of the F<sub>1</sub> generation be necessary, the vaginal cytology in cohort 1B will be followed from the time of pairing until mating evidence is detected.

**Mating and pregnancy**

42. In addition to the standard endpoints (e.g. body weight, food consumption, clinical observations including mortality/morbidity checks), the dates of pairing, the date of insemination and the date of parturition are recorded and the precoital interval (pairing to insemination) and the duration of pregnancy (insemination to parturition) are calculated. The P females should be examined carefully at the time of expected parturition for any signs of dystocia. Any abnormalities in nesting behaviour or nursing performance should be recorded.

**▼ M5**

43. The day on which parturition occurs is lactation day 0 (LD 0) for the dam and postnatal day 0 (PND 0) for the offspring. Alternatively, all comparisons may also be based on post-coital time to eliminate confounding of postnatal development data, by differences in the duration of pregnancy; however, timing relative to parturition should also be recorded. This is especially important when the test chemical exerts an influence on the duration of pregnancy.

**Offspring parameters**

44. Each litter should be examined as soon as possible after parturition (PND 0 or 1) to establish the number and sex of pups, stillbirths, live births, and the presence of gross anomalies (externally visible abnormalities, including cleft palate; subcutaneous haemorrhages; abnormal skin colour or texture; presence of umbilical cord; lack of milk in stomach; presence of dried secretions). In addition, the first clinical examination of the neonates should include a qualitative assessment of body temperature, state of activity and reaction to handling. Pups found dead on PND 0 or at a later time should be examined for possible defects and cause of death. Live pups are counted and weighed individually on PND 0 or PND 1, and regularly thereafter, e.g. at least on PND 4, 7, 14, and 21. Clinical examinations, as applicable for the age of the animals, should be repeated when the offspring are weighed, or more often if case-specific findings have been made at birth. Signs noted could include, but may not be limited to, external abnormalities, changes in skin, fur, eyes, mucous membranes, occurrence of secretions and excretions and autonomic activity. Changes in gait, posture, response to handling, as well as the presence of clonic or tonic movements, stereotypy or bizarre behaviour, should also be recorded.
45. The anogenital distance (AGD) of each pup should be measured on at least one occasion from PND 0 through PND 4. Pup body weight should be collected on the day the AGD is measured and the AGD should be normalised to a measure of pup size, preferably the cube root of body weight (12). The presence of nipples/areolae in male pups should be checked on PND 12 or 13.
46. All selected F<sub>1</sub> animals are evaluated daily for balano-preputial separation or vaginal patency for male/female respectively commencing before the expected day for achievement of these endpoints to detect if sexual maturation occurs early. Any abnormalities of genital organs, such as persistent vaginal thread, hypospadias or cleft penis, should be noted. Sexual maturity of F<sub>1</sub> animals is compared to physical development by determining age and body weight at balano-preputial separation or vaginal opening for male/female respectively (13).

**Assessment of potential developmental neurotoxicity (cohorts 2A and 2B)**

47. Ten male and 10 female cohort 2A animals and 10 male and 10 female cohort 2B animals, from each treatment group (for each cohort: 1 male or 1 female per litter; all litters represented by at least 1 pup; randomly selected) should be used for neurotoxicity assessments. Cohort 2A animals should be subjected to auditory startle, functional observational battery, motor activity (see paragraphs 48-50), and neuropathology assessments (see paragraphs 74-75). Efforts should be made to ensure that variations in all test

**▼ M5**

conditions are minimal and are not systematically related to treatment. Among the variables that can affect behaviour are sound level (e.g. intermittent noise), temperature, humidity, lighting, odours, time of day, and environmental distractions. Results of the neurotoxicity assays should be interpreted in relation to appropriate historical control reference ranges. Cohort 2B animals should be used for neuropathology assessment on PND 21 or PND 22 (see paragraphs 74-75).

48. An auditory startle test should be performed on PND 24 ( $\pm$  1 day) using animals in cohort 2A. The day of testing should be counterbalanced across treated and control groups. Each session consists of 50 trials. In performing the auditory startle test, the mean response amplitude on each block of 10 trials (5 blocks of 10 trials) should be determined, with test conditions optimised to produce intra-session habituation. These procedures should be consistent with test method B.53 (35).
  
49. At an appropriate time between PND 63 and PND 75, the cohort 2A animals are subjected to a functional observational battery and an automated test of motor activity. These procedures should be consistent with test methods B.43 (33) and B.53 (35). The functional observational battery includes a thorough description of the subject's appearance, behaviour and functional integrity. This is assessed through observations in the home cage, after removal to a standard arena for observation (open field) where the animal is moving freely, and through manipulative tests. Testing should proceed from the least to the most interactive. A list of measures is presented in Appendix 1. All animals should be observed carefully by trained observers who are unaware of the animals' treatment status, using standardised procedures to minimise observer variability. Where possible, it is advisable that the same observer evaluates the animals in a given test. If this is not possible, some demonstration of inter-observer reliability is required. For each parameter in the behavioural testing battery, explicit operationally defined scales and scoring criteria are to be used. If possible, objective quantitative measures should be developed for observational endpoints, which involve subjective ranking. For motor activity, each animal is tested individually. The test session should be long enough to demonstrate intra-session habituation for controls. Motor activity should be monitored by an automated activity recording apparatus which should be capable of detecting both increases and decreases in activity, (i.e. baseline activity as measured by the device should not be so low as to preclude detection of decreases, nor so high as to preclude detection of increases in activity). Each device should be tested by standard procedures to ensure, to the extent possible, reliability of operation across devices and across days. To the extent possible, treatment groups should be balanced across devices. Treatment groups should be counter-balanced across test times to avoid confounding by circadian rhythms of activity.
  
50. If existing information indicates the need for other functional testing (e.g. sensory, social, cognitive), these should be integrated without compromising the integrity of the other evaluations conducted in the study. If this testing is performed in the same animals as used for standard auditory startle, functional observational battery and motor activity testing, different tests should be scheduled to minimise the risk of compromising the integrity of these tests. Supplemental procedures may be particularly useful when empirical observation, anticipated effects, or mechanistic/mode-of-action indicate a specific type of neurotoxicity.

**▼ M5****Assessment of potential developmental immunotoxicity (cohort 3)**

51. At PND 56 ( $\pm$  3 days), 10 male and 10 female cohort 3 animals from each treatment group (1 male or 1 female per litter; all litters represented by at least 1 pup; randomly selected) should be used in a T-cell dependant antibody response assay, i.e. the primary IgM antibody response to a T-cell dependent antigen, such as Sheep Red Blood Cells (SRBC) or Keyhole Limpet Hemocyanin (KLH), consistent with current immunotoxicity testing procedures (14) (15). The response may be evaluated by counting specific plaque-forming cells (PFC) in the spleen or by determining the titer of SRBC- or KLH-specific IgM antibody in the serum by ELISA, at the peak of the response. Responses typically peak four (PFC response) or five (ELISA) days after intravenous immunisation. If the primary antibody response is assayed by counting plaque-forming cells, it is permissible to evaluate subgroups of animals on separate days, provided that: subgroup immunisation and sacrifice are timed so that PFCs are counted at the peak of the response; that subgroups contain an equal number of male and female offspring from all dose groups, including controls; and that subgroups are evaluated at approximately the same postnatal age. Exposure to the test chemical will continue until the day before collecting spleens for the PFC response or serum for the ELISA assay.

**Follow-up assessment of potential reproductive toxicity (cohort 1B)**

52. Cohort 1B animals can be maintained on treatment beyond PND 90 and bred to obtain a F<sub>2</sub> generation if necessary. Males and females of the same dose group should be cohabited (avoiding the pairing of siblings) for up to two weeks, beginning on or after PND 90, but not exceeding PND 120. Procedures should be similar to those for the P animals. However, based on a weight of evidence, it may suffice to terminate the litters on PND 4 rather than follow them to weaning or beyond.

**TERMINAL OBSERVATIONS****Clinical biochemistry/Haematology**

53. Systemic effects should be monitored in P animals. Fasted blood samples from a defined site are taken from 10 randomly-selected P males and females per dose group at termination, stored under appropriate conditions and subjected to partial or full-scale haematology, clinical biochemistry, assay of T4 and TSH or other examinations suggested by the known effect profile of the test chemical (see OECD Guidance Document 151(40)). The following haematological parameters should be examined: haematocrit, haemoglobin concentration, erythrocyte count, total and differential leukocyte count, platelet count and blood clotting time/potential. Investigations of plasma or serum should include: glucose, total cholesterol, urea, creatinine, total protein, albumin and at least two enzymes indicative of hepatocellular effects (such as alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, gamma glutamyl transpeptidase and sorbitol dehydrogenase). Measurements of additional enzymes and bile acids may provide useful information under certain circumstances. In addition, blood from all animals may be taken and stored for possible analysis at a later time to help clarify equivocal effects or to generate

**▼ M5**

internal exposure data. If a second mating of P animals is not intended, the blood samples are obtained just prior to, or as part of, the procedure at scheduled sacrifice. In the case animals are retained, blood samples should be collected a few days before the animals are mated for the second time. Unless existing data from repeated-dose studies indicate that the parameter is not affected by the test chemical, urinalysis should be performed prior to termination and the following parameters evaluated: appearance, volume, osmolality or specific gravity, pH, protein, glucose, blood and blood cells, cell debris. Urine may also be collected to monitor excretion of test chemical and/or metabolite(s).

54. Systemic effects should also be monitored in F<sub>1</sub> animals. Fasted blood samples from a defined site are taken from 10 randomly selected cohort 1A males and females per dose group at termination, stored under appropriate conditions and subjected to standard clinical biochemistry, including the assessment of serum levels for thyroid hormones (T4 and TSH), haematology (total and differential leukocyte plus erythrocyte counts) and urinalysis assessments.
55. The surplus pups at PND 4 are subject to gross necropsy and consideration given to measuring serum thyroid hormone (T4) concentrations. If necessary, neonatal (PND 4) blood can be pooled by litters for biochemical/thyroid hormone analyses. Blood is also collected for T4 and TSH analysis from weanlings subject to gross necropsy on PND 22 (F<sub>1</sub> pups not selected for cohorts).

**Sperm parameters**

56. Sperm parameters should be measured in all P generation males unless there is existing data to show that sperm parameters are unaffected in a 90-day study. Examination of sperm parameters should be performed in all cohort 1A males.
57. At termination, testis and epididymis weights are recorded for all P and F<sub>1</sub> (cohort 1A) males. At least one testis and one epididymis are reserved for histopathological examination. The remaining epididymis is used for enumeration of cauda epididymis sperm reserves (16) (17). In addition, sperm from the cauda epididymis (or vas deferens) is collected using methods that minimise damage for evaluation of sperm motility and morphology (18).
58. Sperm motility can either be evaluated immediately after sacrifice or recorded for later analysis. The percentage of progressively motile sperm could be determined either subjectively or objectively by computer-assisted motion analysis (19) (20) (21) (22) (23) (24). For the evaluation of sperm morphology, an epididymal (or vas deferens) sperm sample should be examined as fixed or wet preparations (25) and at least 200 spermatozoa per sample classified as either normal (both head and midpiece/tail appear normal) or abnormal. Examples of morphologic sperm abnormalities would include fusion, isolated heads, and misshapen heads and/or tails (26). Misshapen or large sperm heads may indicate defects in spermiation.
59. If sperm samples are frozen, smears fixed and images for sperm motility analysis recorded at the time of necropsy (27), subsequent analysis may be restricted to control and high-dose males. However, if treatment-related effects are observed, the lower dose groups should also be evaluated.

**▼ M5****Gross necropsy**

60. At the time of termination or premature death, all P and F<sub>1</sub> animals are necropsied and examined macroscopically for any structural abnormalities or pathological changes. Special attention should be paid to the organs of the reproductive system. Pups that are humanely killed in a moribund condition and dead pups should be recorded and, when not macerated, examined for possible defects and/or cause of death and preserved.
61. For adult P and F<sub>1</sub> females, a vaginal smear is examined on the day of necropsy to determine the stage of the oestrous cycle and allow correlation with histopathology in reproductive organs. The uteri of all P females (and F<sub>1</sub> females, if applicable) are examined for the presence and number of implantation sites, in a manner which does not compromise histopathological evaluation.

**Organ weight and tissue preservation — P and F<sub>1</sub> adult animals**

62. At the time of termination, body weights and wet weights of the organs listed below from all P animals and all F<sub>1</sub> adults, from relevant cohorts (as outlined below), are determined as soon as possible after dissection to avoid drying. These organs should then be preserved under appropriate conditions. Unless specified otherwise, paired organs can be weighed individually or combined, consistent with the typical practice of the performing laboratory.
- Uterus (with oviducts and cervix), ovaries
  - Testes, epididymides (total and cauda for the samples used for sperm counts)
  - Prostate (dorsolateral and ventral parts combined). Care should be exercised when trimming the prostate complex to avoid puncture of the fluid filled seminal vesicles. In the event of a treatment-related effect on total prostate weight, the dorsolateral and ventral segments should be carefully dissected after fixation, and weighed separately.
  - Seminal vesicles with coagulating glands and their fluids (as one unit)
  - Brain, liver, kidneys, heart, spleen, thymus, pituitary, thyroid (post-fixation), adrenal glands and known target organs or tissues.
63. In addition to the organs listed above, samples of peripheral nerve, muscle, spinal cord, eye plus optic nerve, gastrointestinal tract, urinary bladder, lung, trachea (with thyroid and parathyroid attached), bone marrow, vas deferens (males), mammary gland (males and females) and vagina should be preserved under appropriate conditions.
64. Cohort 1A animals have all organs weighed and preserved for histopathology.
65. For the investigation of pre- and postnatally induced immunotoxic effects, 10 male and 10 female cohort 1A animals from each treatment group (1 male or 1 female per litter; all litters represented by at least 1 pup; randomly selected) will be subject to the following at termination:
- weighing of the lymph nodes associated with and distant from the route of exposure (in addition to the weight of the adrenal glands, the thymus and the spleen, already performed in all cohort 1A animals)

**▼ M5**

- splenic lymphocyte subpopulation analysis (CD4+ and CD8+ T lymphocytes, B lymphocytes, and natural killer cells) using one half of the spleen, the other half of the spleen being preserved for histopathological evaluation,

Analysis of splenic lymphocyte subpopulations in non-immunised (cohort 1A) animals will determine if exposure is related to a shift in the immunological steady state distribution of ‘helper’ (CD4+) or cytotoxic (CD8+) thymus-derived lymphocytes or natural killer (NK) cells (rapid responses to neoplastic cells and pathogens).

66. Cohort 1B animals should have the following organs weighed and corresponding tissues processed to the block stage:

- Vagina (not weighed)
- Uterus with cervix
- Ovaries
- Testes (at least one)
- Epididymides
- Seminal vesicles and coagulating glands
- Prostate
- Pituitary
- Identified target organs

Histopathology in cohort 1B would be conducted if results from cohort 1A are equivocal or in cases of suspected reproductive or endocrine toxicants.

67. Cohorts 2A and 2B: Developmental neurotoxicity testing (PND 21 or PND 22 and adult offspring). Cohort 2A animals are terminated after behavioural testing, with brain weight recorded and full neurohistopathology for purposes of neurotoxicity assessment. Cohort 2B animals are terminated on PND 21 or PND 22, with brain weight recorded and microscopic examination of the brain for purposes of neurotoxicity assessment. Perfusion fixation is required for cohort 2A animals and optional for cohort 2B animals, as provided in test method B.53 (35).

**Organ weight and tissue preservation — F<sub>1</sub> weanlings**

68. The pups not selected for cohorts, including runts, are terminated after weaning, on PND 22, unless the results indicate the need for further in-life investigations. Terminated pups are subjected to gross necropsy including an assessment of the reproductive organs, as described in paragraphs 62 and 63. For up to 10 pups per sex per group, from as many litters as possible, brain, spleen, and thymus should be weighed and retained under appropriate conditions. In addition, mammary tissues for these male and female pups may be preserved for further microscopic analysis<sup>(1)</sup> (see OECD Guidance Document 151(40)). Gross abnormalities and target tissues should be saved for possible histological examination.

<sup>(1)</sup> Research has shown the mammary gland, especially in early life mammary gland development, to be a sensitive endpoint for oestrogen action. It is recommended that endpoints involving pup mammary glands of both sexes be included in this test method, when validated.



**▼ M5****Histopathology — P animals**

69. Full histopathology of the organs listed in paragraphs 62 and 63 is performed for all high-dose and control P animals. Organs demonstrating treatment-related changes should also be examined in all animals at the lower dose groups to aid in determining a NOAEL. Additionally, reproductive organs of all animals suspected of reduced fertility, e.g. those that failed to mate, conceive, sire, or deliver healthy offspring, or for which oestrous cyclicity or sperm number, motility, or morphology were affected, and all gross lesions should be subjected to histopathological evaluation.

**Histopathology — F<sub>1</sub> animals***Cohort 1 animals*

70. Full histopathology of the organs listed in paragraphs 62 and 63 is performed for all high-dose and control adult cohort 1A animals. All litters should be represented by at least 1 pup per sex. Organs and tissues demonstrating treatment-related changes and all gross lesions should also be examined in all animals in the lower dose groups to aid in determining a NOAEL. For the evaluation of pre- and postnatally induced effects on lymphoid organs also the histopathology on the collected lymph nodes and bone marrow should be evaluated of 10 male and 10 female cohort 1A animals next to histopathological evaluation of the thymus, spleen, and the adrenal glands already performed in all 1A animals.
71. Reproductive and endocrine tissues from all cohort 1B animals, processed to the block stage as described in paragraph 66, should be examined for histopathology in cases of suspected reproductive or endocrine toxicants. Cohort 1B should also undergo histological examination if results from cohort 1A are equivocal.
72. Ovaries of adult females should contain primordial and growing follicles, as well as corpora lutea; therefore, a histopathological examination should be aimed at detecting a quantitative evaluation of primordial and small growing follicles, as well as corpora lutea, in F<sub>1</sub> females; the number of animals, ovarian section selection, and section sample size should be statistically appropriate for the evaluation procedure used. Follicular enumeration may first be conducted on control and high-dose animals, and in the event of an adverse effect in the latter, lower doses should be examined. Examination should include enumeration of the number of primordial follicles, which can be combined with small growing follicles, for comparison of treated and control ovaries (see OECD Guidance Document 151(40)). Corpora lutea assessment should be conducted in parallel with oestrous cyclicity testing so that the stage of the cycle can be taken into account in the assessment. Oviduct, uterus and vagina are examined for appropriate organ-typic development.
73. Detailed testicular histopathology examinations are conducted on the F<sub>1</sub> males in order to identify treatment-related effects on testis differentiation and development and on spermatogenesis (38). When possible, sections of the rete testis should be examined. Caput, corpus, and cauda of the epididymis and the vas deferens are examined for appropriate organ-typic development, as well as for the parameters required for the P males.

**▼ M5***Cohort 2 animals*

74. Neurohistopathology is performed for all high-dose and control cohort 2A animals per sex following completion of neurobehavioral testing (after PND 75, but not to exceed PND 90). Brain histopathology is performed for all high-dose and control cohort 2B animals per sex on PND 21 or PND 22. Organs or tissues demonstrating treatment-related changes should also be examined for the animals in the lower dose groups to aid in determining a NOAEL. For cohort 2A and 2B animals, multiple sections are examined from the brain to allow examination of olfactory bulbs, cerebral cortex, hippocampus, basal ganglia, thalamus, hypothalamus, mid-brain (thecum, tegmentum, and cerebral peduncles), brain-stem and cerebellum. For cohort 2A only, the eyes (retina and optic nerve) and samples of peripheral nerve, muscle and spinal cord are examined. All neurohistological procedures should be consistent with test method B.53 (35).
75. Morphometric (quantitative) evaluations should be performed on representative areas of the brain (homologous sections carefully selected based on reliable microscopic landmarks) and may include linear and/or areal measurements of specific brain regions. At least three consecutive sections should be taken at each landmark (level) in order to select the most homologous and representative section for the specific brain area to be evaluated. The neuropathologist should exercise appropriate judgment as to whether sections prepared for measurement are homologous with others in the sample set and therefore suitable for inclusion, since linear measurements in particular may change over a relatively short distance (28). Non-homologous sections should not be used. While the objective is to sample all animals reserved for this purpose (10/sex/dose level), smaller numbers may still be adequate. However, samples from fewer than 6 animals/sex/dose level would generally not be considered sufficient for the purposes of this test method. Stereology may be used to identify treatment-related effects on parameters such as volume or cell number for specific neuroanatomic regions. All aspects of the preparation of tissue samples, from tissue fixation, through the dissection of tissue samples, tissue processing, and staining of slides, should employ a counterbalanced design, such that each batch contains representative samples from each dose group. When morphometric or stereological analyses are to be used, then brain tissue should be embedded in appropriate media at all dose levels at the same time in order to avoid shrinkage artefacts associated with prolonged storage in fixative.

**REPORTING****Data**

76. Data are reported individually and summarised in tabular form. Where appropriate, for each test group and each generation, the following should be reported: number of animals at the start of the test, number of animals found dead during the test or killed for humane reasons, time of any death or humane kill, number of fertile animals, number of pregnant females, number of females giving birth to a litter, and number of animals showing signs of toxicity. A description of the toxicity, including time of onset, duration, and severity should also be reported.
77. Numerical results should be evaluated by an appropriate, and accepted statistical method. The statistical methods should be selected as part of the study design and should appropriately address non-normal data (e.g. count data), censored data (e.g. limited observation time), non-independence (e.g. litter effects and repeated measures), and unequal

**▼ M5**

variances. Generalised linear mixed models and dose-response models cover a broad class of analytical tools that may be appropriate for the data generated under this test method. The report should include sufficient information on the method of analysis and the computer program employed, so that an independent reviewer/statistician can evaluate/re-evaluate the analysis.

**Evaluation of results**

78. The findings should be evaluated in terms of the observed effects, including necropsy and microscopic findings. The evaluation includes the relationship, or lack thereof, between the dose and the presence, incidence, and severity of abnormalities, including gross lesions. Target organs, fertility, clinical abnormalities, reproductive and litter performance, body weight changes, mortality and any other toxic and developmental effects should also be assessed. Special attention should be given to sex-specific changes. The physico-chemical properties of the test chemical, and when available, TK data, including placental transfer and milk excretion, should be taken into consideration when evaluating the test results.

**Test report**

79. The test report should include the following information obtained in the present study from P, F<sub>1</sub> animals and F<sub>2</sub> animals (where relevant):

*Test chemical:*

- All relevant available information on the chemical, toxicokinetic and toxicodynamic properties of the test chemical;
- Identification data;
- Purity;

*Vehicle (if appropriate):*

- Justification for choice of vehicle if other than water;

*Test animals:*

- Species/strain used;
- Number, age and sex of animals;
- Source, housing conditions, diet, nesting materials, etc.;
- Individual weights of animals at the start of the test;
- Vaginal smear data for P females before initiation of treatment (if data are collected at that time);
- P generation pairing records indicating male and female partner of a mating and mating success;
- Litter of origin records for adult F<sub>1</sub> generation animals;

*Test conditions:*

- Rationale for dose level selection;
- Details of test chemical formulation/diet preparation, achieved concentrations;

**▼ M5**

- Stability and homogeneity of the preparation in the vehicle or carrier (e.g. diet, drinking water), in the blood and/or milk under the conditions of use and storage between uses;
- Details of the administration of the test chemical;
- Conversion from diet/drinking water test chemical concentration (ppm) to the achieved dose (mg/kg body weight/day), if applicable;
- Details of food and water quality (including diet composition, if available);
- Detailed description of the randomisation procedures to select pups for culling and to assign pups to test groups;
- Environmental conditions;
- List of study personnel, including professional training;

*Results (summary and individual data by sex and dose):*

- Food consumption, water consumption if available, food efficiency (body weight gain per gram of food consumed, except for the period of cohabitation and during lactation), and test chemical consumption (for dietary/drinking water administration) for P and F<sub>1</sub> animals;
- Absorption data (if available);
- Body weight data for P animals;
- Body weight data for the selected F<sub>1</sub> animals postweaning;
- Time of death during the study or whether animals survived to termination;
- Nature, severity and duration of clinical observations (whether reversible or not);
- Haematology, urinalysis and clinical chemistry data including TSH and T4;
- Phenotypic analysis of spleen cells (T-, B-, NK-cells);
- Bone marrow cellularity;
- Toxic response data;
- Number of P and F<sub>1</sub> females with normal or abnormal oestrous cycle and cycle duration;
- Time to mating (precoital interval, the number of days between pairing and mating);
- Toxic or other effects on reproduction, including numbers and percentages of animals that accomplished mating, pregnancy, parturition and lactation, of males inducing pregnancy, of females with signs of dystocia/prolonged or difficult parturition;
- Duration of pregnancy and, if available, parturition;
- Numbers of implantations, litter size and percentage of male pups;
- Number and percent of post-implantation loss, live births and stillbirths;

**▼ M5**

- Litter weight and pup weight data (males, females and combined), the number of runts if determined;
- Number of pups with grossly visible abnormalities;
- Toxic or other effects on offspring, postnatal growth, viability, etc.;
- Data on physical landmarks in pups and other postnatal developmental data;
- Data on sexual maturation of F<sub>1</sub> animals;
- Data on functional observations in pups and adults, as applicable;
- Body weight at sacrifice and absolute and relative organ weight data for the P and adult F<sub>1</sub> animals;
- Necropsy findings;
- Detailed description of all histopathological findings;
- Total cauda epididymal sperm number, percent progressively motile sperm, percent morphologically normal sperm, and percent of sperm with each identified abnormality for P and F<sub>1</sub> males;
- Numbers and maturational stages of follicles contained in the ovaries of P and F<sub>1</sub> females, where applicable;
- Enumeration of corpora lutea in the ovaries of F<sub>1</sub> females;
- Statistical treatment of results, where appropriate;

*Cohort 2 parameters:*

- Detailed description of the procedures used to standardise observations and procedures as well as operational definitions for scoring observations;
- List of all test procedures used, and justification for their use;
- Details of the behavioural/functional, neuropathological and morphometric procedures used, including information and details on automated devices;
- Procedures for calibrating and ensuring the equivalence of devices and the balancing of treatment groups in testing procedures;
- Short justification explaining any decisions involving professional judgment;
- Detailed description of all behavioural/functional, neuropathological and morphometric findings by sex and dose group, including both increases and decreases from controls;
- Brain weight;
- Any diagnoses derived from neurological signs and lesions, including naturally-occurring diseases or conditions;
- Images of exemplar findings;
- Low-power images to assess homology of sections used for morphometry;

**▼ M5**

- Statistical treatment of results, including statistical models used to analyse the data, and the results, regardless of whether they were significant or not;
- Relationship of any other toxic effects to a conclusion about the neurotoxic potential of the test chemical, by sex and dose group;
- Impact of any toxicokinetic information on the conclusions;
- Data supporting the reliability and sensitivity of the test method (i.e. positive and historical control data);
- Relationships, if any, between neuropathological and functional effects;
- NOAEL or benchmark dose for dams and offspring, by sex and dose group;
- Discussion of the overall interpretation of the data based on the results, including a conclusion of whether or not the chemical caused developmental neurotoxicity and the NOAEL;

*Cohort 3 parameters:*

- Serum IgM antibody titres (sensitisation to SRBC or KLH), or splenic IgM PFC units (sensitisation to SRBC);
- Performance of the TDAR method should be confirmed as part of the optimisation process by laboratory setting up the assay for the first time, and periodically (e.g. yearly) by all laboratories;
- Discussion of the overall interpretation of the data based on the results, including a conclusion of whether or not the chemical caused developmental immunotoxicity and the NOAEL;

*Discussion of results**Conclusions, including NOAEL values for parental and offspring effects*

All information not obtained during the study, but useful for the interpretation of the results (e.g. similarities of effects to any known neurotoxicants), should also be provided.

**Interpretation of Results**

80. An Extended One-Generation Reproductive Toxicity Study will provide information on the effects of repeated exposure to a chemical during all phases of the reproductive cycle, as necessary. In particular, the study provides information on the reproductive system, and on development, growth, survival, and functional endpoints of offspring up to PND 90.
81. Interpretation of the results of the study should take into account all available information on the chemical, including physico-chemical, TK and toxicodynamic properties, available relevant information on structural analogues, and results of previously-conducted toxicity studies with the test chemical (e.g. acute toxicity, toxicity after repeated application, mechanistic studies and studies assessing if there are substantial qualitative and quantitative species differences in *in vivo/in vitro* metabolic properties). Gross necropsy and organ weight results should be assessed in context with observations made in other repeat-dose studies, when feasible. Decreases in offspring growth might be considered in relationship to an influence of the test chemical on milk composition (29).

▼ **M5***Cohort 2 (Developmental neurotoxicity)*

82. Neurobehavioral and neuropathology results should be interpreted in the context of all findings, using a weight-of-evidence approach with expert judgment. Patterns of behavioural or morphological findings, if present, as well as evidence of dose-response should be discussed. The evaluation of developmental neurotoxicity, including human epidemiological studies or case reports, and experimental animal studies (e.g. toxicokinetic data, structure-activity information, data from other toxicity studies) should be included in this characterisation. Evaluation of data should include a discussion of both the biological and statistical significance. The evaluation should include the relationship, if any, between observed neuropathological and behavioural alterations. For guidance on the interpretation of developmental neurotoxicity results, refer to test method B.53 (35) and Tyl et al., 2008 (31).

*Cohort 3 (Developmental immunotoxicity)*

83. Suppression or enhancement of immune function as assessed by TDAR (T-cell dependent antibody response), should be evaluated in the context of all observations made. Significance of the outcome of TDAR may be supported by other effects on immunologically-related indicators (e.g. bone marrow cellularity, weight and histopathology of lymphoid tissues, lymphocyte subset distribution). Effects established by TDAR may be less meaningful in case of other toxicities observed at lower exposure concentrations.
84. OECD Guidance Document 43 should be consulted for aid in the interpretation of reproduction and neurotoxicity results (26).

**LITERATURE**

- (1) Cooper, R.L., J.C. Lamb, S.M. Barlow, K. Bentley, A.M. Brady, N. Doerr, D.L. Eisenbrandt, P.A. Fenner-Crisp, R.N. Hines, L.F.H. Irvine, C.A. Kimmel, H. Koeter, A.A. Li, S.L. Makris, L.P. Sheets, G.J.A. Speijers and K.E. Whitby (2006), 'A Tiered Approach to Life Stages Testing for Agricultural Chemical Safety Assessment', *Critical Reviews in Toxicology*, 36, 69-98.
- (2) Thigpen, J.E., K.D.R. Setchell, K.B. Ahlmark, J. Locklear, T. Spahr, G.F. Leviness, M.F. Goelz, J.K. Haseman, R.R. Newbold, and D.B. Forsythe (1999), 'Phytoestrogen Content of Purified Open and Closed Formula Laboratory Animal Diets', *Lab. Anim. Sci.*, 49, 530- 536.
- (3) Zoetis, T. and I. Walls (2003), *Principles and Practices for Direct Dosing of Pre-Weaning Mammals in Toxicity Testing and Research*, ILSI Press, Washington, DC.
- (4) Moser, V.C., I. Walls and T. Zoetis (2005), 'Direct Dosing of Prewaning Rodents in Toxicity Testing and Research: Deliberations of an ILSI RSI Expert Working Group', *International Journal of Toxicology*, 24, 87-94.
- (5) Conolly, R.B., B.D. Beck, and J.I. Goodman (1999), 'Stimulating Research to Improve the Scientific Basis of Risk Assessment', *Toxicological Sciences*, 49, 1-4.
- (6) Ulbrich, B. and A.K. Palmer (1995), 'Detection of Effects on Male Reproduction — a Literature Survey', *Journal of the American College of Toxicologists*, 14, 293-327.
- (7) Mangelsdorf, I., J. Buschmann and B. Orthen (2003), 'Some Aspects Relating to the Evaluation of the Effects of Chemicals on Male Fertility', *Regulatory Toxicology and Pharmacology*, 37, 356-369.

## ▼ M5

- (8) Sakai, T., M. Takahashi, K. Mitsumori, K. Yasuhara, K. Kawashima, H. Mayahara and Y. Ohno (2000). 'Collaborative work to evaluate toxicity on male reproductive organs by repeated dose studies in rats — overview of the studies', *Journal of Toxicological Sciences*, 25, 1-21.
- (9) Creasy, D.M. (2003), 'Evaluation of Testicular Toxicology: A Synopsis and Discussion of the Recommendations Proposed by the Society of Toxicologic Pathology', *Birth Defects Research, Part B*, 68, 408-415.
- (10) Goldman, J.M., A.S. Murr, A.R. Buckalew, J.M. Ferrell and R.L. Cooper (2007), 'The Rodent Estrous Cycle: Characterization of Vaginal Cytology and its Utility in Toxicological Studies', *Birth Defects Research, Part B*, 80 (2), 84-97.
- (11) Sadleir, R.M.F.S. (1979), 'Cycles and Seasons', in C.R. Auston and R.V. Short (eds.), *Reproduction in Mammals: I. Germ Cells and Fertilization*, Cambridge, New York.
- (12) Gallavan, R.H. Jr, J.F. Holson, D.G. Stump, J.F. Knapp and V.L. Reynolds (1999), 'Interpreting the Toxicologic Significance of Alterations in Anogenital Distance: Potential for Confounding Effects of Progeny Body Weights', *Reproductive Toxicology*, 13: 383-390.
- (13) Korenbrot, C.C., I.T. Huhtaniemi and R.I. Weiner (1977), 'Preputial Separation as an External Sign of Pubertal Development in the Male Rat', *Biological Reproduction*, 17, 298-303.
- (14) Ladics, G.S. (2007), 'Use of SRBC Antibody Responses for Immunotoxicity Testing', *Methods*, 41, 9-19.
- (15) Gore, E.R., J. Gower, E. Kurali, J.L. Sui, J. Bynum, D. Enmulat and D.J. Herzyk (2004), 'Primary Antibody Response to Keyhole Limpet Hemocyanin in Rat as a Model for Immunotoxicity Evaluation', *Toxicology*, 197, 23-35.
- (16) Gray, L.E., J. Ostby, J. Ferrell, G. Rehnberg, R. Linder, R. Cooper, J. Goldman, V. Slott and J. Laskey (1989), 'A Dose-Response Analysis of Methoxychlor-Induced Alterations of Reproductive Development and Function in the Rat', *Fundamental and Applied Toxicology*, 12, 92-108.
- (17) Robb, G.W., R.P. Amann and G.J. Killian (1978), 'Daily Sperm Production and Epididymal Sperm Reserves of Pubertal and Adult Rats', *Journal of Reproduction and Fertility*, 54, 103-107.
- (18) Klinefelter, G.R., L.E. Jr Gray and J.D. Suarez (1991), 'The Method of Sperm Collection Significantly Influences Sperm Motion Parameters Following Ethane Dimethanesulfonate Administration in the Rat'. *Reproductive Toxicology*, 5, 39-44.
- (19) Seed, J., R.E. Chapin, E.D. Clegg, L.A. Dostal, R.H. Foote, M.E. Hurtt, G.R. Klinefelter, S.L. Makris, S.D. Perreault, S. Schrader, D. Seyler, R. Sprando, K.A. Treinen, D.N. Veeramachaneni and L.D. Wise (1996), 'Methods for Assessing Sperm Motility, Morphology, and Counts in the Rat, Rabbit, and Dog: a Consensus Report', *Reproductive Toxicology*, 10, 237- 244.
- (20) Chapin, R.E., R.S. Filler, D. Gulati, J.J. Heindel, D.F. Katz, C.A. Mebus, F. Obasaju, S.D. Perreault, S.R. Russell and S. Schrader (1992), 'Methods for Assessing Rat Sperm Motility', *Reproductive Toxicology*, 6, 267-273.
- (21) Klinefelter, G.R., N.L. Roberts and J.D. Suarez (1992), 'Direct Effects of Ethane Dimethanesulphonate on Epididymal Function in Adult Rats: an *In Vitro* Demonstration', *Journal of Andrology*, 13, 409-421.
- (22) Slott, V.L., J.D. Suarez and S.D. Perreault (1991), 'Rat Sperm Motility Analysis: Methodologic Considerations', *Reproductive Toxicology*, 5, 449-458.



## ▼ M5

- (23) Slott, V.L., and S.D. Perreault (1993), 'Computer-Assisted Sperm Analysis of Rodent Epididymal Sperm Motility Using the Hamilton-Thorn Motility Analyzer', *Methods in Toxicology, Part A*, Academic, Orlando, Florida. pp. 319-333.
- (24) Toth, G.P., J.A. Stober, E.J. Read, H. Zenick and M.K. Smith (1989), "The Automated Analysis of Rat Sperm Motility Following Subchronic Epichlorohydrin Administration: Methodologic and Statistical Considerations", *Journal of Andrology*, 10, 401-415.
- (25) Linder, R.E., L.F. Strader, V.L. Slott and J.D. Suarez (1992), 'Endpoints of Spermatotoxicity in the Rat After Short Duration Exposures to Fourteen Reproductive Toxicants', *Reproductive Toxicology*, 6, 491-505.
- (26) OECD (2008), *Guidance Document on Mammalian Reproductive Toxicity Testing and Assessment*, Series on Testing and Assessment, No 43, ENV/JM/MONO(2008)16, OECD, Paris.
- (27) Working, P.K., M. Hurtt (1987), 'Computerized Videomicrographic Analysis of Rat Sperm Motility', *Journal of Andrology*, 8, 330-337.
- (28) Bolin, B., R. Garman, K. Jensen, G. Krinke, B. Stuart, and an ad Hoc Working Group of the STP Scientific and Regulatory Policy Committee (2006), 'A "Best Practices" Approach to Neuropathologic Assessment in Developmental Neurotoxicity Testing — for Today', *Toxicological Pathology*, 34, 296-313.
- (29) Stütz, N., B. Bongiovanni, M. Rassetto, A. Ferri, A.M. Evangelista de Duffard, and R. Duffard (2006), 'Detection of 2,4-dichlorophenoxyacetic Acid in Rat Milk of Dams Exposed During Lactation and Milk Analysis of their Major Components', *Food Chemicals Toxicology*, 44, 8-16.
- (30) Thigpen, JE, K.D.R. Setchell, J.K. Haseman, H.E. Saunders, G.F. Caviness, G.E. Kissling, M.G. Grant and D.B. Forsythe (2007), 'Variations in Phytoestrogen Content between Different Mill Dates of the Same Diet Produces Significant Differences in the Time of Vaginal Opening in CD-1 Mice and F344 Rats but not in CD Sprague Dawley Rats', *Environmental health perspectives*, 115(12), 1717-1726.
- (31) Tyl, R.W., K. Crofton, A. Moretto, V. Moser, L.P. Sheets and T.J. Sobotka (2008), 'Identification and Interpretation of Developmental Neurotoxicity Effects: a Report from the ILSI Research Foundation/Risk Science Institute Expert Working Group on Neurodevelopmental Endpoints', *Neurotoxicology and Teratology*, 30: 349-381.
- (32) OECD (1996), *Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test*, OECD Guideline for Testing of Chemicals, No 422, OECD, Paris.
- (33) Chapter B.43 of this Annex, Neurotoxicity Study in Rodents
- (34) OECD (2000), *Guidance Document on the recognition, assessment, and use of clinical signs as humane endpoints for experimental animals used in safety evaluations*, Series on Testing and Assessment, No 19, ENV/JM/MONO(2000)7, OECD, Paris.
- (35) Chapter B.53 of this Annex, Developmental Neurotoxicity Study
- (36) Chapter B.54 of this Annex, Uterotrophic Bioassay in Rodents: A short-term Screening Test for Oestrogenic Properties
- (37) Chapter B.55 of this Annex, Hershberger Bioassay in Rats: A Short-term Screening Assay for (Anti)Androgenic Properties

**▼M5**

- (38) OECD (2009), Guidance Document for Histologic Evaluation of Endocrine and Reproductive Test in Rodents, Series on Testing and Assessment, No 106, OECD, Paris.
- (39) OECD (2011), Guidance Document on the Current Implementation of Internal Triggers in the Extended One Generation Reproductive Toxicity Study in the United States and Canada, Series on Testing and Assessment, No 117, ENV/JM/MONO(2011)21, OECD, Paris.
- (40) OECD (2013), Guidance Document supporting TG 443: Extended One Generation Reproductive Toxicity Study, Series on Testing and Assessment, No 151, OECD, Paris.

▼ **M5***Appendix 1***Measures and observations included in the functional observational battery (Cohort 2A)**

Home Cage & Open Field	Manipulative	Physiologic
Posture	Ease of removal	Temperature
Involuntary Clonic & Tonic	Ease of handling	Body weight
Palpebral Closure	Muscle Tone	Pupil response
Piloerection	Approach Response	Pupil size
Salivation	Touch Response	
Lacrimation	Auditory Response	
Vocalisations	Tail Pinch Response	
Rearing	Righting Response	
Gait Abnormalities	Landing Foot Splay	
Arousal	Forelimb Grip Strength	
Stereotypy	Hindlimb Grip Strength	
Bizarre Behaviour		
Stains		
Respiratory Abnormalities		

▼ M5

*Appendix 2*

DEFINITIONS:

**Chemical:** A substance or a mixture.

**Test Chemical:** Any substance or mixture tested using this test method.

▼ **M5****B.57 H295R STEROIDOGENESIS ASSAY**

## INTRODUCTION

1. This test method is equivalent to OECD Test Guideline (TG) 456 (2011). The OECD initiated a high-priority activity in 1998 to revise existing, and to develop new, test guidelines for the screening and testing of potential endocrine disrupting chemicals. The 2002 OECD Conceptual Framework for Testing and Assessment of Endocrine Disrupting Chemicals comprises five levels, each level corresponding to a different level of biological complexity (1). The *in vitro* H295R Steroidogenesis Assay (H295R) described in this test method utilises a human adeno-carcinoma cell line (NCI-H295R cells) and constitutes a level 2 '*in vitro* assay, providing mechanistic data', to be used for screening and prioritisation purposes. Development and standardisation of the assay as a screen for chemical effects on steroidogenesis, specifically the production of 17 $\beta$ -oestradiol (E2) and testosterone (T), was carried out in a multi-step process. The H295R assay has been optimised and validated (2) (3) (4) (5).
2. The objective of the H295R Steroidogenesis Assay is to detect chemicals that affect production of E2 and T. The H295R assay is intended to identify xenobiotics that have as their target site(s) the endogenous components that comprise the intracellular biochemical pathway beginning with the sequence of reactions from cholesterol to the production of E2 and/or T. The H295R assay is not intended to identify chemicals that affect steroidogenesis due to effects on the hypothalamic-pituitary-gonadal (HPG) axis. The goal of the assay is to provide a YES/NO answer with regard to the potential of a chemical to induce or inhibit the production of T and E2; however, quantitative results may be obtained in some cases (see paragraphs 53 and 54). The results of the assay are expressed as relative changes in hormone production compared with the solvent controls (SCs). The assay does not aim to provide specific mechanistic information concerning the interaction of the test chemical with the endocrine system. Research has been conducted using the cell line to identify effects on specific enzymes and intermediate hormones such as progesterone (2).
3. Definitions and abbreviations used in this test method are described in the Appendix. A detailed protocol including instructions on how to prepare solutions, cultivate cells and perform various aspects of the test is available as Appendix I-III to the OECD document 'Multi-Laboratory Validation of the H295R Steroidogenesis Assay to Identify Modulators of Testosterone and Estradiol Production' (4).

## INITIAL CONSIDERATIONS AND LIMITATIONS

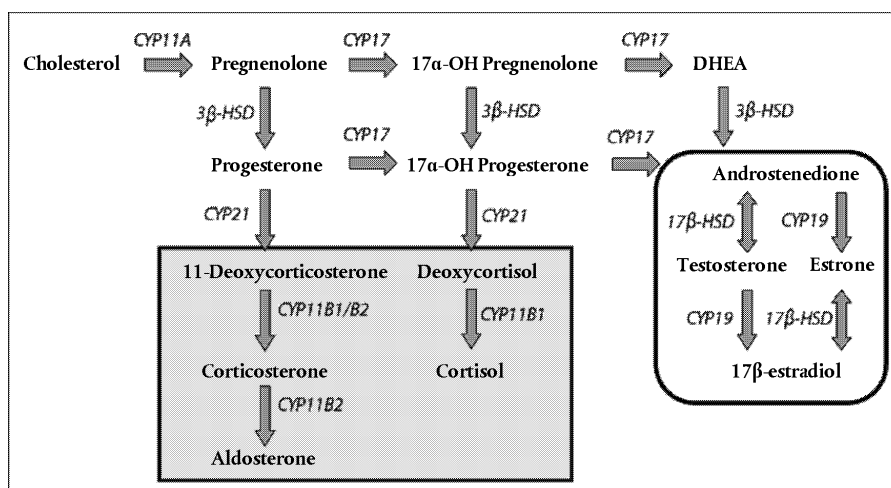
4. Five different enzymes catalysing six different reactions are involved in sex steroid hormone biosynthesis. Enzymatic conversion of cholesterol to pregnenolone by the cytochrome P450 (CYP) cholesterol side-chain cleavage enzyme (CYP11A) constitutes the initial step in a series of biochemical reactions that culminate in synthesis of steroid end-products. Depending upon the order of the next two reactions, the steroidogenic pathway splits into two paths, the  $\Delta^5$ -hydroxysteroid pathway and  $\Delta^4$ -ketosteroid pathway, which converge in the production of androstenedione (Figure 1).
5. Androstenedione is converted to testosterone (T) by 17 $\beta$ -hydroxysteroid dehydrogenase (17 $\beta$ -HSD). Testosterone is both an intermediate and end-hormone product. In the male, T can be converted to dihydrotestosterone (DHT) by 5 $\alpha$ -reductase, which is found in the cellular membranes, nuclear envelope, and endoplasmic reticulum of target tissues of androgenic action such as prostate and seminal vesicles. DHT is significantly more potent as an androgen than T and is also considered an end-product hormone. The H295R assay does not measure DHT (see paragraph 10).

## ▼ M5

6. The enzyme in the steroidogenic pathway which converts androgenic chemicals into oestrogenic chemicals is aromatase (CYP19). CYP19 converts T into 17 $\beta$ -oestradiol (E2) and androstenedione into oestrone. E2 and T are considered end-product hormones of the steroidogenic pathway.
7. The specificity of the lyase activity of CYP17 differs for the intermediate substrates among species. In the human, the enzyme favours substrates of the  $\Delta^5$ -hydroxysteroid pathway (pregnenolone), whereas substrates in the  $\Delta^4$ -ketosteroid pathway (progesterone) are favoured in the rat (19). Such differences in the CYP17 lyase activity may explain some species-dependent differences in response to chemicals that alter steroidogenesis *in vivo* (6). The H295 cells have been shown to most closely reflect human adult adrenal enzyme expression and steroid production pattern (20), but are known to express enzymes for both the  $\Delta^5$ -hydroxysteroid and  $\Delta^4$ -ketosteroid pathways for androgen synthesis (7) (11) (13) (15).

Figure 1

## Steroidogenic pathway in H295R cells.



## Note:

Enzymes are in italics, hormones are bolded and arrows indicate the direction of synthesis. Gray background indicates corticosteroid pathways/products. Sex steroid pathways/products are circled. CYP = cytochrome P450; HSD = hydroxysteroid dehydrogenase; DHEA = dehydroepiandrosterone.

8. The human H295R adreno-carcinoma cell line is a useful *in vitro* model for the investigation of effects on steroid hormone synthesis (2) (7) (8) (9) (10). The H295R cell line expresses genes that encode for all the key enzymes for steroidogenesis noted above (11) (15) (Figure 1). This is a unique property because *in vivo* expression of these genes is tissue and developmental stage-specific with typically no one tissue or one developmental stage expressing all of the genes involved in steroidogenesis (2). H295R cells have physiological characteristics of zonally undifferentiated human foetal adrenal cells (11). The cells represent a unique *in vitro* system in that they have the ability to produce all of the steroid hormones found in the adult adrenal cortex and the gonads, allowing testing for effects on both corticosteroid synthesis and the production of sex steroid hormones such as androgens and oestrogens, although the assay was validated only to detect T and E2. Changes recorded by the test system in the form of alteration in the

**▼ M5**

production of T and E2 can be the result of a multitude of different interactions of test chemicals with steroidogenic functions that are expressed by the H295R cells. These include modulation of the expression, synthesis or function of enzymes involved in the production, transformation, or elimination of steroid hormones (12) (13) (14). Inhibition of hormone production can be due to direct competitive binding to an enzyme in the pathway, impact on co-factors such as NADPH (Nicotinamide Adenine Dinucleotide Phosphate) and cAMP (cyclic Adenosine Monophosphate), and/or increase in steroid metabolism or suppression of gene expression of certain enzymes in the steroidogenesis pathway. While inhibition can be a function of both direct or indirect processes involved with hormone production, induction is typically of an indirect nature, such as by affecting co-factors such as NADPH and cAMP (as in the case of forskolin), decreasing steroid metabolism (13), and or up-regulating steroidogenic gene expression.

9. The H295R assay has several advantages:
  - It allows for the detection of both increases and decreases in the production of both T and E2;
  - It permits the direct assessment of the potential impact of a chemical on cell viability/cytotoxicity. This is an important feature as it allows for the discrimination between effects that are due to cytotoxicity from those due to the direct interaction of chemicals with steroidogenic pathways, which is not possible in tissue explants systems that consist of multiple cell types of varying sensitivities and functionalities;
  - It does not require the use of animals;
  - The H295R cell line is commercially available.
10. The principle limitations of the assay are as follows:
  - Its metabolic capability is unknown but probably quite limited; therefore, chemicals that need to be metabolically activated will probably be missed in this assay.
  - Being derived from adrenal tissue, the H295R possesses the enzymes capable of producing the gluco-, and mineral-corticoids as well as the sex hormones; therefore, effects on the production of gluco-, and mineral corticoids could influence the levels of T and E2 observed in the assay.
  - It does not measure DHT and, therefore, would not be expected to detect chemicals that inhibit 5 $\alpha$ -reductase in which case the Hersherberger assay (16) can be used.
  - The H295R assay will not detect chemicals that interfere with steroidogenesis by affecting the hypothalamic-pituitary-gonadal axis (HPG) axis as this can only be studied in intact animals.

**PRINCIPLE OF THE TEST**

11. The purpose of the assay is the detection of chemicals that affect T and E2 production. T is also an intermediate in the pathway to produce E2. The assay can detect chemicals that typically inhibit or induce the enzymes of the steroidogenesis pathway.

**▼ M5**

12. The assay is usually performed under standard cell culture conditions in 24-well culture plates. Alternatively, other plate sizes can be used for conducting the assay; however, seeding and experimental conditions should be adjusted accordingly to maintain adherence to the performance criteria.
  
13. After an acclimation period of 24 h in multi-well plates, cells are exposed for 48 h to seven concentrations of the test chemical in at least triplicate. Solvent and a known inhibitor and inducer of hormone production are run at a fixed concentration as negative and positive controls. At the end of the exposure period, the medium is removed from each well. Cell viability in each well is analysed immediately after removal of medium. Concentrations of hormones in the medium can be measured using a variety of methods including commercially available hormone measurement kits and/or instrumental techniques such as liquid chromatography-mass spectrometry (LC-MS). Data are expressed as fold change relative to the solvent control and the Lowest-Observed-Effect-Concentration (LOEC). If the assay is negative, the highest concentration tested is reported as the No-Observed-Effect-Concentration (NOEC). Conclusions regarding the ability of a chemical to affect steroidogenesis should be based on at least two independent test runs. The first test run may function as a range finding run with subsequent adjustment of concentrations for runs 2 and 3, if applicable, if solubility or cytotoxicity problems are encountered or the activity of the chemical seems to be at the end of the range of concentrations tested.

**CULTURE PROCEDURE****Cell Line**

14. The NCI-H295R cells are commercially available from the American Type Culture Collections (ATCC) upon signing a Material Transfer Agreement (MTA) <sup>(1)</sup>.

**Introduction**

15. Due to changes in the E2 producing capacity of the cells with increasing age/passages (2), cells should be cultured following a specific protocol before they are used and the number of passages since the cells were defrosted as well as the passage number at which the cells were frozen and placed in liquid nitrogen storage should be noted. The first number indicates the actual cell passage number and the second number describes the passage number at which the cells were frozen and placed in storage. For example, cells that were frozen after passage five and defrosted and then were split three times (4 passages counting the freshly thawed cells as passage 1) after they were cultured again would be labelled passage 4.5. An example of a numbering scheme is illustrated in Appendix I to the validation report (4).
  
16. Stock medium is used as the base for the supplemented and freezing media. Supplemented medium is a necessary component for culturing cells. Freezing medium is specifically designed to allow for impact-free freezing of cells for long-term storage. Prior to use, Nu-serum (or a comparable serum of equal properties that has been demonstrated to produce data that meets the test performance and Quality Control (QC) requirements), which is a constituent of supplemented media, should be analysed for background T and E2 concentrations. The preparation of these solutions is described in Appendix II to the validation report (4).

<sup>(1)</sup> ATCC CRL-2128; ATCC, Manassas, VA, USA, [<http://www.lgcstandards-atcc.org/>].



**▼ M5**

17. After initiation of an H295R cell culture from an original ATCC batch, cells should be grown for five passages (i.e. the cells are split 4 times). Passage five cells are then frozen in liquid nitrogen for storage. Prior to freezing the cells, a sample of the previous passage four cells is run in a QC plate (See paragraph 36 and 37) to verify whether the basal production of hormones and the response to positive control chemicals meet the assay quality control criteria as defined in Table 5.
  
18. H295R cells need to be cultured, frozen and stored in liquid nitrogen to make sure that there are always cells of the appropriate passage/age available for culture and use. The maximum number of passages after taking a new <sup>(1)</sup> or frozen <sup>(2)</sup> batch of cells into culture that is acceptable for use in the H295R assay should not exceed 10. For example, acceptable passages for cultures of cells from a batch frozen at passage 5 would be 4.5 through 10.5. For cells started from these frozen batches, the procedure described in paragraph 19 should be followed. These cells should be cultured for at least four (4) additional passages (passage 4.5) prior to their use in testing.

**Starting Cells from the Frozen Stock**

19. The procedure for starting the cells from frozen stock is to be used when a new batch of cells is removed from liquid nitrogen storage for the purpose of culture and testing. Details for this procedure are set forth in Appendix III to the validation report (4). Cells are removed from liquid nitrogen storage, thawed rapidly, placed in supplemented medium in a centrifuge tube, centrifuged at room temperature, re-suspended in supplemented medium, and transferred to a culture flask. The medium should be changed the following day. The H295R cells are cultivated in an incubator at 37 °C with 5 % CO<sub>2</sub> in air atmosphere and the medium is renewed 2-3 times per week. When the cells are approximately 85-90 % confluent, they should be split. Splitting of the cells is necessary to ensure the health and growth of the cells and to maintain cells for performing bioassays. The cells are rinsed three times with phosphate-buffered saline (PBS, without Ca<sup>2+</sup> Mg<sup>2+</sup>) and freed from the culture flask by the addition of an appropriate detachment enzyme, e.g. trypsin, in PBS (without Ca<sup>2+</sup> Mg<sup>2+</sup>). Immediately after the cells detach from the culture flask, the enzyme action should be stopped with the addition of supplemented medium at a ratio of 3× the volume used for the enzyme treatment. Cells are placed into a centrifuge tube, centrifuged at room temperature, the supernatant is removed and the pellet of cells is re-suspended in supplemented medium. The appropriate amount of cell solution is placed in the new culture flask. The amount of cell solution should be adjusted so that the cells are confluent within 5-7 days. The recommended sub-cultivation ratio is 1:3 to 1:4. The plate should be carefully labelled. The cells are now ready to be used in the assay and excess cells should be frozen in liquid nitrogen as described in paragraph 20.

<sup>(1)</sup> 'New batch' refers to a fresh batch of cells received from ATCC.

<sup>(2)</sup> 'Frozen batch' refers to cells that have been previously cultured and then frozen at a laboratory other than ATCC.

**▼ M5****Freezing H295R Cells (preparing cells for liquid nitrogen storage)**

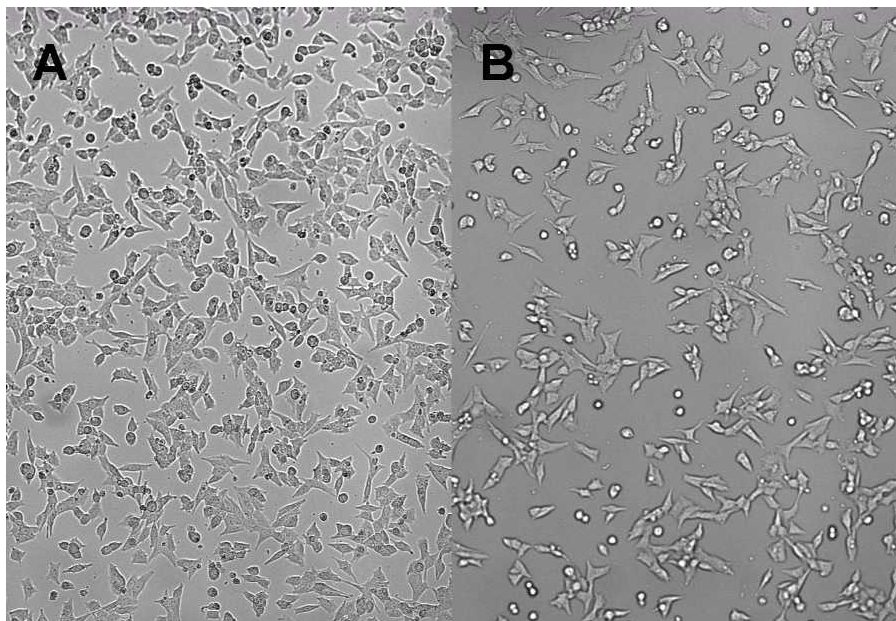
20. To prepare H295R cells for freezing, the procedure described above for splitting cells should be followed until the step for re-suspending the pellet of cells in the bottom of the centrifuge tube. Here, the pellet of cells is re-suspended in freezing medium. The solution is transferred to a cryogenic vial, labelled appropriately, and frozen at  $-80^{\circ}\text{C}$  for 24 hours after which the cryogenic vial is transferred to liquid nitrogen for storage. Details for this procedure are set forth in Appendix III to the validation report (4).

**Plating and Pre-incubation of Cells for Testing**

21. The number of 24-well plates, prepared as outlined in paragraph 19, that will be needed depends on the number of chemicals to be tested and the confluency of the cells in the culture dishes. As a general rule, one culture flask ( $75\text{ cm}^2$ ) of 80-90 % confluent cells will supply sufficient cells for one to 1,5 (24-well) plates at a target density of 200 000 to 300 000 cells per ml of medium resulting in approximately 50-60 % confluency in the wells at 24 hours (Figure 2). This is typically the optimal cell density for hormone production in the assay. At higher densities, T as well as E2 production patterns are altered. Before conducting the assay the first time, it is recommended that different seeding densities between 200 000 and 300 000 cells per ml be tested, and the density resulting in 50-60 % confluency in the well at 24 hours be selected for further experiments.

*Figure 2*

**Photomicrograph of H295R cells at a seeding density of 50 % in a 24 well culture plate at 24 hours taken at the edge (A) and centre (B) of a well**



22. The medium is pipetted off the culture flask, and the cells are rinsed 3 times with sterile PBS (without  $\text{Ca}^{2+}$  $\text{Mg}^{2+}$ ). An enzyme solution (in PBS) is added to detach the cells from the culture flask. Following an appropriate time for detachment of the cells, the enzyme action should be stopped with the addition of supplemented medium at a ratio of  $3 \times$  the volume used for the enzyme treatment. Cells are placed into a centrifuge tube, centrifuged at

**▼ M5**

room temperature, the supernatant is removed, and the pellet of cells is re-suspended in supplemented medium. The cell density is calculated using e.g. a haemocytometer or cell counter. The cell solution should be diluted to the desired plating density and thoroughly mixed to assure homogenous cell density. The cells should be plated with 1 ml of the cell solution/well and the plates and wells labelled. The seeded plates are incubated at 37 °C under 5 % CO<sub>2</sub> in air atmosphere for 24 hours to allow the cells to attach to the wells.

**QUALITY CONTROL REQUIREMENTS**

23. It is critical that exact volumes of solutions and samples are delivered into the wells during dosing because these volumes determine the concentrations used in the calculations of assay results.
24. Prior to the initiation of cell culture and any subsequent testing, each laboratory should demonstrate the sensitivity of its hormone measurement system (paragraphs 29-31).
25. If antibody-based hormone measurement assays are to be used, the chemicals to be tested should be analysed for their potential to interfere with the measurement system used to quantify T and E2 as outlined in paragraph 32 prior to initiating testing.
26. DMSO is the recommended solvent for the assay. If an alternative solvent is utilised, the following should be determined:
  - The solubility of the test chemical, forskolin and prochloraz in the solvent; and
  - The cytotoxicity as a function of the concentration of solvent.

It is recommended that the maximum allowable solvent concentration should not exceed a 10 × dilution of the least cytotoxic concentration of the solvent.

27. Prior to conducting testing for the first time, the laboratory should conduct a qualifying experiment demonstrating that the laboratory is capable of maintaining and achieving appropriate cell culture and experimental conditions required for chemical testing as described in paragraphs 33-35.
28. When initiating testing using a new batch, a control plate should be run before using a new batch of cells to evaluate the performance of the cells as described in paragraphs 36 and 37.

**Performance of the Hormone Measurement System**

*Method sensitivity, accuracy, precision and cross-reactivity with sample matrix*

29. Each laboratory may use a hormone measurement system of its choice for the analysis of the production of T and E2 by H295R cells so long as it meets performance criteria, including the Limit of Quantification (LOQ). Nominally these are 100 pg/ml for T and 10 pg/ml for E2, which are based on the basal hormone levels observed in the validation studies. However, greater or lower levels may be appropriate depending upon the basal hormone levels achieved in the performing laboratory. Prior to initiation of QC plate and test runs, the laboratory should demonstrate that the

**▼ M5**

hormone assay to be used can measure hormone concentrations in supplemented medium with sufficient accuracy and precision to meet the QC criteria specified in Tables 1 and 5 by analysing supplemented medium spiked with an internal hormone control. Supplemented medium should be spiked with at least three concentrations of each hormone (e.g. 100, 500 and 2 500 pg/ml of T; 10, 50 and 250 pg/ml of E2; or the lowest possible concentrations based upon the detection limits of the chosen hormone measurement system can be used for the lowest spike concentrations for T and E2) and analysed. Measured hormone concentrations of non-extracted samples should be within 30 % of nominal concentrations, and variation between replicate measurements of the same sample should not exceed 25 % (see also Table 8 for additional QC criteria). If these QC criteria are fulfilled it is assumed that the selected hormone measurement assay is sufficiently accurate, precise and does not cross-react with components in the medium (sample matrix) such that a significant influence on the outcome of the assay would be expected. In this case, no extraction of samples prior to measurement of hormones is required.

30. In the case that the QC criteria in Tables 1 and 8 are not fulfilled, a significant matrix effect may be occurring, and an experiment with extracted spiked medium should be conducted. An example of an extraction procedure is described in Appendix II to the validation report (4). Measurements of the hormone concentrations in the extracted samples should be made in triplicate.<sup>(1)</sup> If it can be shown that after extraction the components of the medium do not interfere with the hormone detection method as defined by the QC criteria, all further experiments should be conducted using extracted samples. If the QC criteria cannot be met after extraction, the utilised hormone measurement system is not suitable for the purpose of the H295R Steroidogenesis Assay, and an alternative hormone detection method should be used.

*Standard curve*

31. The hormone concentrations of the solvent controls (SC) should be within the linear portion of the standard curve. Preferably, the SC values should fall close to the centre of the linear portion to ensure that induction and inhibition of hormone synthesis can be measured. Dilutions of medium (or extracts) to be measured are to be selected accordingly. The linear relationship is to be determined by a suitable statistical approach.

*Chemical interference test*

32. If antibody-based assays such as Enzyme-Linked Immunosorbent Assays (ELISAs) and Radio-Immuno Assays (RIAs) are going to be used to measure hormones, each chemical should be tested for potential interference with the hormone measurement system to be utilised prior to initiation of the actual testing of chemicals (Appendix III to the validation report (4)) because some chemicals can interfere with these tests (17). If interference occurs that is  $\geq 20$  % of basal hormone production for T and/or E2 as determined by hormone analysis, the Chemical Hormone Assay Interference Test (such as described in Appendix III to the validation report (4) section 5.0) should be run on all test chemical stock solution dilutions to identify the threshold dose at which significant ( $\geq 20$  %) interference

<sup>(1)</sup> Note: If extraction is required, three replicate measurements are made for each extract. Each sample will be extracted only once.

▼ **M5**

occurs. If interference is less than 30 %, results may be corrected for the interference. If interference exceeds 30 %, the data are invalid and the data at these concentrations should be discarded. If significant interference of a test chemical with a hormone measurement system occurs at more than one non-cytotoxic concentration, a different hormone measurement system should be used. In order to avoid interference from contaminating chemicals it is recommended that hormones are extracted from the medium using suitable solvent, possible methods can be found in the validation report (4).

Table 1

**Performance criteria for hormone measurement systems**

Parameter	Criterion
Measurement Method Sensitivity	Limit of Quantification (LOQ) T: 100 pg/ml; E2: 10 pg/ml <sup>(a)</sup>
Hormone Extraction Efficiency (only when extraction is needed)	The average recovery rates (based on triplicate measures) for the spiked amounts of hormone should not deviate more than 30 % from amount that was added.
Chemical Interference (only antibody based systems)	No substantial ( $\geq 30$ % of basal hormone production of the respective hormone) cross-reactivity with any of the hormones produced by the cells should occur <sup>(b)</sup> <sup>(c)</sup>

<sup>(a)</sup> *Note:* Method measurement limits are based on the basal hormone production values provided in Table 5, and are performance based. If greater basal hormone production can be achieved the limit can be greater.

<sup>(b)</sup> Some T and E2 antibodies may cross-react with androstenedione and oestrone, respectively, at a greater percentage. In such cases it is not possible to accurately determine effects on 17 $\beta$ -HSD. However, the data can still provide useful information regarding the effects on oestrogen or androgen production in general. In such cases data should be expressed as androgen/oestrogen responses rather than E2 and T.

<sup>(c)</sup> These include: cholesterol, pregnenolone, progesterone, 11-deoxycorticosterone, corticosterone, aldosterone, 17 $\alpha$ -pregnenolone, 17 $\alpha$ -progesterone, deoxycortisol, cortisol, DHEA, androstenedione, oestrone.

**Laboratory Proficiency Test**

33. Before testing unknown chemicals, a laboratory should demonstrate that it is capable of achieving and maintaining appropriate cell culture and test conditions required for the successful conduct of the assay by running the laboratory proficiency test. As the performance of an assay is directly linked to the laboratory personnel conducting the assay, these procedures should be partly repeated if a change in laboratory personnel occurs.
34. This proficiency test will be conducted under the same conditions listed in paragraphs 38 through 40 by exposing cells to 7 increasing concentrations of strong, moderate and weak inducers and inhibitors as well as a negative chemical (see Table 2). Specifically, chemicals to be tested include the strong inducer forskolin (CAS No 66575-29-9); the strong inhibitor prochloraz (CAS No 67747-09-5); the moderate inducer atrazine

▼ **M5**

(CAS No 1912-24-9); the moderate inhibitor aminoglutethimide (CAS No 125-84-8); the weak inducer (E2 production) and weak inhibitor (T production) bisphenol A (CAS No 80-05-7); and the negative chemical human chorionic gonadotropin (HCG) (CAS No 9002-61-3) as shown in Table 2. Separate plates are run for all chemicals using the format as shown in Table 6. One QC plate (Table 4, paragraphs 36-37) should be included with each daily run for the proficiency chemicals.

Table 2

**Proficiency chemicals and exposure concentrations**

Proficiency chemical	Test Concentrations [ $\mu$ M]
Prochloraz	0 <sup>(a)</sup> , 0,01, 0,03, 0,1, 0,3, 1, 3, 10
Forskolin	0 <sup>(a)</sup> , 0,03, 0,1, 0,3, 1, 3, 10, 30
Atrazine	0 <sup>(a)</sup> , 0,03, 0,1, 1, 3, 10, 30, 100
Aminoglutethimide	0 <sup>(a)</sup> , 0,03, 0,1, 1, 3, 10, 30, 100
Bisphenol A	0 <sup>(a)</sup> , 0,03, 0,1, 1, 3, 10, 30, 100
HCG	0 <sup>(a)</sup> , 0,03, 0,1, 1, 3, 10, 30, 100

<sup>(a)</sup> Solvent (DMSO) control (0), 1  $\mu$ l DMSO/well.

Exposure of H295R to proficiency chemicals should be conducted in 24 well plates during the laboratory proficiency test. Dosing is in  $\mu$ M for all test chemical doses. Doses should be administered in DMSO at 0,1 % v/v per well. All test concentrations should be tested in triplicate wells (Table 6). Separate plates are run for each chemical. One QC plate is included with each daily run.

35. Cell viability and hormone analyses should be conducted as provided in paragraphs 42 through 46. The threshold value (lowest observed effect concentration, LOEC) and classification decision should be reported and compared with the values in Table 3. The data are considered acceptable if they meet the LOEC and decision classification in Table 3.

Table 3

**Threshold values (LOECs) and decision classifications for Proficiency Chemicals**

	CAS No	LOEC [ $\mu$ M]		Decision Classification	
		T	E2	T	E2
Prochloraz	67747-09-5	$\leq 0,1$	$\leq 1,0$	+ <sup>(a)</sup> (Inhibition)	+ (Inhibition)
Forskolin	66575-29-9	$\leq 10$	$\leq 0,1$	+ (Induction)	+ (Induction)
Atrazine	1912-24-9	$\leq 100$	$\leq 10$	+ (Induction)	+ (Induction)
Aminoglu-tethimide	125-84-8	$\leq 100$	$\leq 100$	+ (Inhibition)	+ (Inhibition)

## ▼ M5

	CAS No	LOEC [ $\mu$ M]		Decision Classification	
		T	E2	T	E2
Bisphenol A	80-05-7	$\leq 10$	$\leq 10$	+ (Inhibition)	+ (Induction)
HCG	9002-61-3	n/a	n/a	Negative	Negative

(<sup>a</sup>) +, positive

n/a: not applicable as no changes should occur after exposure to non-cytotoxic concentrations of negative control.

#### Quality Control Plate

36. The quality control (QC) plate is used to verify the performance of the H295R cells under standard culture conditions, and to establish a historical database for hormone concentrations in solvent controls, positive and negative controls, as well as other QC measures over time.

— H295R cell performance should be assessed using a QC plate for each new ATCC batch or after using a previously frozen stock of cells for the first time unless the laboratory proficiency test (paragraphs 32-34) has been run with that batch of cells.

— A QC plate provides a complete assessment of the assay conditions (e.g. cell viability, solvent controls, negative and positive controls, as well as intra- and inter-assay variability) when testing chemicals and should be part of each test run.

37. The QC test is conducted in a 24-well plate and follows the same incubation, dosing, cell viability/cytotoxicity, hormone extraction and hormone analysis procedures described in paragraphs 38 through 46 for testing chemicals. The QC plate contains blanks, solvent controls, and two concentrations of a known inducer (forskolin, 1, 10  $\mu$ M) and inhibitor (prochloraz, 0,1, 1  $\mu$ M) of E2 and T synthesis. In addition, MeOH is used in select wells as a positive control for the viability/cytotoxicity assay. A detailed description of the plate layout is provided in Table 4. The criteria to be met on the QC plate are listed in Table 5. The minimum basal hormone production for T and E2 should be met in both the solvent control and blank wells.

Table 4

**Quality control plate layout for testing performance of unexposed H295R cells and cells exposed to known inhibitors (PRO = prochloraz) and stimulators (FOR = forskolin) of E2 and T production. After termination of the exposure experiment and removal of medium, a 70 % methanol solution will be added to all MeOH wells to serve as a positive control for cytotoxicity (see cytotoxicity assay in Appendix III to the validation report (4))**

	1	2	3	4	5	6
A	Blank ( <sup>a</sup> )	Blank ( <sup>a</sup> )	Blank ( <sup>a</sup> )	Blank ( <sup>a</sup> ) (+ MeOH) ( <sup>b</sup> )	Blank ( <sup>a</sup> ) (+ MeOH) ( <sup>b</sup> )	Blank ( <sup>a</sup> ) (+ MeOH) ( <sup>b</sup> )
B	DMSO ( <sup>c</sup> ) 1 $\mu$ l	DMSO ( <sup>c</sup> ) 1 $\mu$ l	DMSO ( <sup>c</sup> ) 1 $\mu$ l	DMSO ( <sup>c</sup> ) 1 $\mu$ l (+ MeOH) ( <sup>b</sup> )	DMSO ( <sup>c</sup> ) 1 $\mu$ l (+ MeOH) ( <sup>b</sup> )	DMSO ( <sup>c</sup> ) 1 $\mu$ l (+ MeOH) ( <sup>b</sup> )

## ▼ M5

	1	2	3	4	5	6
C	FOR 1 μM	FOR 1 μM	FOR 1 μM	PRO 0,1 μM	PRO 0,1 μM	PRO 0,1 μM
D	FOR 10 μM	FOR 10 μM	FOR 10 μM	PRO 1 μM	PRO 1 μM	PRO 1 μM

(a) Cells in Blank wells receive medium only (i.e. no solvent).

(b) Methanol (MeOH) will be added **after** the exposure is terminated and the medium is removed from these wells.

(c) DMSO solvent control (1 μl/well).

Table 5

## Performance criteria for the Quality Control Plate

	T	E2
Basal Production of hormone in the solvent control (SC)	≥ 5 times the LOQ	≥ 2,5 times the LOQ
Induction (10 μM forskolin)	≥ 1,5 times the SC	≥ 7,5 times the SC
Inhibition (1 μM prochloraz)	≤ 0,5 times the SC	≤ 0,5 times the SC

## CHEMICAL EXPOSURE PROCEDURE

38. The pre-incubated cells are removed from the incubator (paragraph 21) and checked under a microscope to assure that they are in good condition (attachment, morphology) prior to dosing.
39. The cells are placed in a bio-safety cabinet and the supplemented medium removed and replaced with new supplemented medium (1 ml/well). DMSO is the preferred solvent for this test method. However, if there are reasons for using other solvents the scientific rationale should be described. Cells are exposed to the test chemical by adding 1 μl of the appropriate stock solution in DMSO (see Appendix II to the validation report (4)) per 1 ml supplemented medium (well volume). This results in a final concentration of 0,1 % DMSO in the wells. To assure adequate mixing it is generally preferred that the appropriate stock solution of the test chemical in DMSO is mixed with supplemented medium to yield the desired final concentration for each dose, and the mixture added to each well immediately after removal of old medium. If this option is used, the concentration of DMSO (0,1 %) should remain consistent among all wells. The wells containing the greatest two concentrations are visually assessed for formation of precipitates or cloudiness as an indication of incomplete solubility of the test chemical by using a stereo microscope. If such conditions (cloudiness, formation precipitates) are observed, wells containing the next lesser concentrations are examined as well (and so forth) and concentrations that did not completely go into solution are to be excluded from further evaluation and analysis. The plate is returned to the incubator at 37 °C under a 5 % CO<sub>2</sub> in air atmosphere for 48 hours. The test chemical plate layout is shown in Table 6. Stocks 1 -7 show placement of increasing doses of test chemical.



## ▼M5

Table 6

**Dosing schematic for the exposure of H295R cells to test chemicals in a 24 well plate**

	1	2	3	4	5	6
A	DMSO	DMSO	DMSO	Stock 4	Stock 4	Stock 4
B	Stock 1	Stock 1	Stock 1	Stock 5	Stock 5	Stock 5
C	Stock 2	Stock 2	Stock 2	Stock 6	Stock 6	Stock 6
D	Stock 3	Stock 3	Stock 3	Stock 7	Stock 7	Stock 7

40. After 48 hours the exposure plates are removed from the incubator and every well is checked under the microscope for cell condition (attachment, morphology, degree of confluence) and signs of cytotoxicity. The medium from each well is split into two equal amounts (approximately 490 µl each) and transferred to two separate vials appropriately labelled (i.e. one aliquot to provide a spare sample for each well). To prevent cells from drying out, medium is removed a row or column at a time and replaced with the medium for the cell viability/cytotoxicity assay. If cell viability/cytotoxicity is not to be measured immediately, 200 µl PBS with Ca<sup>2+</sup> and Mg<sup>2+</sup> is added to each well. The media are frozen at – 80 °C until further processing to analyse hormone concentrations (see paragraphs 44-46). While T and E2 in medium kept at – 80 °C are generally stable for at least 3 months, hormone stability during storage should be documented within each laboratory.
41. Immediately after removing the medium, cell viability/cytotoxicity is determined for each exposure plate.

**Cell Viability Determination**

42. A cell viability/cytotoxicity assay of choice can be used to determine the potential impact of the test chemical on cell viability. The assay should be able to provide a true measure of the percentage of viable cells present in a well, or it should be demonstrated that it is directly comparable to (a linear function of) the Live/Dead® Assay (see Appendix III to the validation report (4)). An alternative assay that has been shown to work equally well is the MTT [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide] test (18). The assessment of cell viability using the above methods is a relative measurement that does not necessarily exhibit linear relationships with the absolute number of cells in a well. Therefore, a subjective parallel visual assessment of each well by the analyst should be conducted, and digital pictures of the SCs and the two greatest non-cytotoxic concentrations are to be taken and archived to enable later assessment of true cell density if this should be required. If by visual inspection or as demonstrated by the viability/cytotoxicity assay there appears to be an increase in cell number, the apparent increase needs to be verified. If an increase in cell numbers is verified, this should be stated in the test report. Cell viability will be expressed relative to the average response in the SCs, which is considered 100 % viable cells, and is calculated as appropriate for the cell viability/cytotoxicity assay that is used. For the MTT assay, the following formula may be used:

▼ **M5**

**% viable cells** = (response in well – average response in MeOH treated [= 100 % dead] wells) ÷ (average response in SC wells – average response in MeOH treated [= 100 % dead] wells)

43. Wells with viability lower than 80 %, relative to the average viability in the SCs (= 100 % viability), should not be included in the final data analysis. Inhibition of steroidogenesis occurring in the presence of almost 20 % cytotoxicity should be carefully evaluated to ensure that cytotoxicity is not the cause for the inhibition.

**Hormone Analysis**

44. Each laboratory can use a hormone measurement system of its choice for the analysis of T and E2. Spare aliquots of medium from each treatment group may be used to prepare dilutions to bring the concentration within the linear part of the standard curve. As noted in paragraph 29, each laboratory should demonstrate the conformance of their hormone measurement system (e.g. ELISA, RIA, LC-MS, LC-MS/MS) with the QC criteria by analysing supplemented medium spiked with an internal hormone control prior to conducting QC runs or testing of chemicals. In order to ensure that the components of the test system do not interfere with measurement of hormones, the hormones may need to be extracted from the media prior to their measurement (see paragraph 30 for the conditions under which an extraction is or is not required). It is recommended to conduct extraction following the procedures in Appendix III to the validation report (4).
45. If a commercial test kit is being used to measure the hormone production, the hormone analysis should be conducted as specified in the manuals provided by the test kit manufacturer. Most manufacturers have a unique procedure by which the hormone analyses are conducted. Dilutions of samples need to be adjusted such that expected hormone concentrations for the solvent controls fall within the centre of the linear range of the standard curve of the individual assay (Appendix III to the validation report (4)). Values outside of the linear portion of the standard curve should be rejected.
46. Final hormone concentrations are calculated as follows:

Example:

Extracted:	450 µl medium
Reconstituted in:	250 µl assay buffer
Dilution in Assay:	1:10 (to bring the sample within the linear range of the standard curve)
Hormone Concentration in Assay:	150 pg/ml (already adjusted to concentration per ml sample assayed)
Recovery:	89 %
Final hormone concentration =	(Hormone concentration (per ml) ÷ recovery) (dilution factor)
Final hormone concentration =	(150 pg/ml) ÷ (0,89) × (250 µl/450 µl) × 10 = 936,3 pg/ml

▼ **M5****Selection of test concentrations**

47. A minimum of two independent runs of the assay should be conducted. Unless prior information such as information on solubility limits or cytotoxicity provides a basis for selecting test concentrations, it is recommended that the test concentrations for the initial run be spaced at  $\log_{10}$  intervals with  $10^{-3}$  M being the maximum concentration. If the chemical is soluble, and not cytotoxic at any of the tested concentrations, and the first run was negative for all concentrations, then it is to be confirmed in one more run using the same conditions as the first run was conducted (Table 7). If the results of the first run are *equivocal* (i.e. the fold-change is statistically significant from the SC at only one concentration) or positive (i.e. the fold change at two or more adjacent concentrations is statistically significant), the test should be repeated as indicated in Table 7 by refining the selected test concentrations. Test concentrations in runs two and three (if applicable) should be adjusted on the basis of the results of the initial run bracketing concentrations that elicited an effect using 1/2-log concentration spacing (e.g. if the original run of 0,001, 0,01, 0,1, 1, 10, 100, 1 000  $\mu\text{M}$  resulted in inductions at 1 and 10  $\mu\text{M}$ , the concentrations tested in the second run should be 0,1, 0,3, 1, 3, 10, 30, 100  $\mu\text{M}$ ), unless lower concentrations need to be employed to achieve a LOEC. In the latter case, at least five concentrations below the lowest concentration tested in the first run should be used in the second run using a 1/2-log scale. If the second run does not confirm the first run (i.e. statistical significance does not occur at the previously positively tested Live/Deadconcentration  $\pm 1$  concentration-increment), a third experiment is to be conducted using the original testing conditions. Equivocal results in the first run are considered negative if the observed effect could not be confirmed in any of the two subsequent runs. Equivocal results are considered as positive responses (effect) when the response can be confirmed in at least one more run within a  $\pm 1$  concentration increment (see section 55 for the Data Interpretation Procedure).

Table 7

**Decision matrix for possible outcome scenarios**

Run 1	Run 2		Run 3		Decision	
	Decision	Scenario	Decision	Scenario	Positive	Negative
Negative	Confirm <sup>(a)</sup>	Negative	Stop			X
Negative	Confirm <sup>(a)</sup>	Positive	Refine <sup>(b)</sup>	Negative		X
Equivocal <sup>(c)</sup>	Refine <sup>(b)</sup>	Negative	Confirm <sup>(a)</sup>	Negative		X
Equivocal <sup>(c)</sup>	Refine <sup>(b)</sup>	Negative	Confirm <sup>(a)</sup>	Positive	X	
Equivocal <sup>(c)</sup>	Refine <sup>(b)</sup>	Positive			X	
Positive	Refine <sup>(b)</sup>	Negative	Confirm <sup>(a)</sup>	Positive	X	

## ▼ M5

Run 1	Run 2		Run 3		Decision	
Scenario	Decision	Scenario	Decision	Scenario	Positive	Negative
Negative	Confirm <sup>(a)</sup>	Positive	Refine <sup>(b)</sup>	Positive	X	
Positive	Refine <sup>(b)</sup>	Positive	Stop		X	

<sup>(a)</sup> Confirm previous run using the same experimental design.

<sup>(b)</sup> Re-run assay at 1/2-log concentration spacing (bracketing the concentration that tested significantly different in the preceding experiment).

<sup>(c)</sup> Fold-change at one concentration is statistically significantly different from the SC

#### Quality Control of the Test Plate

48. In addition to meeting the criteria for the QC plate, other quality criteria that pertain to acceptable variation between replicate wells, replicate experiments, linearity and sensitivity of hormone measurement systems, variability between replicate hormone measures of the same sample, and percentage recovery of hormone spikes after extraction of medium (if applicable; see Paragraph 30 regarding extraction requirements) should be met and are provided in Table 8. Data should fall within the acceptable ranges defined for each parameter to be considered for further evaluation. If these criteria are not met, the spreadsheet should note that QC criteria were not met for the sample in question, and the sample should be re-analysed or dropped from the data set.

Table 8

#### Acceptable ranges and/or variation (%) for H295R assay test plate parameters.

(LOQ: Limit of Quantification of the hormone measurement system. CV: Coefficient of variation; SC: Solvent Control; DPM: Disintegrations per minute)

	Comparison Between	T	E2
Basal hormone production in SCs	Fold-greater than LOQ	≥ 5-fold	≥ 2,5-fold
Exposure Experiments — Within Plate CV for SCs (Replicate Wells)	Absolute Concentrations	≤ 30 %	≤ 30 %
Exposure Experiments — Between Plate CV for SCs (Replicate Experiments)	Fold-Change	≤ 30 %	≤ 30 %
Hormone Measurement System — Sensitivity	Detectable fold-decrease relative to SC	≥ 5-fold	≥ 2,5-fold
Hormone Measurement System — Replicate Measure CV for SCs <sup>(a)</sup>	Absolute Concentrations	≤ 25 %	≤ 25 %
Medium Extraction — Recovery of Internal 3H Standard (If Applicable)	DPM	≥ 65 % Nominal	

<sup>(a)</sup> Refers to replicate measures of the same sample

**▼ M5**

## DATA ANALYSIS AND REPORTING

**Data Analysis**

49. To evaluate the relative increase/decrease in chemically altered hormone production, the results should be normalised to the mean SC value of each test plate, and results expressed as changes relative to the SC in each test plate. All data are to be expressed as mean  $\pm$  1 standard deviation (SD).
50. Only hormone data from wells where cytotoxicity was less than 20 % should be included in the data analysis. Relative changes should be calculated as follows:

**Relative Change** = (Hormone concentration in each well)  $\div$  (Mean hormone concentration in all solvent control well).

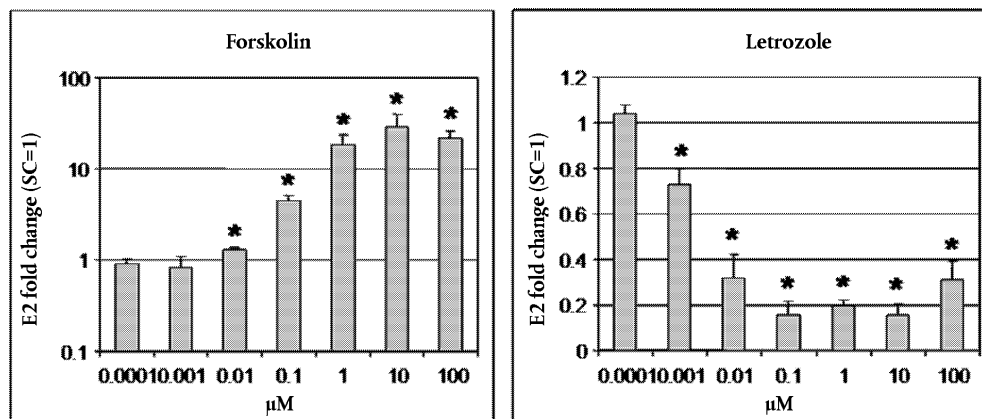
51. If by visual inspection of the well or as demonstrated by the viability/cytotoxicity assay described in paragraph 42 there appears to be an increase in cell number, the apparent increase needs to be verified. If an increase in cell numbers is verified, this should be stated in the test report.
52. Prior to conducting statistical analyses, the assumptions of normality and variance homogeneity should be evaluated. Normality should be evaluated using standard probability plots or other appropriate statistical method (e.g. Shapiro-Wilk's test). If the data (fold changes) are not normally distributed, transformation of the data should be attempted to approximate a normal distribution. If the data are normally distributed or approximate a normal distribution, differences between chemical concentration groups and SCs should be analysed using a parametric test (e.g. Dunnett's Test) with *concentration* being the independent, and *response* (fold-change) being the dependent variable. If data are not normally distributed, an appropriate non-parametric test should be used (e.g. Kruskal Wallis, Steel's Many-one rank test). Differences are considered significant at  $p \leq 0,05$ . Statistical evaluations are done based on average values for each well that represent independent replicate data points. It is anticipated that due to the large spacing of doses in the first run ( $\log_{10}$  scale) in many cases it will not be possible to describe clear concentration-response relationships where the two greatest doses will be on the linear portion of the sigmoid curve. Therefore, for the first run or any other data sets where this condition occurs (e.g. where no maximum efficacy can be estimated) type I fixed variable statistics as described above will be applied.
53. If more than two data points lie on the linear portion of the curve and where maximum efficacies can be calculated — as is anticipated for some of the 2<sup>nd</sup> runs that are conducted using a semi-log spacing of exposure concentrations — a probit, logit or other appropriate regression model should be utilised to calculate effective concentrations (e.g. EC50 and EC20).
54. Results should be provided both in graphical (bar graphs representing mean  $\pm$  1 SD) and tabular (LOEC/NOEC, direction of effect, and strength of maximum response that is part of the dose-response portion of the data) formats (see Figure 3 for an example). Data assessment is only considered valid if it has been based on at least two independently conducted runs. An experiment or run is considered independent if it has been conducted at a different date using a new set of solutions and controls. The concentration range used in runs 2 and 3 (if necessary) may be tailored on the basis of the results of run 1 to better define the dose response range containing the LOEC (see paragraph 47).

## ▼ M5

Figure 3

Example of the presentation and evaluation of data obtained during the conduct of the H295R Assay in graphical and tabular format.

Asterisks indicate statistically significant differences from the solvent control ( $p < 0,05$ ). LOEC: Lowest observed effective concentration; Max Change: Maximum strength of the response observed at any concentration relative to the average SC response (= 1).



Chemical	LOEC	Max Change
Forskolol	0,01	0,15 fold
Letrozole	0,001	29 fold

#### Data Interpretation Procedure

55. A test chemical is judged to be positive if the fold induction is statistically different ( $p \leq 0,05$ ) from the solvent control at two adjacent concentrations in at least two independent runs (Table 7). A test chemical is judged to be negative following two independent negative runs, or in three runs, comprising two negative runs and one equivocal or positive run. If the data generated in three independent experiments does not meet the decision criteria listed in Table 7, the experimental results are not interpretable. Results at concentrations exceeding the limits of solubility or at cytotoxic concentrations should not be included in the interpretation of results.

#### Test Report

56. The test report should include the following information:

##### Testing facility

- Name of facility and location;
- Study director and other personnel and their study responsibilities;
- Dates the study began and ended;

**▼ M5***Test chemical, reagents and controls*

- Identity (name/CAS No as appropriate), source, lot/batch number, purity, supplier, and characterisation of test chemical, reagents, and controls;
- Physical nature and relevant physicochemical properties of test chemical;
- Storage conditions and the method and frequency of preparation of test chemicals, reagents and controls;
- Stability of test chemical;

*Cells*

- Source and type of cells;
- Number of cell passages (cell passage identifier) of cells used in test;
- Description of procedures for maintenance of cell cultures;

*Pre-test requirements (if applicable)*

- Description and results of chemical hormone-assay interference test;
- Description and results of hormone extraction efficiency measurements;
- Standard and calibration curves for all analytical assays to be conducted;
- Detection limits for the selected analytical assays;

*Test conditions*

- Composition of media;
- Concentration of test chemical;
- Cell density (estimated or measured cell concentrations at 24 hours and 48 hours)
- Solubility of test chemical (limit of solubility, if determined);
- Incubation time and conditions;

*Test results*

- Raw data for each well for controls and test chemicals—each replicate measure in form of the original data provided by the instrument utilised to measure hormone production (e.g. OD, fluorescence units, DPM, etc.);
- Validation of normality or explanation of data transformation;
- Mean responses  $\pm$  1 SD for each well measured;
- Cytotoxicity data (test concentrations that caused cytotoxicity);
- Confirmation that QC requirements were met;

**▼ M5**

- Relative change compared with solvent control corrected for cytotoxicity;
- A bar graph showing relative (fold change) at each concentration, SD and statistical significance as stated in paragraph 49-54;

*Data interpretation*

- Apply the data interpretation procedure to the results and discuss findings;

*Discussion*

- Are there any indications from the study regarding the possibility that the T/E2 data could be influenced by indirect effects on the gluco-, and mineral-corticoid pathways?

*Conclusions***LITERATURE**

- (1) OECD (2002), OECD Conceptual Framework for the Testing and Assessment of Endocrine Disrupting Chemicals, in Appendix 2 to Chapter B.54 of this Annex
- (2) Hecker, M., Newsted, J.L., Murphy, M.B., Higley, E.B., Jones, P.D., Wu, R. and Giesy, J.P. (2006), Human adrenocarcinoma (H295R) cells for rapid in vitro determination of effects on steroidogenesis: Hormone production, *Toxicol. Appl. Pharmacol.*, 217, 114-124.
- (3) Hecker, M., Hollert, H., Cooper, R., Vinggaard, A.-M., Akahori, Y., Murphy, M., Nellemann, C., Higley, E., Newsted, J., Wu, R., Lam, P., Laskey, J., Buckalew, A., Grund, S., Nakai, M., Timm, G., and Giesy, J. P. (2007), The OECD validation program of the H295R steroidogenesis assay for the identification of in vitro inhibitors or inducers of testosterone and estradiol production, Phase 2: inter laboratory pre-validation studies. *Env. Sci. Pollut. Res.*, 14, 23-30.
- (4) OECD (2010), *Multi-Laboratory Validation of the H295R Steroidogenesis Assay to Identify Modulators of Testosterone and Estradiol Production*, OECD Series of Testing and Assessment No 132, ENV/JM/MONO(2010)31, Paris. Available at [[http://www.oecd.org/document/30/0,3746,en\\_2649\\_34377\\_1916638\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/30/0,3746,en_2649_34377_1916638_1_1_1_1,00.html)]
- (5) OECD (2010), *Peer Review Report of the H295R Cell-Based Assay for Steroidogenesis*, OECD Series of Testing and Assessment No 133, ENV/JM/MONO(2010)32, Paris. Available at: [[http://www.oecd.org/document/30/0,3746,en\\_2649\\_34377\\_1916638\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/30/0,3746,en_2649_34377_1916638_1_1_1_1,00.html)]
- (6) Battelle (2005), Detailed Review Paper on Steroidogenesis, Available at: [[http://www.epa.gov/endo/pubs/edmvs/steroidogenesis\\_drp\\_final\\_3\\_29\\_05.pdf](http://www.epa.gov/endo/pubs/edmvs/steroidogenesis_drp_final_3_29_05.pdf)]
- (7) Hilscherova, K., Jones, P. D., Gracia, T., Newsted, J. L., Zhang, X., Sanderson, J. T., Yu, R. M. K., Wu, R. S. S. and Giesy, J. P. (2004), Assessment of the Effects of Chemicals on the Expression of Ten Steroidogenic Genes in the H295R Cell Line Using Real-Time PCR, *Toxicol. Sci.*, 81, 78-89.
- (8) Sanderson, J. T., Boerma, J., Lansbergen, G. and Van den Berg, M. (2002), Induction and inhibition of aromatase (CYP19) activity by various classes of pesticides in H295R human adrenocortical carcinoma cells, *Toxicol. Appl. Pharmacol.*, 182, 44-54.
- (9) Breen, M.S., Breen, M., Terasaki, N., Yamazaki, M. and Conolly, R.B. (2010), Computational model of steroidogenesis in human H295R cells to predict biochemical response to endocrine-active chemicals: Model development for metyrapone, *Environ. Health Perspect.*, 118: 265-272.



▼ **M5**

- (10) Higley, E.B., Newsted, J.L., Zhang, X., Giesy, J.P. and Hecker, M. (2010), Assessment of chemical effects on aromatase activity using the H295R cell line, *Environ. Sci. Poll. Res.*, 17:1137-1148.
- (11) Gazdar, A. F., Oie, H. K., Shackleton, C. H., Chen, T. R., Triche, T. J., Myers, C. E., Chrousos, G. P., Brennan, M. F., Stein, C. A. and La Rocca, R. V. (1990), Establishment and characterization of a human adrenocortical carcinoma cell line that expresses Multiple pathways of steroid biosynthesis, *Cancer Res.*, 50, 5488-5496.
- (12) He, Y.H., Wiseman, S.B., Zhang, X.W., Hecker, M., Jones, P.D., El-Din, M.G., Martin, J.W. and Giesy, J.P. (2010), Ozonation attenuates the steroidogenic disruptive effects of sediment free oil sands process water in the H295R cell line, *Chemosphere*, 80:578-584.
- (13) Zhang, X.W., Yu, R.M.K., Jones, P.D., Lam, G.K.W., Newsted, J.L., Gracia, T., Hecker, M., Hilscherova, K., Sanderson, J.T., Wu, R.S.S. and Giesy, J.P. (2005), Quantitative RT-PCR methods for evaluating toxicant-induced effects on steroidogenesis using the H295R cell line, *Environ. Sci. Technol.*, 39:2777-2785.
- (14) Higley, E.B., Newsted, J.L., Zhang, X., Giesy, J.P. and Hecker, M. (2010), Differential assessment of chemical effects on aromatase activity, and E2 and T production using the H295R cell line, *Environ. Sci. Pol. Res.*, 17:1137-1148.
- (15) Rainey, W. E., Bird, I. M., Sawetawan, C., Hanley, N. A., Mccarthy, J. L., Mcgee, E. A., Wester, R. and Mason, J. I. (1993), Regulation of human adrenal carcinoma cell (NCI-H295) production of C19 steroids, *J. Clin. Endocrinol. Metab.*, 77, 731-737.
- (16) Chapter B.55 of this Annex: Hershberger Bioassay in Rats: A short-term Screening Assay for (Anti)Androgenic Properties.
- (17) Shapiro, R., and Page, L.B. (1976), Interference by 2,3-dimercapto-1-propanol (BAL) in angiotensin I radioimmunoassay, *J. Lab. Clin. Med.*, 2, 222-231.
- (18) Mosmann, T. (1983), Rapid colorimetric assay for growth and survival: application to proliferation and cytotoxicity assays, *J. Immunol. Methods.*, 65, 55-63.
- (19) Brock, B.J., Waterman, M.R. (1999). Biochemical differences between rat and human cytochrome P450c17 support the different steroidogenic needs of these two species, *Biochemistry*. 38:1598-1606.
- (20) Oskarsson, A., Ulleras, E., Plant, K., Hinson, J. Goldfarb, P.S., (2006), Steroidogenic gene expression in H295R cells and the human adrenal gland: adrenotoxic effects of lindane in vitro, *J. Appl. Toxicol.*, 26:484-492.

▼ **M5***Appendix*

## DEFINITIONS:

**Confluency** refers to the coverage or proliferation that the cells are allowed over or throughout the culture medium.

**Chemical** means a substance or a mixture.

**CV** refers to the coefficient of variation, and is defined as the ratio of the standard deviation of a distribution to its arithmetic mean.

**CYP** stands for cytochrome P450 mono-oxygenases, a family of genes and the enzymes produced from them that are involved in catalysing a wide variety of biochemical reactions including the synthesis and metabolism of steroid hormones.

**DPM** are disintegration per minute. It is the number of atoms in a given quantity of radioactive material that is detected to have decayed in one minute.

**E2** is 17 $\beta$ -oestradiol, the most important oestrogen in mammalian systems.

**H295R** cells are human adreno-carcinoma cells which have the physiological characteristics of zonally undifferentiated human foetal adrenal cells and which express all of the enzymes of the steroidogenesis pathway. They are available from the ATCC.

**Freeze medium** is used to freeze and to store frozen cells. It consists of stock medium plus BD NuSerum and dimethyl sulfoxide.

**Linear Range** is the range within the standard curve for a hormone measurement system where the results are proportional to the concentration of the analyte present in the sample.

**LOQ** stands for '*Limit of Quantification*', and is the lowest quantity of a chemical that can be distinguished from the absence of that chemical (a blank value) within a stated confidence limit. For the purpose of this method, the LOQ is typically defined by the manufacturer of the test systems if not specified differently.

**LOEC** is the Lowest Observed Effect Concentration, the lowest concentration level at which the assay response is statistically different from that of the solvent control.

**NOEC** is the No Observed Effect Concentration, which is the highest concentration tested if the assay does not provide a positive response.

**Passage** is the number of times that cells are split after initiation of a culture from frozen stock. The initial passage that was started from the frozen stock is assigned the number one (1). Cells that were split 1 time are labelled passage 2, etc.

**PBS** is Dulbecco's phosphate buffered saline.

**Quality Control**, abbreviated QC, refers to the measures needed to assure valid data.

**Quality control plate** is a 24 well plate containing two concentrations of the positive and negative controls to monitor the performance of a new batch of cells or to provide the positive controls for the assay when testing chemicals.

**Run** is an independent experiment characterised by a new set of solutions and controls.

**Stock medium** is the base for the preparation of other reagents. It consists of a 1:1 mixture of Dulbecco's Modified Eagle's Medium and Ham's F-12 Nutrient mixture (DMEM/F12) in 15 mM HEPES buffer without phenol red or sodium bicarbonate. Sodium bicarbonate is added as the buffer, see Appendix II to the validation report (4).

**▼ M5**

**Supplemented medium** consist of stock medium plus BD Nu-Serum and ITS+ premium mix, see Appendix II to the validation report (4).

**Steroidogenesis** is the synthetic pathway leading from cholesterol to the various steroid hormones. Several intermediates in the steroid synthesis pathway such as progesterone and testosterone are important hormones in their own right but also serve as precursors to hormones farther down the synthetic pathway.

**T** stands for testosterone, one of the two most important androgens in mammalian systems.

**Test chemical** is any substance or mixture tested using this test method.

**Test plate** is the plate on which H295R cells are exposed to test chemicals. Test plates contain the solvent control and the test chemical at seven concentration levels in triplicate.

**Trypsin 1X** is a dilute solution of the enzyme trypsin, a pancreatic serine protease, used to loosen cells from a cell cultivation plate, see Appendix III to the validation report (4).

**▼ M5****B.58 TRANSGENIC RODENT SOMATIC AND GERM CELL GENE MUTATION ASSAYS**

## INTRODUCTION

1. This test method is equivalent to OECD Test Guideline (TG) 488 (2013). EU test methods are available for a wide range of *in vitro* mutation assays that are able to detect chromosomal and/or gene mutations. There are test methods for *in vivo* endpoints (i.e. chromosomal aberrations and unscheduled DNA synthesis); however, these do not measure gene mutations. Transgenic Rodent (TGR) mutation assays fulfil the need for practical and widely available *in vivo* tests for gene mutations.
2. The TGR mutation assays have been reviewed extensively (24) (33). They use transgenic rats and mice that contain multiple copies of chromosomally integrated plasmid or phage shuttle vectors. The transgenes contain reporter genes for the detection of various types of mutations induced *in vivo* by test chemicals.
3. Mutations arising in a rodent are scored by recovering the transgene and analysing the phenotype of the reporter gene in a bacterial host deficient for the reporter gene. TGR gene mutation assays measure mutations induced in genetically neutral genes recovered from virtually any tissue of the rodent. These assays, therefore, circumvent many of the existing limitations associated with the study of *in vivo* gene mutation in endogenous genes (e.g. limited tissues suitable for analysis, negative/positive selection against mutations).
4. The weight of evidence suggests that transgenes respond to mutagens in a similar manner to endogenous genes, especially with regard to the detection of base pair substitutions, frameshift mutations, and small deletions and insertions(24).
5. The International Workshops on Genotoxicity Testing (IWGT) have endorsed the inclusion of TGR gene mutation assays for *in vivo* detection of gene mutations, and have recommended a protocol for their implementation (15) (29). This test method is based on these recommendations. Further analysis supporting the use of this protocol can be found in (16).
6. It is anticipated that in the future it may be possible to combine a TGR gene mutation assay with a repeat dose toxicity study (Chapter B.7 of this Annex). However, data are required to ensure that the sensitivity of TGR gene mutation assays is unaffected by the shorter one day period of time between the end of the administration period and the sampling time, as used in the repeat dose toxicology study, compared to 3 days used in TGR gene mutation assays. Data are also required to indicate that the performance of the repeat dose assay is not adversely affected by using a transgenic rodent strain rather than traditional rodent strains. When these data are available, this test method will be updated.
7. Definitions of key terms are set out in the Appendix.

**▼ M5**

## INITIAL CONSIDERATIONS

8. TGR gene mutation assays for which sufficient data are available to support their use in this test method are: *lacZ* bacteriophage mouse (Muta<sup>TM</sup>Mouse); *lacZ* plasmid mouse; *gpt* delta (*gpt* and *Spi*<sup>-</sup>) mouse and rat; *lacI* mouse and rat (Big Blue®), as performed under standard conditions. In addition, the *cII* positiveselection assay can be used for evaluating mutations in the Big Blue® and Muta<sup>TM</sup>Mouse models. Mutagenesis in the TGR models is normally assessed as mutant frequency; if required, however, molecular analysis of the mutations can provide additional information (see paragraph 24).
  
9. These rodent *in vivo* gene mutation tests are especially relevant to assessing mutagenic hazard in that the assays' responses are dependent upon *in vivo* metabolism, pharmacokinetics, DNA repair processes, and translesion DNA synthesis, although these may vary among species, among tissues and among the types of DNA damage. An *in vivo* assay for gene mutations is useful for further investigation of a mutagenic effect detected by an *in vitro* system, and for following up results of tests using other *in vivo* endpoints (24). In addition to being causally associated with the induction of cancer, gene mutation is a relevant endpoint for the prediction of mutation-based non-cancer diseases in somatic tissues (12) (13) as well as diseases transmitted through the germline.
  
10. If there is evidence that the test chemical, or a relevant metabolite, will not reach any of the tissues of interest, it is not appropriate to perform a TGR gene mutation assay.

## PRINCIPLE OF THE TEST

11. In the assays described in paragraph 8, the target gene is bacterial or bacteriophage in origin, and the means of recovery from the rodent genomic DNA is by incorporation of the transgene into a  $\lambda$  bacteriophage or plasmid shuttle vector. The procedure involves the extraction of genomic DNA from the rodent tissue of interest, *in vitro* processing of the genomic DNA (i.e. packaging of  $\lambda$  vectors, or ligation and electroporation of plasmids to recover the shuttle vector), and subsequent detection of mutations in bacterial hosts under suitable conditions. The assays employ neutral transgenes that are readily recoverable from most tissues.
  
12. The basic TGR gene mutation experiment involves treatment of the rodent with a chemical over a period of time. Chemicals may be administered by any appropriate route, including implantation (e.g. medical device testing). The total period during which an animal is dosed is referred to as the administration period. Administration is usually followed by a period of time, prior to sacrifice, during which the chemical is not administered and during which unrepaired DNA lesions are fixed into stable mutations. In the literature, this period has been variously referred to as the manifestation time, fixation time or expression time; the end of this period is the sampling time (15) (29). After the animal is sacrificed, genomic DNA is isolated from the tissue(s) of interest and purified.

▼ **M5**

13. Data for a single tissue per animal from multiple packaging/ligations are usually aggregated, and mutant frequency is generally evaluated using a total of between  $10^5$  and  $10^7$  plaque-forming or colony-forming units. When using positive selection methods, total plaque-forming units are determined with a separate set of non-selective plates.
14. Positive selection methods have been developed to facilitate the detection of mutations in both the *gpt* gene [*gpt* delta mouse and rat, *gpt*<sup>-</sup> phenotype (20) (22) (28)] and the *lacZ* gene [Muta™Mouse or *lacZ* plasmid mouse (3) (10) (11) (30)]; whereas, *lacI* gene mutations in Big Blue® animals are detected through a non-selective method that identifies mutants through the generation of coloured (blue) plaques. Positive selection methodology is also in place to detect point mutations arising in the *cII* gene of the  $\lambda$  bacteriophage shuttle vector [Big Blue® mouse or rat, and Muta™Mouse (17)] and deletion mutations in the  $\lambda$  *red* and *gam* genes [*Spi*<sup>-</sup> selection in *gpt* delta mouse and rat (21) (22) (28)]. Mutant frequency is calculated by dividing the number of plaques/plasmids containing mutations in the transgene by the total number of plaques/plasmids recovered from the same DNA sample. In TGR gene mutation studies, the mutant frequency is the reported parameter. In addition, a mutation frequency can be determined as the fraction of cells carrying independent mutations; this calculation requires correction for clonal expansion by sequencing the recovered mutants (24).
15. The mutations scored in the *lacI*, *lacZ*, *cII* and *gpt* point mutation assays consist primarily of base pair substitution mutations, frameshift mutations and small insertions/deletions. The relative proportion of these mutation types among spontaneous mutations is similar to that seen in the endogenous *Hprt* gene. Large deletions are detected only with the *Spi*<sup>-</sup> selection and the *lacZ* plasmid assays (24). Mutations of interest are *in vivo* mutations that arise in the mouse or rat. *In vitro* and *ex vivo* mutations, which may arise during phage/plasmid recovery, replication or repair, are relatively rare, and in some systems can be specifically identified, or excluded by the bacterial host/positive selection system.

## DESCRIPTION OF THE METHOD

**Preparations***Selection of animal species*

16. A variety of transgenic mouse gene mutation detection models are currently available, and these systems have been more widely used than transgenic rat models. If the rat is clearly a more appropriate model than the mouse (e.g. when investigating the mechanism of carcinogenesis for a tumour seen only in rats, to correlate with a rat toxicity study, or if rat metabolism is known to be more representative of human metabolism) the use of transgenic rat models should be considered.

*Housing and feeding conditions*

17. The temperature in the experimental animal room ideally should be 22 °C ( $\pm$  3 °C). Although the relative humidity should be at least 30 % and preferably not exceed 70 % other than during room cleaning, the goal should be to maintain a relative humidity of 50-60 %. Lighting should be artificial, with a daily sequence of 12 hours light, followed by 12 hours dark. For feeding, conventional laboratory diets may be used with an

**▼M5**

unlimited supply of drinking water. The choice of diet may be influenced by the need to ensure a suitable admixture of a test chemical when administered by this route. Animals should be housed in small groups (no more than five) of the same sex if no aggressive behaviour is expected. Animals may be housed individually if scientifically justified.

*Preparation of the animals*

18. Healthy young sexually mature adult animals (8-12 weeks old at start of treatment) are randomly assigned to the control and treatment groups. The animals are identified uniquely. The animals are acclimated to the laboratory conditions for at least five days. Cages should be arranged in such a way that possible effects due to cage placement are minimised. At the commencement of the study, the weight variation of animals should be minimal and not exceed  $\pm 20\%$  of the mean weight of each sex.

*Preparation of doses*

19. Solid test chemicals should be dissolved or suspended in appropriate solvents or vehicles or admixed in diet or drinking water prior to dosing of the animals. Liquid test chemicals may be dosed directly or diluted prior to dosing. For inhalation exposures, test chemicals can be administered as gas, vapour or a solid/liquid aerosol, depending on their physicochemical properties. Fresh preparations of the test chemical should be employed unless stability data demonstrate the acceptability of storage.

**Test Conditions***Solvent/vehicle*

20. The solvent/vehicle should not produce toxic effects at the dose volumes used, and should not be suspected of chemical reaction with the test chemical. If other than well-known solvents/vehicles are used, their inclusion should be supported with reference data indicating their compatibility. It is recommended that wherever possible, the use of an aqueous solvent/vehicle should be considered first.

*Positive Controls*

21. Concurrent positive control animals should normally be used. However, for laboratories that have demonstrated competency (see paragraph 23) and routinely use these assays, DNA from previous positive control treated animals may be included with each study to confirm the success of the method. Such DNA from previous experiments should be obtained from the same species and tissues of interest, and properly stored (see paragraph 36). When concurrent positive controls are used, it is not necessary to administer them by the same route as the test chemical; however, the positive controls should be known to induce mutations in one or more tissues of interest for the test chemical. The doses of the positive control chemicals should be selected so as to produce weak or moderate effects that critically assess the performance and sensitivity of the assay. Examples of positive control chemicals and some of their target tissues are included in Table 1.

## ▼M5

Table 1

## Examples of positive control chemicals and some of their target tissues

Positive control chemical and CAS No	EINECS name and EINECS No	Characteristics	Mutation Target Tissue	
			Rat	Mouse
N-Ethyl-N-nitrosourea [CAS No 759-73-9]	N-Ethyl-N-nitrosourea [212-072-2]	Direct acting mutagen	Liver, lung	Bone marrow, colon, colonic epithelium, intestine, liver, lung, spleen, kidney, ovarian granulosa cells, male germ cells
Ethyl carbamate (urethane) [CAS No 51-79-6]	Urethane [200-123-1]	Mutagen, requires metabolism but produces only weak effects		Bone marrow, fore-stomach, small intestine, liver, lung, spleen
2,4-Diaminotoluene [CAS No 95-80-7]	4-Methyl-m-phenylenediamine [202-453-1]	Mutagen, requires metabolism, also positive in the Spi <sup>+</sup> assay	Liver	Liver
Benzo[a]pyrene [CAS No 50-32-8]	Benzo[def]chrysene [200-028-5]	Mutagen, requires metabolism	Liver, omenta,	Bone marrow, breast, colon, fore-stomach, glandular stomach, heart, liver, lung, male germ cells

*Negative controls*

22. Negative controls, treated with solvent or vehicle alone, and otherwise treated in the same way as the treatment groups, should be included for every sampling time. In the absence of historical or published control data showing that no deleterious or mutagenic effects are induced by the chosen solvent/vehicle, untreated controls should also be included for every sampling time in order to establish acceptability of the vehicle control.

*Verification of laboratory proficiency*

23. Competency in these assays should be established by demonstrating the ability to reproduce expected results from published data (24) for: 1) mutant frequencies with positive control chemicals (including weak responses) such as those listed in Table 1, non-mutagens, and vehicle controls; and 2) transgene recovery from genomic DNA (e.g. packaging efficiency).

**Sequencing of mutants**

24. For regulatory applications, DNA sequencing of mutants is not required, particularly where a clear positive or negative result is obtained. However, sequencing data may be useful when high inter-individual variation is observed. In these cases, sequencing can be used to rule out the possibility of jackpots or clonal events by identifying the proportion of unique mutants from a particular tissue. Sequencing approximately 10 mutants per tissue per animal should be sufficient for simply determining if clonal mutants contribute to the mutant frequency; sequencing as many as 25 mutants may



**▼ M5**

be necessary to correct mutant frequency mathematically for clonality. Sequencing of mutants also may be considered when small increases in mutant frequency (i.e. just exceeding the untreated control values) are found. Differences in the mutant spectrum between the mutant colonies from treated and untreated animals may lend support to a mutagenic effect (29). Also, mutation spectra may be useful for developing mechanistic hypotheses. When sequencing is to be included as part of the study protocol, special care should be taken in the design of such studies, in particular with respect to the number of mutants sequenced per sample, to achieve adequate power according to the statistical model used (see paragraph 43).

**PROCEDURE****Number and Sex of Animals**

25. The number of animals per group should be predetermined to be sufficient to provide statistical power necessary to detect at least a doubling in mutant frequency. Group sizes will consist of a minimum of five animals; however, if the statistical power is insufficient, the number of animals should be increased as required. Male animals should normally be used. There may be cases where testing females alone would be justified; for example, when testing human female-specific drugs, or when investigating female-specific metabolism. If there are significant differences between the sexes in terms of toxicity or metabolism, then both males and females will be required.

**Administration Period**

26. Based on observations that mutations accumulate with each treatment, a repeated-dose regimen is necessary, with daily treatments for a period of 28 days. This is generally considered acceptable both for producing a sufficient accumulation of mutations by weak mutagens, and for providing an exposure time adequate for detecting mutations in slowly proliferating organs. Alternative treatment regimens may be appropriate for some evaluations, and these alternative dosing schedules should be scientifically justified in the protocol. Treatments should not be shorter than the time required for the complete induction of all the relevant metabolising enzymes, and shorter treatments may necessitate the use of multiple sampling times that are suitable for organs with different proliferation rates. In any case, all available information (e.g. on general toxicity or metabolism and pharmacokinetics) should be used when justifying a protocol, especially when deviating from the above standard recommendations. While it may increase sensitivity, treatment times longer than 8 weeks should be explained clearly and justified, since long treatment times may produce an apparent increase in mutant frequency through clonal expansion (29).

**Dose Levels**

27. Dose levels should be based on the results of a dose range-finding study measuring general toxicity that was conducted by the same route of exposure, or on the results of pre-existing sub-acute toxicity studies. Non-transgenic animals of the same rodent strain may be used for determining dose ranges. In the main test, in order to obtain dose response information, a complete study should include a negative control group (see paragraph 22) and a minimum of three, appropriately-spaced dose

**▼ M5**

levels, except where the limit dose has been used (see paragraph 28). The top dose should be the Maximum Tolerated Dose (MTD). The MTD is defined as the dose producing signs of toxicity such that higher dose levels, based on the same dosing regimen, would be expected to produce lethality. Chemicals with specific biological activities at low non-toxic doses (such as hormones and mitogens), and chemicals which exhibit saturation of toxicokinetic properties may be exceptions to the dose-setting criteria and should be evaluated on a case-by-case basis. The dose levels used should cover a range from the maximum to little or no toxicity.

**Limit Test**

28. If dose range-finding experiments, or existing data from related rodent strains, indicate that a treatment regime of at least the limit dose (see below) produces no observable toxic effects, and if genotoxicity would not be expected based upon data from structurally related chemicals, then a full study using three dose levels may not be considered necessary. For an administration period of 28 days (i.e. 28 daily treatments), the limit dose is 1 000 mg/kg body weight/day. For administration periods of 14 days or less, the limit dose is 2 000 mg/kg/body weight/day (dosing schedules differing from 28 daily treatments should be scientifically justified in the protocol; see paragraph 26).

**Administration of Doses**

29. The test chemical is usually administered by gavage using a stomach tube or a suitable intubation cannula. In general, the anticipated route of human exposure should be considered when designing an assay. Therefore, other routes of exposure (such as drinking water, subcutaneous, intravenous, topical, inhalation, intratracheal, dietary, or implantation) may be acceptable where they can be justified. Intraperitoneal injection is not recommended since it is not a physiologically relevant route of human exposure. The maximum volume of liquid that can be administered by gavage or injection at one time depends on the size of the test animal. The volume should not exceed 2 ml/100 g body weight. The use of volumes greater than this should be justified. Except for irritating or corrosive chemicals, which will normally reveal exacerbated effects at higher concentrations, variability in test volume should be minimised by adjusting the concentration to ensure a constant volume at all dose levels.

**Sampling Time***Somatic Cells*

30. The sampling time is a critical variable because it is determined by the period needed for mutations to be fixed. This period is tissue-specific and appears to be related to the turnover time of the cell population, with bone marrow and intestine being rapid responders and the liver being much slower. A suitable compromise for the measurement of mutant frequencies in both rapidly and slowly proliferating tissues is 28 consecutive daily treatments (as indicated in paragraph 26) and sampling three days after the final treatment; although the maximum mutant frequency may not manifest itself in slowly proliferating tissues under these conditions. If slowly proliferating tissues are of particular importance, then a later sampling time of 28 days following the 28 day administration period may be more appropriate (16) (29). In such cases, the later sampling time would replace the 3 day sampling time, and would require scientific justification.

**▼ M5***Germ Cells*

31. TGR assays are well-suited for the study of gene mutation induction in male germ cells (7) (8) (27), in which the timing and kinetics of spermatogenesis have been well-defined (27). The low numbers of ova available for analysis, even after super-ovulation, and the fact that there is no DNA synthesis in the oocyte, preclude the determination of mutation in female germ cells using transgenic assays (31).
  
32. The sampling times for male germ cells should be selected so that the range of exposed cell types throughout germ cell development is sampled, and so that the stage targeted in the sampling has received sufficient exposure. The time for the progression of developing germ cells from spermatogonial stem cells to mature sperm reaching the vas deferens/cauda epididymis ~ 49 days for the mouse (36) and ~70 days for the rat (34) (35). Following a 28-day exposure with a subsequent three day sampling period, accumulated sperm collected from the vas deferens/cauda epididymis (7)(8) will represent a population of cells exposed during approximately the latter half of spermatogenesis, which includes the meiotic and postmeiotic period, but not the spermatogonial or stem cell period. In order to adequately sample cells in the vas deferens/cauda epididymis that were spermatogonial stem cells during the exposure period, an additional sampling time at a minimum of 7 weeks (mice) or 10 weeks (rat), after the end of treatment is required.
  
33. Cells extruded from seminiferous tubules after a 28 + 3 day regimen comprise a mixed population enriched for all stages of developing germ cells (7) (8). Sampling these cells for gene mutation detection does not provide as precise an assessment of the stages at which germ cell mutations are induced as can be obtained from sampling spermatozoa from the vas deferens/cauda epididymis (since there is a range of germ cell types sampled from the tubules, and there will be some somatic cells contaminating this cell population). However, sampling cells from seminiferous tubules in addition to spermatozoa from the vas deferens/cauda epididymis following only a 28 + 3 day sampling regimen would provide some coverage of cells exposed across the majority of phases of germ cell development, and may be useful for detecting some germ cell mutagens.

**Observations**

34. General clinical observations should be made at least once a day, preferably at the same time(s) each day and considering the peak period of anticipated effects after dosing. The health condition of the animals should be recorded. At least twice daily, all animals should be observed for morbidity and mortality. All animals should be weighed at least once a week, and at sacrifice. Measurements of food consumption should be made at least weekly. If the test chemical is administered via the drinking water, water consumption should be measured at each change of water and at least weekly. Animals exhibiting non-lethal indicators of excess toxicity should be euthanatised prior to completion of the test period (23).

**▼ M5****Tissue Collection**

35. The rationale for tissue collection should be defined clearly. Since it is possible to study mutation induction in virtually any tissue, the selection of tissues to be collected should be based upon the reason for conducting the study and any existing mutagenicity, carcinogenicity or toxicity data for the chemical under investigation. Important factors for consideration should include the route of administration (based on likely human exposure route(s)), the predicted tissue distribution, and the possible mechanism of action. In the absence of any background information, several somatic tissues as may be of interest should be collected. These should represent rapidly proliferating, slowly proliferating and site of contact tissues. In addition, spermatozoa from the vas deferens/cauda epididymis and developing germ cells from the seminiferous tubules (as described in paragraphs 32 and 33) should be collected and stored in case future analysis of germ cell mutagenicity is required. Organ weights should be obtained, and for larger organs, the same area should be collected from all animals.

**Storage of Tissues and DNA**

36. Tissues (or tissue homogenates) should be stored at or below  $-70^{\circ}\text{C}$  and be used for DNA isolation within 5 years. Isolated DNA, stored refrigerated at  $4^{\circ}\text{C}$  in appropriate buffer, should be used optimally for mutation analysis within 1 year.

**Selection of Tissues for Mutant Analysis**

37. The choice of tissues should be based on considerations such as: 1) the route of administration or site of first contact (e.g. glandular stomach if administration is oral, lung if administration is through inhalation, or skin if topical application has been used); and 2) pharmacokinetic parameters observed in general toxicity studies, which indicate tissue disposition, retention or accumulation, or target organs for toxicity. If studies are conducted to follow up carcinogenicity studies, target tissues for carcinogenicity should be considered. The choice of tissues for analysis should maximise the detection of chemicals that are direct-acting *in vitro* mutagens, rapidly metabolised, highly reactive or poorly absorbed, or those for which the target tissue is determined by route of administration (6).
38. In the absence of background information and taking into consideration the site of contact due to route of administration, the liver and at least one rapidly dividing tissue (e.g. glandular stomach, bone marrow) should be evaluated for mutagenicity. In most cases, the above requirements can be achieved from analyses of two carefully selected tissues, but in some cases, three or more would be needed. If there are reasons to be specifically concerned about germ cell effects, including positive responses in somatic cells, germ cell tissues should be evaluated for mutations.

**Methods of Measurement**

39. Standard laboratory or published methods for the detection of mutants are available for the recommended transgenic models: *lacZ* lambda bacteriophage and plasmid (30); *lacI* mouse (2) (18); *gpt* delta mouse (22); *gpt* delta rat (28); *cII* (17). Modifications should be justified and properly documented. Data from multiple packagings can be aggregated and used to reach an adequate number of plaques or colonies. However, the need for a large number of packaging reactions to reach the appropriate number of plaques may be an indication of poor DNA quality. In such cases, data should be considered cautiously because they may be unreliable. The optimal total number of plaques or colonies per DNA sample is governed

**▼ M5**

by the statistical probability of detecting sufficient numbers of mutants at a given spontaneous mutant frequency. In general, a minimum of 125 000 to 300 000 plaques is required if the spontaneous mutant frequency is in the order of  $3 \times 10^{-5}$  (15). For the Big Blue® *lacI* assay, it is important to demonstrate that the whole range of mutant colour phenotypes can be detected by inclusion of appropriate colour controls concurrent with each plating. Tissues and the resulting samples (items) should be processed and analysed using a block design, where items from the vehicle/solvent control group, the positive control group (if used) or positive control DNA (where appropriate), and each treatment group are processed together.

**DATA AND REPORTING****Treatment of Results**

40. Individual animal data should be presented in tabular form. The experimental unit is the animal. The report should include the total number of plaque-forming units (pfu) or colony-forming units (cfu), the number of mutants, and the mutant frequency for each tissue from each animal. If there are multiple packaging/rescue reactions, the number of reactions per DNA sample should be reported. While data for each individual reaction should be retained, only the total pfu or cfu need be reported. Data on toxicity and clinical signs as per paragraph 34 should be reported. Any sequencing results should be presented for each mutant analysed, and resulting mutation frequency calculations for each animal and tissue should be shown.

**Statistical Evaluation and Interpretation of Results**

41. There are several criteria for determining a positive result, such as a dose-related increase in the mutant frequency, or a clear increase in the mutant frequency in a single dose group compared to the solvent/vehicle control group. At least three treated dose groups should be analysed in order to provide sufficient data for dose-response analysis. While biological relevance of the results should be the primary consideration, appropriate statistical methods may be used as an aid in evaluating the test results (4) (14) (15) (25) (26). Statistical tests used should consider the animal as the experimental unit.
42. A test chemical for which the results do not meet the above criteria in any tissue is considered non-mutagenic in this assay. For biological relevance of a negative result, tissue exposure should be confirmed.
43. For DNA sequencing analyses, a number of statistical approaches are available to assist in interpreting the results (1) (5) (9) (19).
44. Consideration of whether the observed values are within or outside of the historical control range can provide guidance when evaluating the biological significance of the response (32).

**▼ M5****Test report**

45. The test report should include the following information:

*Test chemical:*

- identification data and CAS n<sup>o</sup>, if known;
- source, lot number if available;
- physical nature and purity;
- physiochemical properties relevant to the conduct of the study;
- stability of the test chemical, if known;

*Solvent/vehicle:*

- justification for choice of vehicle;
- solubility and stability of the test chemical in the solvent/vehicle, if known;
- preparation of dietary, drinking water or inhalation formulations;
- analytical determinations on formulations (e.g. stability, homogeneity, nominal concentrations);

*Test animals:*

- species/strain used and justification for the choice;
- number, age and sex of animals;
- source, housing conditions, diet, etc.;
- individual weight of the animals at the start of the test, including body weight range, mean and standard deviation for each group;

*Test conditions:*

- positive and negative (vehicle/solvent) control data;
- data from the range-finding study;
- rationale for dose level selection;
- details of test chemical preparation;
- details of the administration of the test chemical;
- rationale for route of administration;
- methods for measurement of animal toxicity, including, where available, histopathological or haematological analyses and the frequency with which animal observations and body weights were taken;
- methods for verifying that the test chemical reached the target tissue, or general circulation, if negative results are obtained;

**▼ M5**

- actual dose (mg/kg body weight/day) calculated from diet/drinking water test chemical concentration (ppm) and consumption, if applicable;
- details of food and water quality;
- detailed description of treatment and sampling schedules and justifications for the choices;
- method of euthanasia;
- procedures for isolating and preserving tissues;
- methods for isolation of rodent genomic DNA, rescuing the transgene from genomic DNA, and transferring transgenic DNA to a bacterial host;
- source and lot numbers of all cells, kits and reagents (where applicable);
- methods for enumeration of mutants;
- methods for molecular analysis of mutants and use in correcting for clonality and/or calculating mutation frequencies, if applicable;

*Results:*

- animal condition prior to and throughout the test period, including signs of toxicity;
- body and organ weights at sacrifice;
- for each tissue/animal, the number of mutants, number of plaques or colonies evaluated, mutant frequency;
- for each tissue/animal group, number of packaging reactions per DNA sample, total number of mutants, mean mutant frequency, standard deviation;
- dose-response relationship, where possible;
- for each tissue/animal, the number of independent mutants and mean mutation frequency, where molecular analysis of mutations was performed;
- concurrent and historical negative control data with ranges, means and standard deviations;
- concurrent positive control (or non-concurrent DNA positive control) data;
- analytical determinations, if available (e.g. DNA concentrations used in packaging, DNA sequencing data);
- statistical analyses and methods applied;

*Discussion of the results**Conclusion*

## ▼ M5

## LITERATURE

- (1) Adams, W.T. and T.R. Skopek (1987), 'Statistical Test for the Comparison of Samples from Mutational Spectra', *J. Mol. Biol.*, 194: 391-396.
- (2) Bielas, J.H. (2002), 'A more Efficient Big Blue® Protocol Improves Transgene Rescue and Accuracy in an Adduct and Mutation Measurement', *Mutation Res.*, 518: 107-112.
- (3) Boerrigter, M.E., M.E. Dollé, H.-J. Martus, J.A. Gossen and J. Vijg (1995), 'Plasmid-based Transgenic Mouse Model for Studying *in vivo* Mutations' *Nature*, 377(6550): 657-659
- (4) Carr, G.J. and N.J. Gorelick (1995), 'Statistical Design and Analysis of Mutation Studies in Transgenic Mice', *Environ. Mol. Mutagen*, 25(3): 246-255.
- (5) Carr, G.J. and N.J. Gorelick (1996), 'Mutational Spectra in Transgenic Animal Research: Data Analysis and Study Design Based upon the Mutant or Mutation Frequency', *Environ. Mol. Mutagen*, 28: 405-413.
- (6) Dean, S.W., T.M. Brooks, B. Burlinson, J. Mirsalis, B. Myhr, L. Recio and V. Thybaud (1999), 'Transgenic Mouse Mutation Assay Systems can Play an important Role in Regulatory Mutagenicity Testing *in vivo* for the Detection of Site-of-contact Mutagens', *Mutagenesis*, 14(1): 141-151.
- (7) Douglas, G.R., J. Jiao, J.D. Gingerich, J.A. Gossen and L.M. Soper(1995), 'Temporal and Molecular Characteristics of Mutations Induced by Ethylnitrosourea in Germ Cells Isolated from Seminiferous Tubules and in Spermatozoa of *lacZ* Transgenic Mice', *Proc. Natl. Acad. Sci. USA*, 92: 7485-7489.
- (8) Douglas, G.R., J.D. Gingerich, L.M. Soper and J. Jiao (1997), 'Toward an Understanding of the Use of Transgenic Mice for the Detection of Gene Mutations in Germ Cells', *Mutation Res.*, 388(2-3): 197-212.
- (9) Dunson, D.B. and K.R. Tindall (2000), 'Bayesian Analysis of Mutational Spectra', *Genetics*, 156: 1411-1418.
- (10) Gossen, J.A., W.J. de Leeuw, C.H. Tan, E.C. Zwarthoff, F. Berends, P.H. Lohman, D.L. Knook and J. Vijg(1989), 'Efficient Rescue of Integrated Shuttle Vectors from Transgenic Mice: a Model for Studying Mutations *in vivo*', *Proc. Natl. Acad. Sci. USA*, 86(20): 7971-7975.
- (11) Gossen, J.A. and J. Vijg (1993), 'A Selective System for *lacZ*-Phage using a Galactose-sensitive *E. coli* Host', *Biotechniques*, 14(3): 326, 330.
- (12) Erikson, R.P. (2003), 'Somatic Gene Mutation and Human Disease other than Cancer', *Mutation Res.*, 543: 125-136.
- (13) Erikson, R.P. (2010), 'Somatic Gene Mutation and Human Disease other than Cancer: an Update', *Mutation Res.*, **705: 96-106**.
- (14) Fung, K.Y., G.R. Douglas and D. Krewski (1998), 'Statistical Analysis of *lacZ* Mutant Frequency Data from Muta™Mouse Mutagenicity Assays', *Mutagenesis*, 13(3): 249-255.
- (15) Heddle, J.A., S. Dean, T. Nohmi, M. Boerrigter, D. Casciano, G.R. Douglas, B.W. Glickman, N.J. Gorelick, J.C. Mirsalis, H.-J. Martus, T.R. Skopek, V. Thybaud, K.R.Tindall and N. Yajima (2000), 'In vivo Transgenic Mutation Assays', *Environ. Mol. Mutagen.*, 35: 253-259.
- (16) Heddle, J.A., H.-J. Martus and G.R. Douglas (2003), 'Treatment and Sampling Protocols for Transgenic Mutation Assays', *Environ. Mol. Mutagen.*, 41: 1-6.



## ▼ M5

- (17) Jakubczak, J.L., G. Merlino, J.E. French, W.J. Muller, B. Paul, S. Adhya and S. Garges (1996), 'Analysis of Genetic Instability during Mammary Tumor Progression using a novel Selection-based Assay for *in vivo* Mutations in a Bacteriophage  $\lambda$  Transgene Target', *Proc. Natl. Acad. Sci. USA*, 93(17): 9073–9078.
- (18) Kohler, S.W., G.S. Provost, P.L. Kretz, A. Fieck, J.A. Sorge and J.M. Short (1990), 'The Use of Transgenic Mice for Short-term, *in vivo* Mutagenicity Testing', *Genet. Anal. Tech. Appl.*, 7(8): 212–218.
- (19) Lewis P.D., B. Manshian, M.N. Routledge, G.B. Scott and P.A. Burns (2008), 'Comparison of Induced and Cancer-associated Mutational Spectra using Multivariate Data Analysis', *Carcinogenesis*, 29(4): 772-778.
- (20) Nohmi, T., M. Katoh, H. Suzuki, M. Matsui, M. Yamada, M. Watanabe, M. Suzuki, N. Horiya, O. Ueda, T. Shibuya, H. Ikeda and T. Sofuni (1996), 'A new Transgenic Mouse Mutagenesis Test System using Spi<sup>-</sup> and 6-thioguanine Selections', *Environ. Mol. Mutagen.*, 28(4): 465–470.
- (21) Nohmi, T., M. Suzuki, K. Masumura, M. Yamada, K. Matsui, O. Ueda, H. Suzuki, M. Katoh, H. Ikeda and T. Sofuni (1999), 'Spi<sup>-</sup> Selection: an Efficient Method to Detect  $\gamma$ -ray-induced Deletions in Transgenic Mice', *Environ. Mol. Mutagen.*, 34(1): 9–15.
- (22) Nohmi, T., T. Suzuki and K.I. Masumura (2000), 'Recent Advances in the Protocols of Transgenic Mouse Mutation Assays', *Mutation Res.*, 455(1–2): 191–215.
- (23) OECD (2000), Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation, Series on Testing and Assessment, N° 19, ENV/JM/MONO(2000)7, OECD, Paris.
- (24) OECD (2009), Detailed Review Paper on Transgenic Rodent Mutation Assays, Series on Testing and Assessment, N° 103, ENV/JM/MONO(2009)7, OECD, Paris.
- (25) Piegorsch, W.W., B.H. Margolin, M.D. Shelby, A. Johnson, J.E. French, R.W. Tennant and K.R. Tindall (1995), 'Study Design and Sample Sizes for a lacI Transgenic Mouse Mutation Assay', *Environ. Mol. Mutagen.*, 25(3): 231–245.
- (26) Piegorsch, W.W., A.C. Lockhart, G.J. Carr, B.H. Margolin, T. Brooks, ... G.R. Douglas, U.M. Liegibel, T. Suzuki, V. Thybaud, J.H. van Delft and N.J. Gorelick (1997), 'Sources of Variability in Data from a Positive Selection lacZ Transgenic Mouse Mutation Assay: an Interlaboratory Study', *Mutation. Res.*, 388(2–3): 249–289.
- (27) Singer, T.M., I.B. Lambert, A. Williams, G.R. Douglas and C.L. Yauk (2006), 'Detection of Induced Male Germline Mutation: Correlations and Comparisons between Traditional Germline Mutation Assays, Transgenic Rodent Assays and Expanded Simple Tandem Repeat Instability Assays', *Mutation. Res.*, 598: 164-193.
- (28) Toyoda-Hokaiwado, N., T. Inoue, K. Masumura, H. Hayashi, Y. Kawamura, Y. Kurata, M. Takamune, M. Yamada, H. Sanada, T. Umemura, A. Nishikawa and T. Nohmi (2010), 'Integration of *in vivo* Genotoxicity and Short-term Carcinogenicity Assays using F344 gpt delta Transgenic Rats: *in vivo* Mutagenicity of 2,4-diaminotoluene and 2,6-diaminotoluene Structural Isomers', *Toxicol. Sci.*, 114(1): 71-78.
- (29) Thybaud, V., S. Dean, T. Nohmi, J. de Boer, G.R. Douglas, B.W. Glickman, N.J. Gorelick, J.A. Heddle, R.H. Heflich, I. Lambert, H.-J. Martus, J.C. Mirsalis, T. Suzuki and N. Yajima (2003), '*In vivo* Transgenic Mutation Assays', *Mutation Res.*, 540: 141-151.
- (30) Vijg, J. and G.R. Douglas (1996), 'Bacteriophage  $\lambda$  and Plasmid lacZ Transgenic Mice for studying Mutations *in vivo*' in: G. Pfeifer (ed.), *Technologies for Detection of DNA Damage and Mutations, Part II*, Plenum Press, New York, NY, USA, pp. 391–410.

▼ M5

- (31) Yauk, C.L., J.D. Gingerich, L. Soper, A. MacMahon, W.G. Foster and G.R. Douglas (2005), 'A *lacZ* Transgenic Mouse Assay for the Detection of Mutations in Follicular Granulosa Cells', *Mutation Res.*, 578(1-2): 117-123.
- (32) Hayashi, M., K. Dearfield, P. Kasper, D. Lovell, H.-J. Martus, V. Thybaud (2011), 'Compilation and Use of Genetic Toxicity Historical Control Data', *Mutation Res.*, doi:10.1016/j.mrgentox.2010.09.007.
- (33) OECD (2011), *Retrospective Performance Assessment of OECD Test Guideline on Transgenic Rodent Somatic and Germ Cell Gene Mutation Assays*, Series on Testing and Assessment, N° 145, ENV/JM/MONO(2011)20, OECD, Paris.
- (34) Clermont, Y. (1972), 'Kinetics of spermatogenesis in mammals seminiferous epithelium cycle and spermatogonial renewal'. *Physiol. Rev.* 52: 198-236.
- (35) Robaire, B., Hinton, B.T., and Oregbin-Crist, M.-C. (2006), 'The Epididymis', in Neil, J.D., Pfaff, D.W., Chalis, J.R.G., de Kretser, D.M., Richards, J.S., and P. M. Wassarman (eds.), *Physiology of Reproduction*, Elsevier, the Netherlands, pp. 1071-1148.
- (36) Russell, L.B. (2004), 'Effects of male germ-cell stage on the frequency, nature, and spectrum of induced specific-locus mutations in the mouse', *Genetica*, 122: 25-36.

▼ M5*Appendix*

## DEFINITIONS:

**Administration period:** the total period during which an animal is dosed.

**Base pair substitution:** a type of mutation that causes the replacement of a single DNA nucleotide base with another DNA nucleotide base.

**Capsid:** the protein shell that surrounds a virus particle.

**Chemical:** a substance or a mixture.

**Clonal expansion:** the production of many cells from a single (mutant) cell.

**Colony-forming unit (cfu):** a measure of viable bacterial numbers.

**Concatamer:** a long continuous biomolecule composed of multiple identical copies linked in series.

**Cos site:** a 12-nucleotide segment of single-stranded DNA that exists at both ends of the bacteriophage lambda's double-stranded genome.

**Deletion:** a mutation in which one or more (sequential) nucleotides is lost by the genome.

**Electroporation:** the application of electric pulses to increase the permeability of cell membranes.

**Endogenous gene:** a gene native to the genome.

**Extrabinomial variation:** greater variability in repeat estimates of a population proportion than would be expected if the population had a binomial distribution.

**Frameshift mutation:** a genetic mutation caused by insertions or deletions of a number of nucleotides that is not evenly divisible by three within a DNA sequence that codes for a protein/peptide.

**Insertion:** the addition of one or more nucleotide base pairs into a DNA sequence.

**Jackpot:** a large number of mutants that arose through clonal expansion from a single mutation.

**Large deletions:** deletions in DNA of more than several kilobases (which are effectively detected with the Spi<sup>-</sup> selection and the lacZ plasmid assays).

**Ligation:** the covalent linking of two ends of DNA molecules using DNA ligase.

**Mitogen:** a chemical that stimulates a cell to commence cell division, triggering mitosis (i.e. cell division).

**Neutral gene:** a gene that is not affected by positive or negative selective pressures.

**Packaging:** the synthesis of infective phage particles from a preparation of phage capsid and tail proteins and a concatamer of phage DNA molecules. Commonly used to package DNA cloned onto a lambda vector (separated by cos sites) into infectious lambda particles.

**Packaging efficiency:** the efficiency with which packaged bacteriophages are recovered in host bacteria.

**▼ M5**

**Plaque forming unit (pfu):** a measure of viable bacteriophage numbers.

**Point mutation:** a general term for a mutation affecting only a small sequence of DNA including small insertions, deletions, and base pair substitutions.

**Positive selection:** a method that permits only mutants to survive.

**Reporter gene:** a gene whose mutant gene product is easily detected.

**Sampling time:** the end of the period of time, prior to sacrifice, during which the chemical is not administered and during which unprocessed DNA lesions are fixed into stable mutations.

**Shuttle vector:** a vector constructed so that it can propagate in two different host species; accordingly, DNA inserted into a shuttle vector can be tested or manipulated in two different cell types or two different organisms.

**Test chemical:** Any substance or mixture tested using this test method.

**Transgenic:** of, relating to, or being an organism whose genome has been altered by the transfer of a gene or genes from another species.

**▼B****PART C: METHODS FOR THE DETERMINATION OF ECOTOXICITY**

## TABLE OF CONTENTS

C.1.	ACUTE TOXICITY FOR FISH
C.2.	<i>DAPHNIA</i> SP. ACUTE IMMOBILISATION TEST
C.3.	FRESHWATER ALGAE AND CYANOBACTERIA, GROWTH INHIBITION TEST
C.4.	DETERMINATION OF 'READY' BIODEGRADABILITY
PART I.	GENERAL CONSIDERATIONS
PART II.	DOC DIE-AWAY TEST (Method C.4-A)
PART III.	MODIFIED OECD SCREENING TEST (Method C.4-B)
PART IV.	CO <sub>2</sub> EVOLUTION TEST (Method C.4-C)
PART V.	MANOMETRIC RESPIROMETRY TEST (Method C.4-D)
PART VI.	CLOSED BOTTLE TEST (Method C.4-E)
PART VII.	M.I.T.I. TEST (Method C.4-F)
C.5.	DEGRADATION — BIOCHEMICAL OXYGEN DEMAND
C.6.	DEGRADATION — CHEMICAL OXYGEN DEMAND
C.7.	DEGRADATION — ABIOTIC DEGRADATION: HYDROLYSIS AS A FUNCTION OF PH
C.8.	TOXICITY FOR EARTHWORMS
C.9.	BIODEGRADATION — ZAHN-WELLENS TEST
C.10.	SIMULATION TESTAEROBIC SEWAGE TREATMENT: C.10-A: ACTIVATED SLUDGE UNITS — C.10-B: BIOFILMS
C.11.	BIODEGRADATION — ACTIVATED SLUDGE RESPIRATION INHIBITION
C.12.	BIODEGRADATION — MODIFIED SCAS TEST
C.13.	BIOCONCENTRATION: FLOW-THROUGH FISH TEST
C.14.	FISH JUVENILE GROWTH TEST
C.15.	FISH, SHORT-TERM TOXICITY TEST ON <i>EMBRYO</i> AND SAC-FRY STAGES
C.16.	HONEYBEES — ACUTE ORAL TOXICITY TEST
C.17.	HONEYBEES — ACUTE CONTACT TOXICITY TEST
C.18.	ADSORPTION/DESORPTION USING A BATCH EQUILIBRIUM METHOD

**▼ B**

- C.19. ESTIMATION OF THE ADSORPTION COEFFICIENT ( $K_{OC}$ ) ON SOIL AND ON SEWAGE SLUDGE USING HIGH PERFORMANCE LIQUID CHROMATOGRAPHY (HPLC)
- C.20. *DAPHNIA MAGNA* REPRODUCTION TEST
- C.21. SOIL MICROORGANISMS: NITROGEN TRANSFORMATION TEST
- C.22. SOIL MICROORGANISMS: CARBON TRANSFORMATION TEST
- C.23. AEROBIC AND ANAEROBIC TRANSFORMATION IN SOIL
- C.24. AEROBIC AND ANAEROBIC TRANSFORMATION IN AQUATIC SEDIMENT SYSTEMS
- C.25. AEROBIC MINERALISATION IN SURFACE WATER — SIMULATION BIODEGRADATION TEST
- C.26. *LEMNA* SP. GROWTH INHIBITION TEST

**▼ M4**

- C.27. SEDIMENT-WATER CHIRONOMID TOXICITY TEST USING SPIKED SEDIMENT
- C.28. SEDIMENT-WATER CHIRONOMID TOXICITY TEST USING SPIKED WATER
- C.29. READY BIODEGRADABILITY — CO<sub>2</sub> IN SEALED VESSELS (HEADSPACE TEST)
- C.30. BIOACCUMULATION IN TERRESTRIAL OLIGOCHAETES

**▼B****C.1. ACUTE TOXICITY FOR FISH****1. METHOD****1.1. INTRODUCTION**

The purpose of this test is to determine the acute lethal toxicity of a substance to fish in fresh water. It is desirable to have, as far as possible, information on the water solubility, vapour pressure, chemical stability, dissociation constants and biodegradability of the substance to help in the selection of the most appropriate test method (static, semi-static or flow-through) for ensuring satisfactorily constant concentrations of the test substance over the period of the test.

Additional information (for instance structural formula, degree of purity, nature and percentage of significant impurities, presence and amounts of additives, and n-octanol/water partition coefficient) should be taken into consideration in both the planning of the test and interpretation of the results.

**1.2. DEFINITIONS AND UNITS**

Acute toxicity is the discernible adverse effect induced in an organism within a short time (days) of exposure to a substance. In the present test, acute toxicity is expressed as the median lethal concentration ( $LC_{50}$ ) that is the concentration in water which kills 50 % of a test batch of fish within a continuous period of exposure which must be stated.

All concentrations of the test substance are given in weight by volume (milligrams per litre). They may also be expressed as weight by weight ( $mg/kg^{-1}$ ).

**1.3. REFERENCE SUBSTANCES**

A reference substance may be tested as a means of demonstrating that under the laboratory test conditions the response of tested species have not changed significantly.

No reference substances are specified for this test.

**1.4. PRINCIPLE OF THE TEST METHOD**

A limit test may be performed at 100 mg per litre in order to demonstrate that the  $LC_{50}$  is greater than this concentration.

The fish are exposed to the test substance added to water at a range of concentrations for a period of 96 hours. Mortalities are recorded at least at 24-hour intervals, and the concentrations killing 50 % of the fish ( $LC_{50}$ ) at each observation time are calculated where possible.

**1.5. QUALITY CRITERIA**

The quality criteria shall apply to the limit test as well as the full test method.

The mortality in the controls must not exceed 10 % (or one fish if less than ten are used) by the end of the test.

The dissolved oxygen concentration must have been more than 60 % of the air-saturation value throughout.

**▼B**

The concentrations of the test substance shall be maintained to within 80 % of the initial concentrations throughout the duration of the test.

For substances which dissolve easily in the test medium, yielding stable solutions i.e. those which will not to any significant extent volatilise, degrade, hydrolyze or adsorb, the initial concentration can be taken as being equivalent to the nominal concentration. Evidence shall be presented that the concentrations have been maintained throughout the test and that the quality criteria have been satisfied.

For substances that are:

- (i) poorly soluble in the test medium, or
- (ii) capable of forming stable emulsions or dispersions, or
- (iii) not stable in aqueous solutions,

the initial concentration shall be taken as the concentration measured in solution (or, if technically not possible, measured in the water column) at the start of the test. The concentration shall be determined after a period of equilibration but before the introduction of the test fish.

In any of these cases, further measurements must be made during the test to confirm the actual exposure concentrations or that the quality criteria have been met.

The pH should not vary by more than 1 unit.

## 1.6. DESCRIPTION OF THE TEST METHOD

Three types of procedure can be used:

### *Static test:*

Toxicity test in which no flow of test solution occurs. (Solutions remain unchanged throughout the duration of the test.)

### *Semi-static test:*

Test without flow of test solution, but with regular batchwise renewal of test solutions after prolonged periods (e.g. 24 hours).

### *Flow-through test:*

Toxicity test in which the water is renewed constantly in the test chambers, the chemical under test being transported with the water used to renew the test medium.

## 1.6.1. Reagents

### 1.6.1.1. Solutions of test substances

Stock solutions of the required strength are prepared by dissolving the substance in deionised water or water according to 1.6.1.2.

The chosen test concentrations are prepared by dilution of the stock solution. If high concentrations are tested, the substance may be dissolved in the dilution water directly.



**▼ B**

The substances should normally only be tested up to the limit of solubility. For some substances (e.g. substances having low solubility in water, or high  $P_{ow}$ , or those forming stable dispersion rather than true solution in water), it is acceptable to run a test concentration above the solubility limit of the substance to ensure that the maximum soluble/stable concentration has been obtained. It is important, however, that this concentration will not otherwise disturb the test system (e.g. film of the substance on the water surface preventing the oxygenation of the water, etc.).

Ultrasonic dispersion, organic solvents, emulsifiers or dispersants may be used as an aid to prepare stock solutions of substances with low aqueous solubility or to help to disperse these substances in the test medium. When such auxiliary substances are used, all test concentrations should contain the same amount of auxiliary substance, and additional control fish should be exposed to the same concentration of the auxiliary substance as that used in the test series. The concentration of such auxiliaries should be minimised, but in no case should exceed 100 mg per litre in the test medium.

The test should be carried out without adjustment of the pH. If there is evidence of marked change in the pH, it is advised that the test should be repeated with pH adjustment and the results reported. In that case, the pH value of the stock solution should be adjusted to the pH value of the dilution water unless there are specific reasons not to do so. HCl and NaOH are preferred for this purpose. This pH adjustment should be made in such a way that the concentration of test substance in the stock solution is not changed to any significant extent. Should any chemical reaction or physical precipitation of the test compound be caused by the adjustment, this should be reported.

**1.6.1.2. Holding and dilution water**

Orinking-water supply (uncontaminated by potentially harmful concentrations of chlorine, heavy metals or other substances), good-quality natural water or reconstituted water (See Appendix 1) may be used. Waters with a total hardness of between 10 and 250 mg per litre (as  $CaCO_3$ ) and with a pH from 6,0 to 8,5 are preferred.

**1.6.2. Apparatus**

All apparatus must be made of chemically inert material:

- automatic dilution system (for flow-through test),
- oxygen meter,
- equipment for determination of hardness of water,
- adequate apparatus for temperature control,
- pH meter.

**1.6.3. Test fish**

The fish should be in good health and free from any apparent malformation.

**▼ B**

The species used should be selected on the basis of practical criteria, such as their ready availability throughout the year, ease of maintenance, convenience for testing, relative sensitivity to chemicals, and any economic, biological or ecological factors which have any bearing. The need for comparability of the data obtained and existing international harmonisation (reference 1) should also be borne in mind when selecting the fish species.

A list of fish species which are recommended for the performance of this test is given in Appendix 2; Zebra fish and rainbow trout are the preferred species.

1.6.3.1.  *Holding*

Test fish should preferably come from a single stock of similar length and age. The fish must be held for at least 12 days, in the following conditions:

 *loading:*

appropriate to the system (recirculation or flow-through) and the fish species,

 *water:*

see 1.6.1.2,

 *light:*

12 to 16 hours illumination daily,

 *dissolved oxygen concentration:*

at least 80 % of air-saturation value,

 *feeding:*

three times per week or daily, ceasing 24 hours before the start of the test.

1.6.3.2.  *Mortality*

Following a 48-hour settling-in period, mortalities are recorded and the following criteria applied:

— greater than 10 % of population in seven days:

rejection of entire batch,

— between 5 and 10 % of population:

holding period continued for seven additional days.

If no further mortalities occur, the batch is acceptable, otherwise it must be rejected,

— less than 5 % of population:

acceptance of the batch.

1.6.4.  **Adaptation**

All fish must be exposed to water of the quality and the temperature to be used in the test for at least seven days before they are used.

**▼B****1.6.5. Test procedure**

A range-finding test can precede a definitive test, in order to obtain information about the range of concentrations to be used in the main test.

One control without the test substance is run and, if relevant, one control containing the auxiliary substance is also run, in addition to the test series.

Depending on the physical and chemical properties of the test compound, a static, semi-static, or a flow-through test should be selected as appropriate, to fulfil the quality criteria.

Fish are exposed to the substance as described below:

- duration: 96 hours,
- number of animals: at least seven per concentration,
- tanks: of suitable capacity in relation to the recommended loading,
- loading: maximum loading of 1 g per litre for static and semi-static tests is recommended; for flow-through systems, higher loading is acceptable,
- test concentration: At least five concentrations differing by a constant factor not exceeding 2,2 and as far as possible spanning the range of 0 to 100 % mortality,
- water: see 1.6.1.2,
- light: 12 to 16 hours illumination daily,
- temperature: appropriate to the species (Appendix 2) but within  $\pm 1$  °C within any particular test,
- dissolved oxygen concentration: not less than 60 % of the air-saturation value at the selected temperature,
- feeding: none.

The fish are inspected after the first two to four hours and at least at 24-hour intervals. Fish are considered dead if touching of the caudal peduncle produces no reaction, and no breathing movements are visible. Dead fish are removed when observed and mortalities are recorded. Records are kept of visible abnormalities (e.g. loss of equilibrium, changes in swimming behaviour, respiratory function, pigmentation, etc.).

Measurements of pH, dissolved oxygen and temperature must be carried out daily.

*Limit test*

Using the procedures described in this test method, a limit test may be performed at 100 mg per litre in order to demonstrate that the  $LC_{50}$  is greater than this concentration.

If the nature of the substance is such that a concentration of 100 mg per litre in the test water cannot be attained, the limit test should be performed at a concentration equal to the solubility of the substance (or the maximum concentration forming a stable dispersion) in the medium used (see also point 1.6.1.1).

**▼ B**

The limit test should be performed using seven to 10 fish, with the same number in the control(s). (Binomial theory dictates that when 10 fish are used with zero mortality, there is a 99,9 % confidence that the  $LC_{50}$  is greater than the concentration used in the limit test. With 7, 8 or 9 fish, the absence of mortality provides at least 99 % confidence that the  $LC_{50}$  is greater than the concentration used.)

If mortalities occur, a full study must be carried out. If sublethal effects are observed, these should be recorded.

**2. DATA AND EVALUATION**

For each period where observations were recorded (24, 48, 72 and 96 hours), plot percentage mortality for each recommended exposure period against concentration on logarithmic-probability paper.

When possible and for each observation time, the  $LC_{50}$  and the confidence limits ( $p = 0,05$ ) should be estimated using standard procedures; these values should be rounded off to one, or at most two significant figures (examples of rounding off to two figures: 170 for 173,5; 0,13 for 0,127; 1,2 for 1,21).

In those cases where the slope of the concentration/percentage response curve is too steep to permit calculation of the  $LC_{50}$ , a graphical estimate of this value is sufficient.

When two consecutive concentrations, at a ratio of 2,2 give only 0 and 100 % mortality, these two values are sufficient to indicate the range within which the  $LC_{50}$  falls.

If it is observed that the stability or homogeneity of the test substance cannot be maintained, this should be reported and care should be taken in the interpretation of the results.

**3. REPORTING**

The test report shall, if possible, include the following information:

- information about test fish (scientific name, strain, supplier, any pretreatment, size and number used in each test concentration),
- dilution-water source and major chemical characteristics (pH, hardness, temperature),
- in the case of a substance of low aqueous solubility, the method of preparation of stock and test solutions,
- concentration of any auxiliary substances,
- list of the concentrations used and any available information on the stability at the concentrations of the tested chemical in the test solution,
- if chemical analyses are performed, methods used and results obtained,
- results of the limit test if conducted,
- reasons for the choice and details of the test procedure used (e.g. static, semi-static, dosing rate, flow-through rate, whether aerated, fish loading, etc.),

**▼B**

- description of test equipment,
- lighting regime,
- dissolved oxygen concentrations, pH values and temperatures of the test solutions every 24 hours,
- evidence that the quality criteria have been fulfilled,
- a table showing the cumulative mortality at each concentration and the control (and control with the auxiliary substance if required) at each of the recommended observation times,
- graph of the concentration/percentage response curve at the end of the test,
- if possible, the LC<sub>50</sub> values at each of the recommended observation times (with 95 % confidence limits),
- statistical procedures used for determining the LC<sub>50</sub> values,
- if a reference substance is used, the results obtained,
- highest test concentration causing no mortality within the period of the test,
- lowest test concentration causing 100 % mortality within the period of the test.

**4. REFERENCES**

- (1) OECD, Paris, 1981, Test Guideline 203, Decision of the Council C(81) 30 final and updates.
- (2) AFNOR — Determination of the acute toxicity of a substance to *Brachydanio rerio* — Static and Flow Through methods — NFT 90-303 June 1985.
- (3) AFNOR- Determination of the acute toxicity of a substance to *Salmo gairdneri* — Static and Flow — Through methods — NFT 90-305 June 1985.
- (4) ISO 7346/1,2 and/3 — Water Quality — Determination of the acute lethal toxicity of substances to a fresh water fish (*Brachydanio rerio* Hamilton-Buchanan-Teleostei, Cyprinidae). Part 1: Static method. Part 2: Semi-static method. Part 3: Flow-through method.
- (5) Eidgenössisches Department des Innern, Schweiz: Richtlinien für Probenahme und Normung von Wasseruntersuchungsmethoden -Part II 1974.
- (6) DIN Testverfahren mit Wasserorganismen, 38 412 (11) und 1 (15).
- (7) JIS K 0102, Acute toxicity test for fish.
- (8) NEN 6506- Water -Bepaling van de akute toxiciteit met behulp van *Poecilia reticulata*, 1980.
- (9) Environmental Protection Agency, Methods for the acute toxicity tests with fish, macroinvertebrates and amphibians. The Committee on Methods for Toxicity Tests with Aquatic Organisms, Ecological Research Series EPA-660-75-009, 1975.
- (10) Environmental Protection Agency, Environmental monitoring and support laboratory, Office of Research and Development, EPA-600/4-78-012, January 1978.

**▼B**

- 11) Environmental Protection Agency, Toxic Substance Control, Part IV, 16 March 1979.
- 12) Standard methods for the examination of water and wastewater, fourteen edition, APHA-AWWA-WPCF, 1975.
- 13) Commission of the European Communities, Inter-laboratory test programme concerning the study of the ecotoxicity of a chemical substance with respect to the fish. EEC Study D.8368, 22 March 1979.
- 14) Verfahrensvorschlag des Umweltbundesamtes zum akuten Fisch-Test. Rudolph, P. und Boje, R. Okotoxikologie, Grundlagen für die okotoxikologische Bewertung von Umweltchemikalien nach dem Chemikaliengesetz, ecomed 1986.
- 15) Litchfield, J. T. and Wilcoxon, F., A simplified method for evaluating dose effects experiments, J. Pharm. tExp. Therap., 1949, vol. 96,99.
- 16) Finney, D.J. Statistical Methods in Biological Assay. Griffin, Weycombe, U.K., 1978.
- 17) Sprague, J.B. Measurement of pollutant toxicity to fish. Bioassay methods for acute toxicity. Water Res., 1969, vol. 3,793-821.
- 18) Sprague, J.B. Measurement of pollutant toxicity to fish. II Utilising and applying bioassay results. Water Res. 1970, vol. 4, 3-32.
- 19) Stephan, C.E. Methods for calculating an LC<sub>50</sub>. In Aquatic Toxicology and Hazard Evaluation (edited by F.I. Mayer and J.L. Hamelink). American Society for Testing and Materials, ASTM STP 634, 1977, 65-84.
- 20) Stephan, C.E., Busch, K.A., Smith, R., Burke, J. and Andrews, R.W. A computer program for calculating an LC<sub>50</sub>. US EPA.

**▼ B***Appendix 1***Reconstituted water***Example of a suitable dilution water*

All chemicals must be of analytical grade.

The water should be good-quality distilled water, or deionised water with a conductivity of less than  $5 \mu\text{Scm}^{-1}$ .

Apparatus for distillation of water must not contain any parts made of copper.

*Stock solutions*

$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$  (calcium chloride dihydrate): 11,76 g

Dissolve in, and make up to 1 litre with water.

$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$  (magnesium sulphate heptahydrate): 4,93 g

Dissolve in, and make up to 1 litre with water.

$\text{NaHCO}_3$  (sodium hydrogen carbonate): 2,59 g

Dissolve in, and make up to 1 litre with water.

KCl (potassium chloride): 0,23 g

Dissolve in, and make up to 1 litre with water.

*Reconstituted dilution water*

Mix 25 ml of each of the four stock solutions and make up to 1 litre with water.

Aerate until the dissolved oxygen concentration equals the air-saturation value.

The pH should be  $7,8 \pm 0,2$ .

If necessary adjust the pH with NaOH (sodium hydroxide) or HCl (hydrochloric acid).

The dilution water so prepared is set aside for about 12 hours and must not be further aerated.

The sum of the Ca and Mg ions in this solution is 2,5 mmol per litre. The ratio of Ca:Mg ions is 4:1 and of Na:K ions is 10:1. The total alkalinity of this solution is 0,8 mmol per litre.

Any deviation in the preparation of the dilution water must not change the composition or properties of the water.

▼ **B**

## Appendix 2

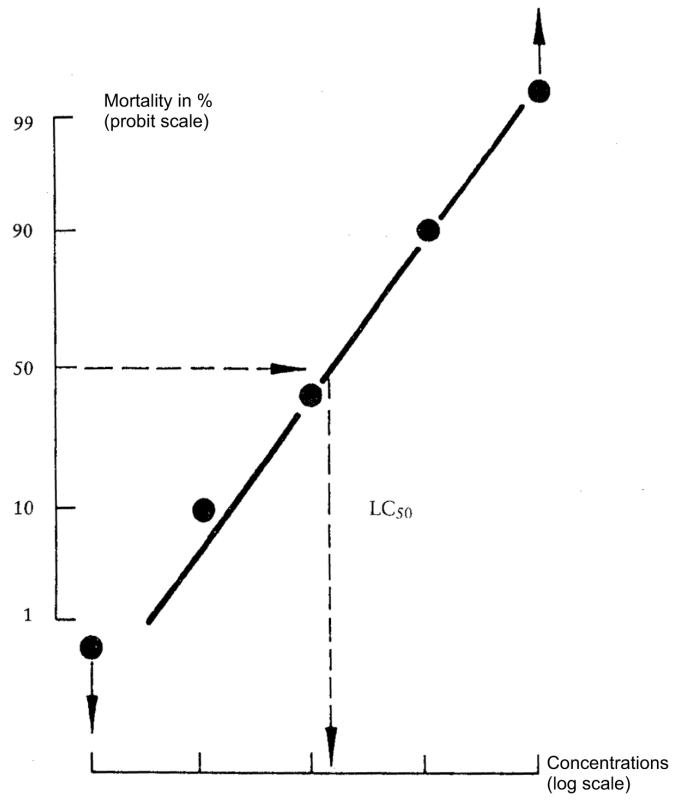
**Fish species recommended for testing**

Recommended species	Recommended range of test temperature (°C)	Recommended total length of test animal (cm)
<i>Brachydanio rerio</i> ( <i>Teleostei, Cyprinidae</i> ) (Hamilton-Buchanan) Zebra-fish	20 to 24	3,0 ± 0,5
<i>Pimephales promelas</i> ( <i>Teleostei, Cyprinidae</i> ) (Rafinesque) Fathead minnow	20 to 24	5,0 ± 2,5
<i>Cyprinus carpio</i> ( <i>Teleostei, Cyprinidae</i> ) (Linnaeus 1758) Common carp	20 to 24	6,0 ± 2,0
<i>Oryzias latipes</i> ( <i>Teleostei, Poeciliidae</i> ) Cyprinodontidae (Tomminck and Schlege 1850) Red killifish	20 to 24	3,0 ± 1,0
<i>Poecilia reticulata</i> ( <i>Teleostei, Poeciliidae</i> ) (Peters 1859) Guppy	20 to 24	3,0 ± 1,0
<i>Lepomis macrochirus</i> ( <i>Teleostei, Centrarchidae</i> ) (Rafinesque Linnaeus 1758) Bluegill	20 to 24	5,0 ± 2,0
<i>Onchorhynchus mykiss</i> ( <i>Teleostei, Salmonidae</i> ) (Walbaum 1988) Rainbow trout	12 to 17	6,0 ± 2,0
<i>Leuciscus idus</i> ( <i>Teleostei, Cyprinidae</i> ) (Linnaeus 1758) Golden Orfe	20 to 24	6,0 ± 2,0

**Collection**

The fish listed above are easy to rear and/or are widely available throughout the year. They are capable of being bred and cultivated either in fish farms or in the laboratory, under disease — and parasite — controlled conditions, so that the test animal will be healthy and of known parentage. These fish are available in many parts of the world.



**▼ B***Appendix 3***Example of concentration: percentage mortality**Example of determination of  $LC_{50}$  using log-probit paper

**▼B****C.2. DAPHNIA SP. ACUTE IMMOBILISATION TEST****1. METHOD**

This acute immobilisation testing method is equivalent to the OECD TG 202 (2004).

**1.1. INTRODUCTION**

This method describes an acute toxicity test to assess effects of chemicals towards daphnids. Existing test methods were used to the extent possible (1)(2)(3).

**1.2. DEFINITIONS**

In the context of this method, the following definitions are used:

**EC<sub>50</sub>**: is the concentration estimated to immobilise 50 % of the daphnids within a stated exposure period. If another definition is used, this must be reported, together with its reference.

**Immobilisation**: those animals that are not able to swim within 15 seconds, after gentle agitation of the test vessel are considered to be immobilised (even if they can still move their antennae).

**1.3. PRINCIPLE OF THE TEST METHOD**

Young daphnids, aged less than 24 hours at the start of the test, are exposed to the test substance at a range of concentrations for a period of 48 hours. Immobilisation is recorded at 24 hours and 48 hours and compared with control values. The results are analysed in order to calculate the EC<sub>50</sub> at 48h (see Section 1.2 for definitions). Determination of the EC<sub>50</sub> at 24h is optional.

**1.4. INFORMATION ON THE TEST SUBSTANCE**

The water solubility and the vapour pressure of the test substance should be known and a reliable analytical method for the quantification of the substance in the test solutions with reported recovery efficiency, and limit of determination should be available. Useful information includes the structural formula, purity of the substance, stability in water or light, P<sub>ow</sub> and results of a test for ready biodegradability (see method C.4).

Note: guidance for testing substances with physical chemical properties that made them difficult to test is provided in (4).

**1.5. REFERENCE SUBSTANCES**

A reference substance may be tested for EC<sub>50</sub> as a means of assuring that the test conditions are reliable. Toxicants used in international ring-tests (1)(5) are recommended for this purpose (1). Test(s) with a reference substance should be done preferably every month and at least twice a year.

(1) The results of these inter laboratory tests and a Technical Corrigendum to ISO 6341 give an EC<sub>50</sub> — 24 h of the potassium dichromate (K<sub>2</sub>Cr<sub>2</sub>O<sub>7</sub>) within the range 0,6 mg/l to 1,7 mg/l.

**▼ B**

## 1.6. QUALITY CRITERIA

For a test to be valid, the following performance criteria apply:

- in the controls, including the control containing the solubilising agent, not more than 10 % of the daphnids should have been immobilised;
- the dissolved oxygen concentration at the end of the test should be  $\geq 3$  mg/l in control and test vessels.

Note: For the first criterion, not more than 10 % of the control daphnids should show immobilisation or other signs of disease or stress, for example, discoloration, unusual behaviour such as trapping at surface of water.

## 1.7. DESCRIPTION OF THE TEST METHOD

1.7.1. **Apparatus**

Test vessels and other apparatus that will come into contact with the test solutions should be made entirely of glass or other chemically inert material. Test vessels will normally be glass test tubes or beakers; they should be cleaned before each use using standard laboratory procedures. Test vessels should be loosely covered to reduce the loss of water due to evaporation and to avoid the entry of dust into the solutions. Volatile substances should be tested in completely filled closed vessels, large enough to prevent oxygen becoming limiting or too low (see Section 1.6 and first paragraph of Section 1.8.3).

In addition some or all of the following equipment will be used: oxygen-meter (with microelectrode or other suitable equipment for measuring dissolved oxygen in low volumes samples); pH-meter; adequate apparatus for temperature control; equipment for the determination of total organic carbon concentration (TOC); equipment for the determination of chemical oxygen demand (COD); equipment for the determination of hardness, etc.

1.7.2. **Test organism**

*Daphnia magna* Straus is the preferred test species although other suitable *Daphnia* species can be used in this test (e.g. *Daphnia pulex*). At the start of the test, the animals should be less than 24 hours old and to reduce variability, it is strongly recommended they are not first brood progeny. They should be derived from a healthy stock (i.e. showing no signs of stress such as high mortality, presence of males and ephippia, delay in the production of the first brood, discoloured animals, etc.). All organisms used for a particular test should have originated from cultures established from the same stock of daphnids. The stock animals must be maintained in culture conditions (light, temperature, medium) similar to those to be used in the test. If the daphnids culture medium to be used in the test is different from that used for routine daphnids culture, it is good practice to include a pre-test acclimation period. For that, brood daphnids should be maintained in dilution water at the test temperature for at least 48 hours prior to the start of the test.

**▼B****1.7.3. Holding and dilution water**

Natural water (surface or ground water), reconstituted water or dechlorinated tap water are acceptable as holding and dilution water if daphnids will survive in it for the duration of the culturing, acclimation and testing without showing signs of stress. Any water which conforms to the chemical characteristics of an acceptable dilution water as listed in Appendix 1 is suitable as a test water. It should be of constant quality during the period of the test. Reconstituted water can be made up by adding specific amounts of reagents of recognised analytical grade to deionised or distilled water. Examples of reconstituted water are given in (1) (6) and in Appendix 2. Note that media containing known chelating agents, such as M4 and M7 media in Appendix 2, should be avoided for testing substances containing metals. The pH should be in the range of 6 to 9. Hardness between 140 and 250 mg/l (as CaCO<sub>3</sub>) is recommended for *Daphnia magna*, while lower hardness may be also appropriate for other *Daphnia* species. The dilution water may be aerated prior to use for the test so that the dissolved oxygen concentration has reached saturation.

If natural water is used, the quality parameters should be measured at least twice a year or whenever it is suspected that these characteristics may have changed significantly (see previous paragraph and Appendix 1). Measurements of heavy metals (e.g. Cu, Pb, Zn, Hg, Cd, Ni) should also be made. If dechlorinated tap water is used, daily chlorine analysis is desirable. If the dilution water is from a surface or ground water source, conductivity and total organic carbon (TOC) or chemical oxygen demand (COD) should be measured.

**1.7.4. Test solutions**

Test solutions of the chosen concentrations are usually prepared by dilution of a stock solution. Stock solutions should preferably be prepared by dissolving the test substance in the dilution water. As far as possible, the use of solvents, emulsifiers or dispersants should be avoided. However, such compounds may be required in some cases in order to produce a suitably concentrated stock solution. Guidance for suitable solvents, emulsifiers and dispersants is given in (4). In any case, the test substance in the test solutions should not exceed the limit of solubility in the dilution water.

The test should be carried out without the adjustment of pH. If the pH does not remain in the range 6-9, then a second test could be carried out, adjusting the pH of the stock solution to that of the dilution water before addition of the test substance. The pH adjustment should be made in such a way that the stock solution concentration is not changed to any significant extent and that no chemical reaction or precipitation of the test substance is caused. HCl and NaOH are preferred.

**▼ B**

## 1.8. PROCEDURE

1.8.1. **Conditions of exposure**1.8.1.1. *Test groups and controls*

Test vessels are filled with appropriate volumes of dilution water and solutions of test substance. Ratio of air/water volume in the vessel should be identical for test and control group. Daphnids are then placed into test vessels. At least 20 animals, preferably divided into four groups of five animals each, should be used at each test concentration and for the controls. At least 2 ml of test solution should be provided for each animal (i.e. a volume of 10 ml for five daphnids per test vessel). The test may be carried out using semi-static renewal or flow-through system when the concentration of the test substance is not stable.

One dilution-water control series and also, if relevant, one control series containing the solubilising agent must be run in addition to the treatment series.

1.8.1.2. *Test concentrations*

A range-finding test may be conducted to determine the range of concentrations for the definitive test unless information on toxicity of the test substance is available. For this purpose, the daphnids are exposed to a series of widely spaced concentrations of the test substance. Five daphnids should be exposed to each test concentration for 48 hours or less, and no replicates are necessary. The exposure period may be shortened (e.g. 24 hours or less) if data suitable for the purpose of the range-finding test can be obtained in less time.

At least five test concentrations should be used. They should be arranged in a geometric series with a separation factor preferably not exceeding 2,2. Justification should be provided if fewer than five concentrations are used. The highest concentration tested should preferably result in 100 % immobilisation, and the lowest concentration tested should preferably give no observable effect.

1.8.1.3. *Incubation conditions*

The temperature should be within the range of 18 °C and 22 °C, and for each single test it should be constant within  $\pm 1$  °C. A 16-hour light and eight-hour dark cycle is recommended. Complete darkness is also acceptable, especially for the test substances unstable in light.

The test vessels must not be aerated during the test. The test is carried out without adjustment of pH. The daphnids should not be fed during the test.

1.8.1.4. *Duration*

The test duration is 48 hours.

1.8.2. **Observations**

Each test vessel should be checked for immobilised daphnids at 24 and 48 hours after the beginning of the test (see Section 1.2 for definitions). In addition to immobility, any abnormal behaviour or appearance should be reported.

**▼B****1.8.3. Analytical measurements**

The dissolved oxygen and pH are measured at the beginning and end of the test in the control(s) and in the highest test substance concentration. The dissolved oxygen concentration in controls should be in compliance with the validity criterion (see Section 1.6). The pH should normally not vary by more than 1,5 units in any one test. The temperature is usually measured in control vessels or in ambient air and it should be recorded preferably continuously during the test or, as a minimum, at the beginning and end of the test.

The concentration of the test substance should be measured, as a minimum, at the highest and lowest test concentration, at the beginning and end of the test (4). It is recommended that results be based on measured concentrations. However, if evidence is available to demonstrate that the concentration of the test substance has been satisfactorily maintained within  $\pm 20\%$  of the nominal or measured initial concentration throughout the test, then the results can be based on nominal or measured initial values.

**1.9. LIMIT TEST**

Using the procedures described in this Method, a limit test may be performed at 100 mg/l of test substance or up to its limit of solubility in the test medium (whichever is the lower) in order to demonstrate that the  $EC_{50}$  is greater than this concentration. The limit test should be performed using 20 daphnids (preferably divided into four groups of five), with the same number in the control(s). If any immobilisation occurs, a full study should be conducted. Any observed abnormal behaviour should be recorded.

**2. DATA**

Data should be summarised in tabular form, showing for each treatment group and control, the number of daphnids used, immobilisation at each observation. The percentages immobilised at 24 hours and 48 hours are plotted against test concentrations. Data are analysed by appropriate statistical methods (e.g. probit analysis, etc.) to calculate the slopes of the curves and the  $EC_{50}$  with 95 % confidence limits ( $p = 0,05$ ) (7) (8).

Where the standard methods of calculating the  $EC_{50}$ , are not applicable to the data obtained, the highest concentration causing no immobility and the lowest concentration producing 100 % immobility should be used as an approximation for the  $EC_{50}$  (this being considered the geometric mean of these two concentrations).

**3. REPORTING****3.1. TEST REPORT**

The test report must include the following:

Test substance:

— physical nature and relevant physical-chemical properties,

**▼ B**

- chemical identification data, including purity.

Test species:

- source and species of *Daphnia*, supplier of source (if known) and the culture conditions used (including source, kind and amount of food, feeding frequency).

Test conditions:

- description of test vessels: type of vessels, volume of solution, number of daphnids per test vessel, number of test vessels (replicates) per concentration,
- methods of preparation of stock and test solutions including the use of any solvent or dispersants, concentrations used,
- details of dilution water: source and water quality characteristics (pH, hardness, Ca/Mg ratio, Na/K ratio, alkalinity, conductivity, etc.); composition of reconstituted water if used,
- incubation conditions: temperature, light intensity and periodicity, dissolved oxygen, pH, etc.

Results:

- the number and percentage of daphnids that were immobilised or showed any adverse effects (including abnormal behaviour) in the controls and in each treatment group, at each observation time and a description of the nature of the effects observed,
- results and date of test performed with reference substance, if available,
- the nominal test concentrations and the result of all analyses to determine the concentration of the test substance in the test vessels; the recovery efficiency of the method and the limit of determination should also be reported,
- all physical-chemical measurements of temperature, pH and dissolved oxygen made during the test,
- the EC<sub>50</sub> at 48 h for immobilisation with confidence intervals and graphs of the fitted model used for their calculation, the slopes of the dose-response curves and their standard error; statistical procedures used for determination of EC<sub>50</sub>; (these data items for immobilisation at 24 h should also be reported when they were measured),
- explanation for any deviation from the Testig Method and whether the deviation affected the test results.

#### 4. REFERENCES

- (1) ISO 6341. (1996). Water quality — Determination of the inhibition of the mobility of *Daphnia magna* Straus (Cladocera, Crustacea) — Acute toxicity test. Third edition, 1996.
- (2) EPA OPPTS 850.1010. (1996). Ecological Effects Test Guidelines — Aquatic Invertebrate Acute Toxicity Test, Freshwater Daphnids.

**▼ B**

- (3) Environment Canada. (1996) Biological test method. Acute Lethality Test Using *Daphnia* spp. EPS 1/RM/11. Environment Canada, Ottawa, Ontario, Canada.
- (4) Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures. OECD Environmental Health and Safety Publication. Series on Testing and Assessment. No 23. Paris 2000.
- (5) Commission of the European Communities. Study D8369. (1979). Inter-laboratory Test Programme concerning the study of the ecotoxicity of a chemical substance with respect to *Daphnia*.
- (6) OECD Guidelines for the Testing of Chemicals. Guideline 211: *Daphnia magna* Reproduction Test, adopted September 1998.
- (7) Stephan C.E. (1977). Methods for calculating an LC<sub>50</sub>. In Aquatic Toxicology and Hazard Evaluation (edited by F.I. Mayer and J.L. Hamelink). ASTM STP 634 — American Society for Testing and Materials. p. 65-84
- (8) Finney D.J. (1978). Statistical Methods in Biological Assay. 3<sup>rd</sup> ed. London. Griffin, Weycombe, UK.



**▼B***Appendix 1***SOME CHEMICAL CHARACTERISTICS OF AN ACCEPTABLE DILUTION WATER**

Substance	Concentration
Particulate matter	< 20 mg/l
Total organic carbon	< 2 mg/l
Unionised ammonia	< 1 µg/l
Residual chlorine	< 10 µg/l
Total organophosphorus pesticides	< 50 ng/l
Total organochlorine pesticides plus polychlorinated biphenyls	< 50 ng/l
Total organic chlorine	< 25 ng/l



## Appendix 2

## EXAMPLES OF SUITABLE RECONSTITUTED TEST WATER

## ISO Test water (I)

Stock solutions (single substance)		To prepare the reconstituted water, add the following volumes of stock solutions to 1 litre water (*)
Substance	Amount added to 1 litre water (*)	
Calcium chloride CaCl <sub>2</sub> , 2H <sub>2</sub> O	11,76 g	25 ml
Magnesium sulfate MgSO <sub>4</sub> , 7H <sub>2</sub> O	4,93 g	25 ml
Sodium bicarbonate NaHCO <sub>3</sub>	2,59 g	25 ml
Potassium chloride KCl	0,23 g	25 ml

(\*) Water of suitable purity, e.g. deionised, distilled or reverse osmosis with conductivity preferably not exceeding 10  $\mu\text{S}\cdot\text{cm}^{-1}$ .

## Elendt M7 and M4 medium

## Acclimation to Elendt M4 and M7 medium

Some laboratories have experienced difficulty in directly transferring *Daphnia* to M4 and M7 media. However, some success has been achieved with gradual acclimation, i.e. moving from own medium to 30 % Elendt, then to 60 % Elendt and then to 100 % Elendt. The acclimation periods may need to be as long as one month.

## Preparation

## Trace element

Separate stock solutions (I) of individual trace elements are first prepared in water of suitable purity, e.g. deionised, distilled or reverse osmosis. From these different stock solutions (I) a second single stock solution (II) is prepared, which contains all trace elements (combined solution), i.e.:

Stock solution(s) I (single substance)	Amount added to water (mg/l)	Concentration (related to medium M4)	To prepare the combined stock solution II, add the following amount of stock solution I to water (ml/l)	
			M4	M7
H <sub>3</sub> BO <sub>3</sub>	57 190	20 000-fold	1,0	0,25
MnCl <sub>2</sub> .4H <sub>2</sub> O	7 210	20 000-fold	1,0	0,25
LiCl	6 120	20 000-fold	1,0	0,25

## ▼ B

Stock solution(s) I (single substance)	Amount added to water (mg/l)	Concentration (related to medium M4)	To prepare the combined stock solution II, add the following amount of stock solution I to water (ml/l)	
			M4	M7
RbCl	1 420	20 000-fold	1,0	0,25
SrCl <sub>2</sub> ·6H <sub>2</sub> O	3 040	20 000-fold	1,0	0,25
NaBr	320	20 000-fold	1,0	0,25
Na <sub>2</sub> MoO <sub>4</sub> ·2H <sub>2</sub> O	1 230	20 000-fold	1,0	0,25
CuCl <sub>2</sub> ·2H <sub>2</sub> O	335	20 000-fold	1,0	0,25
ZnCl <sub>2</sub>	260	20 000-fold	1,0	1,0
CoCl <sub>2</sub> ·6H <sub>2</sub> O	200	20 000-fold	1,0	1,0
KI	65	20 000-fold	1,0	1,0
Na <sub>2</sub> SeO <sub>3</sub>	43,8	20 000-fold	1,0	1,0
NH <sub>4</sub> VO <sub>3</sub>	11,5	20 000-fold	1,0	1,0
Na <sub>2</sub> EDTA·2H <sub>2</sub> O	5 000	2 000-fold	—	—
FeSO <sub>4</sub> ·7H <sub>2</sub> O	1 991	2 000-fold	—	—

Both Na<sub>2</sub> EDTA and FeSO<sub>4</sub> solutions are prepared singly, poured together and autoclaved immediately.

This gives:

2 l Fe-EDTA solution		1 000-fold	20,0	5,0
----------------------	--	------------	------	-----

*M4 and M7 media*

M4 and M7 media are prepared using stock solution II, the macro-nutrients and vitamin as follows:

	Amount added to water (mg/l)	Concentration (related to medium M4)	Amount of stock solution II added to prepare medium (ml/l)	
			M4	M7
Stock solution II (combined trace elements)		20-fold	50	50
Macro nutrient stock solutions (single substance)				
CaCl <sub>2</sub> · 2H <sub>2</sub> O	293 800	1 000-fold	1,0	1,0
MgSO <sub>4</sub> · 7H <sub>2</sub> O	246 600	2 000-fold	0,5	0,5
KCl	58 000	10 000-fold	0,1	0,1
NaHCO <sub>3</sub>	64 800	1 000-fold	1,0	1,0
Na <sub>2</sub> SiO <sub>3</sub> · 9H <sub>2</sub> O	50 000	5 000-fold	0,2	0,2
NaNO <sub>3</sub>	2 740	10 000-fold	0,1	0,1
KH <sub>2</sub> PO <sub>4</sub>	1 430	10 000-fold	0,1	0,1
K <sub>2</sub> HPO <sub>4</sub>	1 840	10 000-fold	0,1	0,1

**▼B**

	Amount added to water (mg/l)	Concentration (related to medium M4)	Amount of stock solution II added to prepare medium (ml/l)	
			M4	M7
Combined Vitamin stock	—	10 000-fold	0,1	0,1

The combined vitamin stock solution is prepared by adding the 3 vitamin to 1 litre water, as shown below:

Thiamine hydrochloride	750	10 000-fold		
Cyanocobalamine (B <sub>12</sub> )	10	10 000-fold		
Biotine	7,5	10 000-fold		

The combined vitamin stock is stored frozen in small aliquots. Vitamins are added to the media shortly before use.

*N.B:* to avoid precipitation of salts when preparing the complete media, add the aliquots of stock solutions to about 500-800 ml deionised water and then fill up to 1 litre.

*N.B:* the first publication of the M4 medium can be found in Elendt, B. P. (1990). Selenium deficiency in crustacea; an ultrastructural approach to antennal damage in *Daphnia magna* Straus. *Protoplasma*, 154, 25-33.

**▼ M1****C.3. FRESHWATER ALGAE AND CYANOBACTERIA, GROWTH INHIBITION TEST****1. METHOD**

This method is equivalent to OECD TG 201 (2006) (1).

**1.1. INTRODUCTION**

Testing Methods are periodically reviewed and updated in the light of scientific progress. Testing Method C.3 needed to be revised to include additional species and to meet the requirements for hazard assessment and classification of chemicals. The revision has been completed on the basis of extensive practical experience, scientific progress in the field of algal toxicity studies, and extensive regulatory use, which has occurred since the original adoption.

**1.2. DEFINITIONS**

The following definitions and abbreviations are used for the purposes of this Testing Method:

**Biomass:** is the dry weight of living matter present in a population expressed in terms of a given volume; e.g. mg algae/litre test solution. Usually 'biomass' is defined as a mass, but in this test this word is used to refer to mass per volume. Also in this test, surrogates for biomass, such as cell counts, fluorescence, etc. are typically measured and the use of the term 'biomass' thus refers to these surrogate measures as well.

**Coefficient of variation:** is a dimensionless measure of the variability of a parameter, defined as the ratio of the standard deviation to the mean. This can also be expressed as a percentage value. The mean coefficient of variation of average specific growth rate in replicate control cultures should be calculated as follows:

1. Calculate % CV of average specific growth rate out of the daily/section-by-section growth rates for the respective replicate.
2. Calculate the mean value of all the values calculated in point 1 to get the mean coefficient of variation of the daily/section-by-section specific growth rate in replicate control cultures.

**EC<sub>x</sub>:** is the concentration of the test substance dissolved in the test medium that results in an x % (e.g. 50 %) reduction in growth of the test organism within a stated exposure period (to be mentioned explicitly if deviating from full or normal test duration). To unambiguously denote an EC value deriving from the growth rate or from the yield, the symbols 'E<sub>r</sub>C' and 'E<sub>y</sub>C' are used respectively.

**Growth medium:** is the complete synthetic culture medium in which test algae grow when exposed to the test substance. The test substance will normally be dissolved in the test medium.

**Growth rate (average specific growth rate):** is the logarithmic increase in biomass during the exposure period.

**▼ M1**

**Lowest Observed Effect Concentration (LOEC):** is the lowest tested concentration at which the substance is observed to have a statistically significant reducing effect on growth (at  $p < 0,05$ ) when compared with the control, within a given exposure time. However, all test concentrations above the LOEC must have a harmful effect equal to or greater than those observed at the LOEC. When these two conditions cannot be satisfied, a full explanation must be given for how the LOEC (and hence the NOEC) has been selected.

**No Observed Effect Concentration (NOEC):** is the test concentration immediately below the LOEC.

**Response variable:** is a variable for the estimation of toxicity derived from any measured parameters describing biomass by different methods of calculation. For this method, growth rates and yield are response variables derived from measuring biomass directly or any of the surrogates mentioned.

**Specific growth rate:** is a response variable defined as the quotient of the difference of the natural logarithms of a parameter of observation (in this Testing Method, biomass) and the respective time period.

**Yield:** is the value of a measurement variable at the end of the exposure period minus the measurement variable's value at the start of the exposure period to express biomass increase during the test.

### 1.3. APPLICABILITY OF THE TEST

This Testing Method is most easily applied to water-soluble substances which, under the conditions of the test, are likely to remain in the water. For testing of substances that are volatile, strongly adsorbing, coloured, having a low solubility in water or substances that may affect the availability of nutrients or minerals in the test medium, certain modifications of the described procedure may be required (e.g. closed system, conditioning of the test vessels). Guidance on some appropriate modifications is given in (2), (3) and (4).

### 1.4. PRINCIPLE OF THE TEST

The purpose of this test is to determine the effects of a substance on the growth of freshwater microalgae and/or cyanobacteria. Exponentially growing test organisms are exposed to the test substance in batch cultures over a period of normally 72 hours. In spite of the relatively brief test duration, effects over several generations can be assessed.

The system response is the reduction of growth in a series of algal cultures (test units) exposed to various concentrations of a test substance. The response is evaluated as a function of the exposure concentration in comparison with the average growth of replicate, unexposed control cultures. For full expression of the system response to toxic effects (optimal sensitivity), the cultures are allowed unrestricted exponential growth under sufficient nutrient conditions and continuous light for a sufficient period of time to measure reduction of the specific growth rate.

**▼ M1**

Growth and growth inhibition are quantified by measurements of the algal biomass as a function of time. Algal biomass is defined as the dry weight per volume, e.g. mg algae/litre test solution. However, dry weight is difficult to measure and therefore surrogate parameters are used. Of these surrogates, cell counts are most often used. Other surrogate parameters include cell volume, fluorescence, optical density, etc. A conversion factor between the measured surrogate parameter and biomass should be known.

The test endpoint is inhibition of growth, expressed as the logarithmic increase in biomass (average specific growth rate) during the exposure period. From the average specific growth rates recorded in a series of test solutions, the concentration bringing about a specified x % inhibition of growth rate (e.g. 50 %) is determined and expressed as the  $E_rC_x$  (e.g.  $E_rC_{50}$ ).

For the application of this method within the EU regulatory framework, calculation of results should be based on an average specific growth rate for the reasons described in section 2.2 below. An additional response variable used in this Testing Method is yield, which may be needed to fulfil specific regulatory requirements in some countries. It is defined as biomass at the end of the exposure period minus the biomass at the start of the exposure period. From the yield recorded in a series of test solutions, the concentration bringing about a specified x % inhibition of yield (e.g. 50 %) is calculated and expressed as the  $E_yC_x$  (e.g.  $E_yC_{50}$ ).

In addition, the lowest observed effect concentration (LOEC) and the no observed effect concentration (NOEC) may be statistically determined.

**1.5. INFORMATION ON THE TEST SUBSTANCE**

Information on the test substance, which may be useful in establishing the test conditions, includes the structural formula, purity, stability in light, stability under the conditions of the test, light absorption properties, pKa, and results of studies of transformation including biodegradability in water.

The water solubility, octanol water partition coefficient ( $P_{ow}$ ) and the vapour pressure of the test substance should be known, and a validated method for the quantification of the substance in the test solutions with reported recovery efficiency and limit of detection should be available.

**1.6. REFERENCE SUBSTANCE**

Reference substance(s), such as 3,5-dichlorophenol used in the international ring test (4), may be tested as a means of checking the test procedure. Potassium dichromate can also be used as a reference substance for green algae. It is desirable to test a reference substance at least twice a year.

**1.7. VALIDITY OF THE TEST**

For the test to be valid, the following performance criteria should be met:

**▼ M1**

- The biomass in the control cultures should have increased exponentially by a factor of at least 16 within the 72-hour test period. This corresponds to a specific growth rate of  $0,92 \text{ day}^{-1}$ . For the most frequently used species, the growth rate is usually substantially higher (see Appendix 1). This criterion may not be met when species that grow slower than those listed in Appendix 1 are used. In this case, the test period should be extended to obtain at least a 16-fold growth in control cultures, while the growth has to be exponential throughout the test period. The test period may be shortened to at least 48 h to maintain unlimited exponential growth during the test, as long as the minimum multiplication factor of 16 is reached.
- The mean coefficient of variation for section-by-section specific growth rates (days 0-1, 1-2 and 2-3, for 72-hour tests) in the control cultures (see section 1.2 under 'coefficient of variation') must not exceed 35 %. See the second paragraph in section 2.2.1 for the calculation of section-by-section specific growth rate. This criterion applies to the mean value of coefficients of variation calculated for replicate control cultures.
- The coefficient of variation of average specific growth rates during the whole test period in replicate control cultures must not exceed 7 % in tests with *Pseudokirchneriella subcapitata* and *Desmodesmus subspicatus*. For other less frequently tested species, the value should not exceed 10 %.

## 1.8. DESCRIPTION OF THE METHOD

1.8.1. **Apparatus**

Test vessels and other apparatus, which will come into contact with the test solutions should be made entirely of glass or other chemically inert material. The items should be thoroughly washed to ensure that no organic or inorganic contaminants may interfere with the algal growth or composition of the test solutions.

The test vessels will normally be glass flasks of dimensions that allow a sufficient volume of culture for measurements during the test and a sufficient mass transfer of  $\text{CO}_2$  from the atmosphere (see the second paragraph in section 1.8.9). Note that the liquid volume must be sufficient for analytical determinations (see the fifth paragraph in section 1.8.11).

In addition, some or all of the following equipment will be required:

- Culturing apparatus: a cabinet or chamber is recommended, in which the chosen incubation temperature can be maintained at  $\pm 2 \text{ }^\circ\text{C}$ .
- Light measurement instruments: it is important to note that the method of measurement of light intensity, and in particular the type of receptor (collector), will affect the measured value. Measurements should preferably be made using a spherical ( $4 \pi$ ) receptor (which responds to direct and reflected light from all angles above and below the plane of measurement), or a  $2 \pi$  receptor (which responds to light from all angles above the measurement plane).



**▼ M1**

- Apparatus to determine algal biomass. Cell count, which is the most frequently used surrogate parameter for algal biomass, may be made using an electronic particle counter, a microscope with counting chamber, or a flow cytometer. Other biomass surrogates can be measured using a flow cytometer, fluorimeter, spectrophotometer or colorimeter. A conversion factor relating cell count to dry weight is useful to calculate. In order to provide useful measurements at low biomass concentrations when using a spectrophotometer, it may be necessary to use cuvettes with a light path of at least 4 cm.

**1.8.2. Test organisms**

Several species of non-attached microalgae and cyanobacteria may be used. The strains listed in Appendix 1 have been shown to be suitable using the test procedure specified in this Testing Method.

If other species are used, the strain and/or origin should be reported. It has to be confirmed that exponential growth of the selected test alga can be maintained throughout the test period under the prevailing conditions.

**1.8.3. Growth medium**

Two alternative growth media, the OECD and the AAP medium, are recommended. The compositions of these media are shown in Appendix 2. Note that the initial pH value and the buffering capacity (regulating pH increase) of the two media are different. Therefore the results of the tests may be different depending on the medium used, particularly when testing ionising substances.

Modification of the growth media may be necessary for certain purposes, e.g. when testing metals and chelating agents or testing at different pH values. Use of a modified medium must be described in detail and justified (3)(4).

**1.8.4. Initial biomass concentration**

The initial biomass in the test cultures must be the same in all test cultures and sufficiently low to allow exponential growth throughout the incubation period without risk of nutrient depletion. The initial biomass should not exceed 0,5 mg/l as dry weight. The following initial cell concentrations are recommended:

<i>Pseudokirchneriella subcapitata</i>	$5 \times 10^3$ - $10^4$	cells/ml
<i>Desmodesmus subspicatus</i>	$2$ - $5 \times 10^3$	cells/ml
<i>Navicula pelliculosa</i>	$10^4$	cells/ml
<i>Anabaena flos-aquae</i>	$10^4$	cells/ml
<i>Synechococcus leopoliensis</i>	$5 \times 10^4$ - $10^5$	cells/ml

**1.8.5. Concentrations of test substance**

The concentration range in which effects are likely to occur may be determined on the basis of results from range-finding tests. For the final definitive test, at least five concentrations arranged in a geometric series with a factor not exceeding 3,2 should be selected. For test substances showing a flat concentration response curve, a higher factor may be justified. The concentration series should preferably cover the range causing 5-75 % inhibition of algal growth rate.

**▼ M1****1.8.6. Replicates and controls**

The test design should include three replicates at each test concentration. If determination of the NOEC is not required, the test design may be altered to increase the number of concentrations and reduce the number of replicates per concentration. The number of control replicates must be at least three, and ideally should be twice the number of replicates used for each test concentration.

A separate set of test solutions may be prepared for analytical determinations of test substance concentrations (see the fourth and sixth paragraphs in section 1.8.11).

When a solvent is used to solubilise the test substance, additional controls containing the solvent at the same concentration as used in the test cultures must be included in the test design.

**1.8.7. Preparation of inoculum culture**

In order to adapt the test alga to the test conditions and ensure that the algae are in the exponential growth phase when used to inoculate the test solutions, an inoculum culture in the test medium is prepared 2-4 days before start of the test. The algal biomass should be adjusted in order to allow exponential growth to prevail in the inoculum culture until the test starts. The inoculum culture shall be incubated under the same conditions as the test cultures. Measure the increase in biomass in the inoculum culture to ensure that growth is within the normal range for the test strain under the culturing conditions. An example of the procedure for algal culturing is described in Appendix 3. To avoid synchronous cell divisions during the test, a second propagation step of the inoculum culture may be required.

**1.8.8. Preparation of test solutions**

All test solutions must contain the same concentrations of growth medium and initial biomass of test algae. Test solutions of the chosen concentrations are usually prepared by mixing a stock solution of the test substance with growth medium and inoculum culture. Stock solutions are normally prepared by dissolving the substance in test medium.

Solvents, e.g. acetone, t-butyl alcohol and dimethyl formamide, may be used as carriers to add substances of low water solubility to the test medium (2)(3). The solvent concentration should not exceed 100 µl/l, and the same concentration of solvent should be added to all cultures (including controls) in the test series.

**1.8.9. Incubation**

Cap the test vessels with air-permeable stoppers. The vessels are shaken in the culturing apparatus. During the test it is necessary to keep the algae in suspension and to facilitate transfer of CO<sub>2</sub>. To this end, constant shaking or stirring should be used. The cultures should be maintained at a temperature in the range of 21 to 24 °C, controlled at ± 2 °C. For species other than those listed in Appendix 1, e.g. tropical species, higher temperatures may be appropriate, providing that the validity criteria can be fulfilled. It is recommended to place the flasks randomly and to reposition them daily in the incubator.

**▼ M1**

The pH of the control medium should not increase by more than 1,5 units during the test. For metals and compounds that partly ionise at a pH around the test pH, it may be necessary to limit the pH drift to obtain reproducible and well defined results. A drift of < 0,5 pH units is technically feasible and can be achieved by ensuring an adequate CO<sub>2</sub> mass transfer rate from the surrounding air to the test solution, e.g. by increasing the shaking rate. Another possibility is to reduce the demand for CO<sub>2</sub> by reducing the initial biomass or the test duration.

The surface where the cultures are incubated should receive continuous, uniform fluorescent illumination e.g. of 'cool-white' or 'daylight' type. Strains of algae and cyanobacteria vary in their light requirements. The light intensity should be selected to suit the test organism used. For the recommended species of green algae, the light intensity at the level of the test solutions shall be selected from the range of 60-120  $\mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$  when measured in the photosynthetically effective wavelength range of 400-700 nm using an appropriate receptor. Some species, in particular *Anabaena flos-aquae*, grow well at lower light intensities and may be damaged at high intensities. For such species an average light intensity in the range 40-60  $\mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$  should be selected. (For light-measuring instruments calibrated in lux, an equivalent range of 4 440-8 880 lux for cool white light corresponds approximately to the recommended light intensity 60-120  $\mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ ). The light intensity shall not vary more than  $\pm 15\%$  from the average light intensity over the incubation area.

**1.8.10. Test duration**

The test duration is normally 72 hours. However, shorter or longer test durations may be used, provided that all validity criteria in section 1.7 can be met.

**1.8.11. Measurements and analytical determinations**

The algal biomass in each flask is determined at least daily during the test period. If measurements are made on small volumes removed from the test solution by pipette, these should not be replaced.

Measurement of biomass is done by manual cell counting by microscope or an electronic particle counter (by cell counts and/or biovolume). Alternative techniques, e.g. flow cytometry, in vitro or in vivo chlorophyll fluorescence (6)(7), or optical density can be used providing a satisfactory correlation with biomass can be demonstrated over the range of biomass occurring in the test.

The pH of the solutions shall be measured at the beginning and at the end of the test.

Provided an analytical procedure for determination of the test substance in the concentration range used is available, the test solutions should be analysed to verify the initial concentrations and maintenance of the exposure concentrations during the test.

**▼ M1**

Analysis of the concentration of the test substance at the start and end of the test of a low and high test concentration, and a concentration around the expected  $EC_{50}$  may be sufficient where it is likely that exposure concentrations will vary less than 20 % from nominal values during the test. Analysis of all test concentrations at the start and end of the test is recommended where concentrations are unlikely to remain within 80-120 % of nominal. For volatile, unstable or strongly adsorbing test substances, additional sampling for analysis at 24 hour intervals during the exposure period is recommended in order to better define loss of the test substance. For these substances, extra replicates will be needed. In all cases, determination of test substance concentrations need only be performed on one replicate vessel at each test concentration (or on the contents of the vessels pooled by replicate).

Test media prepared specifically for analysis of exposure concentrations during the test should be treated identically to those used for testing, i.e. they should be inoculated with algae and incubated under identical conditions. If analysis of the dissolved test substance concentration is required, it may be necessary to separate algae from the medium. Separation should preferably be made by centrifugation at a low g-force, sufficient to settle the algae.

If there is evidence that the concentration of the substance being tested has been satisfactorily maintained within  $\pm 20$  % of the nominal or measured initial concentration throughout the test, analysis of the results can be based on nominal or measured initial values. If the deviation from the nominal or measured initial concentration is greater than  $\pm 20$  %, analysis of the results should be based on geometric mean concentration during exposure or on models describing the decline of the concentration of test substance (3)(8).

The alga growth inhibition test is a more dynamic test system than most other short-term aquatic toxicity tests. As a consequence, the actual exposure concentrations may be difficult to define, especially for adsorbing substances tested at low concentrations. In such cases, disappearance of the substance from solution by adsorption to the increasing algal biomass does not mean that it is lost from the test system. When the result of the test is analysed, it should be checked whether a decrease in concentration of the test substance in the course of the test is accompanied by a decrease in growth inhibition. If this is the case, application of a suitable model describing the decline of the concentration of test substance (8) may be considered. If not, it may be appropriate to base the analysis of the results on the initial (nominal or measured) concentrations.

**1.8.12. Other observations**

Microscopic observation should be performed to verify a normal and healthy appearance of the inoculum culture and to observe any abnormal appearance of the algae (as may be caused by exposure to the test substance) at the end of the test.

**▼ M1****1.8.13. Limit test**

Under some circumstances, e.g. when a preliminary test indicates that the test substance has no toxic effects at concentrations up to 100 mg·l<sup>-1</sup> or up to its limit of solubility in the test medium (whichever is the lower), a limit test involving a comparison of responses in a control group and one treatment group (100 mg·l<sup>-1</sup> or a concentration equal to the limit of solubility), may be undertaken. It is strongly recommended that this be supported by analysis of the exposure concentration. All previously described test conditions and validity criteria apply to a limit test, with the exception that the number of treatment replicates should be at least six. The response variables in the control and treatment group may be analysed using a statistical test to compare means, e.g. a Student's t-test. If variances of the two groups are unequal, a t-test adjusted for unequal variances should be performed.

**1.8.14. Modification for strongly coloured substances**

The irradiation (light intensity) should be in the highest end of the range prescribed in this Testing Method: 120 μE m<sup>-2</sup> s<sup>-1</sup> or higher.

The light path should be shortened by reduction of the volume of the test solutions (in the range of 5-25 ml).

Sufficient agitation (for example by moderate shaking) should be performed in order to obtain a high frequency of exposure of the algae to high irradiation at the surface of the culture.

**2. DATA****2.1. PLOTTING GROWTH CURVES**

The biomass in the test vessels may be expressed in units of the surrogate parameter used for measurement (e.g. cell number, fluorescence).

Tabulate the estimated biomass concentration in test cultures and controls together with the concentrations of test material and the times of measurement, recorded with a resolution of at least whole hours, to produce plots of growth curves. Both logarithmic scales and linear scales can be useful at this first stage, but logarithmic scales are mandatory and generally give a better presentation of variations in growth pattern during the test period. Note that exponential growth produces a straight line when plotted on a logarithmic scale, and that the inclination of the line (slope) indicates the specific growth rate.

Using the plots, examine whether control cultures grow exponentially at the expected rate throughout the test. Critically examine all data points and the appearance of the graphs, and check raw data and procedures for possible errors. Check in particular any data point that seems to deviate by a systematic error. If it is obvious that procedural mistakes can be identified and/or considered highly likely, the specific data point is marked as an outlier and not included in subsequent statistical analysis. (A zero algal concentration in one out of two or three replicate vessels may indicate the vessel was not inoculated correctly, or was improperly cleaned). Reasons for rejection of a data point as an outlier must be clearly stated in the test report. Accepted reasons are only (rare) procedural mistakes and not just bad precision. Statistical procedures for outlier identification are of limited use for this type of problem and cannot replace expert judgement. Outliers (marked as such) should preferably be retained among the data points shown in any subsequent graphical or tabular data presentation.

▼ **M1**

## 2.2. RESPONSE VARIABLES

The purpose of the test is to determine the effects of the test substance on the growth of algae. This Testing Method describes two response variables, as member countries have different preferences and regulatory needs. In order for the test results to be acceptable in all member countries, the effects should be evaluated using both response variables (a) and (b) described below.

- (a) Average specific growth rate: this response variable is calculated on the basis of the logarithmic increase of biomass during the test period, expressed per day.
- (b) Yield: this response variable is the biomass at the end of the test minus the starting biomass.

For the application of this method within the EU regulatory framework, calculation of results should be based on an average specific growth rate for the reasons described below. It should be noted that toxicity values calculated by using these two response variables are not comparable and this difference must be recognised when using the results of the test.  $EC_x$  values based upon average specific growth rate ( $E_rC_x$ ) will generally be higher than results based upon yield ( $E_yC_x$ ) if the test conditions of this Testing Method are adhered to, due to the mathematical basis of the respective approaches. This should not be interpreted as a difference in sensitivity between the two response variables, simply that the values are different mathematically. The concept of average specific growth rate is based on the general exponential growth pattern of algae in non-limited cultures, where toxicity is estimated on the basis of the effects on the growth rate, without being dependent on the absolute level of the specific growth rate of the control, on the slope of the concentration-response curve or on test duration. In contrast, results based upon the yield response variable are dependent upon all these other variables.  $E_yC_x$  is dependent on the specific growth rate of the algal species used in each test and on the maximum specific growth rate that can vary between species and even different algal strains. This response variable should not be used for comparing the sensitivity to toxicants among algal species or even different strains. While the use of average specific growth rate for estimating toxicity is scientifically preferred, toxicity estimates based on yield are also included in this Testing Method so as to satisfy current regulatory requirements in some countries.

2.2.1. **Average growth rate**

The average specific growth rate for a specific period is calculated as the logarithmic increase in biomass from the equation for each single vessel of controls and treatments:

$$\mu_{i-j} = \frac{\ln X_j - \ln X_i}{t_j - t_i} \text{ (day}^{-1}\text{)}$$

where:

$\mu_{i-j}$ : is the average specific growth rate from time  $i$  to  $j$ ;

$X_i$ : is the biomass at time  $i$ ;

$X_j$ : is the biomass at time  $j$ .

**▼ M1**

For each treatment group and control group, calculate a mean value for growth rate along with variance estimates.

Calculate average specific growth rate over the entire test duration (normally days 0-3), using the nominally inoculated biomass as the starting value rather than a measured starting value, because in this way greater precision is normally obtained. If the equipment used for biomass measurement allows sufficiently precise determination of the low inoculum biomass (e.g. flow cytometer) then the measured initial biomass concentration can be used. Also assess the section-by-section growth rate, calculated as the specific growth rates for each day during the course of the test (days 0-1, 1-2 and 2-3) and examine whether the control growth rate remains constant (See validity criteria, section 1.7). A significantly lower specific growth rate on day one than the total average specific growth rate may indicate a lag phase. While a lag phase can be minimised and practically eliminated in control cultures by proper propagation of the pre-culture, a lag phase in exposed cultures may indicate recovery after initial toxic stress or reduced exposure due to loss of test substance (including sorption onto the algal biomass) after initial exposure. Hence the section-by-section growth rate may be assessed in order to evaluate effects of the test substance occurring during the exposure period. Substantial differences between the section-by-section growth rate and the average growth rate indicate deviation from constant exponential growth and that close examination of the growth curves is warranted.

Calculate the percent inhibition of growth rate for each treatment replicate from the equation:

$$\%I_r = \frac{\mu_C - \mu_T}{\mu_C} \times 100$$

where:

$\%I_r$ : percent inhibition in average specific growth rate;

$\mu_C$ : mean value for average specific growth rate ( $\mu$ ) in the control group;

$\mu_T$ : average specific growth rate for the treatment replicate.

When solvents are used to prepare the test solutions, the solvent controls rather than the controls without solvents should be used in calculation of percent inhibition.

### 2.2.2. Yield

The yield is calculated as the biomass at the end of the test minus the starting biomass for each single vessel of controls and treatments. For each test concentration and control, calculate a mean value for the yield along with variance estimates. The percentage inhibition in yield ( $\%I_y$ ) may be calculated for each treatment replicate as follows:

$$\%I_y = \frac{(Y_C - Y_T)}{Y_C} \times 100$$

**▼ M1**

where:

%I<sub>y</sub>: percentage inhibition of yield;

Y<sub>C</sub>: mean value for yield in the control group;

Y<sub>T</sub>: value for yield for the treatment replicate.

### 2.3. PLOTTING CONCENTRATION RESPONSE CURVE

Plot the percentage of inhibition against the logarithm of the test substance concentration and examine the plot closely, disregarding any such data point that was singled out as an outlier in the first phase. Fit a smooth line through the data points by eye or by computerised interpolation to get a first impression of the concentration response relationship, then proceed with a more detailed method, preferably a computerised statistical method. Depending on the intended usage of the data, the quality (precision) and amount of data, as well as the availability of data analysis tools, it may be decided (and sometimes well justified) to stop the data analysis at this stage and simply read the key figures EC<sub>50</sub> and EC<sub>10</sub> (and/or EC<sub>20</sub>) from the eye fitted curve (also see section below on stimulatory effects). Valid reasons for not using a statistical method may include:

- data are not appropriate for computerised methods to produce any more reliable results than can be obtained by expert judgement — in such situations some computer programs may even fail to produce a reliable solution (iterations may not converge etc.),
- stimulatory growth responses cannot be handled adequately using available computer programs (see below).

### 2.4. STATISTICAL PROCEDURES

The aim is to obtain a quantitative concentration-response relationship by regression analysis. It is possible to use a weighted linear regression after having performed a linearising transformation of the response data — for instance into probit or logit or Weibull units (9), but non-linear regression procedures are preferred techniques that better handle unavoidable data irregularities and deviations from smooth distributions. Approaching either zero or total inhibition, such irregularities may be magnified by the transformation, interfering with the analysis (9). It should be noted that standard methods of analysis using probit, logit, or Weibull transforms are intended for use on quantal (e.g. mortality or survival) data, and must be modified to accommodate growth or biomass data. Specific procedures for determination of EC<sub>x</sub> values from continuous data can be found in (10)(11) and (12). The use of non-linear regression analysis is further detailed in Appendix 4.



**▼ M1**

For each response variable to be analysed, use the concentration-response relationship to calculate point estimates of  $EC_x$  values. When possible, the 95 % confidence limits for each estimate should be determined. Goodness of fit of the response data to the regression model should be assessed either graphically or statistically. Regression analysis should be performed using individual replicate responses, not treatment group means. If, however, nonlinear curve fitting is difficult or fails because of too great a scatter in the data, the problem may be circumvented by performing the regression on group means as a practical way of reducing the influence of suspected outliers. Use of this option should be identified in the test report as a deviation from normal procedure because curve fits with individual replicates did not produce a good result.

$EC_{50}$  estimates and confidence limits may also be obtained using linear interpolation with bootstrapping (13), if available regression models/methods are unsuitable for the data.

For estimation of the LOEC and hence the NOEC, and for effects of the test substance on growth rate, it is necessary to compare treatment means using analysis of variance (ANOVA) techniques. The mean for each concentration must then be compared with the control mean using an appropriate multiple comparison or trend test method. Dunnett's or Williams' test may be useful (14)(15)(16)(17)-(18). It is necessary to assess whether the ANOVA assumption of homogeneity of variance holds. This assessment may be performed graphically or by a formal test (18). Suitable tests are Levene's or Bartlett's. Failure to meet the assumption of homogeneity of variances can sometimes be corrected by logarithmic transformation of the data. If heterogeneity of variance is extreme and cannot be corrected by transformation, analysis by methods such as step-down Jonkheere trend tests should be considered. Additional guidance on determining the NOEC can be found in (12).

Recent scientific developments have led to a recommendation of abandoning the concept of NOEC and replacing it with regression based point estimates  $EC_x$ . An appropriate value for  $x$  has not been established for this algal test. A range of 10 to 20 % appears to be appropriate (depending on the response variable chosen), and preferably both the  $EC_{10}$  and  $EC_{20}$  should be reported.

## 2.5. GROWTH STIMULATION

Growth stimulation (negative inhibition) at low concentrations is sometimes observed. This can result from either hormesis (toxic stimulation) or from addition of stimulating growth factors with the test material to the minimal medium used. Note that the addition of inorganic nutrients should not have any direct effect because the test medium should maintain a surplus of nutrients throughout the test. Low dose stimulation can usually be ignored in  $EC_{50}$  calculations unless it is extreme. However, if it is extreme, or an  $EC_x$  value for low  $x$  is to be calculated, special procedures may be needed. Deletion of stimulatory responses from the data analysis should be avoided if possible, and if available curve fitting software cannot accept minor stimulation, linear interpolation with bootstrapping can be used. If stimulation is extreme, use of a hormesis model may be considered (19).

**▼ M1**

## 2.6. NON-TOXIC GROWTH INHIBITION

Light absorbing test materials may give rise to a growth rate reduction because shading reduces the amount of available light. Such physical types of effects should be separated from toxic effects by modifying the test conditions and the former should be reported separately. Guidance may be found in (2) and (3).

3. **REPORTING**

## 3.1. TEST REPORT

The test report must include the following:

Test substance:

- physical nature and relevant physiochemical properties, including water solubility limit,
- chemical identification data, including purity.

Test species:

- the strain, supplier or source and the culture conditions used.

Test conditions:

- date of start of the test and its duration,
- description of test design: test vessels, culture volumes, biomass density at the beginning of the test,
- composition of the medium,
- test concentrations and replicates (e.g. number of replicates, number of test concentrations and geometric progression used),
- description of the preparation of test solutions, including use of solvents etc.,
- culturing apparatus,
- light intensity and quality (source, homogeneity),
- temperature,
- concentrations tested: the nominal test concentrations and any results of analyses to determine the concentration of the test substance in the test vessels. The recovery efficiency of the method and the limit of quantification in the test matrix should be reported,
- all deviations from this Testing Method,
- method for determination of biomass and evidence of correlation between the measured parameter and dry weight.

Results:

- pH values at the start and end of the test at all treatments,
- biomass for each flask at each measuring point and method for measuring biomass,

**▼ M1**

- growth curves (plot of biomass versus time),
- calculated response variables for each treatment replicate, with mean values and coefficient of variation for replicates,
- graphical presentation of the concentration/effect relationship,
- estimates of toxicity for response variables e.g. EC<sub>50</sub>, EC<sub>10</sub>, EC<sub>20</sub> and associated confidence intervals. If calculated, LOEC and NOEC and the statistical methods used for their determination,
- if ANOVA has been used, the size of the effect which can be detected (e.g. the least significant difference),
- any stimulation of growth found in any treatment,
- any other observed effects, e.g. morphological changes of the algae,
- discussion of the results, including any influence on the outcome of the test resulting from deviations from this Testing Method.

**4. LITERATURE**

- (1) OECD TG 201 (2006). Freshwater Alga and Cyanobacteria, Growth Inhibition Test.
- (2) ISO 1998: Water quality — Guidance for algal growth inhibition tests with poorly soluble materials, volatile compounds, metals and waster water. ISO/DIS 14442.
- (3) OECD 2000: Guidance Document on Aquatic Toxicity Testing of Difficult Substances and mixtures. Environmental Health and Safety Publications. Series on Testing and Assessment, no. 23.
- (4) ISO 1998: Water quality — Sampling — Part 16: General Guidance for Biotesting. ISO 5667-16.
- (5) ISO 1993: Water quality — Algal growth inhibition test. ISO 8692.
- (6) Mayer, P., Cuhel, R. and Nyholm, N. (1997). A simple in vitro fluorescence method for biomass measurements in algal growth inhibition tests. *Water Research* 31: 2525-2531.
- (7) Slovacey, R.E. and Hanna, P.J. In vivo fluorescence determinations of phytoplankton chlorophyll, *Limnology & Oceanography* 22,5 (1977), pp.919-925.
- (8) Simpson, S.L., Roland, M.G.E., Stauber, J.L. and Batley, G.E. (2003). Effect of declining toxicant concentrations on algal bioassay endpoints. *Environ. Toxicol. Chem* 22, 2073-2079.
- (9) Christensen, E.R., Nyholm, N. (1984): Ecotoxicological Assays with Algae: Weibull Dose-Response Curves. *Env. Sci. Technol.* 19, 713-718.

**▼ M1**

- (10) Nyholm, N. Sørensen, P.S., Kusk, K.O. and Christensen, E.R. (1992): Statistical treatment of data from microbial toxicity tests. *Environ. Toxicol. Chem.* 11, 157-167.
- (11) Bruce, R.D., and Versteeg, D.J. (1992). A statistical procedure for modelling continuous toxicity data. *Env. Toxicol. Chem.* 11:1485-1494.
- (12) OECD. (2004). *Guidance Document on Statistical Analysis of Ecotoxicity Data.*
- (13) Norberg-King T.J. (1988). An interpolation estimate for chronic toxicity: The ICp approach. National Effluent Toxicity Assessment Center Technical Report 05-88. USEPA, Duluth, MN.
- (14) Dunnett, C.W. (1955). A multiple comparisons procedure for comparing several treatments with a control. *J. Amer. Statist. Assoc.* 50: 1096-1121
- (15) Dunnett, C.W. (1964). New tables for multiple comparisons with a control. *Biometrics* 20: 482-491.
- (16) Williams, D.A. (1971). A test for differences between treatment means when several dose levels are compared with a zero dose control. *Biometrics* 27: 103-117.
- (17) Williams, D.A. (1972). The comparison of several dose levels with a zero dose control. *Biometrics* 28: 510-531.
- (18) Draper, N.R. and Smith, H. (1981). *Applied Regression Analysis*, second edition. Wiley, New York.
- (19) Brain P. and Cousens R. (1989). An equation to describe dose-responses where there is stimulation of growth at low doses. *Weed Research*, 29, 93-96.

**▼ M1***Appendix 1***Strains shown to be suitable for the test****Green algae**

— *Pseudokirchneriella subcapitata*, (formerly known as *Selenastrum capricornutum*), ATCC 22662, CCAP 278/4, 61.81 SAG

— *Desmodesmus subspicatus* (formerly known as *Scenedesmus subspicatus*)  
86.81 SAG

**Diatoms**

— *Navicula pelliculosa*, UTEX 664

**Cyanobacteria**

— *Anabaena flos-aquae*, UTEX 1444, ATCC 29413, CCAP 1403/13A

— *Synechococcus leopoliensis*, UTEX 625, CCAP 1405/1

**Sources of Strains**

The strains recommended are available in unialgal cultures from the following collections (in alphabetical order):

ATCC: American Type Culture Collection  
10801 University Boulevard  
Manassas, Virginia 20110-2209  
UNITED STATES

CCAP, Culture Collection of Algae and Protozoa  
Institute of Freshwater Ecology,  
Windermere Laboratory  
Far Sawrey, Amblerside  
Cumbria  
LA22 0LP  
UNITED KINGDOM

SAG: Collection of Algal Cultures  
Inst. Plant Physiology  
University of Göttingen  
Nicholausberger Weg 18  
3400 Göttingen  
GERMANY

UTEX Culture Collection of Algae  
Section of Molecular, Cellular and Developmental Biology  
School of Biological Sciences  
the University of Texas at Austin  
Austin, Texas 78712  
USA

▼ **M1****Appearance and characteristics of recommended species**

	<i>P. subcapitata</i>	<i>D. subspicatus</i>	<i>N. pelliculosa</i>	<i>A. flos-aquae</i>	<i>S. leopoliensis</i>
Appearance	Curved, twisted single cells	Oval, mostly single cells	Rods	Chains of oval cells	Rods
Size (L × W) μm	8-14 × 2-3	7-15 × 3-12	7,1 × 3,7	4,5 × 3	6 × 1
Cell volume (μm <sup>3</sup> /cell)	40-60 <sup>(1)</sup>	60-80 <sup>(1)</sup>	40-50 <sup>(1)</sup>	30-40 <sup>(1)</sup>	2,5 <sup>(2)</sup>
Cell dry weight (mg/cell)	2-3 × 10 <sup>-8</sup>	3-4 × 10 <sup>-8</sup>	3-4 × 10 <sup>-8</sup>	1-2 × 10 <sup>-8</sup>	2-3 × 10 <sup>-9</sup>
Growth rate <sup>(3)</sup> day <sup>-1</sup>	1,5-1,7	1,2-1,5	1,4	1,1-1,4	2,0-2,4

<sup>(1)</sup> Measured with electronic particle counter

<sup>(2)</sup> Calculated from size

<sup>(3)</sup> Most frequently observed growth rate in OECD medium at light intensity approx. 70 μE·m<sup>-2</sup>·s<sup>-1</sup> and 21 °C

**Specific Recommendations on Culturing and Handling of Recommended Test Species***Pseudokirchneriella subcapitata* and *Desmodesmus subspicatus*

These green algae are generally easy to maintain in various culture media. Information on suitable media is available from the culture collections. The cells are normally solitary, and cell density measurements can easily be performed using an electronic particle counter or microscope.

*Anabaena flos-aquae*

Various growth media may be used for keeping a stock culture. It is particularly important to avoid allowing the batch culture to go past log phase growth when renewing, recovery is difficult at this point.

*Anabaena flos-aquae* develops aggregates of nested chains of cells. The size of these aggregates may vary with culturing conditions. It may be necessary to break up these aggregates when microscope counting or an electronic particle counter is used for determination of biomass.

Sonication of sub-samples may be used to break up chains to reduce count variability. Longer sonication than required for breaking up chains into shorter lengths may destroy the cells. Sonication intensity and duration must be identical for each treatment.

Count enough fields on the hemocytometer (at least 400 cells) to help compensate for variability. This will improve reliability of microscopic density determinations.

An electronic particle counter can be used for determination of total cell volume of *Anabaena* after breaking up the cell chains by careful sonification. The sonification energy has to be adjusted to avoid disruption of the cells.

Use a vortex mixer or similar appropriate method to make sure the algae suspension used to inoculate test vessels is well mixed and homogeneous.

**▼ M1**

Test vessels should be placed on an orbital or reciprocate shaker table at about 150 revolutions per minute. Alternatively, intermittent agitation may be used to reduce the tendency of *Anabaena* to form clumps. If clumping occurs, care must be taken to achieve representative samples for biomass measurements. Vigorous agitation before sampling may be necessary to disintegrate algal clumps.

*Synechococcus leopoliensis.*

Various growth media may be used for keeping a stock culture. Information on suitable media is available from the culture collections.

*Synechococcus leopoliensis* grows as solitary rod-shaped cells. The cells are very small, which complicates the use of microscope counting for biomass measurements. Electronic particle counters equipped for counting particles down to a size of approximately 1 µm are useful. In vitro fluorometric measurements are also applicable.

*Navicula pelliculosa*

Various growth media may be used for keeping a stock culture. Information on suitable media is available from the culture collections. Note that silicate is required in the medium.

*Navicula pelliculosa* may form aggregates under certain growth conditions. Due to production of lipids the algal cells sometimes tend to accumulate in the surface film. Under those circumstances special measures have to be taken when subsamples are taken for biomass determination in order to obtain representative samples. Vigorous shaking, e.g. using a vortex mixer may be required.

▼ **M1***Appendix 2***Growth media**

One of the following two growth media may be used:

OECD medium: Original medium of OECD TG 201, also according to ISO 8692

US. EPA medium AAP also according to ASTM.

When preparing these media, reagent or analytical-grade chemicals should be used and deionised water.

**Composition of The AAP-medium (US. EPA) and the OECD TG 201 medium**

Component	EPA		OECD	
	mg/l	mM	mg/l	mM
NaHCO <sub>3</sub>	15,0	0,179	50,0	0,595
NaNO <sub>3</sub>	25,5	0,300		
NH <sub>4</sub> Cl			15,0	0,280
MgCl <sub>2</sub> ·6(H <sub>2</sub> O)	12,16	0,0598	12,0	0,0590
CaCl <sub>2</sub> ·2(H <sub>2</sub> O)	4,41	0,0300	18,0	0,122
MgSO <sub>4</sub> ·7(H <sub>2</sub> O)	14,6	0,0592	15,0	0,0609
K <sub>2</sub> HPO <sub>4</sub>	1,044	0,00599		
KH <sub>2</sub> PO <sub>4</sub>			1,60	0,00919
FeCl <sub>3</sub> ·6(H <sub>2</sub> O)	0,160	0,000591	0,0640	0,000237
Na <sub>2</sub> EDTA·2(H <sub>2</sub> O)	0,300	0,000806	0,100	0,000269 (*)
H <sub>3</sub> BO <sub>3</sub>	0,186	0,00300	0,185	0,00299
MnCl <sub>2</sub> ·4(H <sub>2</sub> O)	0,415	0,00201	0,415	0,00210
ZnCl <sub>2</sub>	0,00327	0,000024	0,00300	0,0000220
CoCl <sub>2</sub> ·6(H <sub>2</sub> O)	0,00143	0,000006	0,00150	0,00000630
Na <sub>2</sub> MoO <sub>4</sub> ·2(H <sub>2</sub> O)	0,00726	0,000030	0,00700	0,0000289
CuCl <sub>2</sub> ·2(H <sub>2</sub> O)	0,000012	0,00000007	0,00001	0,00000006
pH	7,5		8,1	

(\*) The molar ratio of EDTA to iron slightly exceed unity. This prevents iron precipitation and, at the same time, chelation of heavy metal ions is minimised.

In the test with the diatom *Navicula pelliculosa*, both media must be supplemented with Na<sub>2</sub>SiO<sub>3</sub>·9H<sub>2</sub>O to obtain a concentration of 1,4 mg Si/l.

The pH of the medium is obtained at equilibrium between the carbonate system of the medium and the partial pressure of CO<sub>2</sub> in atmospheric air. An approximate relationship between pH at 25 °C and the molar bicarbonate concentration is:

$$PH_{eq} = 11,30 + \log [HCO_3]$$



**▼ M1**

With 15 mg NaHCO<sub>3</sub>, pH<sub>eq</sub> = 7,5 (U.S. EPA medium) and with 50 mg NaHCO<sub>3</sub>/l, pH<sub>eq</sub> = 8,1 (OECD medium).

**Element composition of test media**

Element	EPA	OECD
	mg/l	mg/l
C	2,144	7,148
N	4,202	3,927
P	0,186	0,285
K	0,469	0,459
Na	11,044	13,704
Ca	1,202	4,905
Mg	2,909	2,913
Fe	0,033	0,017
Mn	0,115	0,115

**Preparation of OECD medium**

Nutrient	Concentration in stock solution
Stock solution 1: macronutrients	
NH <sub>4</sub> Cl	1,5 g·l <sup>-1</sup>
MgCl <sub>2</sub> ·6H <sub>2</sub> O	1,2 g·l <sup>-1</sup>
CaCl <sub>2</sub> ·2H <sub>2</sub> O	1,8 g·l <sup>-1</sup>
MgSO <sub>4</sub> ·7H <sub>2</sub> O	1,5 g·l <sup>-1</sup>
KH <sub>2</sub> PO <sub>4</sub>	0,16 g·l <sup>-1</sup>
Stock solution 2: iron	
FeCl <sub>3</sub> ·6H <sub>2</sub> O	64 mg·l <sup>-1</sup>
Na <sub>2</sub> EDTA·2H <sub>2</sub> O	100 mg·l <sup>-1</sup>
Stock solution 3: trace elements	
H <sub>3</sub> BO <sub>3</sub>	185 mg·l <sup>-1</sup>
MnCl <sub>2</sub> ·4H <sub>2</sub> O	415 mg·l <sup>-1</sup>
ZnCl <sub>2</sub>	3 mg·l <sup>-1</sup>
CoCl <sub>2</sub> ·6H <sub>2</sub> O	1,5 mg·l <sup>-1</sup>
CuCl <sub>2</sub> ·2H <sub>2</sub> O	0,01 mg·l <sup>-1</sup>
Na <sub>2</sub> MoO <sub>4</sub> ·2H <sub>2</sub> O	7 mg·l <sup>-1</sup>
Stock solution 4: bicarbonate	
NaHCO <sub>3</sub>	50 g·l <sup>-1</sup>
Na <sub>2</sub> SiO <sub>3</sub> ·9H <sub>2</sub> O	

Sterilise the stock solutions by membrane filtration (mean pore diameter 0,2 µm) or by autoclaving (120 °C, 15 min). Store the solutions in the dark at 4 °C.

Do not autoclave stock solutions 2 and 4, but sterilise them by membrane filtration.

**▼ M1**

Prepare a growth medium by adding an appropriate volume of the stock solutions 1-4 to water:

Add to 500 ml of sterilised water:

— 10 ml of stock solution 1

— 1 ml of stock solution 2

— 1 ml of stock solution 3

— 1 ml of stock solution 4

Make up to 1 000 ml with sterilised water

Allow sufficient time for equilibrating the medium with the atmospheric CO<sub>2</sub>, if necessary by bubbling with sterile filtered air for some hours.

#### Preparation of AAP medium

A1.1. Add 1 mL of each stock solution in A1.2.1-A1.2.7 to approximately 900 mL of deionised or distilled water and then dilute to 1 L.

A1.2. Macronutrient stock solutions are made by dissolving the following into 500 mL of deionised or distilled water. Reagents A1.2.1, A1.2.2, A1.2.3, and A1.2.4 can be combined into one stock solution.

A1.2.1.  $NaNO_3$ —12,750 g.

A1.2.2.  $MgCl_2 \cdot 6H_2O$ —6,082 g.

A1.2.3.  $CaCl_2 \cdot 2H_2O$ —2,205 g.

A1.2.4. *Micronutrient Stock Solution*—(see A1.3).

A1.2.5.  $MgSO_4 \cdot 7H_2O$ —7,350 g.

A1.2.6.  $K_2HPO_4$ —0,522 g.

A1.2.7.  $NaHCO_3$ —7,500 g.

A1.2.8.  $Na_2SiO_3 \cdot 9H_2O$ —See Note A1.1.

*Note A1.1* — Use for diatom test species only. May be added directly (202,4 mg) or by way of stock solution to give 20 mg/L Si final concentration in medium.

A1.3. The micronutrient stock solution is made by dissolving the following into 500 mL of deionised or distilled water:

A1.3.1.  $H_3BO_3$ —92,760 mg.

A1.3.2.  $MnCl_2 \cdot 4H_2O$ —207,690 mg.

A1.3.3.  $ZnCl_2$ —1,635 mg.

A1.3.4.  $FeCl_3 \cdot 6H_2O$ —79,880 mg.

A1.3.5.  $CoCl_2 \cdot 6H_2O$ —0,714 mg.

A1.3.6.  $Na_2MoO_4 \cdot 2H_2O$ —3,630 mg.

A1.3.7.  $CuCl_2 \cdot 2H_2O$ —0,006 mg.

A1.3.8.  $Na_2EDTA \cdot 2H_2O$ —150,000 mg.

[Disodium (Ethylenedinitrilo) tetraacetate].

**▼ M1**

A1.3.9.  $Na_2SeO_4 \cdot 5H_2O$ —0,005 mg See Note A1.2.

Note A1.2 — Use only in medium for stock cultures of diatom species.

A1.4. Adjust pH to  $7,5 \pm 0,1$  with 0,1 *N* or 1,0 *N* NaOH or HCl.

A1.5. Filter the media into a sterile container through either a 0,22- $\mu\text{m}$  membrane filter if a particle counter is to be used or a 0,45- $\mu\text{m}$  filter if a particle counter is not to be used.

A1.6. Store medium in the dark at approximately 4 °C until use.

**▼ M1***Appendix 3***Example of a procedure for the culturing of algae****General observations**

The purpose of culturing on the basis of the following procedure is to obtain algal cultures for toxicity tests.

Suitable methods must be used to ensure that the algal cultures are not infected with bacteria. Axenic cultures may be desirable but unialgal cultures must be established and used.

All operations must be carried out under sterile conditions in order to avoid contamination with bacteria and other algae.

**Equipment and materials**

See under Testing Method: Apparatus.

**Procedures for obtaining algal cultures***Preparation of nutrient solutions (media):*

All nutrient salts of the medium are prepared as concentrated stock solutions and stored dark and cold. These solutions are sterilised by filtration or by autoclaving.

The medium is prepared by adding the correct amount of stock solution to sterile distilled water, taking care that no infections occur. For solid medium 0,8 % of agar is added.

*Stock culture:*

The stock cultures are small algal cultures that are regularly transferred to fresh medium to act as initial test material. If the cultures are not used regularly, they are streaked out on sloped agar tubes. These are transferred to fresh medium at least once every two months.

The stock cultures are grown in conical flasks containing the appropriate medium (volume about 100 ml). When the algae are incubated at 20 °C with continuous illumination, a weekly transfer is required.

During transfer an amount of 'old' culture is transferred with sterile pipettes into a flask of fresh medium, so that with the fast-growing species the initial concentration is about 100 times smaller than in the old culture.

The growth rate of a species can be determined from the growth curve. If this is known, it is possible to estimate the density at which the culture should be transferred to new medium. This must be done before the culture reaches the death phase.

*Pre-culture:*

The pre-culture is intended to give an amount of algae suitable for the inoculation of test cultures. The pre-culture is incubated under the conditions of the test and used when still exponentially growing, normally after an incubation period of 2 to 4 days. When the algal cultures contain deformed or abnormal cells, they must be discarded.

**▼ M1***Appendix 4***Data analysis by nonlinear regression****General considerations**

The response in algal tests and other microbial growth tests — growth of biomass is by nature a continuous or metric variable — a process rate if growth rate is used, and its integral over time if biomass is selected. Both are referenced to the corresponding mean response of replicate non-exposed controls showing maximum response for the conditions imposed — with light and temperature as primary determining factors in the algal test. The system is distributed or homogenous and the biomass can be viewed as a continuum without consideration of individual cells. The variance distribution of the type of response for such a system relates solely to experimental factors (described typically by the log-normal or normal distributions of error). This is by contrast to typical bioassay responses with quantal data, for which the tolerance (typically binomially distributed) of individual organisms is often assumed to be the dominant variance component. Control responses here are zero or background level.

In the uncomplicated situation, the normalised or relative response,  $r$ , decreases monotonically from 1 (zero inhibition) to 0 (100 per cent inhibition). Note that all responses have an associated error, and that apparent negative inhibitions can be calculated as a result of random error only.

**Regression analysis***Models*

A regression analysis aims at quantitatively describing the concentration-response curve in the form of a mathematical regression function  $Y = f(C)$  or more frequently  $F(Z)$  where  $Z = \log C$ . Used inversely  $C = f^{-1}(Y)$  allows the calculation of  $EC_x$  figures including the  $EC_{50}$ ,  $EC_{10}$  and  $EC_{20}$ , and their 95 % confidence limits. Several simple mathematical functional forms have proved to successfully describe concentration-response relationships obtained in algal growth inhibition tests. Functions include, for instance, the logistic equation, the non-symmetrical Weibul equation and the log normal distribution function, which are all sigmoid curves asymptotically approaching one for  $C \rightarrow 0$ , and zero for  $C \rightarrow \text{infinity}$ .

The use of continuous threshold function models (e.g. the Kooijman model 'for inhibition of population growth' Kooijman et al. 1996) is a recently proposed or alternative to asymptotic models. This model assumes no effects at concentrations below a certain threshold,  $EC_{0+}$ , that is estimated by extrapolation of the response concentration relationship to intercept the concentration axis using a simple continuous function that is not differentiable in the starting point.

Note that the analysis can be a simple minimisation of sums of residual squares (assuming constant variance) or weighted squares if variance heterogeneity is compensated.

▼ **M1***Procedure*

The procedure can be outlined as follows: select an appropriate functional equation,  $Y = f(C)$ , and fit it to the data by non-linear regression. Preferably use the measurements from each individual flask rather than the mean values of the replicates, in order to extract as much information from the data as possible. If the variance is high, on the other hand, practical experience suggests that the mean values of the replicates may provide a more robust mathematical estimation, less influenced by systematic errors in the data, than with each individual data point retained.

Plot the fitted curve and the measured data and examine whether the curve fit is appropriate. Analysis of residuals may be a particularly helpful tool for this purpose. If the chosen functional relationship to fit the concentration response does not describe the whole curve or some essential part of it, such as the response at low concentrations well, choose another curve fit option — e.g. a non-symmetrical curve like the Weibul function, instead of a symmetrical one. Negative inhibitions may be a problem with, for instance, the log-normal distribution function, likewise demanding an alternative regression function. It is not recommended to assign a zero or a small positive value to such negative values because this distorts the error distribution. It may be appropriate to make separate curve fits on parts of the curve such as the low inhibition part to estimate  $EC_{low\ x}$  figures. Calculate from the fitted equation (by ‘inverse estimation’,  $C = f^{-1}(Y)$ ), characteristic point estimates  $EC_x$ ’s, and report as a minimum the  $EC_{50}$  and one or two  $EC_{low\ x}$  estimates. Experience from practical testing has shown that the precision of the algal test normally allows a reasonably accurate estimation at the 10 % inhibition level if data points are sufficient — unless stimulation occurs at low concentrations as a confounding factor. The precision of an  $EC_{20}$  estimate is often considerably better than that of an  $EC_{10}$ , because the  $EC_{20}$  is usually positioned on the approximately linear part of the central concentration response curve. Sometimes  $EC_{10}$  can be difficult to interpret because of growth stimulation. So, while the  $EC_{10}$  is normally obtainable with a sufficient accuracy, it is also recommended to report always the  $EC_{20}$ .

*Weighting factors*

The experimental variance is not generally constant and typically includes a proportional component, a weighted regression is therefore advantageously carried out routinely. Weighting factors for such an analysis are normally assumed inversely proportional to the variance:

$$W_i = 1/\text{Var}(r_i)$$

Many regression programs allow the option of weighted regression analysis with weighting factors listed in a table. Conveniently, weighting factors should be normalised by multiplying them by  $n/\sum w_i$  ( $n$  is the number of data points) so that their sum equals one.

*Normalising responses*

Normalising by the mean control response gives some principle problems and gives rise to a rather complicated variance structure. Dividing the responses by the mean control response for obtaining the percentage of inhibition, one introduces an additional error caused by the error on the control mean. Unless this error is negligibly small, weighting factors in the regression and confidence limits must be corrected for the covariance with the control (17). Note that high precision on the estimated mean control response is important in order to minimise the overall variance for the relative response. This variance is as follows:

**▼ M1**

(subscript i refers to concentration level i and subscript 0 to the controls)

$$Y_i = \text{Relative response} = r_i/r_0 = I - I = f(C_i)$$

with a variance:

$$\text{Var}(Y_i) = \text{Var}(r_i/r_0) \cong (\partial Y_i / \partial r_i)^2 \cdot \text{Var}(r_i) + (\partial Y_i / \partial r_0)^2 \cdot \text{Var}(r_0)$$

and since

$$(\partial Y_i / \partial r_i) = 1/r_0 \text{ and } (\partial Y_i / \partial r_0) = r_i/r_0^2$$

with normally distributed data and  $m_i$  and  $m_0$  replicates:

$$\text{Var}(r_i) = \sigma^2/m_i$$

the total variance of the relative response,  $Y_i$  thus becomes:

$$\text{Var}(Y_i) = \sigma^2/(r_0^2 m_i) + r_i^2 \cdot \sigma^2/r_0^4 m_0$$

The error on the control mean is inversely proportional to the square root of the number of control replicates averaged, and sometimes it can be justified to include historical data and in this way greatly reduce the error. An alternative procedure is not to normalise the data and fit the absolute responses, including the control response data, but introducing the control response value as an additional parameter to be fitted by non-linear regression. With a usual 2 parameter regression equation, this method necessitates the fitting of 3 parameters, and therefore demands more data points than non-linear regression on data that are normalised using a pre-set control response.

*Inverse confidence intervals*

The calculation of non-linear regression confidence intervals by inverse estimation is rather complex and not an available standard option in ordinary statistical computer program packages. Approximate confidence limits may be obtained with standard non-linear regression programs with re-parameterisation (Bruce and Versteeg, 1992), which involves rewriting the mathematical equation with the desired point estimates, e.g. the  $EC_{10}$  and the  $EC_{50}$  as the parameters to be estimated. (Let the function be  $I = f(\alpha, \beta, \text{concentration})$  and utilise the definition relationships  $f(\alpha, \beta, EC_{10}) = 0,1$  and  $f(\alpha, \beta, EC_{50}) = 0,5$  to substitute  $f(\alpha, \beta, \text{concentration})$  with an equivalent function  $g(EC_{10}, EC_{50}, \text{concentration})$ ).

A more direct calculation (Andersen et al, 1998) is performed by retaining the original equation and using a Taylor expansion around the means of  $r_i$  and  $r_0$ .

Recently 'boot strap methods' have become popular. Such methods use the measured data and a random number generator directed frequent re-sampling to estimate an empirical variance distribution.

**Literature**

Kooijman, S.A.L.M.; Hanstveit, A.O.; Nyholm, N. (1996): No-effect concentrations in algal growth inhibition tests. *Water Research*, 30, 1625-1632.

Bruce, R.D. and Versteeg, D.J.(1992) A Statistical Procedure for Modelling Continuous Ecotoxicity Data. *Env. Toxicol. Chem.*11, 1485-1494.

Andersen, J.S., Holst, H., Spliid, H., Andersen, H., Baun, A. & Nyholm, N. (1998): Continuous ecotoxicological data evaluated relative to a control response. *Journal of Agricultural, Biological and Environmental Statistics*, 3, 405-420.

**▼B****C.4. DETERMINATION OF 'READY' BIODEGRADABILITY****PART I. GENERAL CONSIDERATIONS****I.1. INTRODUCTION**

Six test methods are described that permit the screening of chemicals for ready biodegradability in an aerobic aqueous medium:

- (a) Dissolved Organic Carbon (DOC) Die-Away (Method C.4-A)
- (b) Modified OECD Screening — DOC Die-Away (Method C.4-B)
- (c) Carbon dioxide (CO<sub>2</sub>) Evolution (Modified Sturm Test) (Method C.4-C)
- (d) Manometric Respirometry (Method C.4-D)
- (e) Closed Bottle (Method C.4-E)
- (f) MITI (Ministry of International Trade and Industry — Japan) (Method C.4-F)

General and common considerations to all six tests are given in Part I of the method. Items specific for individual methods are given in Parts II to VII. The appendices contain definitions, formulas and guidance material.

An OECD inter-laboratory comparison exercise, done in 1988, has shown that the methods give consistent results. However, depending on the physical characteristics of the substance to be tested, one or other of the methods may be preferred.

**I.2. SELECTION OF THE APPROPRIATE METHOD**

In order to select the most appropriate method, information on the chemical's solubility, vapour pressure and adsorption characteristics is essential. The chemical structure or formula should be known in order to calculate theoretical values and/or check measured values of parameters, e.g. ThOD, ThCO<sub>2</sub>, DOC, TOC, COD (see Appendices 1 and 2).

Test chemicals which are soluble in water to at least 100 mg/l may be assessed by all methods, provided they are non-volatile and non-adsorbing. For those chemicals which are poorly soluble in water, volatile or adsorbing, suitable methods are indicated in Table 1. The manner in which poorly water-soluble chemicals and volatile chemicals can be dealt with is described in Appendix 3. Moderately volatile chemicals may be tested by the DOC Die-Away method if there is sufficient gas space in the test vessels (which should be suitably stoppered). In this case, an abiotic control must be set up to allow for any physical loss.





Table 1:

#### Applicability of test methods

Test	Analytical Method	Suitability for substances which are:		
		poorly soluble	volatile	adsorbing
DOC Die-Away	Dissolved organic carbon	–	–	+/-
Mod. OECD Die-Away	Dissolved organic carbon	–	–	+/-
CO <sub>2</sub> Evolution	Respirometry: CO <sub>2</sub> evolution	+	–	+
Manometric Respirometry	Manometric respirometry: oxygen consumption	+	+/-	+
Closed Bottle	Respirometry: dissolved oxygen	+/-	+	+
MITI	Respirometry: oxygen consumption	+	+/-	+

Information on the purity or the relative proportions of major components of the test material is required to interpret the results obtained, especially when the results are low or marginal.

Information on the toxicity of the test chemical to bacteria (Appendix 4) may be very useful for selecting appropriate test concentrations and may be essential for the correct interpretation of low biodegradation values.

### I.3. REFERENCE SUBSTANCES

In order to check the procedure, reference chemicals which meet the criteria for ready biodegradability are tested by setting up an appropriate flask in parallel to the normal test runs.

Suitable chemicals are aniline (freshly distilled), sodium acetate and sodium benzoate. These reference chemicals all degrade in these methods even when no inoculum is deliberately added.

It was suggested that a reference chemical should be sought which was readily biodegradable but required the addition of an inoculum. Potassium hydrogen phthalate has been proposed but more evidence needs to be obtained with this substance before it can be accepted as a reference substance.

In the respirometric tests, nitrogen-containing compounds may affect the oxygen uptake because of nitrification (see Appendices 2 and 5).

### I.4. PRINCIPLE OF THE TEST METHODS

A solution, or suspension, of the test substance in a mineral medium is inoculated and incubated under aerobic conditions in the dark or in diffuse light. The amount of DOC in the test solution due to the inoculum should be kept as low as possible compared to the amount of DOC due to the test substance. Allowance is made for the endogenous activity of the inoculum by running parallel blank tests with inoculum but without test substance, although the endogenous activity of cells in the presence of the substance will not exactly match that in the endogenous control. A reference substance is run in parallel to check the operation of the procedures.

**▼ B**

In general, degradation is followed by the determination of parameters, such as DOC, CO<sub>2</sub> production and oxygen uptake, and measurements are taken at sufficiently frequent intervals to allow the identification of the beginning and end of biodegradation. With automatic respirometers the measurement is continuous. DOC is sometimes measured in addition to another parameter but this is usually done only at the beginning and the end of the test. Specific chemical analysis can also be used to assess primary degradation of the test substance, and to determine the concentration of any intermediate substances formed (obligatory in the MITI test).

Normally, the test lasts for 28 days. Tests however may be ended before 28 days, i.e. as soon as the biodegradation curve has reached a plateau for at least three determinations. Tests may also be prolonged beyond 28 days when the curve shows that biodegradation has started but that the plateau has not been reached day 28.

**I.5. QUALITY CRITERIA****I.5.1. Reproducibility**

Because of the nature of biodegradation and of the mixed bacterial populations used as inocula, determinations should be carried out at least in duplicate.

It is common experience that the larger the concentration of micro-organisms initially added to the test medium, the smaller will be the variation between replicates. Ring tests have also shown that there can be large variations between results obtained by different laboratories, but good agreement is normally obtained with easily biodegradable compounds.

**I.5.2. Validity of the test**

A test is considered valid if the difference of extremes of replicate values of the removal of test chemical at the plateau, at the end of the test or at the end of the 10-day window, as appropriate, is less than 20 % and if the percentage degradation of the reference substance has reached the level for ready biodegradability by 14 days. If either of these conditions is not met, the test should be repeated. Because of the stringency of the methods, low values do not necessarily mean that the test substance is not biodegradable under environmental conditions, but indicates that more work will be necessary to establish biodegradability.

If in a toxicity test, containing both the test substance and a reference chemical, less than 35 % degradation (based on DOC) or less than 25 % (based on ThOD or ThCO<sub>2</sub>) occurred in 14 days, the test chemicals can be assumed to be inhibitory (see also Appendix 4). The test series should be repeated, if possible using a lower concentration of test chemical and/or a higher concentration of inoculum, but not greater than 30 mg solids/litre.

**I.6. GENERAL PROCEDURES AND PREPARATIONS**

General conditions applying to the tests are summarised in Table 2. Apparatus and other experimental conditions pertaining specifically to an individual test are described later under the heading for that test.



Table 2

## Test conditions

Test	DOC Die-Away	CO <sub>2</sub> Evolution	Manometric Respirometry	Modified OECD Screening	Closed Bottle	MITI (l)
Concentration of Test Substance						
as mg/l			100		2-10	100
mg DOC/l	10-40	10-20		10-40		
mg ThoD/l			50-100		5-10	
Concentration of Inoculum (in cells/l, approximately)	≤ 30 mg/l SS or ≤ 100 ml effluent/l (10 <sup>7</sup> -10 <sup>8</sup> )			0,5 ml secondary effluent/l (10 <sup>5</sup> )	≤ 5 ml of effluent/l (10 <sup>4</sup> -10 <sup>6</sup> )	30 mg/l SS (10 <sup>7</sup> -10 <sup>8</sup> )
Concentration of elements in mineral medium (in mg/l):						
P						116
N						1,3
Na						86
K						122
Mg						2,2
Ca						9,9
Fe						0,05 - 0,1
pH	7,4 ± 0,2					preferably 7,0
Temperature	22 ± 2 °C					25 ± 1 °C
DOC = Dissolved organic Carbon			ThoD = Theoretical Oxygen Demand		SS = Suspended Solids	

## I.6.1. Dilution water

Deionised or distilled water, free from inhibitory concentrations of toxic substances (e.g. Cu<sup>++</sup> ions) is used. It must contain no more than 10 % of the organic carbon content introduced by the test material. The high purity of the test water is necessary to eliminate high blank values. Contamination may result from inherent impurities and also from the ion-exchange resins and lysed material from bacterial and algae. For each series of tests use only one batch of water, checked beforehand by DOC analysis. Such a check is not necessary for the closed bottle test, but the oxygen consumption of the water must be low.

**▼B****I.6.2. Stock solutions of mineral components**

To make up the test solutions, stock solutions of appropriate concentrations of mineral components are made up. The following stock solutions may be used (with different dilution factors) for the methods DOC Die-Away, Modified OECD Screening, CO<sub>2</sub> Evolution, Manometric Respirometry, Closed Bottle test.

The dilution factors and, for the MITI test, the specific preparation of the mineral medium are given under the headings of the specific tests.

*Stock solutions:*

Prepare the following stock solutions, using analytical grade reagents.

(a) Monopotassium dihydrogen orthophosphate, KH<sub>2</sub>PO<sub>4</sub> 8,50 g

Dipotassium monohydrogen orthophosphate, K<sub>2</sub>HPO<sub>4</sub> 21,75 g

Disodium monohydrogen orthophosphate dihydrate Na<sub>2</sub>HPO<sub>4</sub> · 2 H<sub>2</sub>O 33,40 g

Ammonium chloride, NH<sub>4</sub>Cl 0,50 g

Dissolve in water and make up to 1 litre The pH of the solution should be 7,4.

(b) Calcium chloride, anhydrous, CaCl<sub>2</sub> 27,50 g

or Calcium chloride dihydrate, CaCl<sub>2</sub> · 2 H<sub>2</sub>O 36,40 g

Dissolve in water and make up to 1 litre

(c) Magnesium sulphate heptahydrate, MgSO<sub>4</sub> · 7 H<sub>2</sub>O 22,50 g

Dissolve in water and make up to 1 litre.

(d) Iron (III) chloride hexahydrate, FeCl<sub>3</sub> · 6H<sub>2</sub>O 0,25 g

Dissolve in water and make up to 1 litre.

*Note:* in order to avoid having to prepare this solution immediately before use add one drop of conc. HCL or 0,4 g ethylenediaminetetraacetic acid disodium salt (EDTA) par litre.

**I.6.3. Stock solutions of chemicals**

For example, dissolve 1-10 g, as appropriate, of test or reference chemical in deionised water and make up to 1 litre when the solubility exceeds 1 g/l. Otherwise, prepare stock solutions in the mineral medium or add the chemical direct to the mineral medium. For the handling of less soluble chemicals, see Appendix 3, but in the MITI test (Method C.4-F), neither solvents nor emulsifying agents are to be used.

**▼B****I.6.4. Inocula**

The inoculum may be derived from a variety of sources: activated sludge, sewage effluents (unchlorinated), surface waters and soils or from a mixture of these. For the DOC Die-Away, CO<sub>2</sub> Evolution and Manometric Respirometry tests, if activated sludge is used, it should be taken from a treatment plant or laboratory-scale unit receiving predominantly domestic sewage. Inocula from other sources have been found to give higher scattering of results. For the Modified OECD Screening and the Closed Bottle tests a more dilute inoculum without sludge flocs is needed and the preferred source is a secondary effluent from a domestic waste water treatment plant or laboratory-scale unit. For the MITI test the inoculum is derived from a mixture of sources and is described under the heading of this specific test.

**I.6.4.1. *Inoculum from activated sludges***

Collect a sample of activated sludge freshly from the aeration tank of a sewage treatment plant or laboratory-scale unit treating predominantly domestic sewage. Remove coarse particles if necessary by filtration through a fine sieve and keep the sludge aerobic thereafter.

Alternatively, settle or centrifuge (e.g. at 100 g for 10 min) after removal of any coarse particles. Discard the supernatant. The sludge may be washed in the mineral medium. Suspend the concentrated sludge in mineral medium to yield a concentration of 3-5 g suspended solids/l and aerate until required.

Sludge should be taken from a properly working conventional plant. If sludge has to be taken from a high rate treatment plant, or is thought to contain inhibitors, it should be washed. Settle or centrifuge the re-suspended sludge after thorough mixing, discard the supernatant and again re-suspend the washed sludge in a further volume of mineral medium. Repeat this procedure until the sludge is considered to be free from excess substrate or inhibitor.

After complete re-suspension is achieved, or with untreated sludge, withdraw a sample just before use for the determination of the dry weight of the suspended solids.

A further alternative is to homogenise activated sludge (3-5 g suspended solids/l). Treat the sludge in a mechanical blender for two min at medium speed. Settle the blended sludge for 30 min or longer if required and decant liquid for use as inoculum at the rate of 10 mill of mineral medium.

**I.6.4.2. *Other sources of inoculum***

It can be derived from the secondary effluent of a treatment plant or laboratory-scale unit receiving predominantly domestic sewage. Collect a fresh sample and keep it aerobic during transport. Allow to settle for 1 h. or filter through a coarse filter paper and keep the decanted effluent or filtrate aerobic until required. Up to 100 ml of this type of inoculum may be used per litre of medium.

**▼B**

A further source for the inoculum is surface water. In this case, collect a sample of an appropriate surface water, e.g. river, lake, and keep aerobic until required. If necessary, concentrate the inoculum by filtration or centrifugation.

**1.6.5. Pre-conditioning of inocula**

Inocula may be pre-conditioned to the experimental conditions, but not pre-adapted to the test chemical. Pre-conditioning consists of aerating activated sludge in mineral medium or secondary effluent for five to seven days at the test temperature. Pre-conditioning sometimes improves the precision of the test methods by reducing blank values. It is considered unnecessary to pre-condition MITI inoculum.

**1.6.6. Abiotic controls**

When required, check for the possible abiotic degradation of the test substance by determining the removal of DOC, oxygen uptake or carbon dioxide evolution in sterile controls containing no inoculum. Sterilise by filtration through a membrane (0,2 - 0,45 micrometre) or by the addition of a suitable toxic substance at an appropriate concentration. If membrane filtration is used, take samples aseptically to maintain sterility. Unless adsorption of the test chemical has been ruled out beforehand, tests which measure biodegradation as the removal of DOC, especially with activated sludge inocula, should include an abiotic control which is inoculated and poisoned.

**1.6.7. Number of flasks**

The number of flasks in a typical run is described under the headings of each tests.

The following type of flask may be used:

- test suspension: containing test substance and inoculum
- inoculum blank: containing only inoculum
- procedure control: containing reference substance and inoculum
- abiotic sterile control: sterile, containing test substance (see 1.6.6)
- adsorption control: containing test substance, inoculum and sterilising agent
- toxicity control: containing test substance, reference substance and inoculum

It is mandatory that determination in test suspension and inoculum blank is made in parallel. It is advisable to make the determinations in the other flasks in parallel as well.

This may, however, not always be possible. Ensure that sufficient samples or readings are taken to allow the percentage removal in the 10-day window to be assessed.

**▼B****I.7. DATA AND EVALUATION**

In the calculation of  $D_t$ , percentage degradation, the mean values of the duplicate measurement of the parameter in both test vessels and inoculum blank are used. The formulas are set out in the sections below on specific tests. The course of degradation is displayed graphically and the 10-day window is indicated. Calculate and report the percentage removal achieved at the end of the 10-day window and the value at the plateau or at the end of the test, whichever is appropriate.

In respirometric tests nitrogen-containing compounds may affect the oxygen uptake because of nitrification (see Appendices 2 and 5).

**I.7.1. Degradation measured by means of DOC determination**

The percentage degradation  $D_t$  at each time a sample was taken should be calculated separately for the flasks containing test substance using mean values of duplicate DOC measurements in order that the validity of the test can be assessed (see 1.5.2). It is calculated using the following equation:

$$D_t = \left( 1 - \frac{C_t - C_{bt}}{C_o - C_{bo}} \right) \times 100$$

where:

$D_t$  = % degradation at time t

$C_o$  = mean starting concentration of DOC in the inoculated culture medium containing the test substance (mg DOC/l)

$C_t$  = mean concentration of DOC in the inoculated culture medium containing test substance at time t (mg DOC/l)

$C_{bo}$  = mean starting concentration of DOC in blank inoculated mineral medium (mg DOC/l)

$C_{bt}$  = mean concentration of DOC blank inoculated mineral medium at time t (mg DOC/l).

All concentrations are measured experimentally.

**I.7.2. Degradation measured by means of specific analysis**

When specific analytical data are available, calculate primary biodegradation from:

$$D_t = \frac{S_b - S_a}{S_b} \times 100$$

where:

$D_t$  = % degradation at time t, normally 28 days,

$S_a$  = residual amount of test substance in inoculated medium at end of test (mg),

$S_b$  = residual amount of test substance in the blank test with water/medium to which only the test substance was added (mg).

**▼ B****1.7.3. Abiotic degradation**

When an abiotic sterile control is used, calculate the percentage abiotic degradation using:

$$\% \text{ abiotic degradation} = \frac{C_{s(0)} - C_{s(t)}}{C_{s(0)}} \times 100$$

Where:

$C_{s(0)}$  = DOC Concentration in sterile control at day 0

$C_{s(t)}$  = DOC Concentration in sterile control at day t.

**1.8. REPORTING**

The test report shall, if possible, contain the following:

- test and reference chemicals, and their purity,
- test conditions,
- inoculum: nature and sampling site(s), concentration and any pre-conditioning treatment,
- proportion and nature of industrial waste present in sewage if known,
- test duration and temperature,
- in the case of poorly soluble test chemicals, treatment given,
- test method applied; scientific reasons and explanation should be given for any change of procedure,
- data sheet,
- any observed inhibition phenomena,
- any observed abiotic degradation,
- specific chemical analytical data, if available,
- analytical data on intermediates, if available,
- the graph of percentage degradation against time for the test and reference substances; the lag phase, degradation phase, 10-day window and slope should be clearly indicated (Appendix 1). If the test has complied with the validity criteria, the mean of the degradation percentages of the flasks containing test substance may be used for the graph,
- percentage removal after 10-day window, and at plateau or at end of the test.

**PART II. DOC DIE-AWAY TEST (Method C.4-A)****II.1. PRINCIPLE OF THE METHOD**

A measured volume of inoculated mineral medium containing a known concentration of the test substance (10-40 mg DOC/l) as the nominal sole source of organic carbon is aerated in the dark or diffused light at  $22 \pm 2$  °C.



**▼ B**

Degradation is followed by DOC analysis at frequent intervals over a 28-day period. The degree of biodegradation is calculated by expressing the concentration of DOC removed (corrected for that in the blank inoculum control) as a percentage of the concentration initially present. The degree of primary biodegradation may also be calculated from supplemental chemical analysis made at the beginning and end of incubation.

**II.2. DESCRIPTION OF THE METHOD****II.2.1. Apparatus**

- (a) Conical flasks, e. g. 250 ml to 2 l, depending on the volume needed for DOC analysis;
- (b) shaking machine to accommodate the conical flasks, either with automatic temperature control or used in a constant temperature room; and of sufficient power to maintain aerobic conditions in all flasks;
- (c) filtration apparatus, with suitable membranes;
- (d) DOC analyser;
- (e) apparatus for determining dissolved oxygen;
- (f) centrifuge.

**II.2.2. Preparation of mineral medium**

For the preparation of the stock solutions, see I.6.2.

Mix 10 ml of solution (a) with 800 ml dilution water, add 1 ml of solutions (b) to (d) and make up to 1 l with dilution water.

**II.2.3. Preparation and pre-conditioning of inoculum**

The inoculum may be derived from a variety of sources: activated sludge; sewage effluents; surface waters; soils or from a mixture of these.

See I.6.4., I.6.4.1., I.6.4.2. and I.6.5.

**II.2.4. Preparation of flasks**

As an example, introduce 800 ml portions of mineral medium into 2 l conical flasks and add sufficient volumes of stock solutions of the test and reference substances to separate flasks to give a concentration of chemical equivalent to 10-40 mg DOC/l. Check the pH values and adjust, if necessary, to 7.4. Inoculate the flasks with activated sludge or other source of inocula (see I.6.4.), to give a final concentration not greater than 30 mg suspended solids/l. Also prepare inoculum controls in the mineral medium but without test or reference chemical.

If needed, use one vessel to check the possible inhibitory effect of the test chemical by inoculating a solution containing, in the mineral medium, comparable concentrations of both the test and a reference chemical.

Also, if required, set up a further, sterile flask to check whether the test chemical is degraded abiotically by using an uninoculated solution of the chemical (see I.6.6).

**▼B**

Additionally, if the test chemical is suspected of being significantly adsorbed on to glass, sludge, etc., make a preliminary assessment to determine the likely extent of adsorption and thus the suitability of the test for the chemical (see Table 1). Set up a flask containing the test substance, inoculum and sterilising agent.

Make up the volumes in all flasks to 11 with mineral medium and, after mixing, take a sample from each flask to determine the initial concentration of DOC (see Appendix 2.4). Cover the openings of the flasks, e.g. with aluminium foil, in such a way as to allow free exchange of air between the flask and the surrounding atmosphere. Then insert the vessels into the shaking machine for starting the test.

**II.2.5. Number of flasks in typical run**

Flasks 1 and 2: Test suspension

Flasks 3 and 4: Inoculum blank

Flask 5: Procedure control

preferably and when necessary:

Flask 6: Abiotic sterile control

Flask 7: Adsorption control

Flask 8: Toxicity control

See also I.6.7.

**II.2.6. Performance of the test**

Throughout the test, determine the concentrations of DOC in each flask in duplicate at known time intervals, sufficiently frequently to be able to determine the start of the 10-day window and the percentage removal at the end of the 10-day window. Take only the minimal volume of test suspension necessary for each determination.

Before sampling make good evaporation losses from the flasks by adding dilution water (I.6.1) in the required amount if necessary. Mix the culture medium thoroughly before withdrawing a sample and ensure that material adhering to the walls of the vessels is dissolved or suspended before sampling. Membrane-filter or centrifuge (see Appendix 2.4) immediately after the sample has been taken. Analyse the filtered or centrifuged samples on the same day, otherwise store at 2-4 °C for a maximum of 48 h, or below - 18 °C for a longer period.

**II.3. DATA AND REPORTING****II.3.1. Treatment of results**

Calculate the percentage degradation at time t as given under I.7.1 (DOC determination) and, optionally, under I.7.2 (specific analysis).

Record all results on the data sheets provided.

**▼ B****II.3.2. Validity of results**

See I.5.2.

**II.3.3. Reporting**

See I.8.

**II.4. DATA SHEET**

An example of a data sheet is given hereafter.

**DOC DIE-AWAY TEST****1. LABORATORY****2. DATE AT START OF TEST****3. TEST SUBSTANCE**

Name:

Stock solution concentration: ... mg/l as chemical

Initial concentration in medium, to: ... mg/l as chemical

**4. INOCULUM**

Source:

Treatment given:

Pre-conditioning, if any:

Concentration of suspended solids in reaction mixture: mg/l

**5. CARBON DETERMINATIONS**

Carbon analyser:

	Flask nr		DOC after n-days (mg/l)				
			0	n <sub>1</sub>	n <sub>2</sub>	n <sub>3</sub>	n <sub>x</sub>
Test chemical plus inoculum	1	a <sub>1</sub>					
		a <sub>2</sub>					
		a, mean C <sub>a(t)</sub>					
	2	b <sub>1</sub>					
		b <sub>2</sub>					
		b, men C <sub>b(t)</sub>					

**▼ B**

	Flask nr		DOC after n-days (mg/l)				
			0	n <sub>1</sub>	n <sub>2</sub>	n <sub>3</sub>	n <sub>x</sub>
Blank inoculum without test chemical	3	C <sub>1</sub>					
		C <sub>2</sub>					
		C, mean C <sub>c(t)</sub>					
	4	d <sub>1</sub>					
		d <sub>2</sub>					
		d, mean C <sub>d(t)</sub>					
$C_{bl(t)} = \frac{C_{c(t)} + C_{d(t)}}{2}$							

**6. EVALUATION OF RAW DATA**

Flask nr		% degradation after n days				
		0	n <sub>1</sub>	n <sub>2</sub>	n <sub>3</sub>	n <sub>x</sub>
1	$D_1 = \left(1 - \frac{C_{a(t)} - C_{bl(t)}}{C_{a(0)} - C_{bl(0)}}\right) \times 100$	0				
2	$D_2 = \left(1 - \frac{C_{b(t)} - C_{bl(t)}}{C_{b(0)} - C_{bl(0)}}\right) \times 100$	0				
Mean (*)	$D = \frac{D_1 - D_2}{2}$	0				

(\*) D<sub>1</sub> and D<sub>2</sub> should not be averaged if there is a considerable difference.

Note: similar formats may be used for the reference chemical and toxicity controls.

**7. ABIOTIC CONTROL (optional)**

	Time (days)	
	0	t
DOC conc. (mg/l) in sterile control	C <sub>s(0)</sub>	C <sub>s(t)</sub>

$$\% \text{ abiotic degradation} = \frac{C_{s(0)} - C_{s(t)}}{C_{s(0)}} \times 100$$

**8. SPECIFIC CHEMICAL ANALYSIS (optional)**

	residual amount of test chemical at end of test (mg/l)	% primary degradation
Sterile control	S <sub>b</sub>	

**▼ B**

	residual amount of test chemical at end of test (mg/l)	% primary degradation
Inoculated test medium	$S_a$	$\frac{S_b - S_a}{S_b} \times 100$

**PART III. MODIFIED OECD SCREENING TEST**  
(Method C.4-B)

III.1. PRINCIPLE OF THE METHOD

A measured volume of mineral medium containing a known concentration of the test substance (10-40 mg DOC/litre) as the nominal sole source of organic carbon is inoculated with 0,5 ml effluent per litre of medium. The mixture is aerated in the dark or diffused light at  $22 \pm 2$  °C.

Degradation is followed by DOC analysis at frequent intervals over a 28-day period. The degree of biodegradation is calculated by expressing the concentration of DOC removed (corrected for that in the blank inoculum control) as a percentage of the concentration initially present. The degree of primary biodegradation may also be calculated from supplemental chemical analysis made at the beginning and end of incubation.

III.2. DESCRIPTION OF THE METHOD

III.2.1. Apparatus

- (a) Conical flasks, e.g. 250 ml to 2 litres, depending on the volume needed for DOC analysis;
- (b) shaking machine — to accommodate the conical flasks, either with automatic temperature control or used in a constant temperature room, and of sufficient power to maintain aerobic conditions in all flasks;
- (c) filtration apparatus, with suitable membranes;
- (d) DOC analyser;
- (e) apparatus for determining dissolved oxygen;
- (f) centrifuge.

III.2.2. **Preparation of mineral medium**

For the preparation of the stock solutions, see I.6.2.

Mix 10 ml of solution (a) with 80 ml dilution water, add 1 ml of solutions (b) to (d) and make up to 1 litre with dilution water.

This method uses only 0,5 ml effluent/litre as inoculum and therefore the medium may need to be fortified with trace elements and growth factors. This is done by adding 1 ml each of the following solutions per litre of final medium:

**▼ B**

Trace element solution:

Manganese sulfate tetrahydrate, $\text{MnSO}_4 \cdot 4\text{H}_2\text{O}$	39,9 mg
Boric acid, $\text{H}_3\text{BO}_3$	57,2 mg
Zinc sulfate heptahydrate, $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$	42,8 mg
Ammonium heptamolybdate $(\text{NH}_4)_6 \text{Mo}_7\text{O}_{24}$	34,7 mg
Fe-chelate ( $\text{FeCl}_3$ ethylenediamine-tetra-acetic acid)	100,0 mg
Dissolve in, and make up to 1 000 ml with dilution water	

Vitamin solution:

Yeast extract	15,0 mg
---------------	---------

Dissolve the yeast extract in 100 ml water. Sterilise by passage through a 0,2 micron membrane, or make up freshly.

### III.2.3. Preparation and pre-conditioning of inoculum

The inoculum is derived from the secondary effluent of a treatment plant or laboratory scale unit receiving predominantly domestic sewage. See I.6.4.2. and I.6.5.

0,5 ml per litre of mineral medium is used.

### III.2.4. Preparation of flasks

As an example, introduce 800 ml portions of mineral medium into 2-litre conical flasks and add sufficient volumes of stock solutions of the test and reference substances to separate flasks to give a concentration of chemical equivalent to 10-40 mg DOC/litre. Check the pH value and adjust, if necessary, to 7,4. Inoculate the flasks with sewage effluent at 0,5 ml/litre (see I.6.4.2). Also prepare inoculum controls in the mineral medium but without test or reference chemical.

If needed, use one vessel to check the possible inhibitory effect of the test chemical by inoculating a solution containing, in the mineral medium, comparable concentrations of both the test and a reference chemical.

Also, if required, set up a further, sterile flask to check whether the test chemical is degraded abiotically by using an uninoculated solution of the chemical (see I.6.6).

Additionally, if the test chemical is suspected of being significantly adsorbed on to glass, sludge, etc., make a preliminary assessment to determine the likely extent of adsorption and thus the suitability of the test for the chemical (see Table 1). Set up a flask containing the test substance, inoculum and sterilising agent.

Make up the volumes in all flasks to 1 litre with mineral medium and, after mixing, take a sample from each flask to determine the initial concentration of DOC (see Appendix 2.4). Cover the openings of the flasks, e.g. with aluminium foil, in such a way as to allow free exchange of air between the flask and the surrounding atmosphere. Then insert the vessels into the shaking machine for starting the test.

**▼B****III.2.5. Number of flasks in typical run**

Flasks 1 and 2: test suspension

Flasks 3 and 4: inoculum blank

Flask 5: procedure control

and preferably and when necessary:

Flask 6: abiotic sterile control

Flask 7: adsorption control

Flask 8: toxicity control

See also 1.6.7.

**III.2.6. Performance of the test**

Throughout the test, determine the concentrations of DOC in each flask in duplicate at known time intervals, sufficiently frequently to be able to determine the start of the 10-day window and the percentage removal at the end of the 10-day window. Take only the minimal volume of test suspension necessary for each determination.

Before sampling make good evaporation losses from the flasks by adding dilution water (I.6.1) in the required amount if necessary. Mix the culture medium thoroughly before withdrawing a sample and ensure that material adhering to the walls of the vessels is dissolved or suspended before sampling. Membrane-filter or centrifuge (see Appendix 2.4) immediately after the sample has been taken. Analyse the filtered or centrifuged samples on the same day, otherwise store at 2-4 °C for a maximum of 48 h, or below 18 °C for a longer period.

**III.3. DATA AND REPORTING****III.3.1. Treatment of results**

Calculate the percentage degradation at time t as given under I.7.1 (DOC determination) and, optionally, under I.7.2 (specific analysis).

Record all results on the data sheets provided.

**III.3.2. Validity of results**

See I.5.2.

**III.3.3. Reporting**

See I.8.

**III.4. DATA SHEET**

An example of a data sheet is given hereafter

**MODIFIED OECD SCREENING TEST****1. LABORATORY****2. DATE AT START OF TEST**

**▼ B****3. TEST SUBSTANCE**

Name:

Stock solution concentration: ... mg/litre as chemical

Initial concentration in medium, to: ... mg/litre as chemical

**4. INOCULUM**

Source:

Treatment given:

Pre-conditioning, if any:

Concentration of suspended solids in reaction mixture: mg/l

**5. CARBON DETERMINATIONS**

Carbon analyser:

	Flask nr		DOC after n-days (mg/l)				
			0	n <sub>1</sub>	n <sub>2</sub>	n <sub>3</sub>	n <sub>x</sub>
Test chemical plus inoculum	1	a <sub>1</sub>					
		a <sub>2</sub>					
		a, mean C <sub>a(t)</sub>					
	2	b <sub>1</sub>					
		b <sub>2</sub>					
		b, men C <sub>b(t)</sub>					
Blank inoculum without test chemical	3	C <sub>1</sub>					
		C <sub>2</sub>					
		C, mean C <sub>c(t)</sub>					
	4	d <sub>1</sub>					
		d <sub>2</sub>					
		d, mean C <sub>d(t)</sub>					
	$C_{bl(t)} = \frac{C_{c(t)} + C_{d(t)}}{2}$						

**6. EVALUATION OF RAW DATA**

Flask nr		% degradation afer n days				
		0	n <sub>1</sub>	n <sub>2</sub>	n <sub>3</sub>	n <sub>x</sub>
1	$D_1 = \left( 1 - \frac{C_{a(t)} - C_{bl(t)}}{C_{a(o)} - C_{bl(o)}} \right) \times 100$	0				



**▼ B**

Flask nr		% degradation afer n days				
		0	n <sub>1</sub>	n <sub>2</sub>	n <sub>3</sub>	n <sub>x</sub>
2	$D_2 = \left( 1 - \frac{C_{b(t)} - C_{bl(t)}}{C_{b(o)} - C_{bl(o)}} \right) \times 100$	0				
Mean (*)	$D = \frac{D_1 - D_2}{2}$	0				

(\*) D<sub>1</sub> and D<sub>2</sub> should not be averaged if there is a considerable difference.

Note: similar formats may be used for the reference chemical and toxicity controls.

#### 7. ABIOTIC CONTROL (optional)

	Time (days)	
	0	t
DOC conc. (mg/l) in sterile control	C <sub>s(o)</sub>	C <sub>s(t)</sub>

$$\% \text{ abiotic degradation} = \frac{C_{s(o)} - C_{s(t)}}{C_{s(o)}} \times 100$$

#### 8. SPECIFIC CHEMICAL ANALYSIS (optional)

	residual amount of test chemical at end of test (mg/l)	% primary degradation
Sterile control	S <sub>b</sub>	
Inoculated test medium	S <sub>a</sub>	$\frac{S_b - S_a}{S_b} \times 100$

### PART IV. CO<sub>2</sub> EVOLUTION TEST (Method C.4-C)

#### IV.1. PRINCIPLE OF THE METHOD

A measured volume of inoculated mineral medium containing a known concentration of the test chemical (10-20 mg DOC or TOC/l) as the nominal sole source of organic carbon is aerated by the passage of carbon, dioxide-free air at a controlled rate in the dark or in diffuse light. Degradation is followed over 28 days by determining the carbon dioxide produced, which is trapped in barium or sodium hydroxide and which is measured by titration of the residual hydroxide or as inorganic carbon. The amount of carbon dioxide produced from the test chemical (corrected for that derived from the blank inoculum) is expressed as a percentage of ThCO<sub>2</sub>. The degree of biodegradation may also be calculated from supplemental DOC analysis made at the beginning and end of incubation.

**▼ B**

## IV.2. DESCRIPTION OF THE METHOD

IV.2.1. **Apparatus**

- (a) Flasks, 2-5 litres, each fitted with an aeration tube reaching nearly the bottom of the vessel and an outlet;
- (b) magnetic stirrers, when assessing poorly soluble chemicals;
- (c) gas-absorption bottles;
- (d) device for controlling and measuring airflow;
- (e) apparatus for carbon dioxide scrubbing, for preparation of air which is free from carbon dioxide; alternatively, a mixture of CO<sub>2</sub>-free oxygen and CO<sub>2</sub>-free nitrogen, from gas cylinders, in the correct proportions (20 % O<sub>2</sub>: 80 % N<sub>2</sub>) may be used;
- (f) device for determination of carbon dioxide, either titrimetrically or by some form of inorganic carbon analyser;
- (g) membrane filtration device (optional);
- (h) DOC analyser (optional).

IV.2.2. **Preparation of mineral medium**

For the preparation of the stock solutions, see I.6.2.

Mix 10 ml of solution (a) with 800 ml dilution water, add 1 ml of solutions (b) to (d) and make up to 1 l with dilution water.

IV.2.3. **Preparation and pre-conditioning of inoculum**

The inoculum may be derived from a variety of sources: activated sludge; sewage effluents; surface waters; soils or from a mixture of these.

See I.6.4., I.6.4.1., I.6.4.2. and I.6.5.

IV.2.4. **Preparation of flasks**

As an example the following volumes and weights indicate the values for 5-litre flasks containing 3 l of suspension. If smaller volumes are used modify the values accordingly, but ensure that the carbon dioxide formed can be measured accurately.

To each 5-litre flask add 2400 ml mineral medium. Add an appropriate volume of the prepared activated sludge (see I.6.4.1 and I.6.5) to give a concentration of suspended solids of not more than 30 mg/l in the final 3 l of inoculated mixture. Alternatively first dilute the prepared sludge to give a suspension of 500-1000 mg/l in the mineral medium before adding an aliquot to the contents of the 5 litre flask to attain a concentration of 30 mg/l; this ensures greater precision. Other sources of inoculum may be used (see I.6.4.2.).

Aerate these inoculated mixtures with CO<sub>2</sub>-free air overnight to purge the system of carbon dioxide.

**▼B**

Add the test material and reference substance, separately, as known volume of stock solutions, to replicate flasks to yield concentrations, contributed by the added chemicals, of 10 to 20 mg DOC or TOC/l; leave some flasks without addition of chemicals as inoculum controls. Add poorly soluble test substances directly to the flasks on a weight or volume basis or handle as described in Appendix 3.

If required, use one flask to check the possible inhibitory effect of the test chemical by adding both the test and reference chemicals at the same concentrations as present in the other flasks.

Also, if required, use a sterile flask to check whether the test chemical is degraded abiotically by using an uninoculated solution of the chemical (see I.6.6). Sterilise by the addition of a toxic substance at an appropriate concentration.

Make up the volumes of suspensions in all flasks to 3 l by the addition of mineral medium previously aerated with CO<sub>2</sub>-free air. Optionally, samples may be withdrawn for analysis of DOC (see Appendix 2.4.) and/or specific analysis. Connect the absorption bottles to the air outlets of the flasks.

If barium hydroxide is used, connect three absorption bottles, each containing 100 ml of 0,0125 M barium hydroxide solution, in series to each 5-litre flask. The solution must be free of precipitated sulphate and carbonate and its strength must be determined immediately before use. If sodium hydroxide is used, connect two traps, the second acting as a control to demonstrate that all the carbon dioxide was absorbed in the first. Absorption bottles fitted with serum bottle closures are suitable. Add 200 ml 0,05 M sodium hydroxide to each bottle, which is sufficient to absorb the total quantity of carbon dioxide evolved when the test chemical is completely degraded. The sodium hydroxide solution, even when freshly prepared, will contain traces of carbonates; this is corrected by deduction of the carbonate in the blank.

**IV.2.5. Number of flasks in a typical run**

Flasks 1 and 2: Test suspension

Flasks 3 and 4: Inoculum blank

Flask 5: Procedure control

and, preferably and when necessary:

Flask 6: Abiotic sterile control

Flask 7: Toxicity control

See also I.6.7.

**IV.2.6. Performance of the test**

Start the test by bubbling CO<sub>2</sub>-free air through the suspensions at a rate of 30-100 ml/min. Take samples of the carbon dioxide absorbent periodically for analysis of the CO<sub>2</sub>-content. During the first ten days it is recommended that analyses should be made every second or third day and then every fifth day until the 28th day so that the 10-day window period can be identified.

**▼ B**

On the 28th day, withdraw samples (optionally) for DOC and/or specific analysis, measure the pH of the suspensions and add 1 ml of concentrated hydrochloric acid to each flask; aerate the flasks overnight to drive off the carbon dioxide present in the test suspensions. On day 29 make the last analysis of evolved carbon dioxide.

On the days of measurement of CO<sub>2</sub>, disconnect the barium hydroxide absorber closest to the flask and titrate the hydroxide solution with HCl 0,05 M using phenolphthalein as the indicator. Move the remaining absorbers one place closer to the flask and place a new absorber containing 100 ml fresh 0,0125 M barium hydroxide at the far end of the series. Make titrations as needed, for example, when substantial precipitation is seen in the first trap and before any is evident in the second, or at least weekly. Alternatively, with NaOH as absorbent, withdraw with a syringe a small sample (depending on the characteristics of the carbon analyser used) of the sodium hydroxide solution in the absorber nearer to the flask. Inject the sample into the IC part of the carbon analyser for analysis of evolved carbon dioxide directly.

Analyse the contents of the second trap only at the end of the test to correct for any carryover of carbon dioxide.

### IV.3. DATA AND REPORTING

#### IV.3.1. Treatment of results

The amount of CO<sub>2</sub> trapped in an absorber when titrated is given by:

$$\text{mgCO}_2 = (100 \times C_B - 0,5 \times V \times C_A) \times 44$$

where:

V = volume of HCl used for titration of the 100 ml in the absorber (ml)

C<sub>B</sub> = concentration of the barium hydroxide solution (M)

C<sub>A</sub> = concentration of the hydrochloric acid solution (M)

if C<sub>B</sub> is 0,0125 M and C<sub>A</sub> is 0,05 M, the titration for 100 ml barium hydroxide is 50 ml and the weight of CO<sub>2</sub> is given by:

$$\frac{0,05}{2} \times 44 \times \text{ml HCl titrated} = 1,1 \times \text{ml HCl}$$

Thus, in this case, to convert volume of HCl titrated to mg CO<sub>2</sub> produced the factor is 1,1.

Calculate the weights of CO<sub>2</sub> produced from the inoculum alone and from the inoculum plus test chemical using the respective titration values and the difference is the weight of CO<sub>2</sub> produced from the test chemical alone.

For example, if the inoculum alone gives a titration of 48 ml and inoculum plus test chemical gives 45 ml,

$$\text{CO}_2 \text{ from inoculum} = 1,1 \times (50-48) = 2,2 \text{ mg}$$

**▼ B**

$\text{CO}_2$  from inoculum plus test chemical =  $1,1 \times (50-45) = 5,5$  mg

and thus the weight of  $\text{CO}_2$  produced from the test chemical is 3,3 mg.

The percentage biodegradation is calculated from:

$$\% \text{ degradation} = \frac{\text{mg CO}_2 \text{ produced} \times 100}{\text{ThCO}_2 \times \text{mg test chemical added}}$$

or,

$$\% \text{ degradation} = \frac{\text{mg CO}_2 \text{ produced} \times 100}{\text{mg TO added in test} \times 3,67}$$

3,67 being the conversion factor (44/12) for carbon to carbon dioxide.

Obtain the percentage degradation after any time interval by adding the percentage of  $\text{ThCO}_2$  values calculated for each of the days, up to that time, on which it was measured.

For sodium hydroxide absorbers, calculate the amount of carbon dioxide produced, expressed as IC (mg), by multiplying the concentration of IC in the absorbent by the volume of the absorbent.

Calculate the percentage degradation from:

$$\% \text{ of ThCO}_2 = \frac{\text{mg IC flask} - \text{mg IC blank}}{\text{MG TOC added as test chemical}} \times 100$$

Calculate DOC removals (optional) as described under I.7. Record these and all other results on the data sheets provided.

#### IV.3.2. **Validity of results**

The IC content of the test chemical suspension in the mineral medium at the beginning of the test must be less than 5 % of the TC, and the total  $\text{CO}_2$  evolution in the inoculum blank at the end of the test should not normally exceed 40 mg/l medium. If values greater than 70 mg  $\text{CO}_2$ /litre are obtained, the data and experimental technique should be examined critically.

See also I.5.2.

#### IV.3.3. **Reporting**

See I.8.

#### IV.4. DATA SHEET

An example of a data sheet is given hereafter.

#### CARBON DIOXIDE EVOLUTION TEST

1. **LABORATORY**
2. **DATE AT START OF TEST**
3. **TEST SUBSTANCE**

Name:

Stock solution concentration: ... mg/litre as chemical

**▼B**

Initial conc. in medium: ... mg/litre as chemical

Total C added to flask: ... mg C

ThCO<sub>2</sub>: mg CO<sub>2</sub>

**4. INOCULUM**

Source:

Treatment given:

Pre-conditioning if any:

Concentration of suspended solids in reaction mixture: mg/litre

Time (day)	CO <sub>2</sub> formed Test (mg)		CO <sub>2</sub> formed blank (mg)		CO <sub>2</sub> formed cumulative (mg) (test minus blank mean)		ThCO <sub>2</sub> cumulative $\frac{\text{CO}_2}{\text{ThCO}_2} \times 100$		
	1 2	mean	3 4	mean	1	2	1	2	mean
0									
n <sub>1</sub>									
n <sub>2</sub>									
n <sub>3</sub>									
28									

*Note:* similar formats may be used for the reference chemical and toxicity controls.

**6. CARBON ANALYSIS (optional)**

Carbon analyser:

Time (day)	Blank mg/l	Test chemical mg/l
0	C <sub>b(0)</sub>	C <sub>o</sub>
28 (*)	C <sub>b(t)</sub>	C <sub>t</sub>

(\*) Or at end of incubation

**▼ B**

$$\% \text{ DOC removed} = \left( 1 - \frac{C_t - C_{b(t)}}{C_o - C_{b(o)}} \right) \times 100$$

**7. ABIOTIC DEGRADATION** (optional)

$$\% \text{ abiotic degradation} = \frac{\text{CO}_2 \text{ formation in sterile in flask after 28 day (mg)}}{\text{ThCO}_2 \text{ (mg)}} \times 100$$

**PART V. MANOMETRIC RESPIROMETRY TEST**  
(Method C.4-D)**V.1. PRINCIPLE OF THE METHOD**

A measured volume of inoculated mineral medium, containing a known concentration of test chemical (100 mg/litre of the test substance, to give at least 50-100 mg ThOD/litre) as the nominal sole source of organic carbon, is stirred in a closed flask at a constant temperature ( $\pm 1^\circ \text{C}$  or closer) for up to 28 days. The consumption of oxygen is determined either by measuring the quantity of oxygen (produced electrolytically) required to maintain constant gas volume in the respirometer flask, or from the change in volume or pressure (or a combination of the two) in the apparatus. Evolved carbon dioxide is absorbed in a solution of potassium hydroxide or another suitable absorbent. The amount of oxygen taken up by the test chemical (corrected for uptake by blank inoculum, run in parallel) is expressed as a percentage of ThOD or COD. Optionally, primary biodegradation may also be calculated from supplemental specific analysis made at the beginning and end of incubation, and ultimate biodegradation by DOC analysis.

**V.2. DESCRIPTION OF THE METHOD****V.2.1. Apparatus**

- (a) suitable respirometer;
- (b) temperature control, maintaining  $\pm 1^\circ \text{C}$  or better;
- (c) membrane-filtration assembly (optional);
- (d) carbon analyser (optional).

**V.2.2. Preparation of mineral medium**

For the preparation of the stock solutions, see I.6.2.

Mix 10 ml of solution (a) with 800 ml dilution water, add 1 ml of solutions (b) to (d) and make up to 1 litre with dilution water.

**V.2.3. Preparation and pre-conditioning of inoculum**

The inoculum may be derived from a variety of sources: activated sludge; sewage effluents; surface waters and soils or from a mixture of these.

See I.6.4., I.6.4.1., I.6.4.2. and I.6.5.

**V.2.4. Preparation of flasks**

Prepare solutions of the test and reference chemicals, in separate batches, in mineral medium equivalent to a concentration, normally, of 100 mg chemical/litre (giving at least 50-100 mg ThOD/litre), using stock solutions.

**▼ B**

Calculate the ThOD on the basis of formation of ammonium salts unless nitrification is anticipated, when the calculation should be based on nitrate formation (see Appendix 2.2)

Determine the pH values and if necessary adjust to  $7,4 \pm 0,2$ .

Poorly soluble substances should be added at a later stage (see below).

If the toxicity of the test chemical is to be determined, prepare a further solution in mineral medium containing both test and reference chemicals at the same concentrations as in the individual solutions.

If measurement of the physico-chemical uptake of oxygen is required, prepare a solution of the test chemical at, normally, 100 mg ThOD/litre which has been sterilised by the addition of a suitable toxic substance (see I.6.6).

Introduce the requisite volume of solutions of test and reference chemicals, respectively, into at least duplicate flasks. Add to further flasks mineral medium only (for inoculum controls) and, if required, the mixed test/reference chemical solution and the sterile solution.

If the test chemical is poorly soluble, add it directly at this stage on a weight or volume basis or handle it as described in Appendix 3. Add potassium hydroxide, soda lime pellets or other absorbent to the CO<sub>2</sub>-absorber compartments.

**V.2.5. Number of flasks in a typical run**

Flasks 1 and 2: test suspension

Flasks 3 and 4: inoculum blank

Flask 5: procedure control

preferably, and when necessary:

Flask 6: sterile control

Flask 7: toxicity control

See also I.6.7.

**V.2.6. Performance of the test**

Allow the vessels to reach the desired temperature and inoculate appropriate vessels with prepared activated sludge or other source of inoculum to give a concentration of suspended solids not greater than 30 mg/litre. Assemble the equipment, start the stirrer and check for air-tightness, and start the measurement of oxygen uptake. Usually no further attention is required other than taking the necessary readings and making daily checks to see that the correct temperature and adequate stirring are maintained.

Calculate the oxygen uptake from the readings taken at regular and frequent intervals, using the methods given by the manufacturer of the equipment. At the end of incubation, normally 28 days, measure the pH of the contents of the flasks, especially if oxygen uptakes are low or greater than ThODNH<sub>4</sub> (for nitrogen-containing compounds).



**▼B**

If required, withdraw samples from the respirometer flasks, initially and finally, for analysis of DOC or specific chemical (see Appendix 2.4). At the initial withdrawal, ensure that the volume of test suspension remaining in the flask is known. When oxygen is taken up by N-containing test substance, determine the increase in concentration of nitrite and nitrate over 28 days and calculate the correction for the oxygen consumed by nitrification (Appendix 5).

## V.3. DATA AND REPORTING

V.3.1. **Treatment of results**

Divide the oxygen uptake (mg) of the test chemical after a given time (corrected for that by the blank inoculum control after the same time) by the weight of the test chemical used. This yields the BOD expressed as mg oxygen/mg test chemical, that is:

$$\text{BOD} = \frac{(\text{mg O}_2 \text{ uptake by test chemical} - \text{mg O}_2 \text{ uptake by blank})}{(\text{mg test chemical in flask})}$$

$$= \text{mg O}_2 \text{ per mg test chemical}$$

calculate the percentage biodegradation either from:

$$\% \text{ biodegradation} = \% \text{ ThOD} = \frac{\text{BOD (mg O}_2\text{/mg chemical)}}{\text{ThOD (mg O}_2\text{ chemical)}} \times 100$$

or form

$$\% \text{ COD} = \frac{\text{BOD (mg O}_2\text{/mg chemical)}}{\text{COD (mg O}_2\text{ chemical)}} \times 100$$

It should be noted that these two methods do not necessarily give the same value; it is preferable to use the former method.

For test substances containing nitrogen, use the appropriate ThOD (NH<sub>4</sub> or NO<sub>3</sub>) according to what is known or expected about the occurrence of nitrification (Appendix 2.2). If nitrification occurs but is not complete, calculate a correction for the oxygen consumed by nitrification from the changes in concentration of nitrite and nitrate (Appendix 5).

When optional determinations of organic carbon and/or specific chemical are made, calculate the percentage degradation, as described under I.7.

Record all results on the data sheets attached.

V.3.2. **Validity of results**

The oxygen uptake of the inoculum blank is normally 20-30 mg O<sub>2</sub>/litre and should not be greater than 60 mg/litre in 28 days. Values higher than 60 mg/litre require critical examination of the data and experimental techniques. If the pH value is outside the range 6-8,5 and the oxygen consumption by the test chemical is less than 60 %, the test should be repeated with a lower concentration of test chemical.

See also I.5.2.



**▼ B**

		Time (Days)											
		0		7		14			21			28	
% degradation $\frac{\text{BOD}}{\text{ThOD}} \times 100$	D <sub>1</sub> (a <sub>1</sub> )												
	D <sub>2</sub> (a <sub>2</sub> )												
	Mean (*)												

V = volume of medium in test flask

(\*) D<sub>1</sub> and D<sub>2</sub> should not be averaged if there is a considerable difference.

*N.B.:* Similar formats may be used for the reference chemical and the toxicity controls.

**6. CORRECTION FOR NITRIFICATION** (see Annex V)

Day	0	28	Difference
(i) Concentration of nitrate (mg N/litre)			(N)
(ii) Oxygen equivalent ( $4,57 \times N \times V$ ) (mg)	—	—	
(iii) Concentration of nitrite (mg N/litre)			(N)
(iv) Oxygen equivalent ( $3,43 \times N \times V$ ) (mg)	—	—	
(ii + iv) Total oxygen equivalent	—	—	

**7. CARBON ANALYSIS** (optional)

Carbon analyser:

Time (day)	Blank mg/litre	Test chemical mg/litre
0	(C <sub>blo</sub> )	(C <sub>o</sub> )
28 (*)	(C <sub>blt</sub> )	(C <sub>t</sub> )

(\*) Or at end of incubation

$$\% \text{ DOC removed} = \left( 1 - \frac{C_t - C_{blt}}{C_o - C_{blo}} \right) \times 100$$

**8. SPECIFIC CHEMICAL** (optional)

S<sub>b</sub> = concentration in physico-chemical (sterile) control at 28 days

S<sub>a</sub> = concentration in inoculated flask at 28 days,

$$\% \text{ biodegradation} = \frac{S_b - S_a}{S_b} \times 100$$

**9. ABIOTIC DEGRADATION** (optional)

a = oxygen consumption in sterile flasks after 28 days, (mg)

$$\text{oxygen consumption per mg test chemical} = \frac{a \times 100}{C_o V}$$

**▼B**

(see Sections 1 and 3)

$$\% \text{ abiotic degradation} = \frac{a \times 100}{C_0 V \times \text{ThOD}}$$

**PART VI. CLOSED BOTTLE TEST (Method C.4-E)****VI.1 PRINCIPLE OF THE TEST METHOD**

The solution of the test chemical in mineral medium, usually at 2-5 mg/litre, is inoculated with a relatively small number of micro-organisms from a mixed population and kept in completely full, closed bottles in the dark at constant temperature. Degradation is followed by analysis of dissolved oxygen over a 28-day period. The amount of oxygen taken up by the test chemical, corrected for uptake by the blank inoculum run in parallel, is expressed as a percentage of ThOD or COD.

**VI.2 DESCRIPTION OF THE METHOD****VI.2.1. Apparatus**

- a) BOD bottles, with glass stoppers, e.g. 250-300 ml;
- b) water bath or incubator, for keeping bottles at constant temperature ( $\pm 1$  °C or better) with the exclusion of light;
- c) large glass bottles (2-5 litres) for the preparation of media and for filling the BOD bottles;
- d) oxygen electrode and meter, or equipment and reagents for Winkler titration.

**VI.2.2. Preparation of mineral medium**

For the preparation of the stock solutions, see I.6.2.

Mix 1 (one) ml of solution (a) to (d) and make up to 1 litre with dilution water.

**VI.2.3. Preparation of the inoculum**

The inoculum is normally derived from the secondary effluent of a treatment plant or laboratory-scale unit receiving predominantly domestic sewage. An alternative source for the inoculum is surface water. Normally use from one drop (0,05 ml) to 5 ml of filtrate per litre of medium; trials may be needed to discover the optimum volume for a given effluent (See I.6.4.2 and I.6.5).

**VI.2.4. Preparation of flasks**

Strongly aerate mineral medium for at least 20 min. Carry out each test series with mineral medium derived from the same batch. Generally, the medium is ready for use after standing for 20 h, at the test temperature. Determine the concentration of dissolved oxygen for control purposes; the value should be about 9 mg/litre at 20 °C. Conduct all transfer and filling operations of the air-saturated medium bubble-free, for example, by the use of siphons.

**▼B**

Prepare parallel groups of BOD bottles for the determination of the test and reference chemicals in simultaneous experimental series. Assemble a sufficient number of BOD bottles, including inoculum blanks, to allow at least duplicate measurements of oxygen consumption to be made at the desired test intervals, for example, after 0, 7, 14, 21 and 28 days. To ensure being able to identify the 10-day window, more bottles may be required.

Add fully aerated mineral medium to large bottles so that they are about one-third full. Then add sufficient of the stock solutions of the test chemical and reference chemical to separate large bottles so that the final concentration of the chemicals is normally not greater than 10 mg/litre. Add no chemicals to the blank control medium contained in a further large bottle.

In order to ensure that the inoculum activity is not limited, the concentration of dissolved oxygen must not fall below 0,5 mg/litre in the BOD bottles. This limits the concentration of test chemical to about 2 mg/litre. However, for poorly degradable compounds and those with a low ThOD, 5-10 mg/litre can be used. In some cases, it would be advisable to run parallel series of test chemical at two different concentrations, for example, 2 and 5 mg/litre. Normally, calculate the ThOD on the basis of formation of ammonium salts but, if nitrification is expected or known to occur, calculate on the basis of the formation of nitrate (ThOD<sub>NO<sub>3</sub></sub>: see Appendix 2.2). However, if nitrification is not complete but does occur, correct for the changes in concentration of nitrite and nitrate, determined by analysis, (see Appendix 5).

If the toxicity of the test chemical is to be investigated (in the case, for example, of a previous low biodegradability value having been found), another series of bottles is necessary.

Prepare another large bottle to contain aerated mineral medium (to about one-third of its volume) plus test chemical and reference chemical at final concentrations normally the same as those in the other large bottles.

Inoculate the solutions in the large bottles with secondary effluent (one drop or about 0,05 ml, to 5 ml/litre) or with another source such as river water (see I.6.4.2.). Finally, make up the solutions to volume with aerated mineral medium using a hose which reaches down to the bottom of the bottle to achieve adequate mixing.

**VI.2.5. Number of flasks in a typical run**

In a typical run the following bottles are used:

- at least 10 containing test chemical and inoculum (test suspension),
- at least 10 containing only inoculum (inoculum blank),
- at least 10 containing reference chemical and inoculum (procedure control),

**▼ B**

- and, when necessary, six bottles containing test chemical, reference chemical and inoculum (toxicity control). However, to ensure being able to identify the 10-day window, about twice as many bottles would be necessary.

#### VI.2.6. Performance of the test

Dispense each prepared solution immediately into the respective group of BOD bottles by hose from the lower quarter (not the bottom) of the appropriate large bottle, so that all the BOD bottles are completely filled. Tap gently to remove any air bubbles. Analyse the zero-time bottles immediately for dissolved oxygen by the Winkler or electrode methods. The contents of the bottles can be preserved for later analysis by the Winkler method by adding manganese (II) sulfate and sodium hydroxide (the first Winkler reagent). Store the carefully stoppered bottles, containing the oxygen fixed as brown manganese (III) hydrated oxide, in the dark at 10-20 °C for no longer than 24 hours before proceeding with the remaining steps of the Winkler method. Stopper the remaining replicate bottles ensuring that no air bubbles are enclosed, and incubate at 20 °C in the dark. Each series must be accompanied by a complete parallel series for the determination of the inoculated blank medium. Withdraw at least duplicate bottles of all series for dissolved oxygen analysis at time intervals (at least weekly) over the 28 days incubation.

Weekly samples should allow the assessment of percentage removal in a 14-day window, whereas sampling every 3-4 days should allow the 10-day window to be identified, which would require about twice as many bottles.

For N-containing test substances, corrections for uptake of oxygen by any nitrification occurring should be made. To do this, use the O<sub>2</sub>-electrode method for determining the concentration of dissolved oxygen and then withdraw a sample from the BOD bottle for analysis for nitrite and nitrate. From the increase in concentration of nitrite and nitrate, calculate the oxygen used (see Annex V).

### VI.3. DATA AND REPORTING

#### VI.3.1. Treatment of results

First calculate the BOD exerted after each time period by subtracting the oxygen depletion (mg O<sub>2</sub>/litre) of the inoculum blank from that exhibited by the test chemical. Divide this corrected depletion by the concentration (mg/litre) of the test chemical, to obtain the specific BOD as mg oxygen per mg test chemical. Calculate the percentage biodegradability by dividing the specific BOD by the specific ThOD (calculated according to Appendix 2.2) or COD (determined by analysis, see Appendix 2.3), thus:

$$\text{BOD} = \frac{(\text{mg O}_2 \text{ uptake by test chemical} - \text{mg O}_2 \text{ uptake by blank})}{(\text{mg test chemical in flask})}$$

**▼ B**

= mg O<sub>2</sub> per mg test chemical

$$\% \text{ degradation} = \frac{\text{BOD (mg O}_2\text{/mg test chemical)}}{\text{ThOD (mg O}_2\text{/mg test chemical)}} \times 100$$

Or

$$\% \text{ degradation} = \frac{\text{BOD (mg O}_2\text{/mg test chemical)}}{\text{COD (mg O}_2\text{/mg test chemical)}} \times 100$$

It should be noted that these two methods do not necessarily give same value; it is preferable to use the former method.

For test substances containing nitrogen, use the appropriate ThOD (NH<sub>4</sub> or NO<sub>3</sub>) according to what is known or expected about the occurrence of nitrification (Appendix 2.2). If nitrification occurs but is not complete, calculate a correction for the oxygen consumed by nitrification from the changes in concentration of nitrite and nitrate (Appendix 5).

#### VI.3.2. **Validity of results**

Oxygen depletion in the inoculum blank should not exceed 1,5 mg dissolved oxygen/litre after 28 days. Values higher than this require investigation of the experimental techniques. The residual concentration of oxygen in the test bottles should not fall below 0,5 mg/litre at any time. Such low oxygen levels are valid only if the method of determining dissolved oxygen used is capable of measuring such levels accurately.

See also I.5.2.

#### VI.3.3. **Reporting**

See I.8.

#### VI.4. DATA SHEET

An example of a data sheet is given hereafter.

##### CLOSED BOTTLE TEST

1. **LABORATORY**
2. **DATE AT START OF TEST**
3. **TEST SUBSTANCE**

Name:

Stock solution concentration: ... mg/litre

Initial concentration in bottle: ... mg/litre

ThOD or COD: ... mg O<sub>2</sub>/mg test substance

4. **INOCULUM**

Source:

Treatment given:

**▼ B**

Pre-conditioning if any:

Concentration in the reaction mixture: ... mg/litre

**5. DO DETERMINATION**

Method: Winkler/electrode

**Flask analyses**

Time of incubation (d)			DO (mg/l)			
			0	n <sub>1</sub>	n <sub>2</sub>	
Blank (without chemical)	1	C <sub>1</sub>				
	2	C <sub>2</sub>				
Mean	$m_b = \frac{C_1 + C_2}{2}$					
Test chemical	1	a <sub>1</sub>				
	2	a <sub>2</sub>				
Mean	$m_t = \frac{a_1 + a_2}{2}$					

*Note:* Similar format may be used for reference and toxicity control.**6. CORRECTION FOR NITRIFICATION (see Annex V)**

Time of incubation (d)		0	n <sub>1</sub>	n <sub>2</sub>	n <sub>3</sub>
(i)	Concentration of nitrate (mg N/litre)				
(ii)	Change in nitrate concentration (mg N/litre)	—			
(iii)	Oxygen equivalent (mg/litre)	—			
(iv)	Concentration of nitrite (mg N/litre)				
(v)	Change in nitrite concentration (mg N/litre)	—			
(vi)	Oxygen equivalent (mg/litre)	—			
(iii + vi)	Total oxygen equivalent (mg/litre)	—			

**7. DO DEPLETION: % DEGRADATION**

	Depletion after n days (mg/litre)			
	n <sub>1</sub>	n <sub>2</sub>	n <sub>3</sub>	
FLASK 1: (m <sub>t0</sub> - m <sub>tx</sub> ) - (m <sub>b0</sub> - m <sub>bx</sub> )				
FLASK 2: (m <sub>t0</sub> - m <sub>tx</sub> ) - (m <sub>b0</sub> - m <sub>bx</sub> )				



**▼ B**

	Depletion after n days (mg/litre)			
	n <sub>1</sub>	n <sub>2</sub>	n <sub>3</sub>	
FLASK 1: $\% D_1 = \frac{[(m_{t_0} - m_{t_x}) - (m_{b_0} - m_{b_x})] \times 100}{\text{conc. of test} \times \text{ThOD chemical}}$				
FLASK 2: $\% D_2 = \frac{[(m_{t_0} - m_{t_x}) - (m_{b_0} - m_{b_x})] \times 100}{\text{conc. of test} \times \text{ThOD chemical}}$				
$\% D \text{ mean } (*) = \frac{D_1 - D_2}{2}$				
(*) Do not take mean if there is considerable difference between duplicates.				

$m_{t_0}$  = value in the flask at time 0

$m_{t_x}$  = value in the flask at time x

$m_{b_0}$  = mean blank value at time 0

$m_{b_x}$  = mean blank value at time x

Apply also correction for nitrification from iii + vi in section 6.

## 8. BLANK DO DEPLETIONS

Oxygen consumption by blank: ( $m_{b_0} - m_{b_{28}}$ ) mg/litre. This consumption is important for the validity of the test. It should be less than 1,5 mg/litre.

## PART VII. M.I.T.I. TEST (Method C.4-F)

### VII.1. PRINCIPLE OF THE METHOD

The oxygen uptake by a stirred solution, or suspension, of the test chemical in a mineral medium, inoculated with specially grown, unadapted micro-organisms, is measured automatically over a period of 28 days in a darkened, enclosed respirometer at  $25 \pm 1$  °C. Evolved carbon dioxide is absorbed by soda lime. Biodegradability is expressed as the percentage oxygen uptake (corrected for blank uptake) of the theoretical uptake (ThOD). The percentage of primary biodegradability is also calculated from supplemental specific chemical analysis made at the beginning and end of incubation and, optionally, by DOC analysis.

### VII.2. DESCRIPTION OF THE METHOD

#### VII.2.1. Apparatus

- (a) Automatic electrolytic BOD meter or respirometer normally equipped with six bottles, 300 ml each and equipped with cups to contain CO<sub>2</sub> absorbent;

**▼B**

- (b) constant temperature room and/or water-bath at  $25\text{ °C} \pm 1\text{ °C}$  or better;
- (c) membrane-filtration assembly (optional);
- (d) carbon analyser (optional).

**VII.2.2. Preparation of mineral medium**

Prepare the following stock solutions, using analytical grade reagents and water (I.6.1.):

- |   |         |
|---|---------|
| (a) Monopotassium dihydrogen ortho phosphate, $\text{KH}_2\text{PO}_4$                                    | 8,50 g  |
| Dipotassium monohydrogen ortho phosphate, $\text{K}_2\text{HPO}_4$  | 21,75 g |
| Disodium monohydrogen ortho phosphate dodecahydrate $\text{Na}_2\text{HPO}_4 \cdot 12\text{ H}_2\text{O}$ | 44,60 g |
| Ammonium chloride, $\text{NH}_4\text{Cl}$   | 1,70 g  |
| Dissolve in water and make up to 1 litre  |         |
| The pH value of the solution should be 7,2  |         |
| (b) Magnesium sulphate heptahydrate, $\text{MgSO}_4 \cdot 7\text{ H}_2\text{O}$                           | 22,50 g |
| Dissolve in water and make up to 1 litre  |         |
| (c) Calcium chloride anhydrous, $\text{CaCl}_2$   | 27,50 g |
| Dissolve in water and make up to 1 litre  |         |
| (d) Iron (III) chloride hexahydrate, $\text{FeCl}_3 \cdot 6\text{ H}_2\text{O}$                           | 0,25 g  |
| Dissolve in water and make up to 1 litre  |         |

Take 3 ml of each solution (a), (b), (c) and (d) and make up to 1 litre.

**VII.2.3. Preparation of inoculum**

Collect fresh samples from no fewer than ten sites, mainly in areas where a variety of chemicals are used and discharged. From sites such as sewage treatment works, industrial waste-water treatment, rivers, lakes, seas, collect 11 samples of sludge, surface soil, water, etc. and mix thoroughly together. After removing floating matter and allowing to stand, adjust the supernatant to  $\text{pH } 7 \pm 1$  with sodium hydroxide or phosphoric acid.

Use an appropriate volume of the filtered supernatant to fill a fill-and-draw activated sludge vessel and aerate the liquid for about 23 1/2 h. 30 minutes after stopping aeration, discard about one third of the whole volume of supernatant and add an equal volume of a solution (pH 7) containing 0,1 % each of glucose, peptone and monopotassium ortho phosphate, to the settled material and recommence aeration. Repeat this procedure once per day. The sludge unit must be operated according to good practice: effluents should be clear, temperature should be kept at  $25 \pm 2^\circ\text{ C}$ , pH should be  $7 \pm 1$ , sludge should settle well, sufficient aeration to keep the mixture aerobic at all times, protozoa should be present and the activity of the sludge should be tested against a reference substance at least every three months. Do not use sludge as inoculum until after at least one month's operation, but not after more than four months. Thereafter, sample from at least 10 sites at irregular intervals, once every three months.

**▼B**

In order to maintain fresh and old sludge at the same activity, mix the filtered supernatant of an activated sludge in use with an equal volume of the filtered supernatant of a freshly collected ten-source mixture and culture the combined liquor as above. Take sludge for use as inoculum 18-24 h after the unit has been fed.

**VII.2.4. Preparation of flasks**

Prepare the following six flasks:

No 1: test chemical in dilution water at 100 mg/l

No 2, 3 and 4: test chemical in mineral medium at 100 mg/l

No 5: reference chemical (e.g. aniline) in mineral medium at 100 mg/l

No 6: mineral medium only

Add poorly soluble test chemicals directly on a weight or volume basis or handle as described in Appendix 3, except that neither solvents nor emulsifying agents should be used. Add the CO<sub>2</sub> absorbent to all flasks in the special cups provided. Adjust the pH in flasks No 2, 3 and 4 to 7,0.

**VII.2.5. Performance of the test**

Inoculate flasks No 2, 3 and 4 (test suspensions), No 5 (activity control) and No 6 (inoculum blank) with a small volume of the inoculum to give a concentration of 30 mg/l suspended solids. No inoculum is added to flask No 1 which serves as an abiotic control. Assemble the equipment, check for air-tightness, start the stirrers, and start the measurement of oxygen uptake under conditions of darkness. Daily check the temperature, stirrer and coulometric oxygen uptake recorder, and note any changes in colour of the contents of the flasks. Read the oxygen uptakes for the six flasks directly by an appropriate method, for example, from the six-point chart recorder, which produces a BOD curve. At the end of incubation, normally 28 days, measure the pH of the contents of the flasks and determine the concentration of the residual test chemical and any intermediate and, in the case of water soluble substance, the concentration of DOC (Appendix 2.4). Take special care in the case of volatile chemicals. If nitrification is anticipated, determine nitrate and nitrite concentration, if possible.

**VII.3. DATA AND REPORTING****VII.3.1. Treatment of results**

Divide the oxygen uptake (mg) by the test chemical after a given time, corrected for that taken up by the blank inoculum control after the same time, by the weight of the test chemical used. This yields the BOD expressed as mg oxygen/mg test chemical, that is:

$$\text{BOD} = \frac{(\text{mg O}_2 \text{ uptake by test chemical} - \text{mg O}_2 \text{ uptake by blank})}{(\text{mg test chemical in flask})}$$

= mg O<sub>2</sub> per mg test chemical

**▼ B**

The percentage biodegradation is then obtained from:

$$\% \text{ biodegradation} = \% \text{ ThOD} = \frac{\text{BOD (mg O}_2\text{/mg chemical)}}{\text{ThOD (mg O}_2\text{/mg chemical)}} \times 100$$

For mixtures, calculate the ThOD from the elemental analysis, as for simple compound. Use the appropriate ThOD (ThOD<sub>NH4</sub> or ThOD<sub>NO3</sub>) according to whether nitrification is absent or complete (Appendix 2.2). If however, nitrification occurs but is incomplete, make a correction for the oxygen consumed by nitrification calculated from the changes in concentrations of nitrite and nitrate (Appendix 5).

Calculate the percentage primary biodegradation from loss of specific (parent) chemical (see 1.7.2).

$$D_t = \frac{S_b - S_a}{S_b} \times 100 \%$$

If there has been a loss of test chemical in the flask No 1 measuring physico-chemical removal, report this and use the concentration of test chemical ( $S_b$ ) after 28 days in this flask to calculate the percentage biodegradation.

When determinations of DOC are made (optional), calculate the percentage ultimate biodegradation from:

$$D_t = \left( 1 - \frac{C_t - C_{bt}}{C_o - C_{bo}} \right) \times 100 \%$$

as described under point 1.7.1. If there has been a loss of DOC in the flask No 1, measuring physico-chemical removal, use the DOC concentration in this flask to calculate the percentage biodegradation.

Record all results on the data sheets attached.

### VII.3.2. **Validity of results**

The oxygen uptake of the inoculum blank is normally 20-30 mg O<sub>2</sub>/l and should not be greater than 60 mg/l in 28 days. Values higher than 60 mg/l require critical examination of the data and experimental techniques. If the pH value is outside the range 6-8,5 and the oxygen consumption by the test chemical is less than 60 %, the test should be repeated with a lower concentration of test chemical.

See also I.5.2.

If the percentage degradation of aniline calculated from the oxygen consumption does not exceed 40 % after seven days and 65 % after 14 days, the test is regarded as invalid.

### VII.3.3. **Reporting**

See I.8.

### VII.4. **DATA SHEET**

An example of a data sheet is given below.

MITI (I) TEST

1. **LABORATORY**
2. **DATE AT START OF TEST**

**▼ B****3. TEST SUBSTANCE**

Name:

Stock solution concentration: mg/l as chemical

Initial concentration in medium,  $C_0$ : ... mg/l as chemical

Volume of reaction mixture, V: ... ml

ThOD: ... mg O<sub>2</sub>/l**4. INOCULUM**

Sludge sampling sites:

- |         |         |
|---------|---------|
| (1) ... | (6) ... |
| (2) ... | (7) ... |
| (3) ... | (8) ... |
| (4) ... | (9) ... |
| (5) ... | (10)... |

Concentration of suspended solids in activated sludge after acclimatisation with synthetic sewage = ... mg/l

Volume of activated sludge per litre of final medium = ... ml

Concentration of sludge in final medium = ... mg/l

**5. OXYGEN UPTAKE: BIODEGRADABILITY**

Type of respirometer used:

		Time (Days)				
		0	7	14	21	28
O <sub>2</sub> upt. (mg) test chemical	a <sub>1</sub>					
	a <sub>2</sub>					
	a <sub>3</sub>					
O <sub>2</sub> upt. (mg) blank	b					
Corrected BOD (mg)	(a <sub>1</sub> - b <sub>1</sub> ) (a <sub>1</sub> - b <sub>1</sub> ) (a <sub>1</sub> - b <sub>1</sub> )					
BOD per mg test chemical	$\frac{(a - b)}{C_0 V}$	Flask 1				
		Flask 2				
		Flask 3				
% degradation $\frac{BOD}{ThOD} \times 100$		1				
		2				
		3				
		Mean (*)				

(\*) Do not take a mean if there are considerable differences between replicates.

**▼B**

*N.B.:* Similar formats may be used for the reference compound.

### 6. CARBON ANALYSIS (optional)

Carbon analyser:

Flask	DOC			% DOC removed	Mean
	Measured	Corrected			
Water + test substance	a			—	—
Sludge + test substance	b <sub>1</sub>		b <sub>1</sub> - c		
Sludge + test substance	b <sub>2</sub>		b <sub>2</sub> - c		
Sludge + test substance	b <sub>3</sub>		b <sub>3</sub> - c		
Control blank	c		—	—	—

$$\% \text{ DOC removed} : \frac{a_1 - (b - c)}{a} \times 100$$

### 7. SPECIFIC CHEMICAL ANALYTICAL DATA

	Residual amount of test chemical at end of test	% degradation
blank test with water	S <sub>b</sub>	
inoculated medium	S <sub>a1</sub>	
	S <sub>a2</sub>	
	S <sub>a3</sub>	

$$\% \text{ degradation} = \frac{S_b - S_a}{S_b} \times 100$$

Calculate % degradation for flasks a<sub>1</sub>, and a<sub>3</sub> respectively

### 8. REMARKS

BOD curve against time, if available, should be attached.

*Appendix 1***ABBREVIATIONS AND DEFINITIONS**

DO: Dissolved oxygen (mg/l) is the concentration of oxygen dissolved in an aqueous sample.

BOD: Biochemical oxygen demand (g) is the amount of oxygen consumed by micro-organisms when metabolising a test compound; also expressed as g oxygen uptake per g test compound. (See method C.5).

COD: Chemical oxygen demand (g) is the amount of oxygen consumed during oxidation of a test compound with hot, acidic dichromate; it provides a measure of the amount of oxidisable matter present; also expressed as g oxygen consumed per g test compound. (See method C.6).

DOC: Dissolved organic carbon is the organic carbon present in solution or that which passes through a 0,45 micrometre filter or remains in the supernatant after centrifuging at 40 000 m.s<sup>-2</sup> ( $\pm$  4 000 g) for 15 min.

ThOD: Theoretical oxygen demand (mg) is the total amount of oxygen required to oxidise a chemical completely; it is calculated from the molecular formula (see Appendix 2.2) and is also expressed as mg oxygen required per mg test compound.

ThCO<sub>2</sub>: Theoretical carbon dioxide (mg) is the quantity of carbon dioxide calculated to be produced from the known or measured carbon content of the test compound when fully mineralised; also expressed as mg carbon dioxide evolved per mg test compound.

TOC: Total organic carbon of a sample is the sum of the organic carbon in solution and in suspension.

IC: Inorganic carbon

TC: Total carbon, is the sum of the organic and inorganic carbon present in a sample.

*Primary biodegradation:*

is the alteration in the chemical structure of a substance, brought about by biological action, resulting in the loss of specific property of that substance.

*Ultimate biodegradation (aerobic):*

is the level of degradation achieved when the test compound is totally utilised by micro-organisms resulting in the production of carbon dioxide, water, mineral salts and new microbial cellular constituents (biomass).

*Readily biodegradable:*

an arbitrary classification of chemicals which have passed certain specified screening tests for ultimate biodegradability; these tests are so stringent that it is assumed that such compounds will rapidly and completely biodegrade in aquatic environments under aerobic conditions.

**▼B***Inherently biodegradable:*

a classification of chemicals for which there is unequivocal evidence of biodegradation (primary or ultimate) in any recognized test of biodegradability.

*Treatability:*

is the amenability of compounds to removal during biological wastewater treatment without adversely affecting the normal operation of the treatment processes. Generally, readily biodegradable compounds are treatable but not all inherently biodegradable compounds are. Abiotic processes may also operate.

*Lag time*

is the time from inoculation, in a die-away test, until the degradation percentage has increased to at least 10 %. The lag time is often highly variable and poorly reproducible.

*Degradation time*

is the time from the end of the lag time till the time that 90 % of maximum level of degradation has been reached.

*10-day window*

is the 10 days immediately following the attainment of 10 % degradation.



**▼ B***Appendix 2***CALCULATION AND DETERMINATION OF SUITABLE SUMMARY PARAMETERS**

Depending on the method chosen, certain summary parameters will be required. The following section describes the derivation of these values. The use of these parameters is described in the individual methods.

**1. Carbon content**

The carbon content is calculated from the known elemental composition or determined by elemental analysis of the test substance.

**2. Theoretical oxygen demand (ThOD)**

The theoretical oxygen demand (ThOD) may be calculated if the elemental composition is known or determined by elemental analysis. It is for the compound:



without nitrification,

$$\text{ThOD}_{\text{NH}_4} = \frac{16 [2c + 1/2 (h - cl - 3n) + 3s + 5/2p + 1/2na - o]}{\text{MW}} \text{ mg/mg}$$

or with nitrification

$$\text{ThOD}_{\text{NO}_3} = \frac{16 [2c + 1/2 (h - cl) + 5/2n + 3s + 5/2p + 1/2na - o]}{\text{MW}} \text{ mg/mg}$$

**3. Chemical Oxygen Demand (COD)**

The chemical oxygen demand (COD) is determined according to method C.6.

**4. Dissolved organic carbon (DOC)**

Dissolved organic carbon (DOC) is by definition the organic carbon of any chemical or mixture in water passing through a 0,45 micrometre filter.

Samples from the test vessels are withdrawn and filtered immediately in the filtration apparatus using an appropriate membrane filter. The first 20 ml (amount can be reduced when using small filters) of the filtrate are discarded. Volumes of 10-20 ml or lower, if injected (volume depending on the amount required for carbon analyser) are retained for carbon analysis. The DOC concentration is determined by means of an organic carbon analyser which is capable of accurately measuring a carbon concentration equivalent or lower than 10 % of the initial DOC concentration used in the test.

Filtered samples which cannot be analysed on the same working day can be preserved by storage in a refrigerator at 2-4 °C for 48 h, or below - 18 °C for longer periods.

*Remarks:*

Membrane filters are often impregnated with surfactants for hydrophilisation. Thus the filter may contain up to several mg of soluble organic carbon which would interfere in the biodegradability determinations. Surfactants and other soluble organic compounds are removed from the filters by boiling them in deionised water for three periods each of one hour. The filters may then be stored in water for one week. If disposable filter cartridges are used each lot must be checked to confirm that it does not release soluble organic carbon.

**▼B**

Depending on the type of membrane filter the test chemical may be retained by adsorption. It may therefore be advisable to ensure that the test chemical is not retained by the filter.

Centrifugation at  $40\,000\text{ m}\cdot\text{sec}^{-2}$  (4 000 g) for 15 min may be used for differentiation of TOC versus DOC instead of filtration. The method is not reliable at initial concentration of  $< 10\text{ mg DOC/l}$  since either not all bacteria are removed or carbon as part of the bacterial plasma is redissolved.

*BIBLIOGRAPHY*

- Standard Methods for the Examination of Water and Wastewater, 12th ed, Am. Pub. Hlth. Ass., Am. Wat. Poll. Control Fed., Oxygen Demand, 1965, P 65.
- Wagner, R., Von Wasser, 1976, vol. 46, 139.
- DIN-Entwurf 38 409 Teil 41 Deutsche Einheitsverfahren zur Wasser-, Abwasser- und Schlammuntersuchung, Summarische Wirkungs- und Stoffkenngrößen (Gruppe H). Bestimmung des Chemischen Sauerstoffbedarfs (CSB) (H 41), Normenausschuß Wasserwesen (NAW) in DIN Deutsches Institut für Normung e. V.
- Gerike, P., The biodegradability testing of poorly water soluble compounds. Chemosphere, 1984, vol 13 (1), 169.

**▼B***Appendix 3***EVALUATION OF THE BIODEGRADABILITY OF POORLY SOLUBLE SUBSTANCES**

In biodegradability tests with poorly soluble substances the following aspects should receive special attention.

While homogeneous liquids will seldom present sampling problems, it is recommended that solid materials be homogenised by appropriate means to avoid errors due to non-homogeneity. Special care must be taken when representative samples of a few milligrams are required from mixtures of chemicals or substances with large amounts of impurities.

Various forms of agitation during the tests may be used. Care should be taken to use only sufficient agitation to keep the chemical dispersed, and to avoid over-heating, excessive foaming and excessive shear forces.

An emulsifier which gives a stable dispersion of the chemical may be used. It should not be toxic to bacteria and must not be biodegraded or cause foaming under test conditions.

The same criteria apply to solvents as to the emulsifiers.

It is not recommended that solid carriers be used for solid test substances but they may be suitable for only substances.

When auxiliary substances such as emulsifiers, solvents and carriers are used, a blank run containing the auxiliary substance should be performed.

Any of the three respirometric tests CO<sub>2</sub>, BOD, MITI can be used to study the biodegradability of poorly soluble compounds.

*BIBLIOGRAPHY*

- de Morsier, A. et al., Biodegradation tests for poorly soluble compounds. Chemosphere, 1987, vol. 16, 833.
- Gerike, P, The Biodegradability testing of poorly water soluble compounds. Chemosphere, 1984, vol. 13, 169.

**▼B***Appendix 4***EVALUATION OF THE BIODEGRADABILITY OF CHEMICALS  
SUSPECTED TO BE TOXIC TO THE INOCULUM**

When a chemical is subjected to ready biodegradability testing and appears to be non-biodegradable, the following procedure is recommended if a distinction between inhibition and inertness is desired (Reynolds et al., 1987).

Similar or identical inocula should be used for the toxicity and biodegradation tests.

To assess the toxicity of chemicals studied in ready biodegradability tests, the application of one or a combination of the inhibition of Sludge Respiration rate (activated sludge respiration inhibition test — Directive 87/302/EEC), BOD and/or Growth Inhibition methods would seem appropriate.

If inhibition due to toxicity is to be avoided, it is suggested that the test substance concentrations used in ready biodegradability testing should be less than 1/10 of the EC<sub>50</sub> values (or less than EC<sub>20</sub> values) obtained in toxicity testing. Compounds with an EC<sub>50</sub> value of greater than 300 mg/l are not likely to have toxic effects in ready biodegradability testing.

EC<sub>50</sub> values of less than 20 mg/l are likely to pose serious problems for the subsequent testing. Low test concentrations should be employed, necessitating the use of the stringent and sensitive Closed Bottle test or the use of <sup>14</sup>C-labelled material. Alternatively, an acclimatised inoculum may permit higher test substance concentrations to be used. In the latter case, however, the specific criterion of the ready biodegradability test is lost.

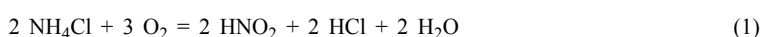
*BIBLIOGRAPHY*

Reynolds, L. et al., Evaluation of the toxicity of substances to be assessed for biodegradability. Chemosphere, 1987, vol. 16, 2259.

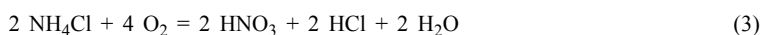
**▼B***Appendix 5***CORRECTION FOR OXYGEN UPTAKE FOR INTERFERENCE BY NITRIFICATION**

Errors due to not considering nitrification in the assessment by oxygen uptake of the biodegradability of test substances not containing N are marginal (not greater than 5 %), even if oxidation of the ammonium-N in the medium occurs erratically as between test and blank vessels. However, for test substances containing N, serious errors can arise.

If nitrification has occurred but is not complete the observed oxygen uptake by the reaction mixture may be corrected for the amount of oxygen used in oxidising ammonium to nitrite and nitrate, if the changes in concentration during incubation of nitrite and nitrate are determined by consideration of the following equations:



Overall:



From equation (1), the oxygen uptake by 28 g of nitrogen contained in ammonium chloride ( $\text{NH}_4\text{Cl}$ ) in being oxidised to nitrite is 96 g, i.e. a factor of 3,43 (96/28). In the same way, from equation (3) the oxygen uptake by 28 g of nitrogen in being oxidised to nitrate is 128 g, i.e. a factor of 4,57 (128/28).

Since the reactions are sequential, being carried out by distinct and different bacterial species, it is possible for the concentration of nitrite to increase or decrease; in the latter case an equivalent concentration of nitrate would be formed. Thus, the oxygen consumed in the formation of nitrate is 4,57 multiplied by the increase in concentration of nitrate, whereas the oxygen associated with the formation of nitrite is 3,43 multiplied by the increase in the concentration of nitrite or with the decrease in its concentration the oxygen loss is - 3,43 multiplied by the decrease in concentration.

That is:

$$\text{O}_2 \text{ consumed in nitrate formation} = 4,57 \times \text{increase in nitrate concentration} \quad (4)$$

and

$$\text{O}_2 \text{ consumed in nitrite formation} = 3,43 \times \text{increase in nitrite concentration} \quad (5)$$

and

$$\text{O}_2 \text{ lost in nitrite disappearance} = - 3,43 \times \text{decrease in nitrite concentration} \quad (6)$$

So that

$$\text{O}_2 \text{ uptake due to nitrification} = \pm 3,43 \times \text{change in nitrite conc.} + 4,57 \times \text{increase in nitrate conc.} \quad (7)$$

and thus

$$\text{O}_2 \text{ uptake due to C oxidation} = \text{total observed uptake} - \text{O}_2 \text{ uptake due to nitrification} \quad (8)$$

Alternatively, if only total oxidised N is determined, the oxygen uptake due to nitrification may be taken to be, as a first approximation,  $4,57 \times \text{increase in oxidised N}$

The corrected value for oxygen consumption due to C oxidation is then compared with ThOD  $\text{NH}_3$ , as calculated in Appendix 2.

**▼B****C.5 DEGRADATION — BIOCHEMICAL OXYGEN DEMAND****1. METHOD****1.1. INTRODUCTION**

The purpose of the method is the measurement of the biochemical oxygen demand (BOD) of solid or liquid organic substances.

Data elaborated with this test pertain to water-soluble compounds; however, volatile compounds and those of low water solubility may also, at least in principle, be tested.

The method is applicable only to those organic test materials which are not inhibitory to bacteria at the concentration used in the test. If the test material is not soluble at the test concentration, special measures, such as the use of ultrasonic dispersion, may have to be employed to achieve good dispersion of test material.

Information on the toxicity of the chemical may be useful to the interpretation of low results and in the selection of appropriate test concentrations.

**1.2. DEFINITION AND UNITS**

The BOD is defined as the mass of dissolved oxygen required by a specified volume of solution of the substance for the process of biochemical oxidation under prescribed conditions.

The results are expressed as grams of BOD per gram of tested substance.

**1.3. REFERENCE SUBSTANCES**

The use of a suitable reference substance to check the activity of the inoculum is desirable.

**1.4. PRINCIPLE OF THE TEST METHOD**

A predetermined amount of the substance, dissolved or dispersed in a well-aerated suitable medium, is inoculated with micro-organisms and incubated at a constant defined ambient temperature in the dark.

The BOD is determined by the difference in dissolved oxygen content at the beginning and at the end of the test. The duration of the test must be at least five days and not more than 28 days.

A blank must be determined in a parallel assay containing no test substance.

**1.5. QUALITY CRITERIA**

The BOD determination cannot be considered as a valid determination of the biodegradability of a substance. This test may only be regarded as a screening test.

**1.6. DESCRIPTION OF THE TEST METHOD**

A preliminary solution or dispersion of the substance is prepared to obtain a BOD concentration compatible with the method used. The BOD is then determined following any suitable national or international standardised method.

**▼B****2. DATA AND EVALUATION**

The BOD contained in the preliminary solution is calculated according to the selected normalised method, and converted into grams of BOD per gram of tested substance.

**3. REPORTING**

The method used shall be stated.

The biochemical oxygen demand should be a mean of at least three valid measurements.

All information and remarks relevant for the interpretation of results have to be reported, especially with regard to impurities, physical state, toxic effects and inherent composition of the substance which would affect the results.

The use of an additive to inhibit biological nitrification must be reported.

**4. REFERENCES**

List of standardised methods, for example:

NF T 90-103: Determination of the biochemical oxygen demand.

NBN 407: Biochemical oxygen demand.

NEN 32355.4: Bepaling van het biochemisch zuurstofverbruik (BZV).

The determination of biochemical oxygen demand, Methods for the examination of water and associated materials, HMSO, London.

ISO 5815: Determination of biochemical oxygen demand after n days.

**▼B****C.6. DEGRADATION — CHEMICAL OXYGEN DEMAND****1. METHOD****1.1. INTRODUCTION**

The purpose of the method is the measurement of the chemical oxygen demand (COD) of solid or liquid organic substances in a standard, arbitrary manner, under fixed laboratory conditions.

Information on the formula of the substance will be useful to conduct this test and interpret the result obtained (e.g. halogen salts, ferrous salts of organic compounds, organochlorine compounds).

**1.2. DEFINITIONS AND UNITS**

The chemical oxygen demand is a measure of the oxidisability of a substance, expressed as the equivalent amount in oxygen of an oxidising reagent consumed by the substance under fixed laboratory conditions.

The result is expressed in grams of COD per gram of tested substance.

**1.3. REFERENCE SUBSTANCES**

Reference substances do not need to be employed in all cases when investigating a new substance. They should serve primarily to calibrate the method from time to time and to allow comparison of results when another method is applied.

**1.4. PRINCIPLE OF THE TEST METHOD**

A predetermined amount of the substance, dissolved or dispersed in water, is oxidised by potassium dichromate in a strong sulphuric acid medium with silver sulphate as a catalyst, under reflux for two hours. The residual dichromate is determined by titration with standardised ferrous ammonium sulphate.

In case of chlorine-containing substances, mercuric sulphate <sup>(1)</sup> is added to reduce chloride interference.

**1.5. QUALITY CRITERIA**

Because of the arbitrary manner of determination, COD is an 'oxidisability indicator' and as such is used as a practical method to measure organic matter.

Chloride can interfere in this test; inorganic reducing or oxidising agents may also interfere with the COD determination.

Some cyclic compounds and many volatile substances (e.g. lower fatty acids) are not fully oxidised by this test.

**1.6. DESCRIPTION OF THE TEST METHOD**

A preliminary solution or dispersion of the substance is prepared to obtain a COD between 250 and 600 mg per litre.

<sup>(1)</sup> After use, solutions containing mercury salts should be treated to avoid dissemination of mercury in the environment.



**▼B***Remarks:*

In the case of poorly soluble and non-dispersible substances, an amount of finely powdered substance or liquid substance corresponding to about 5 mg of COD can be weighed and put in the experimental apparatus with water.

The chemical oxygen demand (COD) is often and especially in case of poorly soluble substances determined advantageously in a variant of the method, i.e., in a closed system with a pressure equaliser (H. Kelkenberg, 1975). In this modification compounds which are only with difficulty determined by the conventional method — e.g. acetic acid — may often be successfully quantified. The method also fails, however, in the case of pyridine. If the potassium dichromate concentration, as prescribed in ref.(1), is raised to 0,25 N (0,0416 M), the direct weighing-in of 5-10 mg of substance is facilitated which is essential for the COD determination of poorly water soluble substances (ref. (2)).

Otherwise, the COD is then determined following any suitable national or international standardised method.

**2. DATA AND EVALUATION**

The COD contained in the experimental flask is calculated following the selected normalised method, and converted to grams of COD per gram of tested substance.

**3. REPORTING**

The reference method used should be stated.

The chemical oxygen demand should be a mean of at least three measurements. All information and remarks relevant to the interpretation of the results have to be reported, especially with regard to impurities, physical state and inherent properties of the substance (if known) which would affect the results.

The use of mercuric sulphate to minimise the chloride interference must be reported.

**4. REFERENCES**

- (1) Kelkenberg, H., Z. von Wasser und Abwasserforschung, 1975, vol. 8, 146.
- (2) Gerike, P. The biodegradability testing of poorly water soluble compounds. Chemosphere, 1984, vol. 13, 169.

List of standardised methods, for example:

NBN T 91-201 Determination of the chemical oxygen demand.

ISBN O 11 7512494 Chemical oxygen demand (dichromate value) of polluted and waste waters.

NF T 90-101 Determination of the chemical oxygen demand.

DS 217 = water analysis Determination of the chemical oxygen demand.

DIN 38409-H-41 Determination of the chemical oxygen demand (COD) within the range above 15 mg per litre.

NEN 3235 5.3 Bepaling van het chemisch zuurstofverbruik.

ISO 6060 Water quality: chemical oxygen demand dichromate methods.

**▼B****C.7. DEGRADATION — ABIOTIC DEGRADATION: HYDROLYSIS  
AS A FUNCTION OF PH****1. METHOD**

This testing method is equivalent to the OECD TG 111 (2004).

**1.1. INTRODUCTION**

Chemicals can enter surface waters by such routes as direct application, spray drift, run-off, drainage, waste disposal, industrial, domestic or agricultural effluent and atmospheric deposition and may be transformed in those waters by chemical (e.g. hydrolysis, oxidation), photochemical and/or microbial processes. This Guideline describes a laboratory test method to assess abiotic hydrolytic transformations of chemicals in aquatic systems at pH values normally found in the environment (pH 4-9) and is based on existing Guidelines (1)(2)(3)(4)(5)(6)(7).

The experiments are performed to determine (i) the rate of hydrolysis of the test substance as a function of pH and (ii) the identity or nature and rates of formation and decline of hydrolysis products to which organisms may be exposed. Such studies may be required for chemicals which are directly applied to water or that are likely to reach the environment by the other routes described above.

**1.2. DEFINITIONS AND UNITS**

See Appendix 2

**1.3. APPLICABILITY OF THE METHOD**

The method is generally applicable to chemical substances (unlabelled or labelled) for which an analytical method with sufficient accuracy and sensitivity is available. It is applicable to slightly volatile and non-volatile compounds of sufficient solubility in water. The test should not be applied to chemicals that are highly volatile from water (e.g. fumigants, organic solvents) and thus cannot be kept in solution under the experimental conditions of this test. The test may be difficult to conduct with substances of minimal solubility in water (8).

**1.4. PRINCIPLE OF THE TEST METHOD**

Sterile aqueous buffer solutions of different pH values (pH 4, 7 and 9) are treated with the test substance and incubated in the dark under controlled laboratory conditions (at constant temperatures). After appropriate time intervals, buffer solutions are analysed for the test substance and for hydrolysis products. With labelled test substance (e.g.  $^{14}\text{C}$ ), a mass balance can be more easily established.

This testing method is designed as a tiered approach which is shown and explained in Appendix 1. Each tier is triggered by the results of the previous tier.

**▼B**

## 1.5. INFORMATION ON THE TEST SUBSTANCE

Non-labelled or labelled test substance can be used to measure the rate of hydrolysis. Labelled material is generally preferred for studying the pathway of hydrolysis and for establishing mass balance; however, in special cases, labelling may not be absolutely necessary. <sup>14</sup>C-labelling is recommended but the use of other isotopes, such as <sup>13</sup>C, <sup>15</sup>N, <sup>3</sup>H, may also be useful. As far as possible, the label should be positioned in the most stable part(s) of the molecule. For example, if the test substance contains one ring, labelling on this ring is required; if the test substance contains two or more rings, separate studies may be needed to evaluate the fate of each labelled ring and to obtain suitable information on formation of hydrolysis products. The purity of the test substance should be at least 95 %.

Before carrying out a hydrolysis test, the following information on the test substance should be available:

- (a) solubility in water [Testing Method A.6];
- (b) solubility in organic solvents;
- (c) vapour pressure [Testing Method A.4] and/or Henry's Law constant;
- (d) n-octanol/water partition coefficient [Testing Method A.8];
- (e) dissociation constant (pK<sub>a</sub>) [OECD Guideline 112] (9);
- (f) direct and indirect phototransformation rate in water where appropriate.

Analytical methods for quantification of the test substance and, if it is relevant, for identification and quantification of hydrolysis products in aqueous solutions should be available (see also Section 1.7.2).

## 1.6. REFERENCE SUBSTANCES

Where possible, reference substances should be used for the identification and quantification of hydrolysis products by spectroscopic and chromatographic methods or other suitably sensitive methods.

## 1.7. QUALITY CRITERIA

## 1.7.1. Recovery

Analysis of, at least, duplicate buffer solutions or of their extracts immediately after the addition of the test substance gives a first indication of the repeatability of the analytical method and of the uniformity of the application procedure for the test substance. Recoveries for later stages of the experiments are given by the respective mass balances (when labelled material is used). Recoveries should range from 90 % to 110 % for labelled and non labelled chemicals (7). In case it is technically difficult to reach this range, a recovery of 70 % for non labelled chemicals is acceptable, but justification should be given.

**▼B****1.7.2. Repeatability and sensitivity of analytical method**

Repeatability of the analytical method(s) used to quantify the test substance and hydrolysis products at later times can be checked by duplicate analysis of the same buffer solutions (or of their extracts) after sufficient quantities of hydrolysis products have formed for quantification

The analytical method should be sufficiently sensitive to quantify test substance concentrations down to 10 % or less of the initial concentration. If relevant, analytical methods should also be sufficiently sensitive to quantify any hydrolysis product representing 10 % or more of applied (at any time during the study) down to 25 % or less of its peak concentration.

**1.7.3. Confidence intervals for hydrolysis kinetic data**

Confidence intervals should be computed and presented for all regression coefficients, rate constants, half-lives, and any other kinetic parameters (e.g. DT50).

**1.8. DESCRIPTION OF THE TEST METHOD****1.8.1. Equipment and apparatus**

The study should be performed in glass containers (e.g. test tubes, small flasks) under dark and sterile conditions, if necessary, unless preliminary information (such as the n-octanol-water partition coefficient) indicates that the test substance may adhere to glass. In such cases, alternative materials (such as Teflon) may have to be considered. It may also be possible to alleviate the problem of adhere to glass by using one or more of the following methods:

- determine the mass of test substance and hydrolysis products sorbed to the test vessel,
- use of an ultrasonic bath,
- ensure a solvent wash of all glassware at each sampling interval,
- use of formulated products,
- use an increased amount of co-solvent for addition of test substance to the system; if a co-solvent is used it should be a co-solvent that does not hydrolyse the test substance.

Temperature-controlled water bath shakers or thermostatically controlled incubators for incubation of the various test solutions are normally required.

Standard laboratory equipment is required, including, in particular, the following:

- pH meter,

**▼B**

- analytical instruments such as GC, HPLC, TLC equipment, including the appropriate detection systems for analysing radio-labelled and non-labelled substances or inverse isotopes dilution method,
- instruments for identification purposes (e.g. MS, GC-MS, HPLC-MS, NMR, etc.),
- liquid scintillation counter,
- separating funnels for liquid-liquid extraction,
- instrumentation for concentrating solutions and extracts (e.g. rotating evaporator),
- temperature control devise (e.g. water bath).

Chemical reagents include, for example:

- organic solvents, analytical grade, such as hexane, dichloromethane, etc.,
- scintillation liquid,
- buffer solutions (for details see Section 1.8.3).

All glassware, reagent-grade water and buffer solutions used in the hydrolysis tests should be sterilised.

#### 1.8.2. **Test substance application**

The test substance should be applied as aqueous solution into the different buffer solutions (see Appendix 3). If it is necessary for adequate dissolution, the use of low amounts of water miscible solvents (such as acetonitrile, acetone, ethanol) is permitted for application and distribution of the test substance but this should not normally exceed 1 % v/v. In case a higher concentration of solvents is considered (e.g. in the case of poorly soluble test substances), this could only be allowed when it can be shown that the solvent has no effect on the hydrolysis of the test substances.

The use of formulated product is not routinely recommended, as it cannot be excluded that the formulation ingredients may influence the hydrolysis process. However, for poorly water-soluble test substances or for substances that adhere to glass (see Section 1.8.1), the use of formulated material may be an appropriate alternative.

One concentration of the test substance should be used; it should not exceed 0,01 M or half of the saturation concentration (see Appendix 1).

**▼ B****1.8.3. Buffer solutions**

The hydrolysis test should be performed at pH values of 4, 7 and 9. For this purpose, buffer solutions should be prepared using reagent grade chemicals and water. Some useful buffer systems are presented in Appendix 3. It should be noted that the buffer system used may influence the rate of hydrolysis and where this is observed an alternate buffer system should be employed <sup>(1)</sup>.

The pH of each buffer solution should be checked with a calibrated pH meter to a precision of at least 0,1 at the required temperature.

**1.8.4. Test conditions****1.8.4.1. Test temperature**

The hydrolysis experiments should be carried out at constant temperatures. For extrapolation purposes, it is important to maintain the temperature to at least  $\pm 0,5$  °C.

A preliminary test (Tier 1) should be conducted at a temperature of 50 °C if the hydrolytic behaviour of the test substance is unknown. Higher Tier kinetic tests should be carried out with a minimum of three temperatures (including the test at 50 °C) unless the test substance is stable to hydrolysis as determined by the Tier 1 testing. A suggested temperature range is 10-70 °C (preferably with at least one temperature below 25 °C utilised), which will encompass the reporting temperature of 25 °C and most of the temperatures encountered in the field.

**1.8.4.2. Light and oxygen**

All of the hydrolysis tests should be carried out using any suitable method to avoid photolytic effects. All suitable measures should be taken to avoid oxygen (e.g. by bubbling helium, nitrogen or argon for five minutes before preparation of the solution).

**1.8.4.3. Test duration**

The preliminary test should be carried out for 5 days whereas the higher Tier tests should be conducted until 90 % hydrolysis of the test substance or for 30 days whichever comes first.

**1.8.5. Performance of the test****1.8.5.1. Preliminary test (Tier 1)**

The preliminary test is performed at  $50 \pm 0,5$  °C and pH 4,0, 7,0 and 9,0. If less than 10 % of hydrolysis is observed after 5 days ( $t_{0,525^{\circ}\text{C}} > 1$  year), the test substance is considered hydrolytically stable and, normally, no additional testing is required. If the substance is known to be unstable at environmentally relevant temperatures <sup>(2)</sup>, the preliminary test is not required. The analytical method must be sufficiently precise and sensitive to detect a reduction of 10 % in the initial concentration.

<sup>(1)</sup> Mabey and Mill recommend the use of borate or acetate buffers instead of phosphate (11).

<sup>(2)</sup> Such information may come from other sources such as hydrolysis data of structurally similar compounds from the literature or from other preliminary, semi-quantitative hydrolysis tests with the test substance at an earlier development stage.

**▼B**1.8.5.2. *Hydrolysis of unstable substances (Tier 2)*

The higher Tier (advanced) test should be performed at the pH values at which the test substance was found unstable as defined by the preliminary test above. The buffered solutions of the test substance should be thermostated at the selected temperatures. To test for first-order behaviour, each reaction solution should be analysed in time intervals which provide a minimum of six spaced data points normally between 10 % and 90 % hydrolysis of the test substance. Individual replicate test samples (a minimum of duplicate samples contained in separate reaction vessels) should be removed and the contents analysed at each of at least six sampling times (for a minimum of twelve replicate data points). The use of a single bulk sample from which individual aliquots of the test solution are removed at each sampling interval is considered to be inadequate, as it does not allow for the analysis of data variability and it may lead to problems with contamination of the test solution. Sterility confirmation tests should be conducted at the end of the higher Tier test (i.e. at 90 % hydrolysis or 30 days). However, if no degradation (i.e. transformation) is observed, sterility tests are not considered necessary.

1.8.5.3. *Identification of hydrolysis products (Tier 3)*

Any major hydrolysis products at least those representing > 10 % of the applied dose should be identified by appropriate analytical methods.

1.8.5.4. *Optional tests*

Additional tests at pH values other than 4, 7 and 9 may be required for a hydrolytically unstable test substance. For example, for physiological purposes a test under more acidic conditions (e.g. pH 1,2) may be required employing a single physiologically relevant temperature (37 °C).

**2. DATA**

The amounts of test substance and of hydrolysis products, if relevant, should be given as % of applied initial concentration and, where appropriate, as mg/L for each sampling interval and for each pH and test temperature. In addition, a mass balance should be given in percentage of the applied initial concentration when labelled test substance has been used.

A graphical presentation of the log-transformed data of the test substance concentrations against time should be reported. Any major hydrolysis products at least those representing  $\geq 10$  % of the applied dose should be identified and their log-transformed concentrations should also be plotted in the same manner as the parent substance to show their rates of formation and decline.

**▼ B**

## 2.1. TREATMENT OF RESULTS

More accurate determinations of half-lives or  $DT_{50}$  values should be obtained by applying appropriate kinetic model calculations. The half-life and/or  $DT_{50}$  values (including confidence limits) should be reported for each pH and temperature together with a description of the model used the order of kinetics and the coefficient of determination ( $r^2$ ). If appropriate, the calculations should also be applied to the hydrolysis products.

In the case of rate studies carried out at different temperatures, the pseudo first-order hydrolysis rate constants ( $k_{obs}$ ) should be described as a function of temperature. The calculation should be based on both the separation of  $k_{obs}$  into rate constants for acid catalysed, neutral, and base catalysed hydrolysis ( $k_H$ ,  $k_{neutral}$ , and  $k_{OH}$  respectively) and the Arrhenius equation:

$$k_{obs} = k_H[H^+] + k_{neutral} + k_{OH}[OH^-] = \sum_{i=H,neutral,OH} A_i e^{-B_i/T}$$

where  $A_i$  and  $B_i$  are regression constants from the intercept and slope, respectively, of the best fit lines generated from linearly regressing  $\ln k_i$  against the reciprocal of the absolute temperature in Kelvin (T). Through the use of the Arrhenius relationships for acid, neutral and base catalysed hydrolysis, pseudo first-order rate constants, and thus half-lives can be calculated for other temperatures for which the direct experimental determination of a rate constant is not practicable (10).

## 2.2. EVALUATION AND INTERPRETATION OF RESULTS

Most hydrolysis reactions follow apparent first order reaction rates and, therefore, half-lives are independent of the concentration (see equation 4 in Appendix 2). This usually permits the application of laboratory results determined at  $10^{-2}$  to  $10^{-3}$  M to environmental conditions ( $\leq 10^{-6}$  M) (10). Several examples of good agreement between rates of hydrolysis measured in both pure and natural waters for a variety of chemicals were reported by Mabey and Mill (11), provided both pH and temperature had been measured.

## 3. REPORTING

## 3.1. TEST REPORT

The test report must include at least the following information:

Test substance:

— common name, chemical name, CAS number, structural formula (indicating position of label when radiolabelled material is used) and relevant physical-chemical properties (see Section 1.5);

— purity (impurities) of test substance,

— label purity of labelled chemical and molar activity (where appropriate).



**▼ B**

- Buffer solutions:
- dates and details of preparation,
- buffers and waters used,
- molarity and pH of buffer solutions.

## Test conditions:

- dates of the performance of the studies,
- amount of test substance applied,
- method and solvents (type and amount) used for application of the test substance,
- volume of buffered test substance solutions incubated,
- description of the incubation system used,
- pH and temperature during the study,
- sampling times,
- method(s) of extraction,
- methods for quantification and identification of the test substance and its hydrolysis products in the buffer solutions,
- number of replicates.

## Results:

- repeatability and sensitivity of the analytical methods used,
- recoveries (% values for a valid study are given in Section 1.7.1),
- replicate data and means in a tabular forms,
- mass balance during and at the end of the studies (when labelled test substance is used),
- results of preliminary test,
- discussion and interpretation of results,
- all original data and figures.

The following information is only required when hydrolysis rate is determined:

- plots of concentrations versus time for the test substances and, where appropriate, for the hydrolysis products at each pH value and temperature;
- tables of results of Arrhenius equation for the temperature 20 °C/-25 °C, with pH, rate constant [ $\text{h}^{-1}$  or  $\text{day}^{-1}$ ], half-life or  $\text{DT}_{50}$ , temperatures [°C] including confidence limits and the coefficients of correlation ( $r^2$ ) or comparable information;
- proposed pathway of hydrolysis.

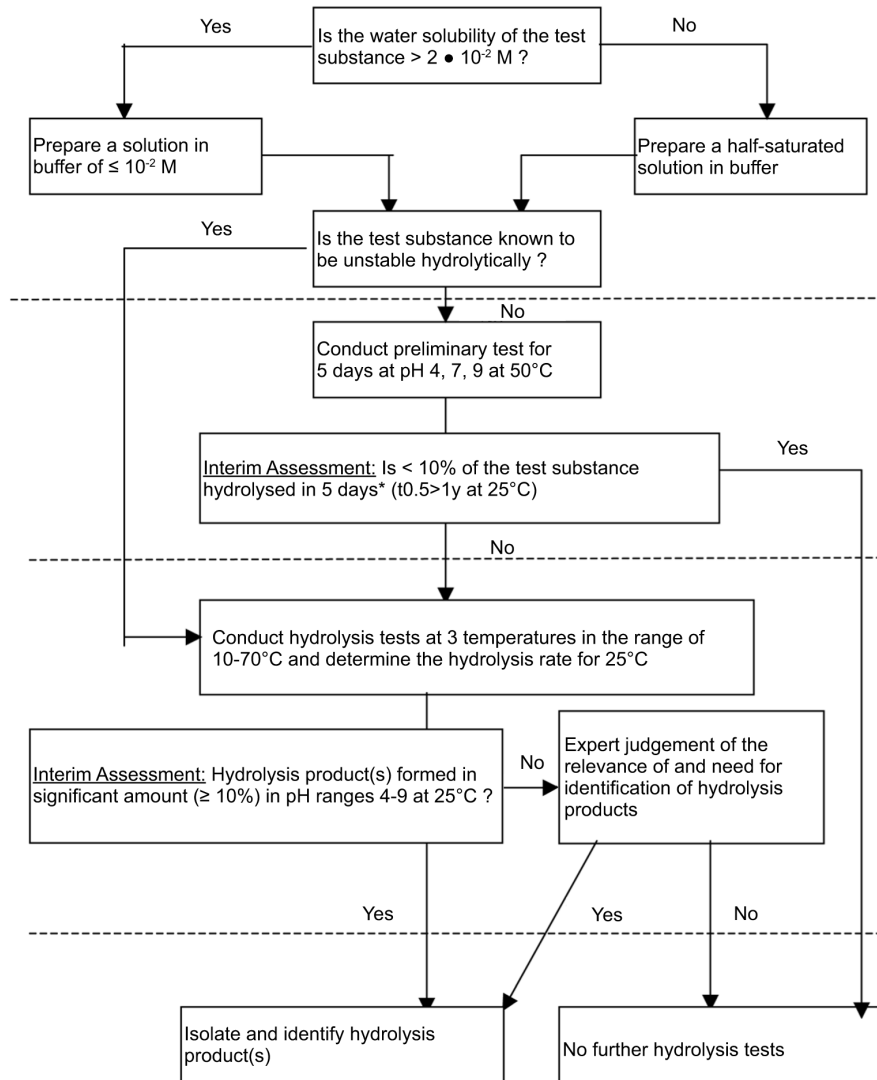
**▼B**4. **REFERENCES**

- (1) OECD, (1981) Hydrolysis as a Function of pH. OECD Guideline for Testing of Chemicals No 111, adopted 12 May 1981.
- (2) US-Environmental Protection Agency, (1982) 40 CFR 796.3500, Hydrolysis as a Function of pH at 25 °C. Pesticide Assessment Guidelines, Subdivision N. Chemistry: Environmental Fate.
- (3) Agriculture Canada, (1987) Environmental Chemistry and Fate Guidelines for registration of pesticides in Canada.
- (4) European Union (EU), (1995) Commission Directive 95/36/EC amending Council Directive 91/414/EEC concerning the placing of plant protection products on the market. Annex V: Fate and Behaviour in the Environment.
- (5) Dutch Commission for Registration of Pesticides, (1991) Application for registration of a pesticide. Section G: Behaviour of the product and its metabolites in soil, water and air.
- (6) BBA, (1980) Merkblatt No 55, Teil I und II: Prüfung des Verhaltens von Pflanzenbehandlungsmitteln im Wasser (October 1980).
- (7) SETAC, (1995) Procedures for Assessing the Environmental Fate and Ecotoxicity of Pesticides. Mark R. Lynch, Ed.
- (8) OECD, (2000) Guidance document on aquatic toxicity testing of difficult substances and mixtures, OECD Environmental Health and Safety Publications Series on Testing and Assessment No 23.
- (9) OECD, (1993) Guidelines for the Testing of Chemicals. Paris. OECD (1994-2000): Addenda 6-11 to Guidelines for the Testing of Chemicals.
- (10) Nelson, H, Laskowski D, Thermes S, and Hendley P., (1997) Recommended changes in pesticide fate study guidelines for improving input to computer models. (Text version of oral presentation at the 14th Annual Meeting of the Society of Environmental Toxicology and Chemistry, Dallas TX, November 1993).
- (11) Mabey, W. and Mill, T., (1978) Critical review of hydrolysis of organic compounds in water under environmental conditions. J. Phys. Chem. Ref. Data 7, p. 383-415.

▼ B

## Appendix 1

## Tiered hydrolysis test scheme



\* 10% hydrolysis of a test substance at 50°C corresponds to a half-life of approx. 30 days which corresponds to a value of approx. 1 year at 20°C.

**▼ B***Appendix 2***Definitions and units**

**Standard International (SI) units** should be used in any case.

**Test substance:** any substance, whether the parent compound or relevant transformation products.

**Transformation products:** all substances resulting from biotic or abiotic transformation reactions of the test substance.

**Hydrolysis products:** all substances resulting from hydrolytic transformation reactions of the test substance.

**Hydrolysis** refers to a reaction of a test substance RX with water, with the net exchange of the group X with OH at the reaction centre:



The rate at which the concentration of RX decreases in this simplified process is given by

$$\text{rate} = k [\text{H}_2\text{O}] [\text{RX}] \quad \text{second order reaction}$$

or

$$\text{rate} = k [\text{RX}] \quad \text{first order reaction}$$

depending on the rate determining step. Because the water is present in great excess compared to the test substance, this type of reaction is usually described as a pseudo-first order reaction in which the observed rate constant is given by the relationship

$$k_{\text{obs}} = k [\text{H}_2\text{O}] \quad [2]$$

and can be determined from the expression (\*)

$$k_{\text{obs}} = \frac{1}{t} \ln \frac{C_0}{C_t} \quad [3]$$

where:

t = time

and  $C_0$ ,  $C_t$  = concentrations of RX at times 0 and t.

The units of this constant have the dimensions of (time)<sup>-1</sup> and the half-life of the reaction (time for 50 % of RX to react) is given by

$$t_{0,5} = \frac{\ln 2}{k_{\text{obs}}} \quad [4]$$

**Half-life:** ( $t_{0,5}$ ) is the time taken for 50 % hydrolysis of a test substance when the reaction can be described by first order kinetics; it is independent of the concentration.

(\*) If the plot of the log-transformed data vs. time does not indicate a linear function (equated with a first-order reaction rate), then the use of equation [3] is not appropriate for determining the hydrolysis rate constant of the test compound.

**▼ B**

**DT<sub>50</sub> (Disappearance Time 50):** is the time within which the concentration of the test substance is reduced by 50 %; it is different from the half-life  $t_{0,5}$  when the reaction does not follow first order kinetics.

**Estimation of k at different temperature**

When the rate constants are known for two temperatures, the rate constants at other temperatures can be calculated using the Arrhenius equation:

$$k = A \times e^{-\frac{E}{R \times T}} \text{ or } \ln k = \frac{-E}{R \times T} + \ln A$$

A plot of  $\ln k$  versus  $1/T$  gives a straight line with a slope of  $-E/R$

where:

k = rate constant, measured at different temperatures

E = activation energy [kJ/mol]

T = absolute temperature [K]

R = gas constant [8,314 J/mol.K]

The activation energy was calculated by regression analysis or the following equation:

$$E = R \times \frac{\ln k_2 - \ln k_1}{\left(\frac{1}{T_1} - \frac{1}{T_2}\right)}$$

where:  $T_2 > T_1$ .

▼ **B***Appendix 3***Buffer Systems****A. CLARK AND LUBS:****Buffer mixtures of CLARK and LUBS (\*)**

Composition	pH
<b>0,2 N HCl and 0,2 N KCl at 20 °C</b>	
47,5 ml. HCl + 25 ml. KCl dil. to 100 ml	1,0
32,25 ml. HCl + 25 ml. KCl dil. to 100 ml	1,2
20,75 ml. HCl + 25 ml. KCl dil. to 100 ml	1,4
13,15 ml. HCl + 25 ml. KCl dil. to 100 ml	1,6
8,3 ml. HCl + 25 ml. KCl dil. to 100 ml	1,8
5,3 ml. HCl + 25 ml. KCl dil. to 100 ml	2,0
3,35 ml. HCl + 25 ml. KCl dil. to 100 ml	2,2
<b>0,1 M potassium biphthalate + 0,1 N HCl at 20 °C</b>	
46,70 ml. 0,1 N HCl + 50 ml. biphthalate to 100 ml	2,2
39,60 ml. 0,1 N HCl + 50 ml. biphthalate to 100 ml	2,4
32,95 ml. 0,1 N HCl + 50 ml. biphthalate to 100 ml	2,6
26,42 ml. 0,1 N HCl + 50 ml. biphthalate to 100 ml	2,8
20,32 ml. 0,1 N HCl + 50 ml. biphthalate to 100 ml	3,0
14,70 ml. 0,1 N HCl + 50 ml. biphthalate to 100 ml	3,2
9,90 ml. 0,1 N HCl + 50 ml. biphthalate to 100 ml	3,4
5,97 ml. 0,1 N HCl + 50 ml. biphthalate to 100 ml	3,6
2,63 ml. 0,1 N HCl + 50 ml. biphthalate to 100 ml	3,8
<b>0,1 M potassium biphthalate + 0,1 N NaOH at 20 °C</b>	
0,40 ml. 0,1 N NaOH + 50 ml. biphthalate to 100 ml	4,0
3,70 ml. 0,1 N NaOH + 50 ml. biphthalate to 100 ml	4,2
7,50 ml. 0,1 N NaOH + 50 ml. biphthalate to 100 ml	4,4
12,15 ml. 0,1 N NaOH + 50 ml. biphthalate to 100 ml	4,6
17,70 ml. 0,1 N NaOH + 50 ml. biphthalate to 100 ml	4,8

(\*) The pH values reported in these tables have been calculated from the potential measurements using Sørensen's standard equations (1909). The corresponding pH values are 0,04 units higher than the tabulated values.

**▼B**

Composition	pH
23,85 ml. 0,1 N NaOH + 50 ml. biphthalate to 100 ml	5,0
29,95 ml. 0,1 N NaOH + 50 ml. biphthalate to 100 ml	5,2
35,45 ml. 0,1 N NaOH + 50 ml. biphthalate to 100 ml	5,4
39,85 ml. 0,1 N NaOH + 50 ml. biphthalate to 100 ml	5,6
43,00 ml. 0,1 N NaOH + 50 ml. biphthalate to 100 ml	5,8
45,45 ml. 0,1 N NaOH + 50 ml. biphthalate to 100 ml	6,0

**Buffer mixtures of CLARK and LUBS (Continued)**

<b>0,1 M monopotassium phosphate + 0,1 N NaOH at 20 °C</b>	
5,70 ml. 0,1 N NaOH + 50 ml. phosphate to 100 ml	6,0
8,60 ml. 0,1 N NaOH + 50 ml. phosphate to 100 ml	6,2
12,60 ml. 0,1 N NaOH + 50 ml. phosphate to 100 ml	6,4
17,80 ml. 0,1 N NaOH + 50 ml. phosphate to 100 ml	6,6
23,45 ml. 0,1 N NaOH + 50 ml. phosphate to 100 ml	6,8
29,63 ml. 0,1 N NaOH + 50 ml. phosphate to 100 ml	7,0
35,00 ml. 0,1 N NaOH + 50 ml. phosphate to 100 ml	7,2
39,50 ml. 0,1 N NaOH + 50 ml. phosphate to 100 ml	7,4
42,80 ml. 0,1 N NaOH + 50 ml. phosphate to 100 ml	7,6
45,20 ml. 0,1 N NaOH + 50 ml. phosphate to 100 ml	7,8
46,80 ml. 0,1 N NaOH + 50 ml. phosphate to 100 ml	8,0
<b>0,1 M H<sub>3</sub>B<sub>3</sub> in 0,1 M KCl + 0,1 N NaOH at 20 °C</b>	
2,61 ml. 0,1 N NaOH + 50 ml. boric acid to 100 ml	7,8
3,97 ml. 0,1 N NaOH + 50 ml. boric acid to 100 ml	8,0
5,90 ml. 0,1 N NaOH + 50 ml. boric acid to 100 ml	8,2
8,50 ml. 0,1 N NaOH + 50 ml. boric acid to 100 ml	8,4
12,00 ml. 0,1 N NaOH + 50 ml. boric acid to 100 ml	8,6
16,30 ml. 0,1 N NaOH + 50 ml. boric acid to 100 ml	8,8
21,30 ml. 0,1 N NaOH + 50 ml. boric acid to 100 ml	9,0
26,70 ml. 0,1 N NaOH + 50 ml. boric acid to 100 ml	9,2

**▼B**

32,00 ml. 0,1 N NaOH + 50 ml. boric acid to 100 ml	9,4
36,85 ml. 0,1 N NaOH + 50 ml. boric acid to 100 ml	9,6
40,80 ml. 0,1 N NaOH + 50 ml. boric acid to 100 ml	9,8
43,90 ml. 0,1 N NaOH + 50 ml. boric acid to 100 ml	10,0

**B. KOLTHOFF AND VLEESCHHOUWER:****Citrate buffers of KOLTHOFF and VLEESCHHOUWER**

Composition	pH
<b>0,1 M monopotassium citrate and 0,1 N HCl at 18 °C (*)</b>	
49,7 ml. 0,1 N HCl + 50 ml. citrate to 100 ml	2,2
43,4 ml. 0,1 N HCl + 50 ml. citrate to 100 ml	2,4
36,8 ml. 0,1 N HCl + 50 ml. citrate to 100 ml	2,6
30,2 ml. 0,1 N HCl + 50 ml. citrate to 100 ml	2,8
23,6 ml. 0,1 N HCl + 50 ml. citrate to 100 ml	3,0
17,2 ml. 0,1 N HCl + 50 ml. citrate to 100 ml	3,2
10,7 ml. 0,1 N HCl + 50 ml. citrate to 100 ml	3,4
4,2 ml. 0,1 N HCl + 50 ml. citrate to 100 ml	3,6
<b>0,1 M monopotassium citrate and 0,1 N NaOH at 18 °C (*)</b>	
2,0 ml. 0,1 N NaOH + 50 ml. citrate to 100 ml	3,8
9,0 ml. 0,1 N NaOH + 50 ml. citrate to 100 ml	4,0
16,3 ml. 0,1 N NaOH + 50 ml. citrate to 100 ml	4,2
23,7 ml. 0,1 N NaOH + 50 ml. citrate to 100 ml	4,4
31,5 ml. 0,1 N NaOH + 50 ml. citrate to 100 ml	4,6
39,2 ml. 0,1 N NaOH + 50 ml. citrate to 100 ml	4,8
46,7 ml. 0,1 N NaOH + 50 ml. citrate to 100 ml	5,0
54,2 ml. 0,1 N NaOH + 50 ml. citrate to 100 ml	5,2
61,0 ml. 0,1 N NaOH + 50 ml. citrate to 100 ml	5,4
68,0 ml. 0,1 N NaOH + 50 ml. citrate to 100 ml	5,6
74,4 ml. 0,1 N NaOH + 50 ml. citrate to 100 ml	5,8
81,2 ml. 0,1 N NaOH + 50 ml. citrate to 100 ml	6,0

(\*) Add tiny crystal of thymol or a similar substance to prevent growth of molds.



## ▼B

## C. SÖRENSEN:

## Borate mixtures of SÖRENSEN

Composition		Sörensen 18 °C	Walbum, pH at		
ml. Borax	ml. HCl/NaOH		10 °C	40 °C	70 °C
<b>0,05 M borax + 0,1 N HCl</b>					
5,25	4,75	7,62	7,64	7,55	7,47
5,50	4,50	7,94	7,98	7,86	7,76
5,75	4,25	8,14	8,17	8,06	7,95
6,00	4,00	8,29	8,32	8,19	8,08
6,50	3,50	8,51	8,54	8,40	8,28
7,00	3,00	8,08	8,72	8,56	8,40
7,50	2,50	8,80	8,84	8,67	8,50
8,00	2,00	8,91	8,96	8,77	8,59
8,50	1,50	9,01	9,06	8,86	8,67
9,00	1,00	9,09	9,14	8,94	8,74
9,50	0,50	9,17	9,22	9,01	8,80
10,00	0,00	9,24	9,30	9,08	8,86
<b>0,05 M borax + 0,1 N NaOH</b>					
10,0	0,0	9,24	9,30	9,08	8,86
9,0	1,0	9,36	9,42	9,18	8,94
8,0	2,0	9,50	9,57	9,30	9,02
7,0	3,0	9,68	9,76	9,44	9,12
6,0	4,0	9,97	10,06	9,67	9,28

## Phosphate mixtures of SÖRENSEN

Composition	pH
<b>0,0667 M Monopotassium phosphate + 0,0667 M Disodium phosphate at 20 °C</b>	
99,2 ml. KH <sub>2</sub> PO <sub>4</sub> + 0,8 ml Na <sub>2</sub> HPO <sub>4</sub>	5,0
98,4 ml. KH <sub>2</sub> PO <sub>4</sub> + 1,6 ml Na <sub>2</sub> HPO <sub>4</sub>	5,2
97,3 ml. KH <sub>2</sub> PO <sub>4</sub> + 2,7 ml Na <sub>2</sub> HPO <sub>4</sub>	5,4
95,5 ml. KH <sub>2</sub> PO <sub>4</sub> + 4,5 ml Na <sub>2</sub> HPO <sub>4</sub>	5,6
92,8 ml. KH <sub>2</sub> PO <sub>4</sub> + 7,2 ml Na <sub>2</sub> HPO <sub>4</sub>	5,8
88,9 ml. KH <sub>2</sub> PO <sub>4</sub> + 11,1 ml Na <sub>2</sub> HPO <sub>4</sub>	6,0
83,0 ml. KH <sub>2</sub> PO <sub>4</sub> + 17,0 ml Na <sub>2</sub> HPO <sub>4</sub>	6,2
75,4 ml. KH <sub>2</sub> PO <sub>4</sub> + 24,6 ml Na <sub>2</sub> HPO <sub>4</sub>	6,4
65,3 ml. KH <sub>2</sub> PO <sub>4</sub> + 34,7 ml Na <sub>2</sub> HPO <sub>4</sub>	6,6

**▼B**

53,4 ml. $\text{KH}_2\text{PO}_4$ + 46,6 ml $\text{Na}_2\text{HPO}_4$	6,8
41,3 ml. $\text{KH}_2\text{PO}_4$ + 58,7 ml $\text{Na}_2\text{HPO}_4$	7,0
29,6 ml. $\text{KH}_2\text{PO}_4$ + 70,4 ml $\text{Na}_2\text{HPO}_4$	7,2
19,7 ml. $\text{KH}_2\text{PO}_4$ + 80,3 ml $\text{Na}_2\text{HPO}_4$	7,4
12,8 ml. $\text{KH}_2\text{PO}_4$ + 87,2 ml $\text{Na}_2\text{HPO}_4$	7,6
7,4 ml. $\text{KH}_2\text{PO}_4$ + 92,6 ml $\text{Na}_2\text{HPO}_4$	7,8
3,7 ml. $\text{KH}_2\text{PO}_4$ + 96,3 ml $\text{Na}_2\text{HPO}_4$	8,0

**▼B****C.8. TOXICITY FOR EARTHWORMS**

## ARTIFICIAL SOIL TEST

**1. METHOD****1.1. INTRODUCTION**

In this laboratory test, the test substance is added to an artificial soil in which worms are placed for 14 days. After this period (and optionally after seven days) the lethal effect of the substance on the earthworms is examined. The test provides a method for relatively short-term screening of the effect of chemicals on earthworms, by dermal and alimentary uptake.

**1.2. DEFINITION AND UNIT**

LC<sub>50</sub>: the concentration of a substance estimated as killing 50 % of the test animals during the test period.

**1.3. REFERENCE SUBSTANCE**

A reference substance is used periodically as a means of demonstration that the sensitivity of the test system has not changed significantly.

Analytical grade chloroacetamide is recommended as the reference substance.

**1.4. PRINCIPLE OF THE TEST**

Soil is a variable medium, so for this test a carefully defined artificial loam soil is used. Adult earthworms of the species *Eisenia foetida* (see note in Appendix) are kept in a defined artificial soil treated with different concentrations of the test substance. The content of the containers is spread on a tray 14 days (and optionally seven days) after the beginning of the test, and the earthworms surviving at each concentration counted.

**1.5. QUALITY CRITERIA**

The test is designed to be as reproducible as possible with respect to the test substrate and organism. Mortality in the controls must not exceed 10 % at the end of the test, or the test is invalid.

**1.6. DESCRIPTION OF THE TEST METHOD****1.6.1. Materials****1.6.1.1. Test substrate**

A defined artificial soil is used as a basic test substrate.

(a) Basic substrate (percentages are in terms of dry weight)

— 10 % sphagnum peat (as close to pH 5,5 to 6,0 as possible with no visible plant remains and finely ground),

**▼B**

- 20 % kaolinite clay with preferably more than 50 % kaolinite,
- about 69 % industrial quartz sand (dominant fine sand with more than 50 % of particle size 0,05 to 0,2 mm). If the substance is not sufficiently dispersible in water, 10 g per test container should be kept available for mixing with the test substance later on,
- about 1 % calcium carbonate (CaCO<sub>3</sub>), pulverised, chemically pure, added to bring the pH to 6,0 ± 0,5.

## (b) Test substrate

The test substrate contains the basic substrate, the test substance and deionised water.

Water content is about 25 to 42 % of the dry weight of the basic substrate. The water content of the substrate is determined by drying a sample to constant weight at 105 °C. The key criterion is that the artificial soil must be wetted to a point where there is no standing water. Care should be taken in mixing to obtain an even distribution of the test substance and the substrate. The way of introducing the test substance to the substrate has to be reported.

## (c) Control substrate

The control substrate contains the basic substrate and water. If an additive agent is used, an additional control should contain the same quantity of the additive agent.

## 1.6.1.2. Test containers

Glass containers of about one litre capacity (adequately covered with plastic lids, dishes or plastic film with ventilation holes) filled with an amount of wet test or control substrate equivalent to 500 g dry weight of substrate.

## 1.6.2. Test conditions

Containers should be kept in climatic chambers at a temperature of 20 ± 2 °C with continuous light. Light intensity should be 400 to 800 lux.

The test period is 14 days, but mortality can be assessed optionally seven days after starting the test.

## 1.6.3. Test procedure

## Test concentrations

Concentrations of the test substance are expressed as weight of substance per dry weight of basic substrate (mg/kg).

## Range finding test

The range of concentrations just causing mortalities of 0 to 100 % may be determined in a range-finding test to provide information on the range of concentrations to be used in the definitive test.

**▼B**

The substance should be tested at the following concentrations: 1 000; 100; 10; 1; 0,1 mg substance/kilogram test substrate (dry weight).

If a full definitive test is to be carried out, one test batch per concentration and one for the untreated control, each with 10 worms, could be sufficient for the range-finding test.

#### Definitive test

The results of the range-finding test are used to choose at least five concentrations in a geometric series just spanning the range 0 to 100 % mortality and differing by a constant factor not exceeding 1,8.

Tests using these series of concentration should allow the  $LC_{50}$  value and its confidence limits to be estimated as precisely as possible.

In the definitive test at least four test batches per concentration and four untreated controls, each with 10 worms, are used. The results of these replicate batches are given as a mean and standard deviation.

When two consecutive concentrations, at a ratio of 1,8, give only 0 % and 100 % mortality, these two values are sufficient to indicate the range within which the  $LC_{50}$  falls.

#### Mixture of the basic test substrate and the test substance

The test substrate should, whenever possible, be made up without any additional agents other than water. Immediately before the start of the test, an emulsion or dispersion of the test substance in deionised water or other solvent is mixed with the basic test substrate, or sprayed evenly over it with a fine chromatographic or similar spray.

If insoluble in water, the test substance can be dissolved in as small a volume as possible of suitable organic solvent (e.g. hexane, acetone or chloroform).

Only agents which volatilise readily may be used to solubilise, disperse or emulsify the test substance. The test substrate must be ventilated before use. The amount of water evaporated must be replaced. The control should contain the same quantity of any additive agent.

If the test substance is not soluble, dispersible or emulsifiable in organic solvents, 10 g of a mixture of fine ground quartz sand and a quantity of test substance necessary to treat 500 g dry weight of artificial soil are mixed with 490 g of dry weight of test substrate.

For each test batch, an amount of wet test substrate equivalent to 500 g dry weight is placed in each glass container and 10 earthworms, which have been conditioned for 24 hours in a similar wet basic substrate and then washed quickly and surplus water absorbed on filter paper before use, are placed on the test substrate surface.

**▼ B**

The containers are covered with perforated plastic lids, dishes or film to prevent the substrate drying and they are kept under the test conditions for 14 days.

The assessments should be made 14 days (and optionally seven days) after setting up the test. The substrate is spread on a plate made of glass or stainless steel. The earthworms are examined and the numbers of surviving earthworms determined. Earthworms are considered dead if they do not respond to a gentle mechanical stimulus to the front end.

When the examination is made at seven days, the container is refilled with the substrate and the surviving earthworms are replaced on the same test substrate surface.

#### 1.6.4. *Test organisms*

Test organisms should be adult *Eisenia foetida* (see note in Appendix) (at least two months old with clitellum) wet weight 300 to 600 mg. (For breeding method see Appendix.)

## 2. **DATA**

### 2.1. TREATMENT AND EVALUATION OF RESULTS

The concentrations of the substance tested are reported with reference to the corresponding percentages of dead earthworms.

When the data are adequate the  $LC_{50}$  value and the confidence limits ( $p = 0,05$ ) should be determined using standard methods (Litchfield and Wilcoxon, 1949, for equivalent method). The  $LC_{50}$  should be given as mg of test substance per kilogram of the test substrate (dry weight).

In those cases where the slope of the concentration curve is too steep to permit calculation of the  $LC_{50}$ , a graphical estimate of this value is sufficient.

When two consecutive concentrations at a ratio of 1,8 give only 0 % and 100 % mortality, the two values are sufficient to indicate the range within which the  $LC_{50}$  falls.

## 3. **REPORTING**

### 3.1. TEST REPORT

The test report shall, if possible, contain the following:

- statement that the test has been carried out in accordance with the abovementioned quality criteria,
- test carried out (range finding test and/or definitive test),
- exact description of the test conditions or statement that the test has been carried out in accordance with the method; any deviations have to be reported,
- exact description of how the test substance has been mixed into the basic test substrate,
- information about test organisms (species, age, mean and range in weight, keeping and breeding conditions, supplier),

**▼B**

- method used for determination of LC<sub>50</sub>,
- test results including all data used,
- description of observed symptoms or changes in behaviour of test organisms,
- mortality in the controls,
- LC<sub>50</sub> or highest tested concentration without mortality and lowest tested concentration with a mortality of 100 %, 14 days (and optionally seven days) after setting up the test,
- plotting of the concentration/response curve,
- results obtained with the reference substance, whether in association with the present test or from previous quality control exercises.

**4. REFERENCES**

- 1) OECD, Paris, 1981, Test Guideline 207, Decision of the Council C(81)30 final.
- 2) Edwards, C. A. and Lofty, J. R., 1977, Biology of Earthworms, Chapman and Hall, London, p. 331.
- 3) Bouche, M. B., 1972, Lombriciens de France, Ecologie et Systematique, Institut National de la Recherche Agronomique, p. 671.
- 4) Litchfield, J. T. and Wilcoxon, F., A simplified method of evaluation dose effect experiments. I. Pharm. Exp. Therap., vol. 96, 1949, p. 99.
- 5) Commission of the European Communities, Development of a standardised laboratory method for assessing the toxicity of chemical substances to earthworms, Report EUR 8714 EN, 1983.
- 6) Umweltbundesamt/Biologische Bundesanstalt für land- und Forstwirtschaft, Berlin, 1984, Verfahrensvorschlag 'Toxizitätstest am Regenwurm *Eisenia foetida* in künstlichem Boden', in: Rudolph/Boje, Ökotoxikologie, ecomed, Landsberg, 1986.

**▼ B***Appendix***Breeding and keeping of the worms before testing**

For breeding the animals, 30 to 50 adult worms, are put in a breeding box with fresh substrate and removed after 14 days. These animals may be used for further breeding batches. The earthworms hatched from the cocoons are used for testing when mature (under the prescribed conditions after two to three months).

**Keeping and breeding conditions**

Climatic chamber: temperature  $20 \pm 2$  °C preferably with continuous light (intensity 400 to 800 lux).

Breeding boxes: suitable shallow containers of 10 to 20 l volume.

Substrate: *Eisenia foetida* may be bred in various animal excrements. It is recommended to use as breeding medium a mixture of 50 % by volume peat and 50 % cow or horse dung. The medium should have a pH value of about 6 to 7 (regulated with calcium carbonate) and a low ionic conductivity (less than 6 mmhos or 0,5 % salt concentration).

The substrate should be moist but not too wet.

Other successful procedures may be used besides the method given above.

*Note: Eisenia foetida* exists in two races which some taxonomists have separated into species (Bouche, 1972). These are morphologically similar but one, *Eisenia foetida foetida*, has typically transverse striping or banding on the segments and the other, *Eisenia foetida andrei*, lacks this and has a variegated reddish colour. Where possible *Eisenia foetida andrei* should be used. Other species may be used if the necessary methodology is available.



**▼B****C.9. BIODEGRADATION****ZAHN-WELLENS TEST****1. METHOD****1.1. INTRODUCTION**

The purpose of the method is the evaluation of the potential ultimate biodegradability of water-soluble, non-volatile organic substances when exposed to relatively high concentrations of micro-organisms in a static test.

Physico-chemical adsorption on the suspended solids may take place and this must be taken into account when interpreting results (see 3.2).

The substances to be studied are used in concentrations corresponding to DOC-values in the range of 50 to 400 mg/litre or COD-values in the range of 100 to 1 000 mg/litre (DOC = dissolved organic carbon; COD = chemical oxygen demand). These relatively high concentrations have the advantage of analytical reliability. Compounds with toxic properties may delay or inhibit the degradation process.

In this method, the measure of the concentration of dissolved organic carbon or the chemical oxygen demand is used to assess the ultimate biodegradability of the test substance.

A simultaneous use of a specific analytical method may allow the assessment of the primary biodegradation of the substance (disappearance of the parent chemical structure).

The method is applicable only to those organic test substances which, at the concentration used in the test:

- are soluble in water under the test conditions,
- have negligible vapour pressure under the test conditions,
- are not inhibitory to bacteria,
- are adsorbed within the test system only to a limited extent,
- are not lost by foaming from the test solution.

Information on the relative proportions of the major components of the test material will be useful in interpreting the results obtained, particularly in those cases where the results are low or marginal.

Information on the toxicity of the substance to micro-organisms is desirable for the interpretation of low results and in the selection of appropriate test concentrations.

**▼ B**

## 1.2. DEFINITIONS AND UNITS

The amount of degradation attained at the end of the test is reported as the 'Biodegradability in the Zahn-Wellens test':

$$D_T(\%) = \left[ 1 - \frac{(C_T - C_B)}{(C_A - C_{BA})} \right] \times 100$$

where:

DT = biodegradation (%) at time T,

CA = DOC (or COD) values in the test mixture measured three hours after the beginning of the test (mg/l) (DOC = dissolved organic carbon, COD = chemical oxygen demand),

CT = DOC or COD values in the test mixture at time of sampling (mg/l),

CB = DOC or COD value of the blank at time of sampling (mg/l),

CBA = DOC or COD value of the blank, measured three hours after the beginning of the test (mg/l).

The extent of degradation is rounded to the nearest full percent.

Percentage degradation is stated as the percentage DOC (or COD) removal of the tested substance.

The difference between the measured value after three hours and the calculated or preferably measured initial value may provide useful information on the elimination of the substance (see 3.2, Interpretation of results).

## 1.3. REFERENCE SUBSTANCES

In some cases when investigating new substances reference substances may be useful; however, specific reference substances cannot yet be recommended.

## 1.4. PRINCIPLE OF THE TEST METHOD

Activated sludge, mineral nutrients and the test material as the sole carbon source in an aqueous solution are placed together in a one to four litre glass vessel equipped with an agitator and an aerator. The mixture is agitated and aerated at 20 to 25 °C under diffuse illumination or in a dark room for up to 28 days. The degradation process is monitored by determination of the DOC (or COD) values in the filtered solution at daily or other appropriate regular time intervals. The ratio of eliminated DOC (or COD) after each interval to the value three hours after the start is expressed as percentage biodegradation and serves as the measure of the extent of degradation at this time. The result is plotted versus time to give the biodegradation curve.

When a specific analytical method is used, changes in the concentration of the parent molecule due to biodegradation can be measured (primary biodegradability).

**▼B**

## 1.5. QUALITY CRITERIA

The reproducibility of this test has been proven to be satisfactory in a ring test.

The sensitivity of the method is largely determined by the variability of the blank and, to a lesser extent, by the precision of the determination of dissolved organic carbon and the level of test compound in the liquor.

## 1.6. DESCRIPTION OF THE TEST PROCEDURE

1.6.1. *Preparations*

## 1.6.1.1. Reagents

Test water: drinking water with an organic carbon content  $< 5$  mg/litre. The concentration of calcium and magnesium ions together must not exceed 2,7 mmole/litre; otherwise adequate dilution with deionised or distilled water is required.

Sulphuric acid, analytical reagent (A.R.):	50 g/l
Sodium hydroxide solution A.R.:	40 g/l
Mineral nutrient solution: dissolve in one litre deionised water:	
ammonium chloride, NH <sub>4</sub> Cl, A.R.:	38,5 g
sodium dihydrogenphosphate, NaH <sub>2</sub> PO <sub>4</sub> ·2H <sub>2</sub> O, A.R.:	33,4 g
potassium dihydrogenphosphate, KH <sub>2</sub> PO <sub>4</sub> , A.R.:	8,5 g
di-potassium mono-hydrogenphosphate, K <sub>2</sub> HPO <sub>4</sub> , A.R.:	21,75 g

The mixture serves both as a nutrient and as buffering system.

## 1.6.1.2. Apparatus

Glass vessels with a volume of one to four litres (e.g. cylindrical vessels).

Agitator with a glass or metal stirrer on a suitable shaft (the stirrer should rotate about 5 to 10 cm above the bottom of the vessel). A magnetic stirrer with a 7 to 10 cm long rod can be used instead.

Glass tube of 2 to 4 mm inner diameter to introduce air. The opening of the tube should be about 1 cm above the bottom of the vessel.

Centrifuge (about 3 550 g).

pH-meter.

Dissolved-oxygen meter.

Paper filters.

**▼B**

Membrane filtration apparatus.

Membrane filters, pore size 0,45 µm. Membrane filters are suitable if it is assured that they neither release carbon nor absorb the substance in the filtration step.

Analytical equipment for determining organic carbon content and chemical oxygen demand.

#### 1.6.1.3. Preparation of the inoculum

Activated sludge from a biological treatment plant is washed by (repeatedly) centrifuging or settling with test water (above).

The activated sludge must be in an appropriate condition. Such sludge is available from a properly working waste-water treatment plant. To get as many different species or strains of bacteria as possible, it may be preferred to mix inocula from different sources (e.g. different treatment plants, soil extracts, river waters, etc.). The mixture is to be treated as described above.

For checking the activity of the activated sludge see 'Functional control', below.

#### 1.6.1.4. Preparation of the test solutions

To the test vessel add 500 ml of test water, 2,5 ml/litre mineral nutrient solution and activated sludge in an amount corresponding to 0,2 to 1,0 g/litre dry matter in the final mixture. Add sufficient stock solution of the substance to be tested so that a DOC concentration of 50 to 400 mg/litre results in the final mixture. The corresponding COD-values are 100 to 1 000 mg/litre. Make up with test water to a total volume of one to four litres. The total volume to be chosen is dependent on the number of samples to be taken for DOC or COD determinations and the volumes necessary for the analytical procedure.

Normally a volume of two litres can be regarded as satisfactory. At least one control vessel (blank) is set up to run in parallel with each test series; it contains only activated sludge and mineral nutrient solution made up with test water to the same total volume as in the test vessels.

#### 1.6.2. *Performance of the test*

The test vessels are agitated with magnetic stirrers or screw propellers under diffuse illumination or in a dark room at 20 to 25 °C. Aeration is accomplished by compressed air cleaned by a cotton-wool strainer and a wash bottle if necessary. It must be ensured that the sludge does not settle and the oxygen concentration does not fall below 2 mg/litre.

The pH-value must be checked at regular intervals (e.g. daily) and adjusted to pH 7 to 8, if necessary.

**▼ B**

Losses from evaporation are made up just before each sampling with deionised or distilled water in the required amounts. A good procedure is to mark the liquid level on the vessel before starting the test. New marks are made after each sampling (without aeration and stirring). The first samples are always taken three hours after the start of the test in order to detect adsorption of test material by the activated sludge.

The elimination of the test material is followed by DOC or COD determinations made daily or at some other regular interval. The samples from the test vessel and the blank are filtered through a carefully washed paper filter. The first 5 ml of test solution filtrate are discarded. Sludges difficult to filter may be removed previously by centrifugation for 10 minutes. DOC and COD determinations are made at least in duplicate. The test is run for up to 28 days.

*Note:* samples remaining turbid are filtered through membrane filters. The membrane filters must not release or adsorb any organic material.

#### Functional control of activated sludge

A vessel containing a known substance should be run in parallel with each test series in order to check the functional capacity of the activated sludge. Diethyleneglycol has been found useful for this purpose.

#### Adaptation

If analyses are carried out at relatively short intervals (e.g. daily), adaptation can be clearly recognised from the degradation curve (see Figure 2). The test should therefore not be started immediately before the weekend.

If the adaptation occurs in the end of the period, the test can be prolonged until the degradation is finished.

*Note:* if a broader knowledge of the behaviour of the adapted sludge is needed, the same activated sludge is exposed once again to the same test material in accordance with the following procedure:

Switch of the agitator and the aerator and allow the activated sludge to settle. Draw off the supernatant liquid, fill up to two litres with test water, stir for 15 minutes and allow to settle again. After the supernatant liquid is drawn off again, use the remaining sludge to repeat the test with the same material in accordance with 1.6.1.4 and 1.6.2, above. The activated sludge can also be isolated by centrifuging instead of settling.

The adapted sludge may be mixed with fresh sludge to a concentration of 0,2 to 1 g dry weight/litre.

**▼B****Analytical means**

Normally samples are filtered through a carefully washed paper filter (for washing use deionised water).

Samples which remain turbid are filtered through membrane filters (0,45 µm).

The DOC concentration is determined in duplicate in the sample filtrates (the first 5 ml are discarded) by means of the TOC instrument. If the filtrate cannot be analysed on the same day, it must be stored in the refrigerator until the next day. Longer storage cannot be recommended.

The COD concentration: is determined in the sample filtrates with a COD analytical set-up by the procedure described in reference (2) below.

**2. DATA AND EVALUATION**

DOC and/or COD concentrations are determined at least in duplicate in the samples according to 1.6.2 above. The degradation at time T is calculated according to the formula (with definitions) given under 1.2 above.

The extent of degradation is rounded to the nearest full percent. The amount of degradation attained at the end of the test is reported as the 'Biodegradability in the Zahn-Wellens test'.

*Note:* if complete degradation is attained before the test time is over and this result is confirmed by a second analysis on the next day, the test can be concluded.

**3. REPORTING****3.1. TEST REPORT**

The test report shall, if possible, contain the following:

- the initial concentration of the substance,
- all other information and the experimental results concerning the tested substance, the reference substance if used, and the blank,
- the concentration after three hours,
- biodegradation: curve with description,
- date and location where test organisms were sampled, status of adaptation, concentration used, etc.,
- scientific reasons for any changes of test procedure.

**3.2. INTERPRETATION OF RESULTS**

Removal of DOC (COD) which takes place gradually over days or weeks indicates that the test substance is being biodegraded.

**▼B**

However, physico-chemical adsorption can, in some cases, play a role and this is indicated when there is complete or partial removal from the outset, within the first three hours, and the difference between control and test supernatant liquors remains at an unexpectedly low level.

Further tests are necessary if a distinction is to be drawn between biodegradation (or partial biodegradation) and adsorption.

This can be done in a number of ways, but the most convincing is to use the supernatant or sludge as inoculum in a base-set test (preferably a respirometric test).

Test substances giving high, non-adsorptive removal of DOC (COD) in this test should be regarded as potentially biodegradable. Partial, non-adsorptive removal indicates that the chemical is at least subject to some biodegradation. Low, or zero removals of DOC (COD) may be due to inhibition of microorganisms by the test substance and this may also be revealed by lysis and loss of sludge, giving turbid supernatants. The test should be repeated using a lower concentration of test substance.

The use of a compound-specific analytical method or of  $^{14}\text{C}$ -labelled test substance may allow greater sensitivity. In the case of  $^{14}\text{C}$  test compound, the recovery of the  $^{14}\text{CO}_2$  will confirm that biodegradation has occurred.

When results are given in terms of primary biodegradation, an explanation should, if possible, be given on the chemical structure change that leads to the loss of response of the parent test substance.

The validation of the analytical method must be given together with the response found on the blank test medium.

#### 4. REFERENCES

- (1) OECD, Paris, 1981, Test Guideline 302 B, Decision of the Council C(81) 30 final.
- (2) Annex V C.9 Degradation: Chemical Oxygen Demand, Commission Directive 84/449/EEC, (OJ L 251, 19.9.1984, p. 1).



## Appendix

## EVALUATION EXAMPLE

Organic compound:	4-Ethoxybenzoic acid
Theoretical test concentration:	600 mg/l
Theoretical DOC:	390 mg/l
Inoculum	Sewage treatment plant of...
Concentration:	1 gram dry material/litre
Adaptation status:	not adapted
Analysis:	DOC-determination
Amount of sample:	3 ml
Control substance:	Diethyleneglycol
Toxicity of compound:	No toxic effects below 1 000 mg/l
	Test used: fermentation tubes test

Test time	Control substance				Test substance		
	Blank DOC <sup>(1)</sup> mg/l	DOC <sup>(1)</sup> mg/l	DOC net mg/l	Degradation %	DOC <sup>(1)</sup> mg/l	DOC net mg/l	Degradation %
0	—	—	300,0	—	—	390,0	—
3 hours	4,0	298,0	294,0	2	371,6	367,6	6
1 day	6,1	288,3	282,2	6	373,3	367,2	6
2 days	5,0	281,2	276,2	8	360,0	355,0	9
5 days	6,3	270,5	264,2	12	193,8	187,5	52
6 days	7,4	253,3	245,9	18	143,9	136,5	65
7 days	11,3	212,5	201,2	33	104,5	93,2	76
8 days	7,8	142,5	134,7	55	58,9	51,1	87
9 days	7,0	35,0	28,0	91	18,1	11,1	97
10 days	18,0	37,0	19,0	94	20,0	2,0	99

<sup>(1)</sup> Mean values of triplicate determinations.



▼B

Figure 1

Examples of biodegradation curves

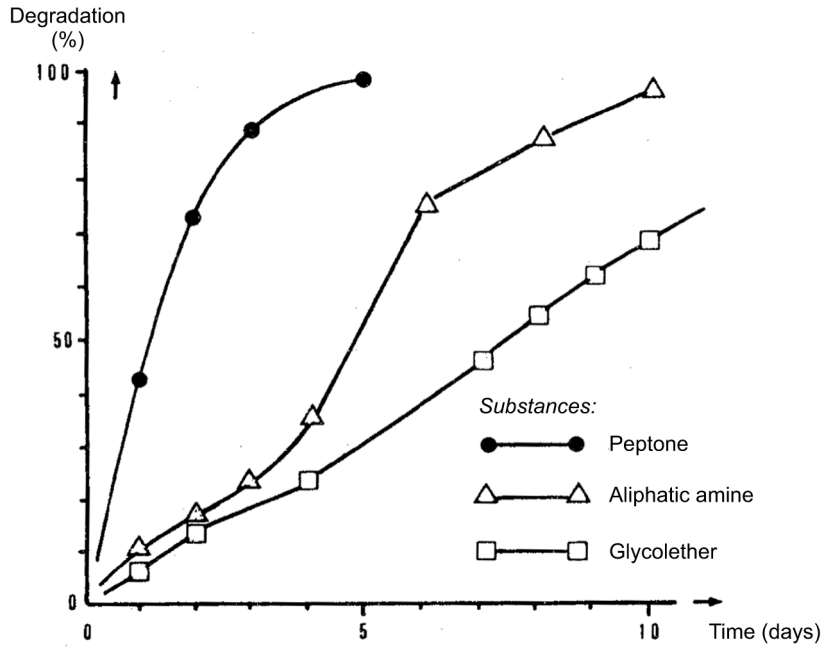
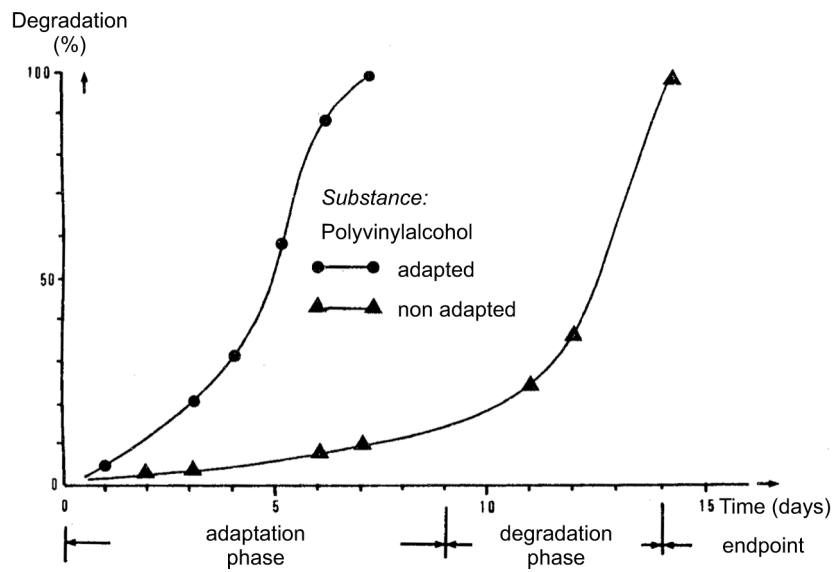


Figure 2

Examples of sludge adaptation



**▼M4****C.10. SIMULATION TEST AEROBIC SEWAGE TREATMENT: C.10-A: ACTIVATED SLUDGE UNITS — C.10-B: BIOFILMS****C.10-A: Activated Sludge Units**

## INTRODUCTION

1. This Test Method is equivalent to OECD Test Guideline (TG) 303 (2001). In the 1950s it was realised that the newly introduced surfactants caused excessive foaming in waste water treatment plants and in rivers. They were not fully removed in the aerobic treatment and in some cases limited the removal of other organic matter. This instigated many investigations into how surfactants could be removed from waste waters and whether new chemicals produced by industry were amenable to waste water treatment. In order to do this, model units were used representing the two main types of aerobic biological waste water treatment (activated sludge and percolating, or trickling, filtration). It would have been impractical and very costly to distribute each new chemical and to monitor large-scale treatment plants, even on a local basis.

## INITIAL CONSIDERATIONS

**Activated sludge units**

2. Model activated sludge units have been described ranging in size from 300 ml up to about 2 000 ml. Some closely mimicked full-scale plants, having sludge settlement tanks with settled sludge being pumped back to the aeration tank, while others provided no settlement facilities e.g. Swisher (1). The size of the apparatus is a compromise; on the one hand, it must be large enough for successful mechanical operation and for the provision of sufficient volume of samples without affecting the operation, while on the other hand it should not be so large that it demands excessive space and materials.
3. Two forms of apparatus which have been extensively and satisfactorily used are the Husmann units (2) and Porous Pot units (3)(4), first used in the study of surfactants; these are described in this Test Method. Others have also been used satisfactorily, e.g. Eckenfelder (5). Because of the relatively high cost and effort of applying this simulation test, simpler and cheaper screening tests, now embodied in chapter C.4 A-F of this Annex (6) were investigated in parallel. Experience with many surfactants and other chemicals has shown that those which passed the screening tests (readily biodegradable) also degraded in the simulation test. Some of those failing the screening tests passed the inherent biodegradability tests (chapters C.12 (7) and C.19 (8) of this Annex) but only some of this latter group were degraded in the simulation test, while those chemicals which failed tests for inherent biodegradability did not degrade in the simulation tests (9)(10)(11).
4. For some purposes simulation tests carried out under a single set of operating conditions are sufficient; the results are expressed as a percentage removal of the test chemical or of dissolved organic carbon (DOC). A description of such a test is given in this test method. However, unlike the previous version of this chapter, which described only one type of apparatus treating synthetic sewage in the coupled mode using a relatively

**▼M4**

crude method of sludge wastage, this text offers a number of variations. Alternatives to the type of apparatus, mode of operation, sewage and sludge wastage removal are described. This text closely follows that of ISO 11733 (12), which was carefully scrutinised during its preparation, though the method has not been subject to a ring test.

5. For other purposes the concentration of the test chemical in the effluent is required to be known more accurately and for this a more extensive method is needed. For example, the sludge wastage rate must be more precisely controlled throughout each day and throughout the period of the test, and units have to be run at a number of wastage rates. For a fully comprehensive method, tests should also be run at two or three different temperatures: such a method is described by Birch (13)(14) and summarised in Appendix 6. However, present knowledge is insufficient to decide which of the kinetic models are applicable to the biodegradation of chemicals in waste water treatment and in the aquatic environment generally. The application of Monod kinetics, given in Appendix 6 as an example, is limited to chemicals present at 1 mg/l and above, but in the opinion of some even this remains to be substantiated. Tests at concentrations more truly reflecting those found in waste waters are indicated, in Appendix 7, but such tests, and those in Appendix 6, are included in Appendices instead of being issued as separate Test Methods.

*Filters*

6. Much less attention has been given to model percolating filters, perhaps because they are more cumbersome and less compact than activated sludge plant models. Gerike et al developed trickling filter units and operated them in the coupled mode ((15). These filters were relatively large (height 2 m; volume 60 l) and each required as much as 2 l/h of sewage. Baumann et al (16), simulated trickling filters by inserting polyester 'fleece' strips into 1 m tubes (14 mm int. diameter) after the strips had been immersed in concentrated activated sludge for 30 min. The test chemical as sole C source in a mineral salts solution was fed down the vertical tube and biodegradation was assessed from measurements of DOC in the effluent and CO<sub>2</sub> in the issuing gas.
7. Biofilters have been simulated in another way (15); the inner surfaces of rotating tubes, inclined at a small angle to the horizontal, were fed with sewage (about 250 ml/h) with and without the test chemical, and the collected effluents analysed for DOC and/or the specific test chemical.

## PRINCIPLE OF THE TEST

8. This method is designed to determine the elimination and the primary and/or ultimate biodegradation of water-soluble organic chemicals by aerobic micro-organisms in a continuously operated test system simulating the activated sludge process. An easily biodegradable organic medium and the organic test chemical are the sources of carbon and energy for the micro-organisms.
9. Two continuously operated test units (activated sludge plants or porous pots) are run in parallel under identical conditions which are chosen to suit the purpose of the test. Normally the mean hydraulic retention time is 6 h and the mean sludge age (sludge retention time) is 6 to 10 days. Sludge is wasted by one of two methods, the test chemical is normally added at a concentration of between 10 mg/l dissolved organic carbon (DOC) and 20 mg/l DOC, to the influent (organic medium) of only one of the units. The second unit is used as a control unit to determine the biodegradation of the organic medium.

**▼M4**

10. In frequently taken samples of the effluents, the DOC, preferably, or chemical oxygen demand (COD) is determined, together with the concentration of the test chemical (if required) by specific analysis, in the effluent from the unit receiving the test chemical. The difference between the effluent concentrations of DOC or COD in the test and control units is assumed to be due to the test chemical or its organic metabolites. This difference is compared with the influent concentration of DOC or COD due to the added test chemical in order to determine the elimination of the test chemical.
11. Biodegradation may normally be distinguished from bioadsorption by careful examination of the elimination-time curve and may usually be confirmed by applying a test for ready biodegradation using an acclimatised inoculum from the unit receiving the test chemical.

## INFORMATION ON THE TEST CHEMICAL

12. The purity, water solubility, volatility and adsorption characteristics of the test chemical should be known to enable correct interpretation of results to be made. Normally volatile and insoluble chemicals cannot be tested unless special precautions are taken (see Appendix 5). The chemical structure, or at least the empirical formula, should also be known in order to calculate theoretical values and/or to check measured values of parameters, e.g. theoretical oxygen demand (ThOD), dissolved organic carbon (DOC) and chemical oxygen demand (COD).
13. Information on the toxicity of the test chemical to micro-organisms (see Appendix 4) may be useful for selecting appropriate test concentrations and may be essential for the correct interpretation of low biodegradation values.

## PASS LEVELS

14. In the original application of this simulation (confirmatory) test to the primary biodegradation of surfactants, a removal of more than 80 % of the specific chemical is required before the surfactant may be marketed. If the value of 80 % is not attained, this simulation (confirmatory) test may be applied and the surfactant may be marketed only if more than 90 % of the specific chemical is removed. With chemicals in general there is no question of pass/fail and the value of percentage removal obtained can be used in proximate calculations of the probable environmental concentration to be used in hazard assessments posed by chemicals. Results tend to follow an all or nothing pattern. In a number of studies of pure chemicals the percentage removal of DOC was found to be > 90 % in more than three quarters and > 80 % in over 90 % of chemicals which showed any significant degree of biodegradability.
15. Relatively few chemicals (e.g. surfactants) are present in sewage at the concentrations (about 10 mg C/l) used in this test. Some chemicals may be inhibitory at these concentrations, while the kinetics of removal of others may be different at low concentrations. A more accurate assessment of the degradation could be made by using modified methods, using realistically low concentrations of the test chemical, and the data collected could be used to calculate kinetic constants. However, the necessary experimental techniques have not yet been fully validated and neither have the kinetic models, which describe the biodegradation reactions, been established (see Appendix 7).

**▼M4**

## REFERENCE CHEMICALS

16. To ensure that the experimental procedure is being carried out correctly, it is useful occasionally to test chemicals whose behaviour is known simultaneously when test chemicals are investigated. Such chemicals include adipic acid, 2-phenyl phenol, 1-naphthol, diphenic acid, 1-naphthoic acid, etc. (9)(10)(11).

## REPRODUCIBILITY OF TEST RESULTS

17. There have been far fewer reports of studies of simulation tests than of tests for ready biodegradability. Reproducibility between (simultaneous) replicates is good (within 10-15 %) for test chemicals degraded by 80 % or more but for less well degraded chemicals variability is greater. Also, with some borderline chemicals widely disparate results (e.g. 10 %, 90 %) have been recorded on different occasions within the 9 weeks allowed in the test.
18. Little difference has been found in results obtained with the two types of apparatus, but some chemicals have been more extensively and consistently degraded in the presence of domestic sewage than with OECD synthetic sewage.

## DESCRIPTION OF THE TEST METHOD

**Apparatus***Test system*

19. The test system for one test chemical consists of a test unit and a control unit; but when only specific analyses are performed (primary biodegradation) only a test unit is required. One control unit can be used for several test units receiving either the same or different test chemicals. In the case of coupling (Appendix 3) each test unit must have its own control unit. The test system may be either an activated sludge plant model, Husmann unit (Appendix 1, Figure 1) or a porous pot (Appendix 1, Figure 2). In both cases storage vessels of sufficient size for the influents and effluents are needed, as well as pumps to dose the influent, either mixed with solution of the test chemical or separately.
20. Each activated sludge plant unit consists of an aeration vessel with a known capacity of about 3 litres of activated sludge and a separator (secondary clarifier) which holds about 1,5 litres; the volumes can, to some extent, be changed by adjusting the height of the separator. Vessels of different sizes are permissible if they are operated with comparable hydraulic loads. If it is not possible to keep the temperature in the test room in the desired range, the use of water-jacketed vessels with temperature controlled water is recommended. An airlift pump or a dosing pump is used to recycle the activated sludge from the separator to the aeration vessel, either continuously or intermittently at regular intervals.
21. The porous pot system consists of an inner, porous cylinder with a conical bottom held in a slightly larger vessel of the same shape, but made of an impervious plastic material. A suitable material for the porous vessel is porous polyethylene of maximum pore size 90 µm and 2 mm thickness. Separation of the sludge from the treated organic medium is effected by differential passage through the porous wall. Effluents collect in the annular space from where it overflows into the collecting vessel. No settlement occurs and hence there is no sludge return. The whole system may be mounted in a thermostatically controlled water-bath. Porous pots become

▼ **M4**

blocked and could overflow in the initial stages. In such a case, replace the porous liner with a clean one by first siphoning the sludge from the pot into a clean bucket and removing the blocked liner. After wiping out the impervious outer cylinder insert a clean liner and return the sludge to the pot. Any sludge adhering to the sides of the blocked liner is also carefully scraped off and transferred. Clean blocked pots first by using a fine jet of water to remove remaining sludge and by soaking in dilute sodium hypochlorite solution, then in water, followed by thoroughly rinsing with water.

22. For aeration of the sludge in the aeration vessels of both systems, suitable techniques are required, for example sintered cubes (diffuser stones) and compressed air. The air shall be cleaned, if necessary, by passing through a suitable filter and washed. Sufficient air must pass through the system to maintain aerobic conditions and to keep sludge flocs in suspension at all times during the test.

*Filtration apparatus or centrifuge*

23. Device for filtration of samples with membrane filters of suitable porosity (nominal aperture diameter 0,45 µm) which adsorb soluble organic chemicals and release organic carbon to a minimum degree. If filters are used which release organic carbon, wash the filters carefully with hot water to remove leachable organic carbon. Alternatively, a centrifuge capable of producing 40 000 m/s<sup>2</sup> may be used.

*Analytical equipment*

24. Apparatus required to determine:
- DOC(dissolved organic carbon) and TOC (total organic carbon), or COD (chemical oxygen demand);
  - specific chemical, if required;
  - suspended solids, pH, oxygen concentration in water;
  - temperature, acidity and alkalinity;
  - ammonium, nitrite and nitrate, if the test is performed under nitrifying conditions.

*Water*

25. Tap water, containing less than 3 mg/l DOC. Determine the alkalinity if not already known.
26. Deionised water, containing less than 2 mg/l DOC.

*Organic medium*

27. Synthetic sewage, domestic sewage or a mixture of both is permissible as the organic medium. It has been shown (11)(14) that the use of domestic sewage alone often gives increased percentage DOC removal and even allows the removal and biodegradation of some chemicals which are not biodegraded when OECD synthetic sewage is used. Also, the constant or intermittent addition of domestic sewage often stabilises the activated sludge, including the crucial ability to settle well. Thus, the use of domestic sewage is recommended. Measure the DOC or COD concentration in each new batch of organic medium. The acidity or alkalinity of the organic medium should be known. The organic medium may require the addition of a suitable buffer (sodium hydrogen carbonate or potassium dihydrogen phosphate) if it is of low acidity or alkalinity, to maintain a pH of about 7,5 ± 0,5 in the aeration vessel during the test. The amount of buffer to be added, and when to add it, has to be decided in each individual case. When mixtures are used either continuously or intermittently, the DOC (or COD) of the mixture must be kept at an approximately constant value, e.g. by dilution with water.

**▼ M4***Synthetic sewage*

28. Dissolve in each litre of tap water: peptone, 160 mg; meat extract, 110 mg; urea, 30 mg; anhydrous dipotassium hydrogen phosphate ( $K_2HPO_4$ ), 28 mg; sodium chloride (NaCl), 7 mg; calcium chloride dihydrate ( $CaCl_2 \cdot 2H_2O$ ), 4 mg; magnesium sulphate heptahydrate ( $Mg_2SO_4 \cdot 7H_2O$ ), 2 mg. This OECD synthetic sewage is an example and gives a mean DOC concentration in the influent of about 100 mg/l. Alternatively, use other compositions, with about the same DOC concentration, which are closer to real sewage. If a less concentrated influent is required, dilute the synthetic sewage, for example 1:1, with tap water to obtain a concentration of about 50 mg/l. Such a weaker influent will allow better growth of nitrifying organisms and this modification should be used if the simulation of nitrifying waste water plants is to be investigated. This synthetic sewage may be made up in distilled water in a concentrated form and stored at about 1 °C for up to one week. When needed, dilute with tap water. (This medium is unsatisfactory e.g. nitrogen concentration is very high, relatively low carbon content, but nothing better has been suggested, except to add more phosphate as buffer and extra peptone).

*Domestic sewage*

29. Use fresh settled sewage collected daily from a treatment works receiving predominantly domestic sewage. It should be collected, prior to primary sedimentation, from the overflow channel of the primary sedimentation tank, or from the feed to the activated sludge plant, and be largely free from coarse particles. The sewage can be used after storage for several days (but generally should not exceed seven days) at about 4 °C, if it is proved that the DOC (or COD) has not significantly decreased (i.e. by less than 20 %) during storage. In order to limit disturbances to the system, the DOC (or COD) of each new batch should be adjusted before use to an appropriate constant value, e.g. by dilution with tap water.

*Activated sludge*

30. Collect activated sludge for inoculation from the aeration tank of a well operated waste water treatment plant or from a laboratory — scale activated sludge unit, treating predominantly domestic sewage.

*Stock solutions of test chemical*

31. For chemicals of adequate solubility, prepare stock solutions at appropriate concentrations (e.g. 1 to 5 g/l) in deionised water, or in the mineral portion of the synthetic sewage. (for insoluble and volatile chemicals, see Appendix 5). Determine the DOC and total organic carbon (TOC) of the stock solution and repeat the measurements for each new batch. If the difference between the DOC and TOC is greater than 20 %, check the water-solubility of the test chemical. Compare the DOC or the concentration of the test chemical measured by specific analysis of the stock solution with the nominal value, to ascertain whether recovery is good enough (normally > 90 % can be expected). Ascertain, especially for dispersions, whether or not DOC can be used as an analytical parameter or if only an analytical technique specific for the test chemical can be used. Centrifugation of the samples is required for dispersions. For each new batch, measure the DOC, COD or the test chemical with specific analysis.

**▼M4**

32. Determine the pH of the stock solution. Extreme values indicate that the addition of the chemical may have an influence on the pH of the activated sludge in the test system. In this case neutralise the stock solution to obtain a pH of  $7 \pm 0,5$  with small amounts of inorganic acid or base, but avoid precipitation of the test chemical.

## PROCEDURE

33. The procedure is described for the activated sludge plant units; it has to be slightly adapted for the porous pot system.

*Preparation of the inoculum*

34. Inoculate the test system at the beginning of the test with either activated sludge or an inoculum containing a low concentration of micro-organisms. Keep the inoculum aerated at room temperature until it is used and use it within 24 h. In the first case, take a sample of activated sludge from the aeration tank of an efficiently operated biological waste water treatment plant, or a laboratory treatment plant, which receives predominantly domestic sewage. If nitrifying conditions are to be simulated, take sludge from a nitrifying waste water treatment plant. Determine the concentration of suspended solids and, if necessary, concentrate the sludge by settling so that the volume added to the test system is minimal. Ensure that the starting concentration of dry matter is about 2,5 g/l.
35. In the second case, use 2 ml/l to 10 ml/l of an effluent from a domestic biological waste water treatment plant as an inoculum. To get as many different species of bacteria as possible, it may be helpful to add inocula from various other sources, for example surface water. In this case, the activated sludge will develop and grow in the test system.

*Dosage of organic medium*

36. Ensure that influent and effluent containers and tubing from influent vessels and to effluent vessels are thoroughly cleaned to remove microbial growths initially and throughout the test. Assemble the test systems in a room where the temperature is controlled (normally in the range 20-25 °C) or use water-jacketed test units. Prepare a sufficient volume of the required organic medium (paragraphs 27-29). Initially fill the aeration vessel and the separator with the organic medium and add the inoculum (paragraphs 34, 35). Start the aeration such that the sludge is kept in suspension and in an aerobic state and begin dosing the influent and recycling the settled sludge. Dose organic medium out of storage vessels into the aeration vessels (paragraphs 20, 21) of the test and control units and collect the respective effluents in similar storage vessels. To get the normal hydraulic retention time of 6 h, the organic medium is pumped at 0,5 l/h. To confirm this rate, measure the daily amount of organic medium dosed by noting the reduction in volumes of the medium in the storage vessels. Other modes of dosing would be necessary for determining the effects of intermittent release and 'shock' loading of chemicals.
37. If the organic medium is prepared for use for a period longer than 1 day, cooling at about 4 °C, or other appropriate methods of conservation are necessary to prevent microbial growth and biodegradation outside the test units (paragraph 29). If synthetic sewage is used, it is possible to prepare, and store at about 4 °C, a concentrated stock solution (e.g. 10-fold the normal concentration, paragraph 28). This stock solution can be well mixed with the appropriate volume of tap water before use; alternatively, it can be pumped directly while the appropriate amount of tap water is pumped separately.



▼ **M4***Dosage of test chemical*

38. Add an appropriate volume of the stock solution of the test chemical (paragraph 31) to the storage vessel of the influent or dose it directly with a separate pump into the aeration vessel. The normal mean test concentration in the influent should be between 10 mg/l and 20 mg/l DOC, with an upper concentration of no more than 50 mg/l. If the water-solubility of the test chemical is low or if toxic effects are likely to occur, reduce the concentration to 5 mg/l DOC or even less, but only if a suitable specific analytical method is available and performed (dispersed test chemicals which are poorly soluble in water may be added using special dosing techniques, see Appendix 5).
39. Start adding the test chemical after a period in which the system has stabilised and is removing DOC of the organic medium efficiently (about 80 %). It is important to check that all units are working equally efficiently before the addition of test chemical; if they are not, it usually helps to mix the individual sludges and to re-disperse equal volumes to individual units. When an inoculum of (about) 2,5 g/l (dry weight) activated sludge is used, the test chemical may be added from the start of the test since directly adding increasing amounts from the beginning has the advantage that the activated sludge may be better able to adapt to the test chemical. In whatever manner the test chemical is added, it is recommended that the relevant flow rate and/or the volumes in the storage vessel(s) are measured at regular intervals.

*Handling of activated sludge*

40. The concentration of activated sludge solids normally stabilises between limits during the test, independent of the inoculum used, in the range 1 to 3 g/l (dry weight) depending on the quality and concentration of the organic medium, operating conditions, the nature of the micro-organisms present and the influence of the test chemical.
41. Either determine the suspended solids in the aeration vessels at least weekly and discard surplus sludge to maintain the concentration at 1 g/l to 3 g/l (dry weight), or control the mean sludge age at a constant value usually in the range 6 days to 10 days. If, for example, a sludge retention time of 8 days is chosen, remove daily 1/8 of the volume of the activated sludge in the aeration vessel and discard it. Carry this out on a daily basis or, preferably, by means of an automatic intermittently operating pump. Maintaining the concentration of suspended solids constant, or within narrow limits, does not maintain a constant sludge retention time (SRT), which is the operating variable that determines the value of the concentration of test chemical in the effluent.
42. Throughout the test, remove, at least daily, any sludge adhering to the walls of the aeration vessel and the separator so that it is resuspended. Check and clean regularly all tubes and tubing to prevent growth of biofilm. Recycle the settled sludge from the separator to the aeration vessel, preferably by intermittent pumping. No recycling takes place in the porous pot system but ensure that clean inner pots are inserted before the volume in the vessel rises significantly (paragraph 21).
43. Poor settlement and loss of sludge may occur in the Husmann plant units. These may be rectified by employing one or more of the actions, listed below, in parallel in test and control units:

— fresh sludge or flocculant (for example 2 ml/vessel of 50 g/l  $\text{FeCl}_3$ ) could be added at regular intervals, e.g. weekly, but ascertain that no reaction or precipitation of the test chemical occurs with  $\text{FeCl}_3$ ;

▼ **M4**

- the air-lift pump could be replaced by a peristaltic pump, thus enabling a sludge recirculation flow which about equals the influent flow to be used and allowing development of an anaerobic zone in the settled sludge (the geometry of the air-lift pump limits the minimum flow rate of returned sludge to be about 12-fold that of the influent);
- sludge could be pumped intermittently from the separator to the aeration vessel (e.g. 5 min. every 2,5 h to recycle 1 l/h to 1,5 l/h);
- a non-toxic, anti-foaming agent at minimal concentration could be used to prevent loss by foaming (e.g. silicone oil);
- air could be passed through the sludge in the separator in short, shock bursts (e.g. 10 sec. every hour);
- the organic medium may be dosed at intervals into the aeration vessel (e.g. 3 min. to 10 min. every hour).

*Sampling and analysis*

44. At regular intervals measure the dissolved oxygen concentration, the temperature and the pH value of the activated sludge in the aeration vessels. Ensure that sufficient oxygen is always available ( $> 2$  mg/l) and that the temperature is kept in the required range (normally  $20$  °C to  $25$  °C). Keep the pH at  $7,5 \pm 0,5$  by dosing small amounts of inorganic base or acid into the aeration vessel or into the influent, or by increasing the buffering capacity of the organic medium (see paragraph 27). When nitrification occurs acid is produced, the oxidation of 1 mg N producing the equivalent of about 7 mg  $\text{CO}_3^-$ . The frequency of measuring depends on the parameter to be measured and the stability of the system, and may vary between daily and weekly measurements.
45. Measure the DOC or COD in the influents to the control and test vessels. Measure the test chemical concentration in the test influent by specific analysis or estimate it from the concentration in the stock solution (paragraph 31), the volume used and the amount of sewage dosed into the test unit. It is recommended that the concentration of the test chemical be calculated in order to reduce the variability of the concentration data.
46. Take suitable samples from the collected effluent (e.g. 24 h composites) and filter through a membrane of pore size  $0,45$   $\mu\text{m}$  or centrifuge them at about  $40,000$   $\text{m/s}^2$  for about 15 min. Centrifuging should be used if filtering is difficult. Determine DOC or COD at least in duplicate to measure ultimate biodegradation and, if required, primary biodegradation by an analysis specific for the test chemical.
47. The use of COD may give rise to analytical problems at low concentrations and is therefore recommended only if a sufficiently high test concentration (about 30 mg/l) is used. Also, for strongly adsorbing chemicals, it is recommended that the amount of adsorbed chemical in the sludge be measured using an analytical technique specific for the test chemical.
48. The frequency of sampling depends on the expected duration of the test. A recommended frequency is three times per week. Once the units are operating efficiently, allow from 1 week to a maximum of 6 weeks after the test chemical has been introduced, for adaptation to reach a steady state. Preferably obtain at least 15 valid values in the plateau phase (paragraph 59), normally lasting 3 weeks, for the evaluation of the test result. The test may be completed if a sufficient degree of elimination is reached (e.g.  $> 90$  %) and these 15 values, which represent analyses carried out each weekday over 3 weeks, are available. Normally, do not exceed a test duration of more than 12 weeks after addition of the test chemical.

**▼ M4**

49. If the sludge nitrifies and if the effects of the test chemical on nitrification are to be studied, analyse samples from the effluent of the test and control units at least once per week for ammonium and/or nitrite plus nitrate.
50. All analyses should be performed as soon as possible, especially the nitrogen determinations. If analyses have to be postponed, store the samples at about 4 °C in the dark in full, tightly stopped bottles. If samples have to be stored for more than 48 h, preserve them by deep-freezing, acidification (e.g. 10 ml/l of a 400 g/l solution of sulphuric acid) or by addition of a suitable toxic substance (e.g. 20 ml/l of a 10 g/l solution of mercury (II) chloride). Ensure that the preservation technique does not influence results of analysis.

*Coupling of test units*

51. If coupling is to be used (Appendix 3), daily exchange the same amount of activated sludge (150 ml to 1 500 ml for aeration vessels containing 3 litres of liquor) between the aeration vessels of the test unit and its control unit. If the test chemical adsorbs strongly onto the sludge, change only the supernatant of the separators. In both cases use a correction factor to calculate the test results (paragraph 55).

**DATA AND REPORTING****Treatment of results**

52. Calculate the percentage of DOC or COD elimination of the test chemical for each timed assessment, using the equation:

$$D_t = \frac{C_s - (E - E_0)}{C_s} \times 100$$

where

$D_t$  = % elimination of DOC or COD at time t

$C_s$  = DOC or COD in the influent due to the test chemical, preferably estimated from the stock solution (mg/l)

$E$  = measured DOC or COD value in the test effluent at time t (mg/l)

$E_0$  = measured DOC or COD value in the control effluent at time t (mg/l)

53. The degree of DOC or COD elimination of the organic medium in the control unit is helpful information in assessing the biodegradative activity of the activated sludge during the test. Calculate the percentage elimination from the equation:

$$D_B = \frac{C_M - E_0}{C_M} \times 100$$

where

$D_B$  = % elimination of DOC or COD of the organic medium in the control unit at time t

$C_M$  = DOC or COD of the organic medium in the control influent (mg/l)

**▼ M4**

Optionally, calculate the percentage elimination DOC or COD due to the organic medium plus test chemical in the test unit from the equation:

$$D_T = \frac{C_T - E}{C_T} \times 100$$

where

$D_T$  = % elimination of total test influent DOC or COD

$C_T$  = DOC or COD of total test influent or calculated from stock solutions (mg/l)

54. Calculate the removal of the test chemical if measured with a specific analytical method at each time assessment from equation:

$$D_{ST} = \frac{S_i - S_e}{S_i} \times 100$$

where

$D_{ST}$  = % primary elimination of test chemical at time t

$S_i$  = measured or estimated test chemical concentration in the test influent (mg/l)

$S_e$  = measured test chemical concentration in test effluent at time t (mg/l)

55. If the coupling mode has been used, compensate the dilution of the test chemical in the aeration vessel by the sludge exchange using a correction factor (see Appendix 3). If a mean hydraulic retention time of 6 h and an exchange of half of the volume of the activated sludge in the aeration vessel have been used, the determined daily elimination values ( $D_t$ , paragraph 52) have to be corrected to obtain the true degree of elimination,  $D_{tc}$ , of the test chemical from the equation:

$$D_{tc} = \frac{4D_t - 100}{3}$$

#### Expression of test results

56. Plot the percentage elimination  $D_t$  (or  $D_{tc}$ ) and  $D_{st}$ , if available, versus time (see Appendix 2). From the shape of the elimination curve of the test chemical (*per se* or as DOC) some conclusions may be drawn about the removal process.

#### Adsorption

57. If a high DOC elimination of the test chemical is observed from the beginning of the test, the test chemical is probably eliminated by adsorption onto the activated sludge solids. It is possible to prove this by determining the adsorbed test chemical by specific analysis. It is not usual for the elimination of DOC of adsorbable chemicals to remain high throughout the test; normally, there is a high degree removal initially which gradually falls to an equilibrium value. If, however, the adsorbable test chemical was able to cause acclimation of the microbial population in some way or other, the DOC elimination of the test chemical would subsequently increase and reach a high plateau value.

**▼ M4***Lag phase*

58. As in static, screening tests, many test chemicals require a lag phase before full biodegradation occurs. In the lag phase, acclimation or adaptation of the degrading bacteria takes place with almost no removal of the test chemical; then the initial growth of these bacteria occurs. This phase ends and the degradation phase is taken to begin when about 10 % of the initial amount of test chemical is removed (after allowing for adsorption, if it occurs). The lag phase is often highly variable and poorly reproducible.

*Plateau phase*

59. The plateau phase of an elimination curve in a continuous test is defined as that phase in which the maximum degradation takes place. The plateau phase should be at least 3 weeks and have about 15 measured valid values.

*Mean degree of elimination of test chemical*

60. Calculate the mean value from the elimination values ( $D_t$ ) of the test chemical at the plateau phase. Rounded to the nearest whole number (1 %), it is the degree of elimination of the test chemical. It is also recommended to calculate the 95 % confidence interval of the mean value.

*Elimination of organic medium*

61. Plot the percentage of elimination of the DOC or COD of the organic medium in the control unit ( $D_B$ ) versus time. Indicate the mean degree of elimination in the same way as for the test chemical (paragraph 60).

*Indication of biodegradation*

62. If the test chemical does not adsorb significantly on to activated sludge and the elimination curve has a typical shape of a biodegradation curve with lag, degradation and plateau phases (paragraphs 58, 59), the measured elimination can safely be attributed to biodegradation. If a high initial removal has taken place, the simulation test cannot differentiate between biological and abiotic elimination processes. In such cases, and in other cases where there is any doubt about biodegradation (e.g. if stripping takes place), analyse adsorbed test chemicals or perform additional static biodegradation tests based on parameters clearly indicating biological processes. Such tests are the oxygen uptake methods (chapter C.4 D, E and F of this Annex (6)) or a test with measurement of carbon dioxide production (chapter C.4 C of this Annex (6)) or the ISO Headspace method (18), using a pre-exposed inoculum from the simulation test. If both the DOC removal and specific chemical removal have been measured, significant differences (the former being lower than the latter) between the percentages removed indicate the presence in the effluents of intermediate organic products which may be more difficult to degrade than the parent chemical.

*Validity of test results*

63. Information on the normal biodegradation behaviour of the inoculum is achieved if the degree of elimination of the organic medium (paragraph 53) in the control unit is determined. Consider the test to be valid if the degree of DOC or COD elimination in the control unit(s) is > 80 % after two weeks and no unusual observations have been made.

**▼ M4**

64. If a readily biodegradable (reference) chemical has been used, the degree of biodegradation ( $D_t$ , paragraph 52) should be  $> 90 \%$ .
65. If the test is performed under nitrifying conditions, the mean concentration in the effluents should be  $< 1 \text{ mg/l}$  ammonia-N and  $< 2 \text{ mg/l}$  nitrite-N.
66. If these criteria (paragraphs 63-65) are not met, repeat the test using an inoculum from a different source, test a reference chemical, and review all experimental procedures.

**Test Report**

67. The test report must include the following:

*Test chemical:*

- identification data;
- physical nature and, where relevant, physical-chemical properties.

*Test conditions:*

- type of test system; any modifications for testing insoluble and volatile chemicals;
- type of organic medium;
- proportion and nature of industrial waste waters in sewage, if known;
- inoculum, nature and sampling site(s), concentration and any pre-treatment;
- test chemical stock solution: DOC and TOC content; how prepared, if suspension; test concentration used; reasons if outside range of 10-20 mg/l DOC; method of addition; date first added; any changes;
- mean sludge age and mean hydraulic retention time; method of sludge wastage; methods of overcoming bulking, loss of sludge, etc.;
- analytical techniques employed;
- test temperature;
- qualities of the sludge-bulking, sludge volume index (SVI), mixed liquor suspended solids (MLSS);
- any deviations from standard procedures and any circumstances which may have affected results.

*Test results:*

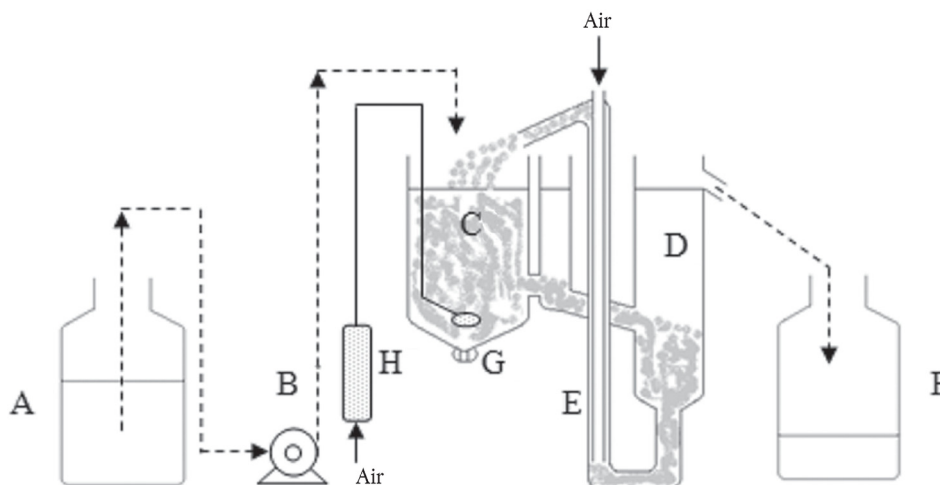
- all measured data (DOC, COD, specific analyses, pH, temperature, oxygen concentration, suspended solids, N chemicals, if relevant);
- all calculated values of  $D_t$  (or  $D_{tc}$ ),  $D_B$ ,  $D_{St}$  obtained in tabular form and the elimination curves;
- information on lag and plateau phases, test duration, the degree of elimination of the test chemical and that of the organic medium in the control unit, together with statistical information and statements of biodegradability and validity of the test;
- discussion of results.

**▼ M4***LITERATURE:*

- (1) Swisher RD (1987). 'Surfactant Biodegradation', 2nd Edn. Marcel Dekker Inc. New York, 1085 pp.
- (2) German Government (1962). Ordinance of the degradability of detergents in washing and cleaning agents. Bundesgesetzblatt, Pt.1 No 49: 698-706.
- (3) Painter HA and King EF (1978a). WRc porous-pot method for assessing biodegradability. Technical Report No 70, Water Research Centre, Medmenham, UK.
- (4) Painter HA and King EF (1978b). The effect of phosphate and temperature on growth of activated sludge and on biodegradation of surfactants. *Wat. Res.* 12: 909-915.
- (5) Eckenfelder, W.W (19) US EPA.
- (6) Chapter C.4 of this Annex, Determination of 'Ready' Biodegradability.
- (7) Chapter C.12 of this Annex, Biodegradation — Modified SCAS Test.
- (8) Chapter C.19 of this Annex, Estimation of the Adsorption Coefficient ( $K_{OC}$ ) on Soil and on Sewage Sludge Using High Performance Liquid Chromatography (HPLC).
- (9) Gerike P and Fischer WK (1979). A correlation study of biodegradability determinations with various chemicals in various tests. *Ecotox. Env. Saf.* 3:157-173.
- (10) Gerike P and Fischer WK (1981), as (9), II Additional results and conclusions. *Ecotox. Env. Saf.* 5: 45-55.
- (11) Painter HA and Bealing D (1989). Experience and data from the OECD activated sludge simulation test. pp. 113-138, In: Laboratory tests for simulation of water treatment processes. CEC Water Pollution Report 18. Eds. Jacobsen BN, Muntau H, Angeletti G.
- (12) ISO 11733 (1995; revised 2004). Evaluation of the elimination and biodegradability of organic substances in an aqueous medium — activated sludge simulation test.
- (13) Birch RR (1982). The biodegradability of alcohol ethoxylates. XIII Jornado Com. Espanol. Deterg.: 33-48.
- (14) Birch RR (1984). Biodegradation of nonionic surfactants. *J.A.O.C.S.* 61 (2): 340-343.
- (15) Gerike P, Fischer WK and Holtmann W (1980). Biodegradability determinations in trickling filter units compared with the OECD confirmatory test. *Wat.Res.* 14: 753-758.
- (16) Baumann U, Kuhn G and Benz M. (1998). Einfache Versuchsanordnung zur Gewinnung gewässerökologisch relevanter Daten, UWSF — *Z. Umweltchem. Ökotox.* 10: 214-220.
- (17) Her Majesty's Stationery Office (1982). Assessment of biodegradability. Methods for the examination of waters and associated materials. pp. 91-98 ISBN 0117516619.
- (18) ISO 14593 (1998). Water Quality — Evaluation in an aqueous medium of the ultimate biodegradability of organic compounds. Method by the analysis of inorganic carbon in sealed vessels.

▼ **M4***Appendix 1**Figure 1***Equipment used for assessment of biodegradability**

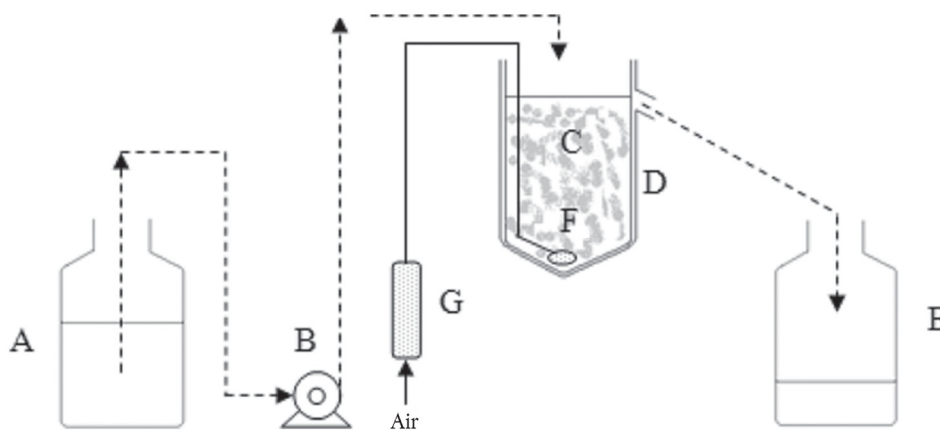
Husmann unit



- |                                   |                      |
|-----------------------------------|----------------------|
| A. Storage vessel                 | E. Air lift pump     |
| B. Dosing pump                    | F. Collection vessel |
| C. Aeration chamber (3l capacity) | G. Aerator           |
| D. Settling vessel                | H. Air flow meter    |

*Figure 2***Equipment used for assessment of biodegradability**

Porous pot

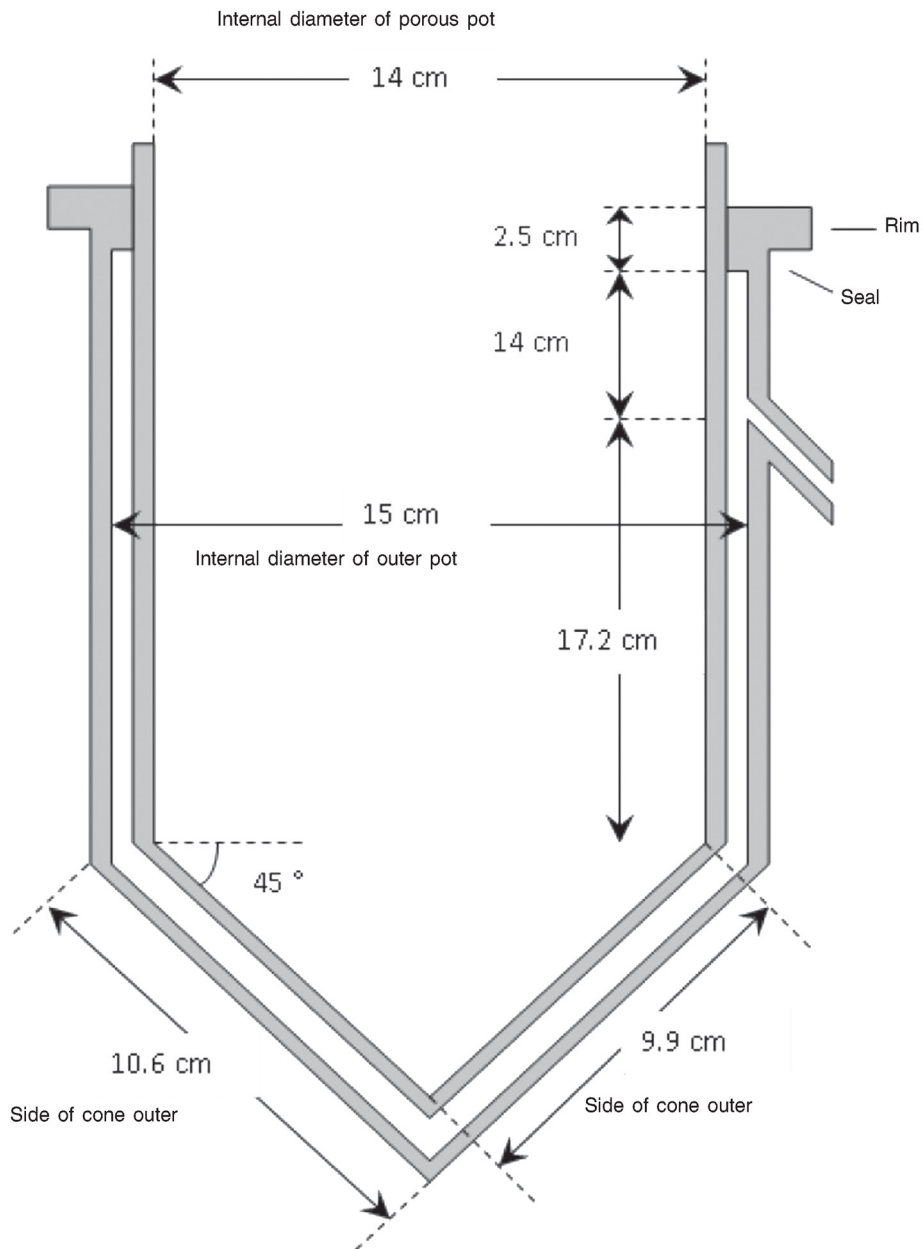


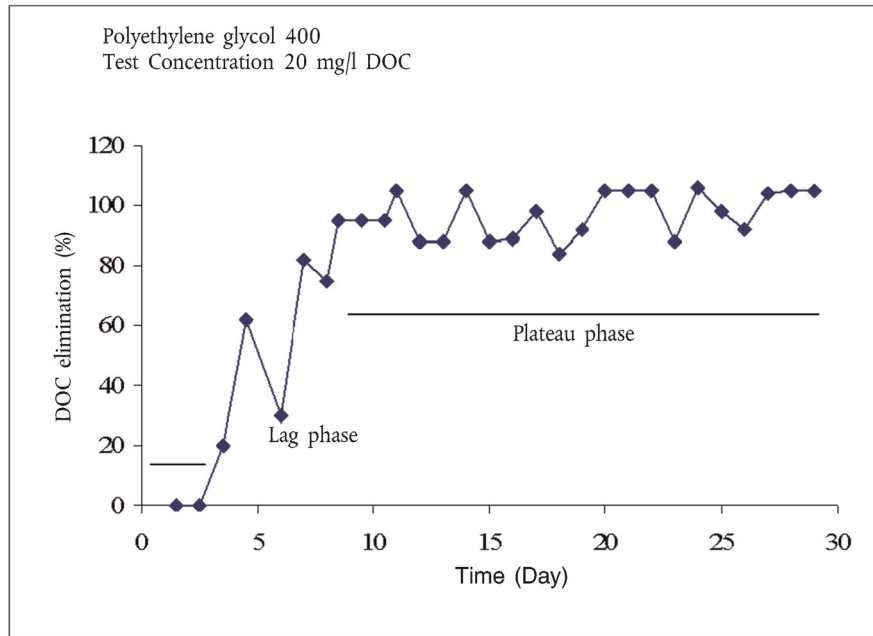
- |                             |                      |
|-----------------------------|----------------------|
| A. Storage vessel           | E. Collection vessel |
| B. Dosing pump              | F. Diffuser          |
| C. Porous aeration vessel   | G. Air flow meter    |
| D. Outer impermeable vessel |                      |



▼ **M4**

Figure 3

**Details of 3 litre porous pot aeration vessel**

▼ M4*Appendix 2***Example of an elimination curve**

▼ **M4***Appendix 3*

[INFORMATIVE]

## COUPLING OF THE TEST UNITS

In order to try to equalise the microbial populations in sludges in a test unit, receiving sewage plus a test chemical, and in a control unit, receiving only sewage, a daily interchange of sludge was introduced (1). The procedure was called coupling and the method is known as coupled units. Coupling was initially performed using Husmann activated sludge units but it has also been done with Porous Pot units (2)(3). No significant differences in results were found as between non-coupled and coupled units, whether Husmann or Porous Pot so there is no advantage in expending the time and energy needed in coupling the units.

Sludge exchanges can give the appearance of quite a considerable removal, since some of the test chemical is transferred and the concentrations of test chemical in the test and control effluents become more nearly equal. Thus, correcting factors have to be used, which depend on the fraction exchanged and the mean hydraulic retention time. More details of the calculation have been published (1).

Calculate the corrected DOC or COD elimination degree using the general formula:

$$D_{tc} = (D_t - 100 \cdot a \cdot r/12)/(1 - a \cdot r/12) \%$$

where

$D_{tc}$  = corrected % DOC or COD elimination

$D_t$  = determined % DOC or COD elimination

$a$  = interchange fraction of the volume of the activated sludge units

$r$  = mean hydraulic retention time (h)

If, for example, half of the volume of the aeration tank is exchanged ( $a = 0,5$ ) and the mean hydraulic retention time is 6 h, the correction formula is:

$$D_{tc} = \frac{4D_t - 100}{3}$$

*LITERATURE:*

- (1) Fischer W, Gerike P, Holtmann W (1975). Biodegradability Determinations via Unspecific Analyses (Chemical Oxygen Demand, DOC) in Coupled Units of the OECD Confirmatory Test. I The test. *Wat. Res.* 9: 1131-1135.
- (2) Painter HA, Bealing DJ (1989). Experience and Data from the OECD Activated Sludge Simulation Test. pp. 113-138. In: *Laboratory Tests for Simulation of Water Treatment Processes* CEC Water Pollution Report 18. Eds. Jacobsen BN, Muntau H, Angeletti G.
- (3) Painter HA, King EF (1978). Water Research Centre Porous Pot Method for Assessing Biodegradability. Technical Report TR70, Water Research Centre, Stevenage, UK.

▼ **M4***Appendix 4*

## EVALUATION OF INHIBITION OF THE ACTIVATED SLUDGE

**Process by test chemicals**

1. A chemical (or a waste water) may not be degraded or removed in the simulation test and may even have an inhibitory effect on the sludge micro-organisms. Other chemicals are biodegraded at low concentrations but are inhibitory at higher concentration (hormesis). Inhibitory effects may have been revealed at an earlier stage or may be determined by applying a toxicity test, using an inoculum similar to or identical with that used in the simulation test (1). Such methods are inhibition of oxygen uptake (chapter C.11 of this Annex (2) and ISO 8192(3)) or inhibition of growth of sludge organisms (ISO 15522 (4)).
2. In the simulation test any inhibition will be manifest by the difference in dissolved organic carbon (DOC) or chemical oxygen demand COD between the effluent from the test vessel and that from the control being greater than the DOC added as test chemical. Expressed in another way, the percentage removal of DOC (and biochemical oxygen demand BOD, chemical oxygen demand COD, and/or  $\text{NH}_4^+$ ) of the organic medium on treatment will be decreased by the presence of the test chemical. If this occurs, the test should be repeated reducing the concentration of the test chemical until a level is reached at which no inhibition occurs and perhaps further reducing the concentration until the test chemical is biodegraded. However, if the test chemical (or waste water) has adverse effects on the process at all concentrations tested, the indications are that the chemical is difficult, if not impossible, to treat biologically, but it may be worth repeating the test with activated sludge from a different source and/or subjecting the sludge to a more gradual acclimation.
3. Conversely, if the test chemical is bioeliminated at the first attempt in the simulation test, its concentration should be increased if it is required to be known whether the chemical could be inhibitory.
4. It should be remembered in trying to determine degrees of inhibition that the activated sludge population can change, so that with time the micro-organisms may develop a tolerance towards an inhibitory chemical.
5. Calculation of degree of inhibition:

The overall percentage removals  $R_o$ , of BOD, DOC, COD etc., for the test and control units can be calculated from:

$$R_o = 100 (I - E)/I \%$$

where:

I = influent concentration of BOD, DOC, COD etc., for test or control vessels (mg/l)

E = respective effluent concentrations (mg/l).

I and E must be corrected for the DOC due to the test chemical in the test units, otherwise the calculations of percentage inhibition will be incorrect.

**▼M4**

The degree of inhibition caused by the presence of the test chemical can be calculated from:

$$\% \text{ inhibition} = 100 (R_c - R_t)/R_c$$

where:

$R_c$  = percentage removal in the control vessels

$R_t$  = percentage removal in the test vessels

*LITERATURE:*

- (1) Reynolds L et al. (1987). Evaluation of the toxicity of substances to be assessed for biodegradability. *Chemosphere* 16: 2259.
- (2) Chapter C.11 of this Annex, Biodegradation — Activated Sludge Respiration Inhibition Test.
- (3) ISO 8192 (2007) Water quality — Test for inhibition of oxygen consumption by activated sludge for carbonaceous and ammonium oxidation.
- (4) ISO 15522 (1999) Water Quality — Determination of the inhibitory effect of water constituents on activated sludge microorganisms.

**▼ M4***Appendix 5***Poorly water-soluble test chemicals — volatile chemicals****Poorly water-soluble chemicals**

Few reports seem to have been published on subjecting poorly water-soluble and insoluble chemicals to tests simulating waste water treatment (1)(2)(3).

There is no single method of dispersal of the test chemical which is applicable to all insoluble chemicals. Two of the four types of method described in ISO 10634 (4) would seem to be suitable for attempting to disperse test chemicals for simulation testing; they are the use of emulsifying agents and/or of ultrasonic energy. The stability over at least 24h periods of the resulting dispersion should be established. Suitably stabilised dispersions, contained in a constantly stirred reservoir (paragraph 38), would then be dosed to the aeration tank separately from the domestic (or synthetic) sewage.

If the dispersions are stable, investigate how the test chemical can be determined in the dispersed form. It is unlikely that DOC will be suitable, so that a specific analytical method for the test chemical would have to be established which could be applied to effluents, effluent solids and activated sludge. The fate of the test chemical in the simulation of the activated sludge process would then be determined in liquid and solid phases. Thus, a 'mass balance' would be established to decide whether the test chemical had been biodegraded. However, this would indicate only primary biodegradation. Demonstration of ultimate biodegradation should be attempted by applying a respirometric test for ready biodegradability (chapter C.4 of this Annex (5) C, F or D) using as inoculum sludge exposed to the test chemical in the simulation test.

**Volatile chemicals**

The application of waste water treatment simulations to volatile chemicals is both debatable and problematic. As with poorly water-soluble test chemicals, very few reports seem to have been published describing simulation tests using volatile chemicals. A conventional type of complete-mixing apparatus is adapted by sealing the aeration and settling tanks, measuring and controlling the air flow using flow-meters and passing the exit gas through traps to collect volatile organic matter. In some cases, a vacuum pump is used to draw the exit gas through a 'cold' trap or a purge-trap containing Tenax and silica gel for gas-chromatographic analyses. The test chemical present in the trap can be determined analytically.

The test is carried out in two parts. The units are first operated without sludge but with the synthetic waste water plus test chemical being pumped into the aeration tank. Influent, effluent and exit gas samples are collected and analysed for the test chemical for a few days. From the data collected, the percentage ( $R_{vs}$ ) of the test chemical stripped from the system may be calculated.

Then the normal biological test (with sludge) is performed under operating conditions identical to those in the stripping study. DOC or COD measurements are also made to check that the units are performing efficiently. Occasional analyses are made to determine the test chemical in the influent, effluent and exit gas in the first part of the test; after acclimation more frequent analyses are made. Again, from the data in the steady state, the percentage of removal of the test chemical from the liquid phase by all processes ( $R_T$ ) (physical and biological) may be calculated, as well as the proportion ( $R_V$ ) stripped from the system.

**▼ M4**

Calculation:

- (a) In the non-biological test, the percentage ( $R_{VP}$ ) of the test material stripped from the system may be calculated from:

$$R_{VP} = \frac{S_{VP}}{S_{IP}} \cdot 100$$

where

$R_{VP}$  = removal of test chemical by volatilisation (%),

$S_{VP}$  = test chemical collected in trap expressed as equivalent concentration in liquid phase (mg/l),

$S_{IP}$  = test chemical concentration in influent (mg/l).

- (b) In the biological test, the percentage ( $R_V$ ) of the test material stripped from the system may be calculated from:

$$R_V = \frac{S_V}{S_I} \cdot 100$$

where

$R_V$  = removal of test chemical by volatilisation in biological test (%),

$S_V$  = test chemical collected in trap in biological test, expressed as equivalent concentration in liquid influent (mg/l),

$S_I$  = test chemical concentration in influent (mg/l).

- (c) In the biological test, the percentage ( $R_T$ ) of the test chemical removed by all processes is given by:

$$R_T = 1 - \frac{S_E}{S_I} \cdot 100$$

where

$S_E$  = concentration of test chemical in the (liquid) effluent (mg/l).

- (d) Thus, the percentage ( $R_{BA}$ ) removed by biodegradation plus adsorption can be calculated from:

$$R_{BA} = (R_T - R_V)$$

Separate tests should be carried out to determine whether the test chemical is adsorbed; if it is, then a further correction may be made.

- (e) A comparison between the proportion of test chemical stripped from the biological ( $R_V$ ) and non-biological test ( $R_{VP}$ ) systems indicates the overall effect that biological treatment has had on the emission of the test chemical into the atmosphere.

*Example:* Benzene

Sludge retention time = 4 days

A synthetic sewage; retention time = 8 h.

$S_{IP} = S_I = 150$  mg/l

$S_{VP} = 150$  mg/l ( $S_{EP} = 0$ )

$S_V = 22,5$  mg/l

$S_E = 50$  µg/l

**▼ M4**

Thus,

$$R_{VP} = 100, R_V = 15$$

$$R_T = 100 \text{ and } R_{BA} = 85.$$

Benzene was assumed not to be adsorbed onto sludge.

*LITERATURE:*

- (1) Horn JA, Moyer JE, Hale JH (1970). Biological degradation of tertiary butyl alcohol. Proc. 25th Ind. Wastes Conference Purdue Univ.: 939-854.
- (2) Pitter P, Chudoba J (1990). Biodegradability of organic substances in the aquatic environment. CRC Press. Boston, USA.
- (3) Stover EL, Kincannon DF (1983). Biological treatability of specific organic compounds found in chemical industry waste waters. J. Wat. Pollut. Control Fed. 55: 97.
- (4) ISO 10634 (1995) Water Quality — Guidance for the preparation and treatment of poorly water-soluble organic compounds for the subsequent evaluation of their biodegradability in an aqueous medium.
- (5) Chapter C.4 of this Annex, Determination of 'Ready' Biodegradability.



**▼ M4***Appendix 6***Effects of sludge retention time (SRT) on treatability of chemicals**

## INTRODUCTION

1. The method described in the main text was designed to ascertain whether the chemicals tested (usually those known to be inherently, but not readily, biodegradable) can be biodegraded within the limits imposed in waste water treatment plants. The results are expressed in terms of percentage removal and percentage biodegradation. The conditions of operation of the activated sludge units and choice of influent allow rather wide variations in concentration of the test chemical in the effluent. Tests are carried out at only one nominal concentration of sludge solids or one nominal sludge retention time (SRT) and the sludge wastage regimes described can cause the value of SRT to vary considerably during the test, both from day to day and during a day.
2. In this variant (1)(2) the SRT is controlled within much narrower limits throughout each 24h period (just as happens on the large-scale) which results in a more constant concentration in effluents. Domestic sewage is recommended since it gives more consistent and higher percentage removals. Also, the effects of a number of SRT values are investigated and in a more detailed study the effects of a range of temperatures on effluent concentration may be determined.
3. There is no general agreement yet on which kinetic models operate when chemicals bio-degrade under conditions in waste water treatment. The Monod model of bacterial growth and substrate utilisation was chosen (1)(2) to be applied to the data collected, since the method was intended to be applied only to chemicals produced in high tonnages, resulting in concentrations in sewage of above 1 mg/l. The validity of the simplified model and the assumptions made was established using a series of alcohol ethoxylates having varying degrees of primary biodegradability (2)(3).

*Note:* This variant method follows closely much of the text of this test method C.10-A and only those details which differ are given hereafter.

## PRINCIPLE OF THE TEST

4. Activated sludge porous-pot units, designed to facilitate the (almost) continuous wastage of mixed liquor allowing very precise control of the sludge retention time (SRT, or  $\theta_s$ ), are operated in the non-coupled mode over a range of SRTs and, optionally, over a range of temperatures. The retention time is usually 2 to 10 days and the temperature between 5 and 20 °C. Sewage, preferably domestic, and a solution of the test chemical are dosed separately to the units at rates to give the required sewage retention time (3 to 6 hours) and the required concentration of test chemical in the influent. Control units receiving no test chemical are operated in parallel for comparative purposes.
5. Other types of apparatus can be used but great care should be exercised to ensure that good control of SRT is achieved. For example, when using plants, which incorporate a settler, allowance for loss of solids via the plant effluent may be necessary. Further, special precautions to avoid errors due to variation in the quantity of sludge in the settler should also be taken.

**▼M4**

6. The units are operated at each selected set of conditions and, after equilibrium has been reached, the average steady state concentrations in the effluents of test chemical and, optionally, DOC are obtained over a period of about three weeks. Besides assessing the percentage removal of test chemical and, optionally, DOC, the relationship between plant-operating conditions and the concentration in the effluent is expressed in graphical form. From this tentative kinetic constants may be calculated and the conditions under which the test chemical can be treated may be predicted.

## INFORMATION ON THE TEST CHEMICAL

7. Chapter C.10 A, paragraphs 12 and 13 apply.

## PASS LEVELS

8. Chapter C.10 A, paragraphs 14 and 15 apply.

## REFERENCE TEST CHEMICAL

9. Chapter C.10 A, paragraph 16 applies.

## REPRODUCIBILITY OF TEST RESULTS

10. Chapter C.10 A, paragraphs 17 and 18 apply.

## DESCRIPTION OF THE METHOD

**Apparatus**

11. A suitable unit is the modified porous pot system (Appendix 6.1). It consists of an inner vessel (or liner) constructed from porous polypropylene of 3,2 mm thickness and pore size of approximately 90 µm, the joint being butt-welded. (This makes a more robust unit than that described in paragraph 21 of this chapter, C.10 A). The liner is fitted into an impervious polyethylene outer vessel, which consists of two parts: a circular base in which holes are bored to accommodate two air lines and a sludge-wastage line, and an upper cylinder which screws on to the base and which has an outlet placed so as to give a known volume (3 l) in the porous pot vessel. One of the air lines is fitted with a diffuser stone and the other is open-ended and set at right-angles to the stone in the pot. This system produces the necessary turbulence to ensure that the contents of the pot are completely mixed, as well as providing concentrations of dissolved oxygen greater than 2 mg/l.
12. The appropriate number of units are maintained at controlled temperatures in the range of 5 to 20 °C ( $\pm 1$  °C), either in water baths or in constant temperature rooms. Pumps are required to dose to the aeration vessels the solution of the test chemical and settled sewage at the required rates (0-1,0 ml/min and 0-25 ml/min, respectively) and a third pump to remove waste sludge from the aeration vessels. The necessary very low flow-rate of waste sludge is achieved by using a pump set at a higher rate and operated intermittently by the use of a timer-switch, e.g. operating for 10 seconds per min, pump delivery rate of 3ml/min yielding a wastage rate of 0,5 ml/min.

*Filtration apparatus or centrifuge*

13. Chapter C10 A, paragraph 23 applies.

*Analytical equipment*

14. Chapter C.10 A, paragraph 24 applies.

*Water*

**▼ M4**

15. Chapter C.10 A, paragraphs 25 and 26 apply.

*Organic medium*

16. Chapter C.10 A, paragraph 27 applies.

*Synthetic sewage*

17. Chapter C.10 A, paragraph 28 applies.

*Domestic sewage*

18. Chapter C.10 A, paragraph 29 applies.

*Activated sludge*

19. Chapter C10 A, paragraph 30 applies.

*Stock solutions of test chemical*

20. Chapter C.10 A, paragraphs 31 and 32 apply.

## PROCEDURE

*Preparation of the inoculum*

21. Chapter C.10 A, paragraph 34 applies only — use activated sludge (about 2,5 g/l).

*Number of test units*

22. For a simple test, ie. to measure percentage removal, only a single SRT is required, but in order to acquire data necessary to calculate tentative kinetic constants 4 or 5 SRT values are required. Values between 2 and 10 days are usually chosen. Practically, it is convenient to perform a test at 4 or 5 SRTs simultaneously at one temperature; in extended studies the same SRT values, or perhaps a different range of values, are used at other temperatures within the range 5 to 20 °C. For primary biodegradation (the main use), only one unit per set of conditions is normally required. However, for ultimate biodegradability a control unit is required, for each set of conditions, which receives sewage but not test chemical. If the test chemical is thought to be present in the sewage used, it would be necessary to use control units when assessing primary biodegradation, and making the necessary correction in the calculations.

*Dosage of organic medium and test chemical*

23. Chapter C.10 A, paragraphs 36 to 39 apply, but note that the test chemical solution is dosed separately and that various sludge wastage rates are used. Also monitor and adjust, if necessary, to within  $\pm 10\%$ , the flow-rates of influents, effluents and sludge wastage frequently, e.g. twice per day. If difficulties are encountered in the analytical methods when domestic sewage is used, carry out the test with synthetic sewage, but it must be assured that different media give comparable kinetic data.

*Handling of activated sludge units*

24. Chapter C.10 A, paragraphs 40 to 43 apply, but control SRT only by 'constant' wastage of sludge.

*Sampling and analysis*

25. Chapter C.10 A, paragraphs 44 to 50 apply, except that the concentration of the test chemical is to be determined and DOC determined optionally; COD should not be used.

**▼M4**

## DATA AND REPORTING

**Treatment of results**

26. Chapter C.10 A, paragraphs 52 to 54 apply.

**Expression of test results**

27. Chapter C.10 A, paragraphs 56 to 62 apply.

**Calculation of kinetic constants**

28. It is more realistic to quote the mean steady — state concentration of the test chemical in the effluent and to describe how this varies with plant-operating conditions than to quote percentage primary biodegradation. This can be done by consideration of equation (6) in Appendix 6.2, which can yield values for  $K_S$ ,  $\mu_m$  and  $\theta_{SC}$ , the critical sludge retention time.

(Alternatively, approximate values of  $K_S$  and  $\mu_m$  may be obtained using a simple computer program to fit the theoretical curve calculated from equation 2 (Appendix 6.2) to the experimental values obtained. Although any given solution will not be unique, a reasonable approximation of  $K_S$  and  $\mu_m$  can be obtained.)

**Variability of results**

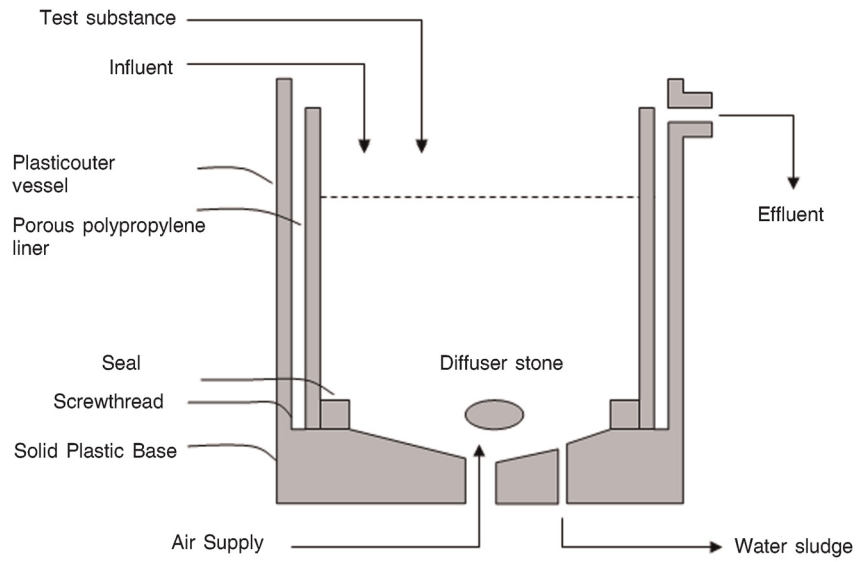
29. It is common experience that variable values of kinetic parameters for individual chemicals are obtained. It is thought that the conditions under which the sludge has been grown, as well as the conditions prevailing in the test used (as in paragraph 5 and in other tests), have a large effect on the resulting values. One aspect of this variability has been discussed by Grady et al (4), who have suggested that the terms 'extant' and 'intrinsic' should be applied to two extreme conditions representing the limits of physiological state a culture may attain during a kinetic experiment. If the state is not allowed to change during the test, the kinetic parameter values reflect the conditions in the environment from which the micro-organisms were taken; these values are called 'extant' or currently existing. At the other extreme, if conditions in the test are such as to permit the full development of the protein-synthesizing system allowing maximum possible growth rate, the kinetic parameters obtained are said to be 'intrinsic', and are dependent only on the nature of the substrate and the types of bacteria in the culture. As a guide, extant values will be obtained by keeping the ratio of concentration of substrate to competent micro-organisms ( $S_0/X_0$ ) low, e.g. 0,025, and intrinsic values arise when the ratio is high e.g. at least 20. In both cases,  $S_0$  should equal or exceed the relevant value of  $K_S$ , the half-saturation constant.
30. Variability and other facets of biodegradation kinetics were discussed at a recent SETAC workshop (5). From such studies, reported and projected, a clearer view of kinetics operating in waste water treatment plants should be forth-coming to enable a better interpretation of existing data to be made, as well as to suggest more relevant designs for future Test Methods.

*LITERATURE:*

- (1) Birch RR (1982). The biodegradability of alcohol ethoxylates. XIII Jornado Com. Espanol Deterg.: 33-48.
- (2) Birch RR (1984). Biodegradation of nonionic surfactants. J.A.O.C.S., 61(2): 340-343.
- (3) Birch RR (1991). Prediction of the fate of detergent chemicals during sewage treatment. J. Chem. Tech. Biotechnol., 50: 411-422.

▼ **M4**

- (4) Grady CPL, Smets BF and Barbeau DS (1996). Variability in kinetic parameter estimates: A review of possible causes and a proposed terminology. *Wat. Res.*, 30 (3): 742-748.
- (5) Biodegradation kinetics: Generation and use of data for regulatory decision making (1997). Workshop at Port Sunlight, UK. Eds. Hales SG, Feitjel T, King H, Fox K, Verstraete W. 4-6th Sept. 1996. SETAC- Europe, Brussels.

▼ **M4***Appendix 6.1***Porous Pot with SRT Control**

▼ **M4***Appendix 6.2***Calculation of Kinetic Constants**

1. By assuming Monod kinetics apply and considering a mass balance of active solids and substrate across the activated sludge system (1), the following steady state expressions can be obtained:

$$\frac{1}{\theta_s} = \frac{\mu_m \cdot S_1}{K_s + S_1} - K_d \quad [1]$$

or

$$S_1 = \frac{K_s \cdot (1 + K_d \cdot \theta_s)}{\theta_s \cdot (\mu_m - K_d) - 1} \quad [2]$$

where:

$S_1$  = concentration of substrate in effluent, (mg/l)

$K_S$  = half-saturation constant, the concentration at which  $\mu = \mu_m/2$  (mg/l)

$\mu$  = specific growth rate ( $d^{-1}$ )

$\mu_m$  = maximum value of  $\mu_m$  ( $d^{-1}$ )

$K_d$  = specific decay rate of active solids ( $d^{-1}$ )

$\theta_S$  = sludge mean retention time, SRT (d)

Examination of this equation leads to the following conclusions:

- (i) The effluent concentration is independent of that in the influent ( $S_0$ ); hence, the percentage biodegradation varies with the influent concentration,  $S_0$ .
- (ii) The only plant-control parameter affecting  $S_1$  is the sludge retention time,  $\theta_S$ .
- (iii) For a given concentration in the influent,  $S_0$ , there will be a critical sludge retention time, such that:

$$\frac{1}{\theta_{SC}} = \frac{\mu_s \cdot S_0}{K_s + S_0} - K_d \quad [3]$$

where:

$\theta_{SC}$  = critical sludge retention time, below which the competent micro-organisms will be washed out of the plant.

- (iv) Since the other parameters in equation (2) are associated with growth kinetics, temperature is likely to affect the effluent substrate level and the critical sludge age, ie. the sludge retention time needed to obtain a certain degree of treatment would increase with decreasing temperature.
2. From a mass balance of solids in the porous pot system, and assuming that the solids concentration in the plant effluent,  $X_2$  is low compared with that in the aeration vessel,  $X_1$ , the sludge retention time

$$\theta_s = \frac{V \cdot X_1}{(Q_0 - Q_1) \cdot X_2 + Q_1 \cdot X_1} \quad [4]$$

**▼ M4**

and

$$\theta_s = \frac{V \cdot X_1}{Q_1 \cdot X_1} = \frac{V}{Q_1}$$

where:

V = volume of the aeration vessel (l)

X<sub>1</sub> = concentration of solids in aeration vessel (mg/l)

X<sub>2</sub> = concentration of solids in effluent (mg/l)

Q<sub>0</sub> = flow rate of influent (l/d)

Q<sub>1</sub> = flow rate of waste sludge (l/d)

Thus, it is possible to control the sludge retention time at any pre-selected value by the control of the waste sludge flow rate, Q<sub>1</sub>.

Conclusions:

3. The main purpose of the test is thus to allow the effluent concentration, and hence the levels of test chemical in the receiving waters, to be predicted.
4. By plotting S<sub>1</sub>, vs. θ<sub>s</sub>, the critical sludge retention time, θ<sub>SC</sub>, can sometimes be readily evaluated, eg. curve 3 in Figure 1. When this is not possible, θ<sub>SC</sub> may be calculated, together with approximate values of μ<sub>m</sub> and K<sub>S</sub>, by plotting S<sub>1</sub>, vs. S<sub>1</sub>•θ<sub>s</sub>.

Rearrangement of equation (1) gives

$$\frac{S_1 \cdot \theta_s}{1 + \theta_s \cdot K_d} = \frac{K_s}{\mu_m} + \frac{S_1}{\mu_m} \quad [5]$$

If K<sub>d</sub> is small, then 1 + θ<sub>s</sub> · K<sub>d</sub> ~ 1 and [5] becomes:

$$S_1 \cdot \theta_s = \frac{K_s}{\mu_m} + \frac{S_1}{\mu_m} \quad [6]$$

Thus, the plot should be a straight line (see Figure 2) of slope 1/μ<sub>m</sub> and intercept K<sub>S</sub>/μ<sub>m</sub>; also θ<sub>S</sub> ~ 1/μ<sub>m</sub>.



▼ M4

Figure 1

Three temperatures; five SRTs

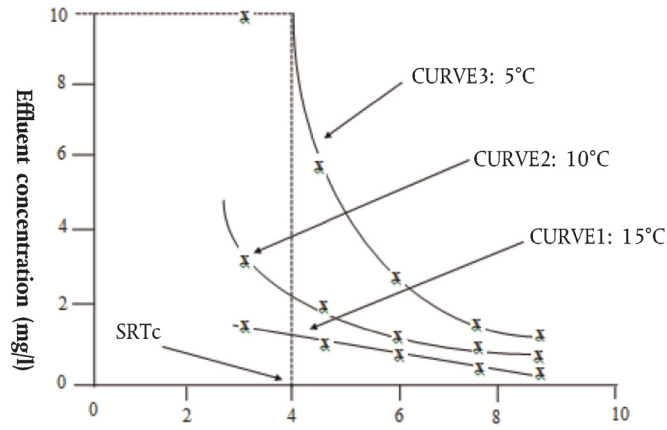
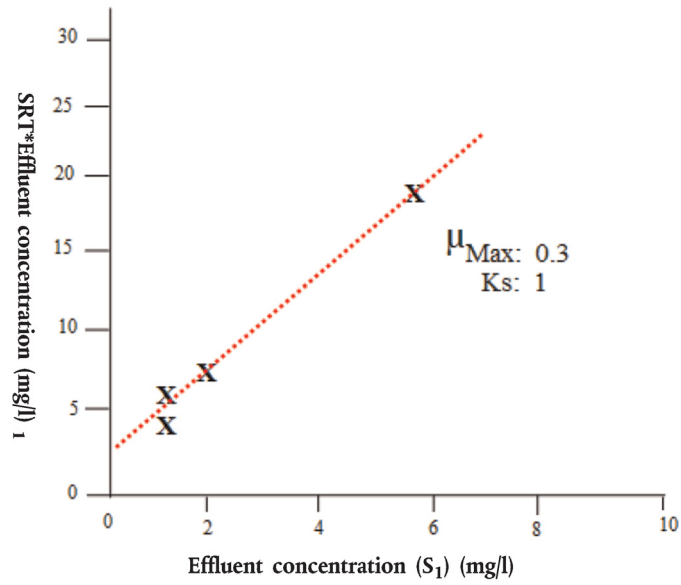


Figure 2

Regression Line  $SRT \cdot S_1$  vs  $S_1$  at  $T = 5^\circ\text{C}$



Glossary:

Effluent concentration:

Curve:

**▼ M4***Appendix 7*TEST AT LOW ( $\mu\text{g/l}$ ) CONCENTRATION RANGE

1. Many chemicals are normally present in the aquatic environment, even in waste waters, at very low concentrations ( $\mu\text{g/l}$ ). At such concentrations, they probably do not serve as primary substrates resulting in growth, but are more likely to degrade as non-growth, secondary substrates, concurrent with a variety of naturally occurring carbon chemicals. Consequently the degradations of such chemicals will not fit the model described in Appendix 6. There are many models which could be applied and, under the conditions prevailing in waste water treatment systems, more than one may be simultaneously operative. Far more research will be necessary to elucidate this problem.
2. Meanwhile the procedure given in the main text (chapter C.10 A) can be followed, but only for primary biodegradability, using suitably low concentrations ( $< 100 \mu\text{g/l}$ ) and a validated analytical procedure. The percentage biodegradation may be calculated (see para. 54 of the Test Method) provided that abiotic processes (adsorption, volatility, etc.) are taken into account. An example is the study by Nyholm and his associates (1)(2) using a 4 h cycle in a fill and draw system. They reported pseudo first-order constants for 5 chemicals added in a synthetic sewage at 5 to  $100 \mu\text{g/l}$ . (For ultimate biodegradability  $^{14}\text{C}$ -labelled test chemicals may be used. A description of this is beyond the scope of this Test Method since there are as yet no agreed procedures, though a proposed method for ISO 14592 (3) contains guidance on the use of  $^{14}\text{C}$ -labelled chemicals.

**SCAS test**

3. Later, a simpler two-stage test was proposed (4)(5)(6); the semi-continuous activated sludge (SCAS) method is followed by short-term kinetic tests on samples withdrawn from the SCAS units. The SCAS system is operated with known sludge wastage rates (unlike the original C.12 test method) and is fed a modified OECD synthetic sewage or domestic sewage. The synthetic sewage was modified (because of changing pH value and poor sludge settleability) by addition of phosphate as buffer, yeast extract, iron (III) chloride and trace element salts, and its COD was increased to about  $750 \text{ mg/l}$  by increasing the concentration of peptone and meat extract. The units were operated on a 24 h cycle: aeration for 23 h, wastage of sludge, settlement, withdrawal of supernatant (effluent) followed by addition of synthetic sewage plus test chemical, up to  $100 \mu\text{g/l}$ , (i.e. at about the same concentration used in the short term test). Once per week 10 % of the total sludge was replaced by fresh sludge in order to maintain a balanced microbial population.
4. The concentrations of test chemical initially and at the end of aeration are measured and the test is continued until a constant removal of test chemical is attained; this takes from one week to several months.

**Short-term test**

5. A short test (e.g. 8 hours) is applied to determine the (pseudo) first order kinetic rate constant for the decay of the test chemical in activated sludge of known but different origins and histories. In particular, sludge samples are taken from the SCAS reactors — at the end of an aeration period when the concentration of organic substrate is low — during the course of an acclimatisation experiment (paragraphs 3, 4). Sludge may also be taken from a parallel SCAS unit not exposed to the test chemical, for comparison.

**▼ M4**

Mixtures of sludge and the test chemical added at two or more concentrations in the range 1-50 µg/l are aerated, without the addition of synthetic sewage or other organic substrate. The test chemical remaining in solution is determined at regular intervals e.g. hourly depending on the degradability of the chemical, for a period not longer than 24h. Samples are centrifuged before appropriate analysis.

**Calculations**

6. Data from the SCAS units are used to calculate the percentage removal of test chemical (paragraph 54). Also, an average rate constant,  $K_1$ , (normalised for concentration of suspended solids) can be calculated from:

$$K_1 = 1/t \cdot \ln \frac{C_e}{C_i} \cdot 1/SS(1/gh)$$

where:

$t$  = aeration time (23h)

$C_e$  = concentration at end of aeration period (µg/l)

$C_i$  = concentration at beginning of aeration (µg/l)

SS = concentration of activated sludge solids (g/l)

7. In the short term test the log % concentration remaining is plotted against time and the slope of the initial part (10-50 % degradation) of the plot is equivalent to  $K_1$ , the (pseudo) first order constant. The constant is normalised with respect to the concentration of sludge solids by dividing the slope by the concentration of sludge solids. The reported result must also include details of initial concentrations of the test chemical and suspended solids, sludge retention time, sludge loading and source, and details of pre-exposure (if any) to the test chemical.

**Variability of results**

8. Variability and other facets of biodegradation kinetics were discussed at a recent SETAC workshop (7). From such studies, reported and projected, a clearer view of kinetics operating in waste water treatment plants should be forth-coming to enable a better interpretation of existing data to be made, as well as to suggest more relevant designs for future Test Methods.

*LITERATURE:*

- (1) Nyholm N, Jacobsen BN, Pedersen BM, Poulsen O, Dambourg A and Schultz B (1992). Removal of micropollutants in laboratory activated sludge reactors. Biodegradability. Wat. Res. 26: 339-353.
- (2) Jacobsen BN, Nyholm N, Pedersen BM, Poulsen O, and Ostfeldt P (1993). Removal of organic micropollutants in laboratory activated sludge reactors under various operating conditions: Sorption. Wat. Res. 27: 1505-1510.
- (3) ISO 14592 (ISO/TC 147/SC5/WG4, N264) (1998). Water Quality — Evaluation of the aerobic biodegradability of organic compounds at low concentrations in water.

**▼ M4**

- (4) Nyholm N, Ingerslev F, Berg UT, Pedersen JP and Frimer-Larsen H (1996). Estimation of kinetic rate constants for biodegradation of chemicals in activated sludge waste water treatment plants using short-term batch experiments and  $\mu\text{g/l}$  range spiked concentrations *Chemosphere* 33 (5): 851-864.
- (5) Berg UT and Nyholm N (1996). Biodegradability simulation Studies in semi-continuous activated sludge reactors with low ( $\mu\text{g/l}$  range) and standard (ppm range) chemical concentrations. *Chemosphere* 33 (4): 711-735.
- (6) Danish Environmental Protection Agency. (1996). Activated sludge biodegradability simulation test. Environmental Project, No 337. Nyholm, N. Berg, UT. Ingerslev, F. Min. of Env. and Energy, Copenhagen.
- (7) Biodegradation kinetics: Generation and use of data for regulatory decision making (1997). Workshop at Port Sunlight, UK. Eds. Hales, SG. Feitjel, T. King, H. Fox, K. and Verstraete, W. 4-6th Sept. 1996. SETAC- Europe, Brussels.

**▼ M4****C.10-B: Biofilms**

## INTRODUCTION

1. Simulation tests are normally applied to chemicals which have failed a screening test for ready biodegradability (Chapter C.4 A to F of this Annex (9)), but have passed a test for inherent biodegradability. Exceptionally simulation tests are also applied to any chemical about which more information is required, especially high-tonnage chemicals, and normally the activated sludge test is applied (C.10 A). In some circumstances, however, specific information is required relating the behaviour of a chemical to methods of waste water treatment involving biofilms, namely, percolating or trickling filters, rotating biological contactors, fluidised beds. To meet this need various devices have been developed.
2. Gerike *et al.* (1) used large, pilot-scale trickling filters which they used in the coupled mode. These filters took up much space and required relatively large volumes of sewage or synthetic sewage. Truesdale *et al.* (2) described smaller filters (6 ft × 6 in. diameter) which were fed surfactant-free natural sewage but still required rather large volumes. As many as 14 weeks were required for the development of a 'mature' biofilm and an additional 4-8 weeks were needed after first introduction of the test surfactant before acclimatisation took place.
3. Baumann *et al.* (3) developed a much smaller filter which used polyester 'fleece' previously steeped in activated sludge as the inert medium supporting the biofilm. The test chemical was used as the sole source of carbon and biodegradability was assessed from measurements of dissolved organic carbon (DOC) in the influent and effluent, and from the amount of CO<sub>2</sub> in the exit gas.
4. A quite different approach was made by Gloyna *et al.* (4) who invented the rotating tubular reactor. On the internal surface of the rotating tube a biofilm was grown on the known surface area by passage of influent introduced at the top end of the tube, inclined at a small angle to the horizontal. The reactor has been used to study the biodegradability of surfactants (5), as well as to investigate the optimal thickness of biofilm and diffusion through the film (6). These latter authors further developed the reactor, including modifying it to be able to determine CO<sub>2</sub> in the exit gas.
5. The rotating tubular reactor has been adopted by the Standing Committee of Analysts (UK) as a standard method for assessing both the biodegradability of chemicals (7) and the treatability and toxicity of waste waters (8). The method described here has the advantages of simplicity, compactness, reproducibility and the need for relatively small volumes of organic medium.

## PRINCIPLE OF THE TEST

6. Synthetic or domestic sewage, and the test chemical, in admixture or alone, are applied to the internal surface of a slowly rotating inclined tube. A layer of microorganisms, similar to those present on bio-filter media, is built up on the internal surface. The conditions of operation of the reactor are chosen to give adequate elimination of organic matter and, if required, oxidation of ammonium.

**▼M4**

7. Effluent from the tube is collected and either settled and/or filtered before analysis for dissolved organic carbon (DOC) and/or the test chemical by a specific method. Control units receiving no test chemical are operated in parallel under the same conditions for comparative purposes. The difference between the concentrations of DOC in the effluent from the test and control units is assumed to be due to the test chemical and its organic metabolites. This difference is compared with the concentration of the added test chemical (as DOC) to calculate the elimination of the test chemical.
8. Biodegradation may normally be distinguished from bio-adsorption by careful examination of the elimination-time curve. Confirmation may usually be obtained by applying a test for ready biodegradation (oxygen uptake or carbon dioxide production) using an acclimated inoculum taken at the end of the test from the reactors receiving the test chemical.

## INFORMATION ON THE TEST CHEMICAL

9. The purity, water solubility, volatile and adsorption characteristics of the test chemical should be known to enable correct interpretation of results to be made.
10. Normally, volatile and poorly soluble chemicals cannot be tested unless special precautions are taken (see Appendix 5 to chapter C.10 A). The chemical structure, or at least the empirical formula, should also be known in order to calculate theoretical values and/or to check measured values of parameters, e.g. theoretical oxygen demand (ThOD), DOC.
11. Information on the toxicity of the test chemical to micro-organisms (see Appendix 4 to chapter C.10 A) may be useful for selecting appropriate test concentrations and may be essential for the correct interpretation of low biodegradation values.

## PASS LEVELS

12. Originally, the primary biodegradation of surfactants was required to reach 80 % or more before the chemical could be marketed. If the value of 80 % is not attained, this simulation (confirmatory) test may be applied and the surfactant may be marketed only if more than 90 % of the specific chemical is removed. With chemicals in general there is no question of a pass/fail level and the value of percentage removed can be used in proximate calculations of the probable environmental concentration to be used in hazard assessments posed by chemicals. In a number of studies of pure chemicals the percentage removal of DOC was found to be > 90 % in more than three-quarters, and > 80 % in over 90 %, of chemicals which showed any significant degree of biodegradability.

## REFERENCE CHEMICALS

13. To ensure that the experimental procedure is being carried out correctly, it is useful occasionally to test reference chemicals whose behaviour is known. Such chemicals include adipic acid, 2-phenyl phenol, 1-naphthol, diphenic acid and 1-naphthoic acid.

## REPRODUCIBILITY OF TEST RESULTS

14. The relative standard deviation within tests was found by a laboratory in the UK to be 3,5 % and between tests to be 5 % (7).

**▼ M4**

## DESCRIPTION OF THE METHOD

**Apparatus***Rotating tubular reactors*

15. The apparatus (see figures 1 and 2 in the Appendix 8 consists of a bank of acrylic tubes each 30,5 cm long and 5 cm internal diameter, supported on rubber-rimmed wheels contained within a metal supporting frame. Each tube has an outside lip, approximately 0,5 cm deep, to retain it on the wheels, the internal surface is roughened with coarse wire wool and there is a 0,5 cm deep internal lip at the upper (feed) end to retain the liquid. The tubes are inclined at an angle of approximately one degree to the horizontal to achieve the required contact time when the test medium is applied to a clean tube. The rubber-tyred wheels are rotated using a slow, variable-speed motor. The temperature of the tubes is controlled by installation in a constant temperature room.
16. By enclosing each tube reactor inside a slightly larger, capped tube and ensuring that connections were gas-tight, exit CO<sub>2</sub> gas could be collected in an alkaline solution for subsequent measurement (6).
17. A 24h supply, for each tube, of organic medium with added test chemical if applicable, is contained in a 20 l storage vessel (A)(see Figure 2). If required, the test chemical solution may be dosed separately. Near the bottom of each storage vessel there is an outlet which is connected by suitable tubing, e.g. silicone rubber, via a peristaltic pump (B) to a glass or acrylic delivery tube which enters 2-4 cm into the higher (feed) end of the inclined tube (C). Effluent is allowed to drip from the lower end of the inclined tube to be collected in another storage vessel (D). The effluent is settled or filtered before analysis.

*Filtration apparatus-centrifuge*

18. Device for filtration of samples with membranes filter of suitable porosity (nominal aperture diameter 0,45 µm) which adsorb organic chemicals or release organic carbon to a minimum degree. If filters are used which release organic carbon, wash them carefully with hot water to remove leachable organic carbon. Alternatively a centrifuge capable of achieving 40 000 m/sec<sup>2</sup> may be used.
19. Analytical equipment for determining:
  - DOC/total organic carbon (TOC), or chemical oxygen demand (COD);
  - specific chemical (HPLC, GC etc.) if required;
  - pH, temperature, acidity, alkalinity;
  - ammonium, nitrite, nitrate, if the tests are performed under nitrifying conditions.

*Water*

20. Tap water, containing less than 3 mg/l DOC.
21. Distilled or deionised water, containing less than 2 mg/l DOC.

**▼ M4***Organic medium*

22. Synthetic sewage, domestic sewage or a mixture of both may be used as the organic medium. It has been shown that the use of domestic sewage alone often gives increased percentage removed of DOC (in activated sludge units) and even allows the biodegradation of some chemicals, which are not biodegraded when OECD synthetic sewage is used. Thus, the use of domestic sewage is recommended. Measure the DOC (or COD) concentration in each new batch of organic medium. The acidity or alkalinity of the organic medium should be known. The medium may require the addition of a suitable buffer (sodium hydrogen carbonate or potassium hydrogen phosphate), if it is of low acidity or alkalinity, to maintain a pH of about  $7,5 \pm 0,5$  in the reactor during the test. The amount of buffer, and when to add it, has to be decided in each individual case.

*Synthetic sewage*

23. Dissolve in each 1 litre of tap water: peptone, 160 mg; meat extract, 110 mg; urea, 30 mg; anhydrous dipotassium hydrogen phosphate, ( $K_2HPO_4$ ), 28 mg; sodium chloride, (NaCl), 7 mg; calcium chloride dihydrate, ( $CaCl_2 \cdot 2H_2O$ ), 4 mg; magnesium sulphate heptahydrate, ( $MgSO_4 \cdot 7H_2O$ ), 2 mg. This OECD synthetic sewage is an example and gives a mean DOC concentration in the influent of about 100 mg/l. Alternatively, use other compositions, with about the same DOC concentrations, which are closer to real sewage. This synthetic sewage may be made up in distilled water in a concentrated form and stored at about 1 °C for up to one week. When needed, dilute with tap water. (This medium is unsatisfactory e.g. nitrogen concentration is very high, relatively low carbon content, but nothing better has been suggested, except to add more phosphate, as buffer, and extra peptone).

*Domestic sewage*

24. Use fresh settled sewage collected daily from a treatment works receiving predominantly domestic sewage. It should be collected from the overflow channel of the primary sedimentation tank, or from the feed to activated sludge plant, and be largely free from coarse particles. The sewage can be used after storage for several days at about 4 °C, if it is proved that the DOC (or COD) has not significantly decreased (i.e. by less than 20 %) during storage. In order to limit disturbances to the system, the DOC (or COD) of each new batch should be adjusted before use to an appropriate constant value, e.g. by dilution with tap water.

*Lubricant*

25. Glycerol or olive oil may be used for lubricating the peristaltic pump rollers: both are suitable for use on silicone-rubber tubing.

*Stocks solutions of test chemical*

26. For chemicals of adequate solubility prepare stock solutions at appropriate concentrations (e.g. 1 to 5 g/l) in deionised water or in the mineral portion of the synthetic sewage. For insoluble chemicals, see Appendix 5 in chapter C.10-A. This method is not suitable for volatile chemicals without modifications to the tubular reactors (paragraph 16). Determine the DOC and TOC of the stock solution and repeat the measurements for each new batch. If the difference between the DOC and TOC is greater than 20 %, check the water-solubility of the test chemical. Compare the DOC or the concentration of the test chemical measured by specific analysis of the stock solution with



**▼M4**

the nominal value to ascertain whether recovery is good enough (normally > 90 % can be expected). Ascertain, especially for dispersions, whether or not DOC can be used as an analytical parameter or if only an analytical technique specific for the test chemical can be used. Centrifugation of the samples is required for dispersions. For each new batch, measure the DOC, COD or the test chemical with specific analysis.

27. Determine the pH of the stock solution. Extreme values indicate that the addition of the chemical may have an influence on the pH of the activated sludge in the test system. In this case neutralise the stock solution to obtain a pH of  $7 \pm 0,5$  with small amounts of inorganic acid or base, but avoid precipitation of the test chemical.

## PROCEDURE

*Preparation of organic medium for dosing*

28. Ensure that all influent and effluent containers and tubing from influent vessels and to effluent vessels are thoroughly cleaned to remove microbial growths, initially and throughout the test.
29. Prepare the synthetic sewage (paragraph 23) freshly each day either from the solids or from the concentrated stock solution by appropriate dilution with tap water. Measure the required amount in a cylinder and add to a clean influent vessel. Also, where necessary, add the required amount of the stock solution of the test chemical or reference chemical to the synthetic sewage before dilution. If it is more convenient or necessary to avoid loss of the test chemical, prepare a separate diluted solution of the test chemical in a separate reservoir and deliver this to the inclined tubes via a different dosing pump.
30. Alternatively (and preferably), use settled domestic sewage (paragraph 24) collected freshly each day if possible.

*Operation of rotating tubular reactors*

31. Two identical tubular reactors are required for the assessment of one test chemical, and they are assembled in a constant temperature room normally at  $22 \pm 2$  °C.
32. Adjust the peristaltic pumps to deliver  $250 \pm 25$  ml/h of the organic medium (without test chemical) into the inclined tubes, which are rotated at  $18 \pm 2$  rpm. Apply the lubricant (paragraph 25) to the pump tubes initially and periodically through the test to ensure proper functioning and to prolong the life of the tubing.
33. Adjust the angle of inclination of the tubes to the horizontal to produce a residence time of  $125 \pm 12,5$  sec. for the feed in a clean tube. Estimate the retention time by adding a non-biological marker (e.g. NaCl, inert dye) to the feed: the time taken to reach peak concentration in the effluent is taken to be the mean retention time (when maximum film is present, the retention time can increase up to about 30 min.).
34. These rates, speeds and times have been found to give adequate removals (> 80 %) of DOC (or COD) and to produce nitrified effluents. The rate of flow should be changed if removal is insufficient or if the performance of a particular treatment plant is to be simulated. In the latter case, adjust the rate of dosing the organic medium until the performance of the reactor matches that of the treatment plant.

**▼M4***Inoculation*

35. Airborne inoculation may be sufficient to start the growth of micro-organisms when synthetic sewage is used, but otherwise add 1 ml/l of settled sewage to the feed for 3 days.

*Measurements*

36. At regular intervals check that the dose-rates and rotating speeds are within the required limits. Also, measure the pH of the effluent, especially if nitrification is expected.

*Sampling and analysis*

37. The method, pattern and frequency of sampling are chosen to suit the purpose of the test. For example, take snap (or grab) samples of influent and effluent, or collect samples over a longer period e.g. 3-6 h. In the first period, without test chemical, take samples twice per week. Filter the samples through membranes or centrifuge at about 40 000 m/sec<sup>2</sup> for about 15 min (paragraph 18). It may be necessary to settle and/or coarse-filter the samples before membrane filtration. Determine DOC (or COD) at least in duplicate and if required BOD, ammonium and nitrite/nitrate.
38. All analyses should be performed as soon as possible after collection and preparation of the samples. If analyses have to be postponed, store the samples at about 4 °C in the dark in full, tightly stoppered bottles. If samples have to be stored for more than 48h, preserve them by deep-freezing, acidification or by addition of a suitable toxic chemical (e.g. 20 ml/l of a 10 g/l solution of mercury (II) chloride). Ensure that the preservation technique does not influence the results of analysis.

*Running-in period*

39. In this period, the surface biofilm grows to reach an optimal thickness, usually taking about 2 weeks and should not exceed 6 weeks. The removal (paragraph 44) of DOC (or COD) increases and reaches a plateau value. When the plateau has been reached at a similar value in both tubes, one is selected to be a control in the remainder of the test, during which their performance should remain consistent.

*Introduction of test chemical*

40. At this stage add the test chemical to the other reactor at the required concentration, usually 10-20 mg C/l. The control continues to receive the organic medium alone.

*Acclimation period*

41. Continue the twice weekly analyses for DOC (or COD) and, if primary biodegradability is to be assessed, also measure the concentration of the test chemical by specific analysis. Allow from one to six weeks (or longer under special conditions) after the test chemical has first been introduced for acclimation to occur. When the percentage removal (paragraphs 43-45) reaches a maximum value, obtain 12-15 valid values in the plateau phase over about 3 weeks for evaluation of the mean percentage removal. The test is considered completed if a sufficiently high degree of elimination is reached. Normally, do not exceed a test duration of more than 12 weeks after the first addition of the test chemical.

**▼ M4***Sloughing of the film*

42. The sudden removal of large quantities of excess film from the tubes, or sloughing, takes place at relatively regular intervals. To ensure that the comparability of results is unaffected, allow tests to cover at least two full cycles of growing and sloughing.

## DATA AND REPORTING

**Treatment of results**

43. Calculate the percentage DOC (or COD) elimination of the test chemical for each timed assessment using the equation:

$$D_t = 100 [C_s - (E - E_o)] / C_s \%$$

where:

$D_t$  = percentage elimination of DOC (or COD) at time  $t$ ;

$C_s$  = concentration of DOC (or COD) in the influent due to the test chemical, preferably estimated from the concentration in, and volume added, of the stock solution (mg/l);

$E$  = measured DOC (or COD) in the test effluent at time  $t$  (mg/l);

$E_o$  = measured DOC (or COD) in the control effluent at time  $t$  (mg/l).

Repeat the calculation for the reference chemical, if tested.

**Performance of the control reactor**

44. The degree of DOC (or COD) elimination ( $D_B$ ) of the organic medium in the control reactors is helpful information in assessing the biodegradative activity of the biofilm during the test. Calculate the percentage elimination from the equation:

$$D_B = 100 (1 - E_o / C_m) \%$$

where:

$C_m$  = DOC (or COD) of the organic medium in the control influent (mg/l).

45. Calculate the removal ( $D_{ST}$ ) of the test chemical, if measured, by a specific analytical method at each time assessment from the equation:

$$D_{ST} = 100 (1 - S_e / S_i) \%$$

where:

$S_i$  = measured or, preferably, estimated concentration of test chemical in the test influent (mg/l)

$S_e$  = measured test chemical concentration in the test effluent at time  $t$  (mg/l)

**▼ M4**

If the analytical method gives a positive value in unamended sewage equivalent to  $S_c$  mg/l, calculate the percentage removal ( $D_{SC}$ ) from:

$$DSC = 100 (S_i - S_e + S_c)/(S_i + S_c) \%$$

**Expression of test results**

46. Plot the percentage elimination  $D_t$  and  $D_{ST}$  (or  $D_{SC}$ ), if available, versus time (see Appendix 2 in chapter C.10- A). Take the mean (expressed to the nearest whole number) and standard deviation of the 12-15 values for  $D_T$  (and for  $D_{ST}$ , if available) obtained in the plateau phase as the percentage removal of the test chemical. From the shape of the elimination curve, some conclusions may be drawn about the removal processes.

**Adsorption**

47. If a high DOC elimination of the test chemical is observed at the beginning of the test, the test chemical is probably eliminated by adsorption on to the biofilm. It may be possible to prove this by determining the adsorbed test chemical on solids sloughed from the film. It is not usual for the elimination of the DOC of adsorbable chemicals to remain high throughout the test; normally, there is an initial high degree of removal which gradually falls to an equilibrium value. If, however, the adsorbed test chemical was able to cause acclimation of the microbial population, the elimination of the test chemical DOC would subsequently increase and reach a high, plateau level.

**Lag phase**

48. As in static, screening tests many test chemicals require a lag phase before full biodegradation occurs. In the lag phase, acclimation (or adaptation) of the competent bacteria takes place with almost no removal of the test chemical; then the initial growth of these bacteria occurs. This phase ends and the degradation phase is arbitrarily taken to begin when about 10 % of the initial amount of test chemical is removed (after allowing for adsorption, if it occurs). The lag phase is often highly variable and poorly reproducible.

**Plateau phase**

49. The plateau phase of an elimination curve in a continuous test is defined as that phase in which the maximum degradation takes place. This phase should last at least 3 weeks and have about 12-15 measured valid values.

**Mean degree of elimination of the test chemical**

50. Calculate the mean value from the elimination values  $D_t$  (and  $D_{st}$ , if available) of the test chemical at the plateau phase. Rounded to the nearest whole number (1 %), it is the degree of elimination of the test chemical. It is also recommended to calculate the 95 % confidence interval of the mean value. In a similar way calculate the mean degree ( $D_B$ ) of elimination of the organic medium in the control vessel.

**▼ M4****Indication of biodegradation**

51. If the test chemical does not adsorb significantly on to the biofilm and the elimination curve has a typical shape of a biodegradation curve with lag, degradation and plateau phases (paragraphs 48, 49), the measured elimination can safely be attributed to biodegradation. If a high initial removal has taken place, the simulation test cannot differentiate between biological and abiotic elimination processes. In such cases, and in other cases where there is any doubt about biodegradation (e.g. if stripping takes place), analyse adsorbed test chemical on samples of the film or perform additional static (screening) tests for biodegradability based on parameters clearly indicating biological processes. Such tests are the oxygen uptake methods (Chapter C.4 of this Annex D, E and F) (9) or a test which measures CO<sub>2</sub> production (Chapter C.4-C of this Annex or the Headspace method) (10); use as inoculum pre-exposed biofilm from the appropriate reactor.
52. If both the DOC removal and specific chemical removal have been measured, significant differences (the former being lower than the latter) between the percentages removed indicate the presence in the effluents of intermediate organic products, which may be more difficult to degrade; these should be investigated.

**Validity of test results**

53. Consider the test to be valid if the degree of DOC (or COD) elimination ( $D_B$ ) in the control units is > 80 % after 2 weeks operation and no unusual observations have been made.
54. If a readily biodegradable (reference) chemical has been tested, the degree of biodegradation should be > 90 % and the difference between duplicate values should not be greater than 5 %. If these two criteria are not met, review the experimental procedures and/or obtain domestic sewage from another source.
55. Similarly, differences between biodegradation values from duplicate units (if used) treating a test chemical should not differ by more than 5 %. If this criterion is not met but the removals are high, continue analysis for a further three weeks. If removal is low, investigate the inhibitory effects of the test chemical if not known and repeat the test at a lower concentration of test chemical, if that is feasible.

**Test Report**

56. The test report must include the following:

*Test chemical:*

- identification data;
- physical nature and, where relevant, physico-chemical properties.

*Test conditions:*

- any modifications to test system, especially if insolubles or volatiles tested;
- type of organic medium;
- proportion and nature of industrial wastes in sewage, if used and if known;
- method of inoculation;

**▼M4**

- test chemical stock solution — DOC (dissolved organic carbon) and TOC (total organic carbon) content; how prepared, if suspension; test concentration(s) used, reasons if outside range 10-20 mg/l DOC; method of addition; date first added; any changes in concentration;
- mean hydraulic retention time (with no growth); rotational speed of tube; approximate angle of inclination, if possible;
- details of sloughing; time and intensity;
- test temperature and range;
- analytical techniques employed.

*Test results:*

- all measured data DOC, COD, specific analyses, pH, temperature, N chemicals, if relevant;
- all calculated data of  $D_t$  (or  $D_{tc}$ ),  $D_B$ ,  $D_s$  obtained in tabular form and elimination curves;
- information on lag and plateau phases, test duration, the degree of elimination of the test chemical, of the reference chemical (if tested) and of the organic medium (in the control unit), together with statistical data and statements of biodegradability and validity of the test;
- discussion of results.

*LITERATURE:*

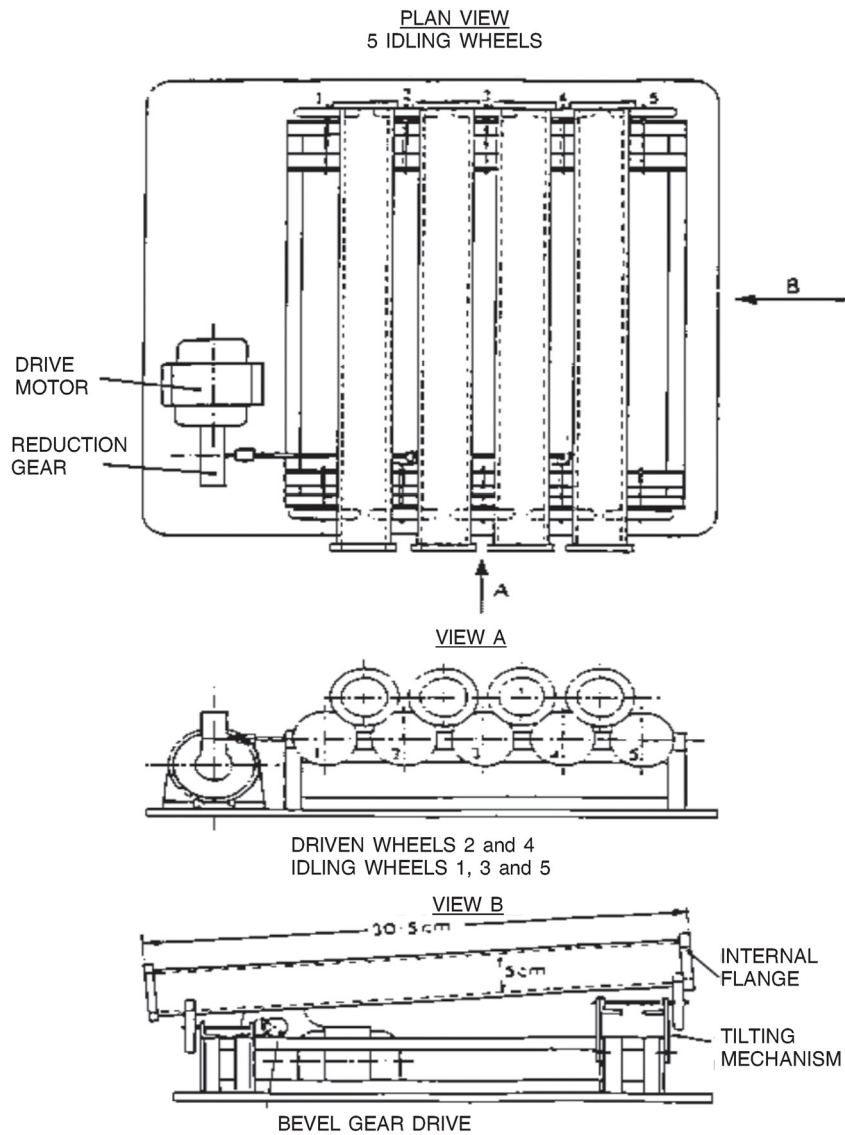
- (1) Gerike P, Fischer W, Holtmann W (1980). Biodegradability determinations in trickling filter units compared with the OECD Confirmatory Test. *Wat. Res.* 14: 753-758.
- (2) Truesdale GA, Jones K, Vandyke KG (1959). Removal of synthetic detergents in sewage treatment processes: Trials of a new biologically attackable material. *Wat. Waste Tr. J.* 7: 441-444.
- (3) Baumann U, Kuhn G and Benz M. (1998) Einfache Versuchsanordnung zur Gewinnung gewässerökologisch relevanter Daten, UWSF — Z. Umweltchem. Ökotox. 10: 214-220.
- (4) Gloyne EF, Comstock RF, Renn CE (1952). Rotary tubes as experimental trickling filters. *Sewage ind. Waste* 24: 1355-1357.
- (5) Kumke GW, Renn CE (1966). LAS removal across an institutional trickling filter. *JAOCS* 43: 92-94.
- (6) Tomlinson TG, Snaddon DHM, (1966). Biological oxidation of sewage by films of micro-organisms. *Int.J. Air Wat. Pollut.* 10: 865-881.
- (7) Her Majesty's Stationery Office (1982). Methods for the examination of waters and associated materials. Assessment of biodegradability, 1981, London.
- (8) Her Majesty's Stationery Office (1984). Methods for the examination of waters and associated materials. Methods for assessing the treatability of chemicals and industrial waste waters and their toxicity to sewage treatment processes, 1982, London.
- (9) Chapter C.4 of this Annex, Determination of 'Ready' Biodegradability A-F.
- (10) ISO 14593 (1998). Water Quality-Evaluation in an aqueous medium of the ultimate biodegradability of organic substances. Method by analysis of released inorganic carbon in sealed vessels.

▼ **M4**

Appendix 8

Figure 1

**Rotating tubes**



*Glossary:*

Plan view:

View A/B:

Driven wheels:

Idling wheels:

Drive motor:

Reduction gear:

Internal flange:

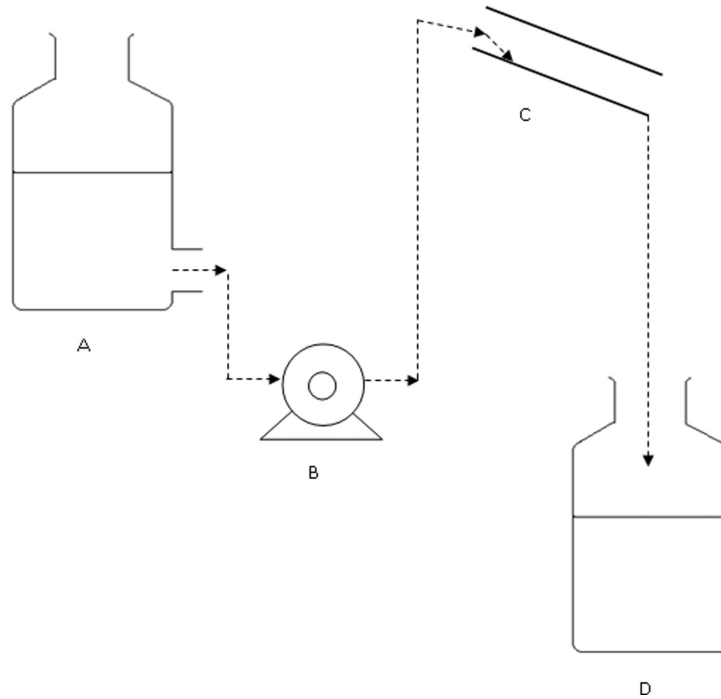
Tilting mechanism:

Bevel gear drive:

▼ M4

Figure 2

## Flow diagram



- A: Feed Tank
- B: Peristaltic Pump
- C: Rotating Tube
- D: Effluent Collection Vessel

*DEFINITIONS:*

Test chemical: Any substance or mixture tested using this Test Method.

Chemicals: It should be noted that the term 'chemical' is used broadly in the UNCED agreements and subsequent documents to include substances, products, mixtures, preparations, or any other terms that may be used in existing systems to denote coverage.





## C.11. BIODEGRADATION

### ACTIVATED SLUDGE RESPIRATION INHIBITION TEST

#### 1. METHOD

##### 1.1. INTRODUCTION

The method described assesses the effect of a test substance on micro-organisms by measuring the respiration rate under defined conditions in the presence of different concentrations of the test substance.

The purpose of this method is to provide a rapid screening method whereby substances which may adversely affect aerobic microbial treatment plants can be identified, and to indicate suitable non-inhibitory concentrations of test substances to be used in biodegradability tests.

A range-finding test may precede a definitive test. It provides information about the range of concentrations to be used in the main test.

Two controls without test substance are included in the test design, one at the start and the other at the end of the test series. Each batch of activated sludge should also be checked using a reference substance.

This method is most readily applied to substances which, due to their water solubility and low volatility, are likely to remain in water.

For substances with limited solubility in the test media, it may not be possible to determine the EC<sub>50</sub>.

Results based on oxygen uptake may lead to erroneous conclusions when the test substance has the propensity to uncouple oxidative phosphorylation.

It is useful to have the following information to perform the test:

- water solubility,
- vapour pressure,
- structural formula,
- purity of the test substance.

#### *Recommendation*

Activated sludge may contain potentially pathogenic organisms and should be handled with care.

##### 1.2. DEFINITIONS AND UNITS

The respiration rate is the oxygen consumption of waste-water micro-organisms in aerobic sludge, expressed generally as mg O<sub>2</sub> per mg of sludge per hour.

**▼ B**

In order to calculate the inhibitory effect of a test substance at a particular concentration, the respiration rate is expressed as a percentage of the mean of the two control respiration rates:

$$\left(1 - \frac{2R_s}{R_{c1} + R_{c2}}\right) \times 100 = \% \text{ inhibition}$$

where:

$R_s$  = oxygen-consumption rate at tested concentration of test substance

$R_{c1}$  = oxygen-consumption rate, control 1

$R_{c2}$  = oxygen-consumption rate, control 2.

$EC_{50}$  in this method is the concentration of the test substance at which the respiration rate is 50 % of that shown by the control under conditions described in this method.

### 1.3. REFERENCE SUBSTANCES

It is recommended that 3,5-dichlorophenol, as a known inhibitor of respiration, be used as a reference substance and tested for  $EC_{50}$  on each batch of activated sludge as a means of checking that the sensitivity of the sludge is not abnormal.

### 1.4. PRINCIPLE OF THE TEST METHOD

The respiration rate of an activated sludge fed with a standard amount of synthetic sewage feed is measured after a contact time of 30 minutes or three hours, or both. The respiration rate of the same activated sludge in the presence of various concentrations of the test substance under otherwise identical conditions is also measured. The inhibitory effect of the test substance at a particular concentration is expressed as a percentage of the mean respiration rates of two controls. An  $EC_{50}$  value is calculated from determinations at different concentrations.

### 1.5. QUALITY CRITERIA

The test results are valid if:

- the two control respiration rates are within 15 % of each other,
- the  $EC_{50}$  (30 minutes and/or three hours) of 3,5-dichlorophenol is in the accepted range 5 to 30 mg/litre.

### 1.6. DESCRIPTION OF THE TEST METHOD

#### 1.6.1. *Reagents*

##### 1.6.1.1. Solutions of the test substance

Solutions of the test substance are freshly prepared at the start of the study using a stock solution. A stock solution concentration of 0,5 g/litre is appropriate if the procedure recommended below is followed.

**▼B**

## 1.6.1.2. Solution of control substance

A solution of 5,5-dichlorophenol can for example be prepared by dissolving 0,5 g 3,5-dichlorophenol in 10 ml of 1M NaOH, diluting to approximately 30 ml with distilled water, adding under stirring 0,5M H<sub>2</sub>SO<sub>4</sub> to the point of incipient precipitation — approximately 8 ml of 0,5M H<sub>2</sub>SO<sub>4</sub> will be required — and finally diluting the mixture to one litre with distilled water. The pH should then be in the range 7 to 8.

## 1.6.1.3. Synthetic sewage

A synthetic sewage feed is made by dissolving the following amounts of substances in one litre of water:

- 16 g peptone,
- 11 g meat extract,
- 3 g urea,
- 0,7 g NaCl,
- 0,4 g CaCl<sub>2</sub>·2H<sub>2</sub>O,
- 0,2 g MgSO<sub>4</sub>·7H<sub>2</sub>O,
- 2,8 g K<sub>2</sub>HPO<sub>4</sub>.

*Note 1:* this synthetic sewage is a 100-fold concentrate of that described in the OECD Technical Report 'Proposed method for the determination of the biodegradability of surfactants used in synthetic detergents' (June 11, 1976), with the addition of dipotassium hydrogen phosphate.

*Note 2:* if the prepared medium is not used immediately, it shall be stored in the dark at 0 to 4 °C, for no longer than one week, in conditions which do not produce any change in its composition. The medium may also be sterilised prior to storage, or the peptone and meat extract may be added shortly before carrying out the test. Before use, it shall be mixed thoroughly and the pH adjusted.

## 1.6.2. Apparatus

Measuring apparatus: the precise design is not critical. However, there should be head space and the probe should fit tightly in the neck of the measuring flask.

Normal laboratory equipment and especially the following is necessary:

- measuring apparatus,
- aeration device,
- pH-electrode and measuring equipment,
- O<sub>2</sub>-electrode.

## 1.6.3. Preparation of the inoculum

Activated sludge from a sewage treatment plant treating predominantly domestic sewage is used as the microbial inoculum for the test.

If necessary, on return to the laboratory, coarse particles may be removed by settling for a short period, e.g. 15 minutes, and decanting the upper layer of finer solids for use. Alternatively, the sludge may be mixed using a blender for a few seconds.

**▼B**

In addition, if it is thought that inhibitory material is present, the sludge should be washed with tap water or an isotonic solution. After centrifuging, the supernatant is decanted (this procedure is repeated three times).

A small amount of the sludge is weighed and dried. From this result, the amount of wet sludge can be calculated which must be suspended in water in order to obtain an activated sludge with a mixed liquor suspended solids range between 2 and 4 g/litre. This level gives a concentration between 0,8 and 1,6 g/litre in the test medium if the procedure recommended below is followed.

If the sludge cannot be used on the day of collection, 50 ml of synthetic sewage is added to each litre of the activated sludge prepared as described above; this is then aerated overnight at  $20 \pm 2$  °C. It is then kept aerated for use during the day. Before use the pH is checked and adjusted, if necessary, to pH 6 to 8. The mixed liquor suspended solids should be determined as described in the preceding paragraph.

If the same batch of sludge is required to be used on subsequent days (maximum four days), a further 50 ml of synthetic sewage feed is added per litre of sludge at the end of each working day.

#### 1.6.4. *Performance of the test*

Duration/contact time:	30 minutes and/or three hours, during which aeration takes place
Water:	Drinking water (dechlorinated if necessary)
Air supply:	Clean, oil-free air. Air flow 0,5 to 1 litre/minute
Measuring apparatus:	Flat bottom flask such as a BOD-flask
Oxygen meter:	Suitable oxygen electrode, with a recorder
Nutrient solution:	Synthetic sewage (see above)
Test substance:	The test solution is freshly prepared at the start of the test
Reference substance:	e.g. 3,5-dichlorophenol (at least three concentrations)
Controls:	Inoculated sample without test substance
Temperature:	$20 \pm 2$ °C

A suggested experimental procedure which may be followed for both the test and reference substance for the three-hour contact period is given below:

Several vessels (e.g. one-litre beakers) are used.

At least five concentrations, spaced by a constant factor preferably not exceeding 3,2, should be used.

At time '0', 16 ml of the synthetic sewage feed are made up to 300 ml with water. 200 ml of microbial inoculum are added and the total mixture (500 ml) poured into a first vessel (first control C<sub>1</sub>).

**▼B**

The test vessels should be aerated continuously such as to ensure that the dissolved O<sub>2</sub> does not fall below 2,5 mg/litre and that, immediately before the measurement of the respiration rate, the O<sub>2</sub> concentration is about 6,5 mg/litre.

At time '15 minutes' (15 minutes is an arbitrary, but convenient, interval) the above is repeated, except that 100 ml of the test substance stock solution are added to the 16 ml of synthetic sewage before adding water to 300 ml and microbial inoculum to make a volume of 500 ml. This mixture is then poured into a second vessel and aerated as above. This process is repeated at 15-minute intervals with different volumes of the test substance stock solution to give a series of vessels containing different concentrations of the test substance. Finally, a second control is prepared (C<sub>2</sub>).

After three hours the pH is recorded, and a well-mixed sample of the contents of the first vessel is poured into the measuring apparatus and the respiration rate is measured over a period of up to 10 minutes.

This determination is repeated on the contents of each vessel at 15-minute intervals, in such a way that the contact time in each vessel is three hours.

The reference substance is tested on each batch of microbial inoculum in the same way.

A different regime (e.g. more than one oxygen meter) will be necessary when measurements are to be made after 30 minutes of contact.

If measurement of the chemical oxygen consumption is required, further vessels are prepared containing test substance, synthetic sewage feed and water, but no activated sludge. Oxygen consumption is measured and recorded after an aeration time of 30 minutes and/or three hours (contact time).

## 2. DATA AND EVALUATION

The respiration rate is calculated from the recorder trace between approximately 6,5 mg O<sub>2</sub>/litre and 2,5 mg O<sub>2</sub>/litre, or over a 10-minute period when the respiration rate is low. The portion of the respiration curve over which the respiration rate is measured should be linear.

If the respiration rates of the two controls are not within 15 % of each other, or the EC<sub>50</sub> (30 minutes and/or three hours) of the reference substance is not in the accepted range (5 to 30 mg/litre for 3,5-dichlorophenol), the test is invalid and must be repeated.

The per - cent inhibition is calculated at each test concentration (see 1.2). The per - cent inhibition is plotted against concentration on log-normal (or log-probability) paper, and an EC<sub>50</sub> value derived.

95 % confidence limits for the EC<sub>50</sub> values can be determined using standard procedures.

**▼B****3. REPORTING****3.1. TEST REPORT**

The test report shall, if possible, contain the following:

- test substance: chemical identification data,
- test system: source, concentration and any pre-treatment of the activated sludge,
- test conditions:
  - pH of the reaction mixture before the respiration measurement,
  - test temperature,
  - test duration,
  - reference substance and its measured EC<sub>50</sub>,
  - abiotic oxygen uptake (if any).
- results:
  - all measured data,
  - inhibition curve and method for calculation of EC<sub>50</sub>,
  - EC<sub>50</sub> and, if possible, 95 % confidence limits, EC<sub>20</sub> and EC<sub>80</sub>,
  - all observations and any deviations from this test method which could have influenced the result.

**3.2. INTERPRETATION OF DATA**

The EC<sub>50</sub> value should be regarded merely as a guide to the likely toxicity of the test substance either to activated sludge sewage treatment or to waste-water microorganisms, since the complex interactions occurring in the environment cannot be accurately simulated in a laboratory test. In addition, test substances which may have an inhibitory effect on ammonia oxidation may also produce atypical inhibition curves. Accordingly, such curves should be interpreted with caution.

**4. REFERENCES**

- (1) International Standard ISO 8192-1986.
- (2) Broecker, B., Zahn, R., Water Research 11,1977, p. 165.
- (3) Brown, D., Hitz, H. R., Schaefer, L., Chemosphere 10, 1981, p. 245.
- (4) ETAD (Ecological and Toxicological Association of Dyestuffs Manufacturing Industries), Recommended Method No 103, also described by:
- (5) Robra, B., Wasserl Abwasser 117, 1976, p. 80.
- (6) Schefer, W., Textilveredlung 6,1977, p. 247.
- (7) OECD, Paris, 1981, Test Guideline 209, Decision of the Council C(84) 30 final.

**▼B****C.12. BIODEGRADATION****MODIFIED SCAS TEST****1. METHOD****1.1. INTRODUCTION**

The purpose of the method is the evaluation of the potential ultimate biodegradability of water-soluble, non-volatile organic substances when exposed to relatively high concentrations of micro-organisms over a long time period. The viability of the microorganisms is maintained over this period by daily addition of a settled sewage feed. (For weekend requirements, the sewage may be stored at 4 °C. Alternatively, the synthetic sewage of the OECD confirmatory test may be used.)

Physico-chemical adsorption on the suspended solids may take place and this must be taken into account when interpreting results (see 3.2).

Because of the long detention period of the liquid phase (36 hours), and the intermittent addition of nutrients, the test does not simulate those conditions experienced in a sewage treatment plant. The results obtained with various test substances indicate that the test has a high biodegradation potential.

The conditions provided by the test are highly favourable to the selection and/or adaptation of micro-organisms capable of degrading the test compound. (The procedure may also be used to produce acclimatised inocula for use in other tests.)

In this method, the measure of the concentration of dissolved organic carbon is used to assess the ultimate biodegradability of the test substances. It is preferable to determine DOC after acidification and purging rather than as the difference of  $C_{\text{total}} - C_{\text{inorganic}}$ .

The simultaneous use of a specific analytical method may allow the assessment of the primary degradation of the substance (disappearance of the parent chemical structure).

The method is applicable only to those organic test substances which, at the concentration used in the test:

- are soluble in water (at least 20 mg dissolved organic carbon/litre),
- have negligible vapour pressure,
- are not inhibitory to bacteria,
- do not significantly adsorb within the test system,
- are not lost by foaming from the test solution.

The organic carbon content of the test material must be established.

**▼ B**

Information on the relative proportions of the major components of the test material will be useful in interpreting the results obtained, particularly in those cases where the results are low or marginal.

Information on the toxicity to microorganisms of the substance may be useful to the interpretation of low results and in the selection of an appropriate test concentration.

## 1.2. DEFINITIONS AND UNITS

$C_T$  = concentration of test compound as organic carbon as present in or added to the settled sewage at the start of the aeration period (mg/litre),

$C_t$  = concentration of dissolved organic carbon found in the supernatant liquor of the test at the end of the aeration period (mg/litre),

$C_c$  = concentration of dissolved organic carbon found in the supernatant liquor of the control at the end of the aeration period (mg/litre).

The biodegradation is defined in this method as the disappearance of the organic carbon. The biodegradation can be expressed as:

1. The percentage removal  $D_{da}$  of the amount of substance added daily:

$$D_{da} = \frac{C_T - (C_t - C_c)}{C_T} \times 100 \quad [1]$$

where

$D_{da}$  = degradation/daily addition.

2. The percentage removal  $D_{ssd}$  of the amount of substance present at the start of each day:

$$D_{ssd} = \frac{2C_T + C_{ti} - C_{ci} - 3C_{t(i+1)} + 3C_{c(i+1)}}{2C_T + C_{ti} - C_{ci}} \times 100 \quad [2 (a)]$$

$$\approx \frac{2C_T - 2(C_t - C_c)}{2C_T + (C_t - C_c)} \times 100 \quad [2 (b)]$$

where

$D_{ssd}$  = degradation/substance start of day;

the indices  $i$  and  $(i + 1)$  refer to the day of measurement.

Equation 2(a) is recommended if effluent DOC varies from day to day, while equation 2(b) may be used when effluent DOC remains relatively constant from day to day.



**▼B**

## 1.3. REFERENCE SUBSTANCES

In some cases, when investigating a new substance, reference substances may be useful; however, no specific reference substance is recommended here.

Data on several compounds evaluated in ring tests are provided (see Appendix 1) primarily so that calibration of the method may be performed from time to time and to permit comparison of results when another method is employed.

## 1.4. PRINCIPLE OF THE TEST METHOD

Activated sludge from a sewage treatment plant is placed in a semi-continuous activated sludge (SCAS) unit. The test compound and settled domestic sewage are added, and the mixture is aerated for 23 hours. The aeration is then stopped, the sludge allowed to settle and the supernatant liquor is removed.

The sludge remaining in the aeration chamber is then mixed with a further aliquot of test compound and sewage and the cycle is repeated.

Biodegradation is established by determination of the dissolved organic carbon content of the supernatant liquor. This value is compared with that found for the liquor obtained from a control tube dosed with settled sewage only.

When a specific analytical method is used, changes in the concentration of the parent molecule due to biodegradation can be measured (primary biodegradability).

## 1.5. QUALITY CRITERIA

The reproducibility of this method based on removal of dissolved organic carbon has not yet been established. (When primary biodegradation is considered, very precise data are obtained for materials that are extensively degraded).

The sensitivity of the method is largely determined by the variability of the blank and to a lesser extent by the precision of the determination of dissolved organic carbon and the level of test compound in the liquor at the start of each cycle.

## 1.6. DESCRIPTION OF THE TEST PROCEDURE

1.6.1. *Preparations*

A sufficient number of clean aeration units, alternatively, the original 1,5 litre SCAS test unit may be used, and air inlet tubes (Figure 1) for each test substance and controls are assembled. Compressed air supplied to the test units, cleaned by a cotton wool strainer, should be free of organic carbon and pre-saturated with water to reduce evaporation losses.

A sample of mixed liquor, containing 1 to 4 g suspended solids/litre, is obtained from an activated sludge plant treating predominantly domestic sewage. Approximately 150 ml of the mixed liquor are required for each aeration unit.

**▼B**

Stock solutions of the test substance are prepared in distilled water; the concentration normally required is 400 mg/litre as organic carbon which gives a test compound concentration of 20 mg/litre carbon at the start of each aeration cycle if no biodegradation is occurring.

Higher concentrations are allowed if the toxicity to microorganisms permits it.

The organic carbon content of the stock solutions is measured.

**1.6.2. *Test conditions***

The test should be performed at 20 to 25 °C.

A high concentration of aerobic microorganisms is used (from 1 to 4 g/litre suspended solids), and the effective detention period is 36 hours. The carbonaceous material in the sewage feed is oxidised extensively, normally within eight hours after the start of each aeration cycle. Thereafter, the sludge respire endogenously for the remainder of the aeration period, during which time the only available substrate is the test compound unless this is also readily metabolised. These features, combined with daily re-inoculation of the test when domestic sewage is used as the medium, provide highly favourable conditions for both acclimatisation and high degrees of biodegradation.

**1.6.3. *Performance of the test***

A sample of mixed liquor from a suitable predominantly domestic activated sludge plant or laboratory unit is obtained and kept aerobic until used in the laboratory. Each aeration unit as well as the control unit are filled with 150 ml of mixed liquor (if the original SCAS test unit is used, multiply the given volumes by 10) and the aeration is started. After 23 hours, aeration is stopped and the sludge is allowed to settle for 45 minutes. The tap of each vessel is opened in turn, and 100 ml portions of the supernatant liquor are withdrawn. A sample of settled domestic sewage is obtained immediately before use, and 100 ml are added to the sludge remaining in each aeration unit. Aeration is started anew. At this stage no test materials are added, and the units are fed daily with domestic sewage only until a clear supernatant liquor is obtained on settling. This usually takes up to two weeks, by which time the dissolved organic carbon in the supernatant liquor at the end of each aeration cycle approaches a constant value.

At the end of this period, the individual settled sludges are mixed, and 50 ml of the resulting composite sludge are added to each unit.

95 ml of settled sewage and 5 ml of water are added to the control units, and 95 ml of the settled sewage plus 5 ml of the appropriate test compound stock solution (400 mg/litre) are added to the test units. Aeration is started again and continued for 23 hours. The sludge is then allowed to settle for 45 minutes and the supernatant drawn off and analysed for dissolved organic carbon content.

The above fill-and-draw procedure is repeated daily throughout the test.

**▼B**

Before settling, it may be necessary to clean the walls of the units to prevent the accumulation of solids above the level of the liquid. A separate scraper or brush is used for each unit to prevent cross contamination.

Ideally, the dissolved organic carbon in the supernatant liquors is determined daily, although less frequent analyses are permissible. Before analysis the liquors are filtered through washed 0,45 µm membrane filters or centrifuged. Membrane filters are suitable if it is assured that they neither release carbon nor absorb the substance in the filtration step. The temperature of the sample must not exceed 40 °C while it is in the centrifuge.

The length of the test for compounds showing little or no biodegradation is indeterminate, but experience suggests that this should be at least 12 weeks in general, but not longer than 26 weeks.

**2. DATA AND EVALUATION**

The dissolved organic carbon values in the supernatant liquors of the test units and the control units are plotted against time.

As biodegradation is achieved, the level found in the test will approach that found in the control. Once the difference between the two levels is found to be constant over three consecutive measurements, such number of further measurements as are sufficient to allow statistical treatment of the data are made and the percentage biodegradation of the test compound is calculated ( $D_{da}$  or  $D_{ssd}$ , see 1.2).

**3. REPORTING****3.1. TEST REPORT**

The test report shall, if possible, contain the following:

- all information on the kind of sewage, the type of unit used and the experimental results concerning the tested substance, the reference substance if used, and the blank,
- the temperature,
- removal curve with description, mode of calculation (see 1.2),
- date and location where the activated sludge and the sewage were sampled, status of adaptation, concentration, etc.,
- scientific reasons for any changes of test procedure,
- signature and date.

**▼B**

## 3.2. INTERPRETATION OF RESULTS

Since the substance to be tested by this method will not be readily biodegradable, any removal of DOC due solely to biodegradation will normally be gradual over days or weeks, except in such cases where acclimatisation is sudden as indicated by an abrupt disappearance occurring after some weeks.

However, physico-chemical adsorption can sometimes play an important role; this is indicated when there is complete or partial removal of the added DOC at the outset. What happens subsequently depends on factors such as the degrees of adsorption and the concentration of suspended solids in the discarded effluent. Usually the difference between the concentration of DOC in the control and test supernatant liquors gradually increases from the initial low value and this difference then remains at the new value for the remainder of the experiment, unless acclimatisation takes place.

If a distinction is to be drawn between biodegradation (or partial biodegradation) and adsorption, further tests are necessary. This can be done in a number of ways, but the most convincing is to use the supernatant liquor, or sludge, as inoculum in a base-set test (preferably a respirometric test).

Test substances giving high, non-adsorptive removal of DOC in this test should be regarded as potentially biodegradable. Partial, non-adsorptive removal indicates that the chemical is at least subject to some biodegradation.

Low, or zero removals of DOC may be due to inhibition of micro-organisms by the test substance and this may also be revealed by lysis and loss of sludge, giving turbid supernatants. The test should be repeated using a lower concentration of test substance.

The use of a specific analytical method or of  $^{14}\text{C}$ -labelled test substance may allow greater sensitivity. In the case of  $^{14}\text{C}$  test compound, the recovery of the  $^{14}\text{CO}_2$  will confirm that biodegradation has occurred.

When results are also given in terms of primary biodegradation, an explanation should, if possible, be given on the chemical structure change that leads to the loss of response of the parent test substance.

The validation of the analytical method must be given together with the response found on the blank test medium.

## 4. REFERENCES

- (1) OECD, Paris, 1981, Test Guideline 302 A, Decision of the Council C(81) 30 final.

**▼ B***Appendix 1***SCAS test: example of results**

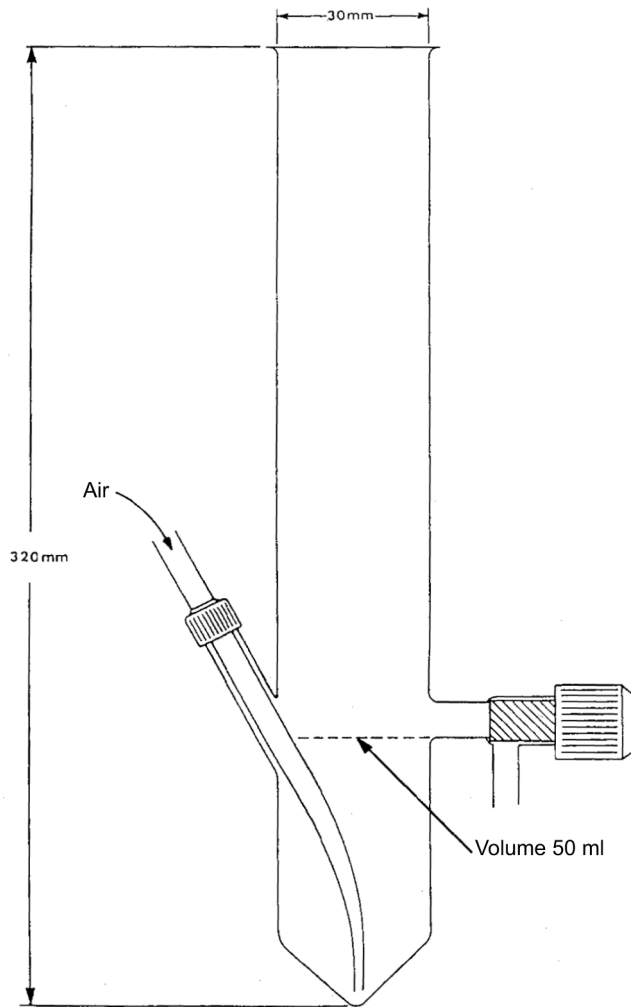
Substance	$C_T$ (mg/l)	$C_t - C_c$ (mg/l)	Percentage biodegradation, $D_{da}$	Test duration (days)
4-acetyl aminobenzene sulphonate	17,2	2,0	85	40
Tetra propylene benzene sulphonate	17,3	8,4	51,4	40
4-nitrophenol	16,9	0,8	95,3	40
Diethylene glycol	16,5	0,2	98,8	40
Aniline	16,9	1,7	95,9	40
Cyclopentane tetra carboxylate	17,9	3,2	81,1	120

▼B

Appendix 2

Example of test apparatus

Figure 1



**▼B****C.13 BIOCONCENTRATION: FLOW-THROUGH FISH TEST****1. METHOD**

This bioconcentration method is a replicate of the OECD TG 305 (1996).

**1.1. INTRODUCTION**

This method describes a procedure for characterising the bioconcentration potential of substances in fish under flow-through conditions. Although flow-through test regimes are much to be preferred, semi-static regimes are permissible, provided that the validity criteria are satisfied.

The method gives sufficient details for performing the test while allowing adequate freedom for adapting the experimental design to the conditions in particular laboratories and for varying characteristics of test substances. It is most validly applied to stable organic chemicals with  $\log P_{ow}$  values between 1,5 and 6,0 (1) but may still be applied to superlipophilic substances (having  $\log P_{ow} > 6,0$ ). The pre-estimate of the bioconcentration factor (BCF), sometimes denoted as  $K_B$ , for such superlipophilic substances will presumably be higher than the steady-state bioconcentration factor ( $BCF_{SS}$ ) value expected to be obtained from laboratory experiments. Preestimates of the bioconcentration factor for organic chemicals with  $\log P_{ow}$  values up to about 9,0 can be obtained by using the equation of Bintein et al (2). The parameters which characterise the bioconcentration potential include the uptake rate constant ( $k_1$ ), the depuration rate constant ( $k_2$ ) and the  $BCF_{SS}$ .

Radio-labelled test substances can facilitate the analysis of water and fish samples and may be used to determine whether degradate identification and quantification should be made. If total radioactive residues are measured (e.g. by combustion or tissue solubilisation), the BCF is based on the parent compound, any retained metabolites and also assimilated carbon. BCFs based on total radioactive residues may not, therefore, be directly comparable to a BCF derived by specific chemical analysis of the parent compound only.

Clean-up procedures may be employed in radiolabelled studies in order to determine BCF based on the parent compound, and the major metabolites may be characterised if deemed necessary. It is also possible to combine a fish metabolism study with a bioconcentration study by analysis and identification of the residues in tissues.

**1.2. DEFINITIONS AND UNITS**

Bioconcentration/Bioaccumulation is the increase in concentration of the test substance in or on an organism (specified tissues thereof) relative to the concentration of test substance in the surrounding medium.

The bioconcentration factor (BCF or  $K_B$ ) at any time during the uptake phase of this accumulation test is the concentration of test substance in/on the fish or specified tissues thereof ( $C_f$  as  $\mu\text{g/g}$  (ppm)) divided by the concentration of the chemical in the surrounding medium ( $C_w$  as  $\mu\text{g/ml}$  (ppm)).

**▼ B**

The steady-state bioconcentration factor ( $BCF_{ss}$  or  $K_B$ ) does not change significantly over a prolonged period of time, the concentration of the test substance in the surrounding medium being constant during this period of time.

A plateau or steady-state is reached in the plot of test substance in fish ( $C_f$ ) against time when the curve becomes parallel to the time axis and three successive analyses of  $C_f$  made on samples taken at intervals of at least two days are within  $\pm 20\%$  of each other, and there are no significant differences among the three sampling periods. When pooled samples are analysed at least four successive analyses are required. For test substances which are taken up slowly the intervals would more appropriately be seven days.

Bioconcentration factors calculated directly from kinetic rate constants ( $k_1/k_2$ ) are termed kinetic concentration factor,  $BCF_K$ .

The octanol-water partition coefficient ( $P_{ow}$ ) is the ratio of a chemical's solubility in n-octanol and water at equilibrium (Method A.8) also expressed as  $K_{ow}$ . The logarithm of  $P_{ow}$  is used as an indication of a chemical's potential for bioconcentration by aquatic organisms.

The exposure or uptake phase is the time during which the fish are exposed to the test chemical.

The uptake rate constant ( $k_1$ ) is the numerical value defining the rate of increase in the concentration of test substance in/on test fish (or specified tissues thereof) when the fish are exposed to that chemical ( $k_1$  is expressed in  $\text{day}^{-1}$ ).

The post-exposure or depuration (loss) phase is the time, following the transfer of the test fish from a medium containing test substance to a medium free of that substance, during which the depuration (or the net loss) of the substance from the test fish (or specified tissue thereof) is studied.

The depuration (loss) rate constant ( $k_2$ ) is the numerical value defining the rate of reduction in the concentration of the test substance in the test fish (or specified tissues thereof) following the transfer of the test fish from a medium containing the test substance to a medium free of that substance ( $k_2$  is expressed in  $\text{day}^{-1}$ ).

### 1.3. PRINCIPLE OF THE TEST METHOD

The test consists of two phases: the exposure (uptake) and post-exposure (depuration) phases. During the uptake phase, separate groups of fish of one species are exposed to at least two concentrations of the test substance. They are then transferred to a medium free of the test substance for the depuration phase. A depuration phase is always necessary unless uptake of the substance during the uptake phase has been insignificant (e.g. the BCF is less than 10). The concentration of the test substance in/on the fish (or specified tissue thereof) is followed through both phases of the test. In addition to the two test concentrations, a control group of fish is held under identical conditions except for the absence of the test substance, to relate possible adverse effects observed in the bioconcentration test to a matching control group and to obtain background concentrations of test substance.



**▼B**

The uptake phase is run for 28 days unless it is demonstrated that equilibrium has been reached earlier. A prediction of the length of the uptake phase and the time to steady-state can be made from equation in Appendix 3. The depuration period is then begun by transferring the fish to the same medium but without the test substance in another clean vessel. Where possible the bioconcentration factor is calculated preferably both as the ratio ( $BCF_{ss}$ ) of concentration of the fish ( $C_f$ ) and in the water ( $C_w$ ) at apparent steady-state and as a kinetic bioconcentration factor,  $BCF_K$  as the ratio of the rate constants of uptake ( $k_1$ ) and depuration ( $k_2$ ) assuming first-order kinetics. If first-order kinetics are obviously not obeyed, more complex models should be employed (Appendix 5).

If a steady-state is not achieved within 28 days, the uptake phase should be extended until steady-state is reached, or 60 days, whichever comes first; the depuration phase is then begun.

The uptake rate constant, the depuration (loss) rate constant (or constants, where more complex models are involved), the bioconcentration factor, and where possible, the confidence limits of each of these parameters are calculated from the model that best describes the measured concentrations of test substance in fish and water.

The BCF is expressed as a function of the total wet weight of the fish. However, for special purposes, specified tissues or organs (e.g. muscle, liver), may be used if the fish are sufficiently large or the fish may be divided into edible (fillet) and non-edible (viscera) fractions. Since, for many organic substances, there is a clear relationship between the potential for bioconcentration and lipophilicity, there is also a corresponding relationship between the lipid content of the test fish and the observed bioconcentration of such substances. Thus, to reduce this source of variability in test results for those substances with high lipophilicity (i.e. with  $\log P_{ow} > 3$ ), bioconcentration should be expressed in relation to lipid content in addition to whole body weight.

The lipid content should be determined on the same biological material as is used to determine the concentration of the test substance, when feasible.

#### 1.4. INFORMATION ON THE TEST SUBSTANCE

Before carrying out the test for bioconcentration, the following information for the test substance should be known:

- solubility in water
- octanol-water partition coefficient  $P_{ow}$  (denoted also as  $K_{ow}$ , determined by an HPLC method in A.8)
- hydrolysis
- phototransformation in water determined under solar or simulated solar irradiation and under the irradiation conditions of the test for bioconcentration (3)

**▼B**

- surface tension (i.e. for substances where the log  $P_{ow}$  cannot be determined)
- vapour pressure
- ready biodegradability (where appropriate)

Other information required is the toxicity to the fish species to be used in the test, preferably the asymptotic  $LC_{50}$  (i.e. time-independent). An appropriate analytical method, of known accuracy, precision and sensitivity, for the quantification of the test substance in the test solutions and in biological material must be available, together with details of sample preparation and storage. Analytical detection limit of test substance in both water and fish tissues should also be known. When  $^{14}C$  labelled test substance is used, the percentage of radioactivity associated with impurities should be known.

#### 1.5. VALIDITY OF THE TEST

The following conditions should apply for a test to be valid:

- the temperature variation is less than  $\pm 2$  °C;
- the concentration of dissolved oxygen does not fall below 60 % saturation;
- the concentration of the test substance in the chambers is maintained within  $\pm 20$  % of the mean of the measured values during the uptake phase;
- the mortality or other adverse effects/disease in both control and treated fish is less than 10 % at the end of the test; where the test is extended over several weeks or months, death or other adverse effects in both sets of fish should be less than 5 % per month and not exceed 30 % in all.

#### 1.6. REFERENCE COMPOUNDS

The use of reference compounds of known bioconcentration potential would be useful in checking the experimental procedure, when required. However, specific substances cannot yet be recommended.

#### 1.7. DESCRIPTION OF THE TEST METHOD

##### 1.7.1. *Apparatus*

Care should be taken to avoid the use of materials, for all parts of the equipment, that can dissolve, sorb or leach and have an adverse effect on the fish. Standard rectangular or cylindrical tanks, made of chemically inert material and of a suitable capacity in compliance with loading rate can be used. The use of soft plastic tubing should be minimised. Teflon®, stainless steel and/or glass tubing are preferably used. Experience has shown that for substances with high adsorption coefficients, such as the synthetic pyrethroids, silanised glass may be required. In these situations the equipment will have to be discarded after use.

**▼B**1.7.2. *Water*

Natural water is generally used in the test and should be obtained from an uncontaminated and uniform quality source. The dilution water must be of a quality that will allow the survival of the chosen fish species for the duration of the acclimation and test periods without them showing any abnormal appearance or behaviour. Ideally, it should be demonstrated that the test species can survive, grow and reproduce in the dilution water (e.g. in laboratory culture or a life-cycle toxicity test). The water should be characterised at least by pH, hardness, total solids, total organic carbon and, preferably also ammonium, nitrite and alkalinity and, for marine species, salinity. The parameters which are important for optimal fish well-being are fully known, but Appendix 1 gives recommended maximum concentrations of a number of parameters for fresh and marine test waters.

The water should be of constant quality during the period of a test. The pH value should be within the range 6,0 to 8,5, but during a given test it should be within a range of  $\pm 0,5$  pH units. In order to ensure that the dilution water will not unduly influence the test result (for example, by complexation of the test substance) or adversely affect the performance of the stock of fish, samples should be taken at intervals for analysis. Determination of heavy metals (e.g. Cu, Pb, Zn, Hg, Cd, Ni), major anions and cations (e.g. Ca, Mg, Na, K, Cl,  $\text{SO}_4$ ), pesticides (e.g. total organophosphorous and total organochlorine pesticides), total organic carbon and suspended solids should be made, for example, every three months where a dilution water is known to be relatively constant in quality. If water quality has been demonstrated to be constant over at least one year, determinations can be less frequent and intervals extended (e.g. every six months).

The natural particle content as well as the total organic carbon (TOC) of the dilution water should be as low as possible to avoid adsorption of the test substance to organic matter which may reduce its bioavailability (4). The maximum acceptable value is 5 mg/l for particulate matter (dry matter, not passing a 0,45  $\mu\text{m}$  filter) and 2 mg/l for total organic carbon (see Appendix 1). If necessary, the water should be filtered before use. The contribution to the organic carbon content in water from the test fish (excreta) and from the food residues should be as low as possible. Throughout the test, the concentration of organic carbon in the test vessel should not exceed the concentration of organic carbon originating from the test substance and, if used, the solubilising agent by more than 10 mg/l ( $\pm 20$  %).

1.7.3. *Test Solutions*

A stock solution of the test substance is prepared at a suitable concentration. The stock solution should preferably be prepared by simply mixing or agitating the test substance in the dilution water. The use of solvents or dispersants (solubilising agents) is not recommended; however this may occur in some cases in order to produce a suitably concentrated stock solution. Solvents which may be used are ethanol, methanol, ethylene glycol monomethyl ether, ethylene glycol dimethyl ether, dimethylformamide and triethylene glycol. Dispersants which may be used are Cremophor RH40, Tween 80, methylcellulose 0,01 % and HCO-40. Care should be taken when using readily biodegradable agents as these can cause problems with bacterial growth in flow-through tests. The test substance may be radio-labelled and should be of the highest purity (e.g. preferably  $> 98$  %).

**▼B**

For flow-through tests, a system which continuously dispenses and dilutes a stock solution of the test substance (e.g. metering pump, proportional diluter, saturator system) is required to deliver the test concentrations to the test chambers. At least five volume replacements through each test chamber per day are preferably allowed. The flow-through mode is to be preferred, but where this is not possible (e.g. when the test organisms are adversely affected) a semi-static technique may be used provided that the validity criteria are satisfied. The flow rates of stock solutions and dilution water should be checked both 48 h before and then at least daily during the test. In this check the determination of the flow-rate through each test chamber is included and ensured that it does not vary by more than 20 % either within or between chambers.

**1.7.4. *Selection of species***

Important criteria in the selection of species are that they are readily available, can be obtained in convenient sizes and can be satisfactorily maintained in the laboratory. Other criteria for selecting fish species include recreational, commercial, ecological importance as well as comparable sensitivity, past successful use etc.

Recommended test species are given in Appendix 2. Other species may be used but the test procedure may have to be adapted to provide suitable test conditions. The rationale for the selection of the species and the experimental method should be reported in this case.

**1.7.5. *Holding of fish***

Acclimate the stock population of fish for at least two weeks in water at the test temperature and feed throughout on a sufficient diet and of the same type to be used during the test.

Following a 48-hour settling-in period, mortalities are recorded and the following criteria applied:

- mortalities of greater than 10 % of population in seven days: reject the entire batch;
  
- mortalities of between 5 and 10 % of population in seven days: acclimate for seven additional days;
  
- mortalities of less than 5 % of population in seven days: accept the batch — if more than 5 % mortality during second seven days reject the entire batch.

Ensure that fish used in tests are free from observable diseases and abnormalities. Discard any diseased fish. Fish should not receive treatment for disease in the two weeks preceding the test, or during the test.

**▼B**

## 1.8. PERFORMANCE OF THE TEST

1.8.1. *Preliminary test*

It may be useful to conduct a preliminary experiment in order to optimise the test conditions of the definitive test, e.g. selection of test substance concentration(s), duration of the uptake and depuration phases.

1.8.2. *Conditions of exposure*

## 1.8.2.1. Duration of uptake phase

A prediction of the duration of the uptake phase can be obtained from practical experience (e.g. from a previous study or an accumulation related chemical) or from certain empirical relationships utilising knowledge of either the aqueous solubility or the octanol/water partition coefficient of the test substance (see Appendix 3).

The uptake phase should be run for 28 days unless it can be demonstrated that equilibrium has been reached earlier. If the steady-state has not been reached by 28 days, the uptake phase should be extended, taking further measurements, until steady-state is reached or 60 days, whichever is shorter.

## 1.8.2.2. Duration of the depuration phase

A period of half the duration of the uptake phase is usually sufficient for an appropriate (e.g. 95 %) reduction in the body burden of the substance to occur (see Appendix 3 for explanation of the estimation). If the time required to reach 95 % loss is impractically long, exceeding for example twice the normal duration of the uptake phase (i.e. more than 56 days) a shorter period may be used (i.e. until the concentration of test substance is less than 10 % of steady-state concentration). However, for substances having more complex patterns of uptake and depuration than are represented by a one-compartment fish model, yielding first order kinetics, allow longer depuration phases for determination of loss rate constants. The period may, however, be governed by the period over which the concentration of test substance in the fish remains above the analytical detection limit.

## 1.8.2.3. Numbers of test fish

Select the numbers of fish per test concentration such that minimum of four fish per sample are available at each sampling. If greater statistical power is required, more fish per sample will be necessary.

If adult fish are used, report whether male or female, or both are used in the experiment. If both sexes are used, differences in lipid content between sexes should be documented to be non-significant before the start of the exposure; pooling all male and all female fish may be necessary.

**▼B**

In any one test, fish of similar weight are selected, such that the smallest are no smaller than two-thirds of the weight of the largest. All should be of the same year-class and come from the same source. Since weight and age of a fish appear sometimes to have a significant effect on BCF values (1) these details are recorded accurately. It is recommended that a sub-sample of the stock of fish is weighed before the test in order to estimate the mean weight.

**1.8.2.4. Loading**

High water-to-fish ratios are used in order to minimise the reduction in  $C_w$  caused by the addition of the fish at the start of the test and also to avoid decreases in dissolved oxygen concentration. It is important that the loading rate is appropriate for the test species used. In any case, a loading rate of 0,1-1,0 g of fish (wet weight) per litre of water per day is normally recommended. High loading rates can be used if it is shown that the required concentration of test substance can be maintained within  $\pm 20$  % limits, and that the concentration of dissolved oxygen does not fall below 60 % saturation.

In choosing appropriate loading regimes, account of the normal habitat of the fish species is taken. For example, bottom-living fish may demand a larger bottom area of the aquarium for the same volume of water than pelagic fish species.

**1.8.2.5. Feeding**

During the acclimation and test periods, fish are fed with an appropriate diet of known lipid and total protein content, in an amount sufficient to keep them in a healthy condition and to maintain body weight. Fish are fed daily throughout the acclimation and test periods at a level of approximately 1 to 2 % of body weight per day; this keeps the lipid concentration in most species of fish at a relatively constant level during the test. The amount of feed should be recalculated, for example, once per week, in order to maintain consistent body weight and lipid content. For this calculation, the weight of the fish in each test chamber can be estimated from the weight of the fish sampled most recently in that chamber. Do not weigh the fish remaining in the chamber.

Uneaten food and faeces are siphoned daily from the test chambers shortly after feeding (30 minutes to 1 hour). Chambers are kept as clean as possible throughout the test so that the concentration of organic matter is kept as low as possible, since the presence of organic carbon may limit the bioavailability of the test substance (1).

Since many feeds are derived from fishmeal, the feed should be analysed for the test substance. It is also desirable to analyse the feed for pesticides and heavy metals.

**▼B**

## 1.8.2.6. Light and temperature

The photoperiod is usually 12 to 16 hours and the temperature ( $\pm 2$  °C) should be appropriate for the test species (see Appendix 2). The type and characteristics of illumination should be known. Caution should be given to the possible phototransformation of the test substance under the irradiation conditions of the study. Appropriate illumination should be used avoiding exposure of the fish to unnatural photoproducts. In some cases it may be appropriate to use a filter to screen out UV irradiation below 290 nm.

## 1.8.2.7. Test concentrations

Fish are exposed under flow-through conditions to at least two concentrations of the test substance in water. Normally, the higher (or highest) concentration of the test substance are selected to be about 1 % of its acute asymptotic LC<sub>50</sub>, and to be at least ten-fold higher than its detection limit in water by the analytical method used.

The highest test concentration can also be determined by dividing the acute 96 h LC<sub>50</sub> by an appropriate acute/chronic ratio (appropriate ratios for some chemicals can be about 3 up to 100). If possible, choose the other concentration(s) such that it differs from the one above by a factor of ten. If this is not possible because of the 1 % of LC<sub>50</sub> criterion and the analytical limit, a lower factor than ten can be used or the use of <sup>14</sup>C labelled test substance should be considered. No concentration used should be above the solubility of the test substance.

Where a solubilising agent is used its concentration should not be greater than 0,1 ml/l and should be the same in all test vessels. Its contribution, together with the test substance, to the overall content of organic carbon in the test water should be known. However, every effort should be made to avoid the use of such materials.

## 1.8.2.8. Controls

One dilution water control or if relevant, one control containing the solubilising agent should be run in addition to the test series, provided that it has been established that the agent has no effects on the fish. If not, both controls should be set up.

1.8.3. *Frequency of water quality measurements*

During the test, dissolved oxygen, TOC, pH and temperature should be measured in all vessels. Total hardness and salinity, if relevant, should be measured in the controls and one vessel at the higher (or highest) concentration. As a minimum, dissolved oxygen and salinity, if relevant, should be measured three times — at the beginning, around the middle and end of the uptake period — and once a week in the depuration period. TOC should be measured at the beginning of the test (24 h and 48 h prior to test initiation of uptake phase) before addition of the fish and at least once a week, during both uptake and depuration phases. Temperature should be measured daily, pH at the beginning and end of each period and hardness once each test. Temperature should preferably be monitored continuously in at least one vessel.

**▼B**1.8.4. *Sampling and analysis of fish and water*

## 1.8.4.1. Fish and water sampling schedule

Water from the test chambers for the determination of test substance concentration is sampled before addition of the fish and during both uptake and depuration phases. As a minimum, the water is sampled at the same time as the fish and before feeding. During the uptake phase, the concentrations of test substance are determined in order to check compliance with the validity criteria.

Fish is sampled on at least five occasions during the uptake phase and at least on four occasions during the depuration phase. Since on some occasions it will be difficult to calculate a reasonably precise estimate of the BCF value based on this number of samples, especially when other than simple first-order depuration kinetics are indicated, it may be advisable to take samples at a higher frequency in both periods (see Appendix 4). The extra samples are stored and analysed only if the results of the first round of analyses prove inadequate for the calculation of the BCF with the desired precision.

An example of an acceptable sampling schedule is given in Appendix 4. Other schedules can readily be calculated using other assumed values of  $P_{ow}$  to calculate the exposure time for 95 % uptake.

Sampling is continued during the uptake phase until a steady-state has been established or for 28 days, whichever is the shorter. If the steady-state has not been reached within 28 days sampling continues until a steady-state has been attained or for 60 days, whichever is shorter. Before beginning the depuration phase the fish are transferred to clean tanks.

## 1.8.4.2. Sampling and Sample Preparation

Water samples for analysis are obtained e.g. by siphoning through inert tubing from a central point in the test chamber. Since neither filtration nor centrifuging appear always to separate the non-bioavailable fraction of the test substance from that which is bioavailable (especially for super-lipophilic chemicals i.e. those chemicals with a  $\log P_{ow} > 5$ ) (1) (5), samples may not be subjected to those treatments.

Instead, measures should be taken to keep the tanks as clean as possible and the content of total organic carbon should be monitored during both the uptake and depuration phases.

An appropriate number of fish (normally a minimum of four) is removed from the test chambers at each sampling time. The sampled fish are rinsed quickly with water, blot 'dry', killed instantly using the most appropriate and humane method, and then weighed.

It is preferable to analyse fish and water immediately after sampling in order to prevent degradation or other losses and to calculate approximate uptake and depuration rates as the test proceeds. Immediate analysis also avoids delay in determining when a plateau has been reached.



**▼B**

Failing immediate analysis, samples are stored by an appropriate method. Before the beginning of the study, information on the proper method of storage for the particular test substance — for example, deep-freezing, holding at 4 °C, duration of storage, extraction, etc. are obtained.

#### 1.8.4.3. Quality of Analytical method

Since the whole procedure is governed essentially by the accuracy, precision and sensitivity of the analytical method used for the test substance, check experimentally that the precision and reproducibility of the chemical analysis, as well as recovery of the test substance from both water and fish are satisfactory for the particular method. Also, check that the test substance is not detectable in the dilution water used.

If necessary, the values of  $C_w$  and  $C_f$  obtained from the test are corrected for the recoveries and background values of controls. Fish and water samples are handled throughout in such a manner as to minimise contamination and loss (e.g. resulting from adsorption by the sampling device).

#### 1.8.4.4. Analysis of Fish Sample

If radiolabelled materials are used in the test, it is possible to analyse for total radio label (i.e. parent and metabolites) or, the samples may be cleaned up so that parent compound can be analysed separately. Also, the major metabolites may be characterised at steady-state or at the end of the uptake phase, whichever is the sooner. If the BCF in terms of total radiolabelled residues is  $\geq 1\,000$  %, it may be advisable, and for certain categories of chemicals such as pesticides strongly recommended, to identify and quantify degradates representing  $\geq 10$  % of total residues in fish tissues at steady state. If degradates representing  $\geq 10$  % of total radiolabelled residues in the fish tissue are identified and quantified, then it is also recommended to identify and quantify degradates in the test water.

The concentration of the test substance should usually be determined for each weighed individual fish. If this is not possible, pooling of the samples on each sampling occasion may be done but pooling does restrict the statistical procedures which can be applied to the data. If a specific statistical procedure and power are important considerations, then an adequate number of fish to accommodate the desired pooling procedure and power should be included in the test (6) (7).

BCF should be expressed both as a function of total wet weight and, for high lipophilic substances, as a function of the lipid content. Lipid content of the fish is determined on each sampling occasion if possible. Suitable methods should be used for determination of lipid content (ref. 8 and 2 of Appendix 3). Chloroform/methanol extraction technique may be recommended as standard method (9). The various methods do not give identical values (10), so it is important to give details of the method used. When possible, the analysis for lipid should be made on the same extract as that produced for analysis for the test substance, since the lipids often have to be removed from the extract before it can be analysed chromatographically. The lipid content of the fish (as mg/kg wet weight) at the end of the experiment should not differ from that at the start by more  $\pm 25$  %. The tissue percent solids should also be reported to allow conversion of lipid concentration from a wet to a dry basis.

**▼ B****2. DATA****2.1. TREATMENT OF RESULTS**

The uptake curve of the test substance is obtained by plotting its concentration in/on fish (or specified tissues) in the uptake phase against time on arithmetic scales. If the curve has reached a plateau, that is, become approximately asymptotic to the time axis, the steady-state  $BCF_{ss}$  is calculated from:

$$\frac{C_f \text{ as steady - state (mean)}}{C_w \text{ as steady - state (mean)}}$$

When no steady-state is reached, it may be possible to calculate a  $BCF_{ss}$  of sufficient precision for hazard assessment from a 'steady-state' at 80 % ( $1,6/k_2$ ) or 95 % ( $3,0/k_2$ ) of equilibrium.

Also the concentration factor ( $BCF_K$ ) is determined, as the ratio  $k_1/k_2$ , the two first-order kinetic constants. The depuration rate constant ( $k_2$ ) is usually determined from the depuration curve (i.e. a plot of the decrease in test substance concentration in the fish with time). The uptake rate constant ( $k_1$ ) is then calculated given  $k_2$  and a value of  $C_f$  which is derived from the uptake curve (see also Appendix 5). The preferred method for obtaining  $BCF_K$  and the rate constants,  $k_1$  and  $k_2$ , is to use non-linear parameter estimation methods on a computer (11). Otherwise, graphical methods may be used to calculate  $k_1$  and  $k_2$ . If the depuration curve is obviously not first-order, then more complex models should be employed (see references in Appendix 3) and advice from a biostatistician sought.

**2.2. INTERPRETATION OF RESULTS**

The results should be interpreted with caution where measured concentrations of test solutions occur at levels near the detection limit of the analytical method.

Clearly defined uptake and loss curves are an indication of good quality bioconcentration data. The variation in uptake/depuration constants between the two test concentrations should be less than 20 %. Observed significant differences in uptake/depuration rates between the two applied test concentrations should be recorded and possible explanations given. Generally the confidence limit of  $BCFs$  from well-designed studies approach  $\pm 20$  %.

**3. REPORTING**

The test report must include the following information:

**3.1. TEST SUBSTANCE:**

- physical nature and, where relevant, physicochemical properties,
- chemical identification data (including the organic carbon content, if appropriate),
- if radio labelled, the precise position of the labelled atom(s) and the percentage of radioactivity associated with impurities.

**▼ B**

## 3.2. TEST SPECIES

- scientific name, strain, source, any pre-treatment, acclimation, age, size-range, etc.

## 3.3. TEST CONDITIONS:

- test procedure used (e.g. flow-through or semi-static),
- type and characteristics of illumination used and photoperiod(s),
- test design (e.g. number and size of test chambers, water volume replacement rate, number of replicates, number of fish per replicate, number of test concentrations, length of uptake and depuration phases, sampling frequency for fish and water samples,
- method of preparation of stock solutions and frequency of renewal (the solubilizing agent, its concentration and its contribution to the organic carbon content of test water must be given, when used),
- the nominal test concentrations, the means of the measured values and their standard deviations in the test vessels and the method by which these were attained,
- source of the dilution water, description of any pre-treatment, results of any demonstration of the ability of test fish to live in the water, and water characteristics: pH, hardness, temperature, dissolved oxygen concentration, residual chlorine levels (if measured), total organic carbon, suspended solids, salinity of the test medium (if appropriate) and any other measurements made,
- water quality within test vessels, pH, hardness, TOC, temperature and dissolved oxygen concentration,
- detailed information on feeding (e.g. type of food, source, composition — at least lipid and protein content if possible, amount given and frequency),
- information on the treatment of fish and water samples, including details of preparation, storage, extraction and analytical procedures (and precision) for the test substance and lipid content (if measured).

## 3.4. RESULTS:

- results from any preliminary study performed,
- mortality of the control fish and the fish in each exposure chamber and any observed abnormal behaviour,
- the lipid content of the fish (if determination on testing occasion),
- curves (including all measured data) showing the uptake and depuration of the test chemical in the fish, the time to steady-state,

**▼B**

- Cf and Cw (with standard deviation and range, if appropriate) for all sampling times (Cf expressed in µg/g wet weight (ppm) of whole body or specified tissues thereof e.g. lipid, and Cw in µg/ml (ppm). Cw values for the control series (background should also be reported),
- the steady-state bioconcentration factor (BCFss) and/or kinetic concentration factor (BCFK) and if applicable, 95 % confidence limits for the uptake and depuration (loss) rate constants (all expressed in relation to the whole body and the total lipid content, if measured, of the animal or specified tissues thereof), confidence limits and standard deviation (as available) and methods of computation/data analysis for each concentration of test substance used,
- where radio-labelled substances are used, and if it is required, the accumulation of any detected metabolites may be presented,
- anything unusual about the test, any deviation from these procedures, and any other relevant information,

minimise results as 'not detected at the limit of detection' by pre-test method development and experimental design, since such results cannot be used for rate constant calculations.

#### 4. REFERENCES

- (1) Connell D.W. (1988) Bioaccumulation behaviour of persistent chemicals with aquatic organisms. *Rev. Environ. Contam. Toxicol.* 102, p. 117-156.
- (2) Bintein S., Devillers J. and Karcher W. (1993) Nonlinear dependence of fish bioconcentration on n-octanol/water partition coefficient. *SAR and QSAR in Environmental Research*, 1, p. 29-390.
- (3) OECD, Paris (1996) Direct Phototransformation of chemicals in water. *Environmental Health and Safety Guidance Document Series on Testing and Assessment of Chemicals*. No 3.
- (4) Kristensen P. (1991) Bioconcentration in fish: Comparison of bioconcentration factors derived from OECD and ASTM testing methods; influence of particulate organic matter to the bioavailability of chemicals. Water Quality Institute, Denmark.
- (5) US EPA 822-R-94-002 (1994) Great Lake Water Quality Initiative Technical Support Document for the Procedure to Determine Bioaccumulation Factors. July 1994.
- (6) US FDA, (Food and Drug Administration) Revision. *Pesticide analytical manual*, 1, 5600 Fisher's Lane, Rockville, MD 20852, July 1975.
- (7) US EPA (1974). Section 5, A(1) Analysis of Human or Animal Adipose Tissue, in *Analysis of Pesticide Residues in Human and Environmental Samples*, Thompson J.F. (ed.) Research Triangle Park, N.C. 27711.
- (8) Compaan H. (1980) in 'The determination of the possible effects of chemicals and wastes on the aquatic environment: degradation, toxicity, bioaccumulation' Ch. 2.3, Part II. Government Publishing Office, The Hague, The Netherlands.

**▼B**

- (9) Gardner et al, (1995) *Limn. & Oceanogr.* 30, p. 1099-1105.
- (10) Randall R.C., Lee H., Ozretich R.J., Lake J.L. and Pruell R.J. (1991). Evaluation of selected lipid methods for normalising pollutant bioaccumulation. *Envir. Toxicol. Chem.* 10, p. 1431-1436.
- (11) CEC, Bioconcentration of chemical substances in fish: the flow-through method-Ring Test Programme, 1984 to 1985. Final report March 1987. Authors: P. Kristensen and N. Nyholm.
- (12) ASTM E-1022-84 (Reapproved 1988) Standard Practice for conducting Bioconcentration Tests with Fishes and Saltwater Bivalve Molluscs.

▼B*Appendix 1***Chemical characteristics of an acceptable dilution water**

	Substance	Limit concentration
1	Particulate Matter	5 mg/l
2	Total Organic Carbon	2 mg/l
3	Un-ionised ammonia	1 µg/l
4	Residual chlorine	10 µg/l
5	Total organophosphorous pesticides	50 ng/l
6	Total organochlorine pesticides plus polychlorinated biphenyls	50 ng/l
7	Total organic chlorine	25 ng/l
8	Aluminium	1µg/l
9	Arsenic	1µg/l
10	Chromium	1µg/l
11	Cobalt	1µg/l
12	Copper	1µg/l
13	Iron	1µg/l
14	Lead	1µg/l
15	Nickel	1µg/l
16	Zinc	1µg/l
17	Cadmium	100 ng/l
18	Mercury	100 ng/l
19	Silver	100 ng/l



## Appendix 2

## Fish species recommended for testing

	Recommended species	Recommended range of test temperature (°C)	Recommended total length of test animal (cm)
1	<i>Danio rerio</i> <sup>(1)</sup> (Teleostei, Cyprinidae) (Hamilton-Buchanan) Zebra — fish	20-25	3,0 ± 0,5
2	<i>Pimephales promelas</i> (Teleostei, Cyprinidae) (Rafinesque) Fathead minnow	20-25	5,0 ± 2,0
3	<i>Cyprinus carpio</i> (Teleostei, Cyprinidae) (Linnaeus) Common Carp	20-25	5,0 ± 3,0
4	<i>Oryzias latipes</i> (Teleostei, Poeciliidae) (Temminck and Schlegel) Ricefish	20-25	4,0 ± 1,0
5	<i>Poecilia reticulata</i> (Teleostei, Poeciliidae) (Peters) Guppy	20-25	3,0 ± 1,0
6	<i>Lepomis macrochirus</i> (Teleostei, Centrarchidae) (Rafinesque) Bluegill	20-25	5,0 ± 2,0
7	<i>Oncorhynchus mykiss</i> (Teleostei, Salmonidae) (Walbaum) Rainbow trout	13-17	8,0 ± 4,0
8	<i>Gasterosteus aculeatus</i> (Teleostei, Gasterosteidae) (Linnaeus) Three-spined stickleback	18-20	3,0 ± 1,0

<sup>(1)</sup> Meyer A., Orti G. (1993) Proc. Royal Society of London, Series B., Vol. 252, p. 231.

Various estuarine and marine species have been used in different countries, for example:

Spot	<i>Leiostomus xanthurus</i>
Sheepshead minnow	<i>Cyprinodon variegatus</i>
Silverside	<i>Menidia beryllina</i>
Shiner perch	<i>Cymatogaster aggregata</i>
English sole	<i>Parophrys vetulus</i>
Staghorn sculpin	<i>Leptocottus armatus</i>
Three-spined stickleback	<i>Gasterosteus aculeatus</i>
Sea bass	<i>Dicentrarchus labrax</i>
Bleak	<i>Alburnus alburnus</i>

## Collection

The fresh water fish listed in the table above are easy to rear and/or are widely available throughout the year, whereas the availability of marine and estuarine species is partially confined to the respective countries. They are capable of being bred and cultivated either in fish farms or in the laboratory, under disease-and parasite-controlled conditions, so that the test animal will be healthy and of known parentage. These fish are available in many parts of the world.

**▼ B***Appendix 3***Prediction of the duration of the uptake and depuration phases**1. *Prediction of the duration of the uptake phase*

Before performing the test, an estimate of  $k_2$  and hence some percentage of the time needed to reach steady-state may be obtained from empirical relationships between  $k_2$  and the n-octanol/water partition coefficient ( $P_{ow}$ ) or  $k_2$  and the aqueous solubility (s).

An estimate of  $k_2$  ( $\text{day}^{-1}$ ) may be obtained, for example from the following empirical relationship (1):

$$\log_{10}k_2 = 0,414 \log_{10}(P_{ow}) + 1,47 \quad (r^2 = 0,95) \quad \text{[equation 1]}$$

For other relationships see Ref. (2).

If the partition coefficient ( $P_{ow}$ ) is not known, an estimate can be made (3) from a knowledge of the aqueous solubility (s) of the substance using:

$$\log_{10}(P_{ow}) = 0,862 \log_{10}(s) + 0,710 \quad (r^2 = 0,994) \quad \text{[equation 2]}$$

where

s = solubility (moles/l): (n=36)

These relationships apply only to chemicals with  $\log P_{ow}$  values between 2 and 6,5 (4).

The time to reach some percentage of steady-state may be obtained, by applying the  $k_2$ -estimate, from the general kinetic equation describing uptake and depuration (first-order kinetics):

$$\frac{dC_f}{dt} = k_1 \cdot C_w - k_2 \cdot C_f$$

or if  $C_w$  is constant:

$$C_f = \frac{k_1}{k_2} \cdot C_w (1 - e^{-k_2 t}) \quad \text{[equation 3]}$$

When steady-state is approached ( $t \rightarrow \infty$ ), equation 3 may be reduced (5) (6) to:

$$C_f = \frac{k_1}{k_2} \cdot C_w \quad \text{or} \quad C_f/C_w = k_1/k_2 = \text{BCF}$$

Then  $k_1/k_2 \cdot C_w$  is an approach to the concentration in the fish at 'steady-state' ( $C_{f,s}$ ).

Equation 3 may be transcribed to:

$$C_f = C_{f,s}(1 - e^{-k_2 t}) \quad \text{or} \quad \frac{C_f}{C_{f,s}} = 1 - e^{-k_2 t} \quad \text{[equation 4]}$$

Applying equation 4, the time to reach some percentage of steady-state may be predicted when  $k_2$  is pre-estimated using equation 1 or 2.



**▼ B**

As a guideline, the statistically optimal duration of the uptake phase for the production of statistically acceptable data ( $BCF_K$ ) is that period which is required for the curve of the logarithm of the concentration of the test substance in fish plotted against linear time to reach its mid-point, or  $1,6/k_2$ , or 80 % of steady-state but not more than  $3,0/k_2$  or 95 % of steady-state (7).

The time to reach 80 % of steady-state is (equation 4):

$$0,80 = 1 - e^{-k_2 t_{80}} \quad \text{or} \quad t_{80} = \frac{1,6}{k_2} \quad \text{[equation 5]}$$

Similarly 95 percent of steady-state is:

$$t_{95} = \frac{3,0}{k_2} \quad \text{[equation 6]}$$

For example, the duration of the uptake phase (up) for a test substance with  $\log P_{ow} = 4$  would be (using equations 1,5,6):

$$\log_{10} k_2 = -0,414 \cdot (4) + 1,47 \quad k_2 = 0,652 \text{ days}^{-1}$$

$$\text{up (80 \%)} = 1,6/0,652, \text{ i.e. } 2,45 \text{ days (59 hours)}$$

$$\text{or up (95 \%)} = 3,0/0,652, \text{ i.e. } 4,60 \text{ days (110 hours)}$$

Similarly, for a test substance with  $s = 10^{-5}$  mol/l ( $\log(s) = -5,0$ ), the duration of up would be (using equations 1,2,5,6):

$$\log_{10} (P_{ow}) = 0,862 \cdot (-5,0) + 0,710 = 5,02$$

$$\log_{10} K_2 = 0,414 \cdot (5,02) + 1,47$$

$$k_2 = 0,246 \text{ days}^{-1}$$

$$\text{up (80 \%)} = 1,6/0,246, \text{ i.e. } 6,5 \text{ days (156 hours)}$$

$$\text{or up (95 \%)} = 3,0/0,246, \text{ i.e. } 12,2 \text{ days (293 hours)}$$

Alternatively, the expression:

$$t_{eq} = 6,54 \times 10^3 P_{ow} + 55,31 \text{ (hours)}$$

may be used to calculate the time for effective steady-state to be reached (4).

## 2. Prediction of the duration of the depuration phase

A prediction of the time needed to reduce the body burden to some percentage of the initial concentration may also be obtained from the general equation describing uptake and depuration (first order kinetics) (1) (8).

For the depuration phase,  $C_w$  is assumed to be zero. The equation may be reduced to:

$$\frac{dC_f}{dt} = -k_2 C_f \quad \text{or} \quad C_f = C_{f,o} \cdot e^{-k_2 t}$$

**▼B**

where  $C_{f,o}$  is the concentration at the start of the depuration period. 50 % depuration will then be reached at the time ( $t_{50}$ ):

$$\frac{C_f}{C_{f,o}} = \frac{1}{2} = e^{-k_2 t_{50}} \text{ or } t_{50} = \frac{0,693}{k_2}$$

Similarly 95 % depuration will be reached at:

$$t_{95} = \frac{3,0}{k_2}$$

If 80 % uptake is used for the first period ( $1,6/k_2$ ) and 95 % loss in the depuration phase ( $3,0/k_2$ ), then the depuration phase is approximately twice the duration of the uptake phase.

It is important to note, however, that the estimations are based on the assumption that uptake and depuration patterns will follow first order kinetics. If first order kinetics are obviously not obeyed, more complex models should be employed (e.g. ref (1)).

**Literature** (of Appendix 3)

- (1) Spacie A. and Hamelink J.L. (1982) Alternative models for describing the bioconcentration of organics in fish. *Environ. Toxicol. and Chem.* 1, pp 309-320.
- (2) Kristensen P. (1991) Bioconcentration in fish: comparison of BCF's derived from OECD and ASTM testing methods; influence of particulate matter to the bioavailability of chemicals. Danish Water Quality Institute.
- (3) Chiou C.T. and Schmedding D.W. (1982) Partitioning of organic compounds in octanol-water systems. *Environ. Sci. Technol.* 16 (1), pp 4-10.
- (4) Hawker D.W. and Connell D.W. (1988) Influence of partition coefficient of lipophilic compounds on bioconcentration kinetics with fish. *Wat. Res.* 22 (6), pp 701-707.
- (5) Branson D.R., Blau G.E., Alexander H.C. and Neely W.B. (1975) *Transactions of the American Fisheries Society*, 104 (4), pp 785-792.
- (6) Ernst W. (1985) Accumulation in Aquatic organisms. In: *Appraisal of tests to predict the environmental behaviour of chemicals*. Ed. by Sheehman P., Korte F., Klein W. and Bourdeau P.H. Part 4.4 pp 243-255. SCOPE, 1985, John Wiley & Sons Ltd. N.Y.
- (7) Reilly P.M., Bajramovic R., Blau G.E., Branson D.R. and Sauerhoff M.W. (1977) Guidelines for the optimal design of experiments to estimate parameters in first order kinetic models, *Can. J. Chem. Eng.* 55, pp 614-622.
- (8) Könemann H. and Van Leeuwen K. (1980) Toxicokinetics in fish: Accumulation and Elimination of six Chlorobenzenes by Guppies. *Chemosphere*, 9, pp 3-19.

▼ B

## Appendix 4

Theoretical example of sampling schedule for bioconcentration tests of substances with  $\log P_{ow} = 4$ 

Fish sampling	Sample time schedule		No of water samples	No of fish per sample
	Minimal required frequency (days)	Additional sampling		
Uptake phase	-1 0		2 (*) 2	add 45-80 fish
1st	0,3	0,4	2 (2)	4 (4)
2nd	0,6	0,9	2 (2)	4 (4)
3rd	1,2	1,7	2 (2)	4 (4)
4th	2,4	3,3	2 (2)	4 (4)
5th	4,7		2	6
Depuration phase				Transfer fish to water free of test chemical
6th	5,0	5,3		4 (4)
7th	5,9	7,0		4 (4)
8th	9,3	11,2		4 (4)
9th	14,0	17,5		6 (4)

(\*) Sample water after minimum of 3 'chamber-volume' have been delivered.

Values in brackets are numbers of samples (water, fish) to be taken if additional sampling is carried out.

Note: pre-test estimate of  $k_2$  for  $\log P_{ow}$  of 4,0 is  $0,652 \text{ days}^{-1}$ . The total duration of the experiment is set to  $3 \times \text{up} = 3 \times 4,6$  days, i.e. 14 days. For the estimation of 'up' refer to Appendix 3.

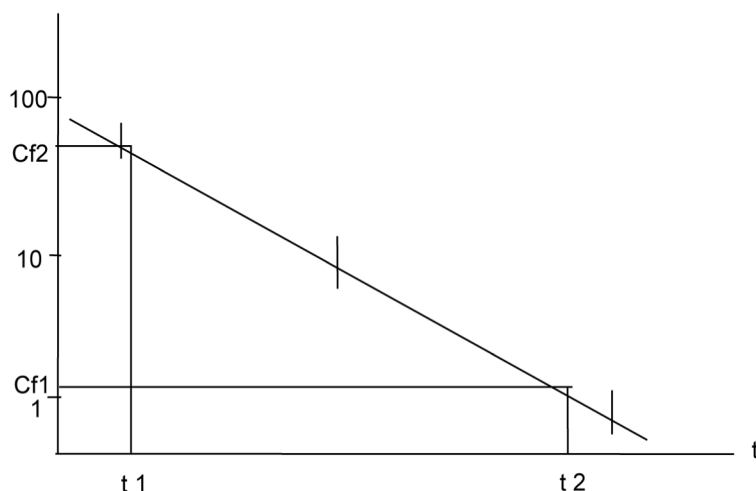
**▼ B***Appendix 5***Model discrimination**

Most bioconcentration data have been assumed to be ‘reasonably’ well described by a simple two-compartment/two-parameter model, as indicated by the rectilinear curve which approximates to the points for concentrations in fish, during the depuration phase, when these are plotted on semi-log paper. (Where these points cannot be described by a rectilinear curve then more complex models should be employed, see, for example, Spacie and Hamelink, Ref 1 in Appendix 3).

**Graphical method for determination of depuration (loss) rate constant  $k_2$** 

Plot the concentration of the test substance found in each sample of fish against sampling time on semi-log paper. The slope of that line is  $k_2$ .

$$k_2 = \frac{\ln (C_{f1}/C_{f2})}{t_2 - t_1}$$



Note that deviations from straight line may indicate a more complex depuration pattern than first order kinetics. A graphical method may be applied for resolving types of depuration deviating from first order kinetics.

**Graphical method for determination of uptake rate constant  $k_1$** 

Given  $K_2$ , calculate  $k_1$  as follows:

$$k_1 = \frac{C_f k_2}{C_w \times (1 - e^{-k_2 t})} \quad \text{[equation 1]}$$

The value of  $C_f$  is read from the midpoint of the smooth uptake curve produced by the data when log concentration is plotted versus time (on an arithmetical scale).

**▼B****Computer method for calculation of uptake and depuration (loss) rate constants**

The preferred means for obtaining the bioconcentration factor and  $k_1$  and  $k_2$  rate constants is to use non-linear parameter estimation methods on a computer. These programs find values for  $k_1$  and  $k_2$  given a set of sequential time concentration data and the model:

$$C_f = C_w \cdot \frac{k_1}{k_2} \times (1 - e^{-k_2 t}) \quad 0 < t < t_c \quad \text{[equation 2]}$$

$$C_f = C_w \cdot \frac{k_1}{k_2} \times (e^{-k_2(t-t_c)} - e^{-k_2 t}) \quad t > t_c \quad \text{[equation 3]}$$

where  $t_c$  = time at the end of the uptake phase.

This approach provides standard deviation estimates of  $k_1$  and  $k_2$ .

As  $k_2$  in most cases can be estimated from the depuration curve with relatively high precision, and because a strong correlation exists between the two parameters  $k_1$  and  $k_2$  if estimated simultaneously, it may be advisable first to calculate  $k_2$  from the depuration data only, and subsequently calculate  $k_1$  from the uptake data using non-linear regression.

**▼B****C.14. FISH JUVENILE GROWTH TEST****1. METHOD**

This growth toxicity test method is a replicate of the OECD TG 215 (2000).

**1.1. INTRODUCTION**

This test is designed to assess the effects of prolonged exposure to chemicals on the growth of juvenile fish. It is based on a method, developed and ring-tested (1)(2) within the European Union, for assessing the effects of chemicals on the growth of juvenile rainbow trout (*Oncorhynchus mykiss*) under flow-through conditions. Other well documented species may be used. For example, experience has been gained from growth tests with zebrafish (*Danio rerio*)<sup>(1)</sup> (3)(4) and ricefish (medaka, *Oryzias latipes*) (5)(6)(7).

See also General introduction Part C.

**1.2. DEFINITIONS**

**Lowest observed effect concentration (LOEC):** is the lowest tested concentration of a test substance at which the substance is observed to have a significant effect (at  $p < 0,05$ ) when compared with the control. However, all test concentrations above the LOEC must have a harmful effect equal to or greater than those observed at the LOEC.

**No observed effect concentration (NOEC):** is the test concentration immediately below the LOEC.

**EC<sub>x</sub>:** in this test method is the concentration of the test substance which causes a x % variation in growth rate of the fish when compared with controls.

**Loading rate:** is the wet weight of fish per volume of water.

**Stocking density:** is the number of fish per volume of water.

**Individual fish specific growth rate:** expresses the growth rate of one individual based on its initial weight.

**Tank-average specific growth rate:** expresses the mean growth rate of a tank population at one concentration.

**Pseudo specific growth rate:** expresses the individual growth rate compared to the mean initial weight of the tank population.

<sup>(1)</sup> Meyer, A., Bierman, C.H. and Orti, G. (1993). The phylogenetic position of the zebrafish (*Danio rerio*), a model system in developmental biology: an invitation to the comparative method. Proc. R. Soc. Lond. B. 252, 231-236.

**▼B**

## 1.3. PRINCIPLE OF THE TEST METHOD

Juvenile fish in exponential growth phase are placed, after being weighted, in test chambers and are exposed to a range of sublethal concentrations of the test substance dissolved in water preferably under flow-through, or, if not possible, under appropriate semi-static (static-renewal) conditions. The test duration is 28 days. Fish are fed daily. The food ration is based on initial fish weights and may be recalculated after 14 days. At the end of the test, the fish are weighed again. Effects on growth rates are analysed using a regression model in order to estimate the concentration that would cause a x % variation in growth rate, i.e.  $EC_x$  (e.g.  $EC_{10}$ ,  $EC_{20}$ , or  $EC_{30}$ ). Alternatively, the data may be compared with control values in order to determine the lowest observed effect concentration (LOEC) and hence the no observed effect concentration (NOEC).

## 1.4. INFORMATION ON THE TEST SUBSTANCE

Results of an acute toxicity test (see Test Method C. 1.) preferably performed with the species chosen for this test, should be available. This implies that the water solubility and the vapour pressure of the test substance are known and a reliable analytical method is available for the quantification of the substance in the test solutions with known and reported accuracy and limit of detection is available.

Useful information includes the structural formula, purity of the substance, stability in water and light,  $pK_a$ ,  $P_{ow}$  and results of a test for ready biodegradability (see Test Method C.4).

## 1.5. VALIDITY OF THE TEST

For the test to be valid the following conditions apply:

- the mortality in the control(s) must not exceed 10 % at the end of the test;
- the mean weight of fish in the control(s) must have increased enough to permit the detection of the minimum variation of growth rate considered as significant. A ring-test (2) has shown that for rainbow trout the mean weight of fish in the controls must have increased by at least the half (i.e. 50 %) of their mean initial weight over 28 days; e.g. initial weight: 1 g/fish (= 100 %), final weight after 28 days:  $\geq 1,5$  g/fish ( $\geq 150$  %);
- the dissolved oxygen concentration must have been at least 60 % of the air saturation value (ASV) throughout the test;
- the water temperature must not differ by more than  $\pm 1$  °C between test chambers at any one time during the test and should be maintained within a range of 2 °C within the temperature ranges specified for the test species (Appendix 1).

## 1.6. DESCRIPTION OF THE TEST METHOD

1.6.1. **Apparatus**

Normal laboratory equipment and especially the following:

- oxygen and pH meters;

**▼B**

- equipment for determination of water hardness and alkalinity;
  
- adequate apparatus for temperature control and preferably continuous monitoring;
  
- tanks made of chemically inert material and of suitable capacity in relation to the recommended loading and stocking density (see Section 1.8.5 and Appendix 1);
  
- suitably accurate balance (i.e. accurate to  $\pm 0,5$  %).

**1.6.2. Water**

Any water in which the test species shows suitable long-term survival and growth may be used as a test water. It should be of constant quality during the period of the test. The pH of the water should be within the range 6,5 to 8,5, but during a given test it should be within a range of  $\pm 0,5$  pH units. Hardness above 140 mg/l (as CaCO<sub>3</sub>) is recommended. In order to ensure that the dilution water will not unduly influence the test result (for example by complexation of test substance), samples should be taken at intervals for analysis. Measurements of heavy metals (e.g. Cu, Pb, Zn, Hg, Cd and Ni), major anions and cations (e.g. Ca, Mg, Na, K, Cl and SO<sub>4</sub>), pesticides (e.g. total organophosphorus and total organochlorine pesticides), total organic carbon and suspended solids should be made, for example, every three months where a dilution water is known to be relatively constant in quality. If water quality has been demonstrated to be constant over at least one year, determinations can be less frequent and intervals extended (e.g. every 6 months). Some chemical characteristics of an acceptable dilution water are listed in Appendix 2.

**1.6.3. Test solutions**

Test solutions of the chosen concentrations are prepared by dilution of a stock solution.

The stock solution should preferably be prepared by simply mixing or agitating the test substance in the dilution water by using mechanical means (e.g. stirring or ultrasonication). Saturation columns (solubility columns) can be used for achieving a suitable concentrated stock solution.

The use of solvents or dispersants (solubilising agents) may be required in some cases in order to produce a suitably concentrated stock solution. Examples of suitable solvents are acetone, ethanol, methanol, dimethylsulfoxide, dimethylformamide and triethyleneglycol. Examples of suitable dispersants are Cremophor RH40, Tween 80, Methylcellulose 0,01 % and HCO-40. Care should be taken when using readily biodegradable agents (e.g. acetone) and/or highly volatile compounds as these can cause problems with bacterial built-up in flow-through tests. When a solubilising agent is used it must have no significant effects on the fish growth nor visible adverse effects on the juvenile as revealed by a solvent-only control.



**▼B**

For flow-through tests, a system which continually dispenses and dilutes a stock solution of the test substance (e.g. metering pump, proportional diluter, saturator system) is required to deliver a series of concentrations to the test chambers. The flow rates of stock solutions and dilution water should be checked at intervals, preferably daily, during the test and should not vary by more than 10 % throughout the test. A ring-test (2) has shown that, for rainbow trout, a frequency of water removal during the test of six litres/g of fish/day is acceptable (see Section 1.8.2.2).

For semi-static (renewal) tests, the frequency of medium renewal will depend on the stability of the test substance, but a daily water renewal is recommended. If, from preliminary stability tests (see Section 1.4), the test substance concentration is not stable (i.e. outside the range 80-120 % of nominal or falling below 80 % of the measured initial concentration) over the renewal period, consideration should be given to the use of a flow-through test.

**1.6.4. Selection of species**

Rainbow trout (*Oncorhynchus mykiss*) is the recommended species for this test since most experience has been gained from ring-test with this species (1)(2). However, other well documented species can be used but the test procedure may have to be adapted to provide suitable test conditions. For example, experience is also available with zebrafish (*Danio rerio*) (3)(4) and ricefish (medaka, *Oryzias latipes*) (5)(6)(7). The rationale for the selection of the species and the experimental method should be reported in this case.

**1.6.5. Holding of fish**

The test fish shall be selected from a population of a single stock, preferably from the same spawning, which has been held for at least two weeks prior to the test under conditions of water quality and illumination similar to those used in the test. They should be fed a minimum ration of 2 % body weight per day and preferably 4 % body weight per day throughout the holding period and during the test.

Following a 48 h setting-in period, mortalities are recorded and the following criteria applied:

- mortalities of greater than 10 % of population in seven days: reject the entire batch;
- mortalities of between 5 % and 10 % of population: acclimation for seven additional days; if more than 5 % mortality during second seven days, reject the entire batch;
- mortalities of less than 5 % of population in seven days: accept the batch.

Fish should not receive treatment for disease in the two weeks preceding the test, or during the test.

**▼ B**

## 1.7. TEST DESIGN

The 'test design' relates to the selection of the number and spacing of the test concentrations, the number of tanks at each concentration level and the number of fish per tank. Ideally, the test design should be chosen with regard to:

- the objective of the study;
- the method of statistical analysis that will be used;
- the availability and cost of experimental resources.

The statement of the objective should, if possible, specify the statistical power at which a given size of difference (e.g. in growth rate) is required to be detected or, alternatively, the precision with which the  $EC_x$  (e.g. with  $x = 10, 20$ , or  $30$ , and preferably not less than  $10$ ) is required to be estimated. Without this, a firm prescription of the size of the study cannot be given.

It is important to recognise that a design which is optimal (makes best use of resources) for use with one method of statistical analysis is not necessarily optimal for another. The recommended design for the estimation of a LOEC/NOEC would not therefore be the same as that recommended for analysis by regression.

In most of cases, regression analysis is preferable to the analysis of variance, for reasons discussed by Stephan and Rogers (8). However, when no suitable regression model is found ( $r^2 < 0,9$ ) NOEC/LOEC should be used.

1.7.1. **Design for analysis by regression**

The important considerations in the design of a test to be analysed by regression are:

- The effect concentration (e.g.  $EC_{10,20,30}$ ) and the concentration range over which the effect of the test substance is of interest, should necessarily be spanned by the concentrations included in the test. The precision with which estimates of effect concentrations can be made, will be best when the effect concentration is in the middle of the range of concentrations tested. A preliminary range-finding test may be helpful in selecting appropriate test concentrations.
- To enable satisfactory statistical modelling, the test should include at least one control tank and five additional tanks at different concentrations. Where appropriate, when a solubilising agent is used, one control containing the solubilising agent at the highest tested concentration should be run in addition to the test series (see Sections 1.8.3 and 1.8.4).
- An appropriate geometric series or logarithmic series (9) (see Appendix 3) may be used. Logarithmic spacing of test concentration is to be preferred.

**▼B**

- If more than six tanks are available, the additional tanks should either be used to provide replication or distributed across the range of concentrations in order to enable closer spacing of the levels. Either of these measures are equally desirable.

**1.7.2. Design for estimation of a NOEC/LOEC using Analysis of Variance (ANOVA)**

There should preferably be replicate tanks at each concentration, and statistical analysis should be at the tank level (10). Without replicate tanks, no allowance can be made for variability between tanks beyond that due to individual fish. However, experience has shown (11) that between-tank variability was very small compared with within-tank (i.e. between-fish) variability in the case examined. Therefore a relatively acceptable alternative is to perform statistical analysis at the level of individual fish.

Conventionally, at least five test concentrations in a geometric series with a factor preferably not exceeding 3,2 are used.

Generally, when tests are performed with replicate tanks, the number of replicate control tanks and therefore the number of fish should be the double of the number in each of the test concentrations, which should be of equal size (12)(13)(14). On the opposite, in absence of replicate tanks, the number of fish in the control group should be the same as the number in each test concentration.

If the ANOVA is to be based on tanks rather than individual fish (which would entail either individual marking of the fish or the use of 'pseudo' specific growth rates (see Section 2.1.2)), there is a need for enough replication of tanks to enable the standard deviation of 'tanks-within-concentrations' to be determined. This means that the degrees of freedom for error in the analysis of variance should be at least 5 (10). If only the controls are replicated, there is a danger that the error variability will be biased because it may increase with the mean value of the growth rate in question. Since growth rate is likely to decrease with increasing concentration, this will tend to lead to an overestimate of the variability.

**1.8. PROCEDURE****1.8.1. Selection and weighing of test fish**

It is important to minimise variation in weight of the fish at the beginning of the test. Suitable size ranges for the different species recommended for use in this test are given in Appendix 1. For the whole batch of fish used in the test, the range in individual weights at the start of the test should ideally be kept to within  $\pm 10\%$  of the arithmetic mean weight and, in any case, should not exceed 25%. It is recommended to weigh a subsample of fish before the test in order to estimate the mean weight.

**▼B**

Food should be withheld from the stock population for 24 h prior to the start of the test. Fish should then be chosen at random. Using a general anaesthetic (e.g. an aqueous solution of 100 mg/l tricaine methane sulphonate (MS 222) neutralised by the addition of two parts of sodium bicarbonate per part of MS 222), fish should be weighted individually as wet weights (blotted dry) to the precision given in Appendix 1. Those fish with weights within the intended range should be retained and then should be randomly distributed between the test vessels. The total wet weight of fish in each test vessel should be recorded. The use of anaesthetics likewise handling of fish (including blotting and weighing) may cause stress and injuries to the juvenile fish, in particular for those species of small size. Therefore handling of juvenile fish must be done with the utmost care to avoid stressing and injuring test animals.

The fish are weighed again on day 28 of the test (see Section 1.8.6). However, if it is deemed necessary to recalculate the food ration, fish can be weighed again on day 14 of the test (see Section 1.8.2.3). Other method as photographic method could be used to determine changes in fish size from which food rations could be adjusted.

## 1.8.2. **Conditions of exposure**

### 1.8.2.1. *Duration*

The test duration is  $\geq 28$  days.

### 1.8.2.2. *Loading rates and stocking densities*

It is important that the loading rate and stocking density is appropriate for the test species used (see Appendix 1). If the stocking density is too high, then overcrowding stress will occur leading to reduced growth rates and possibly to disease. If it is too low, territorial behaviour may be induced which could also affect growth. In any case, the loading rate should be low enough in order that a dissolved oxygen concentration of at least 60 % ASV can be maintained without aeration. A ring-test (2) has shown that, for rainbow trout, a loading rate of 16 trouts of 3-5 g in a 40-litre volume is acceptable. Recommended frequency of water removal during the test is 6 litres/g of fish/day.

### 1.8.2.3. *Feeding*

The fish should be fed with an appropriate food (Appendix 1) at a sufficient rate to induce acceptable growth rate. Care should be taken to avoid microbial growth and water turbidity. For rainbow trout, a rate of 4 % of their body weight per day is likely to satisfy these conditions (2)(15)(16)(17). The daily ration may be divided into two equal portions and given to the fish in two feeds per day, separated by at least 5 h. The ration is based on the initial total fish weight for each test vessel. If the fish are weighted again on day 14, the ration is then recalculated. Food should be withheld from the fish 24 h prior to weighing.

**▼B**

Uneaten food and fecal material should be removed from the test vessels each day by carefully cleaning the bottom of each tank using a suction.

**1.8.2.4. Light and temperature**

The photoperiod and water temperature should be appropriate for the test species (Appendix 1).

**1.8.3. Test concentrations**

Normally five concentrations of the test substance are required, regardless of the test design (see Section 1.7.2). Prior knowledge of the toxicity of the test substance (e.g. from an acute test and/or from range-finding studies) should help in selecting appropriate test concentrations. Justification should be given if fewer than five concentrations are used. The highest tested concentration should not exceed the substance solubility limit in water.

Where a solubilising agent is used to assist in stock solution preparation, its final concentration should not be greater than 0,1 ml/l and should preferably be the same in all test vessels (see Section 1.6.3). However, every effort should be made to avoid use of such materials.

**1.8.4. Controls**

The number of dilution-water controls depends on the test design (see Sections 1.7-1.7.2). If a solubilising agent is used, then the same number of solubilising-agent controls as dilution-water controls should also be included.

**1.8.5. Frequency of analytical determinations and measurements**

During the test, the concentrations of test substance are determined at regular intervals (see below).

In flow-through tests, the flow rates of diluent and toxicant stock solution should be checked at intervals, preferably daily, and should not vary by more than 10 % throughout the test. Where the test substance concentrations are expected to be within  $\pm 20$  % of the nominal values (i.e. within the range 80-120 %; see Sections 1.6.2 and 1.6.3), it is recommended that, as a minimum, the highest and lowest test concentrations be analysed at the start of the test and at weekly intervals thereafter. For the test where the concentration of the test substance is not expected to remain within  $\pm 20$  % of nominal (on the basis of stability data of the test substance), it is necessary to analyse all test concentrations, but following the same regime.

In semi-static (renewal) tests where the concentration of the test substance is expected to remain within  $\pm 20$  % of the nominal values, it is recommended that, as a minimum, the highest and lowest test concentrations be analysed when freshly prepared and immediately prior to renewal at the start of the study and weekly thereafter. For tests where the concentration of the test substance is not expected to remain within  $\pm 20$  % of nominal, all test concentrations must be analysed following the same regime as for more stable substances.

**▼B**

It is recommended that results be based on measured concentrations. However, if evidence is available to demonstrate that the concentration of the test substance in solution has been satisfactorily maintained within  $\pm 20\%$  of the nominal or measured initial concentration throughout the test, then the results can be based on nominal or measured values.

Samples may need to be filtered (e.g. using a 0,45  $\mu\text{m}$  pore size) or centrifuged. Centrifugation is the recommended procedure. However, if the test material does not adsorb to filters, filtration may also be acceptable.

During the test, dissolved oxygen, pH and temperature should be measured in all test vessels. Total hardness, alkalinity and salinity (if relevant) should be measured in the controls and one vessel at the highest concentration. As a minimum, dissolved oxygen and salinity (if relevant) should be measured three times (at the beginning, middle and end of the test). In semi-static tests, it is recommended that dissolved oxygen be measured more frequently, preferably before and after each water renewal or at least once a week. pH should be measured at the beginning and end of each water renewal in static renewal test and at least weekly in flow-through tests. Hardness and alkalinity should be measured once each test. Temperature should preferably be monitored continuously in at least one test vessel.

**1.8.6. Observations**

Weight: at the end of the test all surviving fish must be weighed as wet weights (blotted dry) either in groups by test vessel or individually. Weighing of animals by test vessel is preferred to individual weights which require that fish be individually marked. In the case of the measurement of individual weights for determination of individual fish specific growth rate, the marking technique selected should avoid stressing the animals (alternatives to freeze marking may be appropriate, e.g. the use of coloured fine fishing line).

The fish should be examined daily during the test period and any external abnormalities (such as hemorrhage, discoloration) and abnormal behaviour noted. Any mortalities should be recorded and the dead fish removed as soon as possible. Dead fish are not replaced, the loading rate and stocking density being sufficient to avoid effects on growth through changes in number of fish per tank. However, the feeding rate will need to be adjusted.

**2. DATA AND REPORTING****2.1. TREATMENT OF RESULTS**

It is recommended that a statistician be involved in both the design and analysis of the test since this test method allows for considerable variation in experimental design as for example, in the number of test chambers, number of test concentrations, number of fish, etc. In view of the options available in test design, specific guidance on statistical procedure is not given here.

**▼ B**

Growth rates should not be calculated for test vessels where the mortality exceeds 10 %. However, mortality rate should be indicated for all test concentrations.

Whichever method is used to analyse the data, the central concept is the specific growth rate  $r$  between time  $t_1$  and time  $t_2$ . This can be defined in several ways depending on whether fish are individually marked or not or whether a tank average is required.

$$r_1 = \frac{\log_e w_2 - \log_e w_1}{t_2 - t_1} \times 100$$

$$r_2 = \frac{\log_e w_2 - \log_e w_1}{t_2 - t_1} \times 100$$

$$r_3 = \frac{\log_e w_2 - \log_e w_1}{t_2 - t_1} \times 100$$

where:

$r_1$  = individual fish specific growth rate

$r_2$  = tank-average specific growth rate

$r_3$  = 'pseudo' specific growth rate

$w_1, w_2$  = weights of a particular fish at times  $t_1$  and  $t_2$ , respectively

$\log_e w_1$  = logarithm of the weight of an individual fish at the start of the study period

$\log_e w_2$  = logarithm of the weight of an individual fish at the end of the study period

$\log_e w_1$  = average of the logarithms of the values  $w_1$  for the fish in the tank at the start of the study period

$\log_e w_2$  = average of the logarithms of the values  $w_2$  for the fish in the tank at the end of the study period

$t_1, t_2$  = time (days) at start and end of study period

$r_1, r_2, r_3$  can be calculated for the 0-28 days period and, where appropriate (i.e. when measurement at day 14 has been done) for the 0-14 and 14-28 days periods.

### 2.1.1. Analysis of results by regression (concentration-response modelling)

This method of analysis fits a suitable mathematical relationship between the specific growth rate and concentration, and hence enables the estimation of the 'EC<sub>x</sub>' i.e. any required EC value. Using this method the calculation of  $r$  for individual fish ( $r_1$ ) is not necessary and instead, the analysis can be based on the tank-average value of  $r$  ( $r_2$ ). This last method is preferred. It is also more appropriate in case of the use of smallest species.

The tank-average specific growth rates ( $r_2$ ) should be plotted graphically against concentration, in order to inspect the concentration response relationship.

**▼ B**

For expressing the relationship between  $r_2$  and concentration, an appropriate model should be chosen and its choice must be supported by appropriate reasoning.

If the numbers of fish surviving in each tank are unequal, then the process of model fitting, whether simple or non-linear, should be weighted to allow for unequal sizes of groups.

The method of fitting the model must enable an estimate of, for example, the  $EC_{20}$  and of its dispersion (either standard error or confidence interval) to be derived. The graph of the fitted model should be shown in relation to the data so that the adequacy of the fit of the model can be seen (8)(18)(19)(20).

### 2.1.2. Analysis of results for the estimation of the LOEC

If the test has included replication of tanks at all concentration levels, the estimation of the LOEC could be based on an analysis of variance (ANOVA) of the tank-average specific growth rate (see Section 2.1), followed by a suitable method (e.g. Dunnett's or Williams' test (12)(13)(14)(21)) of comparing the average  $r$  for each concentration with the average  $r$  for the controls to identify the lowest concentration for which this difference is significant at a 0,05 probability level. If the required assumptions for parametric methods are not met — non-normal distribution (e.g. Shapiro-Wilk's test) or heterogeneous variance (Bartlett's test), consideration should be given to transforming the data to homogenise variances prior to performing the ANOVA, or to carrying out a weighted ANOVA.

If the test has not included replication of tanks at each concentration, an ANOVA based on tanks will be insensitive or impossible. In this situation, an acceptable compromise is to base the ANOVA on the 'pseudo' specific growth rate  $r_3$  for individual fish.

The average  $r_3$  for each test concentration may then be compared with the average  $r_3$  for the controls. The LOEC can then be identified as before. It must be recognised that this method provides no allowance for, nor protection against, variability between tanks, beyond that which is accounted for by the variability between individual fish. However, experience has shown (8) that between-tank variability was very small compared with within-tank (i.e. between fish) variability. If individual fish are not included in the analysis, the method of outlier identification and justification for its use must be provided.

## 2.2. INTERPRETATION OF RESULTS

The results should be interpreted with caution where measured toxicant concentrations in test solutions occur at levels near the detection limit of the analytical method or, in semi static tests, when the concentration of the test substance decreases between freshly prepared solution and before renewal.

## 2.3. TEST REPORT

The test report must include the following information:



**▼B**

- 2.3.1. **Test substance:**
- physical nature and relevant physical-chemical properties;
  - chemical identification data including purity and analytical method for quantification of the test substance where appropriate.
- 2.3.2. **Test species:**
- scientific name, possibly
  - strain, size, supplier, any pre-treatment, etc.
- 2.3.3. **Test conditions:**
- test procedure used (e.g. semi-static/renewal, flow-through, loading, stocking density, etc.),
  - test design (e.g. number of test vessels, test concentrations and replicates, number of fish per vessel),
  - method of preparation of stock solutions and frequency of renewal (the solubilising agent and its concentration must be given, when used),
  - the nominal test concentrations, the means of the measured values and their standard deviations in the test vessels and the method by which these were attained and evidence that the measurements refer to the concentrations of the test substance in true solution,
  - dilution water characteristics: pH, hardness, alkalinity, temperature, dissolved oxygen concentration, residual chlorine levels (if measured), total organic carbon, suspended solids, salinity of the test medium (if measured) and any other measurements made,
  - water quality within test vessels: pH, hardness, temperature and dissolved oxygen concentration,
  - detailed information on feeding, (e.g. type of food(s), source, amount given and frequency).
- 2.3.4. **Results:**
- evidence that controls met the validity criterion for survival, and data on mortalities occurring in any of the test concentrations,
  - statistical analytical techniques used, statistics based on replicates or fish, treatment of data and justification of techniques used,
  - tabulated data on individual and mean fish weights on days 0, 14 (if measured) and 28 values of tank-average or pseudo specific growth rates (as appropriate) for the periods 0-28 days or possibly 0-14 and 14-28,
  - results of the statistical analysis (i.e. regression analysis or ANOVA) preferably in tabular and graphical form and the LOEC ( $p = 0,05$ ) and the NOEC or EC<sub>x</sub> with, when possible, standard errors, as appropriate,

**▼B**

— incidence of any unusual reactions by the fish and any visible effects produced by the test substance.

## 3. REFERENCES

- (1) Solbe J.F. de LG (1987) Environmental Effects of Chemicals (CFM 9350 SLD). Report on a UK Ring Test of a Method for Studying the Effects of Chemicals on the Growth rate of Fish. WRc Report No PRD 1388-M/2.
- (2) Ashley S., Mallett M.J. and Grandy N.J., (1990) EEC Ring Test of a Method for Determining the Effects of Chemicals on the Growth Rate of Fish. Final Report to the Commission of the European Communities. WRc Report No EEC 2600-M.
- (3) Crossland N.O. (1985) A method to evaluate effects of toxic chemicals on fish growth. *Chemosphere*, 14, p. 1855-1870.
- (4) Nagel R., Bresh H., Caspers N., Hansen P.D., Market M., Munk R., Scholz N. and Höfte B.B., (1991) Effect of 3,4-dichloroaniline on the early life stages of the Zebrafish (*Brachydanio rerio*): results of a comparative laboratory study. *Ecotox. Environ. Safety*, 21, p. 157-164.
- (5) Yamamoto, Tokio., (1975) Series of stock cultures in biological field. Medaka (killifish) biology and strains. Keigaku Publish. Tokio, Japan.
- (6) Holcombe, G.W., Benoit D.A., Hammermeister, D.E., Leonard, E.N. and Johnson, R.D., (1995) Acute and long-term effects of nine chemicals on the Japanese medaka (*Oryzias latipes*). *Arch. Environ. Conta. Toxicol.* 28, p. 287-297.
- (7) Benoit, D.A., Holcombe, G.W. and Spehar, R.L., (1991) Guidelines for conducting early life toxicity tests with Japanese medaka (*Oryzias latipes*). Ecological Research Series EPA-600/3-91-063. U.S. Environmental Protection Agency, Duluth, Minesota.
- (8) Stephan C.E. and Rogers J.W., (1985) Advantages of using regression analysis to calculate results of chronic toxicity tests. *Aquatic Toxicology and Hazard Assessment: Eighth Symposium, ASTM STP 891*, R C Bahner and D J Hansen, Eds., American Society for Testing and Materials, Philadelphia, p. 328-338.
- (9) Environment Canada, (1992) Biological test method: toxicity tests using early life stages of salmonid fish (rainbow trout, coho salmon, or Atlantic salmon). Conservation and Protection, Ontario, Report EPS 1/RM/28, p. 81.
- (10) Cox D.R., (1958) *Planning of experiments*. Wiley Edt.
- (11) Pack S., (1991) Statistical issues concerning the design of tests for determining the effects of chemicals on the growth rate of fish. Room Document 4, OECD Ad Hoc Meeting of Experts on Aquatic Toxicology, WRc Medmenham, UK, 10-12 December 1991.
- (12) Dunnett C.W., (1955) A Multiple Comparisons Procedure for Comparing Several Treatments with a Control, *J. Amer. Statist. Assoc.*, 50, p. 1096-1121.
- (13) Dunnett C.W., (1964) New tables for multiple comparisons with a control. *Biometrics*, 20, p. 482-491.

**▼B**

- (14) Williams D.A., (1971) A test for differences between treatment means when several dose levels are compared with a zero dose control. *Biometrics* 27, p. 103-117.
- (15) Johnston, W.L., Atkinson, J.L., Glanville N.T., (1994) A technique using sequential feedings of different coloured food to determine food intake by individual rainbow trout, *Oncorhynchus mykiss*: effect of feeding level. *Aquaculture* 120, p. 123-133.
- (16) Quinton, J. C. and Blake, R.W., (1990) The effect of feed cycling and ration level on the compensatory growth response in rainbow trout, *Oncorhynchus mykiss*. *Journal of Fish Biology*, 37, p. 33-41.
- (17) Post, G., (1987) Nutrition and Nutritional Diseases of Fish. Chapter IX in *Testbook of Fish Health*. T.F.H. Publications, Inc. Neptune City, New Jersey, USA. p. 288.
- (18) Bruce, R.D. and Versteeg D.J., (1992) A statistical procedure for modelling continuous toxicity data. *Environ. Toxicol. Chem.* 11, p. 1485-1494.
- (19) DeGraeve, G.M., Cooney, J.M., Pollock, T.L., Reichenbach, J.H., Dean, Marcus, M.D. and McIntyre, D.O., (1989) Precision of EPA seven-day fathead minnow larval survival and growth test; intra and interlaboratory study. report EA-6189 (American Petroleum Institute Publication, No 4468). Electric Power Research Institute, Palo alto, CA.
- (20) Norbert-King T.J., (1988) An interpolation estimate for chronic toxicity: the ICp approach. US Environmental Protection Agency. Environmental Research Lab., Duluth, Minesota. Tech. Rep. No 05-88 of National Effluent Toxicity Assesment Center. Sept. 1988. p. 12.
- (21) Williams D.A., (1972) The comparison of several dose levels with a zero dose control. *Biometrics* 28, p. 510-531.

## FISH SPECIES RECOMMENDED FOR TESTING AND SUITABLE TEST CONDITIONS

Species	Recommended test temperature range (°C)	Photoperiod (hours)	Recommended range for initial fish weight (g)	Required measurement precision	Loading rate (g/l)	Stocking density (per litre)	Food	Test duration (days)
<b>Recommended species:</b> <i>Oncorhynchus mykiss</i> rainbow trout	12,5-16,0	12-16	1-5	to nearest 100 mg	1,2-2,0	4	Dry proprietary salmonid fry food	≥ 28
<b>Other well documented species:</b> <i>Danio rerio</i> zebrafish	21-25	12-16	0,050-0,100	to nearest 1 mg	0,2-1,0	5-10	Live food ( <i>Brachionus Artemia</i> )	≥ 28
<i>Oryzias latipes</i> ricefish (Medaka)	21-25	12-16	0,050-0,100	to nearest 1 mg	0,2-1,0	5-20	Live food ( <i>Brachionus Artemia</i> )	≥ 28

**▼B***Appendix 2***SOME CHEMICAL CHARACTERISTICS OF AN ACCEPTABLE  
DILUTION WATER**

Substance	Concentrations
Particulate matter	< 20 mg/l
Total organic carbon	< 2 mg/l
Unionised ammonia	< 1 µg/l
Residual chlorine	< 10 µg/l
Total organophosphorus pesticides	< 50 ng/l
Total organochlorine pesticides plus polychlorinated biphenyls	< 50 ng/l
Total organic chlorine	< 25 ng/l



## Appendix 3

## Logarithmic series of concentrations suitable for toxicity test (9)

Column (Number of concentrations between 100 and 10, or between 10 and 1) (*)						
1	2	3	4	5	6	7
100	100	100	100	100	100	100
32	46	56	63	68	72	75
10	22	32	40	46	52	56
3,2	10	18	25	32	37	42
1,0	4,6	10	16	22	27	32
	2,2	5,6	10	15	19	24
	1,0	3,2	6,3	10	14	18
		1,8	4,0	6,8	10	13
		1,0	2,5	4,6	7,2	10
			1,6	3,2	5,2	7,5
			1,0	2,2	3,7	5,6
				1,5	2,7	4,2
				1,0	1,9	3,2
					1,4	2,4
					1,0	1,8
						1,3
						1,0

(\*) A series of five (or more) successive concentrations may be chosen from a column. Mid-points between concentrations in column (x) are found in column (2x + 1). The values listed can represent concentrations expressed as percentage per volume or weight (mg/l or µg/l). Values can be multiplied or divided by any power of 10 as appropriate. Column 1 might be used if there was considerable uncertainty on the toxicity level.

**▼B****C.15. FISH, SHORT-TERM TOXICITY TEST ON EMBRYO AND SAC-FRY STAGES****1. METHOD**

This short-term toxicity test method is a replicate of the OECD TG 212 (1998).

**1.1. INTRODUCTION**

This short-term toxicity test on Fish Embryo and Sac-Fry stages is a short-term test in which the life stages from the newly fertilised egg to the end of the sac-fry stage are exposed. No feeding is provided in the embryo and sac-fry test, and the test should thus be terminated while the sac-fry are still nourished from the yolk-sac.

The test is intended to define lethal, and to a limited extent, sub-lethal effects of chemicals on the specific stages and species tested. This test would provide useful information in that it could (a) form a bridge between lethal and sub-lethal tests, (b) be used as a screening test for either a Full Early Life Stage test or for chronic toxicity tests and (c) be used for testing species where husbandry techniques are not sufficiently advanced to cover the period of change from endogenous to exogenous feeding.

It should be borne in mind that only tests incorporating all stages of the life-cycle of fish are generally liable to give an accurate estimate of the chronic toxicity of chemicals to fish, and that any reduced exposure with respect to life stages may reduce the sensitivity and thus underestimate the chronic toxicity. It is therefore expected that the embryo and sac-fry test would be less sensitive than a Full Early Life Stage test, particularly with respect to chemicals with high lipophilicity ( $\log P_{ow} > 4$ ) and chemicals with a specific mode of toxic action. However smaller differences in sensitivity between the two tests would be expected for chemicals with a non-specific, narcotic mode of action (1).

Prior to the publication of this test, most experience with this embryo and sac-fry test has been with the freshwater fish *Danio rerio* Hamilton-Buchanan (Teleostei, Cyprinidae — common name zebrafish). More detailed guidance on test performance for this species is therefore given in Appendix 1. This does not preclude the use of other species for which experience is also available (Table 1).

**1.2. DEFINITIONS**

**Lowest Observed Effect Concentration (LOEC):** is the lowest tested concentration of a test substance at which the substance is observed to have a significant effect (at  $p < 0,05$ ) when compared with the control. However, all test concentrations above the LOEC must have a harmful effect equal to or greater than those observed at the LOEC.

**No Observed Effect Concentration (NOEC):** is the test concentration immediately below the LOEC.

**▼ B**

## 1.3. PRINCIPLE OF THE TEST

The embryo and sac-fry stages of fish are exposed to a range of concentrations of the test substance dissolved in water. Within the protocol a choice is possible between a semi-static and a flow-through procedure. The choice depends on the nature of the test substance. The test is begun by placing fertilised eggs in the test chambers and is terminated just before the yolk-sac of any larvae in any of the test chambers has been completely absorbed or before mortalities by starvation start in controls. Lethal and sub-lethal effects are assessed and compared with control values to determine the lowest observed effect concentration and hence the no observed effect concentration. Alternatively, they may be analysed using a regression model in order to estimate the concentration that would cause a given percentage effect (i.e. LC/EC<sub>x</sub>, where x is a defined % effect).

## 1.4. INFORMATION ON THE TEST SUBSTANCE

Results of an acute toxicity test (see Method C. 1) preferably performed with the species chosen for this test, should be available. The results may be useful in selecting an appropriate range of test concentrations in the early life stages test. Water solubility (including solubility in the test water) and the vapour pressure of the test substance should be known. A reliable analytical method for the quantification of the substance in the test solutions with known and reported accuracy and limit of detection should be available.

Information on the test substance which is useful in establishing the test conditions includes the structural formula, purity of the substance, stability in light, stability under the conditions of the test, pK<sub>a</sub>, P<sub>ow</sub> and results of a test for ready biodegradability (see Method C. 4).

## 1.5. VALIDITY OF THE TEST

For a test to be valid, the following conditions apply:

- overall survival of fertilised eggs in the controls and where relevant, in the solvent-only vessels must be greater than or equal to the limits defined in Appendices 2 and 3
- the dissolved oxygen concentration must be between 60 and 100 % of the air saturation value (ASV) throughout the test
- the water temperature must not differ by more than  $\pm 1,5$  °C between test chambers or between successive days at any time during the test and should be within the temperature ranges specified for the test species (Appendices 2 and 3).



**▼B**

## 1.6. DESCRIPTION OF THE TEST METHOD

1.6.1. **Test chambers**

Any glass or other chemically inert vessels can be used. The dimensions of the vessels should be large enough to allow compliance with the loading rate (see Section 1.7.1.2). It is recommended that test chambers be randomly positioned in the test area. A randomised block design with each treatment being present in each block is preferable to a completely randomised design when there are systematic effects in the laboratory that can be controlled using blocking. Blocking, if used, should be taken account of in the subsequent data analysis. The test chambers should be shielded from unwanted disturbance.

1.6.2. **Selection of fish species**

Recommended fish species are given in Table 1A. This does not preclude the use of other species (examples are given in Table 1B), but the test procedure may have to be adapted to provide suitable test conditions. The rationale for the selection of the species and the experimental method should be reported in this case.

1.6.3. **Holding of the brood fish**

Details on holding the brood stock under satisfactory conditions may be found in OECD TG 210 <sup>(1)</sup> and in references (2)(3)(4)(5)(6).

1.6.4. **Handling of embryos and larvae**

Embryos and larvae may be exposed, within the main vessel, in smaller vessels fitted with mesh sides or ends to permit a flow of test solution through the vessel. Non-turbulent flow through these small vessels may be induced by suspending them from an arm arranged to move the vessel up and down but always keeping the organisms submerged; a siphon-flush system can also be used. Fertilised eggs of salmonid fishes can be supported on racks or meshes with apertures sufficiently large to allow larvae to drop through after hatching. The use of pasteur pipettes is appropriate to remove the embryos and larvae in the semi-static tests with complete daily renewal (see paragraph 1.6.6)

Where egg containers, grids or meshes have been used to hold eggs within the main test vessel, these restraints should be removed after the larvae hatch <sup>(1)</sup>, except that meshes should be retained to prevent the escape of the fish. If there is a need to transfer the larvae, they should not be exposed to the air and nets should not be used to release fish from egg containers (such a caution may not be necessary for some less fragile species, e.g. the carp). The timing of this transfer varies with the species and transfer may not always be necessary. For the semi-static technique, beakers or shallow containers may be used, and, if necessary, equipped with a mesh screen slightly elevated above the bottom of the beaker. If the volume of these containers is sufficient to comply with loading requirements, (see 1.7.1.2) no transfer of embryo or larvae may be necessary.

<sup>(1)</sup> OECD, Paris, 1992, Test Guideline 210, 'Fish, Early-life Stage Toxicity Test'.

**▼B****1.6.5. Water**

Any water which conforms to the chemical characteristics of an acceptable dilution water as listed in Appendix 4 and in which the test species shows control survival at least as good as that described in Appendices 2 and 3 is suitable as a test water. It should be of constant quality during the period of the test. The pH should remain within a range of  $\pm 0,5$  pH units. In order to ensure that the dilution water will not unduly influence the test result (for example by complexation of test substance), or adversely affect the performance of the brood stock, samples should be taken at intervals for analysis. Measurements of heavy metals (e.g. Cu, Pb, Zn, Hg, Cd and Ni), major anions and cations (e.g. Ca, Mg, Na, K, Cl and  $\text{SO}_4$ ), pesticides (e.g. total organophosphorus and total organochlorine pesticides), total organic carbon and suspended solids should be made, for example, every three months, where a dilution water is known to be relatively constant in quality. If water quality has been demonstrated to be constant over at least one year, determinations can be less frequent and intervals extended (e.g. every six months).

**1.6.6. Test solutions**

Test solutions of the chosen concentrations are prepared by dilution of a stock solution.

The stock solution should preferably be prepared by simply mixing or agitating the test substance in the dilution water by using mechanical means (e.g. stirring and ultrasonication). Saturation columns (solubility columns) can be used for achieving a suitable concentrated stock solution. As far as possible, the use of solvents or dispersants (solubilising agents) should be avoided; however, such compounds may be required in some cases in order to produce a suitably concentrated stock solution. Examples of suitable solvents are acetone, ethanol, methanol, dimethylformamide and triethyleneglycol. Examples of suitable dispersants are Cremophor RH40, Tween 80, methylcellulose 0,01 % and HCO-40. Care should be taken when using readily biodegradable agents (e.g. acetone) and/or highly volatile as these can cause problems with bacterial built-up in flow-through tests. When a solubilising agent is used it must have no significant effect on survival nor visible adverse effect on the early-life stages as revealed by a solvent-only control. However, every effort should be made to avoid the use of such materials.

For the semi-static technique, two different renewal procedures may be followed; either (i) new test solutions are prepared in clean vessels and surviving eggs and larvae gently transferred into the new vessels in a small volume of old solution, avoiding exposure to air, or (ii) the test organisms are retained in the vessels whilst a proportion (at least three-quarters) of the test water is changed. The frequency of medium renewal will depend on the stability of the test substance, but a daily water renewal is recommended. If, from preliminary stability tests (see Section 1.4), the test substance concentration is not stable (i.e. outside the range 80-120 % of nominal or falling below 80 % of the measured initial concentration) over the renewal period, consideration should be given to the use of a flow-through test. In any case, care should be taken to avoid stressing the larvae during the water renewal operation.

**▼B**

For flow-through tests, a system which continually dispenses and dilutes a stock solution of the test substance (e.g. metering pump, proportional diluter, saturator system) is required to deliver a series of concentrations to the test chambers. The flow rates of stock solutions and dilution water should be checked at intervals, preferably daily, and should not vary by more than 10 % throughout the test. A flow rate equivalent to at least five test chamber volumes per 24 hours has been found suitable (2).

**1.7. PROCEDURE**

Useful information on the performance of fish embryo and sac-fry toxicity tests is available in the literature, some examples of which are included in the literature section of this text (7)(8)(9).

**1.7.1. Conditions of exposure****1.7.1.1. Duration**

The test should start preferably within 30 minutes after the eggs have been fertilised. The embryos are immersed in the test solution before, or as soon as possible after, commencement of the blastodisc cleavage stage and in any case before the onset of the gastrula stage. For eggs obtained from commercial supplier, it may not be possible to start the test immediately after fertilisation. As the sensitivity of the test may be seriously influenced by delaying the start of the test, the test should be initiated within eight hours after fertilisation. As larvae are not fed during the exposure period, the test should be terminated just before the yolk sac of any larvae in any of the test chambers has been completely absorbed or before mortalities by starvation start in controls. The duration will depend upon the species used. Some recommended durations are given in Appendices 2 and 3.

**1.7.1.2. Loading**

The number of fertilised eggs at the start of the test should be sufficient to meet statistical requirements. They should be randomly distributed among treatments, and at least 30 fertilised eggs, divided equally (or as equally as possible since it can be difficult to obtain equal batches when using some species) between at least three replicate test chambers, should be used per concentration. The loading rate (biomass per volume of test solution) should be low enough in order that a dissolved oxygen concentration of at least 60 % ASV can be maintained without aeration. For flow-through tests, a loading rate not exceeding 0,5 g/l per 24 hours and not exceeding 5 g/l of solution at any time has been recommended (2).

**1.7.1.3. Light and temperature**

The photoperiod and test water temperature should be appropriate for the test species (Appendix 2 and 3). For the purpose of temperature monitoring, it may be appropriate to use an additional test vessel.

**▼B****1.7.2. Test concentrations**

Normally, five concentrations of the test substance spaced by a constant factor not exceeding 3,2 are required. The curve relating LC<sub>50</sub> to period of exposure in the acute study should be considered when selecting the range of test concentrations. The use of fewer than five concentrations, for example in limit tests, and a narrower concentration interval may be appropriate in some circumstances. Justification should be provided if fewer than five concentrations are used. Concentrations of the substance higher than the 96 hour LC<sub>50</sub> or 100 mg/l, whichever is the lower, need not be tested. Substances should not be tested above their solubility limit in the test water.

When a solubilising agent is used to aid preparation of test solutions (see Section 1.6.6), its final concentration in the test vessels should not be greater than 0,1 ml/l and should be the same in all test vessels.

**1.7.3. Controls**

One dilution-water control (replicated as appropriate) and also, if relevant, one control containing the solubilising-agent (replicated as appropriate) should be run in addition to the test series.

**1.7.4. Frequency of analytical determinations and measurements**

During the test, the concentrations of the test substance are determined at regular intervals.

In semi-static tests where the concentration of the test substance is expected to remain within  $\pm 20$  % of the nominal (i.e. within the range 80-120 %; see Section 1.4 and 1.6.6), it is recommended that, as a minimum, the highest and lowest test concentrations be analysed when freshly prepared and immediately prior to renewal on at least three occasions spaced evenly over the test (i.e. analyses should be made on a sample from the same solution — when freshly prepared and at renewal).

For tests where the concentration of the test substance is not expected to remain within  $\pm 20$  % of nominal (on the basis of stability data of the substance), it is necessary to analyse all test concentrations, when freshly prepared and at renewal, but following the same regime (i.e. on at least three occasions spaced evenly over the test). Determination of test substance concentrations prior to renewal need only be performed on one replicate vessel at each test concentration. Determinations should be made no more than seven days apart. It is recommended that results be based on measured concentrations. However, if evidence is available to demonstrate that the concentration of the test substance in solution has been satisfactorily maintained within  $\pm 20$  % of the nominal or measured initial concentration throughout the test, then results can be based on nominal or measured initial values.

For flow-through tests, a similar sampling regime to that described for semi-static tests is appropriate (but measurement of 'old' solutions is not applicable in this case). However, if the test duration is more than seven days, it may be advisable to increase the number of sampling occasions during the first week (e.g. three sets of measurements) to ensure that the test concentrations are remaining stable.

**▼B**

Samples may need to be centrifuged or filtered (e.g. using a 0,45 µm pore size). However, since neither centrifuging nor filtration appears always to separate the non-bioavailable fraction of the test substance from that which is bioavailable, samples may not be subjected to those treatments.

During the test, dissolved oxygen, pH and temperature should be measured in all test vessels. Total hardness and salinity (if relevant) should be measured in the controls and one vessel at the highest concentration. As a minimum, dissolved oxygen and salinity (if relevant) should be measured three times (at the beginning, middle and end of the test). In semi-static tests, it is recommended that dissolved oxygen be measured more frequently, preferably before and after each water renewal or at least once at week. pH should be measured at the beginning and end of each water renewal in semi-static test and at least weekly in flow-through tests. Hardness should be measured once each test. Temperature should be measured daily and it should preferably be monitored continuously in at least one test vessel.

#### 1.7.5. Observations

##### 1.7.5.1. *Stage of embryonic development*

The embryonic stage (i.e. gastrula stage) at the beginning of exposure to the test substance should be verified as precisely as possible. This can be done using a representative sample of eggs suitably preserved and cleared. The literature may also be consulted for the description and illustration of embryonic stages (2)(5)(10)(11).

##### 1.7.5.2. *Hatching and survival*

Observations on hatching and survival should be made at least once daily and numbers recorded. It may be desirable to make more frequent observations at the beginning of the test (e.g. each 30 minutes during the first three hours), since in some cases, survival times can be more relevant than only the number of deaths (e.g. when there are acute toxic effects). Dead embryos and larvae should be removed as soon as observed since they can decompose rapidly. Extreme care should be taken when removing dead individuals not to knock or physically damage adjacent eggs/-larvae, these being extremely delicate and sensitive. Criteria for death vary according to life stage:

- **for eggs:** particularly in the early stages, a marked loss of translucency and change in colouration, caused by coagulation and/or precipitation of protein, leading to a white opaque appearance,
- **for embryos:** absence of body movement and/or absence of heart beat and/or opaque discoloration in species whose embryos are normally translucent,
- **for larvae:** immobility and/or absence of respiratory movement and/or absence of heart-beat and/or white opaque colouration of central nervous system and/or lack of reaction mechanical stimulus.

**▼B**1.7.5.3. *Abnormal appearance*

The number of larvae showing abnormality of body form and/or pigmentation, and the stage of yolk-sac absorption, should be recorded at adequate intervals depending on the duration of the test and the nature of the abnormality described. It should be noted that abnormal embryos and larvae occur naturally and can be of the order of several per cent in the control(s) in some species. Abnormal animals should only be removed from the test vessels on death.

1.7.5.4. *Abnormal behaviour*

Abnormalities, e.g. hyperventilation, uncoordinated swimming, and atypical quiescence should be recorded at adequate intervals depending on the duration of the test. These effects, although difficult to quantify, can, when observed, aid in the interpretation of mortality data, i.e. provide information on the mode of toxic action of the substance.

1.7.5.5. *Length*

At the end of the test, measurement of individual lengths is recommended; standard, fork or total length may be used. If however, caudal fin rot or fin erosion occurs, standard lengths should be used. Generally, in a well-run test, the coefficient of variation for length among replicates in the controls should be  $\leq 20$  %.

1.7.5.6. *Weight*

At the end of the test, individual weights can be measured; dry weights (24 hours at 60 °C) are preferable to wet weights (blotted dry). Generally, in a well-run test, the coefficient of variation for weight among replicates in the controls should be  $\leq 20$  %.

These observations will result in some or all of the following data being available for statistical analysis:

- cumulative mortality,
- numbers of healthy larvae at end of test,
- time to start of hatching and end of hatching (i.e. 90 % hatching in each replicate),
- numbers of larvae hatching each day,
- length (and weight) of surviving animals at end of the test,
- numbers of larvae that are deformed or of abnormal appearance,
- numbers of larvae exhibiting abnormal behaviour.

**▼B****2. DATA AND REPORTING****2.1. TREATMENT OF RESULTS**

It is recommended that a statistician be involved in both the design and analysis of the test since the method allows for considerable variation in experimental design as, for example, in the number of test chambers, number of test concentrations, starting number of fertilised eggs and in the parameters measured. In view of the options available in test design, specific guidance on statistical procedures is not given here.

If LOEC/NOECs are to be estimated, it will be necessary for variations to be analysed within each set of replicates using analysis of variance (ANOVA) or contingency table procedures. In order to make a multiple comparison between the results at the individual concentrations and those for the controls, Dunnett's method may be found useful (12)(13). Other useful examples are also available (14)(15). The size of the effect detectable using ANOVA or other procedures (i.e. the power of the test) should be calculated and reported. It should be noted that not all the observations listed in Section 1.7.5.6 are suitable for statistical analysis using ANOVA. For example, cumulative mortality and numbers of healthy larvae at the end of the test could be analysed using probit methods.

If LC/EC<sub>x</sub>s are to be estimated, (a) suitable curve(s), such as the logistic curve, should be fitted to the data of interest using a statistical method such as least squares or non-linear least squares. The curve(s) should be parameterised so that the LC/EC<sub>x</sub> of interest and its standard error can be estimated directly. This will greatly ease the calculation of the confidence limits around the LC/EC<sub>x</sub>. Unless there are good reasons to prefer different confidence levels, two-sided 95 % confidence should be quoted. The fitting procedure should preferably provide a means for assessing the significance of the lack of fit. Graphical methods for fitting curves can be used. Regression analysis is suitable for all observations listed in Section 1.7.5.6.

**2.2. INTERPRETATION OF RESULTS**

The results should be interpreted with caution where measured toxicant concentrations in test solutions occur at levels near the detection limit of the analytical method. The interpretation of results for concentrations above the water solubility of the substance should also be made with care.

**2.3. THE TEST REPORT**

The test report must include the following information:

**2.3.1. Test substance:**

- physical nature and relevant physical-chemical properties;
- chemical identification data, including purity and analytical method for quantification of the test substance where appropriate.

**▼ B****2.3.2. Test species:**

- scientific name, strain, numbers of parental fish (i.e. how many females were used for providing the required numbers of eggs in the test), source and method of collection of the fertilised eggs and subsequent handling.

**2.3.3. Test conditions:**

- test procedure used (e.g. semi-static or flow-through, time period from fertilisation to start the test, loading, etc),
- photoperiod(s),
- test design (e.g. number of test chambers and replicates, number of embryos per replicate),
- method of preparation of stock solutions and frequency of renewal (the solubilising agent and its concentration must be given, when used),
- the nominal test concentrations, the measured values, their means and their standard deviations in the test vessels and the method by which these were attained and, if the test substance is soluble in water at concentrations below those tested, evidence should be provided that the measurements refer to the concentrations of the test substance in solution,
- dilution water characteristics: pH, hardness, temperature, dissolved oxygen concentration, residual chlorine levels (if measured), total organic carbon, suspended solids, salinity of the test medium (if measured) and any other measurements made,
- water quality within test vessels: pH, hardness, temperature and dissolved oxygen concentration.

**2.3.4. Results:**

- results from any preliminary studies on the stability of the test substance,
- evidence that controls met the overall survival acceptability standard of the test species (Appendices 2 and 3),
- data on mortality/survival at embryo and larval stages and overall mortality/survival,
- days to hatch and numbers hatched,
- data for length (and weight),
- incidence and description of morphological abnormalities, if any,
- incidence and description of behavioural effects, if any,
- statistical analysis and treatment of data,
- for tests analysed using ANOVA, the lowest observed effect concentration (LOEC) at  $p=0,05$  and the no observed effect concentration (NOEC) for each response assessed, including a description of the statistical procedures used and an indication of what size of effect could be detected,



**▼B**

- for tests analysed using regression techniques, the LC/EC<sub>x</sub> and confidence intervals and a graph of the fitted model used for its calculation,
- explanation for any deviation from this testing method.

3. **REFERENCES**

- (1) Kristensen P., (1990) Evaluation of the Sensitivity of Short Term Fish Early Life Stage Tests in Relation to other FELS Test Methods. Final report to the Commission of the European Communities, p. 60. June 1990.
- (2) ASTM (1988) Standard Guide for Conducting Early Life-Stage Toxicity Tests with Fishes. American Society for Testing and Materials. E 1241-88. p. 26
- (3) Brauhn J.L. and Schoettger R.A., (1975) Acquisition and Culture of Research Fish: Rainbow trout, Fathead minnows, Channel Catfish and Bluegills. p. 54, Ecological Research Series, EPA-660/3-75-011, Duluth, Minnesota.
- (4) Brungs W.A. and Jones B.R., (1977) Temperature Criteria for Freshwater Fish: Protocol and Procedures p. 128, Ecological Research Series EPA-600/3-77-061, Duluth, Minnesota.
- (5) Laale H.W., (1977) The Biology and Use of the Zebrafish (*Brachydanio rerio*) in Fisheries Research. A Literature Review. J. Biol. 10, p. 121-173.
- (6) Legault R. (1958) A Technique for Controlling the Time of Daily Spawning and Collecting Eggs of the Zebrafish, *Brachydanio rerio* (Hamilton-Buchanan) Copeia, 4, p. 328-330.
- (7) Dave G., Damgaard B., Grande M., Martelin J.E., Rosander B. and Viktor T., (1987) Ring Test of an Embryo-larval Toxicity Test with Zebrafish (*Brachydanio rerio*) Using Chromium and Zinc as Toxicants. Environmental Toxicology and Chemistry, 6, p. 61-71.
- (8) Birge J.W., Black J.A. and Westerman A.G., (1985) Short-term Fish and Amphibian Embryo-larval Tests for Determining the Effects of Toxicant Stress on Early Life Stages and Estimating Chronic Values for Single Compounds and Complex Effluents. Environmental Toxicology and Chemistry 4, p. 807-821.
- (9) Van Leeuwen C.J., Espeldoorn A. and Mol F. (1986) Aquatic Toxicological Aspects of Dithiocarbamates and Related Compounds. III. Embryolarval Studies with Rainbow Trout (*Salmo gairdneri*). J. Aquatic Toxicology, 9, p. 129-145.
- (10) Kirchen R.V. and W. R. West (1969) Teleostean Development. Carolina Tips 32(4): 1-4. Carolina Biological Supply Company.
- (11) Kirchen R.V. and W. R. West (1976) The Japanese Medaka. Its care and Development. Carolina Biological Supply Company, North Carolina. p. 36.
- (12) Dunnett C.W., (1955) A Multiple Comparisons Procedure for Comparing Several Treatments with Control. J. Amer. Statist. Assoc., 50, p. 1096-1121.
- (13) Dunnett C.W. (1964) New Tables for Multiple Comparisons with a Control. Biometrics, 20, p. 482-491.
- (14) Mc Clave J.T., Sullivan J.H. and Pearson J.G., (1980) Statistical Analysis of Fish Chronic Toxicity Test Data. Proceedings of 4th Aquatic Toxicology Symposium, ASTM, Philadelphia.

## ▼B

- (15) Van Leeuwen C.J., Adema D.M.M. and Hermes J., (1990) Quantitative Structure-Activity Relationships for Fish Early Life Stage Toxicity. *Aquatic Toxicology*, 16, p. 321-334.
- (16) Environment Canada., (1992) Toxicity Tests Using Early Life Stages of Salmonid Fish (Rainbow Trout, Coho Salmon or Atlantic Salmon). Biological Test Method Series. Report EPS 1/RM/28, December 1992, p. 81.
- (17) Dave G. and Xiu R., (1991) Toxicity of Mercury, Nickel, Lead and Cobalt to Embryos and Larvae of Zebrafish, *Brachydanio rerio*. *Arch. of Environmental Contamination and Toxicology*, 21, p. 126-134.
- (18) Meyer A., Bierman C.H. and Orti G., (1993) The phylogenetic position of the Zebrafish (*Danio rerio*), a model system in developmental biology — an invitation to the comparative methods. *Proc. Royal Society of London, Series B*, 252: 231-236.
- (19) Ghillebaert F., Chaillou C., Deschamps F. and Roubaud P., (1995) Toxic Effects, at Three pH Levels, of Two Reference Molecules on Common Carp Embryo. *Ecotoxicology and Environmental Safety* 32, p. 19-28.
- (20) US EPA, (1991) Guidelines for Culturing the Japanese Medaka, *Oryzias latipes*. EPA report EPA/600/3-91/064, Dec. 1991, EPA, Duluth.
- (21) US EPA, (1991) Guidelines for Conducting Early Life Stage Toxicity Tests with Japanese Medaka, (*Oryzias latipes*). EPA report EPA/600/3-91/063, Dec. 1991, EPA, Duluth.
- (22) De Graeve G.M., Cooney J.D., McIntyre D.O., Poccocic T.L., Reichenbach N.G., Dean J.H. and Marcus M.D., (1991) Validity in the performance of the seven-day Fathead minnow (*Pimephales promelas*) larval survival and growth test: an intra- and interlaboratory study. *Environ. Tox. Chem.* 10, p. 1189-1203.
- (23) Calow P., (1993) *Handbook of Ecotoxicology*, Blackwells, Oxford. Vol. 1, chapter 10: Methods for spawning, culturing and conducting toxicity tests with Early Life stages of Estuarine and Marine fish.
- (24) Balon E.K., (1985) *Early life history of fishes: New developmental, ecological and evolutionary perspectives*, Junk Publ., Dordrecht, p. 280.
- (25) Blaxter J.H.S., (1988) Pattern and variety in development, In: W.S. Hoar and D.J. Randall Eds., *Fish Physiology*, vol. XIA, Academic press, p. 1-58.

Table 1A

## Fish species recommended for testing

FRESHWATER
<i>Oncorhynchus mykiss</i> Rainbow trout (9)(16)
<i>Danio rerio</i> Zebrafish (7)(17)(18)
<i>Cyprinus caprio</i> Common carp (8)(19)
<i>Oryzias latipes</i> Japanese ricefish/Medaka (20)(21)
<i>Pimephales promelas</i> Fathead minnow (8)(22)

▼B

Table 1B

## Examples of other well-documented species which have also been used

FRESHWATER	SALTWATER
<i>Carassius auratus</i> Goldfish (8)	<i>Menidia peninsulae</i> Tidewater silverside (23)(24)(25)
<i>Lepomis macrochirus</i> Bluegill (8)	<i>Clupea harengus</i> Herring (24)(25)
	<i>Gadus morhua</i> Cod (24)(25)
	<i>Cyprinodon variegatus</i> Sheepshead minnow (23)(24)(25)

*Appendix 1***GUIDANCE ON PERFORMANCE OF A TOXICITY TEST ON EMBRYOS AND SAC-FRY OF ZEBRAFISH (*BRACHYDANIO RERIO*)**

## INTRODUCTION

The zebrafish originates from the Coromandel coast of India where it inhabits fast-flowing streams. It is a common aquarium fish of the carp family, and information about procedures for its care and culture can be found in standard reference books on tropical fish. Its biology and use in fishery research have been reviewed by Laale (1).

The fish rarely exceeds 45 mm in length. The body is cylindrical with 7-9 dark-blue horizontal silvery stripes. These stripes run into the caudal and anal fins. The back is olive-green. Males are slimmer than females. Females are more silvery and the abdomen is distended, particularly prior to spawning.

Adult fishes are able to tolerate large fluctuations in temperature, pH and hardness. However, in order to get healthy fish which produce eggs of good quality, optimal conditions should be provided.

During spawning the male pursues and butts the female, and as the eggs are expelled they are fertilised. The eggs, which are transparent and non-adhesive, fall to the bottom where they may be eaten by the parents. Spawning is influenced by light. If the morning light is adequate, the fish usually spawns in the early hours following daybreak.

A female can produce batches of several hundreds of eggs at weekly intervals.

## CONDITIONS OF PARENTAL FISH, REPRODUCTION AND EARLY-LIFE STAGES

Select a suitable number of healthy fish and keep these in suitable water (e.g. Annex 4) for at least two weeks prior to the intended spawning. The group of fish should be allowed to breed at least once before producing the batch of eggs used in the test. The density of fish during this period should not exceed 1 gram of fish per litre. Regular changes of water or the use of purification systems will enable the density to be higher. The temperature in the holding tanks should be maintained at  $25 \pm 2$  °C. The fish should be provided with a varied diet, which may consist of, for example, appropriate commercial dry food, live newly hatched *Artemia*, chironomids, *Daphnia*, white worms (*Enchytraeids*).

Two procedures are outlined below, which in practice have led to a sufficient batch of healthy, fertilised eggs for a test to be run:

- i) Eight females and 16 males are placed in a tank containing 50 litres of dilution water, shielded from direct light and left as undisturbed as possible for at least 48 hours. A spawning tray is placed at the bottom of the aquarium in the afternoon the day before start of the test. The spawning tray consists of a frame (plexi-glass or other suitable material), 5-7 cm high with a 2-5 mm coarse net attached at the top and a 10-30 µm fine net at the bottom. A number of 'spawning-trees', consisting of untwisted nylon rope, are attached to the coarse net of the frame. After the fish have been left in dark for 12 hours, a faint light is turned on which will initiate the spawning. Two to four hours after spawning, the spawning tray is removed and the eggs collected. The spawning tray will prevent the fish from eating the eggs and at the same time permit an easy collection of the eggs. The group of fish should have spawned at least once before the spawning from which eggs are used for testing.

**▼B**

- ii) Five to 10 male and female fish are housed individually at least two weeks prior to the intended spawning. After 5-10 days, the abdomens of the females will be distended and their genital papillae visible. Male fish lack papillae. Spawning is performed in spawning tanks equipped with a false mesh bottom (as above). The tank is filled with dilution water, so that the depth of water above the mesh is 5-10 cm. One female and two males are placed in the tank the day before the intended spawning. The water temperature is gradually increased one degree higher than the acclimatisation temperature. The light is turned off and the tank is left as undisturbed as possible. In the morning a faint light is turned on which will initiate spawning. After two to four hours, the fish are removed and the eggs collected. If larger batches of eggs are needed than can be obtained from one female, a sufficient number of spawning tanks may be set-up in parallel. By recording the reproduction success of the individual females prior to the test (size of batch and quality), those females with highest reproduction success may be selected for breeding.

The eggs should be transferred to the test vessels by means of glass tubes (inner diameter not less than 4 mm) provided with a flexible suction bulb. The amount of water accompanying the eggs on their transfer should be as small as possible. The eggs are heavier than water and sink out of the tube. Care should be taken to prevent eggs (and larvae) coming into contact with the air. Microscopic examination of sample(s) of the batch(es) should be carried out to ensure that there are no irregularities in the first developmental stages. Disinfection of the eggs is not allowed.

The mortality rate of the eggs is highest within the first 24 hours after fertilisation. A mortality of 5-40 % is often seen during this period. Eggs degenerate as a result of unsuccessful fertilisation or development failures. The quality of the batch of eggs seems to depend on the female fish, as some females consistently produce good quality eggs, others never will. Also the development rate and the rate of hatching vary from one batch to another. The successfully fertilised eggs and the yolk sac larvae survive well, normally above 90 %. At 25 °C the eggs will hatch three-five days after fertilisation and the yolk sac will be absorbed approximately 13 days after fertilization.

The embryonic development has been well defined by Hisaoka and Battle (2). Due to the transparency of the eggs and post-hatch larvae, the development of the fish may be followed and the presence of malformations may be observed. Approximately four hours after spawning, the non-fertilized eggs may be distinguished from the fertilized (3). For this examination, eggs and larvae are placed in test vessels of small volume and studied under a microscope.

The test conditions, which apply to the early life stages, are listed in Appendix 2. Optimal values for pH values and hardness of the dilution water are 7,8 and 250 mg CaCO<sub>3</sub>/l respectively.

#### CALCULATIONS AND STATISTICS

A two-stage approach is proposed. First, the data on mortality, abnormal development and hatching-time are analysed statistically. Then, for those concentrations at which no adverse effects on any of these parameters have been detected, the body length is statistically evaluated. This approach is advisable since the toxicant may selectively kill smaller fish, delay hatching-time and induce gross malformations, thus leading to biased length measurements. Furthermore, there will be roughly the same number of fish to be measured per treatment, ensuring the validity of the test statistics.

**▼ B**LC<sub>50</sub> AND EC<sub>50</sub> DETERMINATIONS

The percentage of surviving eggs and larvae is calculated and corrected for mortality in the controls in accordance with Abbott's formula (4):

$$P = 100 - \left( \frac{C - P'}{C} \times 100 \right)$$

where:

P = corrected % survival

P' = % survival observed in the test concentration

C = % survival in the control

If possible, the LC<sub>50</sub> is determined by a suitable method at the end of the test.

If the inclusion of morphological abnormalities in the EC<sub>50</sub> statistic is desired, guidance can be found in Stephan (5).

## ESTIMATION OF LOEC AND NOEC

An objective of the egg and sac-fry test is to compare the non-zero concentrations with the control, i.e. to determine the LOEC. Therefore multiple comparison procedures should be utilised (6)(7)(8)(9)(10).

## REFERENCES

- (1) Laale H.W., (1977) The Biology and Use of the Zebrafish (*Brachydanio rerio*) in Fisheries Research. A Literature Review. J. Fish Biol. 10, p. 121-173.
- (2) Hisaoka K.K. and Battle H.I., (1958). The Normal Development Stages of the Zebrafish *Brachydanio rerio* (Hamilton-Buchanan) J. Morph., 102, p. 311.
- (3) Nagel R., (1986). Untersuchungen zur Eiproduktion beim Zebraabräbling (*Brachydanio rerio* Hamilton-Buchanan). Journal of Applied Ichthyology, 2, p. 173-181.
- (4) Finney D.J., (1971). Probit Analysis, 3rd ed., Cambridge University Press, Great Britain, p. 1-333.
- (5) Stephan C.E., (1982). Increasing the Usefulness of Acute Toxicity Tests. Aquatic Toxicology and Hazard Assessment: Fifth Conference, ASTM STP 766, J.G. Pearson, R.B. Foster and W.E. Bishop, Eds., American Society for Testing and Materials, p. 69-81.
- (6) Dunnett C.W. (1955). A Multiple Comparisons Procedure for Comparing Several Treatments with a Control. J. Amer. Statist. Assoc., 50, p. 1096-1121.
- (7) Dunnett C.W., (1964) New Tables for Multiple Comparisons with a Control. Biometrics, 20, p. 482-491.
- (8) Williams D.A., (1971). A Test for Differences Between Treatment Means when Several Dose Levels are Compared with a Zero Dose Control. Biometrics, 27, p. 103-117.
- (9) Williams D.A., (1972). The Comparison of Several Dose Levels with a Zero Dose Control. Biometrics 28, p. 519-531.
- (10) Sokal R.R. and Rohlf F.J., (1981). Biometry, the Principles and Practice of Statistics in Biological Research, W.H. Freeman and Co., San Francisco.

## TEST CONDITIONS, DURATION AND SURVIVAL CRITERIA FOR RECOMMENDED SPECIES

Species	Temp (°C)	Salinity (0/00)	Photo-period (hrs)	Duration of stages (days)		Typical duration of test	Survival of control, (minimum %)	
				Embryo	Sac-fry		Hatching success	Post-hatch
FRESHWATER								
<i>Brachydanio rerio</i> Zebrafish	25 ± 1	—	12-16	3-5	8-10	As soon as possible after fertilisation (early gastrula stage) to 5 days post-hatch (8-10 days)	80	90
<i>Oncorhynchus mykiss</i> Rainbow trout	10 ± 1 <sup>(1)</sup> 12 ± 1 <sup>(2)</sup>	—	0 <sup>(a)</sup>	30-35	25-30	As soon as possible after fertilisation (early gastrula stage) to 20 days post-hatch (50-55 days)	66	70
<i>Cyprinus carpio</i> Common carp	21-25	—	12-16	5	> 4	As soon as possible after fertilisation (early gastrula stage) to 4 days post-hatch (8-9 days)	80	75
<i>Oryzias latipes</i> Japanese ricefish/Medaka	24 ± 1 <sup>(1)</sup> 23 ± 1 <sup>(2)</sup>	—	12-16	8-11	4-8	As soon as possible after fertilisation (early gastrula stage) to 5 days post-hatch (13-16 days)	80	80
<i>Pimephales promelas</i> Fathead minnow	25 ± 2	—	16	4-5	5	As soon as possible after fertilisation (early gastrula stage) to 4 days post-hatch (8-9 days)	60	70

<sup>(1)</sup> For embryos.

<sup>(2)</sup> For larvae.

<sup>(a)</sup> Darkness for embryo and larvae until one week after hatching except when they are being inspected. Then subdued lighting throughout the test.

## Test conditions, duration and survival criteria for other well documented species

Species	Temp ( °C)	Salinity (0/00)	Photo-period (hrs)	Duration of stages (days)		Typical duration of embryo and sac-fry test	Survival of control (minimum %)	
				Embryo	Sac-fry test		Hatching success	Post-hatch
FRESHWATER								
<i>Carassius auratus</i> Goldfish	24 ± 1	—	—	3-4	> 4	As soon as possible after fertilisation (early gastrula stage) to 4 days post-hatch (7 days)	—	80
<i>Leopomis macrochirus</i> Blugill sunfish	21 ± 1	—	16	3	> 4	As soon as possible after fertilisation (early gastrula stage) to 4 days post-hatch (7 days)	—	75
SALTWATER								
<i>Menidia peninsulae</i> Tidewater silverside	22-25	15-22	12	1,5	10	As soon as possible after fertilisation (early gastrula stage) to 5 days post-hatch (6-7 days)	80	60
<i>Clupea harengus</i> Herring	10 ± 1	8-15	12	20-25	3-5	As soon as possible after fertilisation (early gastrula stage) to 3 days post-hatch (23-27 days)	60	80
<i>Gadus morhua</i> Cod	5 ± 1	5-30	12	14-16	3-5	As soon as possible after fertilisation (early gastrula stage) to 3 days post-hatch (18 days)	60	80
<i>Cyprinodon variegatus</i> Sheepshead minnow	25 ± 1	15-30	12	—	—	As soon as possible after fertilisation (early gastrula stage) to 4/7 days post-hatch (28 days)	> 75	80



**▼B***Appendix 4***SOME CHEMICAL CHARACTERISTICS OF AN ACCEPTABLE  
DILUTION WATER**

Substance	Concentrations
Particulate matter	< 20 mg/l
Total organic carbon	< 2 mg/l
Unionised ammonia	< 1 µg/l
Residual chlorine	< 10 µg/l
Total organophosphorus pesticides	< 50 ng/l
Total organochlorine pesticides plus polychlorinated biphenyls	< 50 ng/l
Total organic chlorine	< 25 ng/l

**▼B****C.16. HONEYBEES — ACUTE ORAL TOXICITY TEST****1. METHOD**

This acute toxicity test method is a replicate of the OECD TG 213 (1998).

**1.1. INTRODUCTION**

This toxicity test is a laboratory method, designed to assess the oral acute toxicity of plant protection products and other chemicals, to adult worker honeybees.

In the assessment and evaluation of toxic characteristics of substances, determination of acute oral toxicity in honeybees may be required, e.g. when exposure of bees to a given chemical is likely. The acute oral toxicity test is carried out to determine the inherent toxicity of pesticides and other chemicals to bees. The results of this test should be used to define the need for further evaluation. In particular, this method can be used in step-wise programmes for evaluating the hazards of pesticides to bees, based on sequential progression from laboratory toxicity tests to semi-field and field experiments (1). Pesticides can be tested as active substances (a.s.) or as formulated products.

A toxic standard should be used to verify the sensitivity of the bees and the precision of the test procedure.

**1.2. DEFINITIONS**

**Acute oral toxicity:** is the adverse effects occurring within a maximum period of 96h of an oral administration of a single dose of test substance.

**Dose:** is the amount of test substance consumed. Dose is expressed as mass ( $\mu\text{g}$ ) of test substance per test animal ( $\mu\text{g}/\text{bee}$ ). The real dose for each bee can not be calculated as the bees are fed collectively, but an average dose can be estimated (totally consumed test substance/number of test bees in one cage).

**LD<sub>50</sub> (Median Lethal Dose) oral:** is a statistically derived single dose of a substance that can cause death in 50 % of animals when administered by the oral route. The LD<sub>50</sub> value is expressed in  $\mu\text{g}$  of test substance per bee. For pesticides, the test substance may be either an active substance (a.s.) or a formulated product containing one or more than one active substance.

**Mortality:** an animal is recorded as dead when it is completely immobile.

**1.3. PRINCIPLE OF THE TEST METHOD**

Adult worker honeybees (*Apis mellifera*) are exposed to a range of doses of the test substance dispersed in sucrose solution. The bees are then fed the same diet, free of the test substance. Mortality is recorded daily during at least 48 h and compared with control values. If the mortality rate is increasing between 24 h and 48 h whilst control mortality remains at an accepted level, i.e.  $\leq 10\%$ , it is appropriate to extend the duration of the test to a maximum of 96 h. The results are analysed in order to calculate the LD<sub>50</sub> at 24 h and 48 h and, in case the study is prolonged, at 72 h and 96 h.

**▼B**

## 1.4. VALIDITY OF THE TEST

For a test to be valid, the following conditions apply:

— the average mortality for the total number of controls must not exceed 10 % at the end of the test,

— the LD<sub>50</sub> of the toxic standard meets the specified range.

## 1.5. DESCRIPTION OF THE TEST METHOD

1.5.1. **Collection of bees**

Young adult worker bees of the same race should be used, i.e. bees of the same age, feeding status, etc. Bees should be obtained from adequately fed, healthy, as far as possible disease-free and queen-right colonies with known history and physiological status. They could be collected in the morning of use or in the evening before test and kept under test conditions to the next day. Bees collected from frames without brood are suitable. Collection in early spring or late autumn should be avoided as the bees have a changed physiology during this time. If tests must be conducted in early spring or late autumn, bees can be emerged in an incubator and reared for one week with 'bee bread' (pollen collected from the comb) and sucrose solution. Bees treated with chemical substances, such as antibiotics, anti-varroa products, etc., should not be used for toxicity test for four weeks from the time of the end of the last treatment.

1.5.2. **Housing and feeding conditions**

Easy to clean and well-ventilated cages are used. Any appropriate material can be used, e.g. stainless steel, wire mesh, plastic or disposable wooden cages, etc. Groups of 10 bees per cage are preferred. The size of test cages should be appropriate to the number of bees, i.e. providing adequate space.

The bees should be held in the dark in an experimental room at a temperature of  $25 \pm 2$  °C. The relative humidity, normally around 50-70 %, should be recorded throughout the test. Handling procedures, including treatment and observations may be conducted under (day) light. Sucrose solution in water with a final concentration of 500 g/l (50 % w/v) is used as food. After given test doses, food should be provided *ad libitum*. The feeding system should allow recording food intake for each cage (see Section 1.6.3.1). A glass tube (approximately 50 mm long and 10 mm wide with the open end narrowed to about 2 mm diameter) can be used.

1.5.3. **Preparation of bees**

The collected bees are randomly allocated to test cages, which are randomly placed in the experimental room.

**▼B**

The bees may be starved for up to 2 h before the initiation of the test. It is recommended that the bees are deprived of food prior to treatment so that all bees are equal in terms of their gut contents at the start of the test. Moribund bees should be rejected and replaced by healthy bees before starting the test.

**1.5.4. Preparation of doses**

Where the test substance is a water miscible compound this may be dispersed directly in 50 % sucrose solution. For technical products and substances of low water solubility, vehicles such as organic solvent, emulsifiers or dispersants of low toxicity to bees may be used (e.g. acetone, dimethylformamide, dimethylsulfoxide). The concentration of the vehicle depends on the solubility of the test substance and it should be the same for all concentrations tested. However, a concentration of the vehicle of 1 % is generally appropriate and should not be exceeded.

Appropriate control solutions should be prepared, i.e. where a solvent or a dispersant is used to solubilise the test substance, two separate control groups should be used: a solution in water, and a sucrose solution with the solvent/carrier at the concentration used in dosing solutions.

**1.6. PROCEDURE****1.6.1. Test and control groups**

The number of doses and replicates tested should meet the statistical requirements for determination of LD<sub>50</sub> with 95 % confidence limits. Normally, five doses in a geometric series, with a factor not exceeding 2,2, and covering the range for LD<sub>50</sub>, are required for the test. However, the dilution factor and the number of concentrations for dosage have to be determined in relation to the slope of the toxicity curve (dose versus mortality) and with consideration taken to the statistical method which is chosen for analysis of the results. A range-finding test enables the choice of the appropriate concentrations for dosage.

A minimum of three replicate test groups, each of 10 bees, should be dosed with each test concentration. A minimum of three control batches, each of 10 bees, should be run in addition to the test series. Control batches should also be included for the solvents/-carriers used (see Section 1.5.4).

**1.6.2. Toxic standard**

A toxic standard should be included in the test series. At least three doses should be selected to cover the expected LD<sub>50</sub> value. A minimum of three replicate cages, each containing 10 bees, should be used with each test dose. The preferred toxic standard is dime-thoate, for which the reported oral LD<sub>50</sub>-24 h is in the range 0,10-0,35 µg a.s./bee (2). However, other toxic standards would be acceptable where sufficient data can be provided to verify the expected dose response (e.g. parathion).

**▼B****1.6.3. Exposure****1.6.3.1. Administration of doses**

Each test group of bees must be provided with 100-200 µl of 50 % sucrose solution in water, containing the test substance at the appropriate concentration. A larger volume is required for products of low solubility, low toxicity or low concentration in the formulation, as higher proportions in the sucrose solution have to be used. The amount of treated diet consumed per group should be monitored. Once consumed (usually within 3-4 h), the feeder should be removed from the cage and replaced with one containing sucrose solution alone. The sucrose solutions are then provided *ad libitum*. For some compounds, at higher concentrations rejection of test dose may result in little or no food being consumed. After a maximum of 6 h, unconsumed treated diet should be replaced with the sucrose solution alone. The amount of treated diet consumed should be assessed (e.g. measurement of volume/weight of treated diet remaining).

**1.6.3.2. Duration**

The duration of the test is preferably 48 h after the test solution has been replaced with sucrose solution alone. If mortality continues to rise by more than 10 % after the first 24 h, the test duration should be extended to a maximum of 96 h provided that control mortality does not exceed 10 %.

**1.6.4. Observations**

Mortality is recorded at 4 h after starting the test and thereafter at 24 h and 48 h (i.e. after giving dose). If a prolonged observation period is required, further assessments should be made at 24 h intervals, up to a maximum of 96 h, provided that the control mortality does not exceed 10 %.

The amount of diet consumed per group should be estimated. Comparison of the rates of consumption of treated and untreated diet within the given 6 h can provide information about palatability of the treated diet.

All abnormal behavioural effects observed during the testing period should be recorded.

**1.6.5. Limit test**

In some cases (e.g. when a test substance is expected to be of low toxicity) a limit test may be performed, using 100 µg a.s./bee in order to demonstrate that the LD<sub>50</sub> is greater than this value. The same procedure should be used, including three replicate test groups for the test dose, the relevant controls, the assessment of the amount of treated diet consumed, and the use of the toxic standard. If mortalities occur, a full study should be conducted. If sublethal effects are observed (see Section 1.6.4), these should be recorded.

**▼ B****2. DATA AND REPORTING****2.1. DATA**

Data should be summarised in tabular form, showing for each treatment group, as well as control and toxic standard groups, the number of bees used, mortality at each observation time and number of bees with adverse behaviour. Analyse the mortality data by appropriate statistical methods (e.g. probit analysis, moving-average, binomial probability) (3)(4). Plot dose-response curves at each recommended observation time and calculate the slopes of the curves and the median lethal doses ( $LD_{50}$ ) with 95 % confidence limits. Corrections for control mortality could be made using Abbott's correction (4)(5). Where treated diet is not completely consumed, the dose of test substance consumed per group should be determined.  $LD_{50}$  should be expressed in  $\mu\text{g}$  of test substance per bee.

**2.2. TEST REPORT**

The test report must include the following information:

**2.2.1. Test substance:**

- physical nature and relevant physical-chemical properties (e.g. stability in water, vapour pressure),
- chemical identification data, including structural formula, purity (i.e. for pesticides, the identity and concentration of active substance(s)).

**2.2.2. Test species:**

- scientific name, race, approximate age (in weeks), collection method, date of collection,
- information on colonies used for collection of test bees including health, any adult disease, any pre-treatment, etc.

**2.2.3. Test conditions:**

- temperature and relative humidity of experimental room,
- housing conditions including type, size and material of cages,
- methods of preparation of stock and test solutions (the solvent and its concentration must be given, when used),
- test design, e.g. number and test concentrations used, number of controls; for each test concentration and control, number of replicate cages and number of bees per cage,
- date of test.

**▼ B**2.2.4. **Results:**

- results of preliminary range-finding study if performed,
- raw data: mortality at each dose tested at each observation time,
- graph of the dose-response curves at the end of the test,
- LD<sub>50</sub> values with 95 % confidence limits, at each of the recommended observation times, for test substance and toxic standard;
- statistical procedures used for determining the LD<sub>50</sub>,
- mortality in controls,
- other biological effects observed or measured e.g. abnormal behaviour of the bees (including rejection of the test dose), rate of consumption of diet in treated and untreated groups,
- any deviation from the test procedures described here and any other relevant information.

3. **REFERENCES**

- (1) EPPO/Council of Europe (1993) Decision-Making Scheme for the Environmental Risk Assessment of Plant Protection Products — Honeybees. EPPO bulletin, vol. 23, N.1, 151-165. March 1993.
- (2) Gough, H. J., McIndoe, E.C., Lewis, G.B., (1994) The use of dimethoate as a reference compound in laboratory acute toxicity tests on honeybees (*Apis mellifera* L.) 1981-1992. Journal of Apicultural Research, 22, p. 119-125.
- (3) Litchfield, J.T. and Wilcoxon, F., (1949) A simplified method of evaluating dose-effect experiments. Jour. Pharmacol. and Exper. Ther., 96, p. 99-113.
- (4) Finney, D.J., (1971) Probit Analysis. 3rd ed., Cambridge, London and New-York.
- (5) Abbott, W.S., (1925) A method for computing the effectiveness of an insecticide. Jour. Econ. Entomol., 18, p. 265-267.

**▼B****C.17. HONEYBEES — ACUTE CONTACT TOXICITY TEST****1. METHOD**

This acute toxicity test method is a replicate of the OECD TG 214 (1998).

**1.1. INTRODUCTION**

This toxicity test is a laboratory method, designed to assess the acute contact toxicity of plant protection products and other chemicals to adult worker honeybees.

In the assessment and evaluation of toxic characteristics of substances, determination of acute contact toxicity in honeybees may be required, e.g. when exposure of bees to a given chemical is likely. The acute contact toxicity test is carried out to determine the inherent toxicity of pesticides and other chemicals to bees. The results of this test should be used to define the need for further evaluation. In particular, this method can be used in step-wise programmes for evaluating the hazards of pesticides to bees, based on sequential progression from laboratory toxicity tests to semi-field and field experiments (1). Pesticides can be tested as active substances (a.s.) or as formulated products.

A toxic standard should be used to verify the sensitivity of the bees and the precision of the test procedure.

**1.2. DEFINITIONS**

**Acute contact toxicity:** is the adverse effects occurring within a maximum period of 96 h of a topical application of a single dose of a substance.

**Dose:** is the amount of test substance applied. Dose is expressed as mass ( $\mu\text{g}$ ) of test substance per test animal ( $\mu\text{g}/\text{bee}$ ).

**LD<sub>50</sub> (Median Lethal Dose) contact:** is a statistically derived single dose of a substance that can cause death in 50 % of animals when administered by the contact. The LD<sub>50</sub> value is given in  $\mu\text{g}$  of test substance per bee. For pesticides, the test substance may be either an active substance (a.s.) or a formulated product containing one or more than one active substance.

**Mortality:** an animal is recorded as dead when it is completely immobile.

**1.3. PRINCIPLE OF THE TEST METHOD**

Adult worker honeybees (*Apis mellifera*) are exposed to a range of doses of the test substance dissolved in appropriate carrier, by direct application to the thorax (droplets). The test duration is 48 h. If the mortality rate is increasing between 24 h and 48 h whilst control mortality remains at an accepted level, i.e. < 10 %, it is appropriate to extend the duration of the test to a maximum of 96 h. Mortality is recorded daily and compared with control values. The results are analysed in order to calculate the LD<sub>50</sub> at 24 h and 48 h, and in case the study is prolonged at 72 h and 96 h.



**▼B**

## 1.4. VALIDITY OF THE TEST

For a test to be valid, the following conditions apply:

- the average mortality for the total numbers of controls must not exceed 10 % at the end of the test,
- the LD<sub>50</sub> of the toxic standard meets the specified range.

## 1.5. DESCRIPTION OF THE TEST METHOD

1.5.1. **Collection of bees**

Young adult worker bees should be used, i.e. bees of the same age, feeding status, race etc. Bees should be obtained from adequately fed, healthy, as far as possible disease-free and queen-right colonies with known history and physiological status. They could be collected in the morning of use or in the evening before test and kept under test conditions to the next day. Bees collected from frames without brood are suitable. Collection in early spring or late autumn should be avoided, as the bees have a changed physiology during the time. If tests have to be conducted in early spring or late autumn, bees can be emerged in an incubator and reared for one week with 'bee bread' (pollen collected from the comb) and sucrose solution. Bees treated with chemical substances, such as antibiotics, anti-varroa products, etc., should not be used for toxicity test for four weeks from the time of the end of the last treatment.

1.5.2. **Housing and feeding conditions**

Easy to clean and well-ventilated cages are used. Any appropriate material can be used, e.g. stainless steel, wire mesh, plastic, disposable wooden cages, etc. The size of test cages should be appropriate to the number of bees, i.e. providing adequate space. Groups of 10 bees per cage are preferred.

The bees should be held in the dark in an experimental room at a temperature of  $25 \pm 2$  °C. The relative humidity, normally around 50-70 %, should be recorded throughout the test. Handling procedures, including treatment and observations may be conducted under (day) light. Sucrose solution in water with a final concentration of 500 g/l (50 % w/v) should be used as food and provided *ad libitum* during the test time, using a bee feeder. This can be a glass tube (approximately 50 mm long and 10 mm wide with the open end narrowed to about 2 mm diameter).

1.5.3. **Preparation of bees**

The collected bees may be anaesthetised with carbon dioxide or nitrogen for application of the test substance. The amount of anaesthetic used and time of exposure should be minimised. Moribund bees should be rejected and replaced by healthy bees before starting the test.

1.5.4. **Preparation of doses**

The test substance is to be applied as solution in a carrier, i.e. an organic solvent or a water solution with a wetting agent. As organic solvent, acetone is preferred but other organic solvents of low toxicity to bees may be used (e.g. dimethylformamide, dimethylsulfoxide). For water dispersed formulated products and highly polar organic substances not soluble in organic carrier solvents, solutions may be easier to apply if prepared in a weak solution of a commercial wetting agent (e.g. Agral, Cittowett, Lubrol, Triton, Tween).

**▼B**

Appropriate control solutions should be prepared, i.e. where a solvent or a dispersant is used to solubilise the test substance, two separate control groups should be used, one treated with water, and one treated with the solvent/dispersant.

**1.6. PROCEDURE****1.6.1. Test and control groups**

The number of doses and replicates tested should meet the statistical requirements for determination LD<sub>50</sub> with 95 % confidence limits. Normally five doses in a geometric series, with a factor not exceeding 2,2, and covering the range for LD<sub>50</sub>, are required for the test. However, the number of doses has to be determined in relation to the slope of the toxicity curve (dose versus mortality) and with consideration taken to the statistical method which is chosen for analysis of the results. A range-finding test enables the choice of the appropriate doses.

A minimum of three replicate test groups, each of 10 bees, should be dosed with each test concentration.

A minimum of three control batches, each of 10 bees, should be run in addition to the test series. If an organic solvent or a wetting agent is used three additional control batches of each 10 bees for the solvent or the wetting agent have to be included.

**1.6.2. Toxic standard**

A toxic standard must be included in the test series. At least three doses should be selected to cover the expected LD<sub>50</sub> value. A minimum of three replicate cages, each containing 10 bees, should be used with each test dose. The preferred toxic standard is dime-thoate, for which the reported contact LD<sub>50</sub>-24 h is in the range 0,10-0,30 µg a.s./bee (2). However, other toxic standards would be acceptable where sufficient data can be provided to verify the expected dose response (e.g. parathion).

**1.6.3. Exposure****1.6.3.1. Administration of doses**

Anaesthetised bees are individually treated by topical application. The bees are randomly assigned to the different test doses and controls. A volume of 1 µl of solution containing the test substance at the suitable concentration should be applied with a microapplicator to the dorsal side of the thorax of each bee. Other volumes may be used, if justified. After application, the bees are allocated to test cages and supplied with sucrose solutions.

**1.6.3.2. Duration**

The duration of the test is preferably 48 hours. If mortality increases by more than 10 % between 24 h and 48 h, the test duration should be extended up to a maximum of 96 h provided that control mortality does not exceed 10 %.

**▼ B****1.6.4. Observations**

Mortality is recorded at 4 h after dosing and thereafter at 24 h and 48 h. If a prolonged observation period is required, further assessments should be made, at 24 h intervals, to a maximum of 96 h, provided that the control mortality does not exceed 10 %.

All abnormal behavioural effects observed during the testing period should be recorded.

**1.6.5. Limit test**

In some cases (e.g. when a test substance is expected to be of low toxicity) limit test may be performed, using 100 µg a.s./bee in order to demonstrate that the LD<sub>50</sub> is greater than this value. The same procedure should be used, including three replicate test groups for the test dose, the relevant controls, and the use of the toxic standard. If mortalities occur, a full study should be conducted. If sublethal effects are observed (see Section 1.6.4) these should be recorded.

**2. DATA AND REPORTING****2.1. DATA**

Data should be summarised in tabular form, showing for each treatment group, as well as, control and toxic standard groups, the number of bees used, mortality at each observation time and number of bees with adverse behaviour. Analyse the mortality data by appropriate statistical methods (e.g. probit analysis, moving-average, binomial probability) (3)(4). Plot dose-response curves at each recommended observation time (i.e. 24 h, 48 h and, if relevant, 72 h, 96 h) and calculate the slopes of the curves and the median lethal doses (LD<sub>50</sub>) with 95 % confidence limits. Corrections for control mortality could be made using Abbott's correction (4)(5). LD<sub>50</sub> should be expressed in µg of test substance per bee.

**2.2. TEST REPORT**

The test report must include the following information:

**2.2.1. Test substance:**

- physical nature and physical-chemical properties (e.g. stability in water, vapour pressure),
- chemical identification data, including structural formula, purity (i.e. for pesticides, the identity and concentration of active substance(s)).

**2.2.2. Test species:**

- scientific name, race, approximate age (in weeks), collection method, date of collection,
- information on colonies used for collection of test bees including health, any adult disease, any pre-treatment, etc.

**▼B**2.2.3. **Test conditions:**

- temperature and relative humidity of experimental room,
- housing conditions including type, size and material of cages,
- methods of administration of test substance, e.g. carrier solvent used, volume of test solution applied anaesthetics used,
- test design, e.g. number and test doses used, number of controls; for each test dose and control, number of replicate cages and number of bees per cage,
- date of test.

2.2.4. **Results:**

- results of preliminary range-finding study if performed,
- raw data: mortality at each concentration tested at each observation time,
- graph of the dose-response curves at the end of the test,
- LD<sub>50</sub> values, with 95 % confidence limits, at each of the recommended observation times, for test substance and toxic standard,
- statistical procedures used for determining the LD<sub>50</sub>,
- mortality in controls,
- other biological effects observed or measured and any abnormal responses of the bees,
- any deviation from the test method procedures described here and any other relevant information.

3. **REFERENCES**

- (1) EPPO/Council of Europe (1993) Decision-Making Scheme for the Environmental Risk Assessment of Plant Protection Products — Honeybees. EPPO bulletin, vol. 23, N.1, p. 151-165. March, 1993.
- (2) Gough, H. J., McIndoe, E.C., Lewis, G.B., (1994) The use of dimethoate as a reference compound in laboratory acute toxicity tests on honeybees (*Apis mellifera* L.), 1981-1992. Journal of Apicultural Research 22, p. 119-125.
- (3) Litchfield, J.T. and Wilcoxon, F., (1949) A simplified method of evaluating dose-effect experiments. Jour. Pharmacol. and Exper. Ther., 96, p. 99-113.
- (4) Finney, D.J., (1971) Probit Analysis. 3rd ed., Cambridge, London and New-York.
- (5) Abbott, W.S., (1925) A method for computing the effectiveness of an insecticide. Jour. Econ. Entomol. 18, p. 265-267.

**▼ B****C.18. ADSORPTION/DESORPTION USING A BATCH EQUILIBRIUM METHOD****1. METHOD**

This method is a replicate of the OECD TG 106, for the Determination of Soil Adsorption/Desorption, using a Batch Equilibrium Method (2000).

**1.1. INTRODUCTION**

The method takes into account a ring test and a workshop for soil selection for the development of an adsorption test (1)(2)(3)(4) and also existing guidelines at national level (5)(6)(7)(8)(9)(10)(11).

Adsorption/desorption studies are useful for generating essential information on the mobility of chemicals and their distribution in the soil, water and air compartments of the biosphere (12)(13)(14)(15)(16)(17)(18)(19)(20)(21). The information can be used in the prediction or estimation, for example, of the availability of a chemical for degradation (22)(23), transformation and uptake by organisms (24); leaching through the soil profile (16)(18)(19)(21)(25)(26)(27)(28); volatility from soil (21)(29)(30); run-off from land surfaces into natural waters (18)(31)(32). Adsorption data can be used for comparative and modelling purposes (19)(33)(34)(35).

The distribution of a chemical between soil and aqueous phases is a complex process depending on a number of different factors: the chemical nature of the substance (12)(36)(37)(38)(39)(40), the characteristics of the soil (4)(12)(13)(14)(41)(42)(43)(44)(45)(46)(47)(48)(49), and climatic factors such as rainfall, temperature, sunlight and wind. Thus, the numerous phenomena and mechanisms involved in the process of adsorption of a chemical by soil cannot be completely defined by a simplified laboratory model such as the present method. However, even if this attempt cannot cover all the environmentally possible cases, it provides valuable information on the environmental relevance of the adsorption of a chemical.

See also General Introduction.

**1.2. SCOPE**

The method is aimed at estimating the adsorption/desorption behaviour of a substance on soils. The goal is to obtain a sorption value which can be used to predict partitioning under a variety of environmental conditions; to this end, equilibrium adsorption coefficients for a chemical on various soils are determined as a function of soil characteristics (e.g. organic carbon content, clay content and soil texture and pH). Different soil types have to be used in order to cover as widely as possible the interactions of a given substance with naturally occurring soils.

In this method, adsorption represents the process of the binding of a chemical to surfaces of soils; it does not distinguish between different adsorption processes (physical and chemical adsorption) and such processes as surface catalysed degradation, bulk adsorption or chemical reaction. Adsorption that will occur on colloids particles (diameter < 0,2 µm) generated by the soils is not taken into account.

**▼ B**

The soil parameters that are believed most important for adsorption are: organic carbon content (3)(4)(12)(13)(14)(41)(43)(44)(45)(46)(47)(48); clay content and soil texture (3)(4)(41)(42)(43)(44)(45)(46) (47)(48) and pH for ionisable compounds (3)(4)(42). Other soil parameters which may have an impact on the adsorption/desorption of a particular substance are the effective cation exchange capacity (ECEC), the content of amorphous iron and aluminium oxides, particularly for volcanic and tropical soils (4), as well as the specific surface (49).

The test is designed to evaluate the adsorption of a chemical on different soil types with a varying range of organic carbon content, clay content and soil texture, and pH. It comprises three tiers:

**Tier 1:** preliminary study in order to determine:

- the soil/solution ratio,
- the equilibrium time for adsorption and the amount of test substance adsorbed at equilibrium,
- the adsorption of the test substance on the surfaces of the test vessels and the stability of the test substance during the test period.

**Tier 2:** screening test: the adsorption is studied in five different soil types by means of adsorption kinetics at a single concentration and determination of distribution coefficient  $K_d$  and  $K_{oc}$ .

**Tier 3:** determination of Freundlich adsorption isotherms to determine the influence of concentration on the extent of adsorption on soils.

Study of desorption by means of desorption kinetics/  
Freundlich desorption isotherms (Appendix 1).

## 1.3. DEFINITIONS AND UNITS

Symbol	Definition	Units
$A_{t_i}$	adsorption percentage at the time $t_i$	%
$A_{eq}$	adsorption percentage at adsorption equilibrium	%
$m_s^{ads}(t_i)$	mass of the test substance adsorbed on the soil at the time $t_i$	$\mu\text{g}$
$m_s^{ads}(\Delta t_i)$	mass of the test substance adsorbed on the soil during the time interval $\Delta t_i$	$\mu\text{g}$
$m_s^{ads}(eq)$	mass of the test substance adsorbed on the soil at adsorption equilibrium	$\mu\text{g}$
$m_0$	mass of the test substance in the test tube, at the beginning of the adsorption test	$\mu\text{g}$
$m_m^{ads}(t_i)$	mass of the test substance measured in an aliquot ( $v_a^A$ ) at the time point $t_i$	$\mu\text{g}$
$m_{aq}^{ads}(eq)$	mass of the substance in the solution at adsorption equilibrium	$\mu\text{g}$
$m_{soil}$	quantity of the soil phase, expressed in dry mass of soil	g

▼ B

Symbol	Definition	Units
$C_{st}$	mass concentration of the stock solution of the substance	$\mu\text{g cm}^{-3}$
$C_0$	initial mass concentration of the test solution in contact with the soil	$\mu\text{g cm}^{-3}$
$C_{aq}^{ads}(t_i)$	mass concentration of the substance in the aqueous phase at the time $t_i$ that the analysis is performed	$\mu\text{g cm}^{-3}$
$C_s^{ads}(eq)$	content of the substance adsorbed on soil at adsorption equilibrium an equilibrium	$\mu\text{g g}^{-1}$
$C_{aq}^{ads}(eq)$	mass concentration of the substance in the aqueous phase at adsorption equilibrium	$\mu\text{g cm}^{-3}$
$V_0$	initial volume of the aqueous phase in contact with the soil during the adsorption test	$\text{cm}^3$
$v_a^A$	volume of the aliquot in which the test substance is measured	$\text{cm}^3$
$K_d$	distribution coefficient for adsorption	$\text{cm}^3 \text{g}^{-1}$
$K_{oc}$	organic carbon normalised adsorption coefficient	$\text{cm}^3 \text{g}^{-1}$
$K_{om}$	organic matter normalised distribution coefficient	$\text{cm}^3 \text{g}^{-1}$
$K_F^{ads}$	Freundlich adsorption coefficient	$\mu\text{g}^{-1/n} (\text{cm}^3)^{1/n} \text{g}^{-1}$
$1/n$	Freundlich exponent	
$D_{t_i}$	desorption percentage at a point time $t_i$	%
$D_{\Delta t_i}$	desorption percentage corresponding to a time interval $\Delta t_i$	%
$K_{des}$	apparent desorption coefficient	$\text{cm}^3 \text{g}^{-1}$
$K_F^{des}$	Freundlich desorption coefficient	$\mu\text{g}^{-1/n} (\text{cm}^3)^{1/n} \text{g}^{-1}$
$m_{aq}^{des}(t_i)$	mass of the test substance desorbed from soil at the time $t_i$	$\mu\text{g}$
$m_m^{des}(\Delta t_i)$	mass of the test substance desorbed from soil during the time $\Delta t_i$	$\mu\text{g}$
$m_m^{des}(eq)$	mass of the substance determined analytically in the aqueous phase at desorption equilibrium	$\mu\text{g}$
$m_{aq}^{des}(eq)$	total mass of the test substance desorbed at desorption equilibrium	$\mu\text{g}$
$m_s^{des}(\Delta t_i)$	mass of the substance remaining adsorbed on the soil after the time interval $\Delta t_i$	$\mu\text{g}$
$m_{aq}^A$	mass of the substance left over from the adsorption equilibrium due to incomplete volume replacement	$\mu\text{g}$
$C_s^{des}(eq)$	content of the test substance remaining adsorbed on the soil at desorption equilibrium	$\mu\text{g g}^{-1}$

## ▼ B

Symbol	Definition	Units
$C_{aq}^{des}$ (eq)	mass concentration of the test substance in the aqueous phase at desorption equilibrium	$\mu\text{g cm}^{-3}$
$V_T$	total volume of the aqueous phase in contact with the soil during the desorption kinetics experiment performed with the serial method	$\text{cm}^3$
$V_R$	volume of the supernatant removed from the tube after the attainment of adsorption equilibrium and replaced by the same volume of a 0,01 M $\text{CaCl}_2$ solution	$\text{cm}^3$
$V_a^D$	volume of the aliquot sampled for analytical purpose from the time (i), during the desorption kinetics experiment performed with the serial method	$\text{cm}^3$
$V_r^r$	volume of the solution taken from the tube (i) for the measurement of the test substance, in desorption kinetics experiment (parallel method)	$\text{cm}^3$
$V_r^F$	volume of the solution taken from the tube for the measurement of the test substance, at desorption equilibrium	$\text{cm}^3$
MB	mass balance	%
$m_E$	total mass of the test substance extracted from soil and walls of the test vessel in two steps	$\mu\text{g}$
$V_{rec}$	volume of the supernatant recovered after the adsorption equilibrium	$\text{cm}^3$
$P_{ow}$	octanol/water partition coefficient	
pKa	dissociation constant	
$S_w$	water solubility	$\text{g l}^{-1}$

## 1.4. PRINCIPLE OF THE TEST METHOD

Known volumes of solutions of the test substance, non-labelled or radiolabelled, at known concentrations in 0,01 M  $\text{CaCl}_2$  are added to soil samples of known dry weight which have been pre-equilibrated in 0,01 M  $\text{CaCl}_2$ . The mixture is agitated for an appropriate time. The soil suspensions are then separated by centrifugation and, if so wished, filtration and the aqueous phase is analysed. The amount of test substance adsorbed on the soil sample is calculated as the difference between the amount of test substance initially present in solution and the amount remaining at the end of the experiment (indirect method).

As an option, the amount of the test substance adsorbed can also be directly determined by analysis of soil (direct method). This procedure which involves stepwise soil extraction with appropriate solvent, is recommended in cases where the difference in the solution concentration of the substance cannot be accurately determined. Examples of such cases are: adsorption of the test substance on surface of the test vessels, instability of the test substance in the time scale of the experiment, weak adsorption giving only small concentration change in the solution; and strong adsorption yielding low concentration which cannot be accurately determined. If radiolabelled substance is used, the soil extraction may be avoided by analysis of the soil phase by combustion and liquid scintillation counting. However, liquid scintillation counting is an unspecific technique which cannot differentiate between parental and transformation products; therefore it should be used only if the test chemical is stable for the duration of the study.



**▼B**

## 1.5. INFORMATION ON THE TEST SUBSTANCE

Chemical reagents should be of analytical grade. The use of non-labelled test substances with known composition and preferably at least 95 % purity or of radiolabelled test substances with known composition and radio-purity, is recommended. In the case of short half-life tracers, decay corrections should be applied.

Before carrying out a test for adsorption-desorption, the following information about the test substance should be available:

- (a) water solubility (A.6);
- (b) vapour pressure (A.4) and/or Henry's Law Constant;
- (c) abiotic degradation: hydrolysis as a function of pH (C.7);
- (d) partition coefficient (A.8);
- (e) ready biodegradability (C.4) or aerobic and anaerobic transformation in soil;
- (f) pKa of ionisable substances;
- (g) direct photolysis in water (i.e. UV-vis absorption spectrum in water, quantum yield) and photodegradation on soil.

## 1.6. APPLICABILITY OF THE TEST

The test is applicable to chemical substances for which an analytical method with sufficient accuracy is available. An important parameter that can influence the reliability of the results, especially when the indirect method is followed, is the stability of the test substance in the time scale of the test. Thus, it is a prerequisite to check the stability in a preliminary study; if a transformation in the time scale of the test is observed, it is recommended that the main study be performed by analysing both soil and aqueous phases.

Difficulties may arise in conducting this test for test substances with low water solubility ( $S_w < 10^{-4} \text{ g l}^{-1}$ ), as well as for highly charged substances, due to the fact that the concentration in the aqueous phase cannot be measured analytically with sufficient accuracy. In these cases, additional steps have to be taken. Guidance on how to deal with these problems is given in the relevant sections of this method.

When testing volatile substances, care should be taken to avoid losses during the study.

## 1.7. DESCRIPTION OF THE METHOD

1.7.1. **Apparatus and chemical reagents**

Standard laboratory equipment, especially the following:

- (a) tubes or vessels to conduct the experiments. It is important that these tubes or vessels,
  - fit directly in the centrifuge apparatus in order to minimise handling and transfer errors,
  - be made of an inert material, which minimises adsorption of the test substance on its surface,

**▼B**

- (b) agitation device: overhead shaker or equivalent equipment; the agitation device should keep the soil in suspension during shaking,
- (c) centrifuge: preferably high-speed, e.g. centrifugation forces > 3 000 g, temperature controlled, capable of removing particles with a diameter greater than 0,2 µm from aqueous solution. The containers should be capped during agitation and centrifugation to avoid volatility and water losses; to minimise adsorption on them, deactivated caps such as teflon lined screw caps should be used,
- (d) optional: filtration device; filters of 0,2 µm porosity, sterile, single use. Special care should be taken in the choice of the filter material, to avoid any losses of the test substance on it; for poorly soluble test substances, organic filter material is not recommended,
- (e) analytical instrumentation, suitable for measuring the concentration of the test chemical,
- (f) laboratory oven, capable of maintaining a temperature of 103 °C to 110 °C,

**1.7.2. Characterisation and selection of soils**

The soils should be characterised by three parameters considered to be largely responsible for the adsorptive capacity: organic carbon, clay content and soil texture, and pH. As already mentioned (see Scope) other physico-chemical properties of the soil may have an impact on the adsorption/desorption of a particular substance and should be considered in such cases.

The methods used for soil characterisation are very important and can have a significant influence on the results. Therefore, it is recommended that soil pH should be measured in a solution of 0,01 M CaCl<sub>2</sub> (that is the solution used in adsorption/desorption testing) according to the corresponding ISO method (ISO-10390-1). It is also recommended that the other relevant soil properties be determined according to standard methods (for example ISO 'Handbook of Soil Analysis'); this permits the analysis of sorption data to be based on globally standardised soil parameters. Some guidance for existing standard methods of soil analysis and characterisation is given in references (50-52). For calibration of soil test methods, the use of reference soils is recommended.

Guidance for selection of soils for adsorption/desorption experiments is given in Table 1. The seven selected soils cover soil types encountered in temperate geographical zones. For ionisable test substances, the selected soils should cover a wide range of pH, in order to be able to evaluate the adsorption of the substance in its ionised and unionised forms. Guidance on how many different soils to use at the various stages of the test is given under 'Performance of the test' 1.9.

If other soil types are preferred, they should be characterised by the same parameters and should have similar variation in properties to those described in Table 1, even if they do not match the criteria exactly.



Table 1:

**Guidance for selection of soil samples for adsorption-desorption**

Soil Type	pH range (in 0,01 M CaCl <sub>2</sub> )	Organic carbon content (%)	Clay content (%)	Soil texture <sup>(1)</sup>
1	4,5 - 5,5	1,0 - 2,0	65-80	clay
2	> 7,5	3,5 - 5,0	20-40	clay loam
3	5,5 - 7,0	1,5 - 3,0	15-25	silt loam
4	4,0 - 5,5	3,0 - 4,0	15-30	loam
5	< 4,0 - 6,0 <sup>(2)</sup>	< 0,5 - 1,5 <sup>(2)</sup> <sup>(3)</sup>	< 10-15 <sup>(2)</sup>	loamy sand
6	> 7,0	< 0,5 - 1,0 <sup>(2)</sup> <sup>(3)</sup>	40-65	clay loam/clay
7	< 4,5	> 10	< 10	sand/loamy sand

<sup>(1)</sup> According to FAO and the US system (85).

<sup>(2)</sup> The respective variables should preferably show values within the range given. If, however, difficulties in finding appropriate soil material occur, values below the indicated minimum are accepted.

<sup>(3)</sup> Soils with less than 0,3 % organic carbon may disturb correlation between organic content and adsorption. Thus, it is recommended the use of soils with a minimum organic carbon content of 0,3 %.

### 1.7.3. Collection and storage of soil samples

#### 1.7.3.1. Collection

No specific sampling techniques or tools are recommended; the sampling technique depends on the purpose of the study (53)-(54)(55)(56)(57)(58).

The following should be considered:

- a) detailed information on the history of the field site is necessary; this includes location, vegetation cover, treatments with pesticides and/or fertilisers, biological additions or accidental contamination. Recommendations of the ISO standard on soil sampling (ISO 10381-6) should be followed with respect to the description of the sampling site;
- b) the sampling site has to be defined by UTM (Universal Transversal Mercator-Projection/European Horizontal Datum) or geographical co-ordinates; this could allow recollection of a particular soil in the future or could help in defining soil under various classification systems used in different countries. Also, only A horizon up to a maximum depth of 20 cm should be collected. Especially for the soil type No 7 if a O<sub>h</sub> horizon is present as part of the soil, it should be included in the sampling.

The soil samples should be transported using containers and under temperature conditions which guarantee that the initial soil properties are not significantly altered.

**▼B**1.7.3.2. *Storage*

The use of soils freshly taken from the field is preferred. Only if this is not possible soil can be stored at ambient temperature and should be kept air-dried. No limit on the storage time is recommended, but soils stored for more than three years should be re-analysed prior to the use with respect to their organic carbon content, pH and CEC.

1.7.3.3. *Handling and preparation of soil samples for the test*

The soils are air-dried at ambient temperature (preferably between 20-25 °C). Disaggregation should be performed with minimal force, so that the original texture of the soil will be changed as little as possible. The soils are sieved to a particle size  $\leq 2$  mm; recommendations of the ISO standard on soil sampling (ISO 10381-6) should be followed with respect to the sieving process. Careful homogenisation is recommended, as this enhances the reproducibility of the results. The moisture content of each soil is determined on three aliquots with heating at 105 °C until there is no significant change in weight (approximately 12 h). For all calculations the mass of soil refers to oven dry mass, i.e. the weight of soil corrected for moisture content.

1.7.4. **Preparation of the test substance for application to soil**

The test substance is dissolved in a solution of 0,01 M CaCl<sub>2</sub> in distilled or de-ionised water; the CaCl<sub>2</sub> solution is used as the aqueous solvent phase to improve centrifugation and minimise cation exchange. The concentration of the stock solution should preferably be three orders of magnitude higher than the detection limit of the analytical method used. This threshold safeguards accurate measurements with respect to the methodology followed in this method; additionally, the stock solution concentration should be below water solubility of the test substance.

The stock solution should preferably be prepared just before application to soil samples and should be kept closed in the dark at 4 °C. The storage time depends on the stability of the test substance and its concentration in the solution.

Only for poorly soluble substances ( $S_w < 10^{-4}$  g l<sup>-1</sup>), an appropriate solubilising agent may be needed when it is difficult to dissolve the test substance. This solubilising agent: (a) should be miscible with water such as methanol or acetonitrile; (b) its concentration should not exceed 1 % of the total volume of the stock solution and should constitute less than that in the solution of the test substance which will come in contact with the soil (preferably less than 0,1 %); and (c) should not be a surfactant or undergo solvolytic reactions with the test chemical. The use of a solubilising agent should be stipulated and justified in the reporting of the data.

Another alternative for poorly soluble substances is to add the test substance to the test system by spiking: the test substance is dissolved in an organic solvent, an aliquot of which is added to the system of soil and 0,01 M solution of CaCl<sub>2</sub> in distilled or de-ionised water. The content of organic solvent in the aqueous phase should be kept as low as possible, normally not exceeding 0,1 %. Spiking from an organic solution may suffer from volume unreproducibility. Thus, an additional error may be introduced as the test substance and co-solvent concentration would not be the same in all tests.

**▼B**

## 1.8. PREREQUISITES FOR PERFORMING THE ADSORPTION/-DESORPTION TEST

1.8.1. **The analytical method**

The key parameters that can influence the accuracy of sorption measurements include the accuracy of the analytical method in analysis of both the solution and adsorbed phases, the stability and purity of the test substance, the attainment of sorption equilibrium, the magnitude of the solution concentration change, the soil/solution ratio and changes in the soil structure during the equilibration process (35)(59-62). Some examples bearing upon the accuracy issues are given in Appendix 2.

The reliability of the analytical method used must be checked at the concentration range which is likely to occur during the test. The experimenter should feel free to develop an appropriate method with appropriate accuracy, precision, reproducibility, detection limits and recovery. Guidance on how to perform such a test is given by the experiment below.

An appropriate volume of 0,01 M CaCl<sub>2</sub>, e.g. 100 cm<sup>3</sup>, is agitated during 4 h with a weight of soil, e.g. 20 g, of high adsorbability, i.e. with high organic carbon and clay content; these weights and volumes may vary depending on analytical needs, but a soil/solution ratio of 1:5 is a convenient starting point. The mixture is centrifuged and the aqueous phase may be filtrated. A certain volume of the test substance stock solution is added to the latter to reach a nominal concentration within the concentration range which is likely to occur during the test. This volume should not exceed 10 % of the final volume of the aqueous phase, in order to change as little as possible the nature of the pre-equilibration solution. The solution is analysed.

One blank run consisting of the system soil + CaCl<sub>2</sub> solution (without test substance) must be included, in order to check for artefacts in the analytical method and for matrix effects caused by the soil.

The analytical methods which can be used for sorption measurements include gas-liquid chromatography (GLC), high-performance liquid chromatography (HPLC), spectrometry (e.g. GC/mass spectrometry, HPLC/mass spectrometry) and liquid scintillation counting (for radiolabelled substances). Independent of the analytical method used, it is considered suitable if the recoveries are between 90 % and 110 % of the nominal value. In order to allow for detection and evaluation after partitioning has taken place, the detection limits of the analytical method should be at least two orders of magnitude below the nominal concentration.

The characteristics and detection limits of the analytical method available for carrying out adsorption studies play an important role in defining the test conditions and the whole experimental performance of the test. This method follows a general experimental path and provides recommendations and guidance for alternative solutions where the analytical method and laboratory facilities may impose limitations.

▼ **B**1.8.2. **The selection of optimal soil/solution ratios**

Selection of appropriate soil to solution ratios for sorption studies depends on the distribution coefficient  $K_d$  and the relative degree of adsorption desired. The change of the substance concentration in the solution determines the statistical accuracy of the measurement based on the form of adsorption equation and the limit of the analytical methodology, in detecting the concentration of the chemical in solution. Therefore, in general practice it is useful to settle on a few fixed ratios, for which the percentage adsorbed is above 20 %, and preferably >50 % (62), while care should be taken to keep the test substance concentration in the aqueous phase high enough to be measured accurately. This is particularly important in the case of high adsorption percentages.

A convenient approach to selecting the appropriate soil/water ratios, is based on an estimate of the  $K_d$  value either by preliminary studies or by established estimation techniques (Appendix 3). Selection of an appropriate ratio can then be made based on a plot of soil/solution ratio versus  $K_d$  for fixed percentages of adsorption (Fig.1). In this plot it is assumed that the adsorption equation is linear<sup>(1)</sup>. The applicable relationship is obtained by rearranging equation (4) of the  $K_d$  in the form of equation (1):

$$\frac{V_0}{m_{\text{soil}}} = \left( \frac{m_0}{m_s^{\text{ads}}(\text{eq})} - 1 \right) K_d \quad (1)$$

or in its logarithmic form assuming that  $R = m_{\text{soil}}/V_0$  and  $A_{\text{eq}} \% / 100 = \frac{m_s^{\text{ads}}(\text{eq})}{m_0}$ :

$$\log R = -\log K_d + \log \left[ \frac{(A_{\text{eq}} \% / 100)}{(1 - A_{\text{eq}} \% / 100)} \right] \quad (2)$$

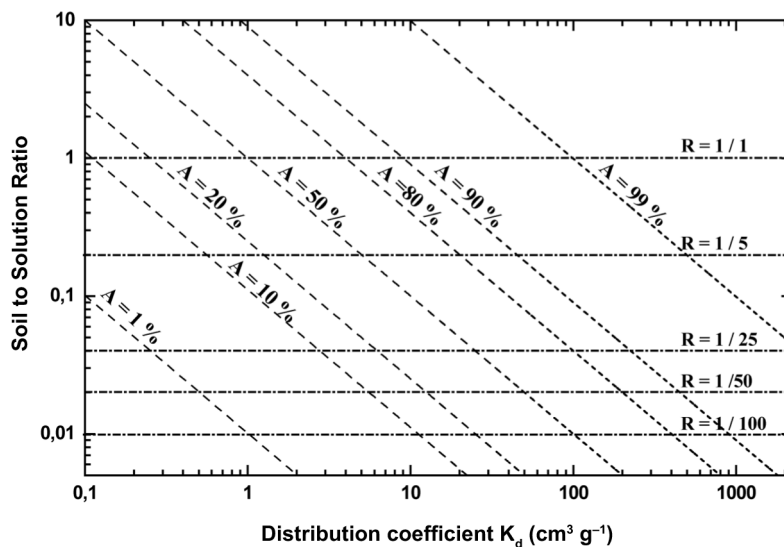


Fig. 1 Relationship between soil to solution ratios and  $K_d$  at various percentages of adsorbed test substance

<sup>(1)</sup>  $C_s^{\text{ads}}(\text{eq}) = K_d \cdot C_{\text{aq}}^{\text{ads}}(\text{eq})$ .

**▼ B**

Fig. 1 shows soil/solution ratios required as a function of  $K_d$  for different levels of adsorption. For example, with a soil/solution ratio of 1:5 and a  $K_d$  of 20, approximately 80 % adsorption would occur. To obtain 50 % adsorption for the same  $K_d$ , a 1:25 ratio must be used. This approach to selecting the appropriate soil/solution ratios gives the investigator the flexibility to meet experimental needs.

Areas which are more difficult to deal with are those where the chemical is highly or very slightly adsorbed. Where low adsorption occurs, a 1:1 soil/solution ratio is recommended, although for some very organic soil types smaller ratios may be necessary to obtain a slurry. Care must be taken with the analytical methodology to measure small changes in solution concentration; otherwise the adsorption measurement will be inaccurate. On the other hand, at very high distribution coefficients  $K_d$ , one can go up to a 1:100 soil/solution ratio in order to leave a significant amount of chemical in solution. However, care must be taken to ensure good mixing, and adequate time must be allowed for the system to equilibrate. An alternative approach to deal with these extreme cases when adequate analytical methodology is missing, is to predict the  $K_d$  value applying estimation techniques based, for example, on  $P_{ow}$  values (Appendix 3). This could be useful especially for low adsorbed/polar chemicals with  $P_{ow} < 20$  and for lipophilic/highly sorptive chemicals with  $P_{ow} > 10^4$ .

## 1.9. PERFORMANCE OF THE TEST

### 1.9.1. Test conditions

All experiments are done at ambient temperature and, if possible, at a constant temperature between 20 °C and 25 °C.

Centrifugation conditions should allow the removal of particles larger than 0,2 µm from the solution. This value triggers the smallest sized particle that is considered as a solid particle, and is the limit between solid and colloid particles. Guidance on how to determine the centrifugation conditions is given in Appendix 4.

If the centrifugation facilities cannot guarantee the removal of particles larger than 0,2 µm, a combination of centrifugation and filtration with 0,2 µm filters could be used. These filters should be made of a suitable inert material to avoid any losses of the test substance on them. In any case, it should be proven that no losses of the test substance occur during filtration.

### 1.9.2. Tier 1 — Preliminary study

The purpose of conducting a preliminary study has already been given in the Scope section. Guidance for setting up such a test is given with the experiment suggested below.

#### 1.9.2.1. Selection of optimal soil/solution ratios

Two soil types and three soil/solution ratios (six experiments) are used. One soil type has high organic carbon and low clay content, and the other low organic carbon and high clay content. The following soil to solution ratios are suggested:

— 50 g soil and 50 cm<sup>3</sup> aqueous solution of the test substance (ratio 1/1),

**▼B**

- 10 g soil and 50 cm<sup>3</sup> aqueous solution of the test substance (ratio 1/5),
  
- 2 g soil and 50 cm<sup>3</sup> aqueous solution of the test substance (ratio 1/25).

The minimum amount of soil on which the experiment can be carried out depends on the laboratory facilities and the performance of analytical methods used. However, it is recommended to use at least 1 g, and preferably 2 g, in order to obtain reliable results from the test.

One control sample with only the test substance in 0,01 M CaCl<sub>2</sub> solution (no soil) is subjected to precisely the same steps as the test systems, in order to check the stability of the test substance in CaCl<sub>2</sub> solution and its possible adsorption on the surfaces of the test vessels.

A blank run per soil with the same amount of soil and total volume of 50 cm<sup>3</sup> 0,01 M CaCl<sub>2</sub> solution (without test substance) is subjected to the same test procedure. This serves as a background control during the analysis to detect interfering substances or contaminated soils.

All the experiments, included controls and blanks, should be performed at least in duplicate. The total number of the samples which should be prepared for the study can be calculated with respect to the methodology which will be followed.

Methods for the preliminary study and the main study are generally the same, exceptions are mentioned where relevant.

The air-dried soil samples are equilibrated by shaking with a minimum volume of 45 cm<sup>3</sup> of 0,01 M CaCl<sub>2</sub> overnight (12 h) before the day of the experiment. Afterwards, a certain volume of the stock solution of the test substance is added in order to adjust the final volume to 50 cm<sup>3</sup>. This volume of the stock solution added: (a) should not exceed 10 % of the final 50 cm<sup>3</sup> volume of the aqueous phase in order to change as little as possible the nature of the pre-equilibration solution; and (b) should preferably result in an initial concentration of the test substance being in contact with the soil ( $C_0$ ) at least two orders of magnitude higher than the detection limit of the analytical method; this threshold safeguards the ability to perform accurate measurements even when strong adsorption occurs (> 90 %) and to determine later the adsorption isotherms. It is also recommended, if possible, that the initial substance concentration ( $C_0$ ) not exceed half of its solubility limit.

An example of how to calculate the concentration of the stock solution ( $C_{st}$ ) is given below. A detection limit of 0,01 µg cm<sup>-3</sup> and 90 % adsorption are assumed; thus, the initial concentration of the test substance in contact with the soil should preferably be 1 µg cm<sup>-3</sup> (two orders of magnitude higher than the detection limit). Supposing that the maximum recommended volume of the stock solution is added, i.e. 5 to 45 cm<sup>3</sup> 0,01 M CaCl<sub>2</sub> equilibration solution (= 10 % of the stock solution to 50 cm<sup>3</sup> total volume of aqueous phase), the concentration of the stock solution should be 10 µg cm<sup>-3</sup>; this is three orders of magnitude higher than the detection limit of the analytical method.

The pH of the aqueous phase should be measured before and after contact with the soil since it plays an important role in the whole adsorption process, especially for ionisable substances.



**▼ B**

The mixture is shaken until adsorption equilibrium is reached. The equilibrium time in soils is highly variable, depending on the chemical and the soil; a period of 24 h is generally sufficient (77). In the preliminary study, samples may be collected sequentially over a 48 h period of mixing (for example at 4, 8, 24, 48 h). However, times of analysis should be considered with flexibility with respect to the work schedule of the laboratory.

There are two options for the analysis of the test substance in the aqueous solution: (a) the parallel method and (b) the serial method. It should be stressed that, although the parallel method is experimentally more tedious, the mathematical treatment of the results is simpler (Appendix 5). However, the choice of the methodology to be followed, is left to the experimenter who will need to consider the available laboratory facilities and resources.

- (a) parallel method: samples with the same soil/solution ratio are prepared, as many as the time intervals at which it is desired to study the adsorption kinetics. After centrifugation and if so wished filtration, the aqueous phase of the first tube is recovered as completely as possible and is measured after, for example, 4 h, that of the second tube after 8 h, that of the third after 24, etc.
- (b) serial method: only a duplicate sample is prepared for each soil/solution ratio. At defined time intervals the mixture is centrifuged to separate the phases. A small aliquot of the aqueous phase is immediately analysed for the test substance; then the experiment continues with the original mixture. If filtration is applied after centrifugation, the laboratory should have facilities to handle filtration of small aqueous aliquots. It is recommended that the total volume of the aliquots taken not exceed 1 % of the total volume of the solution, in order not to change significantly the soil/solution ratio and to decrease the mass of solute available for adsorption during the test.

The percentage adsorption  $A_{t_i}$  is calculated at each time point ( $t_i$ ) on the basis of the nominal initial concentration and the measured concentration at the sampling time ( $t_i$ ), corrected for the value of the blank. Plots of the  $A_{t_i}$  versus time (Fig. 1 Appendix 5) are generated in order to estimate the achievement of equilibrium plateau<sup>(1)</sup>. The  $K_d$  value at equilibrium is also calculated. Based on this  $K_d$  value, appropriate soil/solution ratios are selected from Fig.1, so that the percentage adsorption reaches above 20 % and preferably >50 % (61). All the applicable equations and principles of plotting are given in section on Data and Reporting and in Appendix 5.

#### 1.9.2.2. *Determination of adsorption equilibration time and of the amount of test substance adsorbed at equilibrium*

As already mentioned, plots of  $A_{t_i}$  or  $C_{aq}^{ads}$  versus time permit estimation of the achievement of the adsorption equilibrium and the amount of test substance adsorbed at equilibrium. Figs. 1 and 2 in the Appendix 5 show examples of such plots. Equilibration time is the system needs to reach a plateau.

<sup>(1)</sup> Plots of the concentration of the test substance in the aqueous phase ( $C_{aq}^{ads}$ ) versus time could also be used to estimate the achievement of the equilibrium plateau (see Fig. 2 in Appendix 5).

**▼B**

If, with a particular soil, no plateau but a steady increase is found, this may be due to complicating factors such as biodegradation or slow diffusion. Biodegradation can be shown by repeating the experiment with a sterilised sample of the soil. If no plateau is achieved even in this case, the experimenter should search for other phenomena that could be involved in his specific studies; this could be done with appropriate modifications of the experiment conditions (temperature, shaking times, soil/solution ratios). It is left to the experimenter to decide whether to continue the test procedure in spite of a possible failure to achieve an equilibrium.

1.9.2.3. *Adsorption on the surface of the test vessel and stability of the test substance*

Some information on the adsorption of the test substance on the surface of test vessels, as well as its stability, can be derived by analysing the control samples. If a depletion more than the standard error of the analytical method is observed, abiotic degradation and/or adsorption on the surface of the test vessel could be involved. Distinction between these two phenomena could be achieved by thoroughly washing the walls of the vessel with a known volume of an appropriate solvent and subjecting the wash solution to analysis for the test substance. If no adsorption on the surface of the test vessels is observed, the depletion demonstrates abiotic instability of the test substance. If adsorption is found, changing the material of the test vessels is necessary. However, data on the adsorption on the surface of the test vessels gained from this experiment cannot be directly extrapolated to soil/solution experiment. The presence of soil will affect this adsorption.

Additional information on the stability of the test substance can be derived by determination of the parental mass balance over time. This means that the aqueous phase, extracts of soil and test vessel walls are analysed for the test substance. The difference between the mass of the test chemical added and the sum of the test chemical masses in the aqueous phase, extracts of the soil and test vessel walls is equal to the mass degraded and/or volatilised and/or not extracted. In order to perform a mass balance determination, the adsorption equilibrium should have been reached within the period of the experiment.

The mass balance is performed on both soils and for one soil/solution ratio per soil that gives a depletion above 20 % and preferably > 50 % at equilibrium. When the ratio-finding experiment is completed with the analysis of the last sample of the aqueous phase after 48 h, the phases are separated by centrifugation and, if so wished, filtration. The aqueous phase is recovered as much as possible, and a suitable extraction solvent (extraction coefficient of at least 95 %) is added to the soil to extract the test substance. At least two successive extractions are recommended. The amount of test substance in the soil and test vessel extracts is determined and the mass balance is calculated (equation 10, Data and Reporting). If it is less than 90 %, the test substance is considered to be unstable in the time scale of the test. However, studies could still be continued, taking into account the instability of the test substance; in this case it is recommended to analyse both phases in the main study.

**▼ B**1.9.3. *Tier 2 — Adsorption kinetics at one concentration of the test substance*

Five soils are used, selected from Table 1. There is an advantage to including some or all of the soils used in the preliminary study, if appropriate, among these five soils. In this case, Tier 2 has not to be repeated for the soils used in preliminary study.

The equilibration time, the soil/solution ratio, the weight of the soil sample, the volume of the aqueous phase in contact with the soil and concentration of the test substance in the solution are chosen based on the preliminary study results. Analysis should preferably be done approximately after 2, 4, 6, 8 (possibly also 10) and 24 h contact time; the agitation time may be extended to a maximum of 48 h in case a chemical requires longer equilibration time with respect to ratio-finding results. However, times of analysis could be considered with flexibility.

Each experiment (one soil and one solution) is done at least in duplicate to allow estimation of the variance of the results. In every experiment one blank is run. It consists of the soil and 0,01 M CaCl<sub>2</sub> solution, without test substance, and of weight and volume, respectively, identical to those of the experiment. A control sample with only the test substance in 0,01 M CaCl<sub>2</sub> solution (without soil) is subjected to the same test procedure, serving to safeguard against the unexpected.

The percentage adsorption is calculated at each time point  $A_t$  and/or time interval  $A_{\Delta t}$  (according to the need) and is plotted versus time. The distribution coefficient  $K_d$  at equilibrium, as well as the organic carbon normalised adsorption coefficient  $K_{oc}$  (for non-polar organic chemicals), are also calculated.

## Results of the adsorption kinetics test

The linear  $K_d$  value is generally accurate to describe sorptive behaviour in soil (35)(78) and represents an expression of inherent mobility of chemicals in soil. For example, in general chemicals with  $K_d \leq 1 \text{ cm}^3 \text{ g}^{-1}$  are considered to be qualitatively mobile. Similarly, a mobility classification scheme based on  $K_{oc}$  values has been developed by MacCall *et al.* (16). Additionally, leaching classification schemes exist based on a relationship between  $K_{oc}$  and DT-50<sup>(1)</sup> (32)(79).

Also, according to error analysis studies (61),  $K_d$  values below  $0,3 \text{ cm}^3 \text{ g}^{-1}$  cannot be estimated accurately from a decrease in concentration in the aqueous phase, even when the most favourable (from point of view of accuracy) soil/solution ratio is applied, i.e. 1:1. In this case analysis of both phases, soil and solution, is recommended.

<sup>(1)</sup> DT-50: degradation time for 50 % of the test substance.

**▼ B**

With respect to the above remarks, it is recommended that the study of the adsorptive behaviour of a chemical in soil and its potential mobility be continued by determining Freundlich adsorption isotherms for these systems, for which an accurate determination of  $K_d$  is possible with the experimental protocol followed in this test method. Accurate determination is possible if the value which results by multiplying the  $K_d$  with the soil/solution ratio is  $> 0,3$ , when measurements are based on concentration decrease in the aqueous phase (indirect method), or  $> 0,1$ , when both phases are analysed (direct method) (61).

1.9.4. *Tier 3 — Adsorption isotherms and desorption kinetics/desorption isotherms*

1.9.4.1. Adsorption isotherms

Five test substance concentrations are used, covering preferably two orders of magnitude; in the choice of these concentrations the water solubility and the resulting aqueous equilibrium concentrations should be taken into account. The same soil/solution ratio per soil should be kept along the study. The adsorption test is performed as described above, with the only difference that the aqueous phase is analysed only once at the time necessary to reach equilibrium as determined before in Tier 2. The equilibrium concentrations in the solution are determined and the amount adsorbed is calculated from the depletion of the test substance in the solution or with the direct method. The adsorbed mass per unit mass of soil is plotted as a function of the equilibrium concentration of the test substance (see Data and Reporting).

Results from the adsorption isotherms experiment

Among the mathematical adsorption models proposed so far, the Freundlich isotherm is the one most frequently used to describe adsorption processes. More detailed information on the interpretation and importance of adsorption models is provided in the references (41)(45)(80)(81)(82).

**Note:** it should be mentioned that a comparison of  $K_F$  (Freundlich adsorption coefficient) values for different substances is only possible if these  $K_F$  values are expressed in the same units (83).

1.9.4.2. Desorption kinetics

The purpose of this experiment is to investigate whether a chemical is reversibly or irreversibly adsorbed on a soil. This information is important, since the desorption process also plays an important role in the behaviour of a chemical in field soil. Moreover, desorption data are useful inputs in the computer modelling of leaching and dissolved run-off simulation. If a desorption study is desired, it is recommended that the study described below be carried out on each system for which an accurate determination of  $K_d$  in the preceding adsorption kinetics experiment was possible.

Likewise with the adsorption kinetics study, there are two options to proceed with the desorption kinetics experiment: (a) the parallel method and (b) serial method. The choice of methodology to be followed, is left to the experimenter who will need to consider the available laboratory facilities and resources.

**▼B**

- (a) parallel method: for each soil which is chosen to proceed with the desorption study, samples with the same soil/solution ratio are prepared, as many as the time intervals at which it is desired to study the desorption kinetics. Preferably, the same time intervals as in the adsorption kinetics experiment should be used; however, the total time may be extended as appropriate in order the system to reach desorption equilibrium. In every experiment (one soil, one solution) one blank is run. It consists of the soil and 0,01 M CaCl<sub>2</sub> solution, without test substance, and of weight and volume, respectively, identical to those of the experiment. As a control sample the test substance in 0,01 M CaCl<sub>2</sub> solution (without soil) is subjected to the same test procedure. All the mixtures of the soil with the solution is agitating until to reach adsorption equilibrium (as determined before in Tier 2). Then, the phases are separated by centrifugation and the aqueous phases are removed as much as possible. The volume of solution removed is replaced by an equal volume of 0,01 M CaCl<sub>2</sub> without test substance and the new mixtures are agitated again. The aqueous phase of the first tube is recovered as completely as possible and is measured after, for example, 2 h, that of the second tube after 4 h, that of the third after 6 h, etc. until the desorption equilibrium is reached.
- (b) serial method: after the adsorption kinetics experiment, the mixture is centrifuged and the aqueous phase is removed as much as possible. The volume of solution removed is replaced by an equal volume of 0,01 M CaCl<sub>2</sub> without test substance. The new mixture is agitated until the desorption equilibrium is reached. During this time period, at defined time intervals, the mixture is centrifuged to separate the phases. A small aliquot of the aqueous phase is immediately analysed for the test substance; then, the experiment continues with the original mixture. The volume of each individual aliquot should be less than 1 % of the total volume. The same quantity of fresh 0,01 M CaCl<sub>2</sub> solution is added to the mixture to maintain the soil to solution ratio, and the agitation continues until the next time interval.

The percentage desorption is calculated at each time point ( $D_t$ ) and/or time interval ( $D_{\Delta t}$ ) (according to the needs of the study) and is plotted versus time. The desorption coefficient of  $K_{des}$  at equilibrium is also calculated. All applicable equations are given in Data and Reporting and Appendix 5.

#### Results from desorption kinetics experiment

Common plots of the percentage desorption  $D_t$  and adsorption  $A_t$  versus time, allow estimation of the reversibility of the adsorption process. If the desorption equilibrium is attained even within twice the time of the adsorption equilibrium, and the total desorption is more than 75 % of the amount adsorbed, the adsorption is considered to be reversible.

#### 1.9.4.3. Desorption isotherms

Freundlich desorption isotherms are determined on the soils used in the adsorption isotherms experiment. The desorption test is performed as described in the section 'Desorption kinetics', with the only difference that the aqueous phase is analysed only once, at desorption equilibrium. The amount of the test substance desorbed is calculated. The content of test substance remaining adsorbed on soil at desorption equilibrium is plotted as a function of the equilibrium concentration of the test substance in solution (see Data and Reporting and Appendix 5).

**▼B****2. DATA AND REPORTING**

The analytical data are presented in tabular form (see Appendix 6). Individual measurements and averages calculated are given. Graphical representations of adsorption isotherms are provided. The calculations are made as described below.

For the purpose of the test, it is considered that the weight of 1 cm<sup>3</sup> of aqueous solution is 1g. The soil/solution ratio may be expressed in units of w/w or w/vol with the same figure.

**2.1. ADSORPTION**

The adsorption  $A_{t_i}$  is defined as the percentage of substance adsorbed on the soil related to the quantity present at the beginning of the test, under the test conditions. If the test substance is stable and does not adsorb significantly to the container wall,  $A_{t_i}$  is calculated at each time point  $t_i$ , according to the equation:

$$A_{t_i} = \frac{m_s^{\text{ads}}(t_i) \cdot 100}{m_0} (\%) \quad (3)$$

where:

$A_{t_i}$  = adsorption percentage at the time point  $t_i$  (%);

$m_s^{\text{ads}}(t_i)$  = mass of the test substance adsorbed on the soil at the time  $t_i$  (µg);

$m_0$  = mass of the test substance in the test tube, at the beginning of the test (µg).

Detailed information on how to calculate the percentage of adsorption  $A_{t_i}$  for the parallel and serial methods is given in Appendix 5.

The distribution coefficient  $K_d$  is the ratio between the content of the substance in the soil phase and the mass concentration of the substance in the aqueous solution, under the test conditions, when adsorption equilibrium is reached.

$$K_d = \frac{C_s^{\text{ads}}(\text{eq})}{C_{\text{aq}}^{\text{ads}}(\text{eq})} = \frac{m_s^{\text{ads}}(\text{eq})}{m_{\text{aq}}^{\text{ads}}(\text{eq})} \cdot \frac{V_0}{m_{\text{soil}}} (\text{cm}^3 \text{ g}^{-1}) \quad (4)$$

where:

$C_s^{\text{ads}}(\text{eq})$  = content of the substance adsorbed on the soil at adsorption equilibrium (µg g<sup>-1</sup>);

$C_{\text{aq}}^{\text{ads}}(\text{eq})$  = mass concentration of the substance in the aqueous phase at adsorption equilibrium (µg cm<sup>-3</sup>). This concentration is analytically determined taking into account the values given by the blanks;

$m_s^{\text{ads}}(\text{eq})$  = mass of the substance adsorbed on the soil at adsorption equilibrium (µg);

$m_{\text{aq}}^{\text{ads}}(\text{eq})$  = mass of the substance in the solution at adsorption equilibrium (µg);

$m_{\text{soil}}$  = quantity of the soil phase, expressed in dry mass of soil (g);

$V_0$  = initial volume of the aqueous phase in contact with the soil (cm<sup>3</sup>).

The relation between  $A_{\text{eq}}$  and  $K_d$  is given by:

$$K_d = \frac{A_{\text{eq}}}{100 - A_{\text{eq}}} \cdot \frac{V_0}{m_{\text{soil}}} (\text{cm}^3 \text{ g}^{-1}) \quad (5)$$

**▼ B**

where:

$A_{eq}$  = percentage of adsorption at adsorption equilibrium, %.

The organic carbon normalised adsorption coefficient  $K_{oc}$  relates the distribution coefficient  $K_d$  to the content of organic carbon of the soil sample:

$$K_{oc} = K_d \cdot \frac{100}{\%OC} \text{ (cm}^3 \text{ g}^{-1}\text{)} \quad (6)$$

where:

%OC = percentage of organic carbon in the soil sample (g g<sup>-1</sup>).

$K_{oc}$  coefficient represents a single value which characterises the partitioning mainly of non-polar organic chemicals between organic carbon in the soil or sediment and water. The adsorption of these chemicals is correlated with the organic content of the sorbing solid (7); thus,  $K_{oc}$  values depend on the specific characteristics of the humic fractions which differ considerably in sorption capacity, due to differences in origin, genesis, etc.

### 2.1.1. Adsorption isotherms

The Freundlich adsorption isotherms equation relates the amount of the test substance adsorbed to the concentration of the test substance in solution at equilibrium (equation 8).

The data are treated as under 'Adsorption' and, for each test tube, the content of the test substance adsorbed on the soil after the adsorption test ( $C_s^{ads}(eq)$ , elsewhere denoted as x/m) is calculated. It is assumed that equilibrium has been attained and that  $C_s^{ads}(eq)$  represents the equilibrium value:

$$C_s^{ads}(eq) = \frac{m_s^{ads}(eq)}{m_{soil}} = \frac{[C_0 - C_{aq}^{ads}(eq)] \cdot V_0}{m_{soil}} \text{ (}\mu\text{g g}^{-1}\text{)} \quad (7)$$

The Freundlich adsorption equation is shown in (8):

$$C_s^{ads}(eq) = K_F^{ads} \cdot C_{aq}^{ads}(eq)^{1/n} \text{ (}\mu\text{g g}^{-1}\text{)} \quad (8)$$

or in the linear form:

$$\log C_s^{ads}(eq) = \log K_F^{ads} + 1/n \cdot \log C_{aq}^{ads}(eq) \quad (9)$$

where:

$K_F^{ads}$  = Freundlich adsorption coefficient; its dimension is cm<sup>3</sup> g<sup>-1</sup> only if 1/n = 1; in all other cases, the slope 1/n is introduced in the dimension of  $K_F^{ads}$  ( $\mu\text{g}^{1-1/n} \text{ (cm}^3\text{)}^{1/n} \text{ g}^{-1}$ );

n = regression constant; 1/n generally ranges between 0,7-1,0 indicating that sorption data is frequently slightly non-linear.

Equations (8) and (9) are plotted and the values of  $K_F^{ads}$  and 1/n are calculated by regression analysis using the equation 9. The correlation coefficient  $r^2$  of the log equation is also calculated. An example of such plots is given in Fig. 2.

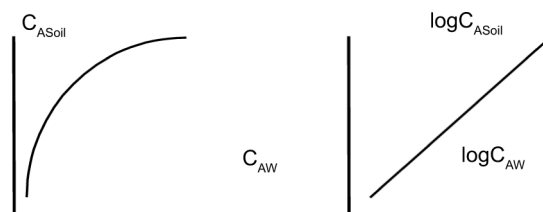
▼ B

Fig. 2 Freundlich Adsorption Plot, normal and linearised

### 2.1.2. Mass balance

The mass balance (MB) is defined as the percentage of substance which can be analytically recovered after an adsorption test versus the nominal amount of substance at the beginning of the test.

The treatment of data will differ if the solvent is completely miscible with water. In the case of water-miscible solvent, the treatment of data described under 'Desorption' may be applied to determine the amount of substance recovered by solvent extraction. If the solvent is less miscible with water, the determination of the amount recovered has to be made.

The mass balance MB for the adsorption is calculated as follows; it is assumed that the term ( $m_E$ ) corresponds to the sum of the test chemical masses extracted from the soil and surface of the test vessel with an organic solvent:

$$MB = \frac{(V_{rec} \cdot C_{aq}^{ads}(eq) + m_E) \cdot 100}{V_0 \cdot C_0} (\%) \quad (10)$$

where:

MB = mass balance (%);

$m_E$  = total mass of test substance extracted from the soil and walls of the test vessel in two steps ( $\mu\text{g}$ );

$C_0$  = initial mass concentration of the test solution in contact with the soil ( $\mu\text{g cm}^{-3}$ );

$V_{rec}$  = volume of the supernatant recovered after the adsorption equilibrium ( $\text{cm}^3$ ).

### 2.2. DESORPTION

The desorption (D) is defined as the percentage of the test substance which is desorbed, related to the quantity of substance previously adsorbed, under the test conditions:

$$D_{t_i} = \frac{m_{aq}^{des}(t_i)}{m_s^{ads}(eq)} \cdot 100(\%) \quad (11)$$

where:

$D_{t_i}$  = desorption percentage at a time point  $t_i$ (%);



**▼ B**

$m_{\text{aq}}^{\text{des}}(t_i)$  = mass of the test substance desorbed from soil at a time point  $t_i$  ( $\mu\text{g}$ );

$m_{\text{s}}^{\text{ads}}(\text{eq})$  = mass of the test substance adsorbed on soil at adsorption equilibrium ( $\mu\text{g}$ ).

Detailed information on how to calculate the percentage of desorption  $D_t$  for the parallel and serial methods is given in Appendix 5.

The apparent desorption coefficient ( $K_{\text{des}}$ ) is, under the test conditions, the ratio between the content of the substance remaining in the soil phase and the mass concentration of the desorbed substance in the aqueous solution, when desorption equilibrium is reached:

$$K_{\text{des}} = \frac{m_{\text{s}}^{\text{ads}}(\text{eq}) - m_{\text{aq}}^{\text{des}}(\text{eq}) V_{\text{T}}}{m_{\text{aq}}^{\text{des}}(\text{eq}) m_{\text{soil}}} \text{ (cm}^3 \text{ g}^{-1}\text{)} \quad (12)$$

where:

$K_{\text{des}}$  = desorption coefficient ( $\text{cm}^3 \text{ g}^{-1}$ );

$m_{\text{aq}}^{\text{des}}(\text{eq})$  = total mass of the test substance desorbed from soil at desorption equilibrium ( $\mu\text{g}$ );

$V_{\text{T}}$  = total volume of the aqueous phase in contact with the soil during the desorption kinetics test ( $\text{cm}^3$ ).

Guidance for calculating the  $m_{\text{aq}}^{\text{des}}(\text{eq})$  is given in Appendix 5 under the heading 'Desorption'.

Remark

If the adsorption test which was preceded, was performed with the parallel method the volume  $V_{\text{T}}$  in the equation (12) is considered to be equal to  $V_0$ .

### 2.2.1. Desorption isotherms

The Freundlich desorption isotherms equation relates the content of the test substance remaining adsorbed on the soil to the concentration of the test substance in solution at desorption equilibrium (equation 16).

For each test tube, the content of the substance remaining adsorbed on soil at desorption equilibrium is calculated as follows:

$$C_{\text{s}}^{\text{des}}(\text{eq}) = \frac{m_{\text{s}}^{\text{ads}}(\text{eq}) - m_{\text{aq}}^{\text{des}}(\text{eq})}{m_{\text{soil}}} \text{ (}\mu\text{g g}^{-1}\text{)} \quad (13)$$

$m_{\text{aq}}^{\text{des}}(\text{eq})$  is defined as:

$$m_{\text{aq}}^{\text{des}}(\text{eq}) = m_{\text{m}}^{\text{des}}(\text{eq}) \cdot \frac{V_0}{V_{\text{r}}} - m_{\text{aq}}^{\text{A}} \text{ (}\mu\text{g)} \quad (14)$$

where:

$C_{\text{s}}^{\text{des}}(\text{eq})$  = content of the test substance remaining adsorbed on the soil at desorption equilibrium ( $\mu\text{g g}^{-1}$ );

$m_{\text{m}}^{\text{des}}(\text{eq})$  = mass of substance determined analytically in the aqueous phase at desorption equilibrium ( $\mu\text{g}$ );

**▼ B**

$m_{\text{aq}}^{\text{A}}$  = mass of the test substance left over from the adsorption equilibrium due to incomplete volume replacement ( $\mu\text{g}$ );

$m_{\text{aq}}^{\text{des}}(\text{eq})$  = mass of the substance in the solution at adsorption equilibrium ( $\mu\text{g}$ );

$$m_{\text{aq}}^{\text{A}} = m_{\text{aq}}^{\text{ads}}(\text{eq}) \cdot \left( \frac{V_0 - V_{\text{R}}}{V_0} \right) \quad (15)$$

$V_{\text{r}}^{\text{F}}$  = volume of the solution taken from the tube for the measurement of the test substance, at desorption equilibrium ( $\text{cm}^3$ );

$V_{\text{R}}$  = volume of the supernatant removed from the tube after the attainment of adsorption equilibrium and replaced by the same volume of a 0,01 M  $\text{CaCl}_2$  solution ( $\text{cm}^3$ );

The Freundlich desorption equation is shown in (16):

$$C_{\text{s}}^{\text{des}}(\text{eq}) = K_{\text{F}}^{\text{des}} \cdot C_{\text{aq}}^{\text{des}}(\text{eq})^{1/n} \quad (\mu\text{g g}^{-1}) \quad (16)$$

or in the linear form:

$$\log C_{\text{s}}^{\text{des}}(\text{eq}) = \log K_{\text{F}}^{\text{des}} + 1/n \cdot \log C_{\text{aq}}^{\text{des}}(\text{eq}) \quad (17)$$

where:

$K_{\text{F}}^{\text{des}}$  = Freundlich desorption coefficient;

$n$  = regression constant;

$C_{\text{aq}}^{\text{des}}(\text{eq})$  = mass concentration of the substance in the aqueous phase at desorption equilibrium ( $\mu\text{g cm}^{-3}$ ).

The equations (16) and (17) can be plotted and the value of  $K_{\text{F}}^{\text{des}}$  and  $1/n$  are calculated by regression analysis using the equation 17.

Remark:

if the Freundlich adsorption or desorption exponent  $1/n$  is equal to 1, the Freundlich adsorption or desorption binding constant ( $K_{\text{F}}^{\text{ads}}$  and  $K_{\text{F}}^{\text{des}}$ ) will be equal to the adsorption or desorption equilibrium constants ( $K_{\text{d}}$  and  $K_{\text{des}}$ ) respectively, and plots of  $C_{\text{s}}$  vs  $C_{\text{aq}}$  will be linear. If the exponents are not equal to 1, plots of  $C_{\text{s}}$  vs  $C_{\text{aq}}$  will be non-linear and the adsorption and desorption constants will vary along the isotherms.

### 2.2.2. Test report

The test report should include the following information:

- complete identification of the soil samples used including,
- geographical reference of the site (latitude, longitude),
- date of sampling,

**▼ B**

- use pattern (e.g. agricultural soil, forest, etc.),
- depth of sampling,
- sand/silt/clay content;
- pH values (in 0,01 M CaCl<sub>2</sub>),
- organic carbon content,
- organic matter content,
- nitrogen content,
- C/N ratio,
- cation Exchange Capacity (mmol/kg),
- all information relating to the collection and storage of soil samples,
- where appropriate, all relevant information for the interpretation of the adsorption — desorption of the test substance,
- reference of the methods used for the determination of each parameter,
- information on the test substance as appropriate,
- temperature of the experiments,
- centrifugation conditions,
- analytical procedure used to analyse the test substance,
- justification for any use of solubilising agent for the preparation of the stock solution of the test substance,
- explanations of corrections made in the calculations, if relevant,
- data according to the form sheet (Appendix 6) and graphical presentations,
- all information and observations helpful for the interpretation of the test results.

### 3. REFERENCES

- (1) Kukowski H. and Brümmer G., (1987) Investigations on the Adsorption and Desorption of Selected Chemicals in Soils. UBA Report 106 02 045, Part II.
- (2) Fränzle O., Kuhnt G. and Vetter L., (1987) Selection of Representative Soils in the EC-Territory. UBA Report 106 02 045, Part I.
- (3) Kuhnt G. and Muntau H. (Eds.) EURO-Soils: Identification, Collection, Treatment, Characterisation. Special Publication No 1.94.60, Joint Research Centre. European Commission, ISPRA, December 1994.
- (4) OECD Test Guidelines Programme, Final Report of the OECD Workshop on Selection of Soils/Sediments, Belgirate, Italy, 18-20 January 1995 (June 1995).
- (5) US-Environment Protection Agency: Pesticide Assessment Guidelines, Subdivision N, Chemistry: Environmental Fate, Series 163-1, Leaching and Adsorption/Desorption Studies, Addendum 6 on Data Reporting, 540/09-88-096, Date: 1/1988.

**▼B**

- (6) US-Environment Protection Agency: Prevention, Pesticides and Toxic Substances, OPPTS Harmonized Test Guidelines, Series 835-Fate, Transport and Transformation Test Guidelines, OPPTS No: 835.1220 Sediment and Soil Adsorption/-Desorption Isotherm. EPA No: 712-C-96-048, April 1996.
- (7) ASTM Standards, E 1195-85, Standard Test Method for Determining a Sorption Constant ( $K_{oc}$ ) for an Organic Chemical in Soil and Sediments.
- (8) Agriculture Canada: Environmental Chemistry and Fate. Guidelines for registration of pesticides in Canada, 15 July 1987.
- (9) Netherlands Commission Registration Pesticides (1995): Application for registration of a pesticide. Section G. Behaviour of the product and its metabolites in soil, water and air.
- (10) Danish National Agency of Environmental Protection (October 1988): Criteria for registration of pesticides as especially dangerous to health or especially harmful to the environment.
- (11) BBA (1990) Guidelines for the Official Testing of Plant Protection Products, Biological Research Centre for Agriculture and Forestry, Braunschweig, Germany.
- (12) Calvet R., (1989) 'Evaluation of adsorption coefficients and the prediction of the mobilities of pesticides in soils', in Methodological Aspects of the Study of Pesticide Behaviour in Soil (ed. P. Jamet), INRA, Paris, (Review).
- (13) Calvet R., (1980) 'Adsorption-Desorption Phenomena' in Interactions between herbicides and the soil. (R.J. Hance ed.), Academic Press, London, p. 83-122.
- (14) Hasset J.J., and Banwart W.L., (1989), 'The sorption of nonpolar organics by soils and sediments' in Reactions and Movement of Organic Chemicals in Soils. Soil Science Society of America (S.S.S.A), Special Publication No 22, p. 31-44.
- (15) van Genuchten M. Th., Davidson J.M., and Wierenga P.J., (1974) 'An evaluation of kinetic and equilibrium equations for the prediction of pesticide movement through porous media'. Soil Sci. Soc. Am. Proc., Vol. 38(1), p. 29-35.
- (16) McCall P.J., Laskowski D.A., Swann R.L., and Dishburger H.J., (1981) 'Measurement of sorption coefficients of organic chemicals and their use, in environmental fate analysis', in Test Protocols for Environmental Fate and Movement of Toxicants. Proceedings of AOAC Symposium, AOAC, Washington DC.
- (17) Lambert S.M., Porter P.E., and Schiefferrstein R.H., (1965) 'Movement and sorption of chemicals applied to the soil'. Weeds, 13, p. 185-190.
- (18) Rhodes R.C., Belasco I.J., and Pease H.L., (1970) 'Determination of mobility and adsorption of agrochemicals in soils'. J.Agric.Food Chem., 18, p. 524-528.
- (19) Russell M.H., (1995), 'Recommended approaches to assess pesticide mobility in soil' in Environmental Behavior of Agrochemicals (ed. T.R. Roberts and P.C. Kearney). John Wiley & Sons Ltd.
- (20) Esser H.O., Hemingway R.J., Klein W., Sharp D.B., Vonk J.W. and Holland P.T., (1988) 'Recommended approach to the evaluation of the environmental behavior of pesticides', IUPAC Reports on Pesticides (24). Pure Appl. Chem., 60, p. 901-932.

**▼B**

- (21) Guth J.A., Burkhard N., and D.O. Eberle, (1976) 'Experimental models for studying the persistence of pesticides in soils'. Proc. BCPC Symposium: Persistence of Insecticides and Herbicides, p. 137-157, BCPC, Surrey, UK.
- (22) Furminge C.G.L., and Osgerby J.M., (1967) 'Persistence of herbicides in soil'. J. Sci. Food Agric., 18, p. 269-273.
- (23) Burkhard N., and Guth J.A., (1981) 'Chemical hydrolysis of 2-Chloro-4,6-bis(alkylamino)-1,3,5-triazine herbicides and their breakdown in soil under the influence of adsorption'. Pestic. Sci. 12, p. 45-52.
- (24) Guth J.A., Gerber H.R., and Schlaepfer T., (1977) 'Effect of adsorption, movement and persistence on the biological availability of soil-applied pesticides'. Proc. Br. Crop Prot. Conf., 3, p. 961-971.
- (25) Osgerby J.M., (1973) 'Process affecting herbicide action in soil'. Pestic. Sci., 4, p. 247-258.
- (26) Guth J.A., (1972) 'Adsorptions- und Einwascheverhalten von Pflanzenschutzmitteln in Böden'. Schr. Reihe Ver. Wass.-Boden-Lufthyg. Berlin-Dahlem, Heft 37, p. 143-154.
- (27) Hamaker J.W., (1975) 'The interpretation of soil leaching experiments', in Environmental Dynamics of Pesticides (eds R. Haque and V.H. Freed), p. 135-172, Plenum Press, NY.
- (28) Helling C.S., (1971) 'Pesticide mobility in soils'. Soil Sci. Soc. Amer. Proc., 35, p. 732-210.
- (29) Hamaker J.W., (1972) 'Diffusion and volatilization' in Organic chemicals in the soil environment (C.A.I. Goring and J.W. Hamaker eds), Vol. I, p. 49-143.
- (30) Burkhard N. and Guth J.A., (1981) 'Rate of volatilisation of pesticides from soil surfaces; Comparison of calculated results with those determined in a laboratory model system'. Pestic. Sci. 12, p. 37-44.
- (31) Cohen S.Z., Creeger S.M., Carsel R.F., and Enfield C.G., (1984) 'Potential pesticide contamination of groundwater from agricultural uses', in Treatment and Disposal of Pesticide Wastes, p. 297-325, Acs Symp. Ser. 259, American Chemical Society, Washington, DC.
- (32) Gustafson D.I., (1989) 'Groundwater ubiquity score: a simple method for assessing pesticide leachability'. J. Environ. Toxic. Chem., 8(4), p. 339-357.
- (33) Leistra M., and Dekkers W.A., (1976) 'Computed effects of adsorption kinetics on pesticide movement in soils'. J. of Soil Sci., 28, p. 340-350.
- (34) Bromilov R.H., and Leistra M., (1980) 'Measured and simulated behavior of aldicarb and its oxydation products in fallow soils'. Pest. Sci., 11, p. 389-395.
- (35) Green R.E., and Karickhoff S.W., (1990) 'Sorption estimates for modeling', in Pesticides in the Soil Environment: Process, Impacts and Modeling (ed. H.H. Cheng). Soil Sci. Soc. Am., Book Series No 2, p. 80-101,
- (36) Lambert S.M., (1967) 'Functional relationship between sorption in soil and chemical structure'. J. Agri. Food Chem., 15, p. 572-576.

**▼B**

- (37) Hance R.J., (1969) 'An empirical relationship between chemical structure and the sorption of some herbicides by soils'. *J. Agri. Food Chem.*, 17, p. 667-668.
- (38) Briggs G.G. (1969) 'Molecular structure of herbicides and their sorption by soils'. *Nature*, 223, p. 1288.
- (39) Briggs G.G. (1981) 'Theoretical and experimental relationships between soil adsorption, octanol-water partition coefficients, water solubilities, bioconcentration factors, and the parachor'. *J. Agric. Food Chem.*, 29, p. 1050-1059.
- (40) Sabljic A., (1984) 'Predictions of the nature and strength of soil sorption of organic polutance by molecular topology'. *J. Agric. Food Chem.*, 32, p. 243-246.
- (41) Bailey G.W., and White J.L., (1970) 'Factors influencing the adsorption, desorption, and movement of pesticides in soil'. *Residue Rev.*, 32, p. 29-92.
- (42) Bailey G.W., J.L. White and Y. Rothberg., (1968) 'Adsorption of organic herbicides by montomorillonite: Role of pH and chemical character of adsorbate'. *Soil Sci. Soc. Amer. Proc.* 32, p. 222-234.
- (43) Karickhoff S.W., (1981) 'Semi-empirical estimation of sorption of hydrophobic pollutants on natural sediments and soils'. *Chemosphere* 10, p. 833-846.
- (44) Paya-Perez A., Riaz M. and Larsen B., (1989) 'Soil Sorption of 6 Chlorobenzenes and 20 PCB Congeners'. *Environ. Toxicol. Safety* 21, p. 1-17.
- (45) Hamaker J.W., and Thompson J.M., (1972) 'Adsorption in organic chemicals' in *Organic Chemicals in the Soil Environment* (Goring C.A.I. and Hamaker J.W., eds), Vol I and II, Marcel Dekker, Inc., New York, NY, 1972, p. 49-143.
- (46) Deli J., and Warren G.F., 1971 'Adsorption, desorption and leaching of diphenamid in soils'. *Weed Sci.* 19, p. 67-69.
- (47) Chu-Huang Wu, Buehring N., Davinson J.M. and Santelmann, (1975) 'Napropamide Adsorption, desorption and Movement in soils'. *Weed Science*, Vol. 23, p. 454-457.
- (48) Haues M.H.B., Stacey M., and Thompson J.M., (1968) 'Adsorption of s-triazine herbicides by soil organic preparations' in *Isotopes and Radiation in Soil Organic Studies*, p. 75, International Atomic Energy Agency, Vienna.
- (49) Pionke H.B., and Deangelis R.J., (1980) 'Methods for distributing pesticide loss in field run-off between the solution and adsorbed phase', CREAMS, in *A Field Scale Model for Chemicals, Run-off and Erosion from Agricultural Management Systems*, Chapter 19, Vol. III: Supporting Documentation, USDA Conservation Research report.
- (50) ISO Standard Compendium Environment: Soil Quality — General aspects; chemical and physical methods of analysis; biological methods of analysis. First Edition (1994).
- (51) Scheffer F., and Schachtschabel, *Lehrbuch der Bodenkunde*, F. Enke Verlag, Stuttgart (1982) 11th edition.
- (52) Black, Evans D.D., White J.L., Ensminger L.E., and Clark F.E., eds. 'Methods of Soil Analysis', Vol 1 and 2, American Society of Agronomy, Madison, WI, 1982.

**▼B**

- (53) ISO/DIS 10381-1 Soil Quality — Sampling — Part 1: Guidance on the design of sampling programmes.
- (54) ISO/DIS 10381-2 Soil Quality — Sampling — Part 2: Guidance on sampling techniques.
- (55) ISO/DIS 10381-3 Soil Quality — Sampling — Part 3: Guidance on safety of sampling.
- (56) ISO/DIS 10381-4 Soil Quality — Sampling — Part 4: Guidance on the investigation of natural and cultivated soils.
- (57) ISO/DIS 10381-5 Soil Quality — Sampling — Part 5: Guidance on the investigation of soil contamination of urban and industrial sites.
- (58) ISO 10381-6, 1993: Soil Quality — Sampling — Part 6: Guidance on the collection, handling and storage of soil for the assessment of aerobic microbial processes in the laboratory.
- (59) Green R.E., and Yamane V.K., (1970) 'Precision in pesticide adsorption measurements'. *Soil Sci. Am. Proc.*, 34, 353-354.
- (60) Grover R., and Hance R.J. (1970) 'Effect of ratio of soil to water on adsorption of linuron and atrazine'. *Soil Sci.*, p. 109-138.
- (61) Boesten, J.J.T.I., 'Influence of soil/liquid ratio on the experimental error of sorption coefficients in pesticide/soil system'. *Pest. Sci.* 1990, 30, p. 31-41.
- (62) Boesten, J.J.T.I. 'Influence of soil/liquid ratio on the experimental error of sorption coefficients in relation to OECD guideline 106' Proceedings of 5th international workshop on environmental behaviour of pesticides and regulatory aspects, Brussels, 26-29 April 1994.
- (63) Bastide J., Cantier J.M., et Coste C., (1980) 'Comportement de substances herbicides dans le sol en fonction de leur structure chimique'. *Weed Res.* 21, p. 227-231.
- (64) Brown D.S., and Flagg E.W., (1981) 'Empirical prediction of organic pollutants sorption in natural sediments'. *J. Environ.Qual.*, 10(3), p. 382-386.
- (65) Chiou C.T., Porter P.E., and Schmedding D.W., (1983) 'Partition equilibria of non-ionic organic compounds between soil organic matter and water'. *Environ. Sci. Technol.*, 17(4), p. 227-231.
- (66) Gerstl Z., and Mingelgrin U., (1984) 'Sorption of organic substances by soils and sediments'. *J. Environm. Sci. Health*, B19 (3), p. 297-312.
- (67) Vowles P.D., and Mantoura R.F.C., (1987), 'Sediment-water partition coefficient and HPLC retention factors of aromatic hydrocarbons'. *Chemosphere*, 16(1), p. 109-116.
- (68) Lyman W.J., Reehl W.F. and Rosenblatt D.H. (1990) *Handbook of Chemical Property Estimation Methods. Environmental Behaviour of Organic Compounds.* American Chemical Society, Washington DC.
- (69) Keniga E.E., and Goring, C.A.I. (1980) 'Relationship between water solubility, soil sorption, octanol-water partitioning and concentration of chemicals in the biota' in *Aquatic Toxicology* (eds J.G. Eaton, et al.), p. 78-115, ASTM STP 707, Philadelphia.

## ▼B

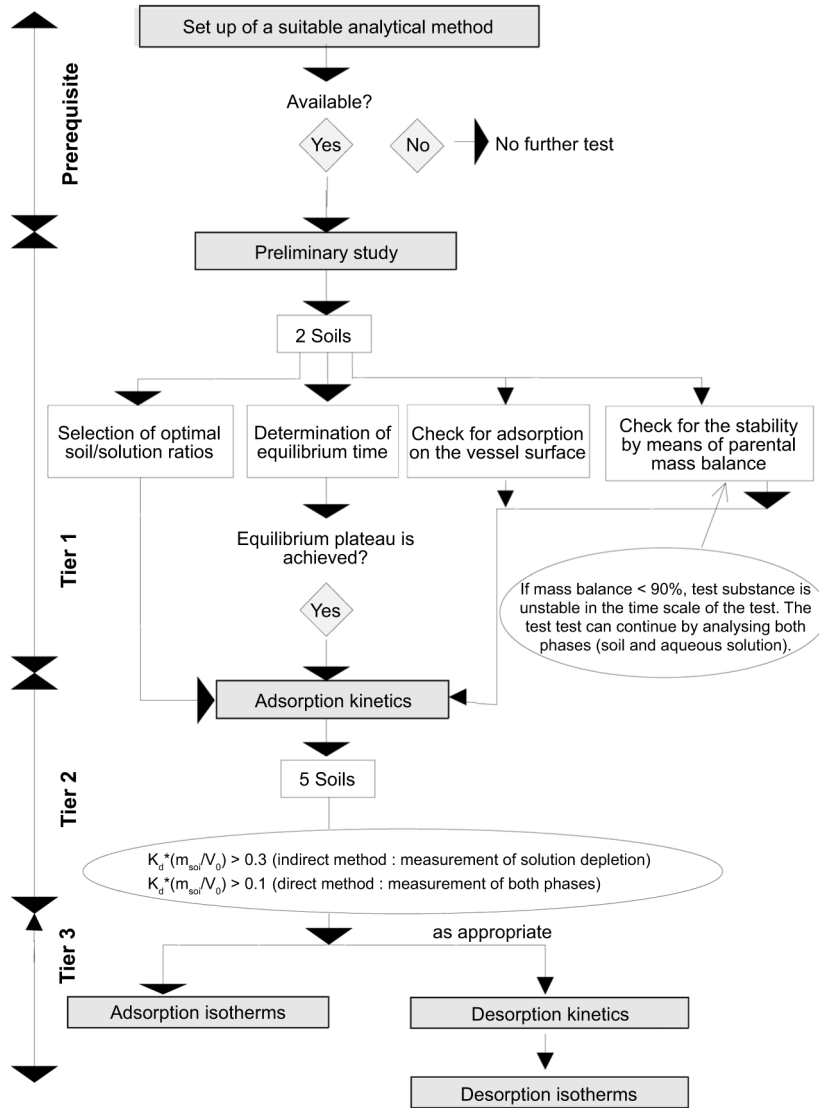
- (70) Chiou C.T., Peters L.J., and Freed V.H., (1979) 'A physical concept of soil-water equilibria for non-ionic organic compounds'. *Science*, Vol. 206, p. 831-832.
- (71) Hassett J.J., Banwart W.I., Wood S.G., and Means J.C., (1981) 'Sorption of *p*-Naphthol: implications concerning the limits of hydrophobic sorption'. *Soil Sci. Soc. Am. J.* 45, p. 38-42.
- (72) Karickhoff S.W., (1981), 'Semi-empirical estimation of sorption of hydrophobic pollutants on natural sediments and soils'. *Chemosphere*, Vol. 10(8), p. 833-846.
- (73) Moreale A., van Bladel R., (1981) 'Adsorption de 13 herbicides et insecticides par le sol. Relation solubilité — réactivité'. *Revue de l'Agric.*, 34 (4), p. 319-322.
- (74) Müller M., Kördel W. (1996) 'Comparison of screening methods for the determination/estimation of adsorption coefficients on soil'. *Chemosphere*, 32(12), p. 2493-2504.
- (75) Kördel W., Kotthoff G., Müller M. (1995) 'HPLC — screening method for the determination of the adsorption coefficient on soil — results of a ring test'. *Chemosphere* 30 (7), p. 1373-1384.
- (76) Kördel W., Stutte J., Kotthoff G. (1993), 'HPLC — screening method for the determination of the adsorption coefficient on soil — comparison of different stationary phases. *Chemosphere* 27 (12), p. 2341-2352.
- (77) Hance, R.J., (1967), 'The speed of Attainment of Sorption Equilibria in Some Systems Involving Herbicides'. *Weed Research*, Vol. 7, p. 29-36.
- (78) Koskinen W.C., and Harper S.S., (1990), 'The retention processes: mechanisms' in *Pesticides in the Soil Environment: Processes, Impacts and Modelling* (ed. H.H. Cheng). *Soil Sci. Soc. Am. Book Series*, No 2, Madison, Wisconsin.
- (79) Cohen S.Z., Creeger S.M., Carsel R.F., and Enfield C.G. (1984), 'Potential pesticide contamination of groundwater from agricultural uses', in *Treatment and Disposal of Pesticide Wastes*, p. 297-325, ACS Symp. Ser. 259, American Chemical Society, Washington, DC.
- (80) Giles C.H., (1970) 'Interpretation and use of sorption isotherms' in *Sorption and Transport Processes in Soils*. S.C.I. Monograph No. 37, p. 14-32.
- (81) Giles, C.H.; McEwan J.H.; Nakhwa, S.N. and Smith, D. (1960) 'Studies in adsorption: XI. A system of classification of solution adsorption isotherms and its use in the diagnosis of adsorption mechanisms and in measurements of pesticides surface areas of soils'. *J. Chem. Soc.*, p. 3973-93.
- (82) Calvet R., Tercé M., and Arvien J.C., (1980) 'Adsorption des pesticides par les sols et leurs constituants: 3. Caractéristiques générales de l'adsorption'. *Ann. Agron.* 31, p. 239-251.
- (83) Bedbur E., (1996) 'Anomalies in the Freundlich equation', *Proc. COST 66 Workshop, Pesticides in soil and the environment*, 13-15 May 1996, Stratford-upon-Avon, U.K.
- (84) Guth, J.A., (1985) 'Adsorption/desorption', in *Joint International Symposium, Physicochemical Properties and their Role in Environmental Hazard Assessment*, July 1-3, Canterbury, UK.
- (85) Soil Texture Classification (US and FAO systems): *Weed Science*, 33, Suppl. 1 (1985) and *Soil Sci. Soc. Amer. Proc.* 26, p. 305 (1962).



▼B

Appendix 1

Testing scheme



## ▼B

## Appendix 2

**INFLUENCE OF ACCURACY OF ANALYTICAL METHOD AND  
CONCENTRATION CHANGE ON ACCURACY OF ADSORPTION  
RESULTS**

From the following table (84) it becomes obvious that when the difference between the initial mass ( $m_0 = 110 \mu\text{g}$ ) and equilibrium mass ( $m_{\text{aq}}^{\text{ads}}(\text{eq}) = 100 \mu\text{g}$ ) of the test substance in the solution is very small, an error of 5 % in the measurement of equilibrium concentration results in an error of 50 % in the calculation of the mass of the substance adsorbed in soil ( $m_{\text{s}}^{\text{ads}}(\text{eq})$ ) and of 52,4 % in the calculation of the  $K_{\text{d}}$ .

Amount of soil  $m_{\text{soil}} = 10 \text{ g}$

Volume of solution  $V_0 = 100 \text{ cm}^3$

	$m_{\text{aq}}^{\text{ads}}(\text{eq})$ ( $\mu\text{g}$ )	$C_{\text{aq}}^{\text{ads}}(\text{eq})$ ( $\mu\text{g cm}^{-3}$ )	R	$m_{\text{s}}^{\text{ads}}(\text{eq})^*$ ( $\mu\text{g}$ )	$C_{\text{s}}^{\text{ads}}(\text{eq})^*$ ( $\mu\text{g g}^{-1}$ )	$R_{\ddagger}$	$K_{\text{d}}^*$	$R_{\ddagger}$
	<b>FOR A = 9 %</b>							
$m_0 = 110 \mu\text{g}$ or $C_0 = 1,100 \mu\text{g/cm}^3$	100	1,000	true value	10	1,00	true value	1	
	101	1,010	1 %	9	0,90	10 %	0,891	10,9 %
	105	1,050	5 %	5	0,50	50 %	0,476	52,4 %
	109	1,090	9 %	1	0,10	90 %	0,092	90,8 %
	<b>FOR A = 55 %</b>							
$m_0 = 110 \mu\text{g}$ or $C_0 = 1,100 \mu\text{g/cm}^3$	50,0	0,500	true value	60,0	6,00	true value	12,00	
	50,5	0,505	1 %	59,5	5,95	0,8 %	11,78	1,8 %
	52,5	0,525	5 %	57,5	5,75	4,0 %	10,95	8,8 %
	55,0	0,550	10 %	55,0	5,50	8,3 %	10,00	16,7 %
	<b>FOR A = 99 %</b>							
$m_0 = 110 \mu\text{g}$ or $C_0 = 1,100 \mu\text{g/cm}^3$	1,100	0,011	true value	108,9	10,89	true value	990	
	1,111	0,01111	1 %	108,889	10,8889	0,01 %	980	1,0 %
	1,155	0,01155	5 %	108,845	10,8845	0,05 %	942	4,8 %
	1,21	0,0121	10 %	108,790	10,8790	0,10 %	899	9,2 %

$$*m_{\text{s}}^{\text{ads}}(\text{eq}) = m_0 - m_{\text{aq}}^{\text{ads}}(\text{eq}), C_{\text{s}}^{\text{ads}}(\text{eq}) = \frac{[C_0 - C_{\text{aq}}^{\text{ads}}(\text{eq})] V_0}{m_{\text{soil}}} \cdot K_{\text{d}} = \frac{m_{\text{s}}^{\text{ads}}(\text{eq})}{m_{\text{aq}}^{\text{ads}}(\text{eq})} \frac{V_0}{m_{\text{soil}}}$$

$m_{\text{s}}^{\text{ads}}(\text{eq})$  = mass of the test substance in the soil phase at equilibrium,  $\mu\text{g}$ ;

$m_{\text{aq}}^{\text{ads}}(\text{eq})$  = mass of the test substance in the aqueous phase at equilibrium,  $\mu\text{g}$ ;

$C_{\text{s}}^{\text{ads}}(\text{eq})$  = content of the test substance in the soil phase at equilibrium,  $\mu\text{g g}^{-1}$ ;

$C_{\text{aq}}^{\text{ads}}(\text{eq})$  = mass concentration of the test substance in the aqueous phase at equilibrium,  $\mu\text{g cm}^{-3}$ ;

R = analytical error in the determination of the  $m_{\text{aq}}^{\text{ads}}(\text{eq})$ ;

$R_{\ddagger}$  = calculated error due to the analytical error R.



## Appendix 3

ESTIMATION TECHNIQUES FOR  $K_D$ 

1. Estimation techniques permit prediction of  $K_d$  based on correlations with, for example,  $P_{ow}$  values (12)(39)(63-68), water solubility data (12)(19)(21)(39)(68-73), or polarity data derived by application of HPLC on reversed phase (74-76). As shown in Tables 1 and 2, is the  $K_{oc}$  or  $K_{om}$  that are calculated from these equations and then, indirectly, the  $K_d$  from the equations:

$$K_{oc} = K_d \cdot \frac{100}{\%oc} \text{ (cm}^3 \text{ g}^{-1}) \quad K_{om} = \frac{K_d}{1,724} \cdot \frac{100}{\%oc} \text{ (cm}^3 \text{ g}^{-1})$$

2. The concept of these correlations is based on two assumptions: (1) it is the organic matter of the soil that mainly influences the adsorption of a substance; and (2) the interactions involved are mainly non-polar. As a result, these correlations: (1) are not, or are only to some extent, applicable to polar substances, and (2) are not applicable in cases where the organic matter content of the soil is very small (12). In addition, although satisfactory correlations have been found between  $P_{ow}$  and adsorption (19), the same cannot be said for the relationship between water solubility and extent of adsorption (19)(21); so far the studies are very contradictory.
3. Some examples of correlations between the adsorption coefficient and the octanol-water partition coefficient, as well as water solubility are given in Tables 1 and 2, respectively.

Table 1

**Examples of correlations between the adsorption distribution coefficient and the octanol-water partition coefficient; for further examples (12) (68).**

Substances	Correlations	Authors
Substituted ureas	$\log K_{om} = 0,69 + 0,52 \log P_{ow}$	Briggs (1981) (39)
Aromatic chlorinated	$\log K_{oc} = - 0,779 + 0,904 \log P_{ow}$	Chiou <i>et al.</i> (1983) (65)
Various pesticides	$\log K_{om} = 4,4 + 0,72 \log P_{ow}$	Gerstl and Mingelgrin (1984) (66)
Aromatic hydrocarbons	$\log K_{oc} = - 2,53 + 1,15 \log P_{ow}$	Vowles and Mantoura (1987) (67)

Table 2

**Examples of correlations between the adsorption distribution coefficient and water solubility; for further examples see (68) (69).**

Compounds	Correlations	Authors
Various pesticides	$\log K_{om} = 3,8 - 0,561 \log S_w$	Gerstl and Mingelgrin (1984) (66)
Aliphatic, aromatic chlorinated substances	$\log K_{om} = (4,040 \pm 0,038) - (0,557 \pm 0,012) \log S_w$	Chiou <i>et al.</i> (1979) (70)
$\alpha$ -naphthol	$\log K_{oc} = 4,273 - 0,686 \log S_w$	Hasset <i>et al.</i> (1981) (71)
Cyclic, aliphatic aromatic substances	$\log K_{oc} = - 1,405 - 0,921 \log S_w - 0,00953 \text{ (mp-25)}$	Karickhoff (1981) (72)
Various compounds	$\log K_{om} = 2,75 - 0,45 \log S_w$	Moreale van Blade (1982) (73)

▼ **B**

## Appendix 4

**CALCULATIONS FOR DEFINING THE CENTRIFUGATION CONDITIONS**

1. The centrifugation time is given by the following formula, assuming spherical particles:

$$t = \frac{9}{2} \left[ \frac{\eta}{\omega^2 r_p^2 (\rho_s - \rho_{aq})} \right] \ln (R_b/R_t) \quad (1)$$

For simplification purposes, all parameters are described in non-SI units (g, cm).

where:

$\omega$  = rotational speed (=2  $\pi$  rpm/60), rad s<sup>-1</sup>

rpm = revolutions per minute

$\eta$  = viscosity of solution, g s<sup>-1</sup> cm<sup>-1</sup>

$r_p$  = particle radius, cm

$\rho_s$  = soil density, g cm<sup>-3</sup>

$\rho_{aq}$  = solution density, g cm<sup>-3</sup>

$R_t$  = distance from the centre of centrifuge rotor to top of solution in centrifuge tube, cm

$R_b$  = distance from the centre of centrifuge rotor to bottom in centrifuge tube, cm

$R_b - R_t$  = length of the soil/solution mixture in the centrifuge tube, cm.

In general practice, double the calculated times is used to ensure complete separation.

2. The equation (1) can be simplified further if we consider the viscosity ( $\eta$ ) and the density ( $\rho_{aq}$ ) of the solution as equal to the viscosity and density of water at 25 °C; thus,  $\eta = 8,95 \times 10^{-3}$  g s<sup>-1</sup> cm<sup>-1</sup> and  $\rho_{aq} = 1,0$  g. cm<sup>-3</sup>.

Then, the centrifugation time is given by the equation (2):

$$t = \frac{3.7}{(\text{rpm})^2 \cdot r_p^2 (\rho_s - 1)} \ln \frac{R_b}{R_t} \quad (2)$$

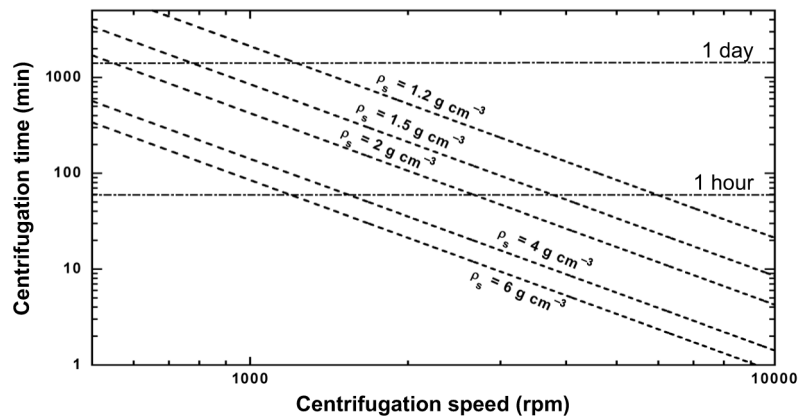
3. From the equation (2) it becomes apparent that two parameters are important in defining the centrifugation condition, i.e. time (t) and speed (rpm), in order to achieve separation of particles with a specific size (in our case 0,1  $\mu\text{m}$  radius): (1) the density of the soil and (2) the length of the mixture in the centrifuge tube ( $R_b - R_t$ ), i.e. the distance which a soil particle covers from the top of the solution to the bottom of the tube; obviously, for a fixed volume the length of the mixture in the tube will depend on the square of the radius of the tube.
4. Fig. 1 presents variations in the centrifugation time (t) versus centrifugation speed (rpm) for different soil densities ( $\rho_s$ ) (Fig.1a) and different lengths of the mixture in the centrifuge tubes (Fig.2a). From Fig.1a the influence of the soil density appears obvious; for example, for a classical centrifugation of 3 000 rpm the centrifugation time is approximately 240 min for 1,2 g cm<sup>3</sup> soil density, while it is only 50 min for 2,0 g cm<sup>3</sup>. Similarly, from Fig 1b, for a classical centrifugation of 3 000 rpm the centrifugation time is approximately 50 min for a length of the mixture of 10 cm and only seven min for a length of 1 cm. However, it is important to find an optimal relation between centrifugation which requires the less length possible and easy handling for the experimenter in separating the phases after centrifugation.

## ▼ B

5. Moreover, when defining the experimental conditions for the separation of soil/solution phases, it is important to consider the possible existence of a third 'pseudo-phase', the colloids. These particles, with a size less than  $0,2 \mu\text{m}$ , can have an important impact on the whole adsorption mechanism of a substance in a soil suspension. When centrifugation is performed as described above, colloids remain in the aqueous phase and are subjected to analysis together with the aqueous phase. Thus, the information about their impact is lost.

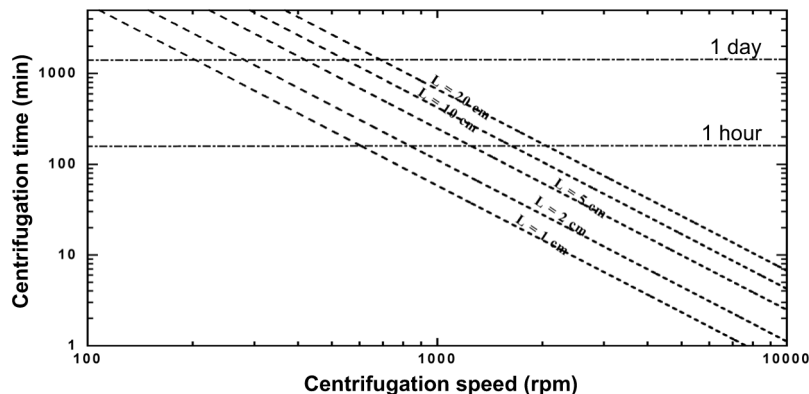
If the conducting laboratory has ultracentrifugation or ultrafiltration facilities, the adsorption/desorption of a substance in soil could be studied more in depth, including information on the adsorption of the substance on the colloids. In this case, an ultracentrifugation at 60 000 rpm/min or an ultrafiltration with filter porosity of 100 000 Daltons should be applied in order to separate the three phases soil, colloids, solution. The test protocol should also be modified accordingly, in order all three phases to be subjected to substance analysis.

Fig. 1a.



Variations of centrifugation time ( $t$ ) versus centrifugation speed (rpm) for different soil densities ( $\rho_s$ ).  $R_t = 10 \text{ cm}$ ,  $R_b - R_t = 10 \text{ cm}$ ,  $\eta = 8,95 \times 10^{-3} \text{ g s}^{-1} \text{ cm}^{-1}$  and  $\rho_{\text{aq}} = 1,0 \text{ g. cm}^{-3}$  at  $25 \text{ }^\circ\text{C}$ .

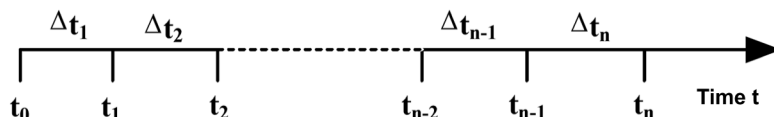
Fig. 1b.



Variations of centrifugation time ( $t$ ) versus centrifugation speed (rpm) for different lengths of the mixture in the centrifuge tube ( $R_b - R_t$ ) =  $L$ ;  $R_t = 10 \text{ cm}$ ,  $\eta = 8,95 \times 10^{-3} \text{ g s}^{-1} \text{ cm}^{-1}$ ,  $\rho_{\text{aq}} = 1,0 \text{ g. cm}^{-3}$  at  $25 \text{ }^\circ\text{C}$  and  $\rho_s = 2,0 \text{ g cm}^{-3}$ .

**▼ B***Appendix 5***CALCULATION OF ADSORPTION A (%) AND DESORPTION D (%)**

The time scheme of the procedure is:



For all the calculations it is assumed that the test substance is stable and does not adsorb significantly to the container walls.

**ADSORPTION A (A%)**a) *Parallel method*

The percentage adsorption is calculated for each test tube (i) at each time point ( $t_i$ ), according to the equation:

$$A_{t_i} = \frac{m_s^{\text{ads}}(t_i) \cdot 100}{m_0} (\%) \quad (1)$$

The terms of this equation may be calculated as follows:

$$m_0 = C_0 \cdot V_0 (\mu\text{g}) \quad (2)$$

$$m_s^{\text{ads}}(t_i) = m_0 - C_{\text{aq}}^{\text{ads}}(t_i) \cdot V_0 (\mu\text{g}) \quad (3)$$

where:

$A_{t_i}$  = adsorption percentage (%) at the time point  $t_i$

$m_s^{\text{ads}}(t_i)$  = mass of the test substance on soil at the time  $t_i$  that the analysis is performed ( $\mu\text{g}$ )

$m_0$  = mass of test substance in the test tube, at the beginning of the test ( $\mu\text{g}$ )

$C_0$  = initial mass concentration of the test solution in contact with the soil ( $\mu\text{g cm}^{-3}$ )

**▼ B**

$C_{aq}^{ads}(t_i)$  = mass concentration of the substance in the aqueous phase at the time  $t_i$  that the analysis is performed ( $\mu\text{g cm}^{-3}$ ); this concentration is analytically determined taking into account the values given by the blanks

$V_0$  = initial volume of the test solution in contact with the soil ( $\text{cm}^3$ ).

The values of the adsorption percentage  $A_{t_i}$  or  $C_{aq}^{ads}(t_i)$  are plotted versus time and the time after which the sorption equilibrium is attained is determined. Examples of such plots are given in Fig. 1 and Fig. 2 respectively.

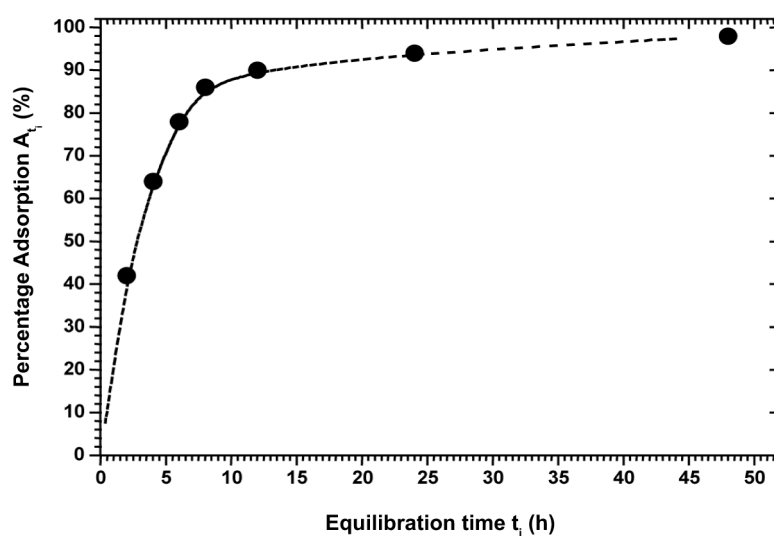


Fig. 1.

Adsorption equilibrium plot

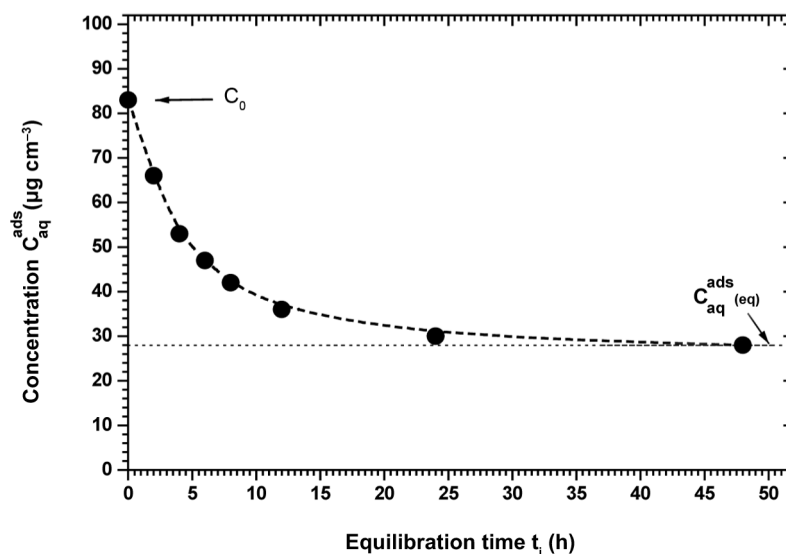


Fig. 2.

Mass concentration of the test substance in the aqueous phase ( $C_{aq}$ ) versus time

**▼ B**b) *Serial method*

The following equations take into account that the adsorption procedure is carried out by measurements of the test substance in small aliquots of the aqueous phase at specific time intervals.

During each time interval the amount of the substance adsorbed on the soil is calculated as follows:

— for the first time interval  $\Delta t_1 = t_1 - t_0$

$$m_s^{\text{ads}}(\Delta t_1) = m_0 - m_m^{\text{ads}}(t_1) \cdot \left( \frac{V_0}{V_a^A} \right) \quad (4)$$

— for the second time interval  $\Delta t_2 = t_2 - t_1$

$$m_s^{\text{ads}}(\Delta t_2) = m_m^{\text{ads}}(t_1) \cdot \left( \frac{V_0}{V_a^A} \right) - m_m^{\text{ads}}(t_2) \cdot \left( \frac{V_0 - V_a^A}{V_a^A} \right) \quad (5)$$

— for the third time interval  $\Delta t_3 = t_3 - t_2$

$$m_s^{\text{ads}}(\Delta t_3) = m_m^{\text{ads}}(t_2) \cdot \left( \frac{V_0 - V_a^A}{V_a^A} \right) - m_m^{\text{ads}}(t_3) \cdot \left( \frac{V_0 - 2 \cdot V_a^A}{V_a^A} \right) \quad (6)$$

— for the nth time interval  $\Delta t_n = t_n - t_{n-1}$

$$m_s^{\text{ads}}(\Delta t_n) = m_m^{\text{ads}}(t_{n-1}) \cdot \left( \frac{V_0 - (n-2) \cdot V_a^A}{V_a^A} \right) - m_m^{\text{ads}}(t_n) \cdot \left( \frac{V_0 - (n-1) \cdot V_a^A}{V_a^A} \right) \quad (7)$$

The percentage of adsorption at each time interval,  $A_{\Delta t_i}$ , is calculated using the following equation:

$$A_{\Delta t_i} = \frac{m_s^{\text{ads}}(\Delta t_i)}{m_0} \cdot 100(\%) \quad (8)$$

while the percentage of adsorption ( $A_{t_i}$ ) at a time point  $t_i$  is given by the equation:

$$A_{t_i} = \frac{\sum_{j=\Delta t_1}^{\Delta t_i} m_s^{\text{ads}}(j)}{m_0} \cdot 100(\%) \quad (9)$$

The values of the adsorption  $A_{t_i}$  or  $A_{\Delta t_i}$  (with respect to the needs of the study) are plotted versus time and the time after which the sorption equilibrium is attained is determined.

At the equilibration time  $t_{\text{eq}}$ :

— the mass of the test substance adsorbed on the soil is:

$$m_s^{\text{ads}}(\text{eq}) = \sum_{\Delta t_i=1}^n m_s^{\text{ads}}(\Delta t_i) \quad (10)$$



**▼ B**

— the mass of the test substance in the solution is:

$$m_{\text{aq}}^{\text{ads}}(\text{eq}) = m_0 - \sum_{\Delta t_i=1}^n m_s^{\text{ads}}(\Delta t_i) \quad (11)$$

— and the percentage of adsorption at equilibrium is:

$$A_{\text{eq}} = \frac{m_s^{\text{ads}}(\text{eq})}{m_0} \cdot 100(\%) \quad (12)$$

The parameters used above are defined as:

$m_s^{\text{ads}}(\Delta t_1), m_s^{\text{ads}}(\Delta t_2), \dots, m_s^{\text{ads}}(\Delta t_n)$  = mass of the substance adsorbed on the soil during the time intervals  $\Delta t_1, \Delta t_2, \dots, \Delta t_n$  respectively ( $\mu\text{g}$ );

$m_m^{\text{ads}}(t_1), m_m^{\text{ads}}(t_2), \dots, m_m^{\text{ads}}(t_n)$  = mass of the substance measured in an aliquot  $v_a^A$  at the time points  $t_1, t_2, t_n$  respectively ( $\mu\text{g}$ );

$m_s^{\text{ads}}(\text{eq})$  = mass of the substance adsorbed on the soil at adsorption equilibrium ( $\mu\text{g}$ );

$m_{\text{aq}}^{\text{ads}}(\text{eq})$  = mass of the substance in the solution at adsorption equilibrium ( $\mu\text{g}$ );

$v_a^A$  = volume of the aliquot in which the test substance is measured ( $\text{cm}^3$ );

$A_{\Delta t_i}$  = percentage of adsorption corresponding at a time interval  $\Delta t_i$  (%);

$A_{\text{eq}}$  = percentage of adsorption at adsorption equilibrium (%).

**DESORPTION D (A %)**

The time  $t_0$  that the desorption kinetics experiment begins, is considered as the moment that the maximal recovered volume of the test substance solution (after that the adsorption equilibrium is attained) is replaced by an equal volume of 0,01 M  $\text{CaCl}_2$  solution.

**(a) Parallel method**

At a time point  $t_i$ , the mass of the test substance is measured in the aqueous phase taken from the tube  $i$  ( $V_r^i$ ), and the mass desorbed is calculated according to the equation:

$$m_{\text{aq}}^{\text{des}}(t_i) = m_m^{\text{des}}(t_i) \cdot \left( \frac{V_0}{V_r^i} \right) - m_{\text{aq}}^A \quad (13)$$

At desorption equilibrium  $t_i = t_{\text{eq}}$  and therefore  $m_{\text{aq}}^{\text{des}}(t_i) = m_{\text{aq}}^{\text{des}}(\text{eq})$

The mass of the test substance desorbed during a time interval ( $\Delta t_i$ ) is given by the equation:

$$m_{\text{aq}}^{\text{des}}(\Delta t_i) = m_{\text{aq}}^{\text{des}}(t_i) - \sum_{j=1}^{i-1} m_{\text{aq}}^{\text{des}}(j) \quad (14)$$

The percentage of desorption is calculated:

at a time point  $t_i$  from the equation:

**▼ B**

$$D_{t_i} = \frac{m_{\text{aq}}^{\text{des}}(t_i)}{m_{\text{s}}^{\text{ads}}(\text{eq})} \cdot 100 \quad (\%) \quad (15)$$

and during a time interval ( $\Delta t_i$ ) from the equation:

$$D_{\Delta t_i} = \frac{m_{\text{aq}}^{\text{des}}(\Delta t_i)}{m_{\text{s}}^{\text{ads}}(\text{eq})} \cdot 100 \quad (\%) \quad (16)$$

where:

$D_{t_i}$  = desorption percentage at a time point  $t_i$  (%)

$D_{\Delta t_i}$  = desorption percentage corresponding to a time interval  $\Delta t_i$  (%)

$m_{\text{aq}}^{\text{des}}(t_i)$  = mass of the test substance desorbed at a time point  $t_i$ , ( $\mu\text{g}$ )

$m_{\text{aq}}^{\text{des}}(\Delta t_i)$  = mass of the test substance desorbed during a time interval  $\Delta t_i$  ( $\mu\text{g}$ )

$m_{\text{m}}^{\text{des}}(t_i)$  = mass of the test substance analytically measured at a time  $t_i$  in a solution volume  $V_r^i$ , which is taken for the analysis ( $\mu\text{g}$ )

$m_{\text{aq}}^{\text{A}}$  = mass of the test substance left over from the adsorption equilibrium due to incomplete volume replacement ( $\mu\text{g}$ )

$$m_{\text{aq}}^{\text{A}} = m_{\text{aq}}^{\text{ads}}(\text{eq}) \cdot \left( \frac{V_0 - V_R}{V_0} \right) \quad (17)$$

$m_{\text{aq}}^{\text{ads}}(\text{eq})$  = mass of the test substance in the solution at adsorption equilibrium ( $\mu\text{g}$ )

$V_R$  = volume of the supernatant removed from the tube after the attainment of adsorption equilibrium and replaced by the same volume of a 0,01 M  $\text{CaCl}_2$  solution ( $\text{cm}^3$ )

$V_r^i$  = volume of the solution taken from the tube (i) for the measurement of the test substance, in desorption kinetics experiment ( $\text{cm}^3$ ).

The values of desorption  $D_{t_i}$  or  $D_{\Delta t_i}$  (according to the needs of the study) are plotted versus time and the time after which the desorption equilibrium is attained is determined.

**(b) Serial method**

The following equations take into account that the adsorption procedure, which was preceded, was carried out by measurement of test substance in small aliquots ( $v_a^{\text{A}}$ ) of the aqueous phase (serial method in 'Performance of the test' 1,9). It is assumed that: (a) the volume of the supernatant removed from the tube after the adsorption kinetics experiment was replaced by the same volume of 0,01 M  $\text{CaCl}_2$  solution ( $V_R$ ) and (b) and the total volume of the aqueous phase in contact with the soil ( $V_T$ ) during the desorption kinetics experiment remains constant and is given by the equation:

$$V_T = V_0 - \sum_{i=1}^n v_a^{\text{A}}(i) \quad (18)$$

**▼ B**

At a time point  $t_i$ :

- The mass of the test substance is measured in a small aliquot ( $v_a^D$ ) and the mass desorbed is calculated, according to the equation:

$$m_{aq}^{des}(t_i) = m_m^{des}(t_i) \cdot \left( \frac{V_T}{v_a^D} \right) - m_{aq}^A \cdot \left( \frac{(V_T - (i-1) \cdot v_a^D)}{V_T} \right) \quad (19)$$

- At desorption equilibrium  $t_i = t_{eq}$  and therefore  $m_{aq}^{des}(t_i) = m_{aq}^{des}(eq)$ .

- The percentage of desorption  $D_{t_i}$  is calculated, from the following equation:

$$D_{t_i} = \frac{m_{aq}^{des}(t_i)}{m_s^{ads}(eq)} \cdot 100 (\%) \quad (20)$$

At a time interval ( $\Delta t_i$ ):

During each time interval the amount of the substance desorbed is calculated as follows:

- for the first time interval  $\Delta t_1 = t_1 - t_0$

$$m_{aq}^{des}(\Delta t_1) = m_m^{des}(t_1) \cdot \left( \frac{V_T}{v_a^D} \right) - m_{aq}^A \quad \text{and} \quad m_s^{des}(t_1) = m_s^{aq}(eq) - m_{aq}^{des}(\Delta t_1) \quad (21)$$

- for the second time interval  $\Delta t_2 = t_2 - t_1$

$$m_{aq}^{des}(\Delta t_2) = m_m^{des}(t_2) \cdot \left( \frac{V_T}{v_a^D} \right) - m_{aq}^{des}(\Delta t_1) \cdot \left( \frac{(V_T - v_a^D)}{V_T} \right) - m_{aq}^A \cdot \left( \frac{(V_T - v_a^D)}{V_T} \right)$$

and

$$m_s^{des}(t_2) = m_s^{ads}(eq) - [m_{aq}^{des}(\Delta t_1) + m_{aq}^{des}(\Delta t_2)] \quad (22)$$

- for the  $n^{\text{th}}$  interval  $\Delta t_n = t_n - t_{n-1}$

$$m_{aq}^{des}(\Delta t_n) = \left[ m_m^{des}(t_n) \cdot \left( \frac{V_T}{v_a^D} \right) - m_{aq}^A \cdot \left( \frac{(V_T - (n-1) \cdot v_a^D)}{V_T} \right) - \sum_{i=1, n \neq 1}^{n-1} \left( \frac{(V_T - (n-i) \cdot v_a^D)}{V_T} \cdot m_{aq}^{des}(\Delta t_i) \right) \right] \quad \text{and}$$

$$m_s^{des}(t_n) = m_s^{ads}(eq) - \sum_{i=1, n \neq 1}^n m_{aq}^{des}(\Delta t_i) \quad (23)$$

Finally, the percentage of desorption at each time interval,  $D_{\Delta t_i}$ , is calculated using the following equation:

$$D_{\Delta t_i} = \frac{m_{aq}^{des}(\Delta t_i)}{m_s^{ads}(eq)} \cdot 100 (\%) \quad (24)$$

while the percentage of desorption  $D_{t_i}$  at a time point  $t_i$  is given by the equation:

$$D_{t_i} = \frac{\sum_{j=\Delta t_1}^{\Delta t_i} m_{aq}^{des}(j)}{m_s^{ads}(eq)} \cdot 100 = \frac{m_{aq}^{des}(t_i)}{m_s^{ads}(eq)} \cdot 100 (\%) \quad (25)$$

**▼ B**

where the above used parameters are defined as:

$m_s^{\text{des}}(\Delta t_1), m_s^{\text{des}}(\Delta t_2), \dots, m_s^{\text{des}}(\Delta t_n)$  = mass of the substance remaining adsorbed on the soil after the time intervals  $\Delta t_1, \Delta t_2, \dots, \Delta t_n$  respectively ( $\mu\text{g}$ )

$m_{\text{aq}}^{\text{des}}(\Delta t_1), m_{\text{aq}}^{\text{des}}(\Delta t_2), \dots, m_{\text{aq}}^{\text{des}}(\Delta t_n)$  = mass of the test substance desorbed during the time intervals  $\Delta t_1, \Delta t_2, \dots, \Delta t_n$  respectively ( $\mu\text{g}$ )

$m_m^{\text{des}}(t_1), m_m^{\text{des}}(t_2), \dots, m_m^{\text{des}}(t_n)$  = mass of the substance measured in an aliquot ( $v_a^{\text{D}}$ ) at time points  $t_1, t_2, \dots, t_n$ , respectively ( $\mu\text{g}$ )

$V_T$  = total volume of the aqueous phase in contact with the soil during the desorption kinetics experiment performed with the serial method ( $\text{cm}^3$ )

$m_{\text{aq}}^{\text{A}}$  = mass of the test substance left over from the adsorption equilibrium due to incomplete volume replacement ( $\mu\text{g}$ )

$$m_{\text{aq}}^{\text{A}} = \left( \frac{\left( V_0 - \sum_{i=1}^n v_a^{\text{A}}(i) \right) - V_R}{\left( V_0 - \sum_{i=1}^n v_a^{\text{A}}(i) \right)} \right) \cdot m_{\text{aq}}^{\text{ads}}(\text{eq}) \quad (26)$$

$V_R$  = volume of the supernatant removed from the tube after the attainment of adsorption equilibrium and replaced by the same volume of a 0,01 M  $\text{CaCl}_2$  solution ( $\text{cm}^3$ )

$v_a^{\text{D}}$  = volume of the aliquot sampled for analytical purpose from the tube (i), during the desorption kinetics experiment performed with the serial method ( $\text{cm}^3$ )

$$v_a^{\text{D}} \leq 0,02 \cdot V_T \quad (27)$$

▼ **B**

## Appendix 6

**ADSORPTION-DESORPTION IN SOILS: DATA REPORTING SHEETS**

Substance tested:

Soil tested:

Dry mass content of the soil (105 °C, 12h): ..... %

Temperature: ..... °C

**Suitability of the analytical method**

Weighed soil	g	
Soil: dry mass	g	
Volume CaCl <sub>2</sub> sol.	cm <sup>3</sup>	
Nominal conc. final sol.	µg cm <sup>-3</sup>	
Analytical conc. final sol.	µg cm <sup>-3</sup>	

Principle of the analytical method used:

Calibration of the analytical method:

Substance tested:

Soil tested

Dry mass content of the soil (105 °C, 12 h): ..... %

Temperature: ..... °C

Analytical methodology followed: Indirect  Parallel  Serial   
 Direct

**Adsorption test: test samples**

	Symbol	Units	Equilibration time	Equilibration time	Equilibration time	Equilibration time
Tube No						
Weighed soil	—	g				
Soil: dry mass	m <sub>soil</sub>	g				
Water volume in weighed soil (calculated)	V <sub>WS</sub>	cm <sup>3</sup>				
Volume 0,01 M CaCl <sub>2</sub> sol. to equilibrate the soil		cm <sup>3</sup>				
Volume of stock solution		cm <sup>3</sup>				
Total volume of aq. phase in contact with soil	V <sub>0</sub>	cm <sup>3</sup>				
Initial concentration Test solution	C <sub>0</sub>	µg cm <sup>-3</sup>				
Mass test subst. at the beginning of the test	m <sub>0</sub>	µg				

**▼ B**

	Symbol	Units	Equilibration time	Equilibration time	Equilibration time	Equilibration time	Equilibration time	Equilibration time	Equilibration time
<b>After agitation and centrifugation</b>									
INDIRECT METHOD									
Parallel method									
Concentration test subst. aq. phase Blank correction included	$C_{aq}^{ads}(t_i)$	$\mu\text{g cm}^{-3}$							
Serial method									
Measured mass test subst. in the aliquot $V_a^A$	$m_m^{ads}(t_i)$	$\mu\text{g}$							
DIRECT METHOD									
Mass test substance adsorbed on soil	$m_s^{ads}(t_i)$	$\mu\text{g}$							
Calculation of adsorption									
Adsorption	$A_{t_i}$	%							
	$A_{\Delta t_i}$	%							
Means									
Adsorption coefficient	$K_d$	$\text{cm}^3 \text{g}^{-1}$							
Means									
Adsorption coefficient	$K_{oc}$	$\text{cm}^3 \text{g}^{-1}$							
Means									

Substance tested:

Soil tested:

Dry mass content of the soil (105 °C, 12 h): ..... %

Temperature: ..... °C

**Adsorption test: blanks and control**

	Symbol	Units	Blank		Blank		Control	
Tube No								
Weighed soils		g					0	0
Water amount in weighed soil (calculated)		$\text{cm}^3$					—	—
Volume of 0,01 M $\text{CaCl}_2$ solution added		$\text{cm}^3$						
Volume of the stock solution of the test substance added		$\text{cm}^3$	0	0				
Total volume of aq. phase (calculated)		$\text{cm}^3$					—	—

**▼ B**

	Symbol	Units	Blank		Blank		Control	
Initial concentration of the test substance in aqueous phase		$\mu\text{g cm}^{-3}$						
<b>After agitation and centrifugation</b>								
Concentration in aqueous phase		$\mu\text{g cm}^{-3}$						

Remark: add columns if necessary

Substance tested:

Soil tested:

Dry mass content of the soil (105 °C 12 h): ..... %

Temperature: ..... °C

**Mass balance**

	Symbol	Units				
Tube No						
Weighed soil	—	g				
Soil: dry mass	$m_{\text{soil}}$	g				
Water volume in weighed soil (calculated)	$V_{\text{WS}}$	ml				
Volume 0,01 M CaCl <sub>2</sub> sol. to equilibrate the soil		ml				
Volume of stock solution		$\text{cm}^3$				
Total volume of aq. phase in contact with soil	$V_0$	$\text{cm}^3$				
Initial concentration test solution	$C_0$	$\mu\text{g cm}^{-3}$				
Equilibration time	—	h				

**After agitation and centrifugation**

Concentr. test subst. aq. phase at adsorption equilibrium blank correction included	$C_{\text{aq}}^{\text{ads}}(\text{eq})$	$\mu\text{g cm}^{-3}$				
Equalibration time	$t_{\text{eq}}$	h				

**1st dilution with solvent**

Removed volume aq. phase	$V_{\text{rec}}$	$\text{cm}^3$				
Added volume of solvent	$\Delta V$	$\text{cm}^3$				

**1st extraction with solvent**

Signal analysed in solvent	$S_{\text{E1}}$	var.				
Conc. test subst. in solvent	$C_{\text{E1}}$	$\mu\text{g cm}^{-3}$				

**▼ B**

	Symbol	Units					
Mass of substance extracted from soil and vessel walls	$m_{E1}$	$\mu\text{g}$					
2nd dilution with solvent							
Removed volume of solvent	$\Delta V_s$	$\text{cm}^3$					
Added volume of solvent	$\Delta V'$	$\text{cm}^3$					
2nd extraction with solvent							
Signal analysed in solvent phase	$S_{E2}$	var.					
Conc. test subst. in solvent	$C_{E2}$	$\mu\text{g cm}^{-3}$					
Mass of substance extracted from soil and vessel walls	$m_{E2}$	$\mu\text{g}$					
Total mass test subst. extracted in two steps	$m_E$	$\mu\text{g}$					
Mass balance	MB	%					

Substance tested:

Soil tested:

Dry mass content of the soil (105 °C, 12 h): ..... %

Temperature: ..... °C

**Adsorption isotherms**

	Symbol	Units							
Tube No									
Weighed soil	—	g							
Soil: dry mass	E	g							
Water volume in weighed soil (calculated)	$V_{WS}$	$\text{cm}^3$							
Volume 0,01 M $\text{CaCl}_2$ sol. to equilibrate the soil		$\text{cm}^3$							
Volume of stock solution added		$\text{cm}^3$							
Total volume of aq. phase in contact with soil (calculated)	$V_0$	$\text{cm}^3$							
Concentration solution	$C_0$	$\mu\text{g cm}^{-3}$							
Equilibration time	—	h							



**▼ B**

	Symbol	Units								
<b>After agitation and centrifugation</b>										
Concentration subst. aq. phase, blank correction included	$C_{aq}^{ads} (eq)$	$\mu g\ cm^{-3}$								
Temperature		$^{\circ}C$								
Adsorb. mass per unit soil	$C_s^{ads} (eq)$	$\mu g\ g^{-1}$								

Regression analysis:

value of  $K_F^{ads}$ :value of  $1/n$ :regression coefficient  $r^2$ :

Substance tested:

Soil tested:

Dry mass content of the soil (105 °C, 12 h): ..... %

Temperature: ..... °C

Analytical methodology followed:      Indirect            Parallel            Serial      **Desorption test**

	Symbol	Units	Time interval	Time interval	Time interval	Time interval
Tube No coming from adsorption step						
Mass of substance adsorbed on soil at adsorption equilibrium	$m_s^{ads} (eq)$	$\mu g$				
Removed volume aq. phase, replaced by 0,01 M CaCl <sub>2</sub>	$V_R$	$cm^3$				
Total volume of aq. phase in contact with soil	PM	$V_0$	$cm^3$			
	SM	$V_T$	$cm^3$			
Mass test subst. left over the adsorption equilibrium due to incomplete volume replacement	$m_{aq}^A$	$\mu g$				

**Desorption kinetics**

Measured mass of substance desorbed from soil at time $t_i$	$m_m^{des} (t_i)$	$\mu g$				
Volume of the solution taken from the tube (i) for the measurement of the test substance	PM	$V_r^i$	$cm^3$			
	SM	$v_a^D$	$cm^3$			
Mass of substance desorbed from soil at time $t_i$ (calculated)	$m_{aq}^{des} (t_i)$	$\mu g$				
Mass of substance desorbed from soil during time interval $\Delta t_i$ (calculated)	$m_{aq}^{des} (\Delta t_i)$	$\mu g$				

**▼ B**

	Symbol	Units	Time interval	Time interval	Time interval	Time interval
<b>Desorption percentage</b>						
Desorption at time $t_i$	$D_{t_i}$	%				
Desorption at time interval $\Delta t_i$	$D_{\Delta t_i}$	%				
Apparent desorption coefficient	$K_{des}$					

PM: parallel method

SM: serial method

**▼ B****C.19. ESTIMATION OF THE ADSORPTION COEFFICIENT ( $K_{OC}$ ) ON SOIL AND ON SEWAGE SLUDGE USING HIGH PERFORMANCE LIQUID CHROMATOGRAPHY (HPLC)****1. METHOD**

This method is a replicate of OECD TG121 (2001).

**1.1. INTRODUCTION**

The sorption behaviour of substances in soils or sewage sludges can be described through parameters experimentally determined by means of the test method C.18. An important parameter is the adsorption coefficient which is defined as the ratio between the concentration of the substance in the soil/sludge and the concentration of the substance in the aqueous phase at adsorption equilibrium. The adsorption coefficient normalised to the organic carbon content of the soil  $K_{oc}$  is a useful indicator of the binding capacity of a chemical on organic matter of soil and sewage sludge and allows comparisons to be made between different chemicals. This parameter can be estimated through correlations with the water solubility and the n-octanol/water partition coefficient (1)(2)-(3)(4)(5)(6)(7).

The experimental method described in this test uses HPLC for the estimation of the adsorption coefficient  $K_{oc}$  in soil and in sewage sludge (8). The estimates are of higher reliability than those from QSAR calculations (9). As an estimation method it cannot fully replace batch equilibrium experiments used in the test method C.18. However, the estimated  $K_{oc}$  may be useful for choosing appropriate test parameters for adsorption/desorption studies according to the test method C.18 by calculating  $K_d$  (distribution coefficient) or  $K_f$  (Freundlich adsorption coefficient) according to the equation 3 (see Section 1.2).

**1.2. DEFINITIONS**

**$K_d$ :** distribution coefficient is defined as the ratio of equilibrium concentrations  $C$  of a dissolved test substance in a two phase system consisting of a sorbent (soil or sewage sludge) and an aqueous phase; it is a dimensionless value when concentrations in both phases are expressed on a weight/weight base. In case the concentration in the aqueous phase is given on a weight/volume base then the units are  $\text{ml} \cdot \text{g}^{-1}$ .  $K_d$  can vary with sorbent properties and can be concentration dependent.

$$K_d = \frac{C_{soil}}{C_{aq}} \text{ or } \frac{C_{sludge}}{C_{aq}} \quad (1)$$

where:

$C_{soil}$  = concentration of test substance in soil at equilibrium ( $\mu\text{g} \cdot \text{g}^{-1}$ )

$C_{sludge}$  = concentration of test substance in sludge at equilibrium ( $\mu\text{g} \cdot \text{g}^{-1}$ )

$C_{aq}$  = concentration of test substance in aqueous phase at equilibrium ( $\mu\text{g} \cdot \text{g}^{-1}$ ,  $\mu\text{g} \cdot \text{ml}^{-1}$ ).

**▼ B**

**K<sub>f</sub>**: Freundlich adsorption coefficient is defined as the concentration of the test substance in soil or sewage sludge ( $x/m$ ) when the equilibrium concentration  $C_{aq}$  in the aqueous phase is equal to one; units are  $\mu\text{g}\cdot\text{g}^{-1}$  sorbent. The value can vary with sorbent properties.

$$\log \frac{x}{m} = \log K_f + \frac{1}{n} \cdot \log C_{aq} \quad (2)$$

where:

$x/m$  = amount of test substance  $x$  ( $\mu\text{g}$ ) adsorbed on amount of sorbent  $m$  (g) at equilibrium

$1/n$  = slope of Freundlich adsorption isotherm

$C_{aq}$  = concentration of test substance in aqueous phase at equilibrium ( $\mu\text{g}\cdot\text{ml}^{-1}$ )

$$\text{At } C_{aq} = 1; \log K_f = \log \frac{x}{m}$$

**K<sub>oc</sub>**: distribution coefficient ( $K_d$ ) or Freundlich adsorption coefficient ( $K_f$ ) normalised to the organic carbon content ( $f_{oc}$ ) of a sorbent; particularly for non-ionised chemicals, it is an approximate indicator for the extent of adsorption between a substance and the sorbent and allows comparisons to be made between different chemicals. Depending on the dimensions of  $K_d$  and  $K_f$ ,  $K_{oc}$  can be dimensionless or have the units  $\text{ml}\cdot\text{g}^{-1}$  or  $\mu\text{g}\cdot\text{g}^{-1}$  organic matter.

$$K_{oc} = \frac{K_d}{f_{oc}} \left( \text{dimensionless or } \text{ml}\cdot\text{g}^{-1} \right) \text{ or } \frac{K_f}{f_{oc}} \left( \mu\text{g}\cdot\text{g}^{-1} \right) \quad (3)$$

The relationship between  $K_{oc}$  and  $K_d$  is not always linear and thus  $K_{oc}$  values can vary from soil to soil but their variability is greatly reduced compared to  $K_d$  or  $K_f$  values.

The adsorption coefficient ( $K_{oc}$ ) is deduced from the capacity factor ( $k'$ ) using a calibration plot of  $\log k'$  versus  $\log K_{oc}$  of the selected reference compounds.

$$k' = \frac{t_R - t_0}{t_0} \quad (4)$$

where:

$t_R$  = HPLC retention time of test and reference substance (minutes);

$t_0$  = HPLC dead time (minutes) (see Section 1.8.2).

**P<sub>ow</sub>**: The octanol-water partition coefficient is defined as the ratio of the concentrations of dissolved substance in n-octanol and water; it is a dimensionless value.

$$P_{ow} = \frac{C_{octanol}}{C_{aq}} (= K_{ow}) \quad (5)$$

### 1.3. REFERENCE SUBSTANCES

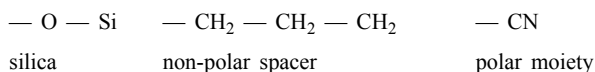
The structural formula, the purity and the dissociation constant (if appropriate) should be known before using the method. Information on solubility in water and organic solvents, on octanol-water partition coefficient and on hydrolysis characteristics is useful.

**▼B**

To correlate the measured HPLC-retention data of a test substance with its adsorption coefficient  $K_{oc}$ , a calibration graph of  $\log K_{oc}$  versus  $\log k'$  has to be established. A minimum of six reference points, at least one above and one below the expected value of the test substance should be used. The accuracy of the method will be significantly improved if reference substances that are structurally related to the test substance are used. If such data are not available, it is up to the user to select the appropriate calibration substances. A more general set of structurally heterogeneous substances should be chosen in this case. Substances and  $K_{oc}$ -values which may be used are listed in the Appendix in Table 1 for sewage sludge and in Table 3 for soil. The selection of other calibration substances should be justified.

## 1.4. PRINCIPLE OF THE TEST METHOD

HPLC is performed on analytical columns packed with a commercially available cyanopropyl solid phase containing lipophilic and polar moieties. A moderately polar stationary phase based on a silica matrix is used:



The principle of the test method is similar to testing method A.8 (Partition coefficient, HPLC method). While passing through the column along with the mobile phase the test substance interacts with the stationary phase. As a result of partitioning between mobile and stationary phases the test substance is retarded. The dual composition of the stationary phase having polar and non-polar sites allows for interaction of polar and non-polar groups of a molecule in a similar way as is the case for organic matter in soil or sewage sludge matrices. This enables the relationship between the retention time on the column and the adsorption coefficient on organic matter to be established.

pH has a significant influence on sorption behaviour in particular for polar substances. For agricultural soils or tanks of sewage treatment plants pH normally varies between pH 5,5 and 7,5. For ionisable substances, two tests should be performed with both ionised and non-ionised forms in appropriate buffer solutions but only in cases where at least 10 % of the test compound will be dissociated within pH 5,5 to 7,5.

Since only the relationship between the retention on the HPLC column and the adsorption coefficient is employed for the evaluation, no quantitative analytical method is required and only the determination of the retention time is necessary. If a suitable set of reference substances is available and standard experimental conditions can be used, the method provides a fast and efficient way to estimate the adsorption coefficient  $K_{oc}$ .

## 1.5. APPLICABILITY OF THE TEST

The HPLC method is applicable to chemical substances (unlabelled or labelled) for which an appropriate detection system (e.g. spectrophotometer, radioactivity detector) is available and which are sufficiently stable during the duration of the experiment. It may be particularly useful for chemicals which are difficult to study in other experimental systems (i.e. volatile substances; substances which are not soluble in water at a concentration which can be measured analytically; substances with a high affinity to the surface of incubation systems). The method can be used for mixtures which give unresolved elution bands. In such a case, upper and lower limits of the  $\log K_{oc}$  values of the compounds of the test mixture should be stated.

**▼B**

Impurities may sometimes cause problems for interpretation of HPLC results, but they are of minor importance as long as the test substance can analytically be clearly identified and separated from the impurities.

The method is validated for the substances listed in Table 1 in the Appendix and was also applied to a variety of other chemicals belonging to the following chemical classes:

- aromatic amines (e.g. trifluralin, 4-chloroaniline, 3,5-dinitroaniline, 4-methylaniline, N-methylaniline, 1-naphthylamine),
- aromatic carboxylic acid esters (e.g. benzoic acid methylester, 3,5-dinitrobenzoic acid ethylester),
- aromatic hydrocarbons (e.g. toluene, xylene, ethylbenzene, nitrobenzene),
- aryloxyphenoxypropionic acid esters (e.g. diclofop-methyl, fenoxaprop-ethyl, fenoxaprop-P-ethyl),
- benzimidazole and imidazole fungicides (e.g. carbendazim, fuberidazole, triazoxide),
- carboxylic acid amides (e.g. 2-chlorobenzamide, N,N-dimethylbenzamide, 3,5-dinitrobenzamide, N-methylbenzamide, 2-nitrobenzamide, 3-nitrobenzamide),
- chlorinated hydrocarbons (e.g. endosulfan, DDT, hexachlorobenzene, quintozone, 1,2,3-trichlorobenzene),
- organophosphorus insecticides (e.g. azinphos-methyl, disulfoton, fenamiphos, isofenphos, pyrazophos, sulprofos, triazophos),
- phenols (e.g. phenol, 2-nitrophenol, 4-nitrophenol, pentachlorophenol, 2,4,6-trichlorophenol, 1-naphthol),
- phenylurea derivatives (e.g. isoproturon, monolinuron, pencycuron),
- pigment dyestuffs (e.g. Acid Yellow 219, Basic Blue 41, Direct Red 81),
- polyaromatic hydrocarbons (e.g. acenaphthene, naphthalene),
- 1,3,5-triazine herbicides (e.g. prometryn, propazine, simazine, terbutryn),
- triazole derivatives (e.g. tebuconazole, triadimefon, tradimenol, triapenthenol).

The method is not applicable for substances which react either with the eluent or the stationary phase. It is also not applicable for substances that interact in a specific way with inorganic components (e.g. formation of cluster complexes with clay minerals). The method may not work for surface active substances, inorganic compounds and moderate or strong organic acids and bases. Log  $K_{oc}$  values ranging from 1,5 to 5,0 can be determined. Ionisable substances must be measured using a buffered mobile phase, but care has to be taken to avoid precipitation of buffer components or test substance.

**▼ B**

## 1.6. QUALITY CRITERIA

1.6.1. **Accuracy**

Normally, the adsorption coefficient of a test substance can be estimated to within +/- 0,5 log unit of the value determined by the batch equilibrium method (see Table 1 in the Appendix). Higher accuracy may be achieved if the reference substances used are structurally related to the test substance.

1.6.2. **Repeatability**

Determinations should be run at least in duplicate. The values of log  $K_{oc}$  derived from individual measurements should be within a range of 0,25 log unit.

1.6.3. **Reproducibility**

Experience gained so far in the application of the method is supportive of its validity. An investigation of the HPLC method, using 48 substances (mostly pesticides) for which reliable data on  $K_{oc}$  on soils were available gave a correlation coefficient of  $R = 0,95$  (10) (11).

An inter-laboratory comparison test with 11 participating laboratories was performed to improve and validate the method (12). Results are given in Table 2 of the Appendix.

## 1.7. DESCRIPTION OF THE TEST METHOD

1.7.1. **Preliminary Estimation of the Adsorption Coefficient**

The octanol-water partition coefficient  $P_{ow}$  (=  $K_{ow}$ ) and, to some extent, the water solubility can be used as indicators for the extent of adsorption, particularly for non-ionised substances, and thus may be used for preliminary range finding. A variety of useful correlations have been published for several groups of chemicals (1)(2)(3)(4)(5)(6)(7).

1.7.2. **Apparatus**

A liquid chromatograph, fitted with a pulse-free pump and a suitable detection device is required. The use of an injection valve with an injection loop is recommended. Commercial cyanopropyl chemically bound resins on a silica base shall be used (e.g. Hypersil and Zorbax CN). A guard column of the same material may be positioned between the injection system and the analytical column. Columns from different suppliers may vary considerably in their separation efficiency. As a guidance, the following capacity factors  $k'$  should be reached:  $\log k' > 0,0$  for  $\log K_{oc} = 3,0$  and  $\log k' > -0,4$  for  $\log K_{oc} = 2,0$  when using methanol/water 55/45 % as mobile phase.

1.7.3. **Mobile phases**

Several mobile phases have been tested and the following two are recommended:

— methanol/water (55/45 % v/v)

— methanol/0,01M citrate-buffer pH 6,0 (55/45 % v/v)

**▼ B**

HPLC grade methanol and distilled water or citrate-buffer are used to prepare the eluting solvent. The mixture is degassed before use. Isocratic elution should be employed. If methanol/water mixtures are not appropriate, other organic solvent/water mixtures may be tried, e.g. ethanol/water or acetonitrile/water mixtures. For ionisable compounds the use of buffer solution is recommended to stabilise pH. Care must be taken to avoid salt precipitation and column deterioration, which may occur with some organic phase/-buffer mixtures.

No additives such as ion pair reagents may be used because they can affect the sorption properties of the stationary phase. Such changes of the stationary phase may be irreversible. For this reason, it is mandatory that experiments using additives are carried out on separate columns.

**1.7.4. Solutes**

Test and reference substances should be dissolved in the mobile phase.

**1.8. PERFORMANCE OF THE TEST****1.8.1. Test condition**

The temperature during the measurements should be recorded. The use of a temperature controlled column compartment is highly recommended to guarantee constant conditions during calibration and estimation runs and measurement of the test substance.

**1.8.2. Determination of dead time  $t_0$** 

For the determination of the dead time  $t_0$  two different methods may be used (see also Section 1.2).

**1.8.2.1. *Determination of the dead time  $t_0$  by means of a homologous series***

This procedure has proven to yield reliable and standardised  $t_0$  values. For details see Testing Method A.8: Partition coefficient (n-octanol/water), HPLC method.

**1.8.2.2. *Determination of the dead time  $t_0$  by inert substances which are not retained by the column***

This technique is based on the injection of solutions of formamide, urea or sodium nitrate. Measurements should be performed at least in duplicate.

**1.8.3. Determination of the retention times  $t_R$** 

Reference substances should be selected as described in Section 1.3. They may be injected as a mixed standard to determine their retention times, provided it has been confirmed that the retention time of each reference standard is unaffected by the presence of the other reference standards. The calibration should be performed at regular intervals at least twice daily in order to account for unexpected changes in column performance. For best practice the calibration injections should be carried out before and after injections of the test substance to confirm retention times have not drifted. The test substances are injected separately in quantities as small as possible (to avoid column overload) and their retention times are determined.



**▼ B**

In order to increase the confidence in the measurement, at least duplicate determinations should be made. The values of  $\log K_{oc}$  derived from individual measurements should fall within a range of 0,25 log unit.

**1.8.4. Evaluation**

The capacity factors  $k'$  are calculated from the dead time  $t_0$  and retention times  $t_R$  of the selected reference substances according to equation 4 (see Section 1.2). The  $\log k'$  data of the reference substances are then plotted against their  $\log K_{oc}$  values from batch equilibrium experiments given in Tables 1 and 3 of the Appendix. Using this plot, the  $\log k'$  value of a test substance is then used to determine its  $\log K_{oc}$  value. If the actual results show that the  $\log K_{oc}$  of the test substance is outside the calibration range the test should be repeated using different, more appropriate reference substances.

**2. DATA AND REPORTING**

The report must include the following information:

- identity of test and reference substances and their purity, and  $pK_a$  values if relevant,
- description of equipment and operating conditions, e.g. type and dimension of analytical (and guard) column, means of detection, mobile phase (ratio of components and pH), temperature range during measurements,
- dead time and the method used for its determination,
- quantities of test and reference substances introduced in the column,
- retention times of reference compounds used for calibration,
- details of fitted regression line ( $\log k'$  vs  $\log K_{oc}$ ) and a graph of the regression line,
- average retention data and estimated  $\log K_{oc}$  value for the test compound,
- chromatograms.

**3. REFERENCES**

- (1) W.J. Lyman, W.F. Reehl, D.H. Rosenblatt (ed), (1990) Handbook of chemical property estimation methods, chapt. 4, McGraw-Hill, New York.
- (2) J. Hodson, N.A. Williams, (1988) The estimation of the adsorption coefficient ( $K_{oc}$ ) for soils by HPLC. Chemosphere, 17, p. 167.
- (3) G.G. Briggs, (1981) Theoretical and experimental relationships between soil adsorption, octanol-water partition coefficients, water solubilities, bioconcentration factors, and the parachor. J. Agric. Food Chem., 29, p. 1050-1059.
- (4) C.T. Chiou, P.E. Porter, D.W. Schmedding, (1983) Partition equilibria of nonionic organic compounds between soil organic matter and water. Environ. Sci. Technol., 17, p. 227-231.
- (5) Z. Gerstl, U. Mingelgrin (1984) Sorption of organic substances by soils and sediment. J. Environm. Sci. Health, B19, p. 297-312.

**▼B**

- (6) C.T. Chiou, L.J. Peters, V. H. Freed, (1979) A physical concept of soil water equilibria for nonionic organic compounds, *Science*, 106, p. 831-832.
- (7) S.W. Karickhoff, (1981) Semi-empirical estimation of sorption of hydrophobic pollutants on natural sediments and soils. *Chemosphere*, 10, p. 833-846.
- (8) W. Kördel, D. Hennecke, M. Herrmann, (1997) Application of the HPLC-screening method for the determination of the adsorption coefficient on sewage sludges. *Chemosphere*, 35(1/2), p. 121-128.
- (9) M. Mueller, W. Kördel (1996) Comparison of screening methods for the estimation of adsorption coefficients on soil. *Chemosphere*, 32(12), p. 2493-2504.
- (10) W. Kördel, J. Stutte, G. Kotthoff (1993) HPLC-screening method for the determination of the adsorption coefficient in soil-comparison of different stationary phases, *Chemosphere*, 27(12), p. 2341-2352.
- (11) B. von Oepen, W. Kördel, W. Klein (1991) Sorption of nonpolar and polar compounds to soils: Processes, measurements and experience with the applicability of the modified OECD Guideline 106, *Chemosphere*, 22, p. 285-304.
- (12) W. Kördel, G. Kotthoff, J. Müller (1995) HPLC-screening method for the determination of the adsorption coefficient on soil-results of a ring test. *Chemosphere*, 30(7), p. 1373-1384.



## Appendix

Table 1

Comparison of  $K_{oc}$  values for soils and sewage sludges, and calculated values by the HPLC screening method <sup>(1)</sup>, <sup>(2)</sup>

substance	CAS-No	log $K_{oc}$ sewage sludges	log $K_{oc}$ HPLC	$\Delta$	log $K_{oc}$ soils	log $K_{oc}$ HPLC	$\Delta$
Atrazine	1912-24-9	1,66	2,14	0,48	1,81	2,20	0,39
Linuron	330-55-2	2,43	2,96	0,53	2,59	2,89	0,30
Fenthion	55-38-9	3,75	3,58	0,17	3,31	3,40	0,09
Monuron	150-68-5	1,46	2,21	0,75	1,99	2,26	0,27
Phenanthrene	85-01-8	4,35	3,72	0,63	4,09	3,52	0,57
Benzoic acid phenylester	93-99-2	3,26	3,03	0,23	2,87	2,94	0,07
Benzamide	55-21-0	1,60	1,00	0,60	1,26	1,25	0,01
4-Nitrobenzamide	619-80-7	1,52	1,49	0,03	1,93	1,66	0,27
Acetanilide	103-84-4	1,52	1,53	0,01	1,26	1,69	0,08
Aniline	62-53-3	1,74	1,47	0,27	2,07	1,64	0,43
2,5-Dichloroaniline	95-82-9	2,45	2,59	0,14	2,55	2,58	0,03

<sup>(1)</sup> W. Kördel, D. Hennecke, M. Herrmann, (1997) Application of the HPLC-screening method for the determination of the adsorption coefficient on sewage sludges. Chemosphere, 35(1/2), p. 121-128.

<sup>(2)</sup> W. Kördel, D. Hennecke, C. Franke, (1997) Determination of the adsorption-coefficients of organic substances on sewage sludges. Chemosphere, 35 (1/2), p. 107-119.

Table 2

Results of a laboratory inter-comparison test (11 participating laboratories) performed to improve and validate the HPLC-method <sup>(1)</sup>

substance	CAS-No	log $K_{oc}$	$K_{oc}$	log $K_{oc}$
		[OECD 106]	[HPLC-method]	
Atrazine	1912-24-9	1,81	78 ± 16	1,89
Monuron	150-68-5	1,99	100 ± 8	2,00
Triapenthenol	77608-88-3	2,37	292 ± 58	2,47
Linuron	330-55-2	2,59	465 ± 62	2,67
Fenthion	55-38-9	3,31	2062 ± 648	3,31

<sup>(1)</sup> W. Kördel, G. Kotthoff, J. Müller (1995). HPLC-screening method for the determination of the adsorption coefficient on soil-results of a ring test. Chemosphere, 30(7), 1373-1384.



Table 3

**Recommended reference substances for the HPLC screening method based on soil adsorption data.**

Reference substance	CAS-No	log K <sub>oc</sub> mean values from batch equilibrium	number of K <sub>oc</sub> data	log S.D.	source
Acetanilide	103-84-4	1,25	4	0,48	(a)
Phenol	108-95-2	1,32	4	0,70	(a)
2-Nitrobenzamide	610-15-1	1,45	3	0,90	(b)
N,N-dimethylbenzamide	611-74-5	1,52	2	0,45	(a)
4-Methylbenzamide	619-55-6	1,78	3	1,76	(a)
Methylbenzoate	93-58-3	1,80	4	1,08	(a)
Atrazine	1912-24-9	1,81	3	1,08	(c)
Isoproturon	34123-59-6	1,86	5	1,53	(c)
3-Nitrobenzamide	645-09-0	1,95	3	1,31	(b)
Aniline	62-53-3	2,07	4	1,73	(a)
3,5-Dinitrobenzamide	121-81-3	2,31	3	1,27	(b)
Carbendazim	10605-21-7	2,35	3	1,37	(c)
Triadimenol	55219-65-3	2,40	3	1,85	(c)
Triazoxide	72459-58-6	2,44	3	1,66	(c)
Triazophos	24017-47-8	2,55	3	1,78	(c)
Linuron	330-55-2	2,59	3	1,97	(c)
Naphthalene	91-20-3	2,75	4	2,20	(a)
Endosulfan-diol	2157-19-9	3,02	5	2,29	(c)
Methiocarb	2032-65-7	3,10	4	2,39	(c)
Acid Yellow 219	63405-85-6	3,16	4	2,83	(a)
1,2,3-Trichlorobenzene	87-61-6	3,16	4	1,40	(a)
γ-HCH	58-89-9	3,23	5	2,94	(a)
Fenthion	55-38-9	3,31	3	2,49	(c)
Direct Red 81	2610-11-9	3,43	4	2,68	(a)
Pyrazophos	13457-18-6	3,65	3	2,70	(c)
α-Endosulfan	959-98-8	4,09	5	3,74	(c)
Diclofop-methyl	51338-27-3	4,20	3	3,77	(c)
Phenanthrene	85-01-8	4,09	4	3,83	(a)
Basic Blue 41 (mix)	26850-47-5	4,89	4	4,46	(a)
	12270-13-2				
DDT	50-29-3	5,63	1	—	(b)

(a) W. Kördel, J. Müller (1994). Bestimmung des Adsorptionskoeffizienten organischer Chemikalien mit der HPLC. UBA R & D Report No. 106 01 044 (1994).

(b) B.V. Oepen, W. Kördel, W. Klein., (1991) Chemosphere, 22, p. 285-304.

(c) Data provided by industry.

**▼B**C.20 ***DAPHNIA MAGNA* REPRODUCTION TEST**1. **METHOD**

This Reproduction toxicity test method is a replicate of the OECD TG 211 (1998).

## 1.1. INTRODUCTION

The primary objective of the test is to assess the effect of chemicals on the reproductive output of *Daphnia magna*.

## 1.2. DEFINITIONS AND UNITS

**Parent Animals:** are those female *Daphnia* present at the start of the test and of which the reproductive output is the object of the study.

**Offspring:** are the young *Daphnia* produced by the parent animals in the course of the test.

**Lowest Observed Effect Concentration (LOEC):** is the lowest tested concentration at which the substance is observed to have a statistically significant effect on reproduction and parent mortality (at  $p < 0,05$ ) when compared with the control, within a stated exposure period. However, all test concentrations above the LOEC must have a harmful effect equal to or greater than those observed at the LOEC. When these two conditions cannot be satisfied, a full explanation must be given for how the LOEC (and hence the NOEC) has been selected.

**No Observed Effect Concentration (NOEC):** is the test concentration immediately below the LOEC, which when compared with the control, has no statistically significant effect ( $p < 0,05$ ), within a stated exposure period.

**EC<sub>x</sub>:** is the concentration of the test substance dissolved in water that results in a x % reduction in reproduction of *Daphnia magna* within a stated exposure period.

**Intrinsic rate of increase:** is a measure of population growth which integrates reproductive output and age-specific mortality (20) (21) (22). In steady state populations it will be zero. For growing populations it will be positive and for shrinking populations it will be negative. Clearly, the latter is not sustainable and ultimately will lead to extinction.

**Limit of Detection:** is the lowest concentration that can be detected but not quantified.

**Limit of Determination:** is the lowest concentration that can be measured quantitatively.

**Mortality:** an animal is recorded as dead when it is immobile, i.e. when it is not able to swim, or if there is no observed movement of appendages or postabdomen, within 15 seconds after gentle agitation of the test container. (If another definition is used, this must be reported together with its reference).

**▼ B**

## 1.3. PRINCIPLE OF THE TEST METHOD

Young female *Daphnia* (the parent animals), aged less than 24 hours at the start of the test, are exposed to the test substance added to water at a range of concentrations. The test duration is 21 days. At the end of the test, the total number of living offspring produced per parent animal alive at the end of the test is assessed. This means that juveniles produced by adults that die during the test are excluded from the calculations. Reproductive output of parent animals can be expressed in other ways (e.g. number of living offspring produced per animal per day from the first day offspring were observed) but these should be reported in addition to the total number of juveniles produced per parent alive at the end of the test. The reproductive output of the animals exposed to the test substance is compared to that of the control(s) in order to determine the lowest observed effect concentration (LOEC) and hence the no observed effect concentration (NOEC). In addition, and as far as possible, the data are analysed using a regression model in order to estimate the concentration that would cause a x % reduction in reproductive output (i.e. the EC<sub>50</sub>, EC<sub>20</sub>, or EC<sub>10</sub>).

The survival of the parent animals and time to production of first brood must also be reported. Other substance-related effects on parameters such as growth (e.g. length) and possibly intrinsic rate of increase, may also be examined.

## 1.4. INFORMATION ON THE TEST SUBSTANCE

Results of an acute toxicity test (see Method C.2, Part I) performed with *Daphnia magna* should be available. The result may be useful in selecting an appropriate range of test concentrations in the reproduction tests. The water solubility and the vapour pressure of the test substance should be known and a reliable analytical method for the quantification of the substance in the test solutions with reported recovery efficiency and limit of determination should be available.

Information on the test substance which may be useful in establishing the test conditions includes the structural formula, purity of the substance, stability in light, stability under the conditions of the test, pK<sub>a</sub>, P<sub>ow</sub> and results of the test for ready biodegradability (see Method C.4).

## 1.5. VALIDITY OF THE TEST

For a test to be valid, the following performance criteria should be met in the control(s):

- the mortality of the parent animals (female *Daphnia*) does not exceed 20 % at the end of the test,
- the mean number of live offspring produced per parent animal surviving at the end of the test is  $\geq 60$ .

## 1.6. DESCRIPTION OF THE TEST METHOD

## 1.6.1. Apparatus

Test vessels and other apparatus which will come into contact with the test solutions should be made entirely of glass or other chemically inert material. The test vessels will normally be glass beakers.

In addition, some or all of the following equipment will be required:

- oxygen meter (with microelectrode or other suitable equipment for measuring dissolved oxygen in low volume samples),

**▼ B**

- adequate apparatus for temperature control,
- pH meter,
- equipment for the determination of the hardness of water,
- equipment for the determination of the total organic carbon concentration (TOC) of water or equipment for the determination of the chemical oxygen demand (COD),
- adequate apparatus for the control of the lighting regime and the measurement of light intensity.

**1.6.2. Test organism**

The species to be used in the test is *Daphnia magna* Straus. Other *Daphnia* species may be used providing they meet the validity criteria as appropriate (the validity criterion relating to the reproductive output in the controls should be relevant for the *Daphnia* species). If other species of *Daphnia* are used they must be clearly identified and their use justified.

Preferably, the clone should have been identified by genotyping. Research (1) has shown that the reproductive performance of Clone A (which originated from IRCHA in France) (3) consistently meets the validity criterion of a mean of  $\geq 60$  offspring per parent animal surviving when cultured under the conditions described in this method. However, other clones are acceptable provided that the *Daphnia* culture is shown to meet the validity criteria for a test.

At the start of the test, the animals should be less than 24 hours old and must not be first brood progeny. They should be derived from a healthy stock (i.e. showing no signs of stress such as high mortality, presence of males and ephippia, delay in the production of the first brood, discoloured animals etc.). The stock animals must be maintained in culture conditions (light, temperature, medium, feeding and animals per unit volume) similar to those to be used in the test. If the *Daphnia* culture medium to be used in the test is different from that used for routine *Daphnia* culture, it is good practice to include a pre-test acclimation period of normally about 3 weeks (i.e. one generation) to avoid stressing the parent animals.

**1.6.3. Test medium**

It is recommended that a fully defined medium be used in this test. This can avoid the use of additives (e.g. seaweed, soil extract etc.), which are difficult to characterise, and therefore improves the opportunities for standardisation between laboratories. Elendt M4 (4) and M7 media (see Appendix 1) have been found to be suitable for this purpose. However, other media (e.g. (5) (6)) are acceptable providing the performance of the *Daphnia* culture is shown to meet the validity criteria for the test.

If media are used which include undefined additives, these additives should be specified clearly and information should be provided in the test report on composition, particularly with regard to carbon content as this may contribute to the diet provided. It is recommended that the total organic carbon (TOC) and/or chemical oxygen demand (COD) of the stock preparation of the organic additive is determined and an estimate of the resulting contribution to the TOC/COD in the test medium made. It is recommended that TOC levels in the medium (i.e. before addition of the algae) be below 2 mg/l (7).

**▼ B**

When testing substances containing metals, it is important to recognise that the properties of the test medium (e.g. hardness, chelating capacity) may have a bearing on the toxicity of the test substance. For this reason, a fully defined medium is desirable. However, at present, the only fully defined media which are known to be suitable for long-term culture of *Daphnia magna* are Elendt M4 and M7. Both media contain the chelating agent EDTA. Work has shown (2) that the 'apparent toxicity' of cadmium is generally lower when the reproduction test is performed in M4 and M7 media than in media containing no EDTA. M4 and M7 are not, therefore, recommended for testing substances containing metals, and other media containing known chelating agents should also be avoided. For metal-containing substances it may be advisable to use an alternative medium such as, for example, ASTM reconstituted hard fresh water (7), which contains no EDTA, with added seaweed extract (8). This combination of ASTM reconstituted hard fresh water and seaweed extract is also suitable for long-term culture and testing of *Daphnia magna* (2), although it still exerts a mild chelating action due to the organic component in the added seaweed extract.

At the beginning and during the test, the dissolved oxygen concentration should be above 3 mg/l. The pH should be within the range 6-9, and normally it should not vary by more than 1,5 units in any one test. Hardness above 140 mg/l (as CaCO<sub>3</sub>) is recommended. Tests at this level and above have demonstrated reproductive performance in compliance with the validity criteria (9) (10).

**1.6.4. Test solutions**

Test solutions of the chosen concentrations are usually prepared by dilution of a stock solution. Stock solutions should preferably be prepared by dissolving the substance in test medium.

The use of organic solvents or dispersants may be required in some cases in order to produce a suitably concentrated stock solution, but every effort should be made to avoid the use of such materials. Examples of suitable solvents are acetone, ethanol, methanol, dimethylformamide and triethylene glycol. Examples of suitable dispersants are Cremophor RH40, methylcellulose 0,01 % and HCO-40. In any case, the test substance in the test solutions should not exceed the limit of solubility in the test medium.

Solvents are used to produce a stock solution which can be dosed accurately into water. At the recommended solvent concentration in the final test medium (i.e.  $\leq 0,1$  ml/l), the solvents listed above will not be toxic and will not increase the water solubility of a substance.

Dispersants may assist in accurate dosing and dispersion. At the recommended concentration in the final test medium ( $\leq 0,1$  ml/l), the dispersants listed above will not be toxic and will not increase the water solubility of a substance.



**▼B**

## 1.7. TEST DESIGN

Treatments should be allocated to the test vessels and all subsequent handling of the test vessels should be done in a random fashion. Failure to do this may result in bias that could be construed as being a concentration effect. In particular, if experimental units are handled in treatment or concentration order, then some time-related effect, such as operator fatigue or other error, could lead to greater effects at the higher concentrations. Furthermore, if the test results are likely to be affected by an initial or environmental condition of the test, such as position in the laboratory, then consideration should be given to blocking the test.

## 1.8. PROCEDURE

1.8.1. **Conditions of exposure**1.8.1.1. *Duration*

The test duration is 21 days.

1.8.1.2. *Loading*

Parent animals are maintained individually, one per test vessel, with 50-100 ml of medium in each vessel.

Larger volumes may sometimes be necessary to meet requirements of the analytical procedure used for determination of the test substance concentration, although pooling of replicates for chemical analysis is also allowable. If volumes greater than 100 ml are used, the ration given to the *Daphnia* may need to be increased to ensure adequate food availability and compliance with the validity criteria. For flow-through tests, alternative designs may, for technical reasons, be considered (e.g. four groups of 10 animals in a larger test volume), but any changes to the test design should be reported.

1.8.1.3. *Number of animals*

For semi-static tests, at least 10 animals individually held at each test concentration and at least 10 animals individually held in the control series.

For flow-through tests, 40 animals divided into four groups of 10 animals at each test concentration has been shown to be suitable (1). A smaller number of test organisms may be used and a minimum of 20 animals per concentration divided into two or more replicates with an equal number of animals (e.g. four replicates each with five daphnids) is recommended. Note that for tests where animals are held in groups, it will not be possible to express the reproductive output as the total number of living offspring produced per parent animal alive at the end of the test, if parent animals die. In these cases reproductive output should be expressed as 'total number of living offspring produced per parent present at the beginning of the test'.

1.8.1.4. *Feeding*

For semi-static tests, feeding should preferably be done daily, but at least three times per week (i.e. corresponding to media changes). Deviations from this (e.g. for flow-through tests) should be reported.

**▼ B**

During the test the diet of the parent animals should preferably be living algal cells of one or more of the following: *Chlorella* sp, *Selenastrum capricornutum* (now *Pseudokirchneriella subcapitata* (11)) and *Scenedesmus subspicatus*. The supplied diet should be based on the amount of organic carbon (C) provided to each parent animal. Research (12) has shown that, for *Daphnia magna*, ration levels of between 0,1 and 0,2 mg C/*Daphnia*/day are sufficient for achieving the required number of offspring to meet the test validity criteria. The ration can be supplied either at a consistent rate throughout the period of the test, or, if desired, a lower rate can be used at the beginning and then increased during the test to take account of growth of the parent animals. In this case, the ration should still remain within the recommended range of 0,1-0,2 mg C/*Daphnia*/day at all times.

If surrogate measures, such as algal cell number or light absorbance, are to be used to feed the required ration level (i.e. for convenience since measurement of carbon content is time consuming), each laboratory must produce its own nomograph relating the surrogate measure to carbon content of the algal culture (see Appendix 2 for advice on nomograph production). Nomographs should be checked at least annually and more frequently if algal culture conditions have changed. Light absorbance has been found to be a better surrogate for carbon content than cell number (13).

A concentrated algal suspension should be fed to the *Daphnia* to minimise the volume of algal culture medium transferred to the test vessels. Concentration of the algae can be achieved by centrifugation followed by resuspension in distilled water, deionised water or *Daphnia* culture medium.

**1.8.1.5. Light**

16 hours light at an intensity not exceeding 15-20  $\mu\text{E} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ .

**1.8.1.6. Temperature**

The temperature of the test media should be within the range 18-22 °C. However, for any one test, the temperature should not, if possible, vary by more than 2 °C within these limits (e.g. 18-20, 19-21 or 20-22 °C). It may be appropriate to use an additional test vessel for the purposes of temperature monitoring.

**1.8.1.7. Aeration**

The test vessels must not be aerated during the test.

**1.8.2. Test concentration**

Normally, there should be at least five test concentrations arranged in a geometric series with a separation factor preferably not exceeding 3.2, and the appropriate number of replicates for each test concentration should be used (see Section 1.8.1.3). Justification should be provided if fewer than five concentrations are used. Substances should not be tested above their solubility limit in the test medium.

**▼B**

In setting the range of concentrations, the following should be borne in mind:

If the aim is to obtain the LOEC/NOEC, the lowest test concentration must be low enough so that the fecundity at that concentration is not significantly lower than that in the control. If this is not the case, the test will have to be repeated with a reduced lowest concentration.

If the aim is to obtain the LOEC/NOEC, the highest test concentration must be high enough so that the fecundity at that concentration is significantly lower than that in the control. If this is not the case, the test will have to be repeated with an increased highest concentration.

If the  $EC_X$  for effects on reproduction is estimated, it is advisable that sufficient concentrations are used to define the  $EC_X$  with an appropriate level of confidence. If the  $EC_{50}$  for effects on reproduction is estimated, it is advisable that the highest test concentration is greater than this  $EC_{50}$ . Otherwise, although it will still be possible to estimate the  $EC_{50}$ , the confidence interval for the  $EC_{50}$  will be very wide and it may not be possible to satisfactorily assess the adequacy of the fitted model.

The range of test concentration should preferably not include any concentrations that have a statistically significant effect on adult survival since this would change the nature of the test from simply a reproduction test to a combined reproduction and mortality test requiring much more complex statistical analysis.

Prior knowledge of the toxicity of the test substance (e.g. from an acute test and/or from range-finding studies) should help in selecting appropriate test concentrations.

Where a solvent or dispersant is used to aid preparation of test solutions (see Section 1.6.4), its final concentration in the test vessels should not be greater than 0,1 ml/l and should be the same in all test vessels.

**1.8.3. Controls**

One test-medium control series and also, if relevant, one control series containing the solvent or dispersant should be run in addition to the test series. When used, the solvent or dispersant concentration should be the same as that used in the vessels containing the test substance. The appropriate number of replicates should be used (see Section 1.8.1.3).

Generally, in a well-run test, the coefficient of variation around the mean number of living offspring produced per parent animal in the control(s) should be  $\leq 25\%$ , and this should be reported for test designs using individually held animals.

**1.8.4. Test medium renewal**

The frequency of medium renewal will depend on the stability of the test substance, but should be at least three times per week. If, from preliminary stability tests (see Section 1.4) the test substance concentration is not stable (i.e. outside the range 80-120 % of nominal or falling below 80 % of the measured initial concentration) over the maximum renewal period (i.e. 3 days), consideration should be given to more frequent medium renewal, or to the use of a flow-through test.

**▼B**

When the medium is renewed in semi-static tests, a second series of test vessels are prepared and the parent animals transferred to them by, for example, a glass pipette of suitable diameter. The volume of medium transferred with the *Daphnia* should be minimised.

**1.8.5. Observations**

The results of the observations made during the test should be recorded on data sheets (see examples in Appendices 3 and 4). If other measurements are required (see 1.3 and 1.8.8) additional observations may be required.

**1.8.6. Offspring**

The offspring produced by each parent animal should preferably be removed and counted daily from the appearance of the first brood, to prevent them consuming food intended for the adult. For the purpose of this method it is only the number of living offspring that needs to be counted, but the presence of aborted eggs or dead offspring should be recorded.

**1.8.7. Mortality**

Mortality among the parent animals should be recorded preferably daily, at least at the same times as offspring are counted.

**1.8.8. Other parameters**

Although this method is designed principally to assess effects on reproduction, it is possible that other effects may also be sufficiently quantified to allow statistical analysis. Growth measurements are highly desirable since they provide information on possible sublethal effects, which may be more useful than reproduction measurement alone; the measurement of the length of the parent animals (i.e. body length excluding the anal spine) at the end of the test is recommended. Other parameters that can be measured or calculated include time to production of first brood (and subsequent broods), number and size of broods per animal, number of aborted broods, presence of males or ephippia and the intrinsic rate of population increase.

**1.8.9. Frequency of analytical determinations and measurements**

Oxygen concentration, temperature, hardness and pH values should be measured at least once a week, in fresh and old media, in the control(s) and in the highest test substance concentration.

During the test, the concentrations of test substance are determined at regular intervals.

In semi-static tests where the concentration of the test substance is expected to remain within  $\pm 20$  % of the nominal (i.e. within the range 80-120 % — see 1.4 and 1.8.4), it is recommended that, as a minimum, the highest and lowest test concentrations be analysed when freshly prepared and at the time of renewal on one occasion during the first week of the test (i.e. analyses should be made on a sample from the same solution — when freshly prepared and at renewal). These determinations should be repeated at least at weekly intervals thereafter.

**▼B**

For tests where the concentration of the test substance is not expected to remain within  $\pm 20\%$  of the nominal, it is necessary to analyse all test concentrations, when freshly prepared and at renewal. However, for those tests where the measured initial concentration of the test substance is not within  $\pm 20\%$  of nominal but where sufficient evidence can be provided to show that the initial concentrations are repeatable and stable (i.e. within the range 80-120 % of initial concentrations), chemical determinations could be reduced in weeks 2 and 3 of the test to the highest and lowest test concentrations. In all cases, determination of test substance concentrations prior to renewal need only be performed on one replicate vessel at each test concentration.

If a flow-through test is used, a similar sampling regime to that described for semi-static tests is appropriate (but measurement of 'old' solutions is not applicable in this case). However, it may be advisable to increase the number of sampling occasions during the first week (e.g. three sets of measurements) to ensure that the test concentrations are remaining stable. In these types of test, the flow-rate of diluent and test substance should be checked daily.

If there is evidence that the concentration of the substance being tested has been satisfactorily maintained within  $\pm 20\%$  of the nominal or measured initial concentration throughout the test, then results can be based on nominal or measured initial values. If the deviation from the nominal or measured initial concentration is greater than  $\pm 20\%$ , results should be expressed in terms of the time-weighted mean (see Appendix 5).

## 2. DATA AND REPORTING

### 2.1. TREATMENT OF RESULTS

The purpose of this test is to determine the effect of the test substance on the total number of living offspring produced per parent animal alive at the end of the test. The total number of offspring per parent animal should be calculated for each test vessel (i.e. replicate). If, in any replicate the parent animal dies during the test or turns out to be male, then the replicate is excluded from the analysis. The analysis will then be based on a reduced number of replicates.

For the estimation of the LOEC, and hence the NOEC, for effects of the chemical on reproductive output, it is necessary to calculate the mean reproductive output across replicates for each concentration and the pooled residual standard deviation, and this can be done using analysis of variance (ANOVA). The mean for each concentration must then be compared with the control mean using an appropriate multiple comparison method. Dunnett's or Williams' tests may be useful (14)(15)(16)(17). It is necessary to check whether the ANOVA assumption of homogeneity of variance holds. It is recommended that this be done graphically rather than via a formal significance test (18); a suitable alternative is to run a Bartlett's test. If this assumption does not hold, then consideration should be given to transforming the data to homogenise variances prior to performing the ANOVA, or to carrying out a weighted ANOVA. The size of the effect detectable using ANOVA (i.e. the least significant difference) should be calculated and reported.

**▼ B**

For the estimation of the concentration which would cause a 50 % reduction in reproductive output (i.e. the EC<sub>50</sub>), a suitable curve, such as the logistic curve, should be fitted to the data using a statistical method such as least squares. The curve could be parameterised so that the EC<sub>50</sub> and its standard error can be estimated directly. This would greatly ease the calculation of the confidence limits about the EC<sub>50</sub>. Unless there are good reasons to prefer different confidence levels, two-sided 95 % confidence limits should be quoted. The fitting procedure should preferably provide a means for assessing the significance of the lack of fit. This can be done graphically or by dividing the residual sum of squares into 'lack of fit' and 'pure error components' and performing a significance test for lack of fit. Since treatments giving high fecundity are likely to have greater variance in the number of juveniles produced than treatments giving low fecundity, consideration to weighting the observed values to reflect the different variances in the different treatment groups should be given (see for background information ref. 18).

In the analysis of the data from the final ring test (2), a logistic curve was fitted using the following model, although other suitable models can be used:

$$Y = \frac{c}{1 + \left(\frac{x}{x_0}\right)^b}$$

where:

Y = the total number of juveniles per parent animal alive at the end of the test (calculated for each vessel)

x = the substance concentration

c = the expected number of juveniles when x = 0

x<sub>0</sub> = the EC<sub>50</sub> in the population

b = the slope parameter

This model is likely to be adequate in a large number of situations, but there will be tests for which it is not appropriate. A check should be made on the validity of the model as suggested above. In some cases, a hormesis model in which low concentrations give enhanced effects may be appropriate (19).

Other Effect Concentrations, such as the EC<sub>10</sub> or EC<sub>20</sub> can also be estimated, although it may be preferable to use a different parameterisation of the model from that used to estimate the EC<sub>50</sub>.

## 2.2. TEST REPORT

The test report must include the following:

### 2.2.1. Test substance:

— physical nature and relevant physicochemical properties,

**▼B**

- chemical identification data, including purity.

**2.2.2. Test species:**

- the clone (whether it has been genetically typed), supplier or source (if known) and the culture conditions used. If a different species to *Daphnia magna* is used, this should be reported and justified.

**2.2.3. Test conditions:**

- test procedure used (e.g. semi-static or flow-through, volume, loading in number of *Daphnia* per litre),
- photoperiod and light intensity,
- test design (e.g. number of replicates, number of parents per replicate),
- details of culture medium used,
- if used, additions of organic material including the composition, source, method of preparation, TOC/COD of stock preparations, estimation of resulting TOC/COD in test medium,
- detailed information on feeding, including amount (in mg C/*Daphnia*/day) and schedule (e.g. type of food(s), including for algae the specific name(species) and, if known, the strain, the culture conditions),
- method of preparation of stock solutions and frequency of renewal (the solvent or dispersant and its concentration must be given, when used).

**2.2.4. Results:**

- results from any preliminary studies on the stability of the test substance,
- the nominal test concentrations and the results of all analyses to determine the concentration of the test substance in the test vessels (see example data sheets in Appendix 4); the recovery efficiency of the method and the limit of determination should also be reported,
- water quality within the test vessels (i.e. pH, temperature and dissolved oxygen concentration, and TOC and/or COD and hardness where applicable) (see example data sheet in Appendix 3),
- the full record of living offspring by each parent animal (see example data sheet in Appendix 3),
- the number of deaths among the parent animals and the day on which they occurred (see example data sheet in Appendix 3),
- the coefficient of variation for control fecundity (based on total number of living offspring per parent animal alive at the end of the test),
- plot of total number of living offspring per parent animal (for each replicate) alive at the end of the test versus concentration of the test substance,
- the Lowest Observed Effect Concentration (LOEC) for reproduction, including a description of the statistical procedures used and an indication of what size of effect could be detected and the No Observed Effect Concentration (NOEC) for reproduction; where appropriate, the LOEC/NOEC for mortality of the parent animals should also be reported,

**▼B**

- where appropriate, the EC<sub>x</sub> for reproduction and confidence intervals and a graph of the fitted model used for its calculation, the slope of the dose-response curve and its standard error,
- other observed biological effects or measurements: report any other biological effects which were observed or measured (e.g. growth of parent animals) including any appropriate justification,
- an explanation for any deviation from the Test Method.

3. **REFERENCES**

- (1) OECD Test Guideline Programme, Report of the Workshop on the *Daphnia magna* Pilot Ring Test, Sheffield University, UK, 20-21 March 1993.
- (2) OECD Environmental Health and Safety Publications. Series on Testing and Assessment No6. Report of the Final Ring Test of the *Daphnia magna* Reproduction Test Paris. 1997.
- (3) Baird D.J., Barber J., Bradley M.C., Soares A.M.V.M. and Calow P. (1991) A comparative study of genotype sensitivity to acute toxic stress using clones of *Daphnia magna* Straus. *Ecotoxicology and Environmental Safety*, 21, p. 257-265.
- (4) Elenđt B.P., (1990) Selenium deficiency in Crustacea; An ultrastructural approach to antennal damage in *Daphnia magna* Straus. *Protoplasma*, 154, p. 25-33.
- (5) EPA (1993) Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms. (Fourth ed.). EPA/600/4-90/027F. C. I. Weber (ed), USEPA, Cincinnati, Ohio.
- (6) Vigano L., (1991) Suitability of commercially available spring waters as standard medium for culturing *Daphnia magna*. *Bull. Environ. Contam. Toxicol.*, 47, p. 775-782.
- (7) ASTM (1988) Standard Guide for Conducting Acute Toxicity Tests with Fishes, Macroinvertebrates and Amphibians. E729-88a. American Society for Testing and Materials, Philadelphia P.A. p. 20.
- (8) Baird D.J., Soares A.M.V.M., Gírling A., Barber J., Bradley M.C. and Calow P., (1989) The long term maintenance of *Daphnia magna* Straus for use in ecotoxicological tests; problems and prospects. In: Proceedings of the 1st European Conference on Ecotoxicology. Copenhagen 1988 (H.Løkke, H. Tyle & F. Bro-Rasmussen. Eds.) p. 144-148.
- (9) Parkhurst B.R., Forte J.L. and Wright G.P., (1981) Reproducibility of a life-cycle toxicity test with *Daphnia magna*. *Bull. Environ. Contam. and Toxicol.*, 26, p. 1-8.
- (10) Cowgill U.M. and Milazzo D.P., (1990) The sensitivity of two cladocerans to water quality variables: salinity and hardness. *Arch. Hydrobiol.*, 120(2), p. 185-196.
- (11) Korshikov (1990) *Pseudokirchneriella subcapitata* Hindak, F-1990. *Biologice Prace*, 36, p. 209.
- (12) Sims I.R., Watson S. and Holmes D., (1993) Toward a standard *Daphnia* juvenile production test. *Environmental Toxicology and Chemistry*, 12, p. 2053-2058.



**▼ B**

- (13) Sims I., (1993) Measuring the growth of phytoplankton: the relationship between total organic carbon with three commonly used parameters of algal growth. *Arch. Hydrobiol.*, 128, 459-466.
- (14) Dunnett C.W., (1955) A multiple comparisons procedure for comparing several treatments with a control. *J. Amer. Statist. Assoc.*, 50, p. 1096-1121.
- (15) Dunnett C.W., (1964) New tables for multiple comparisons with a control. *Biometrics*, 20, p. 482-491.
- (16) Williams D.A., (1971) A test for differences between treatment means when several dose levels are compared with a zero dose control. *Biometrics* 27, p. 103-117.
- (17) Williams D.A., (1972) The comparison of several dose levels with a zero dose control. *Biometrics*, 28, p. 510-531.
- (18) Draper N.R. and Smith H., (1981) *Applied Regression Analysis*, second edition, Wiley, N.Y.
- (19) Brain P. and Cousens R., (1989) An equation to describe dose responses where there is stimulation of growth at low doses. *Weed Research*, 29, p. 93-96.
- (20) Wilson E.O. and Bossert, W.H., (1971) *A Primer of Population Biology*. Sinauer Associates Inc. Publishers.
- (21) Poole R.W., (1974) *An Introduction to quantitative Ecology*. Mc Graw Hill Series in Population Biology, New York, p. 532.
- (22) Meyer J.S., Ingersoll C.G., McDonald L.L. and Boyce M.S., (1986) Estimating uncertainty in population growth rates: Jackknife vs bootstrap techniques. *Ecology*, 67, p. 1156-1166.

## ▼B

## Appendix 1

## PREPARATION OF FULLY DEFINED ELENDT M7 AND M4 MEDIA

## Acclimation to Elendt M7 and M4 media

Some laboratories have experienced difficulty in directly transferring *Daphnia* to M4 (1) and M7 media. However, some success has been achieved with gradual acclimation, i.e. moving from own medium to 30 % Elendt, then to 60 % Elendt and then to 100 % Elendt. The acclimation periods may need to be as long as one month.

## PREPARATION

## Trace elements

Separate stock solutions (I) of individual trace elements are first prepared in water of suitable purity, e.g. deionised, distilled or reverse osmosis. From these different stock solutions (I) a second single stock solution (II) is prepared, which contains all trace elements (combined solution), i.e.:

Stock solutions I (single substance)	Amount added to water mg/l	Concentration (in relation to medium M4) fold	To prepare the combined stock-solution II add the following amount of stock solution I to water ml/l	
			M 4	M 7
H <sub>3</sub> BO <sub>3</sub>	57 190	20 000	1,0	0,25
MnCl <sub>2</sub> * 4 H <sub>2</sub> O	7 210	20 000	1,0	0,25
LiCl	6 120	20 000	1,0	0,25
RbCl	1 420	20 000	1,0	0,25
SrCl <sub>2</sub> * 6 H <sub>2</sub> O	3 040	20 000	1,0	0,25
NaBr	320	20 000	1,0	0,25
Na <sub>2</sub> MoO <sub>4</sub> * 2 H <sub>2</sub> O	1 260	20 000	1,0	0,25
CuCl <sub>2</sub> * 2 H <sub>2</sub> O	335	20 000	1,0	0,25
ZnCl <sub>2</sub>	260	20 000	1,0	1,0
CoCl <sub>2</sub> * 6 H <sub>2</sub> O	200	20 000	1,0	1,0
KI	65	20 000	1,0	1,0
Na <sub>2</sub> SeO <sub>3</sub>	43,8	20 000	1,0	1,0
NH <sub>4</sub> VO <sub>3</sub>	11,5	20 000	1,0	1,0
Na <sub>2</sub> EDTA * 2 H <sub>2</sub> O	5 000	2 000	—	—
FeSO <sub>4</sub> * 7 H <sub>2</sub> O	1 991	2 000	—	—

Both Na<sub>2</sub>EDTA and FeSO<sub>4</sub> solutions are prepared singly, poured together and autoclaved immediately. This gives:

21 Fe-EDTA solution		1 000-fold	20,0	5,0
---------------------	--	------------	------	-----

**▼B****M4 and M7 media**

M4 and M7 media are prepared using stock solution II, the macro-nutrients and vitamins as follows:

	Amount added to water mg/l	Concentration (related to medium M4) fold	Amount of stock solution added to prepare medium ml/l	
			M 4	M 7
Stock solution II combined trace elements		20	50	50

## Macro-nutrient stock solutions (single substance)

CaCl <sub>2</sub> * 2 H <sub>2</sub> O	293 800	1 000	1,0	1,0
MgSO <sub>4</sub> * 7 H <sub>2</sub> O	246 600	2 000	0,5	0,5
KCl	58 000	10 000	0,1	0,1
NaHCO <sub>3</sub>	64 800	1 000	1,0	1,0
Na <sub>2</sub> SiO <sub>3</sub> * 9 H <sub>2</sub> O	50 000	5 000	0,2	0,2
NaNO <sub>3</sub>	2 740	10 000	0,1	0,1
KH <sub>2</sub> PO <sub>4</sub>	1 430	10 000	0,1	0,1
K <sub>2</sub> HPO <sub>4</sub>	1 840	10 000	0,1	0,1
Combined Vitamin stock	—	10 000	0,1	0,1

The combined vitamin stock solution is prepared by adding the 3 vitamins to 1 litre water as show below:

Thiamine hydrochloride	750	10 000	—	—
Cyanocobalamine (B <sub>12</sub> )	10	10 000	—	—
Biotine	7,5	10 000	—	—

The combined vitamin stock is stored frozen in small aliquots. Vitamins are added to the media shortly before use.

*N.B.* To avoid precipitation of salts when preparing the complete media, add the aliquots of stock solutions to about 500-800 ml deionised water and then fill up to 1 litre.

*N.N.B.* The first publication of the M4 medium can be found in Elendt, B.P. (1990). Selenium deficiency in crustacea; an ultrastructural approach to antennal damage in *Daphnia magna* Straus. *Protoplasma*, 154, p. 25-33.



### Appendix 2

#### TOTAL ORGANIC CARBON (TOC) ANALYSIS AND PRODUCTION OF A NOMOGRAPH FOR TOC CONTENT OF ALGAL FEED

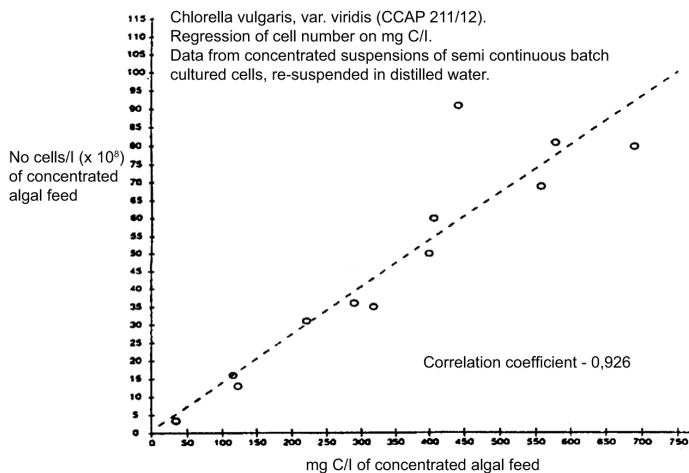
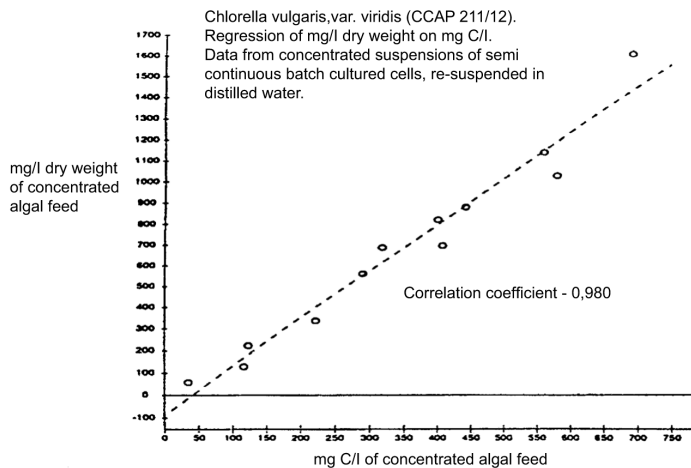
It is recognised that the carbon content of the algal feed will not normally be measured directly but from correlations (i.e. nomographs) with surrogate measures such as algal cell number or light absorbance).

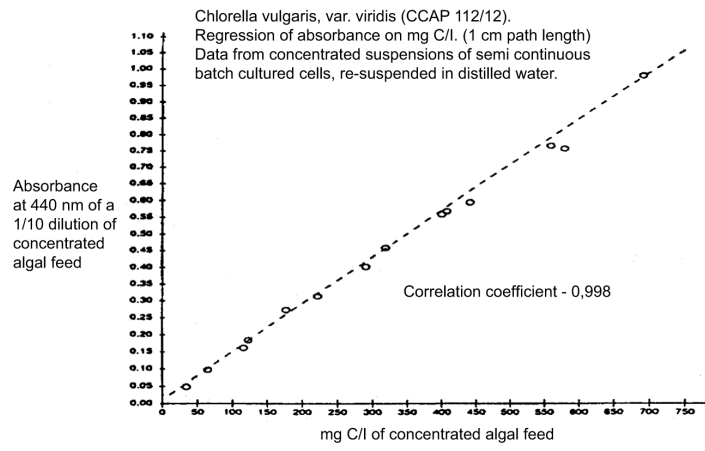
TOC should be measured by high temperature oxidation rather than by UV or persulphate methods. (See: The Instrumental Determination of Total Organic Carbon, Total Oxygen Demand and Related Determinands 1979, HMSO 1980; 49 High Holborn, London WC1V 6HB).

For nomograph production, algae should be separated from the growth medium by centrifugation followed by resuspension in distilled water. Measure the surrogate parameter and TOC concentration in each sample in triplicate. Distilled water blanks should be analysed and the TOC concentration deducted from that of the algal sample TOC concentration.

Nomograph should be linear over the required range of carbon concentrations. Examples are shown below.

*N.B.*: These should not be used for conversions; it is essential that laboratories prepare their own nomographs.



**▼ B**

▼B

Appendix 3

**EXAMPLE DATA SHEET RECORDING MEDIUM RENEWAL, PHYSICAL/CHEMICAL MONITORING DATA, FEEDING, *DAPHNIA* REPRODUCTION AND ADULT MORTALITY**

Experiment No:	Data started:			Clone:			Medium:			Type of food:			Test Substance:			Nominal conc:								
Day	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21		
Medium renewal (tick)																								
PH (°)																								new
																								old
O <sub>2</sub> mg/l (°)																								new
																								old
Temp ( °C) (°)																								new
																								old
Food provided (tick)																								
No live offspring (°)																								Total
Vessel 1																								
2																								
3																								
4																								
5																								

▼**B**

Experiment No:	Data started:			Clone:			Medium:			Type of food:			Test Substance:			Nominal conc:								
Day	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21		
6																								
7																								
8																								
9																								
10																								
																								Total
Cumulative adult mortality <sup>(3)</sup>																								

- <sup>(1)</sup> Indicate which vessel was used for the experiment.
- <sup>(2)</sup> Record mortality of any adult animals as 'M' in relevant box.
- <sup>(3)</sup> Record aborted broods as 'AB' in relevant box.

**▼B***Appendix 4***EXAMPLE DATA SHEET FOR RECORDING RESULTS OF CHEMICAL ANALYSIS****(a) Measured concentrations**

Nominal conc.	Week 1 sample		Week 2 sample		Week 3 sample	
	Fresh	Old	Fresh	Old	Fresh	Old

**(b) Measured concentrations as a percentage of nominal**

Nominal conc.	Week 1 sample		Week 2 sample		Week 3 sample	
	Fresh	Old	Fresh	Old	Fresh	Old



▼ **B***Appendix 5***CALCULATION OF A TIME-WEIGHTED MEAN****Time-weighted mean**

Given that the concentration of the test substance can decline over the period between medium renewals, it is necessary to consider what concentration should be chosen as representative of the range of concentrations experienced by the parent *Daphnia*. The selection should be based on biological considerations as well as statistical ones. For example, if reproduction is thought to be affected mostly by the peak concentration experienced, then the maximum concentration should be used. However, if the accumulated or longer term effect of the toxic substance is considered to be more important, then an average concentration is more relevant. In this case, an appropriate average to use is the time-weighted mean concentration, since this takes account of the variation in instantaneous concentration over time.

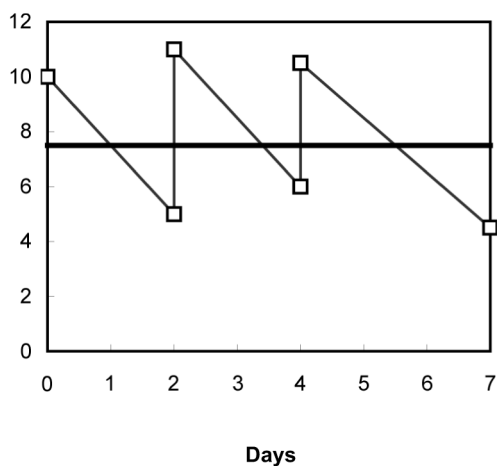
**Figure 1: Example of time-weighted mean**

Figure 1 shows an example of a (simplified) test lasting seven days with medium renewal at Days 0, 2 and 4.

- The thin zig-zag line represents the concentration at any point in time. The fall in concentration is assumed to follow an exponential decay process.
- The six plotted points represent the observed concentrations measured at the start and end of each renewal period.
- The thick solid line indicates the position of the time-weighted mean.

The time-weighted mean is calculated so that the area under the time-weighted mean is equal to the area under the concentration curve. The calculation for the above example is illustrated in Table 1.

**▼B**

Table 1:  
calculation of Time-weighted mean

Renewal No	Days	Conc0	Conc1	Ln(Conc0)	Ln(Conc1)	Area
1	2	10,000	4,493	2,303	1,503	13,767
2	2	11,000	6,037	2,398	1,798	16,544
3	3	10,000	4,066	2,303	1,403	19,781
Total Days: 7					Total Area	50,091
					TW Mean	7,156

*Days* is the number of days in the renewal period

*Conc0* is the measured concentration at the start of each renewal period

*Conc1* is the measured concentration at the end of each renewal period

*Ln(Conc0)* is the natural logarithm of *Conc0*

*Ln(Conc1)* is the natural logarithm of *Conc1*

*Area* is the area under the exponential curve for each renewal period. It is calculated by:

$$Area = \frac{Conc0 - Conc1}{Ln(Conc0) - Ln(Conc1)} \times Days$$

The time-weighted mean (*TW Mean*) is the *total area* divided by the *total days*.

Of course, for the *Daphnia* reproduction test the table would have to be extended to cover 21 days.

It is clear that when observation are taken only at the start and end of each renewal period, it is not possible to confirm that the decay process is, in fact, exponential. A different curve would result in a different calculation for *Area*. However, an exponential decay process is not implausible and is probably the best curve to use in the absence of other information.

However, a work of caution is required if the chemical analysis fails to find any substance at the end of the renewal period. Unless it is possible to estimate how quickly the substance disappeared from the solution, it is impossible to obtain a realistic area under the curve, and hence it is impossible to obtain a reasonable time-weighted mean.

**▼B****C.21. SOIL MICROORGANISMS: NITROGEN TRANSFORMATION TEST****1. METHOD**

This test method is a replicate of OECD TG 216 (2000).

**1.1. INTRODUCTION**

This testing method describes a laboratory method designed to investigate the long-term effects of chemicals, after a single exposure, on nitrogen transformation activity of soil microorganisms. The test is principally based on the recommendations of the European and Mediterranean Plant Protection Organization (1). However, other guideline, including those of the German Biologische Bundesanstalt (2), the US Environmental Protection Agency (3) SETAC (4) and the International Organization for Standardization (5), were also taken into account. An OECD workshop on soil/sediment selection held at Belgirate, Italy, in 1995 (6) agreed on the number and type of soils for use in this test. Recommendations for collection, handling and storage of soil sample are based on an ISO Guidance Document (7) and recommendations from the Belgirate workshop. In the assessment and evaluation of toxic characteristics of test substances, determination of effects on soil microbial activity may be required, e.g. when data on the potential side effects of crop protection products on soil microflora are required or when exposure of soil microorganisms to chemicals other than crop protection products is expected. The nitrogen transformation test is carried out to determine the effects of such chemicals on soil microflora. If agrochemicals (e.g. crop protection products, fertilisers, forestry chemicals) are tested, both nitrogen transformation and carbon transformation tests are conducted. If non-agrochemicals are tested, the nitrogen transformation test is sufficient. However, if EC<sub>50</sub> values of the nitrogen transformation test for such chemicals fall within the range found for commercially available nitrification inhibitors (e.g. nitrapyrin), a carbon transformation test can be conducted to gain further information.

Soils consist of living and non-living components which exist in complex and heterogeneous mixtures. Microorganisms play an important role in break-down and transformation of organic matter in fertile soils with many species contributing to different aspects of soil fertility. Any long-term interference with these biochemical processes could potentially interfere with nutrient cycling and this could alter soil fertility. Transformation of carbon and nitrogen occurs in all fertile soils. Although the microbial communities responsible for these processes differ from soil to soil, the pathways of transformation are essentially the same.

This testing method described is designed to detect long-term adverse effects of a substance on the process of nitrogen transformation in aerobic surface soils. The test method also allows estimation of the effects of substances on carbon transformation by the soil microflora. Nitrate formation takes place subsequent to the degradation of carbon-nitrogen bonds. Therefore, if equal rates of nitrate production are found in treated and control soils, it is highly probable that the major carbon degradation pathways are intact and functional. The substrate chosen for the test (powdered lucerne meal) has a favourable carbon to nitrogen ratio (usually between 12/1 and 16/1). Because of this, carbon starvation is reduced during the test and if microbial communities are damaged by a chemical, they might recover within 100 days.

**▼ B**

The tests from which this testing method was developed were primarily designed for substances for which the amount reaching the soil can be anticipated. This is the case, for example, for crop protection products for which the application rate in the field is known. For agrochemicals, testing of two doses relevant to the anticipated or predicted application rate is sufficient. Agrochemicals can be tested as active ingredients (a.i.) or as formulated products. However, the test is not limited to agrochemicals. By changing both the amounts of test substance applied to the soil, and the way in which the data are evaluated, the test can also be used for chemicals for which the amount expected to reach the soil is not known. Thus, with chemicals other than agrochemicals, the effects of a series of concentrations on nitrogen transformation are determined. The data from these tests are used to prepare a dose-response curve and calculate  $EC_x$  values, where  $x$  is defined % effect.

## 1.2. DEFINITIONS

**Nitrogen transformation:** is the ultimate degradation by microorganisms of nitrogen-containing organic matter, via the process of ammonification and nitrification, to the respective inorganic end-product nitrate.

**$EC_x$  (effective concentration):** is the concentration of the test substance in soil that results in a  $x$  percent inhibition of nitrogen transformation to nitrate.

**$EC_{50}$  (median effective concentration):** is the concentration of the test substance in soil that results in a 50 percent (50 %) inhibition of nitrogen transformation to nitrate.

## 1.3. REFERENCE SUBSTANCES

None.

## 1.4. PRINCIPLE OF THE TEST METHOD

Sieved soil is amended with powdered plant meal and either treated with the test substance or left untreated (control). If agrochemicals are tested, a minimum of two test concentrations are recommended and these should be chosen in relation to the highest concentration anticipated in the field. After 0, 7, 14 days and 28 days of incubation, samples of treated and control soils are extracted with an appropriate solvent, and the quantities of nitrate in the extracts are determined. The rate of nitrate formation in treated samples is compared with the rate in the controls, and the percent deviation of the treated from the control is calculated. All tests run for at least 28 days. If, on the 28th day, differences between treated and untreated soils are equal to or greater than 25 %, measurements are continued to a maximum of 100 days. If non-agrochemicals are tested, a series of concentrations of the test substance are added to samples of the soil, and the quantities of nitrate formed in treated and control samples are measured after 28 days of incubation. Results from tests with multiple concentrations are analysed using a regression model, and the  $EC_x$  values are calculated (i.e.  $EC_{50}$ ,  $EC_{25}$  and/or  $EC_{10}$ ). See definitions.

## 1.5. VALIDITY OF THE TEST

Evaluations of test results with agrochemicals are based on relatively small differences (i.e. average value  $\pm 25$  %) between nitrate concentrations in control and treated soil samples, so large variations in the controls can lead to false results. Therefore, the variation between replicate control samples should be less than  $\pm 15$  %.

**▼ B**

## 1.6. DESCRIPTION OF THE TEST METHOD

1.6.1. **Apparatus**

Test containers made of chemically inert material are used. They should be of a suitable capacity in compliance with the procedure used for incubation of soils, i.e. incubation in bulk or as a series of individual soil samples (see Section 1.7.1.2). Care should be taken both to minimise water loss and to allow gas exchange during the test (e.g. the test containers may be covered with perforated polyethylene foil). When volatile substances are tested, sealable and gas-tight containers should be used. These should be of a size such that approximately one quarter of their volume is filled with the soil sample.

Standard laboratory equipment including the following is used:

- agitation device: mechanical shaker or equivalent equipment;
- centrifuge (3 000 g) or filtration device (using nitrate-free filter paper);
- instrument of adequate sensitivity and reproducibility for nitrate analysis.

1.6.2. **Selection and number of soils**

One single soil is used. The recommended soil characteristics are as follows:

- sand content: not less than 50 % and not greater than 75 %,
- pH: 5,5-7,5,
- organic carbon content: 0,5-1,5 %,
- the microbial biomass should be measured (8)(9) and its carbon content should be at least 1 % of the total soil organic carbon.

In most cases, a soil with these characteristics represents a worst case situation, since adsorption of the test chemical is minimum and its availability to the microflora is maximum. Consequently, tests with other soils are generally unnecessary. However, in certain circumstances, e.g. where the anticipated major use of the test substance is in particular soils such as acidic forest soils, or for electrostatically charged chemicals, it may be necessary to use an additional soil.

**▼B****1.6.3. Collection and storage of soil samples****1.6.3.1. Collection**

Detailed information on the history of the field site from where the test soil is collected should be available. Details include exact location, vegetation cover, dates of treatments with crop protection products, treatments with organic and inorganic fertilisers, additions of biological materials or accidental contaminations. The site chosen for soil collection should be one which allows long-term use. Permanent pastures, fields with annual cereal crops (except maize) or densely sown green manures are suitable. The selected sampling site should not have been treated with crop protection products for a minimum of one year before sampling. Also, no organic fertiliser should have been applied for at least six months. The use of mineral fertiliser is only acceptable when in accordance with the requirements of the crop and soil samples should not be taken until at least three months after fertiliser application. The use of soil treated with fertilisers with known biocidal effects (e.g. calcium cyanamide) should be avoided.

Sampling should be avoided during or immediately following long periods (greater than 30 days) of drought or water logging. For ploughed soils, samples should be taken from a depth of 0 down to 20 cm. For grassland (pasture) or other soils where ploughing does not occur over longer periods (at least one growing season), the maximum depth of sampling may be slightly more than 20 cm (e.g. to 25 cm).

Soil samples should be transported using containers and under temperature conditions which guarantee that the initial soil properties are not significantly altered.

**1.6.3.2. Storage**

The use of soils freshly collected from the field is preferred. If storage in the laboratory cannot be avoided, soils may be stored in the dark at  $4 \pm 2$  °C for a maximum of three months. During the storage of soils, aerobic conditions must be ensured. If soils are collected from areas where they are frozen for at least three months per year, storage for six months at minus 18 °C to minus 22 °C can be considered. The microbial biomass of stored soils is measured prior to each experiment and the carbon in the biomass should be at least 1 % of the total soil organic carbon content (see Section 1.6.2).

**1.6.4. Handling and preparation of soil for the test****1.6.4.1. Pre-incubation**

If the soil was stored (see Section 1.6.3.2), pre-incubation is recommended for a period between two and 28 days. The temperature and moisture content of the soil during pre-incubation should be similar to that used in the test (see Sections 1.6.4.2 and 1.7.1.3).

**▼B**1.6.4.2. *Physical-chemical characteristics*

The soil is manually cleared of large objects (e.g. stones, parts of plants, etc.) and then moist sieved without excess drying to a particle size less than or equal to 2 mm. The moisture content of the soil sample should be adjusted with distilled or deionised water to a value between 40 % and 60 % of the maximum water holding capacity.

1.6.4.3. *Amendment with organic substrate*

The soil should be amended with a suitable organic substrate, e.g. powdered lucerne-grass-green meal (main component: *Medicago sativa*) with a C/N ratio between 12/1 and 16/1. The recommended lucerne-soil ratio is 5 g of lucerne per kilogram of soil (dry weight).

1.6.5. **Preparation of the test substance for the application to soil**

The test substance is normally applied using a carrier. The carrier can be water (for water soluble substances) or an inert solid such as fine quartz sand (particle size: 0,1-0,5mm). Liquid carriers other than water (e.g. organic solvents such as acetone, chloroform) should be avoided since they can damage the microflora. If sand is used as a carrier, it can be coated with the test substance dissolved or suspended in an appropriate solvent. In such cases, the solvent should be removed by evaporation before mixing with the soil. For an optimum distribution of the test substance in soil, a ratio of 10 g of sand per kilogram of soil (dry weight) is recommended. Control samples are treated with an equivalent amount of water and/or quartz sand only.

When testing volatile chemicals, losses during treatment should be avoided as far as possible and an attempt should be made to ensure homogeneous distribution in the soil (e.g. the test substance should be injected into the soil at several places).

1.6.6. **Test concentrations**

If agrochemicals are tested, at least two concentrations should be used. The lower concentration should reflect at least the maximum amount expected to reach the soil under practical conditions whereas the higher concentration should be a multiple of the lower concentration. The concentrations of test substance added to soil are calculated assuming uniform incorporation to a depth of 5 cm and a soil bulk density of 1,5. For agrochemicals that are applied directly to soil, or for chemicals for which the quantity reaching the soil can be predicted, the test concentrations recommended are the maximum Predicted Environmental Concentration (PEC) and five times that concentration. Substances that are expected to be applied to soils several times in one season should be tested at concentrations derived from multiplying the PEC by the maximum anticipated number of applications. The upper concentration tested, however, should not exceed 10 times the maximum single application rate. If non-agrochemicals are tested, a geometric series of at least five concentrations is used. The concentrations tested should cover the range needed to determine the EC<sub>x</sub> values.

**▼B**

## 1.7. PERFORMANCE OF THE TEST

1.7.1. **Conditions of exposure**1.7.1.1. *Treatment and control*

If agrochemicals are tested, the soil is divided into three portions of equal weight. Two portions are mixed with the carrier containing the product, and the other is mixed with the carrier without the product (control). A minimum of three replicates for both treated and untreated soils is recommended. If non-agrochemicals are tested, the soil is divided into six portions of equal weight. Five of the samples are mixed with the carrier containing the test substance, and the sixth sample is mixed with the carrier without the chemical. Three replicates for both treatments and control are recommended. Care should be taken to ensure homogeneous distribution of the test substance in the treated soil samples. During mixing, compacting or balling of the soil should be avoided.

1.7.1.2. *Incubation of soil samples*

Incubation of soil samples can be performed in two ways: as bulk samples of each treated and untreated soil or as a series of individual and equally sized subsamples of each treated and untreated soil. However, when volatile substances are tested, the test should only be performed with a series of individual subsamples. When soils are incubated in bulk, large quantities of each treated and untreated soils are prepared and subsamples to be analysed are taken as needed during the test. The amount initially prepared for each treatment and control depends on the size of the subsamples, the number of replicates used for analysis and the anticipated maximum number of sampling times. Soils incubated in bulk should be thoroughly mixed before subsampling. When soils are incubated as a series of individual soil samples, each treated and untreated bulk soil is divided into the required number of subsamples, and these are utilised as needed. In the experiments where more than two sampling times can be anticipated, enough subsamples should be prepared to account for all replicates and all sampling times. At least three replicate samples of the test soil should be incubated under aerobic conditions (see Section 1.7.1.1). During all tests, appropriate containers with sufficient headspace should be used to avoid development of anaerobic conditions. When volatile substances are tested, the test should only be performed with a series of individual subsamples.

1.7.1.3. *Test conditions and duration*

The test is carried out in the dark at room temperature of  $20 \pm 2$  °C. The moisture content of soil samples should be maintained during the test between 40 % and 60 % of the maximum water holding capacity of the soil (see Section 1.6.4.2) with a range of  $\pm 5$  %. Distilled, deionised water can be added as needed.

The minimum duration of tests is 28 days. If agrochemicals are tested, the rates of nitrate formation in treated and control samples are compared. If these differ by more than 25 % on day 28, the test is continued until a difference equal to or less than 25 % is obtained, or for a maximum of 100 days, whichever is shorter. For non-agrochemicals, the test is terminated after 28 days. On day 28, the quantities of nitrate in treated and control soil samples are determined and the  $EC_x$  values are calculated.



**▼B****1.7.2. Sampling and analysis of soils****1.7.2.1. Soil sampling schedule**

If agrochemicals are tested, soil samples are analysed for nitrate on days 0, 7, 14 and 28. If a prolonged test is required, further measurements should be made at 14 days intervals after day 28.

If non-agrochemicals are tested, at least five test concentrations are used and soil samples are analysed for nitrate at the beginning (day 0) and at the end of the exposure period (28 days). An intermediate measurement, e.g. at day 7, may be added if deemed necessary. The data obtained on day 28 are used to determine  $EC_x$  value for the chemical. If desired, data from day 0 control samples can be used to report the initial quantity of nitrate in the soil.

**1.7.2.2. Analysis of soil samples**

The amount of nitrate formed in each treated and control replicate is determined at each sampling time. Nitrate is extracted from soil by shaking samples with a suitable extraction solvent, e.g. a 0,1 M potassium chloride solution. A ratio of 5 ml of KCl solution per gram dry weight equivalent of soil is recommended. To optimise extraction, containers holding soil and extraction solution should not be more than half full. The mixtures are shaken at 150 rpm for 60 minutes. The mixtures are centrifuged or filtered and the liquid phases are analysed for nitrate. Particle-free liquid extracts can be stored prior to analysis at  $minus\ 20 \pm 5\ ^\circ C$  for up to six months.

**2. DATA****2.1. TREATMENT OF RESULTS**

If tests are conducted with agrochemicals, the quantity of nitrate formed in each replicate soil sample should be recorded, and the mean values of all replicates should be provided in tabular form. Nitrogen transformation rates should be evaluated by appropriate and generally acceptable statistical methods (e.g. F-test, 5 % significance level). The quantities of nitrate formed are expressed in mg nitrate/kg dry weight soil/day. The nitrate formation rate in each treatment is compared with that in the control, and the percent deviation from the control is calculated.

If tests are conducted with non-agrochemicals, the quantity of nitrate formed in each replicate is determined, and a dose-response curve is prepared for estimation of the  $EC_x$  values. The quantities of nitrate (i.e. mg nitrate/kg dry weight soil) found in the treated samples after 28 days are compared to that found in the control. From these data, the % inhibition values for each test concentration are calculated. These percentages are plotted against concentration, and statistical procedures are then used to calculate the  $EC_x$  values. Confidence limits ( $p = 0,95$ ) for the calculated  $EC_x$  are also determined using standard procedures (10)(11)(12).

Test substances that contain high quantities of nitrogen may contribute to the quantities of nitrate formed during the test. If these substances are tested at a high concentration (e.g. chemicals which are expected to be used in repeated applications) appropriate controls must be included in the test (i.e. soil plus test substance but without plant meal). Data from these controls must be accounted for in the  $EC_x$  calculations.

**▼B**

## 2.2. INTERPRETATION OF RESULTS

When results from tests with agrochemicals are evaluated, and the difference in the rates of nitrate formation between the lower treatment (i.e. the maximum predicted concentration) and control is equal to or less than 25 % at any sampling time after day 28, the product can be evaluated as having no long-term influence on nitrogen transformation in soils. When results from tests with chemicals other than agrochemicals are evaluated, the EC<sub>50</sub>, EC<sub>25</sub> and/or EC<sub>10</sub> values are used.

3. **REPORTING**

The test report must include the following information:

Complete identification of the soil used including:

- geographical reference of the site (latitude, longitude),
- information on the history of the site (i.e. vegetation cover, treatments with crop protection products, treatments with fertilisers, accidental contamination, etc.),
- use pattern (e.g. agricultural soil, forest, etc.),
- depth of sampling (cm),
- sand/silt/clay content (% dry weight),
- pH (in water),
- organic carbon content (% dry weight),
- nitrogen content (% dry weight),
- initial nitrate concentration (mg nitrate/kg dry weight),
- cation exchange capacity (mmol/kg),
- microbial biomass in terms of percentage of the total organic carbon,
- reference of the methods used for the determination of each parameter,
- all information relating to the collection and storage of soil samples,
- details of pre-incubation of soil if any.

Test substance:

- physical nature and, where relevant, physical-chemical properties,
- chemical identification data, where relevant, including structural formula, purity (i.e. for crop protection products the percentage of active ingredient), nitrogen content.

Substrate:

- source of substrate,
- composition (i.e. lucerne meal, lucerne-grass-green meal),
- carbon, nitrogen content (% dry weight),
- sieve size (mm).

**▼ B**

Test conditions:

- details of the amendment of soil with organic substrate,
- number of concentrations of test chemical used and, where appropriate, justification of the selected concentrations,
- details of the application of test substance to soil,
- incubation temperature,
- soil moisture content at the beginning and during the test,
- method of soil incubation used (i.e. as bulk or as a series of individual subsamples),
- number of replicates,
- sampling times,
- method used for extraction of nitrate from soil,

Results:

- analytical procedure and equipment used to analyse nitrate,
- tabulated data including individual and mean values for nitrate measurements,
- variation between the replicates in treated and control samples,
- explanations of corrections made in the calculations, if relevant,
- the percent variation in nitrate formation rates at each sampling time or, if appropriate, the EC<sub>50</sub> value with 95 % confidence limit, other EC<sub>x</sub> (i.e. EC<sub>25</sub> or EC<sub>10</sub>) with confidence intervals, and a graph of the dose-response curve,
- statistical treatment of results,
- all information and observations helpful for the interpretation of the results.

#### 4. REFERENCES

- (1) EPPO, (1994) Decision-Making Scheme for the Environmental Risk Assessment of Plant Protection Chemicals. Chapter 7: Soil Microflora. EPPO Bulletin 24: 1-16, 1994.
- (2) BBA, (1990) Effects on the Activity of the Soil Microflora. BBA Guidelines for the Official Testing of Plant Protection Products, VI, 1-1 (2nd eds., 1990).
- (3) EPA (1987) Soil Microbial Community Toxicity Test. EPA 40 CFR Part 797.3700. Toxic Substances Control Act Test Guidelines; Proposed rule. September 28, 1987.
- (4) SETAC-Europe, (1995) Procedures for assessing the environmental fate and ecotoxicity of pesticides, Ed. M.R. Lynch, Pub. SETAC-Europe, Brussels.
- (5) ISO/DIS 14238 (1995) Soil Quality — Determination of Nitrogen Mineralisation and Nitrification in Soils and the Influence of Chemicals on these Processes. Technical Committee ISO/TC 190/SC 4: *Soil Quality — Biological Methods*.

**▼ B**

- (6) OECD, (1995) Final Report of the OECD Workshop on Selection of Soils/Sediments, Belgirate, Italy, 18-20 January 1995.
- (7) ISO 10381-6 (1993) Soil quality — Sampling. Guidance on the collection, handling and storage of soil for the assessment of aerobic microbial processes in the laboratory.
- (8) ISO 14240-1, (1997) Soil quality — Determination of soil microbial biomass — Part 1: Substrate-induced respiration method.
- (9) ISO 14240-2, (1997) Soil quality — Determination of soil microbial biomass — Part 2: Fumigation-extraction method.
- (10) Litchfield, J.T. and Wilcoxon F., (1949) A simplified method of evaluating dose-effect experiments. *Jour. Pharmacol. and Exper. Ther.*, 96, p. 99-113.
- (11) Finney, D.J., (1971) *Probit Analysis*. 3rd ed., Cambridge, London and New-York.
- (12) Finney, D.J., (1978) *Statistical Methods in biological Assay*. Griffin, Weycombe, UK.

**▼B****C.22. SOIL MICROORGANISMS: CARBON TRANSFORMATION TEST****1. METHOD**

This method is a replicate of OECD TG 217 (2000).

**1.1. INTRODUCTION**

This testing method describes a laboratory method designed to investigate long term potential effects of a single exposure of crop protection products and possibly other chemicals on carbon transformation activity of soil microorganisms. The test is principally based on the recommendations of the European and Mediterranean Plant Protection Organization (1). However, other guideline, including those of the German Biologische Bundesanstalt (2), the US Environmental Protection Agency (3) and SETAC (4), were also taken into account. An OECD Workshop on Soil/Sediment Selection held at Belgirate, Italy, in 1995 (5) agreed on the number and type of soils for use in this test. Recommendations for collection, handling and storage of soil sample are based on an ISO Guidance Document (6) and recommendations from the Belgirate Workshop.

In the assessment and evaluation of toxic characteristics of test substances, determination of effects on soil microbial activity may be required, e.g. when data on the potential side effects of crop protection products on soil microflora are required or when exposure of soil microorganisms to chemicals other than crop protection products is expected. The carbon transformation test is carried out to determine the effects of such chemicals on soil microflora. If agrochemicals (e.g. crop protection products, fertilisers, forestry chemicals) are tested, both carbon transformation and nitrogen transformation tests are conducted. If non-agrochemicals are tested, the nitrogen transformation test is sufficient. However, if EC<sub>50</sub> values of the nitrogen transformation test for such chemicals fall within the range found for commercially available nitrification inhibitors (e.g. nitrapyrin), a carbon transformation test can be conducted to gain further information.

Soils consist of living and non-living components which exist in complex and heterogeneous mixtures. Microorganisms play an important role in breakdown and transformation of organic matter in fertile soils with many species contributing to different aspects of soil fertility. Any long-term interference with these biochemical processes could potentially interfere with nutrient cycling and this could alter the soil fertility. Transformation of carbon and nitrogen occurs in all fertile soils. Although the microbial communities responsible for these processes differ from soil to soil, the pathways of transformation are essentially the same.

**▼B**

This testing method is designed to detect long-term adverse effects of a substance on the process of carbon transformation in aerobic surface soils. The test is sensitive to changes in size and activity of microbial communities responsible for carbon transformation since it subjects these communities to both chemical stress and carbon starvation. A sandy soil low in organic matter is used. This soil is treated with the test substance and incubated under conditions that allow rapid microbial metabolism. Under these conditions, sources of readily available carbon in the soil are rapidly depleted. This causes carbon starvation which both kills microbial cells and induces dormancy and/or sporulation. If the test runs for more than 28 days, the sum of these reactions can be measured in (untreated soil) controls as a progressive loss of metabolically active microbial biomass (7). If the biomass in carbon-stressed soil, under the conditions of the test, is affected by the presence of a chemical, it may not return to the same level as the control. Hence, disturbances caused by the test substance at any time during the test will often last until the end of the test.

The tests from which this testing method was developed were primarily designed for substances for which the amount reaching the soil can be anticipated. This is the case, for example, for crop protection products for which the application rate in the field is known. For agrochemicals, testing of two doses relevant to the anticipated or predicted application rate is sufficient. Agrochemicals can be tested as active ingredients (a.i.) or as formulated products. However, the test is not limited to chemicals with predictable environmental concentrations. By changing both the amounts of test substance applied to the soil, and the way in which the data are evaluated, the test can also be used for chemicals for which the amount expected to reach the soil is not known. Thus, with non-agrochemicals, the effects of a series of concentrations on carbon transformation are determined. The data from these tests are used to prepare a dose-response curve and calculate  $EC_x$  values, where x is defined % effect.

## 1.2. DEFINITIONS

**Carbon transformation:** is the degradation by microorganisms of organic matter to form inorganic end-product carbon dioxide.

**$EC_x$  (Effective Concentration):** is the concentration of the test substance in soil that results in a x % inhibition of carbon transformation in carbon dioxide.

**$EC_{50}$  (Median Effective Concentration):** is the concentration of test substance in soil that results in a 50 % inhibition of carbon transformation in carbon dioxide.

## 1.3. REFERENCE SUBSTANCES

None.

**▼ B**

## 1.4. PRINCIPLE OF THE TEST METHOD

Sieved soil is either treated with the test substance or left untreated (control). If agrochemicals are tested, a minimum of two test concentrations are recommended and these should be chosen in relation to the highest concentration anticipated in the field. After 0, 7, 14 and 28 days incubation, samples of treated and control soils are mixed with glucose, and glucose-induced respiration rates are measured for 12 consecutive hours. Respiration rates are expressed as carbon dioxide released (mg carbon dioxide/kg dry soil/h) or oxygen consumed (mg oxygen/kg soil/h). The mean respiration rate in the treated soil samples is compared with that in control and the percent deviation of the treated from the control is calculated. All tests run for at least 28 days. If, on the 28th day, differences between treated and untreated soils are equal to or greater than 25 % measurements are continued in 14 day intervals for a maximum of 100 days. If chemicals other than agrochemicals are tested, a series of concentrations of the test substance are added to samples of the soil, and glucose induced respiration rates (i.e. the mean of the quantities of carbon dioxide formed or oxygen consumed) are measured after 28 days. Results from tests with a series of concentrations are analysed using a regression model, and the EC<sub>x</sub> values are calculated (i.e. EC<sub>50</sub>, EC<sub>25</sub> and/or EC<sub>10</sub>). See definitions.

## 1.5. VALIDITY OF THE TEST

Evaluations of test results with agrochemicals are based on relatively small differences (i.e. average value  $\pm 25$  %) between the carbon dioxide released or the oxygen consumed in (or by) control and treated soil samples, so large variations in the controls can lead to false results. Therefore, the variation between replicate control samples should be less than  $\pm 15$  %.

## 1.6. DESCRIPTION OF THE TEST METHOD

1.6.1. **Apparatus**

Test containers made of chemically inert material are used. They should be of a suitable capacity in compliance with the procedure used for incubation of soils, i.e. incubation in bulk or as a series of individual soil samples (see Section 1.7.1.2). Care should be taken both to minimise water loss and to allow gas exchange during the test (e.g. the test containers may be covered with perforated polyethylene foil). When volatile substances are tested, sealable and gas-tight containers should be used. These should be of a size such that approximately one quarter of their volume is filled with the soil sample.

For determination of glucose-induced respiration, incubation systems and instruments for measurement of carbon dioxide production or oxygen consumption are required. Examples of such systems and instruments are found in the literature (8) (9) (10) (11).

1.6.2. **Selection and number of soils**

One single soil is used. The recommended soil characteristics are as follows:

— sand content: not less than 50 % and not greater than 75 %,

**▼ B**

- pH: 5,5-7,5,
- organic carbon content: 0,5-1,5 %,
- the microbial biomass should be measured (12)(13) and its carbon content should be at least 1 % of the total soil organic carbon.

In most cases, a soil with these characteristics represents a worst case situation, since adsorption of the test chemical is minimised and its availability to the microflora is maximum. Consequently, tests with other soils are generally unnecessary. However, in certain circumstances, e.g. where the anticipated major use of the test substance is in particular soils such as acidic forest soils, or for electrostatically charged chemicals, it may be necessary to substitute an additional soil.

**1.6.3. Collection and storage of soil samples****1.6.3.1. Collection**

Detailed information on the history of the field site from where the test soil is collected should be available. Details include exact location, vegetation cover, dates of treatments with crop protection products, treatments with organic and inorganic fertilisers, additions of biological materials or accidental contaminations. The site chosen for soil collection should be one which allows long-term use. Permanent pastures, fields with annual cereal crops (except maize) or densely sown green manures are suitable. The selected sampling site should not have been treated with crop protection products for a minimum of one year before sampling. Also, no organic fertiliser should have been applied for at least six months. The use of mineral fertiliser is only acceptable when in accordance with the requirements of the crop and soil samples should not be taken until at least three months after fertiliser application. The use of soil treated with fertilisers with known biocidal effects (e.g. calcium cyanamide) should be avoided.

Sampling should be avoided during or immediately following long periods (greater than 30 days) of drought or water logging. For ploughed soils, samples should be taken from a depth of 0 down to 20 cm. For grassland (pasture) or other soils where ploughing does not occur over longer periods (at least one growing season), the maximum depth of sampling may be slightly more than 20 cm (e.g. to 25 cm). Soil samples should be transported using containers and under temperature conditions which guarantee that the initial soil properties are not significantly altered.

**1.6.3.2. Storage**

The use of soils freshly collected from the field is preferred. If storage in the laboratory cannot be avoided, soils may be stored in the dark at  $4 \pm 2$  °C for a maximum of three months. During the storage of soils, aerobic conditions must be ensured. If soils are collected from areas where they are frozen for at least three months per year, storage for six months at minus 18 °C can be considered. The microbial biomass of stored soils is measured prior to each experiment and the carbon in the biomass should be at least 1 % of the total soil organic carbon content (see Section 1.6.2).



**▼B****1.6.4. Handling and preparation of soil for the test****1.6.4.1. Pre-incubation**

If the soil was stored (see Sections 1.6.4.2 and 1.7.1.3), pre-incubation is recommended for a period between two and 28 days. The temperature and moisture content of the soil during pre-incubation should be similar to that used in the test (see Sections 1.6.4.2 and 1.7.1.3).

**1.6.4.2. Physical-chemical characteristics**

The soil is manually cleared of large objects (e.g. stones, parts of plants, etc.) and then moist sieved without excess drying to a particle size less than or equal to 2 mm. The moisture content of the soil sample should be adjusted with distilled or deionised water to a value between 40 % and 60 % of the maximum water holding capacity.

**1.6.5. Preparation of the test substance for the application to soil**

The test substance is normally applied using a carrier. The carrier can be water (for water soluble substances) or an inert solid such as fine quartz sand (particle size: 0,1-0,5 mm). Liquid carriers other than water (e.g. organic solvents such as acetone, chloroform) should be avoided since they can damage the microflora. If sand is used as a carrier, it can be coated with the test substance dissolved or suspended in an appropriate solvent. In such cases, the solvent should be removed by evaporation before mixing with the soil. For an optimum distribution of the test substance in soil, a ratio of 10 g of sand per kilogram of soil (dry weight) is recommended. Control samples are treated with the equivalent amount of water and/or quartz sand only.

When testing volatile chemicals, losses during treatment should be avoided and an attempt should be made to ensure homogeneous distribution in the soil (e.g. the test substance should be injected into the soil at several places).

**1.6.6. Test concentrations**

If crop protection products or other chemicals with predictable environmental concentrations are tested, at least two concentrations should be used. The lower concentration should reflect at least the maximum amount expected to reach the soil under practical conditions whereas the higher concentration should be a multiple of the lower concentration. The concentrations of test substance added to soil are calculated assuming uniform incorporation to a depth of 5 cm and a soil bulk density of 1,5. For agrochemicals that are applied directly to soil, or for chemicals for which the quantity reaching the soil can be predicted, the test concentrations recommended are the Predictable Environmental Concentration (PEC) and five times that concentration. Substances that are expected to be applied to soils several times in one season should be tested at concentrations derived from multiplying the PEC by the maximum anticipated number of applications. The upper concentration tested, however, should not exceed 10 times the maximum single application rate.

If non-agrochemicals are tested, a geometric series of at least five concentrations is used. The concentrations tested should cover the range needed to determine the EC<sub>x</sub> values.

**▼B**

## 1.7. PERFORMANCE OF THE TEST

1.7.1. **Conditions of exposure**1.7.1.1. *Treatment and control*

If agrochemicals are tested, the soil is divided into three portions of equal weight. Two portions are mixed with the carrier containing the product, and the other is mixed with the carrier without the product (control). A minimum of three replicates for both treated and untreated soils is recommended. If non-agrochemicals are tested, the soil is divided into six portions of equal weight. Five of the samples are mixed with the carrier containing the test substance, and the sixth sample is mixed with the carrier without the chemical. Three replicates for both treatments and control are recommended. Care should be taken to ensure homogeneous distribution of the test substance in the treated soil samples. During mixing, compacting or balling of the soil should be avoided.

1.7.1.2. *Incubation of soil samples*

Incubation of soil samples can be performed in two ways: as bulk samples of each treated and untreated soil or as a series of individual and equally sized subsamples of each treated and untreated soil. However, when volatile substances are tested, the test should only be performed with a series of individual subsamples. When soils are incubated in bulk, large quantities of each treated and untreated soils are prepared and subsamples to be analysed are taken as needed during the test. The amount initially prepared for each treatment and control depends on the size of the subsamples, the number of replicates used for analysis and the anticipated maximum number of sampling times. Soils incubated in bulk should be thoroughly mixed before subsampling. When soils are incubated as a series of individual soil samples, each treated and untreated bulk soil is divided into the required number of subsamples, and these are utilised as needed. In the experiments where more than two sampling times can be anticipated, enough subsamples should be prepared to account for all replicates and all sampling times. At least three replicate samples of the test soil should be incubated under aerobic conditions (see Section 1.7.1.1). During all tests, appropriate containers with sufficient headspace should be used to avoid development of anaerobic conditions. When volatile substances are tested, the test should only be performed with a series of individual subsamples.

1.7.1.3. *Test conditions and duration*

The test is carried out in the dark at room temperature of  $20 \pm 2$  °C. The moisture content of soil samples should be maintained during the test between 40 % and 60 % of the maximum water holding capacity of the soil (see Section 1.6.4.2) with a range of  $\pm 5$  %. Distilled, deionised water can be added as needed.

The minimum duration of tests is 28 days. If agrochemicals are tested, the quantities of carbon dioxide released or oxygen consumed in treated and control samples are compared. If these differ by more than 25 % on day 28, the test is continued until a difference equal to or less than 25 % is obtained, or for a maximum of 100 days, whichever is shorter. If non-agrochemicals are tested, the test is terminated after 28 days. On day 28, the quantities of carbon dioxide released or oxygen consumed in treated and control soil samples are determined and the  $EC_x$  values are calculated.

**▼B****1.7.2. Sampling and analysis of soils****1.7.2.1. Soil sampling schedule**

If agrochemicals are tested, soil samples are analysed for glucose-induced respiration rates on days 0, 7, 14 and 28. If a prolonged test is required, further measurements should be made at 14 days intervals after day 28.

If non-agrochemicals are tested, at least five test concentrations are used and soil samples are analysed for glucose-induced respiration at the beginning (day 0) and at the end of the exposure period (28 days). An intermediate measurement, e.g. at day 7, may be added if deemed necessary. The data obtained on day 28 are used to determine  $EC_x$  value for the chemical. If desired, data from day 0 control samples can be used to estimate the initial quantities of metabolically active microbial biomass in the soil (12).

**1.7.2.2. Measurement of glucose-induced respiration rates**

The glucose-induced respiration rate in each treated and control replicate is determined at each sampling time. The soil samples are mixed with a sufficient amount of glucose to elicit an immediate maximum respiratory response. The amount of glucose needed to elicit a maximum respiratory response from a given soil can be determined in a preliminary test using a series of concentrations of glucose (14). However, for sandy soils with 0,5-1,5 % organic carbon, 2 000 mg to 4 000 mg glucose per kg dry weight soil is usually sufficient. The glucose can be ground to a powder with clean quartz sand (10 g sand/kg dry weight soil) and homogeneously mixed with the soil.

The glucose amended soil samples are incubated in a suitable apparatus for measurement of respiration rates either continuously, every hour, or every two hours (see Section 1.6.1) at  $20 \pm 2$  °C. The carbon dioxide released or the oxygen consumed is measured for 12 consecutive hours and measurements should start as soon as possible, i.e. within one to two hours after glucose supplement. The total quantities of carbon dioxide released or oxygen consumed during the 12 hours are measured and mean respiration rates are determined.

**2. DATA****2.1. TREATMENT OF RESULTS**

If agrochemicals are tested, the carbon dioxide released from, or oxygen consumed by each replicate soil sample should be recorded, and the mean values of all replicates should be provided in tabular form. Results should be evaluated by appropriate and generally acceptable statistical methods (e.g. F-test, 5 % significance level). Glucose-induced respiration rates are expressed in mg carbon dioxide/kg dry weight soil/h or mg oxygen/dry weight soil/h. The mean carbon dioxide formation rate or mean oxygen consumption rate in each treatment is compared with that in control, and the percent deviation from the control is calculated.

**▼ B**

If tests are conducted with non-agrochemicals, the quantities of carbon dioxide released or oxygen consumed by each replicate is determined, and a dose-response curve is prepared for estimation of the  $EC_x$  values. The glucose-induced respiration rates (i.e. mg carbon dioxide/kg dry weight soil/h or mg oxygen/dry weight soil/h) found in the treated samples after 28 days are compared to that found in control. From these data, the % inhibition values for each test concentration are calculated. These percentages are plotted against concentration, and statistical procedures are used to calculate the  $EC_x$  values. Confidence limits ( $p = 0,95$ ) for the calculated  $EC_x$  are also determined using standard procedures (15)(16)(17).

## 2.2. INTERPRETATION OF RESULTS

When results from tests with agrochemicals are evaluated, and the difference in respiration rates between the lower treatment (i.e. the maximum predicted concentration) and control is equal to or less than 25 % at any sampling time after day 28, the product can be evaluated as having no long-term influence on carbon transformation in soils. When results from tests with chemicals other than agrochemicals are evaluated, the  $EC_{50}$ ,  $EC_{25}$  and/or  $EC_{10}$  values are used.

## 3. REPORTING

### TEST REPORT

The test report must include the following information:

Complete identification of the soil used including:

- geographical reference of the site (latitude, longitude),
- information on the history of the site (i.e. vegetation cover, treatments with crop protection products, treatments with fertilisers, accidental contamination, etc.),
- use pattern (e.g. agricultural soil, forest, etc.),
- depth of sampling (cm),
- sand/silt/clay content (% dry weight),
- pH (in water),
- organic carbon content (% dry weight),
- nitrogen content (% dry weight);
- cation exchange capacity (mmol/kg),
- initial microbial biomass in terms of percentage of the total organic carbon,
- reference of the methods used for the determination of each parameter,
- all information relating to the collection and storage of soil samples,
- details of pre-incubation of soil if any.

**▼ B**

Test substance:

- physical nature and, where relevant, physical-chemical properties,
- chemical identification data, where relevant, including structural formula, purity (i.e. for crop protection products the percentage of active ingredient), nitrogen content.

Test conditions:

- details of the amendment of soil with organic substrate,
- number of concentrations of test chemical used and, where appropriate, justification of the selected concentrations,
- details of the application of test substance to soil,
- incubation temperature,
- soil moisture content at the beginning and during the test,
- method of soil incubation used (i.e. as bulk or as a series of individual subsamples),
- number of replicates,
- sampling times.

Results:

- method and equipment used for measurement of respiration rates,
- tabulated data including individual and mean values for quantities of carbon dioxide or oxygen,
- variation between the replicates in treated and control samples,
- explanations of corrections made in the calculations, if relevant,
- the percent variation of glucose-induced respiration rates at each sampling time or, if appropriate, the EC<sub>50</sub> with 95 % confidence limit, other EC<sub>x</sub> (i.e. EC<sub>25</sub> or EC<sub>10</sub>) with confidence intervals, and a graph of the dose-response curve,
- statistical treatment of results, where appropriate,
- all information and observations helpful for the interpretation of the results.

#### 4. REFERENCES

- (1) EPPO, (1994) Decision-Making Scheme for the Environmental Risk Assessment of Plant Protection Chemicals. Chapter 7: Soil Microflora. EPPO Bulletin 24: 1-16, 1994.
- (2) BBA, (1990) Effects on the Activity of the Soil Microflora. BBA Guidelines for the Official Testing of Plant Protection Products, VI, 1-1 (2nd eds., 1990).
- (3) EPA, (1987) Soil Microbial Community Toxicity Test. EPA 40 CFR Part 797.3700. Toxic Substances Control Act Test Guidelines; Proposed rule. September 28, 1987.

**▼ B**

- (4) SETAC-Europe, (1995) Procedures for assessing the environmental fate and ecotoxicity of pesticides, Ed. M.R. Lynch, Pub. SETAC-Europe, Brussels.
- (5) OECD, (1995) Final Report of the OECD Workshop on Selection of Soils/Sediments, Belgirate, Italy, 18-20 January 1995.
- (6) ISO 10381-6, (1993) Soil quality — Sampling. Guidance on the collection, handling and storage of soil for the assessment of aerobic microbial processes in the laboratory.
- (7) Anderson, J.P.E., (1987) Handling and Storage of Soils for Pesticide Experiments, in 'Pesticide Effects on Soil Microflora'. Eds. L. Somerville and M.P. Greaves, Chap. 3: 45-60.
- (8) Anderson, J.P.E., (1982) Soil Respiration, in 'Methods of Soil Analysis — Part 2: Chemical and Microbiological Properties'. Agronomy Monograph No 9. Eds. A.L. Page, R.H. Miller and D.R. Keeney. 41: 831- 871.
- (9) ISO 11266-1, (1993) Soil Quality — Guidance on Laboratory Tests for Biodegradation in Soil: Part 1. Aerobic Conditions.
- (10) ISO 14239, (1997E) Soil Quality — Laboratory incubation systems for measuring the mineralisation of organic chemicals in soil under aerobic conditions.
- (11) Heinemeyer O., Insam, H., Kaiser, E.A, and Walenzik, G., (1989) Soil microbial biomass and respiration measurements; an automated technique based on infrared gas analyses. *Plant and Soil*, 116: 77-81.
- (12) ISO 14240-1, (1997) Soil quality — Determination of soil microbial biomass — Part 1: Substrate-induced respiration method.
- (13) ISO 14240-2, (1997) Soil quality — Determination of soil microbial biomass — Part 2: Fumigation-extraction method.
- (14) Malkomes, H.-P., (1986) Einfluß von Glukosemenge auf die Reaktion der Kurzzeit-Atmung im Boden Gegenüber Pflanzenschutzmitteln, Dargestellt am Beispiel eines Herbizide. (Influence of the Amount of Glucose Added to the Soil on the Effect of Pesticides in Short-Term Respiration, using a Herbicide as an Example). *Nachrichtenbl. Deut. Pflanzenschutzd., Braunschweig*, 38: 113-120.
- (15) Litchfield, J.T. and Wilcoxon, F., (1949) A simplified method of evaluating dose-effect experiments. *Jour. Pharmacol. and Exper. Ther.*, 96, 99-113.
- (16) Finney, D.J., (1971) *Probit Analysis*. 3rd ed., Cambridge, London and New-York.
- (17) Finney D.J., (1978) *Statistical Methods in biological Assay*. Griffin, Weycombe, UK.

**▼B****C.23. AEROBIC AND ANAEROBIC TRANSFORMATION IN SOIL****1. METHOD**

This test method is a replicate of the OECD TG 307 (2002)

**1.1. INTRODUCTION**

This test method is based on existing guidelines (1)(2)(3)(4)(5)(6)(7)(8)(9). The method described in this test Method is designed for evaluating aerobic and anaerobic transformation of chemicals in soil. The experiments are performed to determine (i) the rate of transformation of the test substance, and (ii) the nature and rates of formation and decline of transformation products to which plants and soil organisms may be exposed. Such studies are required for chemicals which are directly applied to soil or which are likely to reach the soil environment. The results of such laboratory studies can also be used to develop sampling and analysis protocols for related field studies.

Aerobic and anaerobic studies with one soil type are generally sufficient for the evaluation of transformation pathways (8)(10)(11). Rates of transformation should be determined in at least three additional soils (8)(10).

An OECD Workshop on soil and sediment selection, held at Belgirate, Italy in 1995 (10) agreed, in particular, on the number and types of soils for use in this test. The types of soils tested should be representative of the environmental conditions where use or release will occur. For example, chemicals that may be released in subtropical to tropical climates should be tested with Ferrasols or Nitosols (FAO system). The Workshop also made recommendations relating to collection, handling and storage of soil samples, based on the ISO Guidance (15). The use of paddy (rice) soils is also considered in this method.

**1.2. DEFINITIONS**

**Test substance:** any substance, whether the parent compound or relevant transformation products.

**Transformation products:** all substances resulting from biotic or abiotic transformation reactions of the test substance including CO<sub>2</sub> and products that are in bound residues.

**Bound residues:** 'Bound residues' represent compounds in soil, plant or animal, which persist in the matrix in the form of the parent substance or its metabolite(s)/transformation products after extraction. The extraction method must not substantially change the compounds themselves or the structure of the matrix. The nature of the bond can be clarified in part by matrix-altering extraction methods and sophisticated analytical techniques. To date, for example, covalent ionic and sorptive bonds, as well as entrapments, have been identified in this way. In general, the formation of bound residues reduces the bioaccessibility and the bioavailability significantly (12) [modified from IUPAC 1984 (13)].

**Aerobic transformation:** reactions occurring in the presence of molecular oxygen (14).

**▼B**

**Anaerobic transformation:** reactions occurring under exclusion of molecular oxygen (14).

**Soil:** is a mixture of mineral and organic chemical constituents, the latter containing compounds of high carbon and nitrogen content and of high molecular weights, animated by small (mostly micro-) organisms. Soil may be handled in two states:

- (a) undisturbed, as it has developed with time, in characteristic layers of a variety of soil types;
- (b) disturbed, as it is usually found in arable fields or as occurs when samples are taken by digging and used in this test method (14).

**Mineralisation:** is the complete degradation of an organic compound to CO<sub>2</sub> and H<sub>2</sub>O under aerobic conditions, and CH<sub>4</sub>, CO<sub>2</sub> and H<sub>2</sub>O under anaerobic conditions. In the context of this test method, when <sup>14</sup>C-labelled compound is used, mineralisation means extensive degradation during which a labelled carbon atom is oxidised with release of the appropriate amount of <sup>14</sup>CO<sub>2</sub> (14).

**Half-life:**  $t_{0,5}$ , is the time taken for 50 % transformation of a test substance when the transformation can be described by first-order kinetics; it is independent of the concentration.

**DT<sub>50</sub> (Disappearance Time 50):** is the time within which the concentration of the test substance is reduced by 50 %; it is different from the half-life  $t_{0,5}$  when transformation does not follow first order kinetics.

**DT<sub>75</sub> (Disappearance Time 75):** is the time within which the concentration of the test substance is reduced by 75 %.

**DT<sub>90</sub> (Disappearance Time 90):** is the time within which the concentration of the test substance is reduced by 90 %.

### 1.3. REFERENCE SUBSTANCES

Reference substances should be used for the characterisation and/or identification of transformation products by spectroscopic and chromatographic methods.

### 1.4. APPLICABILITY OF THE TEST

The method is applicable to all chemical substances (non-labelled or radiolabelled) for which an analytical method with sufficient accuracy and sensitivity is available. It is applicable to slightly volatile, non-volatile, water-soluble or water-insoluble compounds. The test should not be applied to chemicals which are highly volatile from soil (e.g. fumigants, organic solvents) and thus cannot be kept in soil under the experimental conditions of this test.



**▼B**

## 1.5. INFORMATION ON THE SUBSTANCE

Non-labelled or labelled test substance can be used to measure the rate of transformation. Labelled material is required for studying the pathway of transformation and for establishing a mass balance. <sup>14</sup>C-labelling is recommended but the use of other isotopes, such as <sup>13</sup>C, <sup>15</sup>N, <sup>3</sup>H, <sup>32</sup>P, may also be useful. As far as possible, the label should be positioned in the most stable part(s) of the molecule<sup>(1)</sup>. The purity of the test substance should be at least 95 %.

Before carrying out a test on aerobic and anaerobic transformation in soil, the following information on the test substance should be available:

- (a) solubility in water (Method A.6)
- (b) solubility in organic solvents;
- (c) vapour pressure (Method A.4) and Henry's law constant;
- (d) n-octanol/water partition coefficient (Method A.8);
- (e) chemical stability in dark (hydrolysis) (Method C.7);
- (f) pK<sub>a</sub> if a molecule is liable to protonation or deprotonation [OECD Guideline 112] (16).

Other useful information may include data on toxicity of the test substance to soil micro-organisms [testing methods C.21 and C.22] (16).

Analytical methods (including extraction and clean-up methods) for quantification and identification of the test substance and its transformation products should be available.

## 1.6. PRINCIPLE OF THE TEST METHOD

Soil samples are treated with the test substance and incubated in the dark in biometer-type flasks or in flow-through systems under controlled laboratory conditions (at constant temperature and soil moisture). After appropriate time intervals, soil samples are extracted and analysed for the parent substance and for transformation products. Volatile products are also collected for analysis using appropriate absorption devices. Using <sup>14</sup>C-labelled material, the various mineralisation rates of the test substance can be measured by trapping evolved <sup>14</sup>CO<sub>2</sub> and a mass balance, including the formation of soil bound residues, can be established.

## 1.7. QUALITY CRITERIA

1.7.1. **Recovery**

Extraction and analysis of, at least, duplicate soil samples immediately after the addition of the test substance gives a first indication of the repeatability of the analytical method and of the uniformity of the application procedure for the test substance. Recoveries for later stages of the experiments are given by the respective mass balances. Recoveries should range from 90 % to 110 % for labelled chemicals (8) and from 70 % to 110 % for non-labelled chemicals (3).

<sup>(1)</sup> For example, if the test substance contains one ring, labelling on this ring is required; if the test substance contains two or more rings, separate studies may be needed to evaluate the fate of each labelled ring and to obtain suitable information on formation of transformation products.

**▼ B****1.7.2. Repeatability and sensitivity of analytical method**

Repeatability of the analytical method (excluding the initial extraction efficiency) to quantify test substance and transformation products can be checked by duplicate analysis of the same extract of the soil, incubated long enough for formation of transformation products.

The limit of detection (LOD) of the analytical method for the test substance and for the transformation products should be at least 0,01 mg·kg<sup>-1</sup> soil (as test substance) or 1 % of applied dose whichever is lower. The limit of quantification (LOQ) should also be specified.

**1.7.3. Accuracy of transformation data**

Regression analysis of the concentrations of the test substance as a function of time gives the appropriate information on the reliability of the transformation curve and allows the calculation of the confidence limits for half-lives (in the case of pseudo first order kinetics) or DT<sub>50</sub> values and, if appropriate, DT<sub>75</sub> and DT<sub>90</sub> values.

**1.8. DESCRIPTION OF THE METHOD****1.8.1. Equipment and chemical reagents**

Incubation systems consist of static closed systems or suitable flow-through systems (7)(17). Examples of suitable flow-through soil incubation apparatus and biometer-type flask are shown in Figures 1 and 2, respectively. Both types of incubation systems have advantages and limitations (7)(17).

Standard laboratory equipment is required and especially the following:

- analytical instruments such as GLC, HPLC, TLC-equipment, including the appropriate detection systems for analysing radio-labelled or non-labelled substances or inverse isotopes dilution method,
- instruments for identification purposes (e.g. MS, GC-MS, HPLC-MS, NMR, etc.),
- liquid scintillation counter,
- oxidiser for combustion of radioactive material,
- centrifuge,
- extraction apparatus (for example, centrifuge tubes for cold extraction and Soxhlet apparatus for continuous extraction under reflux),
- instrumentation for concentrating solutions and extracts (e.g. rotating evaporator),
- water bath,
- mechanical mixing device (e.g. kneading machine, rotating mixer).

**▼B**

Chemical reagents used include, for example:

- NaOH, analytical grade, 2 mol· dm<sup>-3</sup>, or other appropriate base (e.g. KOH, ethanolamine),
- H<sub>2</sub>SO<sub>4</sub>, analytical grade, 0,05 mol· dm<sup>-3</sup>,
- ethylene glycol, analytical grade,
- solid absorption materials such as soda lime and polyurethane plugs,
- organic solvents, analytical grade, such as acetone, methanol, etc.,
- scintillation liquid.

**1.8.2. Test substance application**

For addition to and distribution in soil, the test substance can be dissolved in water (deionised or distilled) or, when necessary, in minimum amounts of acetone or other organic solvents (6) in which the test substance is sufficiently soluble and stable. However, the amount of solvent selected should not have a significant influence on soil microbial activity (see Sections 1.5 and 1.9.2-1.9.3). The use of solvents which inhibit microbial activity, such as chloroform, dichloromethane and other halogenated solvents, should be avoided.

The test substance can also be added as a solid, e.g. mixed in quartz sand (6) or in a small sub-sample of the test soil which has been air-dried and sterilised. If the test substance is added using a solvent the solvent should be allowed to evaporate before the spiked sub-sample is added to the original non-sterile soil sample.

For general chemicals, whose major route of entry into soil is through sewage sludge/farming application, the test substance should be first added to sludge which is then introduced into the soil sample. (see Sections 1.9.2 and 1.9.3)

The use of formulated products is not routinely recommended. However, e.g. for poorly soluble test substances, the use of formulated material may be an appropriate alternative.

**1.8.3. Soils****1.8.3.1. Soil selection**

To determine the transformation pathway, a representative soil can be used; a sandy loam or silty loam or loam or loamy sand (according to FAO and USDA classification (18)), with a pH of 5,5-8,0, an organic carbon content of 0,5-2,5 % and a microbial biomass of at least 1 % of total organic carbon is recommended (10).

For transformation rate studies at least three additional soils should be used representing a range of relevant soils. The soils should vary in their organic carbon content, pH, clay content and microbial biomass (10).

**▼ B**

All soils should be characterised, at least, for texture (% sand, % silt, % clay) [according to FAO and USDA classification (18)], pH, cation exchange capacity, organic carbon, bulk density, water retention characteristic <sup>(1)</sup> and microbial biomass (for aerobic studies only). Additional information on soil properties may be useful in interpreting the results. For determination of the soil characteristics the methods recommended in references (19)(20)(21)(22)(23) can be used. Microbial biomass should be determined by using the substrate-induced respiration (SIR) method (25)(26) or alternative methods (20).

### 1.8.3.2. *Collection, handling, and storage of soils*

Detailed information on the history of the field site from where the test soil is collected should be available. Details include exact location, vegetation cover, treatments with chemicals, treatments with organic and inorganic fertilisers, additions of biological materials or other contamination. If soils have been treated with the test substance or its structural analogues within the previous four years, these should not be used for transformation studies (10)(15).

The soil should be freshly collected from the field (from the A horizon or top 20 cm layer) with a soil water content which facilitates sieving. For soils other than those from paddy fields, sampling should be avoided during or immediately following long periods (> 30 days) of drought, freezing or flooding (14). Samples should be transported in a manner which minimises changes in soil water content and should be kept in the dark with free access of air, as much as possible. A loosely-tied polyethylene bag is generally adequate for this purpose.

The soil should be processed as soon as possible after sampling. Vegetation, larger soil fauna and stones should be removed prior to passing the soil through a 2 mm sieve which removes small stones, fauna and plant debris. Extensive drying and crushing of the soil before sieving should be avoided (15).

When sampling in the field is difficult in winter (soil frozen or covered by layers of snow), it may be taken from a batch of soil stored in the greenhouse under plant cover (e.g. grass or grass-clover mixtures). Studies with soils freshly collected from the field are strongly preferred, but if the collected and processed soil has to be stored prior to the start of the study storage conditions must be adequate and for a limited time only ( $4 \pm 2$  °C for a maximum of three months) to maintain microbial activity <sup>(2)</sup>. Detailed instructions on collection, handling and storage of soils to be used for biotransformation experiments can be found in (8)(10)(15)(26)(27).

<sup>(1)</sup> Water retention characteristic of a soil can be measured as field capacity, as water holding capacity or as water suction tension (pF). For explanations see Appendix 1. It should be reported in the test report whether water retention characteristics and bulk density of soils were determined in undisturbed field samples or in disturbed (processed) samples.

<sup>(2)</sup> Recent research results indicate that soils from temperate zones can also be stored at – 20 °C for more than three months (28)(29) without significant losses of microbial activity.

**▼B**

Before the processed soil is used for this test, it should be pre-incubated to allow germination and removal of seeds, and to re-establish equilibrium of microbial metabolism following the change from sampling or storage conditions to incubation conditions. A pre-incubation period between two and 28 days approximating the temperature and moisture conditions of the actual test is generally adequate (15). Storage and pre-incubation time together should not exceed three months.

**1.9. PERFORMANCE OF THE TEST****1.9.1. Test conditions****1.9.1.1. Test temperature**

During the whole test period, the soils should be incubated in the dark at a constant temperature representative of the climatic conditions where use or release will occur. A temperature of  $20 \pm 2$  °C is recommended for all test substances which may reach the soil in temperate climates. The temperature should be monitored.

For chemicals applied or released in colder climates (e.g. in northern countries, during autumn/winter periods), additional soil samples should be incubated but at a lower temperature (e.g.  $10 \pm 2$  °C).

**1.9.1.2. Moisture content**

For transformation tests under aerobic conditions, the soil moisture content (1) should be adjusted to and maintained at a pF between 2,0 and 2,5 (3). The soil moisture content is expressed as mass of water per mass of dry soil and should be regularly controlled (e.g. in 2 week intervals) by weighing of the incubation flasks and water losses compensated by adding water (preferably sterile-filtered tap water). Care should be given to prevent or minimise losses of test substance and/or transformation products by volatilisation and/or photodegradation (if any) during moisture addition.

For transformation tests under anaerobic and paddy conditions, the soil is water-saturated by flooding.

**1.9.1.3. Aerobic incubation conditions**

In the flow-through systems, aerobic conditions will be maintained by intermittent flushing or by continuously ventilating with humidified air. In the biometer flasks, exchange of air is maintained by diffusion.

**1.9.1.4. Sterile aerobic conditions**

To obtain information on the relevance of abiotic transformation of a test substance, soil samples may be sterilised (for sterilisation methods see references 16 and 29), treated with sterile test substance (e.g. addition of solution through a sterile filter) and aerated with humidified sterile air as described in Section 1.9.1.3. For paddy soils, soil and water should be sterilised and the incubation should be carried out as described in Section 1.9.1.6.

(1) The soil should neither be too wet nor too dry to maintain adequate aeration and nutrition of soil microflora. Moisture contents recommended for optimal microbial growth range from 40-60 % water holding capacity (WHC) and from 0,1-0,33 bar (6). The latter range is equivalent to a pF-range of 2,0-2,5. Typical moisture contents of various soil types are given in Appendix 2.

**▼ B**1.9.1.5. *Anaerobic incubation conditions*

To establish and maintain anaerobic conditions, the soil treated with the test substance and incubated under aerobic conditions for 30 days or one half-life or  $DT_{50}$  (whichever is shorter) is then water-logged (1-3 cm water layer) and the incubation system flushed with an inert gas (e.g. nitrogen or argon) <sup>(1)</sup>. The test system must allow for measurements such as pH, oxygen concentration and redox potential and include trapping devices for volatile products. The biometer-type system must be closed to avoid entrance of air by diffusion.

1.9.1.6. *Paddy incubation conditions*

To study transformation in paddy rice soils, the soil is flooded with a water layer of about 1-5 cm and the test substance applied to the water phase (9). A soil depth of at least 5 cm is recommended. The system is ventilated with air as under aerobic conditions. pH, oxygen concentration and redox potential of the aqueous layer should be monitored and reported. A pre-incubation period of at least two weeks is necessary before commencing transformation studies (see Section 1.8.3.2).

1.9.1.7. *Test duration*

The rate and pathway studies should normally not exceed 120 days <sup>(2)</sup> (3)(6)(8), because thereafter a decrease of the soil microbial activity with time would be expected in an artificial laboratory system isolated from natural replenishment. Where necessary to characterise the decline of the test substance and the formation and decline of major transformation products, studies can be continued for longer periods (e.g. 6 or 12 months) (8). Longer incubation periods should be justified in the test report and accompanied by biomass measurements during and at the end of these periods.

1.9.2. **Performance of the test**

About 50 to 200 g of soil (dry weight basis) are placed into each incubation flask (see Figures 1 and 2 in Appendix 3) and the soil treated with the test substance by one of the methods described in Section 1.8.2. When organic solvents are used for the application of the test substance, they should be removed from soil by evaporation. Then the soil is thoroughly mixed with a spatula and/or by shaking of the flask. If the study is conducted under paddy field conditions, soil and water should be thoroughly mixed after application of the test substance. Small aliquots (e.g. 1 g) of the treated soils should be analysed for the test substance to check for uniform distribution. For alternative method, see below.

<sup>(1)</sup> Aerobic conditions are dominant in surface soils and even in sub-surface soils as shown in an EU sponsored research project [K. Takagi et al. (1992). Microbial diversity and activity in subsoils: Methods, field site, seasonal variation in subsoil temperatures and oxygen contents. Proc. Internat. Symp. Environm. Aspects Pesticides Microbiol., 270-277, 17-21 August 1992, Sigtuna, Sweden]. Anaerobic conditions may only occur occasionally during flooding of soils after heavy rainfalls or when paddy conditions are established in rice fields.

<sup>(2)</sup> Aerobic studies might be terminated much before 120 days provided that ultimate transformation pathway and ultimate mineralisation are clearly reached at that time. Termination of the test is possible after 120 days, or when at least 90 % of the test substance is transformed, but only if at least 5 %  $CO_2$  is formed.

**▼ B**

The treatment rate should correspond to the highest application rate of a crop protection product recommended in the use instructions and uniform incorporation to an appropriate depth in the field (e.g. top 10 cm layer <sup>(1)</sup> of soil). For example, for chemicals foliarly or soil applied without incorporation, the appropriate depth for computing how much chemical should be added to each flask is 2,5 cm. For soil incorporated chemicals, the appropriate depth is the incorporation depth specified in the use instructions. For general chemicals, the application rate should be estimated based on the most relevant route of entry; for example, when the major route of entry in soil is through sewage sludge, the chemical should be dosed into the sludge at a concentration that reflects the expected sludge concentration and the amount of sludge added to the soil should reflect normal sludge loading to agricultural soils. If this concentration is not high enough to identify major transformation products, incubation of separate soil samples containing higher rates may be helpful, but excessive rates influencing soil microbial functions should be avoided (see Sections 1.5 and 1.8.2). Alternatively, a larger batch (i.e. 1 to 2 kg) of soil can be treated with the test substance, carefully mixed in an appropriate mixing machine and then transferred in small portions of 50 to 200 g into the incubation flasks (for example with the use of sample splitters). Small aliquots (e.g. 1 g) of the treated soil batch should be analysed for the test substance to check for uniform distribution. Such a procedure is preferred since it allows for more uniform distribution of the test substance into the soil.

Also untreated soil samples are incubated under the same conditions (aerobic) as the samples treated with the test substance. These samples are used for biomass measurements during and at the end of the studies.

When the test substance is applied to the soil dissolved in organic solvent(s), soil samples treated with the same amount of solvent(s) are incubated under the same conditions (aerobic) as the samples treated with the test substance. These samples are used for biomass measurements initially, during and at the end of the studies to check for effects of the solvent(s) on microbial biomass.

The flasks containing the treated soil are either attached to the flow-through system described in Figure 1 or closed with the absorption column shown in Figure 2 (see Appendix 3).

<sup>(1)</sup> Calculation of the initial concentration on an area basis using the following equation:

$$C_{\text{soil}} [\text{mg}/\text{kg}_{\text{soil}}] = \frac{A [\text{kg}/\text{ha}] \cdot 10^6 [\text{mg}/\text{kg}]}{l [\text{m}] \cdot 10^4 [\text{m}^2/\text{ha}] \cdot d [\text{kg}_{\text{soil}}/\text{m}^3]}$$

$C_{\text{soil}}$  = Initial concentration in soil [ $\text{mg} \cdot \text{kg}^{-1}$ ]

$A$  = Application rate [ $\text{kg} \cdot \text{ha}^{-1}$ ];  $l$  = thickness of field soil layer [m];  $d$  = dry bulk density of soil [ $\text{kg} \cdot \text{m}^{-3}$ ].

As a rule of thumb, an application rate of  $1 \text{ kg} \cdot \text{ha}^{-1}$  results in a soil concentration of approximately  $1 \text{ mg} \cdot \text{kg}^{-1}$  in a 10 cm layer (assuming a bulk density of  $1 \text{ g} \cdot \text{cm}^{-3}$ ).

**▼B****1.9.3. Sampling and measurement**

Duplicate incubation flasks are removed at appropriate time intervals and the soil samples extracted with appropriate solvents of different polarity and analysed for the test substance and/or transformation products. A well-designed study includes sufficient flasks so that two flasks are sacrificed at each sampling event. Also, absorption solutions or solid absorption materials are removed at various time intervals (7-day intervals during the first month and after one month in 17-day intervals) during and at the end of incubation of each soil sample and analysed for volatile products. Besides a soil sample taken directly after application (0-day sample) at least five additional sampling points should be included. Time intervals should be chosen in such a way that pattern of decline of the test substance and patterns of formation and decline of transformation products can be established (e.g. 0, 1, 3, 7 days; 2, 3 weeks; 1, 2, 3 months, etc.).

When using  $^{14}\text{C}$ -labelled test substance, non-extractable radioactivity will be quantified by combustion and a mass balance will be calculated for each sampling interval.

In the case of anaerobic and paddy incubation, the soil and water phases are analysed together for test substance and transformation products or separated by filtration or centrifugation before extraction and analysis.

**1.9.4. Optional tests**

Aerobic, non-sterile studies at additional temperatures and soil moistures may be useful for the estimation of the influence of temperature and soil moisture on the rates of transformation of a test substance and/or its transformation products in soil.

A further characterisation of non-extractable radioactivity can be attempted using, for example, supercritical fluid extraction.

**2. DATA****2.1. TREATMENT OF RESULTS**

The amounts of test substance, transformation products, volatile substances (in % only), and non-extractable should be given as % of applied initial concentration and, where appropriate, as  $\text{mg} \cdot \text{kg}^{-1}$  soil (based on soil dry weight) for each sampling interval. A mass balance should be given in percentage of the applied initial concentration for each sampling interval. A graphical presentation of the test substance concentrations against time will allow an estimation of its transformation half-life or  $\text{DT}_{50}$ . Major transformation products should be identified and their concentrations should also be plotted against time to show their rates of formation and decline. A major transformation product is any product representing  $\geq 10\%$  of applied dose at any time during the study.

The volatile products trapped give some indication of the volatility potential of a test substance and its transformation products from soil.



**▼B**

More accurate determinations of half-lives or DT<sub>50</sub> values and, if appropriate, DT<sub>75</sub> and DT<sub>90</sub> values should be obtained by applying appropriate kinetic model calculations. The half-life and DT<sub>50</sub> values should be reported together with the description of the model used, the order of kinetics and the determination coefficient (r<sup>2</sup>). First order kinetics is favoured unless r<sup>2</sup> < 0,7. If appropriate, the calculations should also be applied to the major transformation products. Examples of appropriate models are described in references 31 to 35.

In the case of rate studies carried out at various temperatures, the transformation rates should be described as a function of temperature within the experimental temperature range using the Arrhenius relationship of the form:

$$k = A \cdot e^{-B/T} \text{ or } \ln k = \ln A - \frac{B}{T},$$

where ln A and B are regression constants from the intercept and slope, respectively, of a best fit line generated from linearly regressing ln k against 1/T, k is the rate constant at temperature T and T is the temperature in Kelvin. Care should be given to the limited temperature range in which the Arrhenius relationship will be valid in case transformation is governed by microbial action.

## 2.2. EVALUATION AND INTERPRETATION OF RESULTS

Although the studies are carried out in an artificial laboratory system, the results will allow estimation of the rate of transformation of the test substance and also of rate of formation and decline of transformation products under field conditions (36)(37).

A study of the transformation pathway of a test substance provides information on the way in which the applied substance is structurally changed in the soil by chemical and microbial reactions.

## 3. REPORTING

### TEST REPORT

The test report must include:

Test substance:

- common name, chemical name, CAS number, structural formula (indicating position of label(s) when radiolabelled material is used) and relevant physical-chemical properties (see Section 1.5),
- purity (impurities) of test substance,
- radiochemical purity of labelled chemical and specific activity (where appropriate),

Reference substances:

- chemical name and structure of reference substances used for the characterisation and/or identification of transformation product,

Test soils:

- details of collection site,
- date and procedure of soil sampling,

**▼B**

- properties of soils, such as pH, organic carbon content, texture (% sand, % silt, % clay), cation exchange capacity, bulk density, water retention characteristic, and microbial biomass,

- length of soil storage and storage conditions (if stored),

Test conditions:

- dates of the performance of the studies,

- amount of test substance applied,

- solvents used and method of application for the test substance,

- weight of soil treated initially and sampled at each interval for analysis,

- description of the incubation system used,

- air flow rates (for flow-through systems only),

- temperature of experimental set-up,

- soil moisture content during incubation,

- microbial biomass initially, during and at the end of the aerobic studies,

- pH, oxygen concentration and redox potential initially, during and at the end of the anaerobic and paddy studies,

- method(s) of extraction,

- methods for quantification and identification of the test substance and major transformation products in soil and absorption materials,

- number of replicates and number of controls.

Results:

- result of microbial activity determination,

- repeatability and sensitivity of the analytical methods used,

- rates of recovery (% values for a valid study are given in Section 1.7.1),

- tables of results expressed as % of applied initial dose and, where appropriate, as  $\text{mg} \cdot \text{kg}^{-1}$  soil (on a dry weight basis),

- mass balance during and at the end of the studies,

- characterisation of non-extractable (bound) radioactivity or residues in soil,

- quantification of released  $\text{CO}_2$  and other volatile compounds,

- plots of soil concentrations versus time for the test substance and, where appropriate, for major transformation products,

- half-life or  $\text{DT}_{50}$ ,  $\text{DT}_{75}$  and  $\text{DT}_{90}$  for the test substance and, where appropriate, for major transformation products including confidence limits,

**▼B**

- estimation of abiotic degradation rate under sterile conditions,
- an assessment of transformation kinetics for the test substance and, where appropriate, for major transformation products,
- proposed pathways of transformation, where appropriate,
- discussion and interpretation of results,
- raw data (i.e. sample chromatograms, sample calculations of transformation rates and means used to identify transformation products).

4. **REFERENCES**

- (1) US — Environmental Protection Agency, (1982) Pesticide Assessment Guidelines, Subdivision N. Chemistry: Environmental Fate.
- (2) Agriculture Canada, (1987) Environmental Chemistry and Fate. Guidelines for registration of pesticides in Canada.
- (3) European Union (EU), (1995) Commission Directive 95/36/EC of 14 July 1995 amending Council Directive 91/414/EEC concerning the placing of plant protection products on the market. Annex II, Part A and Annex III, Part A: Fate and Behaviour in the Environment.
- (4) Dutch Commission for Registration of Pesticides, (1995) Application for registration of a pesticide. Section G: Behaviour of the product and its metabolites in soil, water and air.
- (5) BBA, (1986) Richtlinie für die amtliche Prüfung von Pflanzenschutzmitteln, Teil IV, 4-1. Verbleib von Pflanzenschutzmitteln im Boden — Abbau, Umwandlung und Metabolismus.
- (6) ISO/DIS 11266-1, (1994) Soil Quality -Guidance on laboratory tests for biodegradation of organic chemicals in soil — Part 1: Aerobic conditions.
- (7) ISO 14239, (1997) Soil Quality — Laboratory incubation systems for measuring the mineralisation of organic chemicals in soil under aerobic conditions.
- (8) SETAC, (1995) Procedures for Assessing the Environmental Fate and Ecotoxicity of Pesticides. Mark R. Lynch, Ed.
- (9) MAFF — Japan 2000 — Draft Guidelines for transformation studies of pesticides in soil — Aerobic metabolism study in soil under paddy field conditions (flooded).
- (10) OECD, (1995) Final Report of the OECD Workshop on Selection of Soils/Sediments. Belgirate, Italy, 18-20 January 1995.
- (11) Guth, J.A., (1980) The study of transformations. In Interactions between Herbicides and the Soil (R.J. Hance, Ed.), Academic Press, p. 123-157.
- (12) DFG: Pesticide Bound Residues in Soil. Wiley — VCH (1998).
- (13) T.R. Roberts: Non-extractable pesticide residue in soils and plants. Pure Appl. Chem. 56, 945-956 (IUPAC 1984).
- (14) OECD Test Guideline 304 A: Inherent Biodegradability in Soil (adopted 12 May 1981).

**▼B**

- (15) ISO 10381-6 (1993) Soil Quality — Sampling — Part 6: Guidance on the collection, handling and storage of soil for the assessment of aerobic microbial processes in the laboratory.
- (16) Appendix V to Directive 67/548/EEC.
- (17) Guth, J.A., (1981) Experimental approaches to studying the fate of pesticides in soil. In *Progress in Pesticide Biochemistry*. D.H. Hutson, T.R. Roberts, Eds. J. Wiley & Sons. Vol 1, 85-114.
- (18) Soil Texture Classification (US and FAO systems): *Weed Science*, 33, Suppl. 1 (1985) and *Soil Sci. Soc. Amer. Proc.* 26:305 (1962).
- (19) *Methods of Soil Analysis (1986) Part 1, Physical and Mineralogical Methods*. A. Klute, Ed.) Agronomy Series No 9, 2nd Edition.
- (20) *Methods of Soil Analysis (1982) Part 2, Chemical and Microbiological Properties*. A.L. Page, R.H. Miller and D.R. Keelney, Eds. Agronomy Series No 9, 2nd Edition.
- (21) ISO Standard Compendium Environment (1994) Soil Quality — General aspects; chemical and physical methods of analysis; biological methods of analysis. First Edition.
- (22) Mückenhausen, E., (1975) *Die Bodenkunde und ihre geologischen, geomorphologischen, mineralogischen und petrologischen Grundlagen*. DLG-Verlag, Frankfurt, Main.
- (23) Scheffer, F., Schachtschabel, P., (1975) *Lehrbuch der Bodenkunde*. F. Enke Verlag, Stuttgart.
- (24) Anderson, J.P.E., Domsch, K.H., (1978) A physiological method for the quantitative measurement of microbial biomass in soils. *Soil Biol. Biochem.* 10, p. 215-221.
- (25) ISO 14240-1 and 2 (1997) Soil Quality — Determination of soil microbial biomass — Part 1: Substrate-induced respiration method. Part 2: fumigation-extraction method.
- (26) Anderson, J.P.E., (1987) Handling and storage of soils for pesticide experiments. In *Pesticide Effects on Soil Microflora*. L. Somerville, M.P. Greaves, Eds. Taylor & Francis, 45-60.
- (27) Kato, Yasuhiro. (1998) Mechanism of pesticide transformation in the environment: Aerobic and bio-transformation of pesticides in aqueous environment. *Proceedings of the 16<sup>th</sup> Symposium on Environmental Science of Pesticide*, p. 105-120.
- (28) Keuken O., Anderson J.P.E., (1996) Influence of storage on biochemical processes in soil. In *Pesticides, Soil Microbiology and Soil Quality*, 59-63 (SETAC-Europe).
- (29) Stenberg B., Johansson M., Pell M., Sjö Dahl-Svensson K., Stenström J., Torstensson L. (1996) Effect of freeze and cold storage of soil on microbial activities and biomass. In *Pesticides, Soil Microbiology and Soil Quality*, 68-69 (SETAC-Europe).
- (30) Gennari, M., Negre, M., Ambrosoli, R., (1987) Effects of ethylene oxide on soil microbial content and some chemical characteristics. *Plant and Soil* 102, p. 197-200.
- (31) Anderson, J.P.E. (1975) Einfluss von Temperatur und Feuchte auf Verdampfung, Abbau und Festlegung von Diallat im Boden. *Z. PflKrankh Pflschutz, Sonderheft VII*, p. 141-146.

**▼B**

- (32) Hamaker, J.W., (1976) The application of mathematical modelling to the soil persistence and accumulation of pesticides. Proc. BCPC Symposium: Persistence of Insecticides and Herbicides, p. 181-199.
- (33) Goring, C.A.I., Laskowski, D.A., Hamaker, J.W., Meikle, R.W., (1975) Principles of pesticide degradation in soil. In 'Environmental Dynamics of Pesticides'. R. Haque and V.H. Freed, Eds., p. 135-172.
- (34) Timme, G., Frehse, H., Laska, V., (1986) Statistical interpretation and graphic representation of the degradational behaviour of pesticide residues. II. Pflanzenschutz — Nachrichten Bayer 39, p. 188-204.
- (35) Timme, G., Frehse, H., (1980) Statistical interpretation and graphic representation of the degradational behaviour of pesticide residues. I. Pflanzenschutz — Nachrichten Bayer 33, p. 47-60.
- (36) Gustafson D.I., Holden L.R., (1990) Non-linear pesticide dissipation in soil; a new model based on spatial variability. Environm. Sci. Technol. 24, p. 1032-1041.
- (37) Hurle K., Walker A., (1980) Persistence and its prediction. In Interactions between Herbicides and the Soil (R.J. Hance, Ed.), Academic Press, p. 83-122.



## Appendix 1

**WATER TENSION, FIELD CAPACITY (FC) AND WATER HOLDING CAPACITY (WHC) <sup>(1)</sup>**

Height of Water Column [cm]	pF <sup>(a)</sup>	bar <sup>(b)</sup>	Remarks
10 <sup>7</sup>	7	10 <sup>4</sup>	Dry Soil
1,6 · 10 <sup>4</sup>	4,2	16	Wilting point
10 <sup>4</sup>	4	10	
10 <sup>3</sup>	3	1	
6 · 10 <sup>2</sup>	2,8	0,6	
3,3 · 10 <sup>2</sup>	2,5	0,33 <sup>(c)</sup>	} Range of Field capacity <sup>(d)</sup>
10 <sup>2</sup>	2	0,1	
60	1,8	0,06	
33	1,5	0,033	
10	1	0,01	WHC (approximation)
1	0	0,001	Water saturated soil

<sup>(a)</sup> pF = log of cm water column.

<sup>(b)</sup> 1 bar = 10<sup>5</sup> Pa.

<sup>(c)</sup> Corresponds to an approximate water content of 10 % in sand, 35 % in loam and 45 % in clay.

<sup>(d)</sup> Field capacity is not constant but varies with soil type between pF 1,5 and 2,5.

*Water tension* is measured in cm water column or in bar. Due to the large range of suction tension it is expressed simply as pF value which is equivalent to the logarithm of cm water column.

*Field capacity* is defined as the amount of water which can be stored against gravity by a natural soil two days after a longer raining period or after sufficient irrigation. It is determined in undisturbed soil in situ in the field. The measurement is thus not applicable to disturbed laboratory soil samples. FC values determined in disturbed soils may show great systematic variances.

*Water holding capacity* (WHC) is determined in the laboratory with undisturbed and disturbed soil by saturating a soil column with water by capillary transport. It is particularly useful for disturbed soils and can be up to 30 % greater than field capacity (1). It is also experimentally easier to determine than reliable FC-values.

## Notes

<sup>(1)</sup> Mückenhausen, E., (1975) Die Bodenkunde und ihre geologischen, geomorphologischen, mineralogischen und petrologischen Grundlagen. DLG-Verlag, Frankfurt, Main.

**▼B***Appendix 2***SOIL MOISTURE CONTENTS (g water per 100 g dry soil) OF VARIOUS SOIL TYPES FROM VARIOUS COUNTRIES**

Soil type	Country	Soil moisture content at		
		WHC <sup>(1)</sup>	pF = 1,8	pF = 2,5
Sand	Germany	28,7	8,8	3,9
Loamy sand	Germany	50,4	17,9	12,1
Loamy sand	Switzerland	44,0	35,3	9,2
Silt loam	Switzerland	72,8	56,6	28,4
Clay loam	Brazil	69,7	38,4	27,3
Clay loam	Japan	74,4	57,8	31,4
Sandy loam	Japan	82,4	59,2	36,0
Silt loam	USA	47,2	33,2	18,8
Sandy loam	USA	40,4	25,2	13,3

<sup>(1)</sup> Water holding capacity.

▼B

## Appendix 3

Figure 1

Example of a flow-through apparatus to study transformation of chemicals in soil <sup>(1)</sup> <sup>(2)</sup>

- |  |   |  |
|--|---|--|
| 1: needle valve  | 4: soil metabolism flask (water-logged only for anaerobic and paddy conditions) | 7, 8: sodium hydroxide trap for CO <sub>2</sub> & other acidic volatiles |
| 2: gas washing bottle containing water                       | 5: ethylene glycol trap for organic volatile compounds                          | 9: flow meter.   |
| 3: ultramembrane (sterile conditions only), pore size 0.2 μm | 6: sulphuric acid trap for alkaline volatile compounds                          |  |

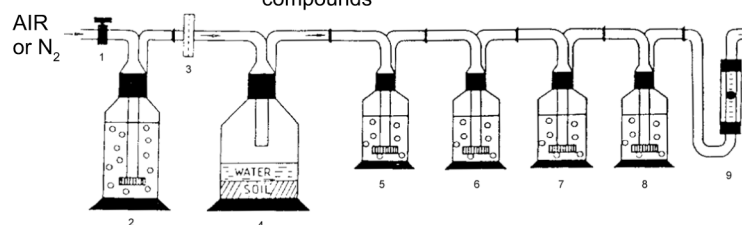
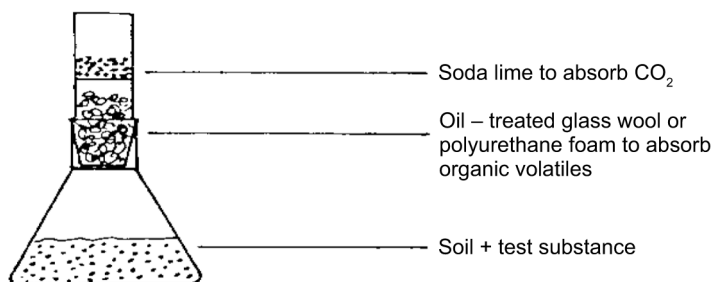


Figure 2

Example of a biometer-type flask for studying the transformation of chemicals in soil <sup>(3)</sup>



<sup>(1)</sup> Guth, J.A., (1980) The study of transformations. In Interactions between Herbicides and the Soil (R.J. Hance, Ed.), Academic Press, p. 123-157.

<sup>(2)</sup> Guth, J.A., (1981) Experimental approaches to studying the fate of pesticides in soil. In Progress in Pesticide Biochemistry. D.H. Hutson, T.R. Roberts, Eds. J. Wiley & Sons. Vol 1, p. 85-114.

<sup>(3)</sup> Anderson, J.P.E., (1975) Einfluss von Temperatur und Feuchte auf Verdampfung, Abbau und Festlegung von Diallyl im Boden. Z. PflKrankh Pflschutz, Sonderheft VII, p. 141-146.



**▼B****C.24. AEROBIC AND ANAEROBIC TRANSFORMATION IN AQUATIC SEDIMENT SYSTEMS****1. METHOD**

This test method is a replicate of the OECD TG 308 (2002).

**1.1. INTRODUCTION**

Chemicals can enter shallow or deep surface waters by such routes as direct application, spray drift, run-off, drainage, waste disposal, industrial, domestic or agricultural effluent and atmospheric deposition. This testing method describes a laboratory method to assess aerobic and anaerobic transformation of organic chemicals in aquatic sediment systems. It is based on existing Guidelines (1)-(2)(3)(4)(5)(6). An OECD Workshop on Soil/Sediment Selection, held in Belgirate, Italy in 1995 (7) agreed, in particular, on the number and type of sediments for use in this test. It also made recommendations relating to collection, handling and storage of sediment samples, based on the ISO Guidance (8). Such studies are required for chemicals which are directly applied to water or which are likely to reach the aqueous environment by the routes described above.

The conditions in natural aquatic sediment systems are often aerobic in the upper water phase. The surface layer of sediment can be either aerobic or anaerobic, whereas the deeper sediment is usually anaerobic. To encompass all of these possibilities both aerobic and anaerobic tests are described in this document. The aerobic test simulates an aerobic water column over an aerobic sediment layer that is underlain with an anaerobic gradient. The anaerobic test simulates a completely anaerobic water-sediment system. If circumstances indicate that it is necessary to deviate significantly from these recommendations, for example by using intact sediment cores or sediments that may have been exposed to the test substance, other methods are available for this purpose (9).

**1.2. DEFINITIONS**

Standard International (SI) units should be used in any case.

**Test substance:** any substance, whether the parent or relevant transformation products.

**Transformation products:** all substances resulting from biotic and abiotic transformation reactions of the test substance including CO<sub>2</sub> and bound residues.

**Bound residues:** 'bound residues' represent compounds in soil, plant or animal that persist in the matrix in the form of the parent substance or its metabolite(s) after extractions. The extraction method must not substantially change the compounds themselves or the structure of the matrix. The nature of the bond can be clarified in part by matrix-altering extraction methods and sophisticated analytical techniques. To date, for example, covalent ionic and sorptive bonds, as well as entrapments, have been identified in this way. In general, the formation of bound residues reduces the bioaccessibility and the bioavailability significantly (10) (modified from IUPAC 1984 (11)).

**▼ B**

**Aerobic transformation:** (oxidising): reactions occurring in the presence of molecular oxygen (12).

**Anaerobic transformation:** (reducing): reactions occurring under exclusion of molecular oxygen (12).

**Natural waters:** are surface waters obtained from ponds, rivers, streams, etc.

**Sediment:** is a mixture of mineral and organic chemical constituents, the latter containing compounds of high carbon and nitrogen content and of high molecular masses. It is deposited by natural water and forms an interface with that water.

**Mineralisation:** is the complete degradation of an organic compound to CO<sub>2</sub>, H<sub>2</sub>O under aerobic conditions, and CH<sub>4</sub>, CO<sub>2</sub> and H<sub>2</sub>O under anaerobic conditions. In the context of this test method, when radiolabelled compound is used, mineralisation means extensive degradation of a molecule during which a labelled carbon atom is oxidised or reduced quantitatively with release of the appropriate amount of <sup>14</sup>CO<sub>2</sub> or <sup>14</sup>CH<sub>4</sub>, respectively.

**Half-life, t<sub>0,5</sub>,** is the time taken for 50 % transformation of a test substance when the transformation can be described by first-order kinetics; it is independent of the initial concentration.

**DT<sub>50</sub> (Disappearance Time 50):** is the time within which the initial concentration of the test substance is reduced by 50 %.

**DT<sub>75</sub> (DISAPPEARANCE TIME 75):** IS THE TIME WITHIN WHICH THE INITIAL CONCENTRATION OF THE TEST SUBSTANCE IS REDUCED BY 75 %.

**DT<sub>90</sub> (Disappearance Time 90):** is the time within which the initial concentration of the test substance is reduced by 90 %.

### 1.3. REFERENCE SUBSTANCES

Reference substances should be used for the identification and quantification of transformation products by spectroscopic and chromatographic methods.

### 1.4. INFORMATION ON THE TEST SUBSTANCE

Non-labelled or isotope-labelled test substance can be used to measure the rate of transformation although labelled material is preferred. Labelled material is required for studying the pathway of transformation and for establishing a mass balance. <sup>14</sup>C-labelling is recommended, but the use of other isotopes, such as <sup>13</sup>C, <sup>15</sup>N, <sup>3</sup>H, <sup>32</sup>P, may also be useful. As far as possible, the label should be positioned in the most stable part(s) of the molecule<sup>(1)</sup>. The chemical and/or radiochemical purity of the test substance should be at least 95 %.

Before carrying out a test, the following information about the test substance should be available:

- (a) solubility in water (Method A.6);
- (b) solubility in organic solvents;
- (c) vapour pressure (Method A.4) and Henry's Law constant;

<sup>(1)</sup> For example, if the substance contains one ring, labelling on this ring is required; if the test substance contains two or more rings, separate studies may be needed to evaluate the fate of each labelled ring and to obtain suitable information on formation of transformation products.

**▼ B**

- (d) n-octanol/water partition coefficient (Method A.8);
- (e) adsorption coefficient ( $K_d$ ,  $K_f$  or  $K_{oc}$ , where appropriate) (Method C.18);
- (f) hydrolysis (Method C.7);
- (g) dissociation constant ( $pK_a$ ) (OECD Guideline 112) (13);
- (h) chemical structure of the test substance and position of the isotope-label(s), if applicable.

*Note:* the temperature at which these measurements were made should be reported.

Other useful information may include data on toxicity of the test substance to microorganisms, data on ready and/or inherent biodegradability, and data on aerobic and anaerobic transformation in soil.

Analytical methods (including extraction and clean-up methods) for identification and quantification of the test substance and its transformation products in water and in sediment should be available (see Section 1.7.2).

#### 1.5. PRINCIPLE OF THE TEST METHOD

The method described in this test employs an aerobic and an anaerobic aquatic sediment (see Appendix 1) system which allows:

the measurement of the transformation rate of the test substance in a water-sediment system,

the measurement of the transformation rate of the test substance in the sediment,

the measurement of the mineralisation rate of the test substance and/or its transformation products (when  $^{14}\text{C}$ -labelled test substance is used),

the identification and quantification of transformation products in water and sediment phases including mass balance (when labelled test substance is used),

the measurement of the distribution of the test substance and its transformation products between the two phases during a period of incubation in the dark (to avoid, for example, algal blooms) at constant temperature. Half-lives,  $DT_{50}$ ,  $DT_{75}$  and  $DT_{90}$  values are determined where the data warrant, but should not be extrapolated far past the experimental period (see Section 1.2).

At least two sediments and their associated waters are required for both the aerobic and the anaerobic studies respectively (7). However, there may be cases where more than two aquatic sediments should be used, for example, for a chemical that may be present in freshwater and/or marine environments.

**▼ B**

## 1.6. APPLICABILITY OF THE TEST

The method is generally applicable to chemical substances (unlabelled or labelled) for which an analytical method with sufficient accuracy and sensitivity is available. It is applicable to slightly volatile, non-volatile, water-soluble or poorly water-soluble compounds. The test should not be applied to chemicals which are highly volatile from water (e.g. fumigants, organic solvents) and thus cannot be kept in water and/or sediment under the experimental conditions of this test.

The method has been applied so far to study the transformation of chemicals in fresh waters and sediments, but in principle can also be applied to estuarine/marine systems. It is not suitable to simulate conditions in flowing water (e.g. rivers) or the open sea.

## 1.7. QUALITY CRITERIA

1.7.1. **Recovery**

Extraction and analysis of, at least, duplicate water and sediment samples immediately after the addition of the test substance gives a first indication of the repeatability of the analytical method and of the uniformity of the application procedure for the test substance. Recoveries for later stages of the experiments are given by the respective mass balances (when labelled material is used). Recoveries should range from 90 % to 110 % for labelled chemicals (6) and from 70 % to 110 % for non-labelled chemicals.

1.7.2. **Repeatability and sensitivity of analytical method**

Repeatability of the analytical method (excluding the initial extraction efficiency) to quantify test substance and transformation products can be checked by duplicate analysis of the same extract of the water or the sediment samples which were incubated sufficiently long enough for formation of transformation products.

The limit of detection (LOD) of the analytical method for the test substance and for the transformation products should be at least 0,01 mg·kg<sup>-1</sup> in water or sediment (as test substance) or 1 % of the initial amount applied to a test system whichever is lower. The limit of quantification (LOQ) should also be specified.

1.7.3. **Accuracy of transformation data**

Regression analysis of the concentrations of the test substance as a function of time gives the appropriate information on the accuracy of the transformation curve and allows the calculation of the confidence limits for half-lives (if pseudo first-order kinetics apply) or DT<sub>50</sub> values and, if appropriate, DT<sub>75</sub> and DT<sub>90</sub> values.

## 1.8. DESCRIPTION OF THE METHOD

1.8.1. **Test system and apparatus**

The study should be performed in glass containers (e.g. bottles, centrifuge tubes), unless preliminary information (such as n-octanol-water partition coefficient, sorption data, etc.) indicates that the test substance may adhere to glass, in which case an alternative material (such as Teflon) may have to be considered. Where the test substance is known to adhere to glass, it may be possible to alleviate this problem using one or more of the following methods:

**▼B**

- determine the mass of test substance and transformation products sorbed to glass,
- ensure a solvent wash of all glassware at the end of the test,
- use of formulated products (see also Section 1.9.2),
- use an increased amount of co-solvent for addition of test substance to the system; if a co-solvent is used it should be a co-solvent that does not solvolyse the test substance.

Examples of typical test apparatus, i.e. gas flow-through and biometer-type systems, are shown in Appendices 2 and 3, respectively (14). Other useful incubation systems are described in reference 15. The design of the experimental apparatus should permit the exchange of air or nitrogen and the trapping of volatile products. The dimensions of the apparatus must be such that the requirements of the test are complied with (see Section 1.9.1). Ventilation may be provided by either gentle bubbling or by passing air or nitrogen over the water surface. In the latter case gentle stirring of the water from above may be advisable for better distribution of the oxygen or nitrogen in the water. CO<sub>2</sub>-free air should not be used as this can result in increases in the pH of the water. In either case, disturbance of the sediment is undesirable and should be avoided as far as possible. Slightly volatile chemicals should be tested in a biometer-type system with gentle stirring of the water surface. Closed vessels with a headspace of either atmospheric air or nitrogen and internal vials for the trapping of volatile products can also be used (16). Regular exchange of the headspace gas is required in the aerobic test in order to compensate for the oxygen consumption by the biomass.

Suitable traps for collecting volatile transformation products include but are not restricted to 1 mol·dm<sup>-3</sup> solutions of potassium hydroxide or sodium hydroxide for carbon dioxide <sup>(1)</sup> and ethylene glycol, ethanolamine or 2 % paraffin in xylene for organic compounds. Volatiles formed under anaerobic conditions, such as methane, can be collected, for example, by molecular sieves. Such volatiles can be combusted, for example, to CO<sub>2</sub> by passing the gas through a quartz tube filled with CuO at a temperature of 900 °C and trapping the CO<sub>2</sub> formed in an absorber with alkali (17).

Laboratory instrumentation for chemical analysis of test substance and transformation products is required (e.g. gas liquid chromatography (GLC), high performance liquid chromatography (HPLC), thin-layer chromatography (TLC), mass spectroscopy (MS), gas chromatography-mass spectroscopy (GC-MS), liquid chromatography-mass spectrometry (LC-MS), nuclear magnetic resonance (NMR), etc.), including detection systems for radiolabelled or non-labelled chemicals as appropriate. When radiolabelled material is used a liquid scintillation counter and combustion oxidiser (for the combustion of sediment samples prior to analysis of radioactivity) will also be required.

Other standard laboratory equipment for physical-chemical and biological determinations (see Section Table 1, Section 1.8.2.2), glassware, chemicals and reagents are required as appropriate.

<sup>(1)</sup> As these alkaline absorption solutions also absorb the carbon dioxide from the ventilation air and that formed by respiration in aerobic experiments, they have to be exchanged in regular intervals to avoid their saturation and thus loss of their absorption capacity.

**▼B****1.8.2. Selection and number of aquatic sediments**

The sampling sites should be selected in accordance with the purpose of the test in any given situation. In selecting sampling sites, the history of possible agricultural, industrial or domestic inputs to the catchment and the waters upstream must be considered. Sediments should not be used if they have been contaminated with the test substance or its structural analogues within the previous four years.

**1.8.2.1. Sediment selection**

Two sediments are normally used for the aerobic studies (7). The two sediments selected should differ with respect to organic carbon content and texture. One sediment should have a high organic carbon content (2,5-7,5 %) and a fine texture, the other sediment should have a low organic carbon content (0,5-2,5 %) and a coarse texture. The difference between the organic carbon contents should normally be at least 2 %. 'Fine texture' is defined as a [clay + silt] <sup>(1)</sup> content of > 50 % and 'coarse texture' is defined as a [clay + silt] content of < 50 %. The difference in [clay + silt] content for the two sediments should normally be at least 20 %. In cases, where a chemical may also reach marine waters, at least one of the water-sediment systems should be of marine origin.

For the strictly anaerobic study, two sediments (including their associated waters) should be sampled from the anaerobic zones of surface water bodies (7). Both the sediment and the water phases should be handled and transported carefully under exclusion of oxygen.

Other parameters may be important in the selection of sediments and should be considered on a case-by-case basis. For example, the pH range of sediments would be important for testing chemicals for which transformation and/or sorption may be pH-dependent. pH-dependency of sorption might be reflected by the  $pK_a$  of the test substance.

**1.8.2.2. Characterisation of water-sediment samples**

Key parameters that must be measured and reported (with reference to the method used) for both water and sediment, and the stage of the test at which those parameters are to be determined are summarised in the Table hereafter. For information, methods for determination of these parameters are given in references (18)(19)(20)(21).

In addition, other parameters may need to be measured and reported on a case by case basis (e.g. for freshwater: particles, alkalinity, hardness, conductivity,  $NO_3/PO_4$  (ratio and individual values); for sediments: cation exchange capacity, water holding capacity, carbonate, total nitrogen and phosphorus; and for marine systems: salinity). Analysis of sediments and water for nitrate, sulfate, bioavailable iron, and possibly other electron acceptors may be also useful in assessing redox conditions, especially in relation to anaerobic transformation.

<sup>(1)</sup> (Clay + silt) is the mineral fraction of the sediment with particle size of < 50  $\mu m$ .


**Measurement of parameters for characterisation of water-sediment samples (7)(22)(23)**

Parameter	Stage of test procedure					
	field sampling	post-handling	start of acclimation	start of test	during test	end of test
<b>Water</b>						
Origin/source	x					
Temperature	x					
pH	x		x	x	x	x
TOC			x	x		x
O <sub>2</sub> concentration*	x		x	x	x	x
Redox Potential*			x	x	x	x
<b>Sediment</b>						
Origin/source	x					
Depth of layer	x					
pH		x	x	x	x	x
Particle size distribution		x				
TOC		x	x	x		x
Microbial biomass (*)		x		x		x
Redox potential (**)	Observation (colour/smell)		x	x	x	x

(\*) Microbial respiration rate method (26), fumigation method (27) or plate count measurements (e.g. bacteria, actinomycetes, fungi and total colonies) for aerobic studies; methanogenesis rate for anaerobic studies.

(\*\*) Recent research results have shown that measurements of water oxygen concentrations and of redox potentials have neither a mechanistic nor a predictive value as far as growth and development of microbial populations in surface waters are concerned (24)(25). Determination of the biochemical oxygen demand (BOD, at field sampling, start and end of test) and of concentrations of micro/macro nutrients Ca, Mg and Mn (at start and end of test) in water and the measurement of total N and total P in sediments (at field sampling and end of test) may be better tools to interpret and evaluate aerobic biotransformation rates and routes.

**1.8.3. Collection, handling and storage**
**1.8.3.1. Collection**

The draft ISO guidance on sampling of bottom sediment (8) should be used for sampling of sediment. Sediment samples should be taken from the entire 5 to 10 cm upper layer of the sediment. Associated water should be collected from the same site or location and at the same time as the sediment. For the anaerobic study, sediment and associated water should be sampled and transported under exclusion of oxygen (28)(see Section 1.8.2.1). Some sampling devices are described in the literature (8)(23).

**▼ B**1.8.3.2. *Handling*

The sediment is separated from the water by filtration and the sediment wet-sieved to a 2 mm-sieve using excess location water that is then discarded. Then known amounts of sediments and water are mixed at the desired ratio (see Section 1.9.1) in incubation flasks and prepared for the acclimation period (see Section 1.8.4). For the anaerobic study, all handling steps have to be done under exclusion of oxygen (29)(30)(31)(32)(33).

1.8.3.3. *Storage*

Use of freshly sampled sediment and water is strongly recommended, but if storage is necessary, sediment and water should be sieved as described above and stored together, water-logged (6-10 cm water layer), in the dark, at  $4 \pm 2^\circ\text{C}$ <sup>4</sup> for a maximum of four weeks (7)(8)(23). Samples to be used for aerobic studies should be stored with free access of air (e.g. in open containers), whereas those for anaerobic studies under exclusion of oxygen. Freezing of sediment and water and drying-out of the sediment must not occur during transportation and storage.

1.8.4. **Preparation of the sediment/water samples for the test**

A period of acclimation should take place prior to adding the test substance, with each sediment/water sample being placed in the incubation vessel to be used in the main test, and the acclimation to be carried out under exactly the same conditions as the test incubation (see Section 1.9.1). The acclimation period is the time needed to reach reasonable stability of the system, as reflected by pH, oxygen concentration in water, redox potential of the sediment and water, and macroscopic separation of phases. The period of acclimation should normally last between one week and two weeks and should not exceed four weeks. Results of determinations performed during this period should be reported.

## 1.9. PERFORMANCE OF THE TEST

1.9.1. **Test conditions**

The test should be performed in the incubation apparatus (see Section 1.8.1) with a water sediment volume ratio between 3:1 and 4:1, and a sediment layer of 2,5 cm ( $\pm 0,5$  cm).<sup>(1)</sup> A minimum amount of 50 g of sediment (dry weight basis) per incubation vessel is recommended.

The test should be performed in the dark at a constant temperature in the range of 10 to 30 °C. A temperature of  $(20 \pm 2)^\circ\text{C}$  is appropriate. Where appropriate, an additional lower temperature (e.g. 10 °C) may be considered on a case-by-case basis, depending on the information required from the test. Incubation temperature should be monitored and reported.

<sup>(1)</sup> Recent studies have shown that storage at 4 °C can lead to a decrease of the organic carbon content of the sediment which may possibly result in a decrease of microbial activity (34).



**▼B****1.9.2. Treatment and application of test substance**

One test concentration of chemical is used <sup>(1)</sup>. For crop protection chemicals applied directly to water bodies, the maximum dosage on the label should be taken as, the maximum application rate calculated on the basis of the surface area of the water in the test vessel. In all other cases, the concentration to be used should be based on predictions from environmental emissions. Care must be taken to ensure that an adequate concentration of test substance is applied in order to characterise the route of transformation and the formation and decline of transformation products. It may be necessary to apply higher doses (e.g. 10 times) in situations where test substance concentrations are close to limits of detection at the start of the study and/or where major transformation products could not readily be detected when present at 10 % of the test substance application rate. However, if higher test concentrations are used they should not have a significant adverse effect on the microbial activity of the water-sediment system. In order to achieve a constant concentration of test substance in vessels of differing dimensions an adjustment to the quantity of the material applied may be considered appropriate, based on the depth of the water column in the vessel in relation to the depth of water in the field (which is assumed to be 100 cm, but other depths can be used). See Appendix 4 for an example calculation.

Ideally the test substance should be applied as an aqueous solution into the water phase of the test system. If unavoidable, the use of low amounts of water miscible solvents (such as acetone, ethanol) is permitted for application and distribution of the test substance, but this should not exceed 1 % v/v and should not have adverse effects on microbial activity of the test system. Care should be exercised in generating the aqueous solution of the test substance — use of generator columns and pre-mixing may be appropriate to ensure complete homogeneity. Following addition of the aqueous solution to the test system, gentle mixing of the water phase is recommended, disturbing the sediment as little as possible.

The use of formulated products is not routinely recommended as the formulation ingredients may affect the distribution of the test substance and/or transformation products between water and sediment phases. However, for poorly water-soluble test substances, the use of formulated material may be an appropriate alternative.

The number of incubation vessels depends on the number of sampling times (see Section 1.9.3). A sufficient number of test systems should be included so that two systems may be sacrificed at each sampling time. Where control units of each aquatic sediment system are employed, they should not be treated with the test substance. The control units can be used to determine the microbial biomass of the sediment and the total organic carbon of the water and sediment at the termination of the study. Two of the control units (i.e. one control unit for each aquatic sediment) can be used to monitor the required parameters in the sediment and water during the acclimation period (see Table in Section 1.8.2.2). Two additional control units have to be included in case the test substance is applied by means of a solvent to measure adverse effects on the microbial activity of the test system.

<sup>(1)</sup> Test with a second concentration can be useful for chemicals that reach surface waters by different entry routes resulting in significantly different concentrations, as long as the lower concentration can be analysed with sufficient accuracy.

**▼ B****1.9.3. Test duration and sampling**

The duration of the experiment should normally not exceed 100 days (6), and should continue until the degradation pathway and water/sediment distribution pattern are established or when 90 % of the test substance has dissipated by transformation and/or volatilisation. The number of sampling times should be at least six (including zero time), with an optional preliminary study (see Section 1.9.4) being used to establish an appropriate sampling regime and the duration of the test, unless sufficient data is available on the test substance from previous studies. For hydrophobic test substances, additional sampling points during the initial period of the study may be necessary in order to determine the rate of distribution between water and sediment phases.

At appropriate sampling times, whole incubation vessels (in replicate) are removed for analysis. Sediment and overlying water are analysed separately<sup>(1)</sup>. The surface water should be carefully removed with minimum disturbance of the sediment. The extraction and characterisation of the test substance and transformation products should follow appropriate analytical procedures. Care should be taken to remove material that may have adsorbed to the incubation vessel or to interconnecting tubing used to trap volatiles.

**1.9.4. Optional preliminary test**

If duration and sampling regime cannot be estimated from other relevant studies on the test substance, an optional preliminary test may be considered appropriate, which should be performed using the same test conditions proposed for the definitive study. Relevant experimental conditions and results from the preliminary test, if performed, should be briefly reported.

**1.9.5. Measurements and analysis**

Concentration of the test substance and the transformation products at every sampling time in water and sediment should be measured and reported (as a concentration and as percentage of applied). In general, transformation products detected at  $\geq 10$  % of the applied radioactivity in the total water-sediment system at any sampling time should be identified unless reasonably justified otherwise. Transformation products for which concentrations are continuously increasing during the study should also be considered for identification, even if their concentrations do not exceed the limits given above, as this may indicate persistence. The latter should be considered on a case by case basis, with justifications being provided in the report.

Results from gases/volatiles trapping systems (CO<sub>2</sub> and others, i.e. volatile organic compounds) should be reported at each sampling time. Mineralisation rates should be reported. Non-extractable (bound) residues in sediment are to be reported at each sampling point.

<sup>(1)</sup> In cases where rapid re-oxidation of anaerobic transformation products may readily occur, anaerobic conditions should be maintained during sampling and analysis.

**▼B****2. DATA****2.1. TREATMENT OF RESULTS**

Total mass balance or recovery (see Section 1.7.1) of added radioactivity is to be calculated at every sampling time. Results should be reported as a percentage of added radioactivity. Distribution of radioactivity between water and sediment should be reported as concentrations and percentages, at every sampling time.

Half-life,  $DT_{50}$  and, if appropriate,  $DT_{75}$  and  $DT_{90}$  of the test substance should be calculated along with their confidence limits (see Section 1.7.3). Information on the rate of dissipation of the test substance in the water and sediment can be obtained through the use of appropriate evaluation tools. These can range from application of pseudo-first order kinetics, empirical curve-fitting techniques which apply graphical or numerical solutions and more complex assessments using, for example, single- or multi-compartment models. Further details can be obtained from the relevant published literature (35)(36)(37).

All approaches have their strengths and weaknesses and vary considerably in complexity. An assumption of first-order kinetics may be an oversimplification of the degradation and distribution processes, but when possible gives a term (the rate constant or half-life) which is easily understood and of value in simulation modelling and calculations of predicted environmental concentrations. Empirical approaches or linear transformations can result in better fits of curves to data and therefore allow better estimation of half-lives,  $DT_{50}$  and, if appropriate,  $DT_{75}$  and  $DT_{90}$  values. The use of the derived constants, however, is limited. Compartment models can generate a number of useful constants of value in risk assessment that describe the rate of degradation in different compartments and the distribution of the chemical. They should also be used for estimation of rate constants for the formation and degradation of major transformation products. In all cases, the method chosen must be justified and the experimenter should demonstrate graphically and/or statistically the goodness of fit.

**3. REPORTING****3.1. TEST REPORT**

The report must include the following information:

Test substance:

— common name, chemical name, CAS number, structural formula (indicating position of the label(s) when radiolabelled material is used) and relevant physical-chemical properties,

— purity (impurities) of test substance,

— radiochemical purity of labelled chemical and molar activity (where appropriate).

**▼ B**

## Reference substances:

- chemical name and structure of reference substances used for the characterisation and/or identification of transformation products.

## Test sediments and waters:

- location and description of aquatic sediment sampling site(s) including, if possible, contamination history,
- all information relating to the collection, storage (if any) and acclimation of water-sediment systems,
- characteristics of the water-sediment samples as listed in Table in section 1.8.2.2.

## Test conditions:

- test system used (e.g. flow-through, biometer, way of ventilation, method of stirring, water volume, mass of sediment, thickness of both water and sediment layer, dimension of test vessels, etc.),
- application of test substance to test system: test concentration used, number of replicates and controls mode of application of test substance (e.g. use of solvent if any), etc.,
- incubation temperature,
- sampling times,
- extraction methods and efficiencies as well as analytical methods and detection limits,
- methods for characterisation/identification of transformation products,
- deviations from the test protocol or test conditions during the study.

## Results:

- raw data figures of representative analyses (all raw data have to be stored in the GLP-archive),
- repeatability and sensitivity of the analytical methods used,
- rates of recovery (% values for a valid study are given in section 1.7.1),
- tables of results expressed as % of the applied dose and in  $\text{mg} \cdot \text{kg}^{-1}$  in water, sediment and total system (% only) for the test substance and, if appropriate, for transformation products and non-extractable radioactivity,
- mass balance during and at the end of the studies,
- a graphical representation of the transformation in the water and sediment fractions and in total system (including mineralisation),
- mineralisation rates,

**▼B**

- half-life, DT<sub>50</sub> and, if appropriate, DT<sub>75</sub> and DT<sub>90</sub> values for the test substance and, where appropriate, for major transformation products including confidence limits in water, sediment and in total system,
- an assessment of the transformation kinetics of the test substance and, where appropriate, the major transformation products,
- a proposed pathway of transformation, where appropriate,
- discussion of results.

4. **REFERENCES**

- (1) BBA-Guidelines for the examination of plant protectors in the registration process., (1990) Part IV, Section 5-1: Degradability and fate of plant protectors in the water/sediment system. Germany.
- (2) Commission for registration of pesticides: Application for registration of a pesticide., (1991) Part G. Behaviour of the product and its metabolites in soil, water and air, Section G.2.1 (a). The Netherlands.
- (3) MAFF Pesticides Safety Directorate., (1992) Preliminary guideline for the conduct of biodegradability tests on pesticides in natural sediment/water systems. Ref No SC 9046. United-Kingdom.
- (4) Agriculture Canada: Environmental chemistry and fate., (1987) Guidelines for registration of pesticides in Canada. Aquatic (Laboratory) — Anaerobic and aerobic. Canada. p. 35-37.
- (5) US-EPA: Pesticide assessment guidelines, Subdivision N. Chemistry: Environmental fate (1982) Section 162-3, Anaerobic aquatic metabolism.
- (6) SETAC-Europe publication., (1995) Procedures for assessing the environmental fate and ecotoxicity of pesticides. Ed. Dr Mark R. Lynch. SETAC-Europe, Brussels.
- (7) OECD Test Guidelines Programme., (1995) Final Report of the OECD Workshop on Selection of Soils/sediments, Belgirate, Italy, 18-20 January 1995.
- (8) ISO/DIS 5667-12., (1994) Water quality — Sampling — Part 12: Guidance on sampling of bottom sediments.
- (9) US-EPA (1998a) Sediment/water microcosm biodegradation test. Harmonised Test Guidelines (OPPTS 835.3180). EPA 712-C-98-080.
- (10) DFG: Pesticide Bound Residues in Soil. Wiley-VCH (1998).
- (11) T.R. Roberts: Non-extractable pesticide residues in soils and plants. Pure Appl. Chem. 56, 945-956 (IUPAC 1984).
- (12) OECD Test Guideline 304A: Inherent Biodegradability in Soil (adopted 12 May 1981).
- (13) OECD (1993): Guidelines for Testing of Chemicals. Paris. OECD (1994-2000): Addenda 6-11 to Guidelines for the Testing of Chemicals.
- (14) Scholz, K., Fritz R., Anderson C. and Spiteller M. (1988) Degradation of pesticides in an aquatic model ecosystem. BCPC — Pests and Diseases, 3B-4, p. 149-158.

**▼B**

- (15) Guth, J.A., (1981) Experimental approaches to studying the fate of pesticides in soil. In Progress in Pesticide Biochemistry (D.H. Hutson, T.R. Roberts, Eds.), Vol. 1, 85-114. J. Wiley & Sons.
- (16) Madsen, T., Kristensen, P. (1997) Effects of bacterial inoculation and non-ionic surfactants on degradation of polycyclic aromatic hydrocarbons in soil. Environ. Toxicol. Chem. 16, p. 631-637.
- (17) Steber, J., Wierich, P. (1987) The anaerobic degradation of detergent range fatty alcohol ethoxylates. Studies with <sup>14</sup>C-labelled model surfactants. Water Research 21, p. 661-667.
- (18) Black, C.A. (1965). Methods of Soil Analysis. Agronomy Monograph No 9. American Society of Agronomy, Madison.
- (19) APHA (1989) Standard Methods for Examination of Water and Wastewater (17<sup>th</sup> edition). American Public Health Association, American Water Works Association and Water Pollution Control Federation, Washington D.C.
- (20) Rowell, D.L. (1994) Soil Science Methods and Applications. Longman.
- (21) Light, T.S., (1972). Standard solution for redox potential measurements. Anal. Chemistry 44, p. 1038-1039.
- (22) SETAC-Europe publication (1991) Guidance document on testing procedures for pesticides in freshwater mesocosms. From the Workshop 'A Meeting of Experts on Guidelines for Static Field Mesocosms Tests', 3-4 July 1991.
- (23) SETAC-Europe publication. (1993) Guidance document on sediment toxicity tests and bioassays for freshwater and marine environments. From the Workshop on Sediment Toxicity Assessment (WOSTA), 8-10 November 1993. Eds.: I.R. Hill, P. Matthiessen and F. Heimbach.
- (24) Vink, J.P.M., van der Zee, S.E.A.T.M. (1997) Pesticide biotransformation in surface waters: multivariate analyses of environmental factors at field sites. Water Research 31, p. 2858-2868.
- (25) Vink, J.P.M., Schraa, G., van der Zee, S.E.A.T.M. (1999) Nutrient effects on microbial transformation of pesticides in nitrifying waters. Environ. Toxicol, p. 329-338.
- (26) Anderson, T.H., Domsch, K.H. (1985) Maintenance carbon requirements of actively-metabolising microbial populations under *in-situ* conditions. Soil Biol. Biochem. 17, p. 197-203.
- (27) ISO-14240-2., (1997) Soil quality — Determination of soil microbial biomass — Part 2: Fumigation-extraction method.
- (28) Beelen, P. Van and F. Van Keulen., (1990), The Kinetics of the Degradation of Chloroform and Benzene in Anaerobic Sediment from the River Rhine. Hydrobiol. Bull. 24 (1), p. 13-21.
- (29) Shelton, D.R. and Tiedje, J.M. (1984) General method for determining anaerobic biodegradation potential. App. Environ. Microbiol. 47, p. 850-857.
- (30) Birch, R.R., Biver, C., Campagna, R., Gledhill, W.E., Pagga, U., Steber, J., Reust, H. and Bontinck, W.J. (1989) Screening of chemicals for anaerobic biodegradation. Chemosphere 19, p. 1527-1550.
- (31) Pagga, U. and Beimborn, D.B. (1993) Anaerobic biodegradation tests for organic compounds. Chemosphere 27, p. 1499-1509.

**▼B**

- (32) Nuck, B.A. and Federle, T.W., (1986) A batch test for assessing the mineralisation of <sup>14</sup>C-radiolabelled compounds under realistic anaerobic conditions. *Environ. Sci. Technol.* 30, p. 3597-3603.
- (33) US-EPA (1998b). Anaerobic biodegradability of organic chemicals. Harmonised Test Guidelines (OPPTS 835.3400). EPA 712-C-98-090.
- (34) Sijm, Haller and Schrap (1997) Influence of storage on sediment characteristics and drying sediment on sorption coefficients of organic contaminants. *Bulletin Environ. Contam. Toxicol.* 58, p. 961-968.
- (35) Timme, G., Frehse H. and Laska V. (1986) Statistical interpretation and graphic representation of the degradational behaviour of pesticide residues II. *Pflanzenschutz — Nachrichten Bayer*, 39, p. 187-203.
- (36) Timme, G., Frehse, H. (1980) Statistical interpretation and graphic representation of the degradational behaviour of pesticide residues I. *Pflanzenschutz — Nachrichten Bayer*, 33, p. 47-60.
- (37) Carlton, R.R., and Allen, R., (1994) The use of a compartment model for evaluating the fate of pesticides in sediment/water systems. Brighton Crop Protection Conference — Pest and Diseases, p. 1349-1354.

**▼B***Appendix 1***GUIDANCE ON THE AEROBIC AND THE ANAEROBIC TEST SYSTEMS****Aerobic test system**

The aerobic test system described in this test method consists of an aerobic water layer (typical oxygen concentrations range from 7 to 10 mg·l<sup>-1</sup>) and a sediment layer, aerobic at the surface and anaerobic below the surface (typical average redox potentials ( $E_h$ ) in the anaerobic zone of the sediment range from — 80 to — 190 mV). Moistened air is passed over the surface of the water in each incubation unit to maintain sufficient oxygen in the head space.

**Anaerobic test system**

For the anaerobic test system, the test procedure is essentially the same as that outlined for the aerobic system with the exception that moistened nitrogen is passed above the surface of the water in each incubation unit to maintain a head space of nitrogen. The sediment and water are regarded as anaerobic once the redox potential ( $E_h$ ) is lower than — 100 mV.

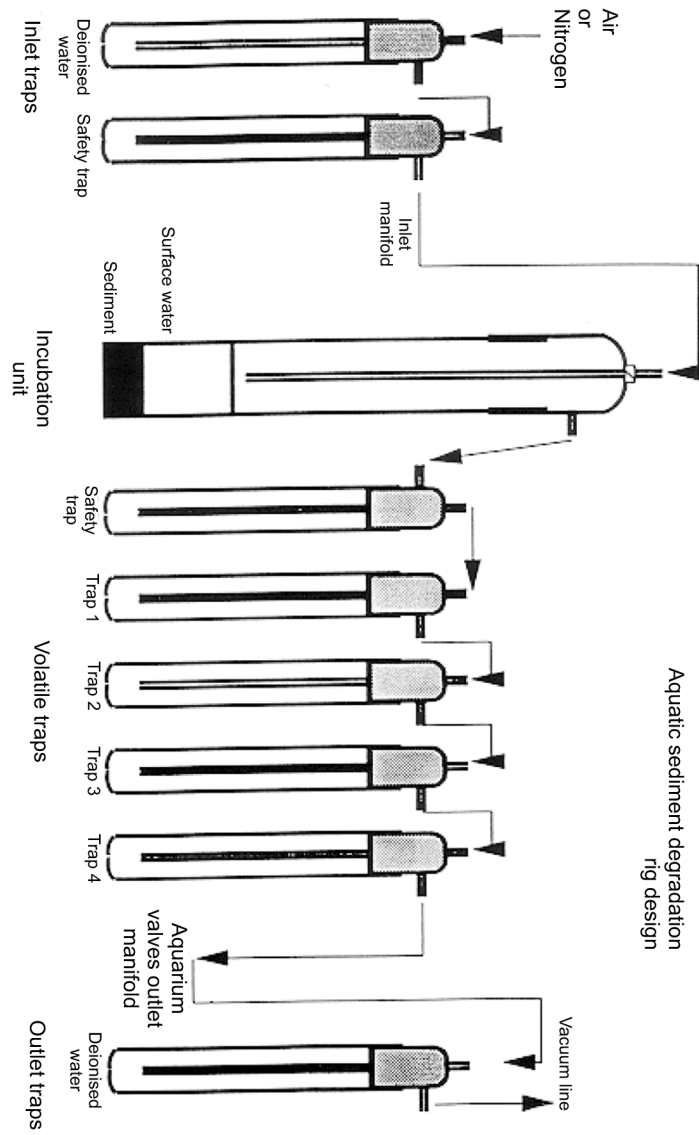
In the anaerobic test, assessment of mineralisation includes measurement of evolved carbon dioxide and methane.



▼B

Appendix 2

EXAMPLE OF A GAS FLOW-THROUGH APPARATUS

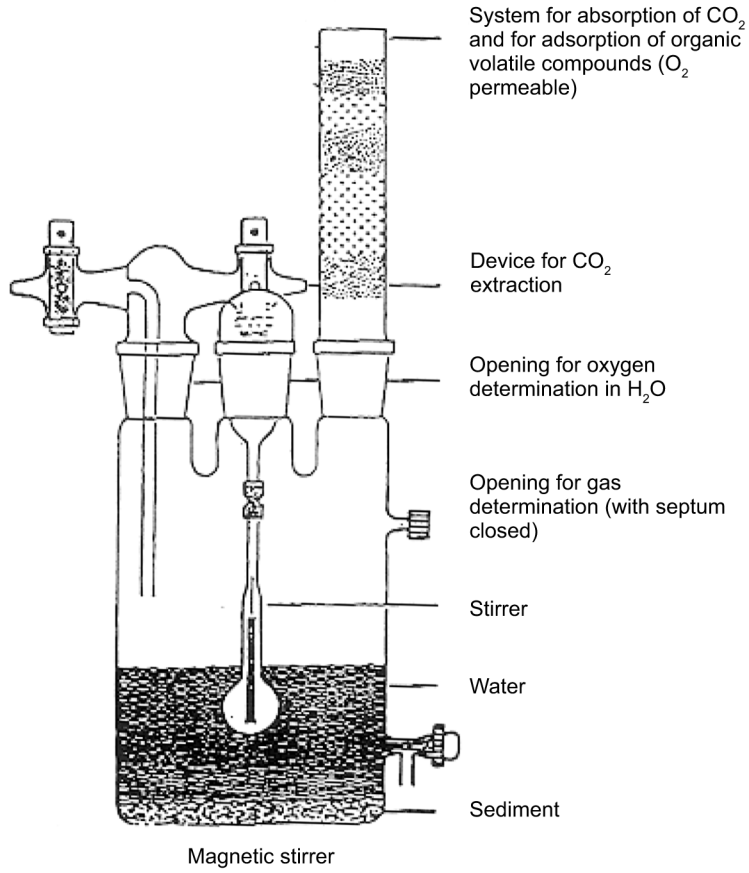


- Safety trap, empty
- Trap 1:  
ethyleneglycol to trap organic volatiles
- Trap 2:  
sulphuric acid 0.1 M to trap alkaline volatiles

▼B

Appendix 3

EXAMPLE OF A BIOMETER APPARATUS



**▼B***Appendix 4***EXAMPLE CALCULATION FOR APPLICATION DOSE TO TEST VESSELS**

Cylinder internal diameter:	= 8 cm
Water column depth not including sediment:	= 12 cm
Surface area: $3,142 \times 4^2$	= 50,3 cm <sup>2</sup>
Application rate: 500 g test substance/ha corresponds to 5 µg/cm <sup>2</sup>	
Total µg: $5 \times 50,3$	= 251,5 µg
Adjust quantity in relation to a depth of 100 cm:	
$12 \times 251,5 \div 100$	= 30,18 µg
Volume of water column: $50,3 \times 12$	= 603 ml
Concentration in water: $30,18 \div 603$	= 0,050 µg/ml or 50 µg/l

**▼ M1****C.25. AEROBIC MINERALISATION IN SURFACE WATER —  
SIMULATION BIODEGRADATION TEST****1. METHOD**

This method is equivalent to OECD TG 309 (2004) (1).

**1.1. INTRODUCTION**

The purpose of this test is to measure the time course of biodegradation of a test substance at low concentration in aerobic natural water and to quantify the observations in the form of kinetic rate expressions. This simulation test is a laboratory shake flask batch test to determine rates of aerobic biodegradation of organic substances in samples of natural surface water (fresh, brackish or marine). It is based on the ISO/DIS 14592-1 (2) and it also includes elements from the testing methods C.23 and C.24 (3)(4). Optionally, with long test times, semi-continuous operation replaces batch operation in order to prevent deterioration of the test microcosm. The principal objective of the simulation test is to determine the mineralisation of the test substance in surface water, and mineralisation constitutes the basis for expressing degradation kinetics. However, an optional secondary objective of the test is to obtain information on the primary degradation and the formation of major transformation products. Identification of transformation products, and if possible quantification of their concentrations, are especially important for substances that are very slowly mineralised (e.g. with half-lives for total residual  $^{14}\text{C}$  exceeding 60 days). Higher concentrations of the test substance (e.g.  $> 100 \mu\text{g/l}$ ) should normally be used for identification and quantification of major transformation products due to analytical limitations.

A low concentration in this test means a concentration (e.g. less than  $1 \mu\text{g/l}$  to  $100 \mu\text{g/l}$ ) which is low enough to ensure that the biodegradation kinetics obtained in the test reflect those expected in the environment. Compared to the total mass of biodegradable carbon substrates available in the natural water used for the test, the test substance present at low concentration will serve as a secondary substrate. This implies that the anticipated biodegradation kinetics is first order ('non-growth' kinetics) and that the test substance may be degraded by 'cometabolism'. First order kinetics implies that the rate of degradation ( $\text{mg/L/day}$ ) is proportional to the concentration of substrate which declines over time. With true first order kinetics the specific degradation rate constant,  $k$ , is independent of time and concentration. That is,  $k$  does not vary appreciably during the course of an experiment and does not change with the added concentration between experiments. By definition, the specific degradation rate constant is equal to the relative change in concentration per time:  $k = (1/C) \cdot (dC/dt)$ . Although first order kinetics are normally expected under the prescribed conditions, there may be certain circumstances where other kinetics are more appropriate. Deviations from first order kinetics may e.g. be observed if mass transfer phenomena such as the diffusion rate, rather than the biological reaction rate, is limiting the rate of biotransformation. However, the data can nearly always be described by pseudo first order kinetics accepting a concentration dependent rate constant.

**▼ M1**

Information on biodegradability of the test substance at higher concentrations (e.g. from standard screening tests) as well as information on abiotic degradability, transformation products and relevant physico-chemical properties should be available prior to the test to help establish the experimental planning and interpret the results. The use of  $^{14}\text{C}$  labelled test substances and the determination of the phase distribution of  $^{14}\text{C}$  at the end of the test, enable ultimate biodegradability to be determined. When non-labelled test substance is used, ultimate biodegradation can only be estimated if a higher concentration is tested and all the major transformation products are known.

## 1.2. DEFINITIONS

**Primary biodegradation:** The structural change (transformation) of a chemical substance by microorganisms resulting in the loss of chemical identity.

**Functional biodegradation:** The structural change (transformation) of a chemical substance by microorganisms resulting in the loss of a specific property.

**Ultimate aerobic biodegradation:** The breakdown of a chemical substance by microorganisms in the presence of oxygen to carbon dioxide, water and mineral salts of any other elements present (mineralisation) and the production of new biomass and organic microbial biosynthesis products.

**Mineralisation:** The breakdown of a chemical substance or organic matter by microorganisms in the presence of oxygen to carbon dioxide, water and mineral salts of any other elements present.

**Lag phase:** The time from the start of a test until adaptation of the degrading micro organisms is achieved and the biodegradation degree of a chemical substance or organic matter has increased to a detectable level (e.g. 10 % of the maximum theoretical biodegradation, or lower, dependent on the accuracy of the measuring technique).

**Maximum level of biodegradation:** The degree of biodegradation of a chemical substance or organic matter in a test, recorded in per cent, above which no further biodegradation takes place during the test.

**Primary substrate:** A collection of natural carbon and energy sources that provide growth and maintenance of the microbial biomass.

**Secondary substrate:** A substrate component present in such a low concentration, that by its degradation, only insignificant amounts of carbon and energy are supplied to the competent microorganisms, as compared to the carbon and energy supplied by the degradation of main substrate components (primary substrates).

**Degradation rate constant:** A first order or pseudo first order kinetic rate constant,  $k$  ( $\text{d}^{-1}$ ), which indicates the rate of degradation processes. For a batch experiment  $k$  is estimated from the initial part of the degradation curve obtained after the end of the lag phase.

**▼ M1**

**Half-life,  $t_{1/2}$  (d):** Term used to characterise the rate of a first order reaction. It is the time interval that corresponds to a concentration decrease by a factor 2. The half-life and the degradation rate constant are related by the equation  $t_{1/2} = \ln 2/k$ .

**Degradation half time,  $DT_{50}$  (d):** Term used to quantify the outcome of biodegradation tests. It is the time interval, including the lag phase, needed to reach a value of 50 % biodegradation.

**Limit of detection (LOD) and limit of quantification (LOQ):** The limit of detection (LOD) is the concentration of a substance below which the identity of the substance cannot be distinguished from analytical artefacts. The limit of quantification (LOQ) is the concentration of a substance below which the concentration cannot be determined with an acceptable accuracy.

**Dissolved organic carbon (DOC):** That part of the organic carbon in a sample of water which cannot be removed by specified phase separation, for example by centrifugation at  $40\,000\text{ ms}^{-2}$  for 15 min. or by membrane filtration using membranes with pores of  $0,2\text{ }\mu\text{m}$ - $0,45\text{ }\mu\text{m}$  diameter.

**Total organic  $^{14}\text{C}$  activity (TOA):** The total  $^{14}\text{C}$  activity associated with organic carbon.

**Dissolved organic  $^{14}\text{C}$  activity (DOA):** The total  $^{14}\text{C}$  activity associated with dissolved organic carbon.

**Particulate organic  $^{14}\text{C}$  activity (POA):** The total  $^{14}\text{C}$  activity associated with particulate organic carbon.

### 1.3. APPLICABILITY OF THE TEST

This simulation test is applicable to non-volatile or slightly volatile organic substances tested at low concentrations. Using flasks open to the atmosphere (e.g. cotton wool plugged), substances with Henry's law constants less than about  $1\text{ Pa}\cdot\text{m}^3/\text{mol}$  (approx.  $10^{-5}\text{ atm}\cdot\text{m}^3/\text{mol}$ ) can be regarded as non-volatile in practice. Using closed flasks with a headspace, it is possible to test slightly volatile substances (with Henry's law constants  $< 100\text{ Pa}\cdot\text{m}^3/\text{mol}$  or  $< 10^{-3}\text{ atm}\cdot\text{m}^3/\text{mol}$ ) without losses from the test system. Loss of  $^{14}\text{C}$ -labelled substances may occur, if the right precautions are not exercised, when the  $\text{CO}_2$  is stripped off. In such situations, it may be necessary to trap  $\text{CO}_2$  in an internal absorber with alkali or to use an external  $\text{CO}_2$  absorber system (direct  $^{14}\text{CO}_2$  determination; see Appendix 3). For the determination of biodegradation kinetics, the concentrations of the test substance must be below its water solubility. It should be noted, however, that literature values of water solubility may be considerably higher than the solubility of the test substance in natural waters. Optionally, the solubility of especially poorly water-soluble test substances may be established by use of the natural waters being tested.

The method can be used for simulating biodegradation in surface water free of coarse particles (pelagic test) or in turbid surface water which, e.g. might exist near a water/sediment interface (suspended sediment test).

**▼ M1**

## 1.4. PRINCIPLE OF THE TEST

The test is performed in batch by incubating the test substance with either surface water only (pelagic test) or surface water amended with suspended solids/sediment of 0,01 to 1 g/L dry weight (suspended sediment test) to simulate a water body with suspended solids or re-suspended sediment. The suspended solids/sediment concentration in the lower range of this interval is typical for most surface waters. The test flasks are incubated in darkness at an environmental temperature under aerobic conditions and agitation. At least two different concentrations of test substance should be used in order to determine the degradation kinetics. The concentrations should differ from each other by a factor of 5 to 10 and should represent the expected range of concentrations in the environment. The maximum concentration of the test substance should not exceed 100 µg/L, but maximum test concentrations below 10 µg/L or less are preferred to ensure that the biodegradation follows first order kinetics. The lowest concentration should not exceed 10 µg/L, but lowest test concentrations of 1-2 µg/L or less than 1 µg/L are preferred. Normally an adequate analysis of such low concentration can be achieved by use of commercially available <sup>14</sup>C-labelled substances. Because of analytical limitations, it is frequently impossible to measure the concentration of the test substance with the required accuracy, if the test substance is applied at a concentration ≤ 100 µg/L (see second paragraph in section 1.7.2). Higher concentrations of test substance (> 100 µg/L and sometimes > 1 mg/L) may be used for the identification and quantification of major transformation products or if a specific analysis method with a low detection limit is not available. If high concentrations of test substance are tested, it may not be possible to use the results to estimate the first order degradation constant and half-life, as the degradation will probably not follow first order kinetics.

Degradation is followed at appropriate time intervals, by measuring either the residual <sup>14</sup>C or the residual concentration of test substance when specific chemical analysis is used. <sup>14</sup>C labelling of the most stable part of the molecule ensures the determination of the total mineralisation, while <sup>14</sup>C labelling of a less stable part of the molecule, as well as the use of specific analysis, enable the assessment of only primary biodegradation. However, the most stable part does not necessarily include the relevant functional moiety of the molecule (that can be related to a specific property such as toxicity, bioaccumulation, etc.). If this is the case, it may be appropriate to use a test substance, which is <sup>14</sup>C-labelled, in the functional part in order to follow the elimination of the specific property.

## 1.5. INFORMATION ON THE TEST SUBSTANCE

Both radiolabelled and non-labelled test substances can be used in this test. <sup>14</sup>C-labelling technique is recommended and labelling should normally be in the most stable part(s) of the molecule (see also section 1.4). For substances containing more than one aromatic ring, one or more carbons in each ring should preferably be <sup>14</sup>C-labelled. In addition, one or more carbons on both sides of easily degradable linkages should preferably be <sup>14</sup>C-labelled. The chemical and/or radiochemical purity of the test substance should be > 95 %. For radiolabelled substances, a specific activity of approx. 50 µCi/mg (1,85 MBq) or more is preferred in order to facilitate <sup>14</sup>C measurements in tests conducted with low initial concentrations. The following information on the test substance should be available:

**▼ M1**

- solubility in water [Method A.6],
- solubility in organic solvent(s) (substances applied with solvent or with low solubility in water),
- dissociation constant (pKa) if the substance is liable to protonation or deprotonation [OECD TG 112] (5),
- vapour pressure [Method A.4] and Henry's law constant,
- chemical stability in water and in the dark (hydrolysis) [Method C.7].

When poorly water-soluble substances are being tested in seawater, it may also be useful to know the salting out constant (or 'Setschenow constant')  $K^s$ , which is defined by the expression:  $\log(S/S') = K^s C_m$ , where  $S$  and  $S'$  are the solubility of the substance in fresh water and seawater, respectively, and  $C_m$  is the molar salt concentration.

If the test is carried out as a 'suspended sediment test' the following information should also be available:

- n-octanol/water partition coefficient [Method A.8],
- adsorption coefficient [Method C.18].

Other useful information may include:

- environmental concentration, if known or estimated,
- toxicity of the test substance to microorganisms [Method C.11],
- ready and/or inherent biodegradability [Methods C.4 A-F, C.12, C.9, OECD TG 302 (5)],
- aerobic or anaerobic biodegradability in soil and sediment/water transformation studies [Methods C.23, C.24].

#### 1.6. REFERENCE SUBSTANCE

A substance, which is normally easily degraded under aerobic conditions (e.g. aniline or sodium benzoate) should be used as reference substance. The expected time interval for degradation of aniline and sodium benzoate is usually less than 2 weeks. The purpose of the reference substances is to ensure that the microbial activity of the test water is within certain limits; i.e. that the water contains an active microbial population.



**▼ M1**

## 1.7. QUALITY CRITERIA

1.7.1. **Recovery**

Immediately after addition of the test substance, each initial test concentration should be verified by measurements of  $^{14}\text{C}$  activity, or by chemical analyses in the case of non-labelled substances, in at least duplicate samples. This provides information on the applicability and repeatability of the analytical method and on the homogeneity of the distribution of the test substance. Normally, the measured initial  $^{14}\text{C}$  activity or test substance concentration is used in the subsequent analyses of data rather than the nominal concentration as losses due to sorption and dosing errors thereby are compensated. For  $^{14}\text{C}$ -labelled test substance, the level of recovery at the end of the experiment is given by mass balance (see last paragraph in section 1.8.9.4). Ideally, the radiolabelled mass balance should range from 90 % to 110 %, whereas the analytical accuracy should lead to an initial recovery of between 70 % and 110 % for non-labelled test substances. These ranges should be interpreted as targets and should not be used as criteria for acceptance of the test. Optionally, the analytical accuracy may be determined for the test substance at a lower concentration than the initial concentration and for major transformation products.

1.7.2. **Repeatability and sensitivity of analytical method**

Repeatability of the analytical method (including the efficiency of the initial extraction) to quantify the test substance, and transformation products, if appropriate, should be checked by five replicate analyses of the individual extracts of the surface water.

The limit of detection (LOD) of the analytical method for the test substance and for the transformation products should be at least 1 % of the initial amount applied to the test system if possible. The limit of quantification (LOQ) should be equal to or less than 10 % of the applied concentration. The chemical analyses of many organic substances and their transformation products frequently require that the test substance is applied at a relatively high concentration, i.e. > 100  $\mu\text{g/L}$ .

## 1.8. DESCRIPTION OF THE TEST METHOD

1.8.1. **Equipment**

The test may be conducted in conical or cylindrical flasks of appropriate capacity (e.g. 0,5 or 1,0 litre) closed with silicone or rubber stoppers, or in serum flasks with  $\text{CO}_2$ -tight lids (e.g. with butyl rubber septa). Another option is to perform the test by use of multiple flasks and to harvest whole flasks, at least in duplicate, at each sample interval (see last paragraph in section 1.8.9.1). For non-volatile test substances that are not radiolabelled, gas-tight stoppers or lids are not required; loose cotton plugs that prevent contamination from air are suitable (see second paragraph in section 1.8.9.1). Slightly volatile substances should be tested in a biometer-type system with gentle stirring of the water surface. To be sure that no bacterial contamination occurs, optionally the vessels can be sterilised by heating or autoclaving prior to use. In addition, the following standard laboratory equipment is used:

— shaking table or magnetic stirrers for continuous agitation of the test flasks,

**▼ M1**

- centrifuge,
  
- pH meter,
  
- turbidimeter for nephelometric turbidity measurements,
  
- oven or microwave oven for dry weight determinations,
  
- membrane filtration apparatus,
  
- autoclave or oven for heat sterilisation of glassware,
  
- facilities to handle <sup>14</sup>C-labelled substances,
  
- equipment to quantify <sup>14</sup>C-activity in samples from CO<sub>2</sub>-trapping solutions and, if required, from sediment samples,
  
- analytical equipment for the determination of the test (and reference) substance if specific chemical analysis is used (e.g. gas chromatograph, high-pressure liquid chromatograph).

**1.8.2. Stock solutions of test substance**

Deionised water is used to prepare stock solutions of the test and reference substances (see first paragraph in section 1.8.7). The deionised water should be free of substances that may be toxic to microorganisms, and dissolved organic carbon (DOC) should be no more than 1 mg/L (6).

**1.8.3. Collection and transport of surface water**

The sampling site for collection of the surface water should be selected in accordance with the purpose of the test in any given situation. In selecting sampling sites, the history of possible agricultural, industrial or domestic inputs must be considered. If it is known that an aquatic environment has been contaminated with the test substance or its structural analogues within the previous four years, it should not be used for the collection of test water, unless investigation of degradation rates in previously exposed sites is the express purpose of the investigator. The pH and temperature of the water should be measured at the site of collection. Furthermore, the depth of sampling and the appearance of the water sample (e.g. colour and turbidity) should be noted (see section 3). Oxygen concentration and/or redox potential in water and in the sediment surface layer should be measured in order to demonstrate aerobic conditions unless this is obvious as judged from appearance and historic experience with the site. The surface water should be transported in a thoroughly cleansed container. During transport, the temperature of the sample should not significantly exceed the temperature used in the test. Cooling to 4 °C is recommended if transport duration exceeds 2 to 3 hours. The water sample must not be frozen.

**▼ M1****1.8.4. Storage and preparation of surface water**

The test should preferably be started within one day after sample collection. Storage of the water, if needed, should be minimised and must in any case not exceed a maximum of 4 weeks. The water sample should be kept at 4 °C with aeration until use. Prior to use, the coarse particles should be removed, e.g. by filtration through a nylon filter with about 100 µm mesh size or with a coarse paper filter, or by sedimentation.

**1.8.5. Preparation of water amended with sediment (optional)**

For the suspended sediment test, surface sediment is added to the flasks containing natural water (filtered to remove coarse particles as described in section 1.8.4) to obtain a suspension; the concentration of suspended solids should be between 0,01 and 1 g/L. The surface sediment should come from the same site as that from which the water sample was taken. Dependent on the particular aquatic environment, the surface sediment may either be characterised by a high organic carbon content (2,5-7,5 %) and a fine texture or by a low organic carbon content (0,5-2,5 %) and a coarse texture (3). The surface sediment can be prepared as follows: extract several sediment cores using a tube of transparent plastic, slice off the upper aerobic layers (from surface to a depth of max. 5 mm) immediately after sampling and pool them together. The resulting sediment sample should be transported in a container with a large air headspace to keep the sediment under aerobic conditions (cool to 4 °C if transport duration exceeds 2-3 hours). The sediment sample should be suspended in the test water at a ratio of 1:10 and kept at 4 °C with aeration until use. Storage of the sediment, if needed, should be minimised and must not in any case exceed a maximum of 4 weeks.

**1.8.6. Semi-continuous procedure (optional)**

Prolonged incubation (several months) may be necessary if a long lag time occurs before a significant degradation of the test substance can be measured. If this is known from previous testing of a substance, the test may be initiated by using a semi-continuous procedure, which allows periodical renewal of a part of the test water or suspension (see Appendix 2). Alternatively, the normal batch test may be changed into a semi-continuous test, if no degradation of the test substance has been achieved during approximately 60 days of testing using the batch procedure (see second paragraph in section 1.8.8.3).

**1.8.7. Addition of the test (or reference) substance**

For substances with high water solubility (> 1 mg/L) and low volatility (Henry's law constants < 1 Pa·m<sup>3</sup>/mol or < 10<sup>-5</sup> atm·m<sup>3</sup>/mol), a stock solution can be prepared in deionised water (see section 1.8.2); the appropriate volume of the stock solution is added to the test vessels to achieve the desired concentration. The volume of any added stock solution should be held to the practical minimum (< 10 % of the final liquid volume, if possible). Another procedure is to dissolve the test substance in a larger volume of the test water, which may be seen as an alternative to the use of organic solvents.

**▼ M1**

If unavoidable, stock solutions of non-volatile substances with poor water-solubility should be prepared by use of a volatile organic solvent, but the amount of solvent added to the test system should not exceed 1 % v/v and should not have adverse effects on the microbial activity. The solvent should not affect the stability of the test substance in water. The solvent should be stripped off to an extremely small quantity so that it does not significantly increase the DOC concentration of the test water or suspension. This should be checked by substance-specific analysis or, if possible, DOC analysis (6). Care must be taken to limit the amount of solvent transferred to what is absolutely necessary, and to ensure that the amount of test substance can dissolve in the final volume of test water. Other techniques to introduce the test substance into the test vessels may be used as described in (7) and (8). When an organic solvent is used for application of the test substance, solvent controls containing the test water (with no additions) and test water with added reference substance should be treated similarly to active test vessels amended with test substance in solvent carrier. The purpose of the solvent controls is to examine possible adverse effects caused by the solvent towards the microbial population as indicated by the degradation of the reference substance.

**1.8.8. Test conditions****1.8.8.1. Test temperature**

Incubation should take place in the dark (preferred) or in diffuse light at a controlled ( $\pm 2$  °C) temperature, which may be the field temperature or a standard temperature of 20-25 °C. Field temperature may be either the actual temperature of the sample at the sampling time or an average field temperature at the sampling site.

**1.8.8.2. Agitation**

Agitation by means of continuous shaking or stirring must be provided to maintain particles and microorganisms in suspension. Agitation also facilitates oxygen transfer from the headspace to the liquid so that aerobic conditions can be adequately maintained. Place the flasks on a shaking table (approx. 100 rpm agitation) or use magnetic stirring. Agitation must be continuous. However, the shaking or stirring should be as gentle as possible, while still maintaining a homogeneous suspension.

**1.8.8.3. Test duration**

The duration of the test should normally not exceed 60 days unless the semi-continuous procedure with periodical renewal of the test suspension is applied (see section 1.8.6 and Appendix 2). However, the test period for the batch test may be extended to a maximum of 90 days, if the degradation of the test substance has started within the first 60 days. Degradation is monitored, at appropriate time intervals, by the determination of the residual  $^{14}\text{C}$  activity or the evolved  $^{14}\text{CO}_2$  (see section 1.8.9.4) and/or by chemical analysis (section 1.8.9.5). The incubation time must be sufficiently long to evaluate the degradation process. The extent of degradation should preferably exceed 50 %; for slowly degradable substances, the extent of degradation must be sufficient (normally greater than 20 % degradation) to ensure the estimation of a kinetic degradation rate constant.

**▼ M1**

Periodic measurements of pH and oxygen concentration in the test system must be conducted unless previous experience from similar tests with water and sediment samples collected from the same site make such measurements unnecessary. Under some conditions, the metabolism of primary substrates at much higher concentrations within the water or sediment could possibly result in enough CO<sub>2</sub> evolution and oxygen depletion to significantly alter the experimental conditions during the test.

**1.8.9. Procedure****1.8.9.1. Preparation of flasks for pelagic test**

Transfer a suitable volume of test water to the test flasks, up to about one third of the flask volume and not less than about 100 ml. If multiple flasks are used (to allow harvesting of whole flasks at each sampling time), the appropriate volume of test water is also about 100 ml, as small sample volumes may influence the length of the lag phase. The test substance is added from a stock solution as described in sections 1.8.2 and 1.8.7. At least two different concentrations of test substance differing by a factor of 5 to 10 should be used in order to determine degradation kinetics and calculate the kinetic degradation rate constant. Both of the selected concentrations should be less than 100 µg/L and preferably in the range of < 1-10 µg/L.

Close the flasks with stoppers or lids impermeable to air and CO<sub>2</sub>. For non-<sup>14</sup>C-labelled non-volatile test chemicals, loose cotton wool plugs that prevent contamination from air are suitable (see section 1.8.1) provided that any major degradation products are known to be non-volatile, and if indirect CO<sub>2</sub> determination is used (see Appendix 3).

Incubate the flasks at the selected temperature (see section 1.8.8.1). Withdraw samples for chemical analysis or <sup>14</sup>C measurements at the beginning of the test (i.e. before biodegradation starts; see section 1.7.1) and then at suitable time intervals during the course of the test. Sampling may be performed by withdrawal of sub-samples (e.g. 5 ml aliquots) from each replicate or by harvest of whole flasks at each sampling time. The mineralisation of the test substance may either be determined indirectly or directly (see Appendix 3). Usually, a minimum of five sampling points are required during the degradation phase (i.e. after ended lag phase) in order to estimate a reliable rate constant, unless it can be justified that three sampling points are sufficient for rapidly degradable substances. For substances that are not rapidly degraded more measurements during the degradation phase can easily be made and, therefore, more data points should be used for the estimation of k. No fixed time schedule for sampling can be stated, as the rate of biodegradation varies; however the recommendation is to sample once a week if degradation is slow. If the test substance is rapidly degradable, sampling should take place once a day during the first three days and then every second or third day. Under certain circumstances, such as with very rapidly hydrolysing substances, it may be necessary to sample at hourly intervals. It is recommended that a preliminary study is conducted prior to the test in order to determine the appropriate sampling intervals. If samples have to be available for further specific analysis, it is advisable to take more samples and then select those to be analysed at the end of the experiment following a backwards strategy, i.e. the last samples are analysed first (see second paragraph in section 1.8.9.5 for guidance on stability of samples during storage).

**▼ M1**1.8.9.2. *Number of flasks and samples*

Set up a sufficient number of test flasks to have:

- test flasks; at least duplicate flasks for each concentration of test substance (preferably a minimum of 3) or multiple test flasks for each concentration, if whole flasks are harvested at each sampling time (symbolised  $F_T$ ),
  
- test flasks for mass balance calculation; at least duplicate flasks for each test concentration (symbolised  $F_M$ ),
  
- blank control, no test substance; at least one blank test flask containing only the test water (symbolised  $F_B$ ),
  
- reference control; duplicate flasks with reference substance (e.g. aniline or sodium benzoate, at 10 µg/l) (symbolised  $F_C$ ). The purpose of the reference control is to confirm a minimum of microbial activity. If convenient, a radiolabelled reference substance may be used, also when the degradation of the test substance is monitored by chemical analyses,
  
- sterile control; one or two flasks containing sterilised test water for examining possible abiotic degradation or other non-biological removal of the test substance (symbolised  $F_S$ ). The biological activity can be stopped by autoclaving (121 °C; 20 min.) the test water or by adding a toxicant (e.g. sodium azide ( $\text{NaN}_3$ ) at 10-20 g/l, mercuric chloride ( $\text{HgCl}_2$ ) at 100 mg/l or formalin at 100 mg/l) or by gamma irradiation. If  $\text{HgCl}_2$  is used, it should be disposed of as toxic waste. For water with sediment added in large amount, sterile conditions are not easy to obtain; in this case repeated autoclaving (e.g. three times) is recommended. It should be considered that the sorption characteristics of the sediment may be altered by autoclaving,
  
- solvent controls, containing test water and test water with reference substance; duplicate flasks treated with the same amount of solvent and by use of the same procedure as that used for application of the test substance. The purpose is to examine possible adverse effects of the solvent by determining the degradation of the reference substance.

In the design of the test, the investigator should consider the relative importance of increased experimental replication versus increased number of sampling times. The exact number of flasks required will depend on the method used for measuring the degradation (see third paragraph in section 1.8.9.1; section 1.8.9.4 and Appendix 3).

**▼ M1**

Two subsamples (e.g. 5 ml aliquots) should be withdrawn from each test flask at each sampling time. If multiple flasks are used to allow harvesting of whole flasks, a minimum of two flasks should be sacrificed at each sampling time (see first paragraph in section 1.8.9.1).

**1.8.9.3. *Preparation of flasks for suspended sediment test [optional]***

Add the necessary volumes of test water and sediment, if required, to the test vessels (see section 1.8.5). The preparation of flasks for suspended sediment test is the same as for the pelagic test (see sections 1.8.9.1 and 1.8.9.2). Use preferably serum bottles or similar shaped flasks. Place the closed flasks horizontally on a shaker. Obviously, open flasks for non-<sup>14</sup>C-labelled, non-volatile substances should be placed in upright position; in this case magnetic stirring and the use of magnetic bars coated with glass are recommended. If necessary, aerate the bottles to maintain proper aerobic conditions.

**1.8.9.4. *Radiochemical determinations***

The evolved <sup>14</sup>CO<sub>2</sub> is measured indirectly and directly (see Appendix 3). The <sup>14</sup>CO<sub>2</sub> is determined indirectly by the difference between the initial <sup>14</sup>C activity in the test water or suspension and the total residual activity at the sampling time as measured after acidifying the sample to pH 2-3 and stripping off CO<sub>2</sub>. Inorganic carbon is thus removed and the residual activity measured derives from organic material. The indirect <sup>14</sup>CO<sub>2</sub> determination should not be used, if major volatile transformation products are formed during the transformation of the test substance (see Appendix 3). If possible, the <sup>14</sup>CO<sub>2</sub> evolution should be measured directly (see Appendix 3) at each sampling time in at least one test flask; this procedure enables both the mass balance and biodegradation process to be checked, but it is restricted to tests conducted with closed flasks.

If the evolved <sup>14</sup>CO<sub>2</sub> is measured directly during the test, more flasks should be set up for this purpose at the start of the test. Direct <sup>14</sup>CO<sub>2</sub> determination is recommended, if major volatile transformation products are formed during the transformation of the test substance. At each measuring point the additional test flasks are acidified to pH 2-3 and the <sup>14</sup>CO<sub>2</sub> is collected in an internal or external absorber (see Appendix 3).

Optionally the concentrations of <sup>14</sup>C-labelled test substance and major transformation products may be determined by use of radiochromatography (e.g. thin layer chromatography, RAD-TLC) or HPLC with radiochemical detection.

Optionally the phase distribution of the remaining radioactivity (see Appendix 1) and residual test substance and transformation products may be determined.

**▼ M1**

At the end of the test the mass balance should be determined by direct  $^{14}\text{CO}_2$  measurement using separate test flasks from which no samples are taken in the course of the test (see Appendix 3).

**1.8.9.5. Specific chemical analysis**

If a sensitive specific analytical method is available, primary biodegradation can be assessed by measuring the total residual concentration of test substance instead of using radiolabelling techniques. If a radiolabelled test substance is used (to measure total mineralisation), specific chemical analyses can be made in parallel to provide useful additional information and check the procedure. Specific chemical analyses may also be used to measure transformation products formed during the degradation of the test substance, and this is recommended for substances that are mineralised with half-lives exceeding 60 days. The concentration of the test substance and the transformation products at every sampling time should be measured and reported (as a concentration and as percentage of applied). In general, transformation products detected at  $\geq 10\%$  of the applied concentration at any sampling time should be identified unless reasonably justified otherwise. Transformation products for which concentrations are continuously increasing during the study should also be considered for identification, even if their concentrations do not exceed the limit given above, as this may indicate persistence. Analyses of transformation products in sterile controls should be considered, if rapid abiotic transformation of the test substance (e.g. hydrolysis) is thought possible. The need for quantification and identification of transformation products should be considered on a case by case basis, with justifications being provided in the report. Extraction techniques with organic solvent should be applied according to directions given in the respective analytical procedure.

All samples should be stored at 2 to 4 °C and air-tight if analysis is carried out within 24 hours (preferred). For longer storage, the samples should be frozen below – 18 °C or chemically preserved. Acidification is not a recommended method to preserve the samples, because acidified samples may be unstable. If the samples are not analysed within 24 hours and are subject to longer storage, a storage stability study should be conducted to demonstrate the stability of chemicals of interest under – 18 °C storage or preserved conditions. If the analytical method involves either solvent extraction or solid phase extraction (SPE), the extraction should be performed immediately after sampling or after storing the sample refrigerated for a maximum of 24 hours.

Depending on the sensitivity of the analytical method, larger sample volumes than those indicated in section 1.8.1 may be necessary. The test can easily be carried out with test volumes of one litre in flasks of 2-3 litre volume, which makes it possible to collect samples of approx. 100 ml.



**▼ M1****2. DATA AND REPORTING****2.1. TREATMENT OF RESULTS****2.1.1. Plot of data**

Round off sampling times to a whole number of hours (unless the substance degrades substantially in a matter of minutes to hours) but not to a whole number of days. Plot the estimates of the residual activity of test substance (for  $^{14}\text{C}$ -labelled substances) or the residual concentration (for non-labelled substances), against time both in a linear and in a semi-logarithmic plot (see Figures 1a, 1b). If degradation has taken place, compare the results from flasks  $F_T$  with those from flasks  $F_S$ . If the means of the results from the flasks with test substance ( $F_T$ ) and the sterile flasks ( $F_S$ ) deviate by less than 10 %, it can be assumed that the degradation observed is predominantly abiotic. If the degradation in flasks  $F_S$  is lower, the figures may be used to correct those obtained with flasks  $F_T$  (by subtraction) in order to estimate the extent of biodegradation. When optional analyses are performed for major transformation products, plots of their formation and decline should be provided in addition to a plot of the decline of the test substance.

Estimate the lag phase duration  $t_L$  from the degradation curve (semi-logarithmic plot) by extrapolating its linear part to zero degradation or alternatively by determining the time for approximately 10 % degradation (see Figures 1a and 1b). From the semi-logarithmic plot, estimate the first order rate constant,  $k$ , and its standard error by linear regression of  $\ln$  (residual  $^{14}\text{C}$  activity or test substance concentration) versus time. With  $^{14}\text{C}$  measurements in particular, use only data belonging to the initial linear part of the curve after the ended lag phase, and give preference to selecting few and representative data rather than selecting a greater number of more uncertain data. Uncertainty includes here errors inherent in the recommended direct use of measured residual  $^{14}\text{C}$  activities (see below). It may sometimes be relevant to calculate two different rate constants, if the degradation follows a biphasic pattern. For this purpose two different phases of the degradation curve are defined. Calculations of the rate constant,  $k$ , and the half-life  $t_{1/2} = \ln 2/k$ , should be carried out for each of the individual replicate flasks, when sub-samples are withdrawn from the same flask, or by using the average values, when whole flasks are harvested at each sampling time (see last paragraph in section 1.8.9.2). When the first-mentioned procedure is used, the rate constant and half-life should be reported for each of the individual replicate flasks and as an average value with a standard error. If high concentrations of test substance have been used, the degradation curve may deviate considerably from a straight line (semi-logarithmic plot) and first order kinetics may not be valid. Defining a half-life has therefore no meaning. However, for a limited data range, pseudo first order kinetics can be applied and the degradation half-time  $DT_{50}$  (time to reach 50 % degradation) estimated. It must be borne in mind, however, that the time course of degradation beyond the selected data range cannot be predicted using the  $DT_{50}$  which is merely a descriptor of a given set of data. Analytical tools to facilitate statistical calculations and curve fitting are easily available and the use of this kind of software is recommended.

**▼ M1**

If specific chemical analyses are made, estimate rate constants and half-lives for primary degradation as above for total mineralisation. If the primary degradation is the limiting process data points from the entire course of degradation may sometimes be used. This is because measurements are direct by contrast to measurements of  $^{14}\text{C}$  activity.

If  $^{14}\text{C}$ -labelled substances are used, a mass balance should be expressed in percentage of the applied initial concentration, at least at the end of the test.

**2.1.2. Residual activity**

When the  $^{14}\text{C}$ -labelled part of an organic substance is biodegraded, the major part of the  $^{14}\text{C}$  is converted to  $^{14}\text{CO}_2$ , while another part is used for growth of biomass and/or synthesis of extra-cellular metabolites. Therefore, complete 'ultimate' biodegradation of a substance does not result in a 100 % conversion of its carbon into  $^{14}\text{CO}_2$ . The  $^{14}\text{C}$  built into products formed by biosynthesis is subsequently released slowly as  $^{14}\text{CO}_2$  due to 'secondary mineralisation'. For these reasons plots of residual organic  $^{14}\text{C}$  activity (measured after stripping off  $\text{CO}_2$ ) or of  $^{14}\text{CO}_2$  produced versus time will show a 'tailing' after degradation has been completed. This complicates a kinetic interpretation of the data and for this purpose, only the initial part of the curve (after the lag phase has ended and before approx. 50 % degradation is reached) should normally be used for the estimation of a degradation rate constant. If the test substance is degraded, the total residual organic  $^{14}\text{C}$  activity is always higher than the  $^{14}\text{C}$  activity associated with the remaining intact test substance. If the test substance is degraded by a first order reaction and a constant fraction  $\alpha$  is mineralised into  $\text{CO}_2$ , the initial slope of the  $^{14}\text{C}$  disappearance curve (total organic  $^{14}\text{C}$  versus time) will be  $\alpha$  times the slope of the corresponding curve for the concentration of test substance (or, to be precise, the part of the test substance labelled with  $^{14}\text{C}$ ). Using measurements of the total organic  $^{14}\text{C}$  activity uncorrected, the calculated degradation rate constant will therefore be conservative. Procedures for estimating the concentrations of the test substance from the measured radiochemical activities based on various simplifying assumptions have been described in the literature (2)(9)(10)(11). Such procedures are most easily applied for rapidly degradable substances.

**2.2. INTERPRETATION OF RESULTS**

If  $k$  is found to be independent of the added concentration (i.e. if the calculated  $k$  is approximately the same at the different concentrations of test substance), it can be assumed that the first order rate constant is representative of the testing conditions used, i.e. the test substance, the water sample and the test temperature. To what extent the results can be generalised or extrapolated to other systems must be evaluated by expert judgement. If a high concentration of test substance is used, and the degradation therefore does not follow first order kinetics, the data cannot be used for direct estimation of a first order rate constant or a corresponding half-life. However, data derived from a test using a high concentration of test substance may still be usable for estimating the degree of total mineralisation and/or detection and quantification of transformation products.

**▼ M1**

If the rates of other loss processes than biodegradation are known (e.g. hydrolysis or volatilisation), they may be subtracted from the net loss rate observed during the test to give an approximated estimate of the biodegradation rate. Data for hydrolysis may, e.g. be obtained from the sterile control or from parallel testing using a higher concentration of the test substance.

The indirect and direct determination of  $^{14}\text{CO}_2$  (section 1.8.9.4 and Appendix 3) can only be used to measure the extent of mineralisation of the test substance to  $\text{CO}_2$ . Radiochromatography (RAD-TLC) or HPLC may be used to analyse the concentrations of  $^{14}\text{C}$ -labelled test substance and the formation of major transformation products (third paragraph in section 1.8.9.4). To enable a direct estimation of the half-life, it is necessary that no major transformation products (defined as  $\geq 10\%$  of the applied amount of test substance) be present. If major transformation products as defined here are present, a detailed evaluation of the data is required. This may include repeated testing and/or identification of the transformation products (see first paragraph in section 1.8.9.5) unless the fate of the transformation products can be reasonably assessed by use of experience (e.g. information on degradation pathway). As the proportion of test substance carbon converted to  $\text{CO}_2$  varies (depending largely on the concentration of test substance and other substrates available, the test conditions and the microbial community), this test does not allow a straightforward estimation of ultimate biodegradation as in a DOC die-away test; but the result is similar to that obtained with a respirometric test. The degree of mineralisation will thus be less than or equal to the minimum level of ultimate biodegradation. To obtain a more complete picture of the ultimate biodegradation (mineralisation and incorporation into biomass), the analysis of the phase distribution of  $^{14}\text{C}$  should be performed at the end of the test (see Appendix 1). The  $^{14}\text{C}$  in the particulate pool will consist of  $^{14}\text{C}$  incorporated into bacterial biomass and  $^{14}\text{C}$  sorbed to organic particles.

### 2.3. VALIDITY OF THE TEST

If the reference substance is not degraded within the expected time interval (for aniline and sodium benzoate, usually less than two weeks), the validity of the test is suspected and must be further verified, or alternatively the test should be repeated with a new water sample. In an ISO ring-test of the method where seven laboratories located around Europe participated, adapted degradation rate constants for aniline ranged from  $0,3$  to  $1,7 \text{ day}^{-1}$  with an average of  $0,8 \text{ d}^{-1}$  at  $20 \text{ }^\circ\text{C}$  and a standard error of  $\pm 0,4 \text{ d}^{-1}$  ( $t_{1/2} = 0,9$  days). Typical lag times were 1 to 7 days. The waters examined were reported to have a bacterial biomass corresponding to  $10^3$  -  $10^4$  colony forming units (CFU) per ml. Degradation rates in nutrient-rich Mid-European waters were greater than in Nordic oligotrophic waters, which may be due to the different trophic status or previous exposure to chemical substances.

The total recovery (mass balance) at the end of the experiment should be between 90 % and 110 % for radiolabelled substances, whereas the initial recovery at the beginning of the experiment should be between 70 % and 110 % for non-labelled substances. However, the indicated ranges should only be interpreted as targets and should not be used as criteria for acceptance of the test.

**▼ M1****3. TEST REPORT**

The type of study, i.e. pelagic or suspended sediment test, must be clearly stated in the test report, which shall also contain at least the following information:

Test substance and reference substance(s):

- common names, chemical names (recommend IUPAC and/or CAS names), CAS numbers, structural formulas (indicating position of  $^{14}\text{C}$  if radiolabelled substance is used) and relevant physico-chemical properties of test and reference substance (see sections 1.5 and 1.6),
- chemical names, CAS numbers, structural formulas (indicating position of  $^{14}\text{C}$  if radiolabelled substance is used) and relevant physico-chemical properties of substances used as standards for identification and quantification of transformation products,
- purity (impurities) of test and reference substances,
- radiochemical purity of labelled chemical and specific activity (where appropriate).

Surface water:

The following minimum information for the water sample taken must be provided:

- location and description of sampling site including, if possible, contamination history,
- date and time of sample collection,
- nutrients (total N, ammonium, nitrite, nitrate, total P, dissolved orthophosphate),
- depth of collection,
- appearance of sample (e.g. colour and turbidity),
- DOC and TOC,
- BOD,
- temperature and pH at the place and time of collection,
- oxygen or redox potential (mandatory only if aerobic conditions are not obvious),
- salinity or conductivity (in the case of sea water and brackish water),
- suspended solids (in case of a turbid sample),
- possibly other relevant information about the sampling location at the time of sampling (e.g. actual or historical data on flow rate of rivers or marine currents, nearby major discharges and type of discharges, weather conditions preceding the sampling time),

and optionally:

- microbial biomass (e.g. acridine orange direct count or colony forming units),

**▼ M1**

- inorganic carbon,
- chlorophyll-a concentration as a specific estimate for algal biomass.

In addition, the following information on the sediment should be provided if the suspended sediment test is conducted:

- depth of sediment collection,
- appearance of the sediment (such as coloured, muddy, silty, or sandy),
- texture (e.g. % coarse sand, fine sand, silt and clay),
- dry weight in g/l of the suspended solids, TOC concentration or weight loss on ignition as a measure of the content of organic matter,
- pH,
- oxygen or redox potential (mandatory only if aerobic conditions are not obvious).

Test conditions:

- delay between collection and use in the laboratory test, sample storage and pre-treatment of the sample, dates of performance of the studies,
- amount of test substance applied, test concentration and reference substance,
- method of application of the test substance including any use of solvents,
- volume of surface water used and sediment (if used) and volume sampled at each interval for analysis,
- description of the test system used.

If dark conditions are not to be maintained, information on the 'diffuse light' conditions:

- information on the method(s) used for establishing sterile controls (e.g. temperature, time and number of autoclavings),
- incubation temperature,
- information on analytical techniques and the method(s) used for radiochemical measurements and for mass balance check and measurements of phase distribution (if conducted),
- number of replicates.

Results:

- percentages of recovery (see section 1.7.1),
- repeatability and sensitivity of the analytical methods used including the limit of detection (LOD) and the limit of quantification (LOQ) (see section 1.7.2),

**▼ M1**

- all measured data (including sampling time points) and calculated values in tabular form and the degradation curves; for each test concentration and for each replicate flask, report the linear correlation coefficient for the slope of the logarithmic plot, the estimated lag phase and a first-order or pseudo-first order rate constant (if possible), and the corresponding degradation half-life (or the half-life period,  $t_{50}$ ),
- report relevant values as the averages of the results observed in individual replicates, e.g. length of lag phase, degradation rate constant and degradation half-life (or  $t_{50}$ ),
- categorise the system as either non-adapted or adapted as judged from the appearance of the degradation curve and from the possible influence of the test concentration,
- the results of the final mass balance check and results on phase distribution measurements (if any),
- the fraction of  $^{14}\text{C}$  mineralised and, if specific analyses are used, the final level of primary degradation,
- the identification, molar concentration and percentage of applied and major transformation products (see first paragraph in section 1.8.9.5), where appropriate,
- a proposed pathway of transformation, where appropriate,
- discussion of results.

**4. LITERATURE**

1. OECD TG 309 (2004) Aerobic Mineralisation in surface water — Simulation Biodegradation Test.
2. ISO/DIS 14592-1 (1999) Water quality — Evaluation of the aerobic biodegradability of organic compounds at low concentrations — Part 1: Shake flask batch test with surface water or surface water/sediment suspensions.
3. Testing Method C.23. Aerobic and anaerobic transformation in soil.
4. Testing Method C.24. Aerobic and anaerobic transformation in aquatic sediments.
5. OECD (1993). Guidelines for the Testing of Chemicals. OECD, Paris.
6. ISO 8245 (1999). Water quality — Guidelines on the determination of total organic carbon (TOC) and dissolved organic carbon (DOC).
7. ISO 10634 (1995). Water quality — Guidance for the preparation and treatment of poorly water-soluble organic compounds for the subsequent evaluation of their biodegradability in an aqueous medium.
8. OECD (2000). Guidance Document on aquatic toxicity testing of difficult substances and mixtures. Environmental Health and Safety Publications. Series on Testing and Assessment. No 22.

**▼ M1**

9. Simkins, S. and Alexander, M. (1984). Models for mineralisation kinetics with the variables of substrate concentration and population density. *Appl. Environ. Microbiol.* 47, 394-401.
10. Ingerslev, F. and N. Nyholm. (2000). Shake-flask test for determination of biodegradation rates of <sup>14</sup>C-labeled chemicals at low concentrations in surface water systems. *Ecotoxicol. Environ. Saf.* 45, 274-283.
11. ISO/CD 14592-1 (1999). Ring test report: Water Quality — Evaluation of the aerobic biodegradability of organic compounds at low concentrations part 1 — report of 1998/1999 ring-test. Shake flask batch test with surface water or surface water/sediment suspensions.

**▼ M1***Appendix 1***Phase distribution of  $^{14}\text{C}$** 

In order to check the procedure, the routine measurements of residual total organic  $^{14}\text{C}$  activity (TOA) should be supplemented by mass balance measurements involving a direct determination of the evolved  $^{14}\text{CO}_2$  after trapping in an absorber (see Appendix 3). In itself, a positive  $^{14}\text{CO}_2$  formation is a direct evidence of biodegradation as opposed to abiotic degradation or other loss mechanisms, such as volatilisation and sorption. Additional useful information characterising the biodegradability behaviour can be obtained from measurements of the distribution of TOA between the dissolved state (dissolved organic  $^{14}\text{C}$  activity, DOA) and the particulate state (particulate organic  $^{14}\text{C}$  activity, POA) after separation of particulate by membrane filtration or centrifugation. POA consists of test substance sorbed onto the microbial biomass and onto other particles in addition to the test substance carbon that has been used for synthesis of new cellular material and thereby incorporated into the particulate biomass fraction. The formation of dissolved  $^{14}\text{C}$  organic material can be estimated as the DOA at the end of biodegradation (plateau on the degradation versus time curve).

Estimate the phase distribution of residual  $^{14}\text{C}$  in selected samples by filtering samples on a 0,22  $\mu\text{m}$  or 0,45  $\mu\text{m}$  membrane filter of a material that does not adsorb significant amounts of the test substance (polycarbonate filters may be suitable). If sorption of test substance onto the filter is too large to be ignored (to be checked prior to the experiment) high-speed centrifugation (2 000 g; 10 min.) can be used instead of filtration.

Proceed with the filtrate or centrifugate as described in Appendix 3 for unfiltered samples. Dissolve membrane filters in a suitable scintillation fluid and count as usually, normally using only the external standard ratio method to correct for quenching, or use a sample oxidiser. If centrifugation has been used, re-suspend the pellet formed of the particulate fraction in 1-2 ml of distilled water and transfer to a scintillation vial. Wash subsequently twice with 1 ml distilled water and transfer the washing water to the vial. If necessary, the suspension can be embedded in a gel for liquid scintillation counting.



▼ **M1***Appendix 2***Semi-continuous procedure**

Prolonged incubation for up to several months may be required in order to achieve a sufficient degradation of recalcitrant substances. The duration of the test should normally not exceed 60 days unless the characteristics of the original water sample are maintained by renewal of the test suspension. However, the test period may be extended to a maximum of 90 days without renewal of the test suspension, if the degradation of the test substance has started within the first 60 days.

During incubation for long periods, the diversity of the microbial community may be reduced due to various loss mechanisms and due to possible depletion of the water sample of essential nutrients and primary carbon substrates. It is therefore recommended that a semi-continuous test is used to adequately determine the degradation rate of slowly degrading substances. The test should be initiated by use of the semi-continuous procedure if, based on previous experience, an incubation period of three months is expected to be necessary to achieve 20 % degradation of the substance. Alternatively, the normal batch test may be changed into a semi-continuous test, if no degradation of the test substance has been achieved during approximately 60 days of testing using the batch procedure. The semi-continuous procedure may be stopped and the test continued as a batch experiment, when a substantial degradation has been recorded (e.g. > 20 %).

In the semi-continuous test, every two weeks, about one third of the volume of the test suspension is replaced by freshly collected water with the test substance added to the initial concentration. Sediment is likewise added to the replacement water to the initial concentration (between 0,01 and 1 g/l), if the optional suspended sediment test is performed. Carrying out the test with suspended sediment solids, it is important that a fully suspended system is maintained also during water renewal, and that the residence time is identical for solids and water, as otherwise the intended similarity to a homogenous aqueous system with no fixed phases can be lost. For these reasons, an initial concentration of suspended sediments in the lower range of the specified interval is preferred, when the semi-continuous procedure is used.

The prescribed addition of test substance implies that the initial concentration of test substance is not exceeded by the partial renewal of the test suspension and, hence, the adaptation, which is frequently seen with high concentrations of a test substance, is avoided. As the procedure comprises both a re-inoculation and a compensation of depleted nutrients and primary substrates, the original microbial diversity is restored, and the duration of the test can be extended to infinity in principle. When the semi-continuous procedure is used, it is important to note that the residual concentration of the test substance must be corrected for the amounts of test substance added and removed at each renewal procedure. The total and the dissolved test substance concentration can be used interchangeably for compounds that sorb little. Sorption is insignificant (< 5 %) under the specified conditions (0,1-1 g solids/l) for substances of  $\log K_{ow} < 3$  (valid for neutral, lipophilic compounds). This is illustrated by the following calculation example. 0,1 g/l of solids roughly corresponds to 10 mg of carbon per litre (fraction of carbon,  $f_C = 0,01$ ). Assuming that:

$\log K_{ow}$  (of the test substance) = 3

$$K_{oc} = 0,42 \times K_{ow}$$

$$\text{Partition coefficient, } K_d = f_C \times K_{oc}$$

then, the dissolved fraction of the total concentration (C-water ( $C_w$ )/C-total ( $C_t$ ) is:

$$C_w/C_t = 1/(1 + K_d \times SS) = 1/(1 + K_{oc} \times f_C \times SS) = 1/(1 + 0,42 \times 10^3 \times 0,01 \times 0,1 \times 10^{-3}) = 0,999$$

▼ **M1***Appendix 3***Determination of  $^{14}\text{CO}_2$** **Indirect  $^{14}\text{CO}_2$  determination**

For routine measurements, the indirect method is normally the least time-consuming and most precise method if the test substance is non-volatile and is not transformed into volatile transformation products. Simply transfer unfiltered samples e.g. 5 ml size to scintillation vials. A suitable activity in samples is 5 000 dpm-10 000 dpm (80-170 Bq) initially, and a minimum initial activity is about 1 000 dpm. The  $\text{CO}_2$  should be stripped off after acidifying to pH 2-3 with 1-2 drops of concentrated  $\text{H}_3\text{PO}_4$  or HCl. The  $\text{CO}_2$  stripping can be performed by bubbling with air for about 1/2-1 hour. Alternatively, vials can be shaken vigorously for 1-2 hours (for instance on a microplate shaker) or with more gentle shaking be left overnight. The efficiency of the  $\text{CO}_2$  stripping procedure must be checked (by prolonging the aeration or shaking period). A scintillation liquid, suitable for counting aqueous samples should then be added, the sample homogenised on a whirling mixer and the radioactivity determined by liquid scintillation counting, subtracting the background activity found in the test blanks ( $F_B$ ). Unless the test water is very coloured or contains a high concentration of particles, the samples will normally show uniform quenching and it will be sufficient to perform quench corrections using an external standard. If the test water is highly coloured, quench correction by means of internal standard addition may be necessary. If the concentration of particles is high it may not be possible to obtain a homogeneous solution or gel, or the quench variation between samples may be large. In that case the counting method described below for test slurries can be used. If the test is carried out as a suspended sediment test, the  $^{14}\text{CO}_2$  measurement could be done indirectly by taking a homogeneous 10-ml sample of the test water/suspension and separating the phases by centrifugation at a suitable speed (e.g. at 40 000  $\text{m/s}^2$  for 15 min.). The aqueous phase should then be then treated as described above. The  $^{14}\text{C}$  activity in the particulate phase (POA) should be determined by re-suspending the sediment into a small volume of distilled water, transferring to scintillation vials, and adding scintillation liquid to form a gel (special scintillation liquids are available for that purpose). Depending on the nature of particles (e.g. their content of organic material), it may be feasible to digest the sample overnight with a tissue solubiliser and then homogenise on a whirling mixer prior to the addition of scintillation liquid. Alternatively, the POA can be determined by combustion in excess of oxygen by use of a sample oxidiser. When counting, internal standards should always be included, and it may be necessary to perform quench corrections using internal standard addition for each individual sample.

**Direct  $^{14}\text{CO}_2$  determination**

If the evolved  $^{14}\text{CO}_2$  is measured directly, it should be done by setting up more flasks at the start of the test, harvesting the test flasks at each measuring point by acidifying the test flasks to pH 2-3 and collecting the  $^{14}\text{CO}_2$  in an internal (placed in each test flask at the start of the test) or external absorber. As absorbing medium either alkali (e.g. 1 N NaOH solution, or a NaOH pellet), ethanolamine or an ethanolamine-based, and commercially available absorbers can be used. For direct measurement of the  $^{14}\text{CO}_2$ , the flasks should be closed with e.g. butyl rubber septa.

▼ M1

Figure 1a

Example of arithmetic plot of data (residual activity versus time)

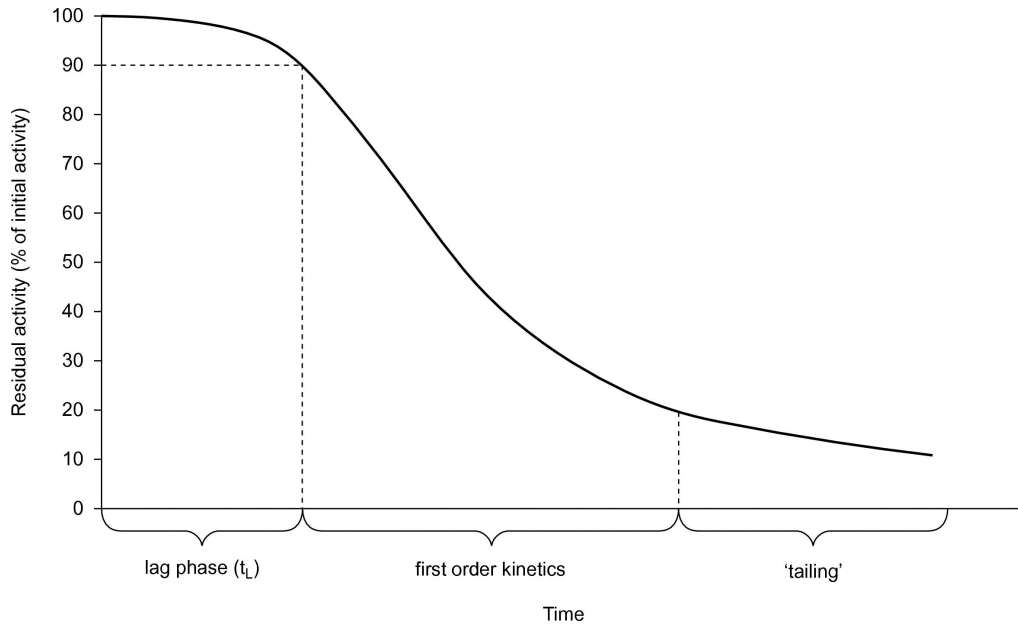
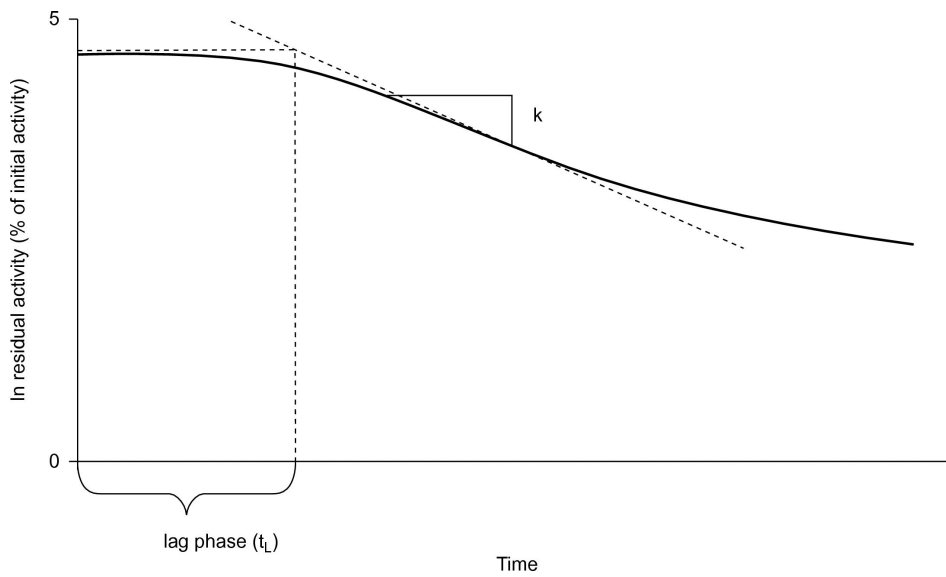


Figure 1b

Example of semi-logarithmic plot of data (ln to residual activity versus time)



**▼ M1****C.26. LEMNA SP. GROWTH INHIBITION TEST****1. METHOD**

This method is equivalent to OECD TG 221 (2006) (1). There has been broad agreement by EU authorities that the Lemna test is a suitable alternative to an algal test for strongly coloured substances (2)(3).

**1.1. INTRODUCTION**

This Testing Method is designed to assess the toxicity of substances to freshwater aquatic plants of the genus *Lemna* (duckweed). It is based on existing guidelines (4)(5)(6)(7)(8)(9) but includes modifications of those methods to reflect recent research and consultation on a number of key issues. The proposed method has been validated by an international ring-test (10).

This Testing Method describes toxicity testing using *Lemna gibba* and *Lemna minor*, both of which have been extensively studied and are the subject of the standards referred to above. The taxonomy of *Lemna* spp. is difficult, being complicated by the existence of a wide range of phenotypes. Although genetic variability in the response to toxicants can occur with *Lemna*, there are currently insufficient data on this source of variability to recommend a specific clone for use with this Testing Method. It should be noted that the test is not conducted axenically, but steps are taken at stages during the test procedure to keep contamination by other organisms to a minimum.

Details on testing with renewal (semi-static and flow-through) and without renewal (static) of the test solution are described. Depending on the objectives of the test and on the regulatory requirements, it is recommended to consider the application of semi-static and flow through methods, e.g. for substances that are rapidly lost from solution as a result of volatilisation, photodegradation, precipitation or biodegradation. Further guidance is given in (11).

**1.2. DEFINITIONS**

The following definitions and abbreviations are used for the purposes of this Testing Method:

**Biomass:** is the dry weight of living matter present in a population. In this test, surrogates for biomass, such as frond counts or frond area, are typically measured and the use of the term 'biomass' thus refers to these surrogate measures as well.

**Chlorosis:** is the yellowing of frond tissue.

**Clone:** is an organism or cell arisen from a single individual by asexual reproduction. Individuals from the same clone are, therefore, genetically identical.

**Colony:** means an aggregate of mother and daughter fronds (usually 2 to 4) attached to each other. Sometimes referred to as a plant.

**▼ M1**

**EC<sub>x</sub>**: is the concentration of the test substance dissolved in test medium that results in an x % (e.g. 50 %) reduction in growth of *Lemna* within a stated exposure period (to be mentioned explicitly if deviating from full or normal test duration). To unambiguously denote an EC value deriving from growth rate or yield, the symbol 'E<sub>r</sub>C' is used for growth rate and 'E<sub>y</sub>C' is used for yield, followed by the measurement variable used, e.g. E<sub>r</sub>C (frond number).

**Flow-through**: is a test in which the test solutions are replaced continuously.

**Frond**: is an individual/single 'leaf-like' structure of a duckweed plant. It is the smallest unit, i.e. individual, capable of reproduction.

**Gibbosity**: means fronds exhibiting a humped or swollen appearance.

**Growth**: is an increase in the measurement variable, e.g. frond number, dry weight, wet weight or frond area, over the test period.

**Growth rate (average specific growth rate)**: is the logarithmic increase in biomass during the exposure period.

**Lowest Observed Effect Concentration (LOEC)**: is the lowest tested concentration at which the substance is observed to have a statistically significant reducing effect on growth (at  $p < 0,05$ ) when compared with the control, within a given exposure time. However, all test concentrations above the LOEC must have a harmful effect equal to or greater than those observed at the LOEC. When these two conditions cannot be satisfied, a full explanation must be given for how the LOEC (and hence the NOEC) has been selected.

**Measurement variables**: are any type of variables which are measured to express the test endpoint using one or more different response variables. In this method frond number, frond area, fresh weight and dry weight are measurement variables.

**Monoculture**: is a culture with one plant species.

**Necrosis**: is dead (i.e. white or water-soaked) frond tissue.

**No Observed Effect Concentration (NOEC)**: is the test concentration immediately below the LOEC.

**Phenotype**: is the observable characteristics of an organism determined by the interaction of its genes with its environment.

**Response variables**: are variables for the estimation of toxicity derived from any measured variables describing biomass by different methods of calculation. For this method, growth rates and yield are response variables derived from measurement variables like frond number, frond area, fresh weight or dry weight.

**Semi-static (renewal) test**: is a test in which the test solution is periodically replaced at specific intervals during the test.

**Static test**: is a test method without renewal of the test solution during the test.

**▼ M1**

**Test endpoint:** describes the general factor that will be changed by the test chemical relative to the control as aim of the test. In this method the test endpoint is inhibition of growth, which may be expressed by different response variables which are based on one or more measurement variables.

**Test medium:** is the complete synthetic growth medium on which test plants grow when exposed to the test substance. The test substance will normally be dissolved in the test medium.

**Yield:** is the value of a measurement variable to express biomass at the end of the exposure period minus the measurement variable at the start of the exposure period.

### 1.3. PRINCIPLE OF THE TEST

Exponentially growing plant cultures of the genus *Lemma* are allowed to grow as monocultures in different concentrations of the test substance over a period of seven days. The objective of the test is to quantify substance-related effects on vegetative growth over this period, based on assessments of selected measurement variables. Frond number is the primary measurement variable. At least one other measurement variable (total frond area, dry weight or fresh weight) is also measured, since some substances may affect other measurement variables much more than frond numbers. To quantify substance-related effects, growth in the test solutions is compared with that of the controls and the concentration bringing about a specified x % inhibition of growth (e.g. 50 %) is determined and expressed as the  $EC_x$  (e.g.  $EC_{50}$ ).

The test endpoint is inhibition of growth, expressed as logarithmic increase in the measurement variable (average specific growth rate) during the exposure period. From the average specific growth rates recorded in a series of test solutions, the concentration bringing about a specified x % inhibition of growth rate (e.g. 50 %) is determined and expressed as the  $E_rC_x$  (e.g.  $E_rC_{50}$ ).

An additional response variable used in this Testing Method is yield, which may be needed to fulfil specific regulatory requirements in some countries. It is defined as the measurement variables at the end of the exposure period minus the measurement variables at the start of the exposure period. From the yield recorded in a series of test solutions, the concentration bringing about a specified x % inhibition of yield (e.g. 50 %) is calculated and expressed as the  $E_yC_x$  (e.g.  $E_yC_{50}$ ).

In addition, the lowest observed effect concentration (LOEC) and the no observed effect concentration (NOEC) may be statistically determined.

### 1.4. INFORMATION ON THE TEST SUBSTANCE

An analytical method, with adequate sensitivity for quantification of the substance in the test medium, should be available.

**▼ M1**

Information on the test substance which may be useful in establishing the test conditions includes the structural formula, purity, water solubility, stability in water and light,  $pK_a$ ,  $K_{ow}$ , vapour pressure and biodegradability. Water solubility and vapour pressure can be used to calculate Henry's Law constant, which will indicate if significant losses of the test substance during the test period are likely. This will help indicate whether particular steps to control such losses should be taken. Where information on the solubility and stability of the test substance is uncertain, it is recommended that these be assessed under the conditions of the test, i.e. growth medium, temperature, lighting regime to be used in the test.

When pH control of the test medium is particularly important, e.g. when testing metals or substances which are hydrolytically unstable, the addition of a buffer to the growth medium is recommended (see first paragraph in section 1.7.4). Further guidance for testing substances with physicochemical properties that make them difficult to test is provided in (11).

**1.5. REFERENCE SUBSTANCE**

Reference substance(s), such as 3,5-dichlorophenol used in the international ring test (10), may be tested as a means of checking the test procedure. It is advisable to test a reference substance at least twice a year or, where testing is carried out at a lower frequency, in parallel to the determination of the toxicity of a test substance.

**1.6. VALIDITY OF THE TEST**

For the test to be valid, the doubling time of frond number in the control must be less than 2,5 days (60 h), corresponding to approximately a seven-fold increase in seven days and an average specific growth rate of  $0,275\text{ d}^{-1}$ . Using the media and test conditions described in this Testing Method, this criterion can be attained using a static test regime (8). It is also anticipated that this criterion will be achievable under semi-static and flow-through test conditions. Calculation of the doubling time is shown in section 2.1.

**1.7. DESCRIPTION OF THE METHOD****1.7.1. Apparatus**

All equipment in contact with the test media should be made of glass or other chemically inert material. Glassware used for culturing and testing purposes should be cleaned of chemical contaminants that might leach into the test medium and should be sterile. The test vessels should be wide enough for the fronds of different colonies in the control vessels to grow without overlapping at the end of the test. It does not matter if the roots touch the bottoms of the test vessels, but a minimum depth of 20 mm and minimum volume of 100 ml in each test vessel is advised. The choice of test vessels is not critical as long as these requirements are met. Glass beakers, crystallising dishes or glass petri dishes of appropriate dimensions have all proved suitable. Test vessels must be covered to minimise evaporation and accidental contamination, while allowing necessary air exchange. Suitable test vessels, and particularly covers, must avoid shadowing or changes in the spectral characteristics of light.

**▼ M1**

The cultures and test vessels should not be kept together. This is best achieved using separate environmental growth chambers, incubators, or rooms. Illumination and temperature must be controllable and maintained at a constant level (see section 1.7.8).

**1.7.2. Test organism**

The organism used for this test is either *Lemna gibba* or *Lemna minor*. Short descriptions of duckweed species that have been used for toxicity testing are given in Appendix 1. Plant material may be obtained from a culture collection, another laboratory or from the field. If collected from the field, plants should be maintained in culture in the same medium as used for testing for a minimum of eight weeks prior to use. Field sites used for collecting starting cultures must be free of obvious sources of contamination. If obtained from another laboratory or from a culture collection they should be similarly maintained for a minimum of three weeks. The source of plant material and the species and clone (if known) used for testing should always be reported.

Monocultures, that are visibly free from contamination by other organisms such as algae and protozoa, should be used. Healthy plants of *L. minor* will consist of colonies comprising between two and five fronds, whilst healthy colonies of *L. gibba* may contain up to seven fronds.

The quality and uniformity of the plants used for the test will have a significant influence on the outcome of the test and should therefore be selected with care. Young, rapidly growing plants without visible lesions or discoloration (chlorosis) should be used. Good quality cultures are indicated by a high incidence of colonies comprising at least two fronds. A large number of single fronds is indicative of environmental stress, e.g. nutrient limitation, and plant material from such cultures should not be used for testing.

**1.7.3. Cultivation**

To reduce the frequency of culture maintenance (e.g. when no *Lemna* tests are planned for a period), cultures can be held under reduced illumination and temperature (4-10 °C). Details of culturing are given in Appendix 2. Obvious signs of contamination by algae or other organisms will require surface sterilisation of a sub-sample of *Lemna* fronds, followed by transfer to fresh medium (see Appendix 2). In this eventuality, the remaining contaminated culture should be discarded.

At least seven days before testing, sufficient colonies are transferred aseptically into fresh sterile medium and cultured for 7-10 days under the conditions of the test.

**1.7.4. Test medium**

Different media are recommended for *Lemna minor* and *Lemna gibba*, as described below. Careful consideration should be given to the inclusion of a pH buffer in the test medium (MOPS (4-morpholinepropane sulphonic acid, CAS No: 1132-61-2; EINECS No: 214-478-5) in *L. minor* medium and NaHCO<sub>3</sub> in *L. gibba* medium) when it is suspected that it might react with the test substance and influence the expression of its toxicity. Steinberg Medium (12) is also acceptable as long as the validity criteria are met.



**▼ M1**

A modification of the Swedish standard (SIS) *Lemna* growth medium is recommended for culturing and testing with *L. minor*. The composition of this medium is given in Appendix 3.

The growth medium, 20X — AAP, as described in Appendix 3, is recommended for culturing and testing with *L. gibba*.

Steinberg medium, as described in Appendix 3, is also suitable for *L. minor*, but may also be used for *L. gibba* as long as the validity criteria are met.

**1.7.5. Test solutions**

Test solutions are usually prepared by dilution of a stock solution. Stock solutions of the test substance are normally prepared by dissolving the substance in growth medium.

The highest tested concentration of the test substance should not normally exceed the water solubility of the substance under the test conditions. It should be noted however that *Lemna* spp. float on the surface and may be exposed to substances that collect at the water-air interface (e.g. poorly water-soluble or hydrophobic substances or surface-active substances). Under such circumstances, exposure will result from material other than in solution and test concentrations may, depending on the characteristics of the test substance, exceed water solubility. For test substances of low water solubility it may be necessary to prepare a concentrated stock solution or dispersion of the substance using an organic solvent or dispersant in order to facilitate the addition of accurate quantities of the test substance to the test medium and aid in its dispersion and dissolution. Every effort should be made to avoid the use of such materials. There should be no phytotoxicity resulting from the use of auxiliary solvents or dispersants. For example, commonly used solvents which do not cause phytotoxicity at concentrations up to 100 µl·l<sup>-1</sup> include acetone and dimethylformamide. If a solvent or dispersant is used, its final concentration should be reported and kept to a minimum ( $\leq 100 \mu\text{l}\cdot\text{l}^{-1}$ ), and all treatments and controls should contain the same concentration of solvent or dispersant. Further guidance on the use of dispersants is given in (11).

**1.7.6. Test and control groups**

Prior knowledge of the toxicity of the test substance to *Lemna*, e.g. from a range-finding test, will help in selecting suitable test concentrations. In the definitive toxicity test, there should normally be at least five test concentrations arranged in a geometric series. Preferably the separation factor between test concentrations should not exceed 3,2, but a larger value may be used where the concentration-response curve is flat. Justification should be provided if fewer than five concentrations are used. At least three replicates should be used at each test concentration.

In setting the range of test concentrations (for range-finding and/or for the definitive toxicity test), the following should be considered:

**▼ M1**

- To determine an  $EC_x$ , test concentrations should bracket the  $EC_x$  value to ensure an appropriate level of confidence. For example, if estimating the  $EC_{50}$ , the highest test concentration should be greater than the  $EC_{50}$  value. If the  $EC_{50}$  value lies outside the range of test concentrations, associated confidence intervals will be large and a proper assessment of the statistical fit of the model may not be possible.
  
- If the aim is to estimate the LOEC/NOEC, the lowest test concentration should be low enough so that growth is not significantly less than that of the control. In addition, the highest test concentration should be high enough so that growth is significantly lower than that in the control. If this is not the case, the test will have to be repeated using a different concentration range (unless the highest concentration is at the limit of solubility or the maximum required limit concentration, e.g.  $100 \text{ mg}\cdot\text{l}^{-1}$ ).

Every test should include controls consisting of the same nutrient medium, number of fronds and colonies, environmental conditions and procedures as the test vessels but without the test substance. If an auxiliary solvent or dispersant is used, an additional control treatment with the solvent/dispersant present at the same concentration as that in the vessels with the test substance should be included. The number of replicate control vessels (and solvent vessels, if applicable) should be at least equal to, ideally twice, the number of vessels used for each test concentration.

If determination of NOEC is not required, the test design may be altered to increase the number of concentrations and reduce the number of replicates per concentration. However, the number of control replicates must be at least three.

**1.7.7. Exposure**

Colonies consisting of 2 to 4 visible fronds are transferred from the inoculum culture and randomly assigned to the test vessels under aseptic conditions. Each test vessel should contain a total of 9 to 12 fronds. The number of fronds and colonies should be the same in each test vessel. Experience gained with this method and ring-test data have indicated that using three replicates per treatment, with each replicate containing 9 to 12 fronds initially, is sufficient to detect differences in growth of approximately 4 to 7 % of inhibition calculated by growth rate (10 to 15 % calculated by yield) between treatments (10).

A randomised design for location of the test vessels in the incubator is required to minimise the influence of spatial differences in light intensity or temperature. A blocked design or random repositioning of the vessels when observations are made (or repositioning more frequently) is also required.

**▼ M1**

If a preliminary stability test shows that the test substance concentration cannot be maintained (i.e. the measured concentration falls below 80 % of the measured initial concentration) over the test duration (7 days), a semi-static test regime is recommended. In this case, the colonies should be exposed to freshly prepared test and control solutions on at least two occasions during the test (e.g. days 3 and 5). The frequency of exposure to fresh medium will depend on the stability of the test substance; a higher frequency may be needed to maintain near-constant concentrations of highly unstable or volatile substances. In some circumstances, a flow-through procedure may be required (11)(13).

The exposure scenario through a foliar application (spray) is not covered in this Testing Method, instead see (14).

**1.7.8. Incubation conditions**

Continuous warm or cool white fluorescent lighting should be used to provide a light intensity selected from the range of 85-135  $\mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$  when measured in a photosynthetically active radiation (400-700 nm) at points the same distance from the light source as the *Lemna* fronds (equivalent to 6 500-10 000 lux). Any differences from the selected light intensity over the test area should not exceed  $\pm 15\%$ . The method of light detection and measurement, in particular the type of sensor, will affect the measured value. Spherical sensors (which respond to light from all angles above and below the plane of measurement) and 'cosine' sensors (which respond to light from all angles above the plane of measurement) are preferred to unidirectional sensors, and will give higher readings for a multi-point light source of the type described here.

The temperature in the test vessels should be  $24 \pm 2$  °C. The pH of the control medium should not increase by more than 1,5 units during the test. However, deviation of more than 1,5 units would not invalidate the test when it can be shown that validity criteria are met. Additional care is needed on pH drift in special cases such as when testing unstable substances or metals. See (11) for further guidance.

**1.7.9. Duration**

The test is terminated 7 days after the plants are transferred into the test vessels.

**1.7.10. Measurements and analytical determinations**

At the start of the test, the frond number in the test vessels is counted and recorded, taking care to ensure that protruding, distinctly visible fronds are accounted for. Frond numbers appearing normal or abnormal, need to be determined at the beginning of the test, at least once every 3 days during the exposure period (i.e. on at least 2 occasions during the 7 day period), and at test termination. Changes in plant development, e.g. in frond size, appearance, indication of necrosis, chlorosis or gibbosity, colony break-up or loss of buoyancy, and in root length and appearance, should be noted. Significant features of the test medium (e.g. presence of undissolved material, growth of algae in the test vessel) should also be noted.

**▼ M1**

In addition to determinations of frond number during the test, effects of the test substance on one (or more) of the following measurement variables are also assessed:

- (i) total frond area;
- (ii) dry weight;
- (iii) fresh weight.

Total frond area has an advantage that it can be determined for each test and control vessel at the start, during, and at the end of the test. Dry or fresh weight should be determined at the start of the test from a sample of the inoculum culture representative of what is used to begin the test, and at the end of the test with the plant material from each test and control vessel. If frond area is not measured, dry weight is preferred over fresh weight.

Total frond area, dry weight and fresh weight may be determined as follows:

- (i) *Total frond area*: The total frond area of all colonies may be determined by image analysis. A silhouette of the test vessel and plants can be captured using a video camera (i.e. by placing the vessel on a light box) and the resulting image digitised. By calibration with flat shapes of known area, the total frond area in a test vessel may then be determined. Care should be taken to exclude interference caused by the rim of the test vessel. An alternative but more laborious approach is to photocopy test vessels and plants, cut out the resulting silhouette of colonies and determine their area using a leaf area analyser or graph paper. Other techniques (e.g. paper weight ratio between silhouette area of colonies and unit area) may also be appropriate.
- (ii) *Dry weight*: All colonies are collected from each of the test vessels and rinsed with distilled or deionised water. They are blotted to remove excess water and then dried at 60 °C to a constant weight. Any root fragments should be included. The dry weight should be expressed to an accuracy of at least 0,1 mg.
- (iii) *Fresh weight*: All colonies are transferred to pre-weighed polystyrene (or other inert material) tubes with small (1 mm) holes in the rounded bottoms. The tubes are then centrifuged at 3 000 rpm for 10 minutes at room temperature. Tubes, containing the now dried colonies, are re-weighed and the fresh weight is calculated by subtracting the weight of the empty tube.

1.7.10.1. *Frequency of measurements and analytical determinations*

If a static test design is used, the pH of each treatment should be measured at the beginning and end of the test. If a semi-static test design is used, the pH should be measured in each batch of 'fresh' test solution prior to each renewal and also in the corresponding 'spent' solutions.

**▼ M1**

Light intensity should be measured in the growth chamber, incubator or room at points the same distance from the light source as the *Lemna* fronds. Measurements should be made at least once during the test. The temperature of the medium in a surrogate vessel held under the same conditions in the growth chamber, incubator or room should be recorded at least daily.

During the test, the concentrations of the test substance are determined at appropriate intervals. In static tests, the minimum requirement is to determine the concentrations at the beginning and the end of the test.

In semi-static tests where the concentration of the test substance is not expected to remain within  $\pm 20\%$  of the nominal concentration, it is necessary to analyse all freshly prepared test solutions and the same solutions at each renewal (see third paragraph in section 1.7.7). However, for those tests where the measured initial concentration of the test substance is not within  $\pm 20\%$  of nominal, but where sufficient evidence can be provided to show that the initial concentrations are repeatable and stable (i.e. within the range 80-120 % of the initial concentration), chemical determinations may be carried out on only the highest and lowest test concentrations. In all cases, determination of test substance concentrations prior to renewal need only be performed on one replicate vessel at each test concentration (or the contents of the vessels pooled by replicate).

If a flow-through test is used, a similar sampling regime to that described for semi-static tests, including analysis at the start, mid-way through and at the end of the test, is appropriate, but measurement of 'spent' solutions is not appropriate in this case. In this type of test, the flow-rate of diluent and test substance or test substance stock solution should be checked daily.

If there is evidence that the concentration of the substance being tested has been satisfactorily maintained within  $\pm 20\%$  of the nominal or measured initial concentration throughout the test, analysis of the results can be based on nominal or measured initial values. If the deviation from the nominal or measured initial concentration is greater than  $\pm 20\%$ , analysis of the results should be based on the geometric mean concentration during exposure or models describing the decline of the concentration of the test substance (11).

**1.7.11. Limit test**

Under some circumstances, e.g. when a preliminary test indicates that the test substance has no toxic effects at concentrations up to  $100\text{ mg}\cdot\text{l}^{-1}$ , or up to its limit of solubility in the test medium (whichever is the lower), a limit test involving a comparison of responses in a control group and one treatment group ( $100\text{ mg}\cdot\text{l}^{-1}$  or a concentration equal to the limit of solubility), may be undertaken. It is strongly recommended that this be supported by analysis of the exposure concentration. All previously described test conditions and validity criteria apply to a limit test, with the exception that the number of treatment replicates should be doubled. Growth in the control and treatment group may be analysed using a statistical test to compare means, e.g. a Student's t-test.

**▼ M1****2. DATA AND REPORTING****2.1. DOUBLING TIME**

To determine the doubling time ( $T_d$ ) of frond number and adherence to this validity criterion by the study (section 1.6), the following formula is used with data obtained from the control vessels:

$$T_d = \ln 2/\mu$$

where  $\mu$  is the average specific growth rate determined as described in first and second paragraph in section 2.2.1.

**2.2. RESPONSE VARIABLES**

The purpose of the test is to determine the effects of the test substance on the vegetative growth of *Lemna*. This Testing Method describes two response variables, as member countries have different preferences and regulatory needs. In order for the test results to be acceptable in all member countries, the effects should be evaluated using both response variables (a) and (b) described below.

- (a) Average specific growth rate: this response variable is calculated on the basis of changes in the logarithms of frond numbers, and in addition, on the basis of changes in the logarithms of another measurement parameter (total frond area, dry weight or fresh weight) over time (expressed per day) in the controls and each treatment group. It is sometimes referred to as relative growth rate (15).
- (b) Yield: this response variable is calculated on the basis of changes in frond number, and in addition, on the basis of changes in another measurement parameter (total frond area, dry weight or fresh weight) in the controls and in each treatment group until the end of the test.

It should be noted that toxicity values calculated by using these two response variables are not comparable and this difference must be recognised when using the results of the test.  $EC_x$  values based upon average specific growth rate ( $E_rC_x$ ) will generally be higher than results based upon yield ( $E_yC_x$ ) if the test conditions of this Testing Method are adhered to, due to the mathematical basis of the respective approaches. This should not be interpreted as a difference in sensitivity between the two response variables, simply that the values are different mathematically. The concept of average specific growth rate is based on the general exponential growth pattern of duckweed in non-limited cultures, where toxicity is estimated on the basis of the effects on the growth rate, without being dependent on the absolute level of the specific growth rate of the control, slope of the concentration-response curve or on test duration. In contrast, results based upon the yield response variable are dependent upon all these other variables.  $E_yC_x$  is dependent on the specific growth rate of the duckweed species used in each test and on the maximum specific growth rate that can vary between species and even different clones. This response variable should not be used for comparing the sensitivity to toxicants among duckweed species or even different clones. While the use of average specific growth rate for estimating toxicity is scientifically preferred, toxicity estimates based on yield are also included in this Testing Method to satisfy current regulatory requirements in some countries.

**▼ M1**

Toxicity estimates should be based on frond number and on one additional measurement variable (total frond area, dry weight or fresh weight), because some substances may affect other measurement variables much more than the frond number. This effect would not be detected by calculating frond number only.

The number of fronds as well as any other recorded measurement variable, i.e. total frond area, dry weight or fresh weight, are tabulated together with the concentrations of the test substance for each measurement occasion. Subsequent data analysis e.g. to estimate a LOEC, NOEC or EC<sub>x</sub> should be based on the values for the individual replicates and not calculated means for each treatment group.

### 2.2.1. Average specific growth rate

The average specific growth rate for a specific period is calculated as the logarithmic increase in the growth variables — frond numbers and one other measurement variable (total frond area, dry weight or fresh weight) — using the formula below for each replicate of control and treatments:

$$\mu_{i-j} = \frac{\ln(N_j) - \ln(N_i)}{t}$$

where:

- $\mu_{i-j}$ : average specific growth rate from time i to j
- $N_i$ : measurement variable in the test or control vessel at time i
- $N_j$ : measurement variable in the test or control vessel at time j
- t: time period from i to j

For each treatment group and control group, calculate a mean value for the growth rate along with variance estimates.

The average specific growth rate should be calculated for the entire test period (time 'i' in the above formula is the beginning of the test and time 'j' is the end of the test). For each test concentration and control, calculate a mean value for average specific growth rate along with the variance estimates. In addition, the section-by-section growth rate should be assessed in order to evaluate effects of the test substance occurring during the exposure period (e.g. by inspecting log-transformed growth curves). Substantial differences between the section-by-section growth rate and the average growth rate indicate deviation from constant exponential growth and that close examination of the growth curves is warranted. In this case, a conservative approach would be to compare specific growth rates from treated cultures during the time period of maximum inhibition to those for controls during the same time period.

**▼ M1**

Percentage inhibition of growth rate ( $I_r$ ) may then be calculated for each test concentration (treatment group) according to the following formula:

$$\% I_r = \frac{(\mu_C - \mu_T)}{\mu_C} \times 100$$

where:

- $\% I_r$ : percentage inhibition in average specific growth rate
- $\mu_C$ : mean value for  $\mu$  in the control
- $\mu_T$ : mean value for  $\mu$  in the treatment group

**2.2.2. Yield**

Effects on yield are determined on the basis of two measurement variables, frond number and one other measurement variable (total frond area, dry weight or fresh weight) present in each test vessel at the start and end of the test. For dry weight or fresh weight, the starting biomass is determined on the basis of a sample of fronds taken from the same batch used to inoculate the test vessels (see second paragraph in section 1.7.3). For each test concentration and control, calculate a mean value for yield along with variance estimates. The mean percentage inhibition in yield ( $\% I_y$ ) may be calculated for each treatment group as follows:

$$\% I_y = \frac{(b_C - b_T)}{b_C} \times 100$$

where:

- $\% I_y$ : percentage reduction in yield
- $b_C$ : final biomass minus starting biomass for the control group
- $b_T$ : final biomass minus starting biomass in the treatment group

**2.2.3. Plotting concentration-response curves**

Concentration-response curves relating mean percentage inhibition of the response variable ( $I_r$ , or  $I_y$  calculated as shown in the last paragraph of section 2.2.1 or in section 2.2.2) and the log concentration of the test substance should be plotted.

**2.2.4. EC<sub>x</sub> estimation**

Estimates of the EC<sub>x</sub> (e.g. EC<sub>50</sub>) should be based upon both average specific growth rate ( $E_r C_x$ ) and yield ( $E_y C_x$ ), each of which should in turn be based upon frond number and one additional measurement variable (total frond area, dry weight, or fresh weight). This is because there are test substances that impact frond number and other measurement variables differently. The desired toxicity parameters are therefore four EC<sub>x</sub> values for each inhibition level  $x$  calculated:  $E_r C_x$  (frond number);  $E_r C_x$  (total frond area, dry weight, or fresh weight);  $E_y C_x$  (frond number); and  $E_y C_x$  (total frond area, dry weight, or fresh weight).



**▼ M1****2.3. STATISTICAL PROCEDURES**

The aim is to obtain a quantitative concentration-response relationship by regression analysis. It is possible to use a weighted linear regression after having performed a linearising transformation of the response data — for instance into probit or logit or Weibull units (16), but non-linear regression procedures are preferred techniques that better handle unavoidable data irregularities and deviations from smooth distributions. Approaching either zero or total inhibition, such irregularities may be magnified by the transformation, interfering with the analysis (16). It should be noted that standard methods of analysis using probit, logit, or Weibull transforms are intended for use on quantal (e.g. mortality or survival) data, and must be modified to accommodate growth rate or yield data. Specific procedures for determination of  $EC_x$  values from continuous data can be found in (17), (18), and (19).

For each response variable to be analysed, use the concentration-response relationship to calculate point estimates of  $EC_x$  values. When possible, the 95 % confidence limits for each estimate should be determined. Goodness of fit of the response data to the regression model should be assessed either graphically or statistically. Regression analysis should be performed using individual replicate responses, not treatment group means.

$EC_{50}$  estimates and confidence limits may also be obtained using linear interpolation with bootstrapping (20), if available regression models/methods are unsuitable for the data.

For estimation of the LOEC and hence the NOEC, it is necessary to compare treatment means using analysis of variance (ANOVA) techniques. The mean for each concentration must then be compared with the control mean using an appropriate multiple comparison or trend test method. Dunnett's or Williams' test may be useful (21)-(22)(23)(24). It is necessary to assess whether the ANOVA assumption of homogeneity of variance holds. This assessment may be performed graphically or by a formal test (25). Suitable tests are Levene's or Bartlett's. Failure to meet the assumption of homogeneity of variances can sometimes be corrected by logarithmic transformation of the data. If heterogeneity of variance is extreme and cannot be corrected by transformation, analysis by methods such as step-down Jonkheere trend tests should be considered. Additional guidance on determining the NOEC can be found in (19).

Recent scientific developments have led to a recommendation of abandoning the concept of NOEC and replacing it with regression based point estimates  $EC_x$ . An appropriate value for  $x$  has not been established for this *Lemna* test. However, a range of 10 to 20 % appears to be appropriate (depending on the response variable chosen), and preferably both the  $EC_{10}$  and  $EC_{20}$  should be reported.

**3. REPORTING****3.1. TEST REPORT**

The test report must include the following:

Test substance:

— physical nature and physicochemical properties, including water solubility limit,

**▼ M1**

- chemical identification data (e.g. CAS Number), including purity.

Test species:

- scientific name, clone (if known) and source.

Test conditions:

- test procedure used (static, semi-static or flow-through),
- date of start of the test and its duration,
- test medium,
- description of the experimental design: test vessels and covers, solution volumes, number of colonies and fronds per test vessel at the beginning of the test,
- test concentrations (nominal and measured as appropriate) and number of replicates per concentration,
- methods of preparation of stock and test solutions including the use of any solvents or dispersants,
- temperature during the test,
- light source, light intensity and homogeneity,
- pH values of the test and control media,
- test substance concentrations and the method of analysis with appropriate quality assessment data (validation studies, standard deviations or confidence limits of analyses),
- methods for determination of frond number and other measurement variables, e.g. dry weight, fresh weight or frond area,
- all deviations from this Testing Method.

Results:

- raw data: number of fronds and other measurement variables in each test and control vessel at each observation and occasion of analysis,
- means and standard deviations for each measurement variable,
- growth curves for each concentration (recommended with log transformed measurement variable, see second paragraph in section 2.2.1),
- doubling time/growth rate in the control based on the frond number,
- calculated response variables for each treatment replicate, with mean values and coefficient of variation for replicates,
- graphical representation of the concentration/effect relationship,
- estimates of toxic endpoints for response variables e.g. EC<sub>50</sub>, EC<sub>10</sub>, EC<sub>20</sub>, and associated confidence intervals. If calculated, LOEC and/or NOEC and the statistical methods used for their determination,

▼ **M1**

- if ANOVA has been used, the size of the effect which can be detected (e.g. the least significant difference),
- any stimulation of growth found in any treatment,
- any visual signs of phytotoxicity as well as observations of test solutions,
- discussion of the results, including any influence on the outcome of the test resulting from deviations from this Testing Method.

4. **LITERATURE**

- (1) OECD TG 221 (2006) Lemna Sp. Growth Inhibition Test.
- (2) The use of Lemna studies for coloured substances is detailed in Section 13.5.3 of the EU Manual of Decisions dated July 2006, at <http://ecb.jrc.ec.europa.eu/new-chemicals>
- (3) Guidance on information requirements and chemical safety assessment — Chapter R.7b: Endpoint specific guidance; Table 7.8.3 Summary of difficult substance testing issues, available at:  
  
[http://guidance.echa.europa.eu/docs/guidance\\_document/-information\\_requirements\\_en.htm?time=1234958685#A](http://guidance.echa.europa.eu/docs/guidance_document/-information_requirements_en.htm?time=1234958685#A)
- (4) ASTM International. (2003). Standard Guide for Conducting Static Toxicity Test With *Lemna gibba* G3. E 1415-91 (Reapproved 1998). pp. 733-742. In, Annual Book of ASTM Standards, Vol. 11.05 Biological Effects and Environmental Fate; Biotechnology; Pesticides, ASTM, West Conshohocken, PA.
- (5) USEPA — United States Environmental Protection Agency. (1996). OPPTS 850.4400 Aquatic Plant Toxicity Test Using Lemna spp., 'Public draft'. EPA 712-C-96-156. 8pp.
- (6) AFNOR — Association Française de Normalisation. (1996). XP T 90-337: Détermination de l'inhibition de la croissance de *Lemna minor*. 10pp.
- (7) SSI — Swedish Standards Institute. (1995). Water quality — Determination of growth inhibition (7-d) *Lemna minor*, duckweed. SS 02 82 13. 15pp. (in Swedish).
- (8) Environment Canada. (1999). Biological Test Method: Test for Measuring the Inhibition of Growth Using the Freshwater Macrophyte, *Lemna minor*. EPS 1/RM/37-120 pp.
- (9) Environment Canada. (1993) Proposed Guidelines for Registration of Chemical Pesticides: Non-Target Plant Testing and Evaluation. Canadian Wildlife Service, Technical Report Series No. 145.
- (10) Sims I., Whitehouse P., and Lacey R. (1999). The OECD Lemna Growth Inhibition Test. Development and Ring-testing of draft OECD Test Guideline. R&D Technical Report EMA 003. WRc plc — Environment Agency.
- (11) OECD (2000). Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures. OECD Environmental Health and Safety Publications, Series on Testing and Assessment No 23.
- (12) ISO DIS 20079. Water Quality — Determination of the Toxic Effect of Water Constituents and Waste Water to Duckweed (*Lemna minor*) — Duckweed Growth Inhibition Test.

▼ M1

- (13) Walbridge C. T. (1977). A flow-through testing procedure with duckweed (*Lemna minor* L.). Environmental Research Laboratory — Duluth, Minnesota 55804. US EPA Report No. EPA-600/3-77 108. September 1977.
- (14) Lockhart W. L., Billeck B. N. and Baron C. L. (1989). Bioassays with a floating plant (*Lemna minor*) for effects of sprayed and dissolved glyphosate. *Hydrobiologia*, 118/119, 353-359.
- (15) Huebert, D.B. and Shay J.M. (1993). Considerations in the assessment of toxicity using duckweeds. *Environmental Toxicology and Chemistry*, 12, 481-483.
- (16) Christensen, E.R., Nyholm, N. (1984): Ecotoxicological Assays with Algae: Weibull Dose-Response Curves. *Env. Sci. Technol.* 19, 713-718.
- (17) Nyholm, N. Sørensen, P.S., Kusk, K.O. and Christensen, E.R. (1992): Statistical treatment of data from microbial toxicity tests. *Environ. Toxicol. Chem.* 11, 157-167.
- (18) Bruce R.D. and Versteeg D.J. (1992). A statistical procedure for modelling continuous toxicity data. *Environmental Toxicology and Chemistry*, 11, 1485-1494.
- (19) OECD. (2004). Guidance Document on Statistical Analysis of Ecotoxicity Data.
- (20) Norberg-King T.J. (1988). An interpolation estimate for chronic toxicity: The ICp approach. National Effluent Toxicity Assessment Center Technical Report 05-88. USEPA, Duluth, MN.
- (21) Dunnett, C.W. (1955). A multiple comparisons procedure for comparing several treatments with a control. *J. Amer. Statist. Assoc.*, 50, 1096-1121.
- (22) Dunnett, C.W. (1964). New tables for multiple comparisons with a control. *Biometrics*, 20, 482-491.
- (23) Williams, D.A. (1971). A test for differences between treatment means when several dose levels are compared with a zero dose control. *Biometrics*, 27: 103-117.
- (24) Williams, D.A. (1972). The comparison of several dose levels with a zero dose control. *Biometrics*, 28: 510-531.
- (25) Brain P. and Cousens R. (1989). An equation to describe dose-responses where there is stimulation of growth at low doses. *Weed Research*, 29, 93-96.

▼ M1

## Appendix 1

**Description of *lemna* spp.**

The aquatic plant commonly referred to as duckweed, *Lemna* spp., belongs to the family *Lemnaceae* which has a number of worldwide species in four genera. Their different appearance and taxonomy have been exhaustively described (1)(2). *Lemna gibba* and *L. minor* are species representative of temperate areas and are commonly used for toxicity tests. Both species have a floating or submerged discoid stem (frond) and a very thin root emanates from the centre of the lower surface of each frond. *Lemna* spp. rarely produce flowers and the plants reproduce by vegetatively producing new fronds (3). In comparison with older plants, the younger ones tend to be paler, have shorter roots and consist of two to three fronds of different sizes. The small size of *Lemna*, its simple structure, asexual reproduction and short generation time makes plants of this genus very suitable for laboratory testing (4)(5).

Because of probable interspecies variation in sensitivity, only comparisons of sensitivity within a species are valid.

**Examples of *Lemna* species which have been used for testing: Species Reference**

*Lemna aequinoctialis*: Eklund, B. (1996). The use of the red alga *Ceramium strictum* and the duckweed *Lemna aequinoctialis* in aquatic ecotoxicological bioassays. Licentiate in Philosophy Thesis 1996:2. Dep. of Systems Ecology, Stockholm University.

*Lemna major*: Clark, N. A. (1925). The rate of reproduction of *Lemna major* as a function of intensity and duration of light. J. phys. Chem., 29: 935-941.

*Lemna minor*: United States Environmental Protection Agency (USEPA). (1996). OPPTS 850.4400 Aquatic Plant Toxicity Test Using *Lemna* spp., 'Public draft'. EPA 712-C-96-156. 8pp.

Association Française de Normalisation (AFNOR). (1996). XP T 90-337: Détermination de l'inhibition de la croissance de *Lemna minor*. 10pp.

Swedish Standards Institute (SIS). (1995). Water quality — Determination of growth inhibition (7-d) *Lemna minor*, duckweed. SS 02 82 13. 15pp. (in Swedish).

*Lemna gibba*: ASTM International. (2003). Standard Guide for Conducting Static Toxicity Test With *Lemna gibba* G3. E 1415-91 (Reapproved 1998). pp. 733-742.

United States Environmental Protection Agency (USEPA). (1996). OPPTS 850.4400 Aquatic Plant Toxicity Test Using *Lemna* spp., 'Public draft'. EPA 712-C-96-156. 8pp.

*Lemna paucicostata*: Nasu, Y., Kugimoto, M. (1981). *Lemna* (duckweed) as an indicator of water pollution. I. The sensitivity of *Lemna paucicostata* to heavy metals. Arch. Environ. Contam. Toxicol., 10:1959-1969.

*Lemna perpusilla*: Clark, J. R. et al. (1981). Accumulation and depuration of metals by duckweed (*Lemna perpusilla*). Ecotoxicol. Environ. Saf., 5:87-96.

*Lemna trisulca*: Huebert, D. B., Shay, J. M. (1993). Considerations in the assessment of toxicity using duckweeds. Environ. Toxicol. and Chem., 12:481-483.

**▼ M1**

*Lemna valdiviana*: Hutchinson, T.C., Czyrska, H. (1975). Heavy metal toxicity and synergism to floating aquatic weeds. Verh.-Int. Ver. Limnol., 19:2102-2111.

**Sources of Lemna species**

University of Toronto Culture Collection of Algae and Cyanobacteria  
Department of Botany, University of Toronto  
Toronto, Ontario, Canada, M5S 3 B2  
Tel. +1-416-978-3641

Fax +1-416-978-5878  
e-mail: jacreman@botany.utoronto.ca  
<http://www.botany.utoronto.ca/utcc>

North Carolina State University  
Forestry Dept  
Duckweed Culture Collection  
Campus Box 8002  
Raleigh, NC 27695-8002  
UNITED STATES  
Tel. 001 (919) 515-7572  
[astomp@unity.ncsu.edu](mailto:astomp@unity.ncsu.edu)

Institute of Applied Environmental Research (ITM) Stockholm University  
SE-106 91 Stockholm  
SWEDEN  
Tel. +46 86747240  
Fax +46 86747636

Federal Environmental Agency (UBA)  
FG III 3.4  
Schichauweg 58  
12307 Berlin  
GERMANY  
e-mail: [lemna@uba.de](mailto:lemna@uba.de)  
<http://www.umweltbundesamt.de/contact.htm>

**Literature**

- (1) Hillman, W.S. (1961). The *Lemnaceae* or duckweeds: A review of the descriptive and experimental literature. *The Botanical Review*, 27:221-287.
- (2) Landolt, E. (1986). Biosystematic investigations in the family of duckweed (*Lemnaceae*). Vol. 2. Geobotanischen Inst. ETH, Stiftung Rubel, Zürich, Switzerland.
- (3) Björndahl, G. (1982). Growth performance, nutrient uptake and human utilization of duckweeds (*Lemnaceae* family). ISBN 82-991150-0-0. The Agricultural Research Council of Norway, University of Oslo.
- (4) Wang, W. (1986). Toxicity tests of aquatic pollutants by using common duckweed. *Environmental Pollution, Ser B*, 11:1-14.
- (5) Wang, W. (1990). Literature review on duckweed toxicity testing. *Environmental Research*, 52:7-22.

**▼ M1***Appendix 2***Maintenance of stock culture**

Stock cultures can be maintained under lower temperatures (4-10 °C) for longer times without needing to be re-established. The *Lemna* growth medium may be the same as that used for testing but other nutrient rich media can be used for stock cultures.

Periodically, a number of young, light-green plants are removed to new culture vessels containing fresh medium using an aseptic technique. Under the cooler conditions suggested here, sub-culturing may be conducted at intervals of up to three months.

Chemically clean (acid-washed) and sterile glass culture vessels should be used and aseptic handling techniques employed. In the event of contamination of the stock culture e.g. by algae or fungi, steps are necessary to eliminate the contaminating organisms. In the case of algae and most other contaminating organisms, this can be achieved by surface sterilisation. A sample of the contaminated plant material is taken and the roots cut off. The material is then shaken vigorously in clean water, followed by immersion in a 0,5 % (v/v) sodium hypochlorite solution for between 30 seconds and 5 minutes. The plant material is then rinsed with sterile water and transferred, as a number of batches, into culture vessels containing fresh growth medium. Many fronds will die as a result of this treatment, especially if longer exposure periods are used, but some of those surviving will usually be free of contamination. These can then be used to re-inoculate new cultures.

▼ **M1**

## Appendix 3

**Media**

Different growth media are recommended for *L. minor* and *L. gibba*. For *L. minor*, a modified Swedish Standard (SIS) medium is recommended whilst for *L. gibba*, 20X AAP medium is recommended. Compositions of both media are given below. When preparing these media, reagent or analytical-grade chemicals and deionised water should be used.

**Swedish Standard (SIS) Lemna growth medium**

- Stock solutions I-V are sterilised by autoclaving (120 °C, 15 minutes) or by membrane filtration (approximately 0,2 µm pore size).
- Stock VI (and optionally VII) are sterilised by membrane filtration only; these should not be autoclaved.
- Sterile stock solutions should be stored under cool and dark conditions. Stocks I-V should be discarded after six months whilst stocks VI (and optionally VII) have a shelf life of one month.

Stock solution No.	Substance	Concentration in stock solution (g·l <sup>-1</sup> )	Concentration in prepared medium (mg·l <sup>-1</sup> )	Prepared medium	
				Element	Concentration (mg·l <sup>-1</sup> )
I	NaNO <sub>3</sub>	8,50	85	Na; N	32; 14
	KH <sub>2</sub> PO <sub>4</sub>	1,34	13,4	K; P	6,0; 2,4
II	MgSO <sub>4</sub> · 7H <sub>2</sub> O	15	75	Mg; S	7,4; 9,8
III	CaCl <sub>2</sub> · 2H <sub>2</sub> O	7,2	36	Ca; Cl	9,8; 17,5
IV	Na <sub>2</sub> CO <sub>3</sub>	4,0	20	C	2,3
V	H <sub>3</sub> BO <sub>3</sub>	1,0	1,00	B	0,17
	MnCl <sub>2</sub> · 4H <sub>2</sub> O	0,20	0,20	Mn	0,056
	Na <sub>2</sub> MoO <sub>4</sub> · 2H <sub>2</sub> O	0,010	0,010	Mo	0,0040
	ZnSO <sub>4</sub> · 7H <sub>2</sub> O	0,050	0,050	Zn	0,011
	CuSO <sub>4</sub> · 5H <sub>2</sub> O	0,0050	0,0050	Cu	0,0013
	Co(NO <sub>3</sub> ) <sub>2</sub> · 6H <sub>2</sub> O	0,010	0,010	Co	0,0020
VI	FeCl <sub>3</sub> · 6H <sub>2</sub> O	0,17	0,84	Fe	0,17
	Na <sub>2</sub> -EDTA·2H <sub>2</sub> O	0,28	1,4	—	—
VII	MOPS (buffer)	490	490	—	—

- To prepare one litre of SIS medium, the following are added to 900 ml of deionised water:
  - 10 ml of stock solution I
  - 5 ml of stock solution II
  - 5 ml of stock solution III
  - 5 ml of stock solution IV



**▼M1**

- 1 ml of stock solution V
- 5 ml of stock solution VI
- 1 ml of stock solution VII (optional)

*Note:* A further stock solution VII (MOPS buffer) may be needed for certain test substances (see last paragraph in section 1.4).

- The pH is adjusted to  $6,5 \pm 0,2$  with either 0,1 or 1 mol HCl or NaOH, and the volume adjusted to one litre with deionised water.

**20X AAP growth medium**

Stock solutions are prepared in sterile distilled or deionised water.

Sterile stock solutions should be stored under cool and dark conditions. Under these conditions the stock solutions will have a shelf life of at least 6-8 weeks.

Five nutrient stock solutions (A1, A2, A3, B and C) are prepared for 20X — AAP medium, using reagent-grade chemicals. 20 ml of each nutrient stock solution are added to approximately 850 ml deionised water to produce the growth medium. The pH is adjusted to  $7,5 \pm 0,1$  with either 0,1 or 1 mol HCl or NaOH, and the volume adjusted to one litre with deionised water. The medium is then filtered through a 0,2  $\mu\text{m}$  (approximate) membrane filter into a sterile container.

The growth medium intended for testing should be prepared 1-2 days before use to allow the pH to stabilise. The pH of the growth medium should be checked prior to use and readjusted if necessary by addition of 0,1 or 1 M NaOH or HCl as described above.

Stock solution No.	Substance	Concentration in stock solution ( $\text{g}\cdot\text{l}^{-1}$ ) (*)	Concentration in prepared medium ( $\text{mg}\cdot\text{l}^{-1}$ ) (*)	Prepared medium	
				Element	Concentration ( $\text{mg}\cdot\text{l}^{-1}$ ) (*)
A1	$\text{NaNO}_3$	26	510	Na; N	190; 84
	$\text{MgCl}_2\cdot 6\text{H}_2\text{O}$	12	240	Mg	58,08
	$\text{CaCl}_2\cdot 2\text{H}_2\text{O}$	4,4	90	Ca	24,04
A2	$\text{MgSO}_4\cdot 7\text{H}_2\text{O}$	15	290	S	38,22
A3	$\text{K}_2\text{HPO}_4\cdot 3\text{H}_2\text{O}$	1,4	30	K; P	9,4;3,7
B	$\text{H}_3\text{BO}_3$	0,19	3,7	B	0,65
	$\text{MnCl}_2\cdot 4\text{H}_2\text{O}$	0,42	8,3	Mn	2,3
	$\text{FeCl}_3\cdot 6\text{H}_2\text{O}$	0,16	3,2	Fe	0,66
	$\text{Na}_2\text{EDTA}\cdot 2\text{H}_2\text{O}$	0,30	6,0	—	—
	$\text{ZnCl}_2$	$3,3 \text{ mg}\cdot\text{l}^{-1}$	$66 \mu\text{g}\cdot\text{l}^{-1}$	Zn	$31 \mu\text{g}\cdot\text{l}^{-1}$
	$\text{CoCl}_2\cdot 6\text{H}_2\text{O}$	$1,4 \text{ mg}\cdot\text{l}^{-1}$	$29 \mu\text{g}\cdot\text{l}^{-1}$	Co	$7,1 \mu\text{g}\cdot\text{l}^{-1}$
	$\text{Na}_2\text{MoO}_4\cdot 2\text{H}_2\text{O}$	$7,3 \text{ mg}\cdot\text{l}^{-1}$	$145 \mu\text{g}\cdot\text{l}^{-1}$	Mo	$58 \mu\text{g}\cdot\text{l}^{-1}$
	$\text{CuCl}_2\cdot 2\text{H}_2\text{O}$	$0,012 \text{ mg}\cdot\text{l}^{-1}$	$0,24 \mu\text{g}\cdot\text{l}^{-1}$	Cu	$0,080 \mu\text{g}\cdot\text{l}^{-1}$
C	$\text{NaHCO}_3$	15	300	Na; C	220; 43

(\*) Unless noted

**▼ M1**

*Footnote:* The theoretically appropriate final bicarbonate concentration (which will avoid appreciable pH adjustment) is 15 mg/l, not 300 mg/l. However, the historical use of 20X-AAP medium, including the ring test for this method, is based upon 300 mg/l. (I. Sims, P. Whitehouse and R. Lacey. (1999) The OECD *Lemna* Growth Inhibition Test. Development and Ring-testing of draft OECD Test Guideline. R & D Technical Report EMA 003. WRc plc — Environment Agency.

**STEINBERG medium (After ISO 20079)***Concentrations and stock solutions*

— The modified Steinberg medium is used in ISO 20079 for *Lemna minor* alone (as only *Lemna minor* is allowed there) but tests showed good results could be reached with *Lemna gibba* too.

— When preparing the medium, reagent- or analytical grade chemicals and deionised water should be used.

— Prepare the nutrient medium from stock solutions or the 10-fold concentrated medium which allows maximum concentration of the medium without precipitation.

Table 1

**pH-stabilised STEINBERG medium (modified acc. to Altenburger)**

Substance		Nutrient medium	
Macroelements	mol weight	mg/l	mmol/l
KNO <sub>3</sub>	101,12	350,00	3,46
Ca(NO <sub>3</sub> ) <sub>2</sub> · 4H <sub>2</sub> O	236,15	295,00	1,25
KH <sub>2</sub> PO <sub>4</sub>	136,09	90,00	0,66
K <sub>2</sub> HPO <sub>4</sub>	174,18	12,60	0,072
MgSO <sub>4</sub> · 7H <sub>2</sub> O	246,37	100,00	0,41
Microelements	mol weight	µg/l	µmol/l
H <sub>3</sub> BO <sub>3</sub>	61,83	120,00	1,94
ZnSO <sub>4</sub> · 7H <sub>2</sub> O	287,43	180,00	0,63
Na <sub>2</sub> MoO <sub>4</sub> · 2H <sub>2</sub> O	241,92	44,00	0,18
MnCl <sub>2</sub> · 4H <sub>2</sub> O	197,84	180,00	0,91
FeCl <sub>3</sub> · 6H <sub>2</sub> O	270,21	760,00	2,81
EDTA Disodium-dihydrate	372,24	1 500,00	4,03

▼ **M1**

Table 2

**Stock solutions (Macroelements)**

1. Macroelements (50-fold concentrated)	g/l
<i>Stock solution 1:</i>	
KNO <sub>3</sub>	17,50
KH <sub>2</sub> PO <sub>4</sub>	4,5
K <sub>2</sub> HPO <sub>4</sub>	0,63
<i>Stock solution 2:</i>	
MgSO <sub>4</sub> · 7H <sub>2</sub> O	5,00
<i>Stock solution 3:</i>	
Ca(NO <sub>3</sub> ) <sub>2</sub> · 4H <sub>2</sub> O	14,75

Table 3

**Stock solutions (Microelements)**

2. Microelements (1 000-fold concentrated)	mg/l
<i>Stock solution 4:</i>	
H <sub>3</sub> BO <sub>3</sub>	120,0
<i>Stock solution 5:</i>	
ZnSO <sub>4</sub> · 7H <sub>2</sub> O	180,0
<i>Stock solution 6:</i>	
Na <sub>2</sub> MoO <sub>4</sub> · 2H <sub>2</sub> O	44,0
<i>Stock solution 7:</i>	
MnCl <sub>2</sub> · 4H <sub>2</sub> O	180,0
<i>Stock solution 8:</i>	
FeCl <sub>3</sub> · 6H <sub>2</sub> O	760,00
EDTA Disodium-dihydrate	1 500,00

— Stock solutions 2 and 3 and separately 4 to 7 may be pooled (taking into account the required concentrations).

— For longer shelf life, treat stock solutions in an autoclave at 121 °C for 20 min. or alternatively carry out a sterile filtration (0,2 µm). For stock solution 8, sterile filtration (0,2 µm) is strongly recommended.

**Preparation of the final concentration of STEINBERG medium (modified)**

— Add 20 ml of stock solutions 1, 2 and 3 (see table 2) to about 900 ml deionised water to avoid precipitation.

— Add 1,0 ml of stock solutions 4, 5, 6, 7 and 8 (see table 3).

— The pH should be 5,5 ± 0,2 (adjust by addition of a minimised volume of NaOH solution or HCl).

**▼ M1**

- Adjust with water to 1 000 ml.
- If stock solutions are sterilised and appropriate water is used no further sterilisation is necessary. If sterilisation is done with the final medium, stock solution 8 should be added after autoclaving (at 121 °C for 20 min).

**Preparation of 10-fold-concentrated STEINBERG medium (modified) for intermediate storage**

- Add to 20 ml of stock solutions 1, 2 and 3 (see table 2) to about 30 ml water to avoid precipitation.
- Add 1,0 ml of stock solutions 4, 5, 6, 7 and 8 (see table 3). Adjust with water to 100 ml.
- If stock solutions are sterilised and appropriate water is used no further sterilisation is necessary. If sterilisation is done with the final medium, stock solution 8 should be added after autoclaving (at 121 °C for 20 min).
- The pH of the medium (final concentration) should be  $5,5 \pm 0,2$ .

**▼M4****C.27 SEDIMENT-WATER CHIRONOMID TOXICITY TEST USING SPIKED SEDIMENT**

## INTRODUCTION

1. This Test Method is equivalent to OECD Test Guideline (TG) 218 (2004). This Test Method is designed to assess the effects of prolonged exposure of chemicals to the sediment-dwelling larvae of the freshwater dipteran *Chironomus* sp. It is based on existing toxicity test protocols for *Chironomus riparius* and *Chironomus tentans* which have been developed in Europe (1)(2)(3) and North America (4)(5)(6)(7)(8) and ring-tested (1)(6)(9). Other well documented chironomid species may also be used, e.g. *Chironomus yoshimatsui* (10)(11).
2. The exposure scenario used in this Test Method is spiking of sediment with the test substance. The selection of the appropriate exposure scenario depends on the intended application of the test. The scenario of spiking sediment is intended to simulate accumulated levels of chemicals persisting in the sediment. This exposure system involves spiking sediment of a sediment-water test system.
3. Substances that need to be tested towards sediment-dwelling organisms usually persist in this compartment over long time periods. The sediment-dwelling organisms may be exposed via a number of routes. The relative importance of each exposure route, and the time taken for each to contribute to the overall toxic effects, is dependent on the physical-chemical properties of the chemical concerned. For strongly adsorbing substances (e.g. with  $\log K_{ow} > 5$ ) or for substances covalently binding to sediment, ingestion of contaminated food may be a significant exposure route. In order not to underestimate the toxicity of highly lipophilic substances, the use of food added to the sediment before application of the test substance may be considered. In order to take all potential routes of exposure into account the focus of this Test Method is on long-term exposure. The test duration is in the range of 20-28 days for *C. riparius* and *C. yoshimatsui*, and 28-65 days for *C. tentans*. If short-term data are required for a specific purpose, for example to investigate the effects of an unstable chemical, additional replicates may be removed after a 10-day period.
4. The measured endpoints are the total number of adults emerged and the time to emergence. It is recommended that measurements of larval survival and growth should only be made after a 10-day period if additional short-term data are required, using additional replicates as appropriate.
5. The use of formulated sediment is recommended. Formulated sediment has several advantages over natural sediments:
  - the experimental variability is reduced because it forms a reproducible 'standardised matrix' and the need to find uncontaminated and clean sediment sources is eliminated;
  - the tests can be initiated at any time without encountering seasonal variability in the test sediment and there is no need to pre-treat the sediment to remove indigenous fauna; the use of formulated sediment also reduces the cost associated with the field collection of sufficient amounts of sediment for routine testing;
  - the use of formulated sediment allows for comparisons of toxicity and ranking substances accordingly.
6. Definitions used are given in Appendix 1.

**▼M4**

## PRINCIPLE OF THE TEST

7. First instar chironomid larvae are exposed to a concentration range of the test chemical in sediment — water systems. The test substance is spiked into the sediment and first instar larvae are subsequently introduced into test beakers in which the sediment and water concentrations have been stabilised. Chironomid emergence and development rate is measured at the end of the test. Larval survival and weight may also be measured after 10 days if required (using additional replicates as appropriate). These data are analysed either by using a regression model in order to estimate the concentration that would cause  $\times$  % reduction in emergence or larval survival or growth (e.g. EC<sub>15</sub>, EC<sub>50</sub> etc.), or by using statistical hypothesis testing to determine a NOEC/LOEC. The latter requires comparison of effect values with control values using statistical tests.

## INFORMATION ON THE TEST SUBSTANCE

8. The water solubility of the test substance, its vapour pressure, measured or calculated partitioning into sediment and stability in water and sediment should be known. A reliable analytical method for the quantification of the test substance in overlying water, pore water and sediment with known and reported accuracy and limit of detection should be available. Useful information includes the structural formula and purity of the test substance. Chemical fate of the test substance (e.g. dissipation, abiotic and biotic degradation, etc.) also is useful information. Further guidance for testing substances with physical-chemical properties that make them difficult to perform the test is provided in (12)

## REFERENCE CHEMICALS

9. Reference chemicals may be tested periodically as a means of assuring that the test protocol and test conditions are reliable. Examples of reference toxicants used successfully in ring-tests and validation studies are: lindane, trifluralin, pentachlorophenol, cadmium chloride and potassium chloride (1)(2)(5)(6)(13).

## VALIDITY OF THE TEST

10. For the test to be valid the following conditions apply:

- the emergence in the controls must be at least 70 % at the end of the test. (1)(6);
- *C. riparius* and *C. yoshimatsui* emergence to adults from control vessels should occur between 12 and 23 days after their insertion into the vessels; for *C. tentans*, a period of 20 to 65 days is necessary.
- at the end of the test, pH and the dissolved oxygen concentration should be measured in each vessel. The oxygen concentration should be at least 60 per cent of the air saturation value (ASV) at the temperature used, and the pH of overlying water should be in the 6-9 range in all test vessels;
- the water temperature should not differ by more than  $\pm 1,0$  °C. The water temperature could be controlled by isothermal room and in that case the room temperature should be confirmed in an appropriate time interval.

▼ **M4****DESCRIPTION OF THE METHOD****Test vessels**

11. The study is conducted in glass 600 ml beakers measuring 8 cm in diameter. Other vessels are suitable, but they should guarantee a suitable depth of overlying water and sediment. The sediment surface should be sufficient to provide 2 to 3 cm<sup>2</sup> per larvae. The ratio of the depth of the sediment layer to the depth of the overlying water should be 1:4. Test vessels and other apparatus that will come into contact with the test system should be made entirely of glass or other chemically inert material (e.g. Teflon).

**Selection of species**

12. The species to be used in the test is preferably *Chironomus riparius*. *Chironomus tentans* is also suitable but more difficult to handle and requires a longer test period. *Chironomus yohimatsui* may also be used. Details of culture methods are given in Appendix 2 for *Chironomus riparius*. Information on culture conditions is also available for other species, i.e. *Chironomus tentans* (4) and *Chironomus yohimatsui* (11). Identification of species must be confirmed before testing but is not required prior to every test if organisms come from an in-house culture.

**Sediment**

13. Formulated sediment (also called reconstituted, artificial or synthetic sediment) should preferably be used. However, if natural sediment is used, it should be characterised (at least pH, organic carbon content, determination of other parameters such as C/N ratio and granulometry are also recommended), and it should be free from any contamination and other organisms that might compete with, or consume the chironomids. It is also recommended that, before it is used in a chironomid toxicity test, the natural sediment be conditioned for seven days under the same conditions which prevail in the subsequent test. The following formulated sediment, based on the artificial soil used in Test Method C.8 (14), is recommended for use in this test (1)(15)(16):

- (a) 4-5 % (dry weight) peat: as close to pH 5,5 to 6,0 as possible; it is important to use peat in powder form, finely ground (particle size ≤ 1 mm) and only air dried.
- (b) 20 % (dry weight) kaolin clay (kaolinite content preferably above 30 %).
- (c) 75-76 % (dry weight) quartz sand (fine sand should predominate with more than 50 per cent of the particles between 50 and 200 µm).
- (d) Deionised water is added to obtain moisture content of the final mixture in a range of 30-50 %.
- (e) Calcium carbonate of chemically pure quality (CaCO<sub>3</sub>) is added to adjust the pH of the final mixture of the sediment to 7,0 ± 0,5. Organic carbon content of the final mixture should be 2 % (± 0,5 %) and is to be adjusted by the use of appropriate amounts of peat and sand, according to (a) and (c).

14. The source of peat, kaolin clay and sand should be known. The sediment components should be checked for the absence of chemical contamination (e.g. heavy metals, organochlorine compounds, organophosphorous compounds, etc.). An example for the preparation of the formulated sediment is described in Appendix 3. Mixing of dry constituents is also acceptable if it is demonstrated that after addition of overlying water a separation of sediment constituents (e.g. floating of peat particles) does not occur, and that the peat or the sediment is sufficiently conditioned.

**▼ M4****Water**

15. Any water which conforms to the chemical characteristics of acceptable dilution water as listed in Appendices 2 and 4 is suitable as test water. Any suitable water, natural water (surface or ground water), reconstituted water (see Appendix 2) or dechlorinated tap water are acceptable as culturing water and test water if chironomids will survive in it for the duration of the culturing and testing without showing signs of stress. At the start of the test, the pH of the test water should be between 6 and 9 and the total hardness not higher than 400 mg/l as CaCO<sub>3</sub>. However, if there is an interaction suspected between hardness ions and the test substance, lower hardness water should be used (and thus, Elendt Medium M4 must not be used in this situation). The same type of water should be used throughout the whole study. The water quality characteristics listed in Appendix 4 should be measured at least twice a year or when it is suspected that these characteristics may have changed significantly.

**Stock solutions — Spiked sediments**

16. Spiked sediments of the chosen concentration are usually prepared by addition of a solution of the test substance directly to the sediment. A stock solution of the test substance dissolved in deionised water is mixed with the formulated sediment by rolling mill, feed mixer or hand mixing. If poorly soluble in water, the test substance can be dissolved in as small a volume as possible of a suitable organic solvent (e.g. hexane, acetone or chloroform). This solution is then mixed with 10 g of fine quartz sand for one test vessel. The solvent is allowed to evaporate and it has to be totally removed from sand; the sand is then mixed with the suitable amount of sediment per test beaker. Only agents which volatilise readily can be used to solubilise, disperse or emulsify the test substance. It should be born in mind that the sand provided by the test substance and sand mixture, has to be taken into account when preparing the sediment (i.e. the sediment should thus be prepared with less sand). Care should be taken to ensure that the test substance added to sediment is thoroughly and evenly distributed within the sediment. If necessary, subsamples can be analysed to determine degree of homogeneity.

**TEST DESIGN**

17. The test design relates to the selection of the number and spacing of the test concentrations, the number of vessels at each concentration and the number of larvae per vessel. Designs for EC point estimation, for estimation of NOEC, and for conducting a limit test are described.

**Design for analysis by regression**

18. The effect concentration (e.g. EC<sub>15</sub>, EC<sub>50</sub>) and the concentration range, over which the effect of the test substance is of interest, should be spanned by the concentrations included in the test. Generally, the accuracy and especially validity, with which estimates of effect concentrations (EC<sub>x</sub>) can be made, is improved when the effect concentration is within the range of concentrations tested. Extrapolating much below the lowest positive concentration or above the highest concentration should be avoided. A preliminary range-finding test is helpful for selecting the range of concentrations to be used (see paragraph 27).



**▼M4**

19. If the  $EC_x$  is to be estimated, at least five concentrations and three replicates for each concentration should be tested. In any case, it is advisable that sufficient test concentrations are used to allow good model estimation. The factor between concentrations should not be greater than two (an exception could be made in cases when the dose response curve has a shallow slope). The number of replicates at each treatment can be reduced if the number of test concentrations with different responses is increased. Increasing the number of replicates or reducing the size of the test concentration intervals tends to lead to narrower confidence intervals for the test. Additional replicates are required if 10-day larval survival and growth are to be estimated.

**Design for estimation of a NOEC/LOEC**

20. If the LOEC or NOEC are to be estimated, five test concentrations with at least four replicates should be used and the factor between concentrations should not be greater than two. The number of replicates should be sufficient to ensure adequate statistical power to detect a 20 % difference from the control at the 5 % level of significance ( $p = 0,05$ ). With the development rate, an Analysis of Variance (ANOVA) is usually appropriate, such as Dunnett-test and Williams-test (17)(18)(19)(20). In the emergence ratio the Cochran-Armitage, Fisher's exact (with Bonferroni correction), or Mantel-Haenszel tests may be used.

**Limit test**

21. A limit test may be performed (one test concentration and control) if no effects were seen in the preliminary range-finding test. The purpose of the limit test is to perform a test at a concentration sufficiently high to enable decision makers to exclude possible toxic effects of the test substance, and the limit is set at a concentration which is not expected to appear in any situation. 1 000 mg/kg (dry weight) is recommended. Usually, at least six replicates for both the treatment and control are necessary. Adequate statistical power to detect a 20 % difference from the control at the 5 % level of significance ( $p = 0,05$ ) should be demonstrated. With metric response (development rate and weight), the t-test is a suitable statistical method if data meet the requirements of this test (normality, homogeneous variances). The unequal-variance t-test or a non parametric test, such as the Wilcoxon-Mann-Whitney test may be used, if these requirements are not fulfilled. With the emergence ratio, the Fisher exact test is appropriate.

**PROCEDURE****Conditions of exposure***Preparation of spiked sediment — water system*

22. The spiking procedure described in Test Method C.8: Toxicity for Earthworms is recommended for application of the test substance (14). The spiked sediments are placed in the vessels and overlying water is added to produce a sediment-water volume ratio of 1:4 (see paragraphs 11 and 15). The depth of the sediment layer should be in the range of 1,5-3 cm. To avoid separation of sediment ingredients and re-suspension of fine material during addition of test water in the water column, the sediment can be covered with a plastic disc while water is poured onto it, and the disc removed immediately afterwards. Other devices may also be appropriate.
23. The test vessels should be covered (e.g. by glass plates). If necessary, during the study the water levels will be topped to the original volume in order to compensate for water evaporation. This should be performed using distilled or deionised water to prevent build-up of salts.

**▼ M4***Stabilisation*

24. Once the spiked sediment with overlying water has been prepared, it is desirable to allow partitioning of the test substance from the aqueous phase to the sediment (3)(4)(6)(13). This should preferably be done under the conditions of temperature and aeration used in the test. Appropriate equilibration time is sediment and chemical specific, and can be in the order of hours to days and in rare cases up to several weeks (4-5 weeks). As this would leave time for degradation of many chemicals, equilibrium is not awaited but an equilibration period of 48 hours is recommended. At the end of this further equilibration period, the concentration of the test substance should be measured in the overlying water, the pore water and the sediment, at least at the highest concentration and a lower one (see paragraph 38). These analytical determinations of the test substance allow for calculation of mass balance and expression of results based on measured concentrations.

*Addition of test organisms*

25. Four to five days before adding the test organisms to the test vessels, egg masses should be taken from the cultures and placed in small vessels in culture medium. Aged medium from the stock culture or freshly prepared medium may be used. If the latter is used, a small amount of food e.g. green algae and/or a few droplets of filtrate from a finely ground suspension of flaked fish food should be added to the culture medium (see Appendix 2). Only freshly laid egg masses should be used. Normally, the larvae begin to hatch a couple of days after the eggs are laid (2 to 3 days for *Chironomus riparius* at 20 °C and 1 to 4 days for *Chironomus tentans* at 23 °C and *Chironomus yoshimatoi* at 25 °C) and larval growth occurs in four instars, each of 4-8 days duration. First instar larvae (2-3 or 1-4 days post hatching) should be used in the test. The instar of midges can possibly be checked using head capsule width (6).
26. Twenty first instar larvae are allocated randomly to each test vessel containing the spiked sediment and water, using a blunt pipette. Aeration of the water has to be stopped while adding the larvae to test vessels and remain so for another 24 hours after addition of larvae (see paragraphs 25 and 32). According to the test design used (see paragraphs 19 and 20), the number of larvae used per concentration is at least 60 for the EC point estimation and 80 for determination of NOEC.

*Test concentrations*

27. A range-finding test may be helpful to determine the range of concentrations for the definitive test. For this purpose a series of widely spaced concentrations of the test substance are used. In order to provide the same density of surface per chironomids, which is to be used for the definitive test, chironomids are exposed to each concentration of the test substance for a period which allows estimation of appropriate test concentrations, and no replicates are required.
28. The test concentrations for the definitive test are decided based on the result of the range-finding test. At least five concentrations should be used and selected as described in paragraphs 18 to 20.

**▼M4***Controls*

29. Control vessels without any test substance but including sediment should be included in the test with the appropriate number of replicates (see paragraphs 19-20). If a solvent has been used for application of test substance (see paragraph 16), a sediment solvent control should be added.

*Test system*

30. Static systems are used. Semi-static or flow-through systems with intermittent or continuous renewal of overlying water might be used in exceptional cases as for instance if water quality specifications become inappropriate for the test organism or affect chemical equilibrium (e.g. dissolved oxygen levels fall too low, the concentration of excretory products rises too high or minerals leach from sediment and affect pH and/or water hardness). However, other methods for ameliorating the quality of overlying water, such as aeration, will normally suffice and be preferable.

*Food*

31. It is necessary to feed the larvae, preferably daily or at least three times per week. Fish-food (a suspension in water or finely ground food, e.g. TetraMin or TetraPhyll; see details in Appendix 2) in the amount of 0,25-0,5 mg (0,35-0,5 mg for *C. yoshimatui*) per larvae per day seems adequate for young larvae for the first 10 days. Slightly more food may be necessary for older larvae: 0,5-1 mg per larvae per day should be sufficient for the rest of the test. The food ration should be reduced in all treatments and control if fungal growth is seen or if mortality is observed in controls. If fungal development cannot be stopped the test is to be repeated. When testing strongly adsorbing substances (e.g. with  $\log K_{ow} > 5$ ), or substances covalently binding to sediment, the amount of food necessary to ensure survival and natural growth of the organisms may be added to the formulated sediment before the stabilisation period. For this, plant material must be used instead of fish food, e.g. addition of 0,5 % (dry weight) finely ground leaves of e.g. stinging nettle (*Urtica dioica*), mulberry (*Morus alba*), white clover (*Trifolium repens*), spinach (*Spinacia oleracea*) or of other plant material (*Cerophyl* or alpha-cellulose) may be used.

*Incubation conditions*

32. Gentle aeration of the overlying water in test vessels is supplied preferably 24 hours after addition of the larvae and is pursued throughout the test (care should be taken that dissolved oxygen concentration does not fall below 60 per cent of ASV). Aeration is provided through a glass Pasteur pipette fixed 2-3 cm above the sediment layer (i.e. one or few bubbles/sec). When testing volatile chemicals, consideration may be given not to aerate the sediment-water system.
33. The test is conducted at a constant temperature of 20 °C ( $\pm 2$  °C). For *C. tentans* and *C. yoshimatui* recommended temperatures are 23 °C and 25 °C ( $\pm 2$  °C), respectively. A 16 hours photoperiod is used and the light intensity should be 500 to 1 000 lux.

**▼ M4***Exposure duration*

34. The exposure commences with the addition of larvae to the spiked and control vessels. The maximum exposure duration is 28 days for *C. riparius* and *C. yoshimatsui*, and 65 days for *C. tentans*. If midges emerge earlier, the test can be terminated after a minimum of five days after emergence of the last adult in the control.

**Observations***Emergence*

35. The development time and the total number of fully emerged male and female midges are determined. Males are easily identified by their plumose antennae.
  
36. The test vessels should be observed at least three times per week to make visual assessment of any abnormal behaviour (e.g. leaving sediment, unusual swimming), compared with the control. During the period of expected emergence a daily count of emerged midges is necessary. The sex and number of fully emerged midges are recorded daily. After identification the midges are removed from the vessels. Any egg masses deposited prior to the termination of the test should be recorded and then removed to prevent re-introduction of larvae into the sediment. The number of visible pupae that have failed to emerge is also recorded. Guidance on measurement of emergence is provided in Appendix 5.

*Growth and survival*

37. If data on 10-day larval survival and growth are to be provided, additional test vessels should be included at the start, so that they may be used subsequently. The sediment from these additional vessels is sieved using a 250 µm sieve to retain the larvae. Criteria for death are immobility or lack of reaction to a mechanical stimulus. Larvae not recovered should also be counted as dead (larvae which have died at beginning of the test may have been degraded by microbes). The (ash free) dry weight of the surviving larvae per test vessel is determined and the mean individual dry weight per vessel calculated. It is useful to determine which instar the surviving larvae belong to; for that measurement of the width of the head capsule of each individual can be used.

**Analytical measurements***Concentration of the test substance*

38. Prior to test commencement (i.e. addition of larvae), samples of bulk sediment are removed from at least one vessel per treatment for the analytical determination of the test substance concentration in the sediment. It is recommended that, as a minimum, samples of the overlying water, the pore water and the sediment be analysed at the start (see paragraph 24) and at the end of the test, at the highest concentration and a lower one. These determinations of test substance concentration inform about the behaviour/partitioning of the test substance in the water-sediment system.
  
39. When intermediate measurements are made (e.g. at day 7) and if the analysis needs large samples which cannot be taken from test vessels without influencing the test system, analytical determinations should be performed on samples from additional test vessels treated in the same way (including the presence of test organisms) but not used for biological observations.

**▼ M4**

40. Centrifugation at e.g. 10 000 g and 4 °C for 30 min. is the recommended procedure to isolate interstitial water. However, if the test substance is demonstrated not to adsorb to filters, filtration may also be acceptable. In some cases it might not be possible to analyse concentrations in the pore water as the sample size is too small.

*Physical-chemical parameters*

41. pH and temperature of the test vessels should be measured in an appropriate manner (see paragraph 10). Hardness and ammonia should be measured in the controls and one test vessel at the highest concentration at the start and the end of the test.

## DATA AND REPORTING

*Treatment of results*

42. The purpose of this test is to determine the effect of the test substance on the development rate and the total number of fully emerged male and female midges, or in the case of the 10-day test effects on survival and weight of the larvae. If there are no indications of statistically different sensitivities of sexes, male and female results may be pooled for statistical analyses. The sensitivity differences between sexes can be statistically judged by e.g. a  $\chi^2$ -r  $\times$  2 table test. Larval survival and mean individual dry weight per vessel must be determined after 10 days where required.
43. Effect concentrations expressed and based on dry weight, are calculated preferably based on measured sediment concentrations at the beginning of the test (see paragraph 38).
44. To compute a point estimate for the EC<sub>50</sub> or any other EC<sub>x</sub>, the per-vessel statistics may be used as true replicates. In calculating a confidence interval for any EC<sub>x</sub> the variability among vessels should be taken into account, or it should be shown that this variability is so small that it can be ignored. When the model is fitted by Least Squares, a transformation should be applied to the per-vessel statistics in order to improve the homogeneity of variance. However, EC<sub>x</sub> values should be calculated after the response is transformed back to the original value.
45. When the statistical analysis aims at determining the NOEC/LOEC by hypothesis testing, the variability among vessels needs to be taken into account, e.g. by a nested ANOVA. Alternatively, more robust tests (21) can be appropriate in situations where there are violations of the usual ANOVA assumptions.

*Emergence ratio*

46. Emergence ratios are quantal data, and can be analyzed by the Cochran-Armitage test applied in step-down manner where a monotonic dose-response is expected and these data are consistent with this expectation. If not, a Fisher's exact or Mantel-Haenszel test with Bonferroni-Holm adjusted p-values can be used. If there is evidence of greater variability between replicates within the same concentration than a binomial distribution would indicate (often referenced as 'extra-binomial' variation), then a robust Cochran-Armitage or Fisher exact test such as proposed in (21), should be used.

**▼M4**

The sum of midges emerged per vessel,  $n_e$ , is determined and divided by the number of larvae introduced,  $n_a$ :

$$ER = \frac{n_e}{n_a}$$

where:

$ER$  = emergence ratio

$n_e$  = number of midges emerged per vessel

$n_a$  = number of larvae introduced per vessel

47. An alternative that is most appropriate for large sample sizes, when there is extra binomial variance, is to treat the emergence ratio as a continuous response and use procedures such as William's test when a monotonic dose-response is expected and is consistent with these ER data. Dunnett's test would be appropriate where monotonicity does not hold. A large sample size is defined here as the number emerged and the number not emerging both exceeding five, on a per replicate (vessel) basis.
48. To apply ANOVA methods values of ER should first be transformed by the arcsin-sqrt-transformation or Freeman-Tukey transformation to obtain an approximate normal distribution and to equalise variances. The Cochran-Armitage, Fisher's exact (Bonferroni), or Mantel-Haenszel tests can be applied when using the absolute frequencies. The arcsin-sqrt transformation is applied by taking the inverse sine ( $\sin^{-1}$ ) of the square root of ER.
49. For emergence ratios,  $EC_x$ -values are calculated using regression analysis (or e.g. probit (22), logit, Weibull, appropriate commercial software etc.). If regression analysis fails (e.g. when there are less than two partial responses), other non-parametric methods such as moving average or simple interpolation are used.

*Development rate*

50. The mean development time represents the mean time span between the introduction of larvae (day 0 of the test) and the emergence of the experimental cohort of midges. (For the calculation of the true development time, the age of larvae at the time of introduction should be considered). The development rate is the reciprocal of the development time (unit: 1/day) and represents that portion of larval development which takes place per day. The development rate is preferred for the evaluation of these sediment toxicity studies as its variance is lower, and it is more homogeneous and closer to normal distribution as compared to development time. Hence, powerful parametric test procedures may be used with development rate rather than with development time. For development rate as a continuous response,  $EC_x$ -values can be estimated by using regression analysis (e.g. (23), (24)).
51. For the following statistical tests, the number of midges observed on inspection day  $x$  are assumed to be emerged at the mean of the time interval between day  $x$  and day  $x - 1$  ( $l$  = length of the inspection interval, usually 1 day). The mean development rate per vessel ( $\bar{x}$ ) is calculated according to:

$$\bar{x} = \sum_{i=1}^m \frac{f_i x_i}{n_e}$$

**▼ M4**

where:

$\bar{x}$ : mean development rate per vessel

$i$ : index of inspection interval

$m$ : maximum number of inspection intervals

$f_i$ : number of midges emerged in the inspection interval  $i$

$n_e$ : total number of midges emerged at the end of experiment ( $= \sum f_i$ )

$x_i$ : development rate of the midges emerged in interval  $i$

$$x_i = \frac{1}{\left(\text{day}_i - \frac{l_i}{2}\right)}$$

where:

$\text{day}_i$ : inspection day (days since application)

$l_i$ : length of inspection interval  $i$  (days, usually 1 day)

**Test report**

52. The test report must at least provide the following information:

*Test substance:*

- physical nature and, where relevant, physical-chemical properties (water solubility, vapour pressure, partition coefficient in soil (or in sediment if available), stability in water, etc.);
- chemical identification data (common name, chemical name, structural formula, CAS number, etc.) including purity and analytical method for quantification of test substance.

*Test species:*

- test animals used: species, scientific name, source of organisms and breeding conditions;
- information on handling of egg masses and larvae;
- age of test animals when inserted into test vessels.

*Test conditions:*

- sediment used, i.e. natural or formulated sediment;
- for natural sediment, location and description of sediment sampling site, including, if possible, contamination history; characteristics: pH, organic carbon content, C/N ratio and granulometry (if appropriate).
- preparation of the formulated sediment: ingredients and characteristics (organic carbon content, pH, moisture, etc. at the start of the test);
- preparation of the test water (if reconstituted water is used) and characteristics (oxygen concentration, pH, conductivity, hardness, etc. at the start of the test);
- depth of sediment and overlying water;
- volume of overlying and pore water; weight of wet sediment with and without pore water;

**▼ M4**

- test vessels (material and size);
- method of spiking sediment: test concentrations used, number of replicates and use of solvent if any;
- stabilisation equilibrium phase of the spiked sediment-water system: duration and conditions;
- incubation conditions: temperature, light cycle and intensity, aeration (frequency and intensity);
- detailed information on feeding including type of food, preparation, amount and feeding regime.

*Results:*

- the nominal test concentrations, the measured test concentrations and the results of all analyses to determine the concentration of the test substance in the test vessel;
- water quality within the test vessels, i.e. pH, temperature, dissolved oxygen, hardness and ammonia;
- replacement of evaporated test water, if any;
- number of emerged male and female midges per vessel and per day;
- number of larvae which failed to emerge as midges per vessel;
- mean individual dry weight of larvae per vessel, and per instar, if appropriate;
- percent emergence per replicate and test concentration (male and female midges pooled);
- mean development rate of fully emerged midges per replicate and treatment rate (male and female midges pooled);
- estimates of toxic endpoints e.g. EC<sub>x</sub> (and associated confidence intervals), NOEC and/or LOEC, and the statistical methods used for their determination;
- discussion of the results, including any influence on the outcome of the test resulting from deviations from this Test Method.

*LITERATURE:*

- (1) BBA (1995). Long-term toxicity test with *Chironomus riparius*: Development and validation of a new test system. Edited by M. Streloke and H.Köpp. Berlin 1995.
- (2) Fleming R *et al.* (1994). Sediment Toxicity Tests for Poorly Water-Soluble Substances. Final Report to them European Commission. Report No: EC 3738. August 1994. WRc, UK.
- (3) SETAC (1993). Guidance Document on Sediment toxicity Tests and Bioassays for Freshwater and Marine Environments. From the WOSTA Workshop held in the Netherlands.
- (4) ASTM International/E1706-00 (2002). Test Method for Measuring the Toxicity of Sediment-Associated Contaminants with Freshwater Invertebrates. pp 1125-1241. In ASTM International 2002 Annual Book of Standards. Volume 11.05. Biological Effects and Environmental Fate;Biotechnology; Pesticides. ASTM. International, West Conshohocken, PA.
- (5) Environment Canada (1997). Test for Growth and Survival in Sediment using Larvae of Freshwater Midges (*Chironomus tentans* or *Chironomus riparius*). Biological Test Method. Report SPE 1/RM/32. December 1997.



▼ **M4**

- (6) US-EPA (2000). Methods for Measuring the Toxicity and Bioaccumulation of Sediment-associated Contaminants with Freshwater Invertebrates. Second edition. EPA 600/R-99/064. March 2000. Revision to the first edition dated June 1994.
- (7) US-EPA/OPPTS 850.1735. (1996): Whole Sediment Acute Toxicity Invertebrates.
- (8) US-EPA/OPPTS 850.1790. (1996): Chironomid Sediment toxicity Test.
- (9) Milani D, Day KE, McLeay DJ, and Kirby RS (1996). Recent intra- and inter-laboratory studies related to the development and standardisation of Environment Canada's biological test methods for measuring sediment toxicity using freshwater amphipods (*Hyalella azteca*) and midge larvae (*Chironomus riparius*). Technical Report. Environment Canada. National Water Research Institute. Burlington, Ontario, Canada.
- (10) Sugaya Y (1997). Intra-specific variations of the susceptibility of insecticides in *Chironomus yoshimatsui*. Jp. J. Sanit. Zool. 48 (4): 345-350.
- (11) Kawai K (1986). Fundamental studies on Chironomid allergy. I. Culture methods of some Japanese Chironomids (Chironomidae, Diptera). Jp. J. Sanit. Zool. 37(1): 47-57.
- (12) OECD (2000). Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures. OECD Environment, Health and Safety Publications, Series on Testing and Assessment No 23.
- (13) Environment Canada (1995). Guidance Document on Measurement of Toxicity Test Precision Using Control Sediments Spiked with a Reference Toxicant. Report EPS 1/RM/30. September 1995.
- (14) Test Method C.8 of this Annex, Toxicity for Earthworms.
- (15) Suedel BC and JH Rodgers (1994). Development of formulated reference sediments for freshwater and estuarine sediment testing. Environ. Toxicol. Chem. 13: 1163-1175.
- (16) Naylor C and C Rodrigues (1995). Development of a test method for *Chironomus riparius* using a formulated sediment. Chemosphere 31: 3291-3303.
- (17) Dunnett CW (1964). A multiple comparisons procedure for comparing several treatments with a control. J. Amer. Statis. Assoc., 50: 1096-1121.
- (18) Dunnett CW (1964). New tables for multiple comparisons with a control. Biometrics, 20: 482-491.
- (19) Williams DA (1971). A test for differences between treatment means when several dose levels are compared with a zero dose control. Biometrics, 27: 103-117.
- (20) Williams DA (1972). The comparison of several dose levels with a zero dose control. Biometrics, 28: 510-531.
- (21) Rao JNK and Scott AJ (1992). A simple method for the analysis of clustered binary data. Biometrics 48: 577-585.
- (22) Christensen ER (1984). Dose-response functions in aquatic toxicity testing and the Weibull model. Water Research 18: 213-221.
- (23) Bruce and Versteeg (1992). A statistical procedure for modelling continuous toxicity data. Environmental Toxicology and Chemistry 11: 1485-1494.
- (24) Slob W (2002). Dose-response modelling of continuous endpoints. Toxicol. Sci. 66: 298-312.

**▼ M4***Appendix 1*

## DEFINITIONS

For the purpose of this Test Method the following definitions are used:

**Formulated sediment** or reconstituted, artificial or synthetic sediment, is a mixture of materials used to mimic the physical components of a natural sediment.

**Overlying water** is the water placed over sediment in the test vessel.

**Interstitial water** or pore water is the water occupying space between sediment and soil particles.

**Spiked sediment** is sediment to which test substance has been added.

**Test chemical:** Any substance or mixture tested using this Test Method.

**▼ M4***Appendix 2***Recommendations for culture of *Chironomus riparius***

1. *Chironomus* larvae may be reared in crystallising dishes or larger containers. Fine quartz sand is spread in a thin layer of about 5 to 10 mm deep over the bottom of the container. Kieselguhr (e.g. Merck, Art 8117) has also been shown to be a suitable substrate (a thinner layer of up to a very few mm is sufficient). Suitable water is then added to a depth of several cm. Water levels should be topped up as necessary to replace evaporative loss, and prevent desiccation. Water can be replaced if necessary. Gentle aeration should be provided. The larval rearing vessels should be held in a suitable cage which will prevent escape of the emerging adults. The cage should be sufficiently large to allow swarming of emerged adults, otherwise copulation may not occur (minimum is ca. 30 × 30 × 30 cm).
2. Cages should be held at room temperature or in a constant environment room at 20 ± 2 °C with a photo period of 16 hour light (intensity ca. 1 000 lux), 8 hours dark. It has been reported that air humidity of less than 60 % RH can impede reproduction.

**Dilution water**

3. Any suitable natural or synthetic water may be used. Well water, dechlorinated tap water and artificial media (e.g. Elendt 'M4' or 'M7' medium, see below) are commonly used. The water has to be aerated before use. If necessary, the culture water may be renewed by pouring or siphoning the used water from culture vessels carefully without destroying the tubes of larvae.

**Feeding larvae**

4. *Chironomus* larvae should be fed with a fish flake food (TetraMin® Tetra-Phyll® or other similar brand of proprietary fish food), at approximately 250 mg per vessel per day. This can be given as a dry ground powder or as a suspension in water: 1,0 g of flake food is added to 20 ml of dilution water and blended to give a homogenous mix. This preparation may be fed at a rate of about 5 ml per vessel per day (shake before use). Older larvae may receive more.
5. Feeding is adjusted according to the water quality. If the culture medium becomes 'cloudy', the feeding should be reduced. Food additions must be carefully monitored. Too little food will cause emigration of the larvae towards the water column, and too much food will cause increased microbial activity and reduced oxygen concentrations. Both conditions can result in reduced growth rates.
6. Some green algae (e.g. *Scenedesmus subspicatus*, *Chlorella vulgaris*) cells may also be added when new culture vessels are set up.

**Feeding emerged adults**

7. Some experimenters have suggested that a cotton wool pad soaked in a saturated sucrose solution may serve as a food for emerged adults.

**▼ M4****Emergence**

8. At  $20 \pm 2$  °C adults will begin to emerge from the larval rearing vessels after approximately 13-15 days. Males are easily distinguished by having plumose antennae.

**Egg masses**

9. Once adults are present within the breeding cage, all larval rearing vessels should be checked three times weekly for deposition of the gelatinous egg masses. If present, the egg masses should be carefully removed. They should be transferred to a small dish containing a sample of the breeding water. Egg masses are used to start a new culture vessel (e.g. 2-4 egg masses/vessel) or are used for toxicity tests.
10. First instar larvae should hatch after 2-3 days.

**Set-up of new culture vessels**

11. Once cultures are established it should be possible to set up a fresh larval culture vessel weekly or less frequently depending on testing requirements, removing the older vessels after adult midges have emerged. Using this system a regular supply of adults will be produced with a minimum of management.

**Preparation of test solutions 'M4' and 'M7'**

12. Elendt (1990) has described the 'M4' medium. The 'M7' medium is prepared as the 'M4' medium except for the substances indicated in Table 1, for which concentrations are four times lower in 'M7' than in 'M4'. A publication on the 'M7' medium is in preparation (Elendt, personal communication). The test solution should not be prepared according to Elendt and Bias (1990) for the concentrations of  $\text{NaSiO}_3 \cdot 5 \text{H}_2\text{O}$ ,  $\text{NaNO}_3$ ,  $\text{KH}_2\text{PO}_4$  and  $\text{K}_2\text{HPO}_4$  given for the preparation of the stock solutions are not adequate.

**Preparation of the 'M7'-medium**

13. Each stock solution (I) is prepared individually and a combined stock solution (II) is prepared from these stock solutions (I) (see Table 1). Fifty ml from the combined stock Solution (II) and the amounts of each macro nutrient stock solution which are given in Table 2 are made up to 1 litre of deionised water to prepare the 'M7' medium. A vitamin stock solution is prepared by adding three vitamins to deionised water as indicated in Table 3, and 0,1 ml of the combined vitamin stock solution are added to the final 'M7' medium shortly before use. (The vitamin stock solution is stored frozen in small aliquots). The medium is aerated and stabilised.

**LITERATURE:**

BBA (1995). Long-term toxicity test with *Chironomus riparius*: Development and validation of a new test system. Edited by M. Streløkke and H. Köpp. Berlin 1995.

▼ **M4**

Table 1

**Stock solutions of trace elements for medium M4 and M7**

Stock solutions (I)	Amount (mg) made up to 1 litre of deionised water	To prepare the combined stock solution (II): mix the following amounts (ml) of stock solutions (I) and make up to 1 litre of deionised water		Final concentrations in test solutions (mg/l)	
		M4	M7	M4	M7
H <sub>3</sub> BO <sub>3</sub> <sup>(1)</sup>	57 190	1,0	0,25	2,86	0,715
MnCl <sub>2</sub> · 4 H <sub>2</sub> O <sup>(1)</sup>	7 210	1,0	0,25	0,361	0,090
LiCl <sup>(1)</sup>	6 120	1,0	0,25	0,306	0,077
RbCl <sup>(1)</sup>	1 420	1,0	0,25	0,071	0,018
SrCl <sub>2</sub> · 6 H <sub>2</sub> O <sup>(1)</sup>	3 040	1,0	0,25	0,152	0,038
NaBr <sup>(1)</sup>	320	1,0	0,25	0,016	0,004
Na <sub>2</sub> MoO <sub>4</sub> · 2 H <sub>2</sub> O <sup>(1)</sup>	1 260	1,0	0,25	0,063	0,016
CuCl <sub>2</sub> · 2 H <sub>2</sub> O <sup>(1)</sup>	335	1,0	0,25	0,017	0,004
ZnCl <sub>2</sub>	260	1,0	1,0	0,013	0,013
CaCl <sub>2</sub> · 6 H <sub>2</sub> O	200	1,0	1,0	0,010	0,010
KI	65	1,0	1,0	0,0033	0,0033
Na <sub>2</sub> SeO <sub>3</sub>	43,8	1,0	1,0	0,0022	0,0022
NH <sub>4</sub> VO <sub>3</sub>	11,5	1,0	1,0	0,00058	0,00058
Na <sub>2</sub> EDTA · 2 H <sub>2</sub> O <sup>(1)</sup> <sup>(2)</sup>	5 000	20,0	5,0	2,5	0,625
FeSO <sub>4</sub> · 7 H <sub>2</sub> O <sup>(1)</sup> <sup>(2)</sup>	1 991	20,0	5,0	1,0	0,249

<sup>(1)</sup> These substances differ in M4 and M7, as indicated above.

<sup>(2)</sup> These solutions are prepared individually, then poured together and autoclaved immediately.

Table 2

**Macro nutrient stock solutions for medium M4 and M7**

	Amount made up to 1 litre of deionised water (mg)	Amount of macro nutrient stock solutions added to prepare medium M4 and M7 (ml/l)	Final concentrations in test solutions M4 and M7 (mg/l)
CaCl <sub>2</sub> · 2 H <sub>2</sub> O	293 800	1,0	293,8
MgSO <sub>4</sub> · 7 H <sub>2</sub> O	246 600	0,5	123,3
KCl	58 000	0,1	5,8
NaHCO <sub>3</sub>	64 800	1,0	64,8
NaSiO <sub>3</sub> · 9 H <sub>2</sub> O	50 000	0,2	10,0
NaNO <sub>3</sub>	2 740	0,1	0,274
KH <sub>2</sub> PO <sub>4</sub>	1 430	0,1	0,143
K <sub>2</sub> HPO <sub>4</sub>	1 840	0,1	0,184

▼ **M4**

Table 3

**Vitamin stock solution for medium M4 and M7. All three vitamin solutions are combined to make a single vitamin stock solution**

	Amount made up to 1 litre of deionised water (mg)	Amount of vitamin stock solution added to prepare medium M4 and M7 (ml/l)	Final concentrations in test solutions M4 and M7 (mg/l)
Thiamine hydrochloride	750	0,1	0,075
Cyanocobalamin (B12)	10	0,1	0,0010
Biotine	7,5	0,1	0,00075

*LITERATURE:*

ElenDt, B.P. (1990). Selenium Deficiency in Crustacean. *Protoplasma* 154: 25-33.

ElenDt, B.P. & W.-R. Bias (1990). Trace Nutrient Deficiency in *Daphnia magna* Cultured in Standard Medium for Toxicity Testing. Effects on the Optimization of Culture Conditions on Life History Parameters of *D. magna*. *Water Research* 24 (9): 1157-1167.

▼ **M4***Appendix 3***PREPARATION OF FORMULATED SEDIMENT****Sediment composition**

The composition of the formulated sediment should be as follows:

Constituent	Characteristics	% of sediment dry weight
Peat	Sphagnum moss peat, as close to pH 5,5-6,0 as possible, no visible plant remains, finely ground (particle size $\leq$ 1 mm) and air dried	4-5
Quartz sand	Grain size: > 50 % of the particles should be in the range of 50-200 $\mu$ m	75-76
Kaolinite clay	Kaolinite content $\geq$ 30 %	20
Organic carbon	Adjusted by addition of peat and sand	2 ( $\pm$ 0,5)
Calcium carbonate	CaCO <sub>3</sub> , pulverised, chemically pure	0,05-0,1
Water	Conductivity $\leq$ 10 $\mu$ S/cm	30-50

**Preparation**

The peat is air dried and ground to a fine powder. A suspension of the required amount of peat powder in deionised water is prepared using a high-performance homogenising device. The pH of this suspension is adjusted to  $5,5 \pm 0,5$  with CaCO<sub>3</sub>. The suspension is conditioned for at least two days with gentle stirring at  $20 \pm 2$  °C, to stabilise pH and establish a stable microbial component. pH is measured again and should be  $6,0 \pm 0,5$ . Then the peat suspension is mixed with the other constituents (sand and kaolin clay) and deionised water to obtain a homogeneous sediment with a water content in a range of 30-50 per cent of dry weight of the sediment. The pH of the final mixture is measured once again and is adjusted to 6,5 to 7,5 with CaCO<sub>3</sub> if necessary. Samples of the sediment are taken to determine the dry weight and the organic carbon content. Then, before it is used in the chironomid toxicity test, it is recommended that the formulated sediment be conditioned for seven days under the same conditions which prevail in the subsequent test.

**Storage**

The dry constituents for preparation of the artificial sediment may be stored in a dry and cool place at room temperature. The formulated (wet) sediment should not be stored prior to its use in the test. It should be used immediately after the 7 days conditioning period that ends its preparation.

*LITERATURE:*

Chapter C.8 of this Annex. Toxicity for Earthworms.

Meller M, Egeler P, Rombke J, Schallnass H, Nagel R, Streit B (1998). Short-term Toxicity of Lindane, Hexachlorobenzene and Copper Sulfate on Tubificid Sludgeworms (Oligochaeta) in Artificial Media. *Ecotox. and Environ. Safety* 39: 10-20.

**▼ M4***Appendix 4***Chemical Characteristics of an Acceptable Dilution Water**

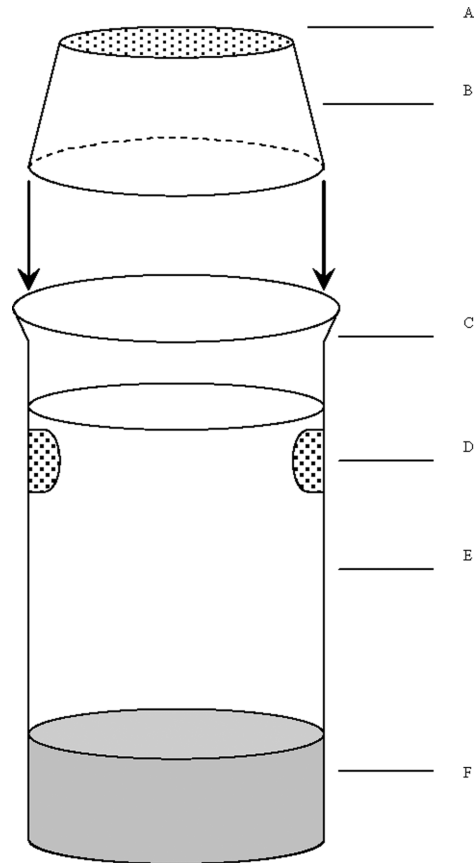
Substance	Concentrations
Particulate matter	< 20 mg/l
Total organic carbon	< 2 mg/l
Unionised ammonia	< 1 µg/l
Hardness as CaCO <sub>3</sub>	< 400 mg/l (*)
Residual chlorine	< 10 µg/l
Total organophosphorus pesticides	< 50 ng/l
Total organochlorine pesticides plus polychlorinated biphenyls	< 50 ng/l
Total organic chlorine	< 25 ng/l

(\*) However, it should be noted that if there is an interaction suspected between hardness ions and the test substance, lower hardness water should be used (and thus, Elendt Medium M4 must not be used in this situation).



▼ **M4***Appendix 5***Guidance for monitoring emergence of chironomid larvae**

Emergence traps are placed on the test beakers. These traps are needed from day 20 to the end of the test. Example of trap used is drawn below:



A: the nylon screen

B: the inverted plastic cups

C: the lipless exposure beaker

D: the water exchange screen ports

E: water

F: sediment

**▼M4****C. 28. SEDIMENT-WATER CHIRONOMID TOXICITY TEST USING SPIKED WATER**

## INTRODUCTION

1. This Test Method is equivalent to OECD TG 219 (2004). This Test Method is designed to assess the effects of prolonged exposure of chemicals to the sediment-dwelling larvae of the freshwater dipteran *Chironomus* sp. It is mainly based on the BBA guideline using a sediment-water test system with artificial soil, and water column exposure scenario (1). It also takes into account existing toxicity test protocols for *Chironomus riparius* and *Chironomus tentans* which have been developed in Europe and North America (2)(3)(4)(5)(6)(7)(8) and ring-tested (1)(6)(9). Other well documented chironomid species may also be used, e.g. *Chironomus yoshimatsui* (10)(11).
2. The exposure scenario used in this Test Method is water spiking. The selection of the appropriate exposure scenario depends on the intended application of the test. The water exposure scenario, involving spiking of the water column, is intended to simulate a pesticide spray drift event and covers the initial peak of concentrations in pore water. It is also useful for other types of exposure (including chemical spills) except accumulation processes lasting longer than the test period.
3. Substances that need to be tested towards sediment-dwelling organisms usually persist in this compartment over long time periods. The sediment-dwelling organisms may be exposed via a number of routes. The relative importance of each exposure route, and the time taken for each to contribute to the overall toxic effects, is dependent on the physical-chemical properties of the chemical concerned. For strongly adsorbing substances (e.g. with  $\log K_{ow} > 5$ ) or for substances covalently binding to sediment, ingestion of contaminated food may be a significant exposure route. In order not to underestimate the toxicity of highly lipophilic substances, the use of food added to the sediment before application of the test substance may be considered. In order to take all potential routes of exposure into account the focus of this Test Method is on long-term exposure. The test duration is in the range of 20-28 days for *C. riparius* and *C. yoshimatsui*, and 28-65 days for *C. tentans*. If short-term data are required for a specific purpose, for example to investigate the effects of unstable chemicals, additional replicates may be removed after a 10-day period.
4. The measured endpoints are the total number of adults emerged and the time to emergence. It is recommended that measurements of larval survival and growth should only be made after a 10-day period if additional short-term data are required, using additional replicates as appropriate.
5. The use of formulated sediment is recommended. Formulated sediment has several advantages over natural sediments:
  - the experimental variability is reduced because it forms a reproducible ‘standardised matrix’ and the need to find uncontaminated and clean sediment sources is eliminated;
  - the tests can be initiated at any time without encountering seasonal variability in the test sediment and there is no need to pre-treat the sediment to remove indigenous fauna; the use of formulated sediment also reduces the cost associated with the field collection of sufficient amounts of sediment for routine testing;

**▼M4**

- the use of formulated sediment allows for comparisons of toxicity and ranking substances accordingly: toxicity data from tests with natural and artificial sediments were comparable for several chemicals (2).

6. Definitions used are given in Appendix 1.

## PRINCIPLE OF THE TEST

7. First instar chironomid larvae are exposed to a concentration range of the test substance in sediment-water systems. The test starts by placing first instar larvae into the test beakers containing the sediment-water system and subsequently spiking the test substance into the water. Chironomid emergence and development rate is measured at the end of the test. Larval survival and weight may also be measured after 10 days if required (using additional replicates as appropriate). These data are analysed either by using a regression model in order to estimate the concentration that would cause x % reduction in emergence, larvae survival or growth (e.g. EC<sub>15</sub>, EC<sub>50</sub>, etc.), or by using statistical hypothesis testing to determine a NOEC/LOEC. The latter requires comparison of effect values with control values using statistical tests.

## INFORMATION ON THE TEST SUBSTANCE

8. The water solubility of the test substance, its vapour pressure, measured or calculated partitioning into sediment and stability in water and sediment should be known. A reliable analytical method for the quantification of the test substance in overlying water, pore water and sediment with known and reported accuracy and limit of detection should be available. Useful information includes the structural formula and purity of the test substance. Chemical fate of the test substance (e.g. dissipation, abiotic and biotic degradation, etc.) also is useful information. Further guidance for testing substances with physical-chemical properties that make them difficult to perform the test is provided in (12).

## REFERENCE CHEMICALS

9. Reference chemicals may be tested periodically as a means of assuring that the test protocol and test conditions are reliable. Examples of reference toxicants used successfully in ring-tests and validation studies are: lindane, trifluralin, pentachlorophenol, cadmium chloride and potassium chloride. (1)(2)(5)(6)(13).

## VALIDITY OF THE TEST

10. For the test to be valid the following conditions apply:

- the emergence in the controls must be at least 70 % at the end of the test. (1)(6);

- *C. riparius* and *C. yoshimatsui* emergence to adults from control vessels should occur between 12 and 23 days after their insertion into the vessels; for *C. tentans*, a period of 20 to 65 days is necessary.

- at the end of the test, pH and the dissolved oxygen concentration should be measured in each vessel. The oxygen concentration should be at least 60 % of the air saturation value (ASV) at the temperature used, and the pH of overlying water should be in the 6-9 range in all test vessels;

**▼ M4**

- the water temperature should not differ by more than  $\pm 1,0$  °C. The water temperature could be controlled by isothermal room and in that case the room temperature should be confirmed in an appropriate time intervals.

## DESCRIPTION OF THE METHOD

**Test vessels**

11. The study is conducted in glass 600 ml beakers measuring 8 cm in diameter. Other vessels are suitable, but they should guarantee a suitable depth of overlying water and sediment. The sediment surface should be sufficient to provide 2 to 3 cm<sup>2</sup> per larvae. The ratio of the depth of the sediment layer to the depth of the overlying water should be 1:4. Test vessels and other apparatus that will come into contact with the test system should be made entirely of glass or other chemically inert material (e.g. Teflon).

**Selection of species**

12. The species to be used in the test is preferably *Chironomus riparius*. *Chironomus tentans* is also suitable but more difficult to handle and requires a longer test period. *Chironomus yohimatsui* may also be used. Details of culture methods are given in Appendix 2 for *Chironomus riparius*. Information on culture conditions is also available for other species, i.e. *Chironomus tentans* (4) and *Chironomus yohimatsui* (11). Identification of species must be confirmed before testing but is not required prior to every test if organisms come from an in-house culture.

**Sediment**

13. Formulated sediment (also called reconstituted, artificial or synthetic sediment) should preferably be used. However, if natural sediment is used, it should be characterised (at least pH, organic carbon content, determination of other parameters such as C/N ratio and granulometry are also recommended), and it should be free from any contamination and other organisms that might compete with, or consume the chironomids. It is also recommended that, before it is used in a chironomid toxicity test, the natural sediment be conditioned for seven days under the same conditions which prevail in the subsequent test. The following formulated sediment, based on the artificial soil used in Test Method C.8 (14), is recommended for use in this test (1)(15)(16):
  - a) 4-5 % (dry weight) peat: as close to pH 5,5 to 6,0 as possible; it is important to use peat in powder form, finely ground (particle size  $\leq 1$  mm) and only air dried.
  - b) 20 % (dry weight) kaolin clay (kaolinite content preferably above 30 %).
  - c) 75-76 % (dry weight) quartz sand (fine sand should predominate with more than 50 % of the particles between 50 and 200  $\mu\text{m}$ ).
  - d) Deionised water is added to obtain moisture of the final mixture in a range of 30-50 %.
  - e) Calcium carbonate of chemically pure quality ( $\text{CaCO}_3$ ) is added adjust the pH of the final mixture of the sediment to  $7,0 \pm 0,5$ .
  - f) Organic carbon content of the final mixture should be 2 % ( $\pm 0,5$  %) and is to be adjusted by the use of appropriate amounts of peat and sand, according to (a) and (c).

**▼ M4**

14. The source of peat, kaolin clay and sand should be known. The sediment components should be checked for the absence of chemical contamination (e.g. heavy metals, organochlorine compounds, organophosphorous compounds, etc.). An example for the preparation of the formulated sediment is described in Appendix 3. Mixing of dry constituents is also acceptable if it is demonstrated that after addition of overlying water a separation of sediment constituents (e.g. floating of peat particles) does not occur, and that the peat or the sediment is sufficiently conditioned.

**Water**

15. Any water which conforms to the chemical characteristics of acceptable dilution water as listed in Appendices 2 and 4 is suitable as test water. Any suitable water, natural water (surface or ground water), reconstituted water (see Appendix 2) or dechlorinated tap water are acceptable as culturing water and test water if chironomids will survive in it for the duration of the culturing and testing without showing signs of stress. At the start of the test, the pH of the test water should be between 6 and 9 and the total hardness not higher than 400 mg/l as CaCO<sub>3</sub>. However, if there is an interaction suspected between hardness ions and the test substance, lower hardness water should be used (and thus, Elendt Medium M4 must not be used in this situation). The same type of water should be used throughout the whole study. The water quality characteristics listed in Appendix 4 should be measured at least twice a year or when it is suspected that these characteristics may have changed significantly.

**Stock solutions — Spiked water**

16. Test concentrations are calculated on the basis of water column concentrations, i.e. the water overlying the sediment. Test solutions of the chosen concentrations are usually prepared by dilution of a stock solution. Stock solutions should preferably be prepared by dissolving the test substance in test medium. The use of solvents or dispersants may be required in some cases in order to produce a suitably concentrated stock solution. Examples of suitable solvents are acetone, ethanol, methanol, ethylene glycol monoethyl ether, ethylene glycol dimethyl ether, dimethylformamide and triethylene glycol. Dispersants which may be used are Cremophor RH40, Tween 80, methylcellulose 0,01 % and HCO-40. The solubilising agent concentration in the final test medium should be minimal (i.e.  $\leq 0,1$  ml/l) and should be the same in all treatments. When a solubilising agent is used, it must have no significant effects on survival or no visible adverse effect on the chironomid larvae as revealed by a solvent-only control. However, every effort should be made to avoid the use of such materials.

**TEST DESIGN**

17. The test design relates to the selection of the number and spacing of the test concentrations, the number of vessels at each concentration and the number of larvae per vessel. Designs for EC point estimation, for estimation of NOEC, and for conducting a limit test are described. The analysis by regression is preferred to the hypothesis testing approach.

*Design for analysis by regression*

18. The effect concentration (e.g. EC<sub>15</sub>, EC<sub>50</sub>) and the concentration range, over which the effect of the test substance is of interest, should be spanned by the concentrations included in the test. Generally, the accuracy and especially validity, with which estimates of effect concentrations (EC<sub>x</sub>) can be made, is improved when the effect concentration is within the range of concentrations

**▼ M4**

tested. Extrapolation much below the lowest positive concentration or above the highest concentration should be avoided. A preliminary range-finding test is helpful for selecting the range of concentrations to be used (see paragraph 27).

19. If the  $EC_x$  is to be estimated, at least five concentrations and three replicates for each concentration should be tested. In any case, it is advisable that sufficient test concentrations are used to allow a good model estimation. The factor between concentrations should not be greater than two (an exception could be made in cases when the dose response curve has a shallow slope). The number of replicates at each treatment can be reduced if the number of test concentrations with different responses is increased. Increasing the number of replicates or reducing the size of the test concentration intervals tends to lead to narrower confidence intervals for the test. Additional replicates are required if 10-day larval survival and growth are to be estimated.

*Design for estimation of a NOEC/LOEC*

20. If the LOEC/NOEC are to be estimated, five test concentrations with at least four replicates should be used and the factor between concentrations should not be greater than two. The number of replicates should be sufficient to ensure adequate statistical power to detect a 20 % difference from the control at the 5 % level of significance ( $p = 0,05$ ). With the development rate, an Analysis of Variance (ANOVA) is usually appropriate, such as Dunnett-test and Williams-test (17)(18)(19)(20). In the emergence ratio the Cochran-Armitage, Fisher's exact (with Bonferroni correction), or Mantel-Haenszel tests may be used.

*Limit test*

21. A limit test may be performed (one test concentration and control) if no effects were seen in the preliminary range-finding test. The purpose of the limit test is to indicate that the toxic value of the test substance is greater than the limit concentration tested. No suggestion for a recommended concentration can be made in this Test Method; this is left to the regulators' judgement. Usually, at least six replicates for both the treatment and control are necessary. Adequate statistical power to detect a 20 % difference from the control at the 5 % level of significance ( $p = 0,05$ ) should be demonstrated. With metric response (development rate and weight), the t-test is a suitable statistical method if data meet the requirements of this test (normality, homogeneous variances). The unequal-variance t-test or a non parametric test, such as the Wilcoxon-Mann-Whitney test may be used, if these requirements are not fulfilled. With the emergence ratio, the Fisher exact test is appropriate.

**PROCEDURE****Conditions of exposure***Preparation of spiked water-sediment system*

22. Appropriate amounts of formulated sediment (see paragraphs 13-14 and Appendix 3) are added in the test vessels to form a layer of at least 1,5 cm. Water is added to a depth of 6 cm (see paragraph 15). The ratio of the depth of the sediment layer and the depth of the water should not exceed 1:4 and the sediment layer should not be deeper than 3 cm. The sediment-water system should be left under gentle aeration for seven days prior to addition of test organisms (see paragraph 14 and Appendix 3). To avoid separation of sediment ingredients and re-suspension of fine material during addition of test water in the water column, the sediment can be covered with a plastic disc while water is poured onto it, and the disc is removed immediately afterwards. Other devices may also be appropriate.

**▼M4**

23. The test vessels should be covered (e.g. by glass plates). If necessary, during the study the water levels will be topped to the original volume in order to compensate for water evaporation. This should be performed using distilled or deionised water to prevent build-up of salts.

*Addition of test organisms*

24. Four to five days before adding the test organisms to the test vessels, egg masses should be taken from the cultures and placed in small vessels in culture medium. Aged medium from the stock culture or freshly prepared medium may be used. If the latter is used, a small amount of food e.g. green algae and/or a few droplets of filtrate from a finely ground suspension of flaked fish food should be added to the culture medium (see Appendix 2). Only freshly laid egg masses should be used. Normally, the larvae begin to hatch a couple of days after the eggs are laid (2 to 3 days for *Chironomus riparius* at 20 °C and 1 to 4 days for *Chironomus tentans* at 23 °C and *Chironomus yoshimatui* at 25 °C) and larval growth occurs in four instars, each of 4-8 days duration. First instar larvae (2-3 or 1-4 days post hatching) should be used in the test. The instar of midges can possibly be checked using head capsule width (6).
25. Twenty first instar larvae are allocated randomly to each test vessel containing the spiked sediment and water, using a blunt pipette. Aeration of the water has to be stopped while adding the larvae to test vessels and remain so for another 24 hours after addition of larvae (see paragraphs 24 and 32). According to the test design used (see paragraphs 19 and 20), the number of larvae used per concentration is at least 60 for the EC point estimation and 80 for determination of NOEC.
26. Twenty-four hours after adding the larvae, the test substance is spiked into the overlying water column, and slight aeration is again supplied. Small volumes of test substance solutions are applied below the surface of the water using a pipette. The overlying water should then be mixed with care not to disturb the sediment.

*Test concentrations*

27. A range-finding test may be helpful to determine the range of concentrations for the definitive test. For this purpose a series of widely spaced concentrations of the test substance are used. In order to provide the same density of surface per chironomids, which is to be used for the definitive test, chironomids are exposed to each concentration of the test substance for a period which allows estimation of appropriate test concentrations, and no replicates are required.
28. The test concentrations for the definitive test are decided based on the result of the range-finding test. At least five concentrations should be used and selected as described in paragraphs 18 to 20.

*Controls*

29. Control vessels without any test substance but including sediment should be included in the test with the appropriate number of replicates (see paragraphs 19-20). If a solvent has been used for application of test substance (see paragraph 16), a sediment solvent control should be added.

▼ **M4***Test system*

30. Static systems are used. Semi-static or flow-through systems with intermittent or continuous renewal of overlying water might be used in exceptional cases as for instance if water quality specifications become inappropriate for the test organism or affect chemical equilibrium (e.g. dissolved oxygen levels fall too low, the concentration of excretory products rises too high or minerals leach from sediment and affect pH and/or water hardness). However, other methods for ameliorating the quality of overlying water, such as aeration, will normally suffice and be preferable.

*Food*

31. It is necessary to feed the larvae, preferably daily or at least three times per week. Fish-food (a suspension in water or finely ground food, e.g. TetraMin or TetraPhyll; see details in Appendix 2) in the amount of 0,25-0,5 mg (0,35-0,5 mg for *C. yoshimatui*) per larvae per day seems adequate for young larvae for the first 10 days. Slightly more food may be necessary for older larvae: 0,5-1 mg per larvae per day should be sufficient for the rest of the test. The food ration should be reduced in all treatments and control if fungal growth is seen or if mortality is observed in controls. If fungal development cannot be stopped the test is to be repeated. When testing strongly adsorbing substances (e.g. with  $\log K_{ow} > 5$ ), or substances covalently binding to sediment, the amount of food necessary to ensure survival and natural growth of the organisms may be added to the formulated sediment before the stabilisation period. For this, plant material must be used instead of fish food, e.g. addition of 0,5 % (dry weight) finely ground leaves of e.g. stinging nettle (*Urtica dioica*), mulberry (*Morus alba*), white clover (*Trifolium repens*), spinach (*Spinacia oleracea*) or of other plant material (*Cerophyl* or alpha-cellulose) may be used.

*Incubation conditions*

32. Gentle aeration of the overlying water in test vessels is supplied preferably 24 hours after addition of the larvae and is pursued throughout the test (care should be taken that dissolved oxygen concentration does not fall below 60 % of ASV). Aeration is provided through a glass Pasteur pipette fixed 2-3 cm above the sediment layer (i.e. one or few bubbles/sec). When testing volatile chemicals, consideration may be given not to aerate the sediment-water system.
33. The test is conducted at a constant temperature of 20 °C ( $\pm 2$  °C). For *C. tentans* and *C. yoshimatui*, recommended temperatures are of 23 °C and 25 °C ( $\pm 2$  °C), respectively. A 16 hours photoperiod is used and the light intensity should be 500 to 1 000 lux.

*Exposure duration*

34. The exposure commences with the addition of larvae to the spiked and control vessels. The maximum exposure duration is 28 days for *C. riparius* and *C. yoshimatsui*, and 65 days for *C. tentans*. If midges emerge earlier, the test can be terminated after a minimum of five days after emergence of the last adult in the control.

## OBSERVATIONS

*Emergence*

35. The development time and the total number of fully emerged male and female midges are determined. Males are easily identified by their plumose antennae.



**▼ M4**

36. The test vessels should be observed at least three times per week to make visual assessment of any abnormal behaviour (e.g. leaving sediment, unusual swimming), compared with the control. During the period of expected emergence a daily count of emerged midges is necessary. The sex and number of fully emerged midges are recorded daily. After identification the midges are removed from the vessels. Any egg masses deposited prior to the termination of the test should be recorded and then removed to prevent re-introduction of larvae into the sediment. The number of visible pupae that have failed to emerge is also recorded. Guidance on measurement of emergence is provided in Appendix 5.

*Growth and survival*

37. If data on 10-day larval survival and growth are to be provided, additional test vessels should be included at the start, so that they may be used subsequently. The sediment from these additional vessels is sieved using a 250 µm sieve to retain the larvae. Criteria for death are immobility or lack of reaction to a mechanical stimulus. Larvae not recovered should also be counted as dead (larvae which have died at beginning of the test may have been degraded by microbes). The (ash free) dry weight of the surviving larvae per test vessel is determined and the mean individual dry weight per vessel calculated. It is useful to determine which instar the surviving larvae belong to; for that measurement of the width of the head capsule of each individual can be used.

**Analytical measurements***Concentration of the test substance*

38. As a minimum, samples of the overlying water, the pore water and the sediment must be analysed at the start (preferably one hour after application of test substance) and at the end of the test, at the highest concentration and a lower one. These determinations of test substance concentration inform on the behaviour/partitioning of the test substance in the water-sediment system. Sampling of sediment at the start of the test may influence the test system (e.g. removing test larvae), thus additional test vessels should be used to perform analytical determinations at the start and during the test if appropriate (see paragraph 39). Measurements in sediment might not be necessary if the partitioning of the test substance between water and sediment has been clearly determined in a water/sediment study under comparable conditions (e.g. sediment to water ratio, type of application, organic carbon content of sediment).
39. When intermediate measurements are made (e.g. at day 7) and if the analysis needs large samples which cannot be taken from test vessels without influencing the test system, analytical determinations should be performed on samples from additional test vessels treated in the same way (including the presence of test organisms) but not used for biological observations.
40. Centrifugation at e.g. 10 000 g and 4 °C for 30 min. is the recommended procedure to isolate interstitial water. However, if the test substance is demonstrated not to adsorb to filters, filtration may also be acceptable. In some cases it might not be possible to analyse concentrations in the pore water as the sample size is too small.

▼ **M4***Physical-chemical parameters*

41. The pH, dissolved oxygen in the test water and temperature of the test vessels should be measured in an appropriate manner (see paragraph 10). Hardness and ammonia should be measured in the controls and one test vessel at the highest concentration at the start and the end of the test.

## DATA AND REPORTING

*Treatment of results*

42. The purpose of this test is to determine the effect of the test substance on the development rate and the total number of fully emerged male and female midges, or in the case of the 10-day test effects on survival and weight of the larvae. If there are no indications of statistically different sensitivities of sexes, male and female results may be pooled for statistical analyses. The sensitivity differences between sexes can be statistically judged by e.g. a  $\chi^2$ -r  $\times$  2 table test. Larval survival and mean individual dry weight per vessel must be determined after 10 days where required.
43. Effect concentrations expressed as concentrations in the overlaying water, are calculated preferably based on measured concentrations at the beginning of the test (see paragraph 38).
44. To compute a point estimate for the  $EC_{50}$  or any other  $EC_x$ , the per-vessel statistics may be used as true replicates. In calculating a confidence interval for any  $EC_x$  the variability among vessels should be taken into account, or it should be shown that this variability is so small that it can be ignored. When the model is fitted by Least Squares, a transformation should be applied to the per-vessel statistics in order to improve the homogeneity of variance. However,  $EC_x$  values should be calculated after the response is transformed back to the original value.
45. When the statistical analysis aims at determining the NOEC/LOEC by hypothesis testing, the variability among vessels needs to be taken into account, e.g. by a nested ANOVA. Alternatively, more robust tests (21) can be appropriate in situations where there are violations of the usual ANOVA assumptions.

*Emergence ratio*

46. Emergence ratios are quantal data, and can be analyzed by the Cochran-Armitage test applied in step-down manner where a monotonic dose-response is expected and these data are consistent with this expectation. If not, a Fisher's exact or Mantel-Haenszel test with Bonferroni-Holm adjusted p-values can be used. If there is evidence of greater variability between replicates within the same concentration than a binomial distribution would indicate (often referenced as 'extra-binomial' variation), then a robust Cochran-Armitage or Fisher exact test such as proposed in (21), should be used.
47. The sum of midges emerged per vessel,  $n_e$ , is determined and divided by the number of larvae introduced,  $n_a$ :

$$ER = \frac{n_e}{n_a}$$

**▼ M4**

where:

$ER$  = emergence ratio

$n_e$  = number of midges emerged per vessel

$n_a$  = number of larvae introduced per vessel

48. An alternative that is most appropriate for large sample sizes, when there is extra binomial variance, is to treat the emergence ratio as a continuous response and use procedures such as William's test when a monotonic dose-response is expected and is consistent with these ER data. Dunnett's test would be appropriate where monotonicity does not hold. A large sample size is defined here as the number emerged and the number not emerging both exceeding five, on a per replicate (vessel) basis.
49. To apply ANOVA methods values of ER should first be transformed by the arcsin square root transformation or Freeman-Tukey transformation to obtain an approximate normal distribution and to equalise variances. The Cochran-Armitage, Fisher's exact (Bonferroni), or Mantel-Haenszel tests can be applied when using the absolute frequencies. The arcsin square root transformation is applied by taking the inverse sine ( $\sin^{-1}$ ) of the square root of ER.
50. For emergence ratios,  $EC_x$ -values are calculated using regression analysis (or e.g. probit (22), logit, Weibull, appropriate commercial software etc.). If regression analysis fails (e.g. when there are less than two partial responses), other non-parametric methods such as moving average or simple interpolation are used.

*Development rate*

51. The mean development time represents the mean time span between the introduction of larvae (day 0 of the test) and the emergence of the experimental cohort of midges. (For the calculation of the true development time, the age of larvae at the time of introduction should be considered). The development rate is the reciprocal of the development time (unit: 1/day) and represents that portion of larval development which takes place per day. The development rate is preferred for the evaluation of these sediment toxicity studies as its variance is lower, and it is more homogeneous and closer to normal distribution as compared to development time. Hence, powerful parametric test procedures may be used with development rate rather than with development time. For development rate as a continuous response,  $EC_x$ -values can be estimated by using regression analysis (e.g. (23)(24)).
52. For the following statistical tests, the number of midges observed on inspection day  $x$  are assumed to be emerged at the mean of the time interval between day  $x$  and day  $x - 1$  ( $l$  = length of the inspection interval, usually 1 day). The mean development rate per vessel ( $\bar{x}$ ) is calculated according to:

$$\bar{x} = \sum_{i=1}^m \frac{f_i x_i}{n_e}$$

where:

$\bar{x}$ : mean development rate per vessel

$i$ : index of inspection interval

$m$ : maximum number of inspection intervals

**▼ M4**

$f_i$ : number of midges emerged in the inspection interval  $i$

$n_e$ : total number of midges emerged at the end of experiment ( $= \sum f_i$ )

$x_i$ : development rate of the midges emerged in interval  $i$

$$x_i = 1 / \left( \text{day}_i - \frac{l_i}{2} \right)$$

where:

$\text{day}_i$ : inspection day (days since application)

$l_i$ : length of inspection interval  $i$  (days, usually 1 day)

**Test report**

53. *The test report must at least provide the following information:*

*Test substance:*

- physical nature and, where relevant, physical-chemical properties (water solubility, vapour pressure, partition coefficient in soil (or in sediment if available), stability in water, etc.);
- chemical identification data (common name, chemical name, structural formula, CAS number, etc.) including purity and analytical method for quantification of test substance.

*Test species:*

- test animals used: species, scientific name, source of organisms and breeding conditions;
- information on handling of egg masses and larvae;
- age of test animals when inserted into test vessels.

*Test conditions:*

- sediment used, i.e. natural or formulated sediment;
- for natural sediment, location and description of sediment sampling site, including, if possible, contamination history; characteristics: pH, organic carbon content, C/N ratio and granulometry (if appropriate).
- preparation of the formulated sediment: ingredients and characteristics (organic carbon content, pH, moisture, etc. at the start of the test);
- preparation of the test water (if reconstituted water is used) and characteristics (oxygen concentration, pH, conductivity, hardness, etc. at the start of the test);
- depth of sediment and overlying water;
- volume of overlying and pore water; weight of wet sediment with and without pore water;
- test vessels (material and size);
- method of preparation of stock solutions and test concentrations;

**▼ M4**

- application of test substance: test concentrations used, number of replicates and use of solvent if any;
- incubation conditions: temperature, light cycle and intensity, aeration (frequency and intensity);
- detailed information on feeding including type of food, preparation, amount and feeding regime.

*Results:*

- the nominal test concentrations, the measured test concentrations and the results of all analyses to determine the concentration of the test substance in the test vessel;
- water quality within the test vessels, i.e. pH, temperature, dissolved oxygen, hardness and ammonia;
- replacement of evaporated test water, if any;
- number of emerged male and female midges per vessel and per day;
- number of larvae which failed to emerge as midges per vessel;
- mean individual dry weight of larvae per vessel, and per instar, if appropriate;
- percent emergence per replicate and test concentration (male and female midges pooled);
- mean development rate of fully emerged midges per replicate and treatment rate (male and female midges pooled);
- estimates of toxic endpoints e.g.  $EC_x$  (and associated confidence intervals), NOEC and/or LOEC, and the statistical methods used for their determination;
- discussion of the results, including any influence on the outcome of the test resulting from deviations from this Test Method.

*LITERATURE:*

- (1) BBA (1995). Long-term toxicity test with *Chironomus riparius*: Development and validation of a new test system. Edited by M. Streløkke and H. Köpp. Berlin 1995.
- (2) Fleming R *et al.* (1994). Sediment Toxicity Tests for Poorly Water-Soluble Substances. Final Report to the European Commission. Report No: EC 3738. August 1994. WRc, UK.
- (3) SETAC (1993). Guidance Document on Sediment toxicity Tests and Bioassays for Freshwater and Marine Environments. From the WOSTA Workshop held in the Netherlands.
- (4) ASTM International/E1706-00 (2002). Test Method for Measuring the Toxicity of Sediment-Associated Contaminants with Freshwater Invertebrates. pp 1125-1241. In ASTM International 2002 Annual Book of Standards. Volume 11.05. Biological Effects and Environmental Fate; Biotechnology; Pesticides. ASTM International, West Conshohocken, PA.
- (5) Environment Canada (1997). Test for Growth and Survival in Sediment using Larvae of Freshwater Midges (*Chironomus tentans* or *Chironomus riparius*). Biological Test Method. Report SPE 1/RM/32. December 1997.
- (6) US-EPA (2000). Methods for Measuring the Toxicity and Bioaccumulation of Sediment-associated Contaminants with Freshwater Invertebrates. Second edition. EPA 600/R-99/064. March 2000. Revision to the first edition dated June 1994.

**▼M4**

- (7) US-EPA/OPPTS 850.1735. (1996): Whole Sediment Acute Toxicity Invertebrates.
- (8) US-EPA/OPPTS 850.1790. (1996): Chironomid Sediment toxicity Test.
- (9) Milani D, Day KE, McLeay DJ, Kirby RS (1996). Recent intra- and inter-laboratory studies related to the development and standardisation of Environment Canada's biological test methods for measuring sediment toxicity using freshwater amphipods (*Hyalella azteca*) and midge larvae (*Chironomus riparius*). Technical Report. Environment Canada. National Water Research Institute. Burlington, Ontario, Canada.
- (10) Sugaya Y (1997). Intra-specific variations of the susceptibility of insecticides in *Chironomus yoshimatsui*. Jp. J. Sanit. Zool. 48 (4): 345-350.
- (11) Kawai K (1986). Fundamental studies on Chironomid allergy. I. Culture methods of some Japanese Chironomids (Chironomidae, Diptera). Jp. J. Sanit. Zool. 37(1): 47-57.
- (12) OECD (2000). Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures. OECD Environment, Health and Safety Publications, Series on Testing and Assessment No 23.
- (13) Environment Canada (1995). Guidance Document on Measurement of Toxicity Test Precision Using Control Sediments Spiked with a Reference Toxicant. Report EPS 1/RM/30. September 1995.
- (14) Chapter C.8 of this Annex, Toxicity for Earthworms,
- (15) Suedel BC and Rodgers JH (1994). Development of formulated reference sediments for freshwater and estuarine sediment testing. Environ. Toxicol. Chem. 13: 1163-1175.
- (16) Naylor C and Rodrigues C (1995). Development of a test method for *Chironomus riparius* using a formulated sediment. Chemosphere 31: 3291-3303.
- (17) Dunnett CW (1964). A multiple comparisons procedure for comparing several treatments with a control. J. Amer. Statist. Assoc. 50: 1096-1121.
- (18) Dunnett CW (1964). New tables for multiple comparisons with a control. Biometrics 20: 482-491.
- (19) Williams DA (1971). A test for differences between treatment means when several dose levels are compared with a zero dose control. Biometrics 27: 103-117.
- (20) Williams DA (1972). The comparison of several dose levels with a zero dose control. Biometrics 28: 510-531.
- (21) Rao JNK and Scott AJ (1992). A simple method for the analysis of clustered binary data. Biometrics 48: 577-585.
- (22) Christensen ER (1984). Dose-response functions in aquatic toxicity testing and the Weibull model. Water Research 18: 213-221.
- (23) Bruce and Versteeg (1992). A statistical procedure for modelling continuous toxicity data. Environmental Toxicology and Chemistry 11:1485-1494.
- (24) Slob W (2002). Dose-response modelling of continuous endpoints. Toxicol. Sci. 66: 298-312.

**▼ M4***Appendix 1*

## DEFINITIONS

For the purpose of this method the following definitions are used:

**Formulated sediment** or reconstituted, artificial or synthetic sediment, is a mixture of materials used to mimic the physical components of a natural sediment.

**Overlying water** is the water placed over sediment in the test vessel.

**Interstitial water** or pore water is the water occupying space between sediment and soil particles.

**Spiked water** is the test water to which test substance has been added.

**Test chemical:** Any substance or mixture tested using this Test Method.

**▼ M4***Appendix 2***Recommendations for culture of *Chironomus riparius***

1. *Chironomus* larvae may be reared in crystallising dishes or larger containers. Fine quartz sand is spread in a thin layer of about 5 to 10 mm deep over the bottom of the container. Kieselguhr (e.g. Merck, Art 8117) has also been shown to be a suitable substrate (a thinner layer of up to a very few mm is sufficient). Suitable water is then added to a depth of several cm. Water levels should be topped up as necessary to replace evaporative loss, and prevent desiccation. Water can be replaced if necessary. Gentle aeration should be provided. The larval rearing vessels should be held in a suitable cage which will prevent escape of the emerging adults. The cage should be sufficiently large to allow swarming of emerged adults, otherwise copulation may not occur (minimum is ca. 30 × 30 × 30 cm).
2. Cages should be held at room temperature or in a constant environment room at 20 ± 2 °C with a photo period of 16 hour light (intensity ca. 1 000 lux), 8 hours dark. It has been reported that air humidity of less than 60 % RH can impede reproduction.

**Dilution water**

3. Any suitable natural or synthetic water may be used. Well water, dechlorinated tap water and artificial media (e.g. Elendt 'M4' or 'M7' medium, see below) are commonly used. The water has to be aerated before use. If necessary, the culture water may be renewed by pouring or siphoning the used water from culture vessels carefully without destroying the tubes of larvae.

**Feeding larvae**

4. *Chironomus* larvae should be fed with a fish flake food (TetraMin<sup>®</sup>, Tetra-Phyll<sup>®</sup> or other similar brand of proprietary fish food), at approximately 250 mg per vessel per day. This can be given as a dry ground powder or as a suspension in water: 1,0 g of flake food is added to 20 ml of dilution water and blended to give a homogenous mix. This preparation may be fed at a rate of about 5 ml per vessel per day (shake before use.) Older larvae may receive more.
5. Feeding is adjusted according to the water quality. If the culture medium becomes 'cloudy', the feeding should be reduced. Food additions must be carefully monitored. Too little food will cause emigration of the larvae towards the water column, and too much food will cause increased microbial activity and reduced oxygen concentrations. Both conditions can result in reduced growth rates.
6. Some green algae (e.g. *Scenedesmus subspicatus*, *Chlorella vulgaris*) cells may also be added when new culture vessels are set up.

**Feeding emerged adults**

7. Some experimenters have suggested that a cotton wool pad soaked in a saturated sucrose solution may serve as a food for emerged adults.



**▼ M4****Emergence**

8. At  $20 \pm 2$  °C adults will begin to emerge from the larval rearing vessels after approximately 13-15 days. Males are easily distinguished by having plumose antennae.

**Egg masses**

9. Once adults are present within the breeding cage, all larval rearing vessels should be checked three times weekly for deposition of the gelatinous egg masses. If present, the egg masses should be carefully removed. They should be transferred to a small dish containing a sample of the breeding water. Egg masses are used to start a new culture vessel (e.g. 2-4 egg masses/vessel) or are used for toxicity tests.
10. First instar larvae should hatch after 2-3 days.

**Set-up of new culture vessels**

11. Once cultures are established it should be possible to set up a fresh larval culture vessel weekly or less frequently depending on testing requirements, removing the older vessels after adult midges have emerged. Using this system a regular supply of adults will be produced with a minimum of management.

**Preparation of test solutions 'M4' and 'M7'**

12. Elendt (1990) has described the 'M4' medium. The 'M7' medium is prepared as the 'M4' medium except for the substances indicated in Table 1, for which concentrations are four times lower in 'M7' than in 'M4'. A publication on the 'M7' medium is in preparation (Elendt, personal communication). The test solution should not be prepared according to Elendt and Bias (1990) for the concentrations of  $\text{NaSiO}_3 \cdot 5 \text{H}_2\text{O}$ ,  $\text{NaNO}_3$ ,  $\text{KH}_2\text{PO}_4$  and  $\text{K}_2\text{HPO}_4$  given for the preparation of the stock solutions are not adequate.

**Preparation of the 'M7'-medium**

13. Each stock solution (I) is prepared individually and a combined stock solution (II) is prepared from these stock solutions (I) (see Table 1). 50 ml from the combined stock Solution (II) and the amounts of each macro nutrient stock solution which are given in Table 2 are made up to 1 l of deionised water to prepare the 'M7' medium. A vitamin stock solution is prepared by adding three vitamins to deionised water as indicated in Table 3, and 0,1 ml of the combined vitamin stock solution are added to the final 'M7' medium shortly before use. (The vitamin stock solution is stored frozen in small aliquots). The medium is aerated and stabilised.

*Table 1***Stock solutions of trace elements for medium M4 and M7**

Stock solutions (I)	Amount (mg) made up to 1 litre of deionised water	To prepare the combined stock solution (II): mix the following amounts (ml) of stock solutions (I) and make up to 1 litre of deionised water		Final concentrations in test solutions (mg/l)	
		M4	M7	M4	M7
$\text{H}_3\text{BO}_3$ <sup>(1)</sup>	57 190	1,0	0,25	2,86	0,715
$\text{MnCl}_2 \cdot 4 \text{H}_2\text{O}$ <sup>(1)</sup>	7 210	1,0	0,25	0,361	0,090
$\text{LiCl}$ <sup>(1)</sup>	6 120	1,0	0,25	0,306	0,077

▼ **M4**

Stock solutions (I)	Amount (mg) made up to 1 litre of deionised water	To prepare the combined stock solution (II): mix the following amounts (ml) of stock solutions (I) and make up to 1 litre of deionised water		Final concentrations in test solutions (mg/l)	
		M4	M7	M4	M7
RbCl <sup>(1)</sup>	1 420	1,0	0,25	0,071	0,018
SrCl <sub>2</sub> · 6 H <sub>2</sub> O <sup>(1)</sup>	3 040	1,0	0,25	0,152	0,038
NaBr <sup>(1)</sup>	320	1,0	0,25	0,016	0,004
Na <sub>2</sub> MoO <sub>4</sub> · 2 H <sub>2</sub> O <sup>(1)</sup>	1 260	1,0	0,25	0,063	0,016
CuCl <sub>2</sub> · 2 H <sub>2</sub> O <sup>(1)</sup>	335	1,0	0,25	0,017	0,004
ZnCl <sub>2</sub>	260	1,0	1,0	0,013	0,013
CaCl <sub>2</sub> · 6 H <sub>2</sub> O	200	1,0	1,0	0,010	0,010
KI	65	1,0	1,0	0,0033	0,0033
Na <sub>2</sub> SeO <sub>3</sub>	43,8	1,0	1,0	0,0022	0,0022
NH <sub>4</sub> VO <sub>3</sub>	11,5	1,0	1,0	0,00058	0,00058
Na <sub>2</sub> EDTA · 2 H <sub>2</sub> O <sup>(1)</sup> <sup>(2)</sup>	5 000	20,0	5,0	2,5	0,625
FeSO <sub>4</sub> · 7 H <sub>2</sub> O <sup>(1)</sup> <sup>(2)</sup>	1 991	20,0	5,0	1,0	0,249

<sup>(1)</sup> These substances differ in M4 and M7, as indicated above.

<sup>(2)</sup> These solutions are prepared individually, then poured together and autoclaved immediately.

Table 2

**Macro nutrient stock solutions for medium M4 and M7**

	Amount made up to 1 litre of deionised water (mg)	Amount of macro nutrient stock solutions added to prepare medium M4 and M7 (ml/l)	Final concentrations in test solutions M4 and M7 (mg/l)
CaCl <sub>2</sub> · 2 H <sub>2</sub> O	293 800	1,0	293,8
MgSO <sub>4</sub> · 7 H <sub>2</sub> O	246 600	0,5	123,3
KCl	58 000	0,1	5,8
NaHCO <sub>3</sub>	64 800	1,0	64,8
NaSiO <sub>3</sub> · 9 H <sub>2</sub> O	50 000	0,2	10,0
NaNO <sub>3</sub>	2 740	0,1	0,274
KH <sub>2</sub> PO <sub>4</sub>	1 430	0,1	0,143
K <sub>2</sub> HPO <sub>4</sub>	1 840	0,1	0,184

▼ **M4**

Table 3

**Vitamin stock solution for medium M4 and M7**

All three vitamin solutions are combined to make a single vitamin stock solution.

	Amount made up to 1 litre of deionised water (mg)	Amount of vitamin stock solution added to prepare medium M4 and M7 (ml/l)	Final concentrations in test solutions M4 and M7 (mg/l)
Thiamine hydrochloride	750	0,1	0,075
Cyanocobalamin (B12)	10	0,1	0,0010
Biotine	7,5	0,1	0,00075

*LITERATURE:*

BBA (1995). Long-term toxicity test with *Chironomus riparius*: Development and validation of a new test system. Edited by M. Streloke and H.Köpp. Berlin 1995.

Elenđt BP (1990). Selenium Deficiency in Crustacean. *Protoplasma* 154: 25-33.

Elenđt BP and Bias W-R (1990). Trace Nutrient Deficiency in *Daphnia magna* Cultured in Standard Medium for Toxicity Testing. Effects on the Optimization of Culture Conditions on Life History Parameters of *D. magna*. *Water Research* 24 (9): 1157-1167.

▼ **M4***Appendix 3***PREPARATION OF FORMULATED SEDIMENT****Sediment composition**

The composition of the formulated sediment should be as follows:

Constituent	Characteristics	% of sediment dry weight
Peat	Sphagnum moss peat, as close to pH 5,5-6,0 as possible, no visible plant remains, finely ground (particle size $\leq$ 1 mm) and air dried	4-5
Quartz sand	Grain size: > 50 % of the particles should be in the range of 50-200 $\mu$ m	75-76
Kaolinite clay	Kaolinite content $\geq$ 30 %	20
Organic carbon	Adjusted by addition of peat and sand	2 ( $\pm$ 0,5)
Calcium carbonate	CaCO <sub>3</sub> , pulverised, chemically pure	0,05-0,1
Water	Conductivity $\leq$ 10 $\mu$ S/cm	30-50

**Preparation**

The peat is air dried and ground to a fine powder. A suspension of the required amount of peat powder in deionised water is prepared using a high-performance homogenising device. The pH of this suspension is adjusted to  $5,5 \pm 0,5$  with CaCO<sub>3</sub>. The suspension is conditioned for at least two days with gentle stirring at  $20 \pm 2$  °C, to stabilise pH and establish a stable microbial component. pH is measured again and should be  $6,0 \pm 0,5$ . Then the peat suspension is mixed with the other constituents (sand and kaolin clay) and deionised water to obtain a homogeneous sediment with a water content in a range of 30-50 per cent of dry weight of the sediment. The pH of the final mixture is measured once again and is adjusted to 6,5 to 7,5 with CaCO<sub>3</sub> if necessary. Samples of the sediment are taken to determine the dry weight and the organic carbon content. Then, before it is used in the chironomid toxicity test, it is recommended that the formulated sediment be conditioned for seven days under the same conditions which prevail in the subsequent test.

**Storage**

The dry constituents for preparation of the artificial sediment may be stored in a dry and cool place at room temperature. The formulated (wet) sediment should not be stored prior to its use in the test. It should be used immediately after the 7 days conditioning period that ends its preparation.

*LITERATURE:*

Chapter C.8 of this Annex, Toxicity for Earthworms

Meller M, Egeler P, Rombke J, Schallnass H, Nagel R and Streit B (1998). Short-term Toxicity of Lindane, Hexachlorobenzene and Copper Sulfate on Tubificid Sludgeworms (Oligochaeta) in Artificial Media. *Ecotox. and Environ. Safety* 39: 10-20.

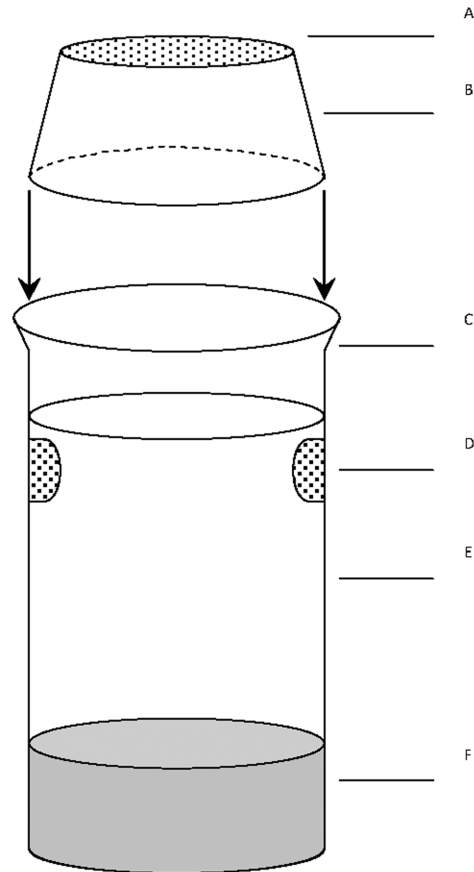
**▼ M4***Appendix 4***Chemical Characteristics of Acceptable Dilution Water**

Substance	Concentrations
Particulate matter	< 20 mg/l
Total organic carbon	< 2 mg/l
Unionised ammonia	< 1 µg/l
Hardness as CaCO <sub>3</sub>	< 400 mg/l (*)
Residual chlorine	< 10 µg/l
Total organophosphorus pesticides	< 50 ng/l
Total organochlorine pesticides plus polychlorinated biphenyls	< 50 ng/l
Total organic chlorine	< 25 ng/l

(\*) However, it should be noted that if there is an interaction suspected between hardness ions and the test substance, lower hardness water should be used (and thus, Elendt Medium M4 must not be used in this situation).

▼ **M4***Appendix 5***Guidance for monitoring emergence of chironomid larvae**

Emergence traps are placed on the test beakers. These traps are needed from day 20 to the end of the test. Example of trap used is drawn below:



A: the nylon screen

B: the inverted plastic cups

C: the lipless exposure beaker

D: the water exchange screen ports

E: water

F: sediment

**▼ M4****C.29. READY BIODEGRADABILITY — CO<sub>2</sub> IN SEALED VESSELS (HEADSPACE TEST)**

## INTRODUCTION

1. This Test Method is equivalent to OECD Test Guideline (TG) 310 (2006). This Test Method is a screening method for the evaluation of ready biodegradability of chemicals and provides similar information to the six test methods described in chapter C.4 of this Annex A to F. Therefore, a chemical that shows positive results in this Test Method can be considered readily biodegradable and consequently rapidly degradable in the environment.
2. The well established carbon dioxide (CO<sub>2</sub>) method (1), based on Sturm's original test (2) for assessing biodegradability of organic chemicals, by the measurement of the carbon dioxide produced by microbial action, has normally been the first choice for testing poorly soluble chemicals and those which strongly adsorb. It is also chosen for soluble (but not volatile) chemicals, since the evolution of carbon dioxide is considered by many to be the only unequivocal proof of microbial activity. Removal of dissolved organic carbon can be effected by physico-chemical processes — adsorption, volatilisation, precipitation, hydrolysis — as well as by microbial action and many non-biological reactions consume oxygen; rarely is CO<sub>2</sub> produced from organic chemicals abiotically. In the original and modified Sturm test (1)(2) CO<sub>2</sub> is removed from the liquid phase to the absorbing vessels by sparging (i.e. bubbling air treated to remove CO<sub>2</sub> through the liquid medium), while in the version of Larson (3)(4) CO<sub>2</sub> is transferred from the reaction vessel to the absorbers by passing CO<sub>2</sub>-free air through the headspace and, additionally, by shaking the test vessel continuously. Only in the Larson modification is the reaction vessel shaken; stirring is specified only for insoluble substances in ISO 9439 (5) and in the original US version (6), both of which specify sparging rather than headspace replacement. In another official US EPA method (7) based on Gledhill's method (8), the shaken reaction vessel is closed to the atmosphere and CO<sub>2</sub> produced is collected in an internal alkaline trap directly from the gaseous phase, as in classical Warburg/Barcroft respirometer flasks.
3. However, inorganic carbon (IC) has been shown to accumulate in the medium during the application of the standard, modified Sturm test to a number of chemicals (9). A concentration of IC as high as 8 mg/l was found during the degradation of 20 mg C/l of aniline. Thus, the collection of CO<sub>2</sub> in the alkaline traps did not give a true reflection of the amount of CO<sub>2</sub> produced microbiologically at intermediate times during the degradation. As a result, the specification that > 60 % theoretical maximum CO<sub>2</sub> production (ThCO<sub>2</sub>) must be collected within a '10-d window' (the 10 days immediately following the attainment of 10 % biodegradation) for a test chemical to be classified as readily biodegraded will not be met for some chemicals which would be so classified using dissolved organic carbon (DOC) removal.
4. When the percentage degradation is a lower value than expected, IC is possibly accumulated in the test solution. Then, the degradability may be assessed with the other ready biodegradability tests.

**▼ M4**

5. Other drawbacks of the Sturm methodology (cumbersome, time-consuming, more prone to experimental error and not applicable to volatile chemicals) had earlier prompted a search for a sealed vessel technique, other than Gledhill's, rather than gas flow-through (10)(11). Boatman et al (12) reviewed the earlier methods and adopted an enclosed headspace system in which the CO<sub>2</sub> was released into the headspace at the end of incubation by acidifying the medium. CO<sub>2</sub> was measured by gas chromatography (GC)/IC analysis in automatically taken samples of the headspace but dissolved inorganic carbon (DIC) in the liquid phase was not taken into account. Also, the vessels used were very small (20 ml) containing only 10 ml of medium, which caused problems e.g. when adding the necessarily very small amounts of insoluble test chemicals, and/or there may be insufficient or no microorganisms present in the inoculated medium that are competent to degrade the test chemicals.
6. These difficulties have been overcome by the independent studies of Struijs and Stoltenkamp (13) and of Birch and Fletcher (14), the latter being inspired by their experience with apparatus used in the anaerobic biodegradation test (15). In the former method (13) CO<sub>2</sub> is measured in the headspace after acidification and equilibration, while in the latter (14) DIC in both the gaseous and liquid phases was measured, without treatment; over 90 % of the IC formed was present in the liquid phase. Both methods had advantages over the Sturm test in that the test system was more compact and manageable, volatile chemicals can be tested and the possibility of delay in measuring CO<sub>2</sub> produced is avoided.
7. The two approaches were combined in the ISO Headspace CO<sub>2</sub> Standard (16), which was ring-tested (17) and it is this Standard which forms the basis of the present Test Method. Similarly, the two approaches have been used in the US EPA method (18). Two methods of measuring CO<sub>2</sub> have been recommended, namely CO<sub>2</sub> in headspace after acidification (13) and IC in the liquid phase after the addition of excess alkali. The latter method was introduced by Peterson during the CONCAWE ring test (19) of this headspace method modified to measure inherent biodegradability. The changes made in the 1992 (20) revision of the methods in chapter C.4 of this Annex for Ready Biodegradability have been incorporated into this Test Method, so that the conditions (medium, duration etc.) are otherwise the same as those in the revised Sturm test (20). Birch and Fletcher (14) have shown that very similar results were obtained with this headspace test as were obtained with the same chemicals in the OECD Ring Test (21) of the revised Test Methods.

**PRINCIPLE OF THE TEST**

8. The test chemical, normally at 20 mg C/l, as the sole source of carbon and energy, is incubated in a buffer-mineral salts medium which has been inoculated with a mixed population of micro-organisms. The test is performed in sealed bottles with a headspace of air, which provides a reservoir of oxygen for aerobic biodegradation. The CO<sub>2</sub> evolution resulting from the ultimate aerobic biodegradation of the test chemical is determined by measuring the IC produced in the test bottles in excess of that produced in blank vessels containing inoculated medium only. The extent of biodegradation is expressed as a percentage of the theoretical maximum IC production (ThIC), based on the quantity of test chemical (as organic carbon) added initially.
9. The DOC removal and/or the extent of primary biodegradation of the test chemical can also be measured (20).



**▼ M4**

## INFORMATION ON THE TEST CHEMICAL

10. The organic carbon content (% w/w) of the test chemical needs to be known, either from its chemical structure or by measurement, so that the percentage degradation may be calculated. For volatile test chemicals, a measured or calculated Henry's law constant is helpful for determining a suitable headspace to liquid volume ratio. Information on the toxicity of the test chemical to micro-organisms is useful in selecting an appropriate test concentration and for interpreting results showing poor biodegradability: it is recommended to include the inhibition control unless it is known that the test chemical is not inhibitory to microbial activities (see paragraph 24).

## APPLICABILITY OF THE METHOD

11. The test is applicable to water-soluble and insoluble test chemicals, though good dispersion of the test chemical should be ensured. Using the recommended headspace to liquid volume ratio of 1:2, volatile chemicals with a Henry's law constant of up to  $50 \text{ Pa}\cdot\text{m}^3\cdot\text{mol}^{-1}$  can be tested as the proportion of test chemical in the headspace will not exceed 1 % (13). A smaller headspace volume may be used when testing chemicals, which are more volatile, but their bioavailability may be limiting especially if they are poorly soluble in water. However, users must ensure that the headspace to liquid volume ratio and the test chemical concentration are such that sufficient oxygen is available to allow complete aerobic biodegradation to occur (e.g. avoid using a high substrate concentration and a small headspace volume). Guidance on this matter can be found in (13)(23).

## REFERENCE CHEMICALS

12. In order to check the test procedure, a reference chemical of known biodegradability should be tested in parallel. For this purpose, aniline, sodium benzoate or ethylene glycol may be used when testing water-soluble test chemicals and 1-octanol for poorly soluble test chemicals (13). Biodegradation of these chemicals must reach > 60 % ThC within 14 days.

## REPRODUCIBILITY

13. In the ISO ring test of the method (17), the following results were obtained using the recommended conditions, including 20 mg C test chemical/l.

Test Chemical	Mean Percentage Biodegradation (28d)	Coefficient of variation (%)	Number of Laboratories
Aniline	90	16	17
1-Octanol	85	12	14

Within-test variability (replicability), using aniline, was low with coefficients of variability not greater than 5 % in nearly all test runs. In the two cases in which the replicability was worse, the greater variability was probably due to high IC production in the blanks. Replicability was worse with 1-octanol but was still less than 10 % for 79 % of test runs. This greater within-test variability may have been due to dosing errors, as a small volume (3 to 4  $\mu\text{l}$ ) of 1-octanol had to be injected into sealed test bottles. Higher coefficients of variation would result when lower concentrations of test chemical are used, especially at concentrations lower than 10 mg C/l. This could be partially overcome by reducing the concentration of total inorganic carbon (TIC) in the inoculum.

▼ **M4**

14. In an EU ring-test (24) of five surfactants added at 10 mg C/l, the following results were obtained:

Test Chemical	Mean Percentage biodegradation (28d)	Coefficient of variation (%)	Number of laboratories
Tetrapropylene Benzene sulphonate	17	45	10
Di-iso-octylsulpho-Succinate (anionic)	72	22	9
Hexadecyl-trimethyl (*) Ammonium chloride (cationic)	75	13	10
Iso-Nonylphenol (ethoxylate) <sub>9</sub> (non-ionic)	41	32	10
Coco-amide-propyl Dimethylhydroxy Sulphobetaine (amphoteric)	60	23	11

(\*) SiO<sub>2</sub> was added to neutralize toxicity.

The results show that generally, the variability was higher for the less well-degraded surfactants. Within-test variability was less than 15 % for over 90 % of cases, the highest reaching 30-40 %.

NOTE: Most surfactants are not single molecular species but are mixtures of isomers, homologues, etc. which degrade after different characteristic lag periods and at different kinetic rates resulting in 'blurred', extenuated curves, so that the 60 % pass value may not be reached within 'the 10-d window', even though each individual molecular species would reach > 60 % within 10 days if tested alone. This may be observed with other complex mixtures as well.

## DESCRIPTION OF THE METHOD

*Apparatus*

15. Normal laboratory apparatus and:
- (a) Glass serum bottles, sealed with butyl rubber stoppers and crimp-on aluminium seals. The recommended size is '125 ml' which have a total volume of around 160 ml (in this case the volume of each bottle should be known to be  $160 \pm 1$  ml). A smaller size of vessel may be used when the results fulfil the conditions described in paragraph 66 and 67;
  - (b) Carbon analyser or other instrument (e.g. gas chromatograph) for measuring inorganic carbon;

**▼M4**

- (c) Syringes of high precision for gaseous and liquid samples;
- (d) Orbital shaker in a temperature-controlled environment;
- (e) A supply of CO<sub>2</sub> free air — this can be prepared by passing air through soda lime granules or by using an 80 % N<sub>2</sub>/20 % O<sub>2</sub> gas mixture (optional) (see paragraph 28);
- (f) Membrane-filtration device of 0,20–0,45 µm porosity (optional);
- (g) Organic carbon analyser (optional).

*Reagents*

16. Use analytical grade reagents throughout.

*Water*

17. Distilled or de-ionised water should be used containing ≤ 1 mg/l as total organic carbon. This represents ≤ 5 % of the initial organic carbon content introduced by the recommended dose of the test chemical.

*Stock solutions for the mineral salts medium*

18. The stock solutions and the mineral salts medium are similar to those in ISO 14593 (16) and C.4 'ready biodegradability' tests (20). The use of a higher concentration of ammonium chloride (2,0 g/l instead of 0,5 g/l) should only be necessary in very exceptional cases, e.g. when the test chemical concentration is > 40 mg C/l. Stock solutions should be stored under refrigeration and disposed of after six months, or earlier if there is evidence of precipitation or microbial growth. Prepare the following stock solutions:

- (a) Potassium dihydrogen phosphate (KH<sub>2</sub>PO<sub>4</sub>) 8,50 g

Dipotassium hydrogen phosphate (K<sub>2</sub>HPO<sub>4</sub>) 21,75 g

Disodium hydrogen phosphate dihydrate (Na<sub>2</sub>HPO<sub>4</sub>·2H<sub>2</sub>O) 33,40 g

Ammonium chloride (NH<sub>4</sub>Cl) 0,50 g

Dissolve in water and make up to 1 litre. The pH of this solution should be 7,4 (± 0,2). If this is not the case, then prepare a new solution.

- (b) Calcium chloride dihydrate (CaCl<sub>2</sub>·2H<sub>2</sub>O) 36,40 g

Dissolve in water and make up to 1 litre.

- (c) Magnesium sulphate heptahydrate (MgSO<sub>4</sub>·7H<sub>2</sub>O) 22,50 g

Dissolve in water and make up to 1 litre.

- (d) Iron (III) chloride hexahydrate (FeCl<sub>3</sub>·6H<sub>2</sub>O) 0,25 g

Dissolve in water and make up to 1 litre and add one drop of concentrated HCl.

*Preparation of mineral medium*

19. Mix 10 ml of solution (a) with approximately 800 ml water (paragraph 17), then add 1 ml of solutions (b), (c) and (d) and make up to 1 litre with water (paragraph 17).

*Other reagents*

20. Concentrated ortho-phosphoric acid (H<sub>3</sub>PO<sub>4</sub>) (> 85 % mass per volume).

**▼M4***Sodium hydroxide solution 7M*

21. Dissolve 280 g of sodium hydroxide (NaOH) in 1 litre of water (paragraph 17). Determine the concentration of DIC of this solution and consider this value when calculating the test result (see paragraphs 55 and 61), especially in the light of the validity criterion in paragraph 66 (b). Prepare a fresh solution if the concentration of DIC is too high.

*Test chemical*

22. Prepare a stock solution of a sufficiently water-soluble test chemical in water (paragraph 17) or in the test medium (paragraph 19) at a concentration preferably 100-fold greater than the final concentration to be used in the test; it may be necessary to adjust the pH of the stock solution. The stock solution should be added to the mineral medium to give a final organic carbon concentration of between 2 and 40 mg C/l, preferably 20 mg C/l. If concentrations lower than these are used, the precision obtained may be impaired. Soluble and insoluble liquid chemicals may be added to the vessels directly using high precision syringes. Poorly soluble and insoluble test chemicals may require special treatment (25). The choices are:

- (a) direct addition of known weighed amounts;
- (b) ultrasonic dispersion before addition;
- (c) dispersion with the aid of emulsifying agents to be required to establish whether they have any inhibitory or stimulatory effects on microbial activity before addition;
- (d) adsorption of liquid test chemicals, or a solution in a suitable volatile solvent, on to an inert medium or support (e.g. glass fibre filter), followed by evaporation of the solvent, if used, and direct addition of known amounts;
- (e) addition of known volume of a solution of the test chemical in an easily volatile solvent to an empty test vessel, followed by evaporation of the solvent.

Agents or solvents used in (c), (d) and (e) have to be tested for any stimulatory or inhibitory effect on microbial activity (see paragraph 42(b).)

*Reference chemical*

23. Prepare a stock solution of the (soluble) reference chemical in water (paragraph 17) at a concentration preferably 100-fold greater than the final concentration to be used (20 mg C/l) in the test.

*Inhibition check*

24. Test chemicals frequently show no significant degradation under the conditions used in ready biodegradation assessments. One possible cause is that the test chemical is inhibitory to the inoculum at the concentration at which it is applied in the test. An inhibition check may be included in the test design to facilitate identification (in retrospect) of inhibition as a possible cause or contributory factor. Alternatively, the inhibition check may rule out such interferences and show that zero or slight degradation is attributable solely to non-amenability to microbial attack under the conditions of the test. In order to obtain information on the toxicity of the test chemical to (aerobic) micro-organisms, prepare a solution in the test medium containing the test chemical and the reference chemical (paragraph 19), each at the same concentrations as added, respectively (see paragraph 22 and 23).

**▼ M4***Inoculum*

25. The inoculum may be derived from a variety of sources: activated sludge; sewage effluent (non-chlorinated); surface waters and soils; or from a mixture of these (20). The biodegradative activity of the source should be checked by using a reference chemical. Whatever the source, micro-organisms previously exposed to the test chemical should not be used if the procedure is to be used as a test for ready biodegradability.

Warning: Activated sludge, sewage and sewage effluent contain pathogenic organisms and must be handled with caution.

26. Based on experience, the optimal volume for the inoculum is that which:
- is sufficient to give adequate biodegradative activity;
  - degrades the reference chemical by the stipulated percentage (see paragraph 66);
  - gives  $10^2$  to  $10^5$  colony-forming units per millilitre in the final mixture;
  - normally gives a concentration of 4 mg/l suspended solids in the final mixture when activated sludge is used, concentrations up to 30 mg/l may be used but may significantly increase  $\text{CO}_2$  production of the blanks (26);
  - contributes less than 10 % of the initial concentration of organic carbon introduced by the test chemical;
  - is generally 1-10 ml of inoculum for 1 litre of test solution.

*Activated sludge*

27. Activated sludge is freshly collected from the aeration tank of a sewage treatment plant or laboratory-scale unit treating predominantly domestic sewage. If necessary, coarse particles should be removed by sieving (e.g. using a 1 mm<sup>2</sup> mesh sieve) and the sludge should be kept aerobic until used.
28. Alternatively, after removal of any coarse particles, settle or centrifuge (e.g. 1 100 × g for 10 minutes). Discard the supernatant liquid. The sludge may be washed in the mineral solution. Suspend the concentrated sludge in mineral medium to yield a concentration of 3-5 g suspended solids/l. Thereafter aerate until required.
29. Sludge should be taken from a properly working conventional treatment plant. If sludge has to be taken from a high rate treatment plant, or is thought to contain inhibitors, it should be washed. Settle or centrifuge the re-suspended sludge after thorough mixing, discard the supernatant liquid and again suspend the washed sludge in a further volume of mineral medium. Repeat this procedure until the sludge is considered to be free from excess substrate or inhibitor.
30. After complete re-suspension is achieved, or with untreated sludge, withdraw a sample just before use for the determination of the dry weight of the suspended solids.
31. A further alternative is to homogenise activated sludge (3-5 g suspended solids/l). Treat the sludge in a Waring blender for 2 minutes at medium speed. Settle the blended sludge for 30 minutes or longer if required and decant liquid for use as inoculum at the rate of about 10 mg/l of mineral medium.

**▼ M4**

32. Still further reduction of the blank CO<sub>2</sub> evolution can be achieved by aerating the sludge overnight with CO<sub>2</sub>-free air. Use 4 mg/l activated sludge solids as the concentration of the inoculum in this test (13).

*Secondary sewage effluent*

33. Alternatively, the inoculum can be derived from the secondary effluent of a treatment plant or laboratory-scale unit receiving predominantly domestic sewage. Maintain the sample under aerobic conditions and use on the day of collection, or pre-condition if necessary. The effluent should be filtered through a coarse filter to remove gross particulate matter and the pH value is measured.
34. To reduce its IC content, the filtrate is sparged with CO<sub>2</sub>-free air (paragraph 15-e) for 1 h while maintaining the pH at 6,5 using orthophosphoric acid (paragraph 20). The pH value is restored to its original value with sodium hydroxide (paragraph 21) and after settling for about 1 h a suitable volume of the supernatant is taken for inoculation. This sparging procedure reduces the IC content of the inoculum. For example, when the maximum recommended volume of filtered sparged effluent (100 ml) per litre was used as inoculum, the amount of IC present in blank control vessels was in the range 0,4 to 1,3 mg/l (14), representing 2-6,5 % of test chemical C at 20 mg C/l and 4-13 % at 10 mg C/l.

*Surface waters*

35. A sample is taken of an appropriate surface water. It should be kept under aerobic conditions and used on the day of collection. The sample should be concentrated, if necessary, by filtration or centrifugation. The volume of inoculum to be used in each test vessel should meet the criteria given in paragraph 26.

*Soils*

36. A sample is taken of an appropriate soil, collected to a depth of up to 20 cm below the soil surface. Stones, plant remains and invertebrates should be removed from the sample of soil before it is sieved through a 2 mm mesh (if the sample is too wet to sieve immediately, then partially air dry to facilitate sieving). It should be kept under aerobic conditions and used on the day of collection (If the sample is transported in a loosely-tied black polythene bag, it can be stored at 2 to 4 °C in the bag for up to one month).

*Preconditioning of inoculum*

37. Inoculum may be pre-conditioned to the experimental conditions, but not pre-adapted to the test chemical. Pre-conditioning can reduce the blank CO<sub>2</sub> evolution. Pre-conditioning consists of aerating activated sludge after diluting in test medium to 30 mg/l with moist CO<sub>2</sub>-free air for up to 5-7 days at the test temperature.

**TEST PROCEDURE***Number of bottles*

38. The number of bottles (paragraph 15-a) needed for a test will depend on the frequency of analysis and the test duration.
39. It is recommended that triplicate bottles be analysed after a sufficient number of time intervals such that the 10-d window may be identified. Also at least five test bottles (paragraph 15-a) from sets (a), (b) and (c) (see paragraph 42) are analysed at the end of the test, to enable 95 % confidence intervals to be calculated for the mean percentage biodegradation value.

**▼ M4***Inoculated medium*

40. The inoculum is used at a concentration of 4 mg/l activated sludge dry solids. Prepare immediately before use sufficient inoculated medium by adding, for example, 2 ml suitably treated activated sludge (paragraphs 27 to 32) at 2 000 mg/l to 1 litre of mineral salts medium (paragraph 19). When secondary sewage effluent is to be used add up to 100 ml effluent (paragraph 33) to 900 ml mineral salts medium (paragraph 19) and dilute to 1 litre with medium.

*Preparation of bottles*

41. Aliquots of inoculated medium are dispensed into replicate bottles to give a headspace to liquid ratio of 1:2 (e.g. add 107 ml to 160 ml-capacity bottles). Other ratios may be used, but see the warning given in paragraph 11. When using either type of inoculum, care must be taken to ensure that the inoculated medium is adequately mixed to ensure that it is uniformly distributed to the test bottles.
42. Sets of bottles (paragraph 15a) are prepared to contain the following:
- (a) Test vessels (denoted  $F_T$ ) containing the test chemical;
  - (b) Blank controls (denoted  $F_B$ ) containing only the test medium plus inoculum; any chemicals, solvents, agents or glass fibre filters used to introduce the test chemical into the test vessels must also be added;
  - (c) Vessels (denoted  $F_C$ ) for checking the procedure containing the reference chemical;
  - (d) If needed, vessels (denoted  $F_I$ ) for checking a possible inhibitory effect of the test chemical containing both the test chemical and reference chemical at the same concentrations (paragraph 24) as in bottles  $F_T$  and  $F_C$ , respectively;
  - (e) Vessels (denoted  $F_S$ ) for checking a possible abiotic degradation as (a) plus 50 mg/l  $HgCl_2$  or sterilised by some other means (e.g. by autoclaving).
43. Water-soluble test chemicals and reference chemicals are added as aqueous stock solutions (paragraphs 22, 23 and 24) to give a concentration of 10 to 20 mg C/l.
44. Insoluble test chemicals and insoluble reference chemicals are added to bottles in a variety of ways (see paragraph 22a-e) according to the nature of the test chemical, either before or after addition of the inoculated medium, depending on the method of treatment of the test chemical. If one of the procedures given in paragraph 22a-e is used, then the blank bottles  $F_B$  (paragraph 42b) should be treated in a similar fashion but excluding the test chemical or reference chemical.
45. Volatile test chemicals should be injected into sealed bottles (paragraph 47) using a micro syringe. The dose is calculated from the volume injected and the density of the test chemical.
46. Water should be added to vessels, where necessary, to give the same liquid volume in each vessel. It must be ensured that the headspace to liquid ratio (usually 1:2) and concentration of the test chemical are such that sufficient oxygen is available in the headspace to allow for complete biodegradation.

**▼ M4**

47. All bottles are then sealed for example, with butyl rubber septa and aluminium caps. Volatile tests chemicals should be added at this stage (paragraph 45). If the decrease in DOC concentration of the test solution is to be monitored and for time zero analyses to be performed for initial IC concentration (sterile controls, paragraph 42e) or other determinands, remove an appropriate sample from the test vessel. The test vessel and its contents are then discarded.
  
48. The sealed bottles are placed on a rotary shaker (paragraph 15d), with a shaking rate sufficient to keep the bottle contents well mixed and in suspension (e.g. 150 to 200 rpm), and incubated in the dark at 20 °C, to be kept within  $\pm 1$  °C.

*Sampling*

49. The pattern of sampling will depend on the lag period and kinetic rate of biodegradation of the test chemical. Bottles are sacrificed for analysis on the day of sampling, which should be at least weekly or more frequently (e.g. twice per week) if a complete degradation curve is required. The requisite number of replicate bottles is taken from the shaker, representing  $F_T$ ,  $F_B$  and  $F_C$  and, if used  $F_I$  and  $F_S$  (see paragraph 42). The test normally runs for 28d. If the biodegradation curve indicates that a plateau has been attained before 28d, the test may be concluded earlier than 28d. Take samples from the five bottles reserved for the 28th day of the test for analysis and use the results to calculate the confidence limits or coefficient of variation of percentage biodegradation. Bottles representing the checks for inhibition and for abiotic degradation need not be sampled as frequently as the other bottles; day 1 and day 28 would be sufficient.

*Inorganic carbon (IC) analysis*

50. CO<sub>2</sub> production in the bottles is determined by measuring the increase in the concentration of inorganic carbon (IC) during incubation. There are two recommended methods available for measuring the amount of IC produced in the test, and these are described immediately below. Since the methods can give slightly different results only one should be used in a test run.
  
51. Method (a) is recommended if the medium is likely to contain remnants of, for example, a glass-filter paper and/or insoluble test chemical. This analysis can be performed using a gas chromatograph if a carbon analyser is not available. It is important that the bottles should be at or close to the test temperature when the headspace gas is analysed. Method (b) can be easier for laboratories using carbon analysers to measure IC. It is important that the sodium hydroxide solution (paragraph 21) used to convert CO<sub>2</sub> to carbonate is either freshly prepared or its IC content is known, so that this can be taken into account when calculating the test results (see paragraph 66-b.)

*Method (a): acidification to pH < 3*

52. Before each batch of analyses, the IC analyser is calibrated using an appropriate IC standard (e.g. 1 % w/w CO<sub>2</sub> in N<sub>2</sub>). Concentrated orthophosphoric acid (paragraph 20) is injected through the septum of each bottle sampled to lower the pH of the medium to < 3 (e.g. add 1 ml to 107 ml test medium). The bottles are placed back on the shaker. After shaking for one hour at the test temperature the bottles are removed from the shaker, aliquots (e.g. 1 ml) of gas are withdrawn from the headspace of each bottle and injected into the IC analyser. The measured IC concentrations are recorded as mg C/l.



**▼M4**

53. The principle of this method is that after acidification to  $\text{pH} < 3$  and equilibration at  $20\text{ }^\circ\text{C}$ , the equilibrium constant for the distribution of  $\text{CO}_2$  between the liquid and gaseous phases in the test bottles is 1,0 when measured as a concentration (13). This should be demonstrated for the test system at least once as follows:

Set up bottles containing 5 and 10 mg/l as IC using a solution of anhydrous sodium carbonate ( $\text{Na}_2\text{CO}_3$ ) in  $\text{CO}_2$ -free water prepared by acidifying water to  $\text{pH} 6,5$  with concentrated ortho-phosphoric acid (paragraph 20), sparging overnight with  $\text{CO}_2$ -free air and raising the  $\text{pH}$  to neutrality with alkali. Ensure that the ratio of the headspace volume to the liquid volume is the same as in the tests (e.g. 1:2). Acidify and equilibrate as described in paragraph 52, and measure the IC concentrations of both the headspace and liquid phases. Check that the two concentrations are the same within experimental error. If they are not, the operator should review the procedures. This check on the distribution of IC between liquid and gaseous phases need not be made every time the test is performed; it could presumably be made while performing the calibration.

54. If DOC removal is to be measured (water-soluble test chemicals only), samples should be taken of the liquid phase from separate (non-acidified) bottles, membrane-filtered and injected into the DOC analyser. These bottles can be used for other analyses as necessary, to measure primary biodegradation.

*Method (b): conversion of  $\text{CO}_2$  to carbonate*

55. Before each batch of analyses, the IC analyser is calibrated using an appropriate standard — for example, a solution of sodium bicarbonate ( $\text{NaHCO}_3$ ) in  $\text{CO}_2$  free water (see paragraph 53) in the range 0 to 20 mg/l as IC. Sodium hydroxide solution (7M, paragraph 21) (e.g. 1 ml to 107 ml medium) is injected through the septum of each bottle sampled and the bottles are shaken for 1 h at the test temperature. Use the same NaOH solution on all bottles sacrificed on a particular day, but not necessarily on all sampling occasions throughout a test. If absolute blank IC values are required at all sampling occasions, IC determinations of the NaOH solution will be required each time it is used. The bottles are removed from the shaker and allowed to settle. Suitable volumes (e.g. 50 to 1 000  $\mu\text{l}$ ) of the liquid phase in each vessel are withdrawn by syringe. The samples are injected into the IC analyser and the concentrations of IC are recorded. It should be ensured that the analyser used is equipped properly to deal with the alkaline samples produced in this method.
56. The principle of this method is that after the addition of alkali and shaking, the concentration of IC in the headspace is negligible. This should be checked for the test system at least once by using IC standards, adding alkali and equilibrating, and measuring the concentration of IC in both the headspace and liquid phases (see paragraph 53). The concentration in the headspace should approach zero. This check on the virtually complete absorption of  $\text{CO}_2$  need not be made every time the test is performed.
57. If DOC removal is to be measured (water-soluble test chemicals only), samples should be taken of the liquid phase from separate bottles (containing no added alkali), membrane filtered and injected into the DOC analyser. These bottles can be used for other analyses, as necessary, to measure primary biodegradability.

**▼ M4****DATA AND REPORTING****Calculating of results**

58. Assuming 100 % mineralisation of the test chemical to CO<sub>2</sub>, the ThIC in excess of that produced in the blank controls equals the TOC added to each test bottle at the start of the test, that is:

$$\text{ThIC} = \text{TOC}$$

The total mass (mg) of inorganic carbon (TIC) in each bottle is:

$$\begin{aligned} \text{TIC} &= (\text{mg C in the liquid} + \text{mg C in the headspace}) && \text{Equation [1]} \\ &= (V_L \times C_L) + (V_H \times C_H) \end{aligned}$$

where:

V<sub>L</sub> = volume of liquid in the bottle (litre);

C<sub>L</sub> = concentration of IC in the liquid (mg/l as carbon);

V<sub>H</sub> = volume of the headspace (litre);

C<sub>H</sub> = concentration of IC in the headspace (mg/l as carbon).

The calculations of TIC for the two analytical methods used for measuring IC in this test are described below in paragraphs 60 and 61. Percentage biodegradation (% D) in each case is given by:

$$\% D = \frac{(\text{TIC}_t - \text{TIC}_b)}{\text{TOC}} \times 100 \quad \text{Equation [2]}$$

where:

TIC<sub>t</sub> = mg TIC in test bottle at time t;

TIC<sub>b</sub> = mean mg TIC in blank bottles at time t;

TOC = mg TOC added initially to the test vessel.

The percentage biodegradation % D is calculated for the test (F<sub>T</sub>), reference (F<sub>C</sub>) and, if included inhibition monitoring control (F<sub>I</sub>) bottles from the respective amounts of TIC produced up to each sampling time.

59. If there has been a significant increase in the TIC content of the sterile controls (F<sub>S</sub>) over the test period, then it may be concluded that abiotic degradation of the test chemical has occurred and this must be taken into account in the calculation of D in Equation [2].

**Acidification to pH < 3**

60. Since acidification to pH < 3 and equilibration results in the equalisation of the concentration of TIC in the liquid and gaseous phases, only the concentration of IC in the gas phase needs to be measured. Thus, from Equation [1] TIC = (V<sub>L</sub> + V<sub>H</sub>) × C<sub>H</sub> = V<sub>B</sub> × C<sub>H</sub>, where V<sub>B</sub> = volume of the serum bottle.

**Conversion of CO<sub>2</sub> to carbonate**

61. In this method calculations are performed as in Equation [1], but the negligible amount of IC in the gaseous phase is ignored, that is V<sub>H</sub> × C<sub>H</sub> = 0, and TIC = V<sub>L</sub> × C<sub>L</sub>.

**▼ M4****Expression of Results**

62. A biodegradation curve is obtained by plotting percentage biodegradation,  $D$ , against time of incubation and if possible, the lag phase, biodegradation phase, 10-d window and plateau phase, that is the phase in which the maximal degradation has been reached and the biodegradation curve has levelled out, are indicated. If comparable results are obtained for parallel test vessels  $F_T$  (< 20 % difference), a mean curve is plotted (see Appendix 2, Fig.1); if not, curves are plotted for each vessel. The mean value of the percentage biodegradation in the plateau phase is determined or the highest value is assessed (e.g. when the curve decreases in the plateau phase), but it is important to assess that in the latter case the value is not an outlier. Indicate this maximum level of biodegradation as 'degree of biodegradation of the test chemical' in the test report. If the number of test vessels was insufficient to indicate a plateau phase, the measured data of the last day of the test are used to calculate a mean value. This last value, the mean of five replicates, serves to indicate the precision with which the percentage biodegradation was determined. Also report the value obtained at the end of the 10-d window.
63. In the same way, a curve for the reference chemical,  $F_C$ , is plotted and, if included, for the abiotic elimination check,  $F_S$  and the inhibition control,  $F_I$ .
64. The amounts of TIC present in the blank controls ( $F_B$ ) are recorded as are those in flasks  $F_S$  (abiotic check), if these vessels were included in the test.
65. Calculate  $D$  for the  $F_I$  vessels, based on the theoretical IC yield anticipated from only the reference component of the mixture. If, at day 28,  $[(D_{FC}^{(1)} - D_{FI}^{(2)})/D_{FC}] \times 100 > 25 \%$ , it may be assumed that the test chemical inhibited the activity of the inoculum, and this may account for low values of  $D_{FT}$  obtained under the conditions of the test. In this case the test could be repeated using a lower test concentration and preferably reducing the DIC in the inoculum and TIC formed in the blank controls, since the lower concentration will otherwise reduce the precision of the method. Alternatively, another inoculum may be used. If in flask  $F_S$  (abiotic) a significant increase (> 10 %) in the amount of TIC is observed, abiotic degradation processes may have occurred.

**Validity of results**

66. A test is considered valid if:
- (a) the mean percentage degradation in vessels  $F_C$  containing the reference chemical is > 60 % by the 14th day of incubation; and
  - (b) the mean amount of TIC present in the blank controls  $F_B$  at the end of the test is > 3 mg C/l.

If these limits are not met, the test should be repeated with an inoculum from another source and/or the procedures used should be reviewed. For example, if high blank IC production is a problem the procedure given in paragraphs 27 to 32 should be followed.

<sup>(1)</sup> The percentage degradation in Vessels  $F_C$  containing the reference substance.

<sup>(2)</sup> The percentage degradation in Vessels  $F_I$ .

**▼M4**

67. If the test chemical does not reach 60 % ThIC and was shown not to be inhibitory (paragraph 65), the test could be repeated with increased concentration of inoculum (up to 30 mg/l activated sludge and 100 ml effluent/l) or inocula from other sources, especially if degradation had been in the range 20 to 60 %.

**Interpretation of results**

68. Biodegradation > 60 % ThIC within the 10-d window in this test demonstrates that the test chemical is readily biodegradable under aerobic conditions.
69. If the pass value of 60 % ThIC is not attained, determine the pH value in media in bottles which have not been made acid or alkaline; a value of less than 6,5 could indicate that nitrification had occurred. In such a case repeat the test with a buffer solution of higher concentration.

**Test Report**

70. Compile a table of % D for each test ( $F_T$ ), reference ( $F_C$ ) and, if included, inhibition control bottle ( $F_I$ ) for each day sampled. If comparable results are obtained for replicate bottles, plot a curve of mean % D against time. Record the amount of TIC in the blanks ( $F_B$ ) and in the sterile controls ( $F_S$ ) DOC and/or other determinands, and their percentage removal.
71. Determine the mean value of % D in the plateau phase, or use the highest value if the biodegradation curve decreases in the plateau phase, and report this as the 'degree of biodegradation of the test chemical'. It is important to ensure that in the latter case the highest value is not an outlier.
72. The test report must include the following information:

*Test chemical:*

- common name, chemical name, CAS number, structural formula and relevant physical-chemical properties;
- purity (impurities) of test chemical.

*Test conditions:*

- reference to this Test Method;
- description of the test system used (e.g. volume of the vessel, head space to liquid ratio, method of stirring, etc.);
- application of test chemical and reference chemical to test system: test concentration used and amount of carbon dosed into each test bottle, any use of solvents;
- details of the inoculum used, any pre-treatment and pre-conditioning;
- incubation temperature;
- validation of the principle of IC analysis;
- main characteristics of the IC analyser employed (and any other analytical methods used);
- number of replicates.

*Results:*

- raw data and calculated values of biodegradability in tabular form;

**▼M4**

- the graph of percentage degradation against time for the test and reference chemicals, the lag phase, degradation phase, 10-d window and slope;
- percentage removal at plateau, at end of test, and after 10-d window;
- reasons for any rejection of the test results;
- any other facts that are relevant to the procedure followed;
- discussion of results.

*LITERATURE:*

- (1) Chapter C.4 of this Annex Determination of 'Ready' Biodegradability — CO<sub>2</sub> Evolution Test (Method C.4-C).
- (2) Sturm RN (1973). Biodegradability of Nonionic surfactants: screening test for predicting rate and ultimate biodegradation. *J.A.,Oil Chem Soc.* 50: 159-167.
- (3) Larson RJ (1979). Estimation of biodegradation potential of xenobiotic organic chemicals. *Appl Env. Microbiol.* 38: 1153-1161.
- (4) Larson RJ, Hansmann MA and Bookland EA (1996). Carbon dioxide recovery in ready biodegradability tests: mass transfer and kinetic constants, *Chemosphere* 33: 1195-1210.
- (5) ISO 9439 (1990; revised 1999). Water Quality — Evaluation of ultimate aerobic biodegradability of organic compounds in aqueous medium — Carbon dioxide evolution Test (Sturm).
- (6) US EPA (1996). Fate, Transport and Transformation Test Guideline. 835.3110 Carbon dioxide evolution test. Office, Prevention Pesticides and Toxic Substances Washington, DC.
- (7) US EPA (1996). Fate, Transport and Transformation Test Guideline. 835.3100. Aerobic aquatic biodegradation. Office, Prevention Pesticides and Toxic Substances Washington, DC.
- (8) Gledhill WE (1975). Screening test for assessment of biodegradability: Linear alkyl benzene sulfonate. *Appl Microbiol.* 30: 922-929.
- (9) Weytjens D, Van Ginneken I and Painter HA (1994). The recovery of carbon dioxide in the Sturm test for ready biodegradability. *Chemosphere* 28: 801-812.
- (10) Ennis DM and Kramer A (1975). A rapid microtechnique for testing biodegradability of nylons and polyamides. *J. Food Sci.* 40: 181-185.
- (11) Ennis DM, Kramer A, Jameson CW, Mazzoccki PH and Bailey PH (1978). *Appl. Env. Microbiol.* 35: 51-53.
- (12) Boatman RJ, Cunningham SL and Ziegler DA (1986). A method for measuring the biodegradation of organic chemicals, *Env. Toxicol. Chem.* 5: 233-243.
- (13) Struijs J and Stoltenkamp J (1990). Head space determination of evolved carbon dioxide in a biodegradability screening test. *Ecotox. Env. Safety* 19: 204-211.
- (14) Birch RR and Fletcher RJ (1991). The application of dissolved inorganic carbon measurements to the study of aerobic biodegradability. *Chemosphere* 23: 507-524.
- (15) Birch RR, Biver C, Campagna R, Gledhill WE, Pagga U, Steber J, Reust H, and Bontinck WJ (1989). Screening of chemicals for anaerobic biodegradation. *Chemosphere* 19: 1527-1550.

**▼ M4**

- (16) ISO 14593, (1999) Water Quality — Evaluation of ultimate aerobic biodegradability of organic compounds in an aerobic medium-method by analysis of inorganic carbon in sealed vessels (CO<sub>2</sub> headspace test).
- (17) Battersby NS (1997). The ISO headspace CO<sub>2</sub> biodegradation test, *Chemosphere* 34: 1813-1822.
- (18) US EPA (1996). Fate, Transport and Transportation. 835.3120. Sealed vessel carbon dioxide production test. Office, Prevention Pesticides and Toxic Substance, Washington, DC.
- (19) Battersby NS, Ciccognani D, Evans MR, King D, Painter HA, Peterson DR and Starkey M (1999). An 'inherent' biodegradability test for oil products: description and results of an international ring test. *Chemosphere* 38: 3219-3235.
- (20) Chapter C.4 of this Annex, Determination of 'Ready' Biodegradability.
- (21) OECD (1988). OECD Ring-test of methods for determining ready biodegradability: Chairman's report (M. Hashimoto; MITI) and final report (M. Kitano and M. Takatsuki; CITI). Paris.
- (22) Chapter C.11 of this Annex, Activated sludge respiration inhibition test.
- (23) Struijs J, Stoltenkamp-Wouterse MJ and Dekkers ALM (1995). A rationale for the appropriate amount of inoculum in ready biodegradability tests. *Biodegradation* 6: 319-327.
- (24) EU (1999). Ring-test of the ISO Headspace CO<sub>2</sub> method: application to surfactants: Surfactant Ring Test-1, Report EU4697, Water Research Centre, May 1999, Medmenham, SL7 2HD, UK.
- (25) ISO 10634 (1996) Water Quality — Guidance for the preparation and treatment of poorly water-soluble organic compounds for the subsequent evaluation of their biodegradability in an aqueous medium.

**▼ M4***Appendix 1*

## ABBREVIATIONS AND DEFINITIONS

**IC:** Inorganic carbon

**ThCO<sub>2</sub>:** Theoretical carbon dioxide (mg) is the quantity of carbon dioxide calculated to be produced from the known or measured carbon content of the test chemical when fully mineralised; also expressed as mg carbon dioxide evolved per mg test chemical.

**DOC:** Dissolved organic carbon is the organic carbon present in solution or that which passes through a 0,45 micrometre filter or remains in the supernatant after centrifuging at approx. 4 000 g (about 40 000 m sec<sup>-2</sup>) for 15 min.

**DIC:** Dissolved inorganic carbon

**ThIC:** Theoretical inorganic carbon

**TIC:** Total inorganic carbon

**Readily biodegradable:** An arbitrary classification of chemicals which have passed certain specified screening tests for ultimate biodegradability; these tests are so stringent that it is assumed that such chemicals will rapidly and completely biodegrade in aquatic environments under aerobic conditions.

**10-d window:** The 10 days immediately following the attainment of 10 % biodegradation.

**Inherent biodegradability:** A classification of chemicals for which there is unequivocal evidence of biodegradation (primary or ultimate) in any test of biodegradability.

**Ultimate aerobic biodegradation:** The level of degradation achieved when the test chemical is totally utilised by micro-organisms resulting in the production of carbon dioxide, water, mineral salts and new microbial cellular constituents (biomass).

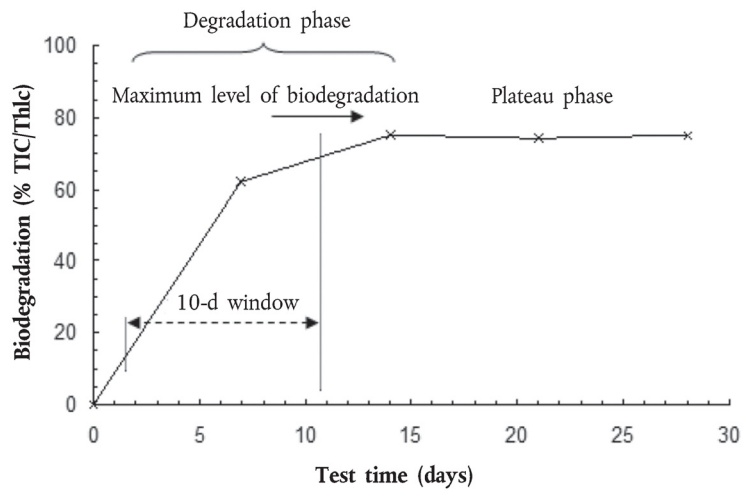
**Mineralisation:** Mineralisation is the complete degradation of an organic chemical to CO<sub>2</sub> and H<sub>2</sub>O under aerobic conditions, and CH<sub>4</sub>, CO<sub>2</sub> and H<sub>2</sub>O under anaerobic conditions.

**Lag phase:** The time from the start of a test until acclimatization and/or adaptation of the degrading microorganisms is achieved and the biodegradation degree of a test chemical or organic matter has increased to a detectable level (e.g. 10 % of the maximum theoretical biodegradation, or lower, dependent on the accuracy of the measuring technique).

**Degradation phase:** The time from the end of the lag period to the time when 90 % of the maximum level of degradation has been reached.

**Plateau phase:** Plateau phase is the phase in which the maximal degradation has been reached and the biodegradation curve has levelled out.

**Test chemical:** Any substance or mixture tested using this Test Method.

▼ **M4***Appendix 2***Example of a biodegradation curve***Figure 1***Biodegradation of 1-octanol in the CO<sub>2</sub> headspace test**

## Glossary

Biodegradation:

Degradation phase:

Maximum level of biodegradation:

Plateau phase:

10-d(ay) window:

Test time (days):



▼ **M4****C. 30. BIOACCUMULATION IN TERRESTRIAL OLIGOCHAETES**

## INTRODUCTION

1. This Test Method is equivalent to OECD Test Guideline (TG) 317 (2010). Among the Test Methods relating to environmental fate, the Bioconcentration: Flow-through Fish Test (chapter C.13 of this Annex (49)) and the Bioaccumulation in Sediment-dwelling Benthic Oligochaetes (53) were published in 1996 and 2008 respectively. The extrapolation of aquatic bioaccumulation data to terrestrial organisms like earthworms is difficult, if possible at all. Model calculations based on a test chemical's lipophilicity, e.g. (14) (37), are currently used for the assessment of bioaccumulation of chemicals in soil, as e.g. in the EU Technical Guidance Document (19). The need for a compartment-specific test method has already been addressed, e.g. (55). Such a method is especially important for the evaluation of secondary poisoning in terrestrial food chains (4). Several national test methods address the issue of bioaccumulation in organisms other than fish e.g. (2) and (72). A method on the measurement of bioaccumulation from contaminated soils in earthworms (*Eisenia fetida*, Savigny) and potworms has been developed by the American Society for Testing and Materials (3). An internationally accepted method for the determination of bioaccumulation in spiked soil will improve the risk assessment of chemicals in terrestrial ecosystems e.g. (25) (29).
  
2. Soil-ingesting invertebrates are exposed to soil bound chemicals. Among these animals, terrestrial oligochaetes play an important role in the structure and function of soils (15) (20). Terrestrial oligochaetes live in soil and partly at the soil surface (especially the litter layer); they frequently represent the most abundant species in terms of biomass (54). By bioturbation of the soil and by serving as prey these animals can have a strong influence on the bioavailability of chemicals to other organisms like invertebrates (e.g. predatory mites and beetles; e.g. (64)) or vertebrate (e.g. foxes and gulls) predators (18) (62). Some species of terrestrial oligochaetes currently used in ecotoxicological testing are described in Appendix 5.
  
3. The ASTM Standard Guide for Conducting Laboratory Soil Toxicity or Bioaccumulation Tests with the Lumbricid Earthworm *Eisenia fetida* and the Enchytraeid Potworm *Enchytraeus albidus* (3) provides many essential and useful details for the performance of the present soil bioaccumulation Test Method. Further documents that are referred to in this Test Method are chapter C.13 of this Annex, Bioconcentration: Flow-through Fish Test (49) and OECD TG 315: Bioaccumulation in Sediment-dwelling Benthic Oligochaetes (53). Practical experience with soil bioaccumulation studies and publications from LITERATURE e.g. (1) (5) (11) (12) (28) (40) (43) (45) (57) (59) (76) (78) (79) are also major sources of information for this Test Method.
  
4. This Test Method is mostly applicable to stable, neutral organic chemicals, which tend to adsorb to soils. Testing for bioaccumulation of soil-associating, stable metallo-organic compounds may be possible with this Test Method. It is also applicable to metals and other trace elements.

## PRE-REQUISITE

5. Tests for measuring the bioaccumulation of a chemical in terrestrial oligochaetes have been performed with heavy metals (see e.g. (63)) and persistent, organic chemicals having log K<sub>ow</sub> values between 3,0 and 6,0, e.g. (40). Such tests also apply to:

— Chemicals that show a log K<sub>ow</sub> of more than 6,0 (super-hydrophobic chemicals);

**▼M4**

- Chemicals which belong to a class of organic chemicals known to have the potential to bioaccumulate in living organisms, e.g. surface active or highly adsorptive chemicals;
  - Chemicals that indicate the potential for bioaccumulation from structural features, e.g. analogues of chemicals with known bioaccumulation potential; and
  - Metals.
6. Information on the test chemical such as common name, chemical name (preferably IUPAC name), structural formula, CAS registry number, purity, safety precautions, proper storage conditions and analytical methods should be obtained before beginning the study. In addition, the following information should be known:
- (a) solubility in water;
  - (b) octanol-water partition coefficient,  $K_{ow}$ ;
  - (c) soil-water partition coefficient, expressed as  $K_{oc}$ ;
  - (d) vapour pressure;
  - (e) degradability (e.g. in soil, water);
  - (f) known metabolites.
7. Radiolabelled or non-radiolabelled test chemicals can be used. However, to facilitate analysis it is recommended to use a radiolabelled test chemical. The decision will be made based on the detection limits or a requirement to measure parent test chemical and metabolites. If a radiolabelled test chemical is used and total radioactive residues are measured, it is important that the radiolabelled residues in both the soil and the test organisms are characterised for percentages of parent test chemical and labelled non-parent, e.g. in samples taken at steady state or at the end of the uptake phase, to allow a bioaccumulation factor (BAF) calculation for the parent test chemical and for the soil metabolites of concern (see paragraph 50). The method described here may have to be modified, e.g. to provide sufficient biomass, for measuring non-radiolabelled organic test chemical or metals. When total radioactive residues are measured (by liquid scintillation counting following extraction, combustion or tissue solubilisation), the bioaccumulation factor is based on the parent test chemical and metabolites. The BAF calculation should preferably be based on the concentration of the parent test chemical in the organisms and total radioactive residues. Subsequently, the biota-soil accumulation factor (BSAF), normalized to the lipid content of worm and organic carbon content (OC) of soil should be calculated from the BAF for reasons of comparability between results from different bioaccumulation tests.
8. Toxicity of the test chemical to the species used in the test should be known, e.g. an effect concentration (EC<sub>x</sub>) or lethal concentration (LC<sub>x</sub>) for the time of the uptake phase (e.g. (19)). The selected concentration of the test chemical should preferably be about 1 % of its acute asymptotic LC<sub>50</sub>, and at least 10-fold higher than its detection limit in soil by the analytical method used. If available, preference should be given to toxicity values derived from long-term studies on sublethal endpoints (51) (52). If such data are not available, an acute toxicity test will provide useful information (see e.g. (23)).

**▼M4**

9. An appropriate analytical method of known accuracy, precision, and sensitivity for the quantification of the chemical in the test solutions, in the soil, and in the biological material should be available, together with details of sample preparation and storage as well as material safety data sheets. Analytical detection limits of the test item in soil and worm tissue should also be known. If a <sup>14</sup>C-labelled test chemical is used, the specific radioactivity (i.e. Bq mol<sup>-1</sup>) and the percentage of radioactivity associated with impurities should be known. The specific radioactivity of the test chemical should be high enough to facilitate analysis, and the test concentrations used should not elicit toxic effects.
  
10. The test can be performed with an artificial soil or with natural soils. Information on characteristics of the natural soil used, e.g. origin of soil or its constituents, pH, organic carbon content, particle size distribution (percent sand, silt, and clay), and water holding capacity (WHC), should be known before the start of the test (3) (48).

## PRINCIPLE OF THE TEST

11. The parameters which characterise the bioaccumulation of a test chemical include the bioaccumulation factor (BAF), the uptake rate constant ( $k_s$ ) and the elimination rate constant ( $k_e$ ). Definitions are provided in Appendix 1.
  
12. The test consists of two phases: the uptake (exposure) phase and the elimination (post-exposure) phase. During the uptake phase, replicated groups of worms are exposed to soil, which has been spiked with the test chemical. In addition to the test animals, groups of control worms are held under identical conditions without the test chemical. The dry weight and lipid content of the test organisms are measured. This can be done using worms of the control group. Analytical background values (blank) can be obtained by analysing samples of the control worms and soil. For the elimination phase, the worms are transferred to a soil free of the test chemical. An elimination phase is always required unless uptake of the test chemical during the exposure phase has been insignificant. An elimination phase provides information on the rate at which the test chemical is excreted by the test organisms (e.g. (27)). If a steady state has not been reached during the uptake phase, the determination of the kinetic parameters – kinetic bioaccumulation factor BAF<sub>k</sub>, uptake and elimination rate constant(s) – should preferably be based on simultaneous fitting of the results of the uptake and elimination phases. The concentration of the test chemical in/on the worms is monitored throughout both phases of the test.
  
13. During the uptake phase, measurements are made at sampling times up to 14 days (enchytraeids) or 21 days (earthworms) until the steady state is reached (11) (12) (67). The steady state occurs when a plot of the concentration in worms against time is parallel to the time axis, and three successive concentration analyses made on samples taken at intervals of at least two days do not vary more than ± 20 % of each other based on statistical comparisons (e.g. analysis of variance, regression analysis).
  
14. The elimination phase consists of transferring the test organisms to vessels containing the same substrate without the test chemical. During the elimination phase, measurements are made at sampling times during 14 days (enchytraeids) or 21 days (earthworms) unless earlier analytical determination showed 90 % reduction of the test chemical residues in worms. The concentration of the test chemical in the worms at the end of the elimination phase is reported as non-eliminated residues. The steady state

**▼ M4**

bioaccumulation factor (BAF<sub>ss</sub>) is calculated preferably both as the ratio of the concentration in worms (C<sub>a</sub>) and in the soil (C<sub>s</sub>) at apparent steady state, and as a kinetic bioaccumulation factor, BAF<sub>K</sub>, as the ratio of the rate constant of uptake from soil (k<sub>s</sub>) and the elimination rate constant (k<sub>e</sub>) (see Appendix 1 for definitions) assuming first-order kinetics (see Appendix 2 for calculations). If first-order kinetics is obviously not applicable, other models should be employed.

15. The uptake rate constant, the elimination rate constant (or constants, where other models are involved), the kinetic bioaccumulation factor (BAF<sub>K</sub>), and where possible, the confidence limits of each of these parameters are calculated from computerised model equations (see Appendix 2 for guidance). The goodness of fit of any model can be determined from e.g. the correlation coefficient or the coefficient of determination (coefficients close to one indicate a good fit) or chi-squared. Also the size of the standard error or confidence limit around the estimated parameters may be indicative of the goodness of fit of the model.
  
16. To reduce variability in test results for test chemicals with high lipophilicity, bioaccumulation factors should be expressed in relation to lipid content and organic carbon content (kg soil organic carbon (OC) kg<sup>-1</sup> worm lipid content). This approach is based on the fact that for some chemical classes, there is a clear relationship between the potential for bioaccumulation and lipophilicity; this has been well established for fish (47). There is a relationship between the lipid content of fish and the bioaccumulation of such chemicals. For benthic organisms, similar correlations have been found e.g. (30) (44). Likewise for terrestrial oligochaetes this correlation has been demonstrated e.g. (5) (6) (7) (14). If sufficient worm tissue is available, the lipid content of the test animals can be determined on the same biological material as the one used to determine the concentration of the test chemical. Alternatively, control animals can be used to measure the lipid content.

#### VALIDITY OF THE TEST

17. For a test to be valid the following criteria should be fulfilled for both controls and treatments:
  - At the end of the test, the overall mortality during uptake and elimination phase should not exceed 10 % (earthworms) or 20 % (enchytraeids) of the total number of the introduced worms.
  
  - For *Eisenia fetida* and *Eisenia andrei*, the mean mass loss as measured at the end of the uptake and at the end of the elimination phase should not exceed 20 % compared to the initial fresh weight (f.w.) at start of each phase.

#### DESCRIPTION OF THE METHOD

##### Test species

18. Several species of terrestrial oligochaetes are recommended for bioaccumulation testing. The most commonly used species *Eisenia fetida* or *Eisenia andrei* (Lumbricidae), or *Enchytraeus albidus*, *Enchytraeus crypticus*, or *Enchytraeus luxuriosus* (Enchytraeidae) are described in Appendix 5.

**▼ M4****Apparatus**

19. Care should be taken to avoid the use of materials, for all parts of the equipment, which can dissolve, adsorb the test chemical or leach other chemicals, and have an adverse effect on the test animals. Standard rectangular or cylindrical vessels, made of chemically inert material and of suitable capacity can be used in compliance with the loading rate, i.e. the number of test worms. Stainless steel, plastic or glass may be used for any equipment having contact with the test media. The test vessels should be appropriately covered to prevent escaping of the worms, while allowing sufficient air supply. For chemicals with high adsorption coefficients, such as synthetic pyrethroids, silanised glass may be required. In these situations the equipment will have to be discarded after use (49). Radiolabelled test items and volatile chemicals should be prevented from escaping. Traps (e.g. glass gas washing bottles) should be employed containing suitable absorbents to retain any residues evaporating from the test vessels.

**Soil**

20. The test soil should be of a quality that will allow the survival and preferably the reproduction of the test organisms for the duration of the acclimation and test periods without them showing any abnormal appearance or behaviour. The worms should burrow in the soil.
  
21. The artificial soil described in the chapter C.8 of this Annex (48) is recommended for use as the substrate in the tests. Preparation of the artificial soil for use in the bioaccumulation tests and recommendations for the storage of artificial soil are given in Appendix 4. Air-dried artificial soil may be stored at room temperature until use.
  
22. However, natural soils from unpolluted sites may serve as test and/or culture soil. Natural soils should be characterised at least by origin (collection site), pH, organic carbon content, particle size distribution (percent sand, silt, and clay), maximum water holding capacity (WHC<sub>max</sub>), and percent water content (3). Analysis of the soil or its constituents for micro-pollutants prior to use should provide useful information. If field soil from agricultural land is used, it should not have been treated with crop protection products or with manure from treated animals as fertilizers for at least one year and with organic fertilizers for at least six months prior to sampling (50). Manipulation procedures for natural soils prior to use in ecotoxicological tests with oligochaetes in the laboratory are described in (3). For natural soils the storage time in the laboratory should be kept as short as possible.

**Application of the test chemical**

23. The test chemical is incorporated into the soil. The physicochemical properties of the test chemical should be taken into consideration. A water-soluble test chemical should be completely dissolved in water before it be mixed with the soil. The recommended spiking procedure for poorly water-soluble test chemical involves coating of one or more of the (artificial) soil constituents with the test chemical. For example, the quartz sand, or a portion thereof, can be soaked with a solution of the test chemical in a suitable organic solvent, which is then slowly evaporated to dryness. The coated fraction can then be mixed into the wet soil. The major advantage of this procedure is that no solvent is introduced into the soil. When a natural soil is used, the test chemical may be added by spiking an

**▼M4**

air-dried portion of the soil as described above for the artificial soil, or by stirring the test chemical into the wet soil, with subsequent evaporating step if a solubilising agent is used. In general, the contact of wet soil with solvents should be avoided as far as possible. The following should be considered (3):

— If a solvent other than water is used, it should be one that is water-miscible and/or can be driven off (for example, evaporated), leaving only the test chemical on the soil.

— If a solvent control is used, there is no need for negative control. The solvent control should contain the highest concentration of solvent added to the soil and should use solvent from the same batch used to make the stock solution. Toxicity and volatility of the solvent, and solubility of the test chemical in the chosen solvent should be the main criteria used for the selection of a suitable solubilising agent.

24. For chemicals that are poorly soluble in water and in organic solvents, 2,0-2,5 g of finely ground quartz sand per test vessel can be mixed with the quantity of test chemical, e.g. using mortar and pestle, to obtain the desired test concentration. This mixture of quartz sand and test chemical is added to the pre-moistened soil and thoroughly mixed with an appropriate amount of de-ionised water to obtain the required moisture content. The final mixture is distributed to the test vessels. The procedure is repeated for each test concentration, and an appropriate control with 2,0-2,5 g of finely ground quartz sand per test vessel is also prepared.

25. The concentration of the test chemical in the soil should be determined after spiking. The homogenous distribution of the test chemical into the soil should be verified before introducing the test organisms. The method used for spiking, and the reasons for choosing a specific spiking procedure should be reported (24).

26. Equilibrium between the soil and the pore-water phase should ideally be established before adding the organisms; a time period of four days at 20 °C is recommended. For many poorly water-soluble organic chemicals the time required to reach a true equilibrium between adsorbed and dissolved fractions can be counted in days or months. Depending on the purpose of the study, for example when the environmental conditions are to be mimicked, the spiked soil may be 'aged' for a longer period, e.g. for metals three weeks at 20 °C (22).

**Culturing of the test organisms**

27. Worms should be preferably kept in permanent laboratory culture. Guidance on laboratory culture methods for *Eisenia fetida* and *Eisenia andrei*, and Enchytraeid species, is provided in Appendix 5 (see also (48) (51) (52)).

28. The worms used in the tests should be free from observable diseases, abnormalities and parasites.

**PERFORMANCE OF THE TEST**

29. The test organisms are exposed to the test chemical during the uptake phase. The uptake phase should be of 14 days (enchytraeids) or 21 days (earthworms) unless it is demonstrated that steady state has been reached.

**▼M4**

30. For the elimination phase, the worms are transferred to a soil free of test chemical. The first sample should be taken at 4-24 h after the start of elimination phase. Examples of sampling schedules for a 21-day uptake phase and a 21-day elimination phase are given in Appendix 3.

**Test organisms**

31. For many species of terrestrial enchytraeids the individual weight is very low (e.g. 5-10 mg wet weight per individual for *Enchytraeus albidus* and less for *Enchytraeus crypticus* or *Enchytraeus luxuriosus*); in order to perform the weight measurements and chemical analysis, it may be necessary to pool the worms of the replicate test vessels (i.e. all the worms of a replicate vessel will be used for obtaining one analytical tissue result). 20 individual enchytraeids are added to each replicate, and at least three replicates should be used. If the analytical detection limit of the test chemical is high, more worms may be necessary. For test species with higher individual weight (*Eisenia fetida* and *Eisenia andrei*), replicate vessels containing one individual can be used.
32. The earthworms used in a test should be of similar weight (e.g. *Eisenia fetida* and *Eisenia andrei* should have an individual weight of 250-600 mg). Enchytraeids (e.g. *Enchytraeus albidus*) should have a length of approximately 1 cm. All worms used in a particular test should come from the same source, and should be adult animals with clitellum (see Appendix 5). Since the weight and age of an animal might have an effect on the BAF-values (e.g. due to varying lipid content and/or presence of eggs), these parameters should be recorded accurately and taken into account in the interpretation of results. In addition, cocoons can be deposited during the exposure period, which will also have an impact on the BAF values. It is recommended that a sub-sample of the test worms be weighed before the test in order to estimate the mean wet and dry weights.
33. A high soil-to-worm ratio should be used in order to minimise the decrease of the test chemical concentration in the soil during the uptake phase. For *Eisenia fetida* and *Eisenia andrei* a minimum amount of 50 g dry weight (d.w.) of soil per worm, and for enchytraeids, a minimum of 10-20 g d.w. of soil per test vessel are recommended. The vessels should contain a soil layer of 2-3 cm (enchytraeids) or 4-5 cm (earthworms).
34. The worms used in a test are removed from the culture (e.g. enchytraeids by using jeweller's tweezers). Adult animals are transferred to non-treated test soil for acclimation, and fed (see paragraph 36). If the test conditions differ from the culture conditions, an acclimation phase of 24-72 h should be sufficient to adapt the worms to the test conditions. After acclimation, earthworms are rinsed by transfer to glass dishes (e.g. petri dishes) containing clean water, and subsequently weighed before they are added to the test soil. Prior to weighing, excess water should be removed from the worms by gently touching them against the edge of the dish or by blotting them cautiously dry by using a slightly moistened paper towel.
35. Burrowing behaviour of the test organisms should be observed and recorded. In tests with earthworms, the animals (control and treatments) normally burrow in the soil within a period of a few hours; this should be checked no later than 24 h after addition of the worms to the test vessels. If the earthworms fail to burrow in the soil (e.g. more than 10 % over more than half of the uptake phase), this indicates that either the test conditions are not appropriate or the test organisms are not healthy. In such

**▼ M4**

in a case the test should be stopped and repeated. Enchytraeids mainly live in the interstitial pores of the soil, and frequently their integument may be only partly in contact with the surrounding substrate; exposure of burrowing and non-burrowing enchytraeids is assumed to be equivalent and non-burrowing of the enchytraeids does not necessarily require the repetition of the test.

**Feeding**

36. Feeding should be envisaged when a soil with low total organic carbon content is used. When an artificial soil is used, a weekly feeding rate (i.e. the worms should be fed once a week) of 7 mg of dried dung per g soil dry weight is recommended for earthworms, and a weekly rate of 2-2,5 mg of ground oat flakes per g soil dry weight is recommended for enchytraeids (11). The first food ration should be mixed with the soil immediately before the test organisms are added. Preferably the same type of food like in the cultures should be used (see Appendix 5).

**Light and temperature**

37. The tests should be carried out under a controlled 16/8 hours light/dark cycle, preferably 400 to 800 lx in the area of the test vessels (3). The test temperature should be  $20 \pm 2$  °C throughout the test.

**Test concentrations**

38. A single concentration is used. Situations where additional concentration(s) is(are) required should be justified. If toxicity (EC<sub>x</sub>) of the test chemical is close to the analytical detection limit, the use of radiolabelled test chemical with high specific radioactivity is recommended. For metals, the concentration should be above the background level in tissue and soil.

**Replicates**

39. For the kinetic measurements (uptake and elimination phase), the minimum number of treated replicate vessels should be three per sampling point. The total number of replicates prepared should be sufficient to cover all sampling times during the uptake and the elimination phase.
40. For the biological observations and measurements (e.g. dry-to-wet weight ratio, lipid content) and for the analysis of background concentrations in worms and soil, at least 12 replicate vessels of a negative control (four sampled at start, four at end of uptake, and four at end of elimination) should be provided if no solvent other than water is used. If any solubilising agent is used for application of the test chemical, a solvent control (four replicate vessels should be sampled at start, four at the end of the uptake phase, and four at the end of the elimination phase) containing all constituents except for test item should be run in addition to the treated replicates. In this case, four additional replicate vessels of a negative control (no solvent) may also be provided for optional sampling at the end of the uptake phase. These replicates can be compared biologically with the solvent control in order to gain information on a possible influence of the solvent on the test organisms. It is recommended establishing a sufficient number of additional reserve replicate vessels (e.g. eight) for treatment and control(s).



**▼M4****Frequency of soil quality measurements**

41. Soil pH, soil moisture content and the temperature (continuously) in the test room should be measured at the start and end of the uptake and elimination phases. Once per week the soil moisture content should be controlled by weighing the test vessels and comparing actual weights with initial weights at test start. Water losses should be compensated by adding deionised water.

**Sampling and analysis of worms and soil**

42. An example of schedule for the uptake and elimination phases in earthworm and enchytraeid bioaccumulation tests is given in Appendix 3.
43. The soil is sampled from the test vessels for the determination of test chemical concentration before inserting the worms, and during the uptake and elimination phases. During the test the concentrations of test chemical are determined in the worms and the soil. In general, total soil concentrations are measured. As an option, concentrations in pore water may be measured; in such case, rationale and appropriate methods should be provided prior to initiation of a study, and included in the report.
44. The worms and soil are sampled at least at six occasions during the uptake and the elimination phases. If the stability of a test chemical is demonstrated, the number of soil analyses can be reduced. It is recommended analysing at least three replicates at the beginning and at the end of the uptake phase. If the concentration in soil measured at the end of the uptake phase deviates from the initial concentration by more than 30 %, the soil samples taken at other dates should also be analysed.
45. Remove the worms of a given replicate from the soil at each sampling time (e.g. after spreading the soil of the replicate on a shallow tray and picking the worms using soft jewellers' tweezers), rinse them quickly with water in a shallow glass or steel tray. Remove excess water (see paragraph 34). Transfer the worms carefully to a pre-weighed vessel, weigh them instantly, including gut content.
46. The earthworms (*Eisenia* sp.) should then be allowed to purge their gut overnight e.g. on a moist filter paper in a covered petri dish (see paragraph 34). After purging, the weight of the worms should be determined in order to assess a possible decrease in biomass during the test (see validity criteria in paragraph 17). Weighing and tissue analysis of Enchytraeids is carried out without purging, as this is technically difficult due to the small size of these worms. After final weight determination, the worms should be killed immediately, using the most appropriate method (e.g. using liquid nitrogen, or freezing at temperatures below – 18 °C).
47. During the elimination phase, the worms replace contaminated gut contents with clean soil. This means, measurements in un-purged worms (enchytraeids in this context) sampled immediately before the elimination phase include contaminated gut soil. For aquatic oligochaetes it is assumed that after the initial 4-24 h of the elimination phase, most of the contaminated gut content has been replaced by clean sediment e.g. (46). Similar findings have been reported for earthworms in studies on the accumulation of radiolabelled cadmium and zinc (78). In the non-purged enchytraeids, the concentration of this first sample of the elimination phase may be considered as the tissue concentration after gut purge. To account for dilution of the test item concentration by uncontaminated soil during the elimination phase, the weight of the gut content may be estimated from worm wet weight/worm ash weight or worm dry weight/worm ash weight ratios.

**▼ M4**

48. The soil and worm samples should be preferably analysed immediately after removal (i.e. within 1-2 days) in order to prevent degradation or other losses, and it is recommended calculating the approximate uptake and elimination rates as the test proceeds. If the analysis is delayed, the samples should be stored by an appropriate method, e.g. by deep-freezing ( $\leq -18$  °C).
  
49. It should be checked that the precision and reproducibility of the chemical analysis, as well as the recovery of the test chemical from soil and worm samples are satisfactory for the given method; the extraction efficiency, the limit of detection (LOD) and the limit of quantification (LOQ) should be reported. Likewise it should be checked that the test chemical is not detectable in the control vessels in concentrations higher than background. When the concentration of the test chemical in the test organism  $C_a$  is  $> 0$  in the control worms, this should be included in the calculation of the kinetic parameters (see Appendix 2). All samples should be handled throughout the test to minimise contamination and loss (e.g. resulting from adsorption of the test chemical on the sampling device).
  
50. When working with radiolabelled test chemicals, it is possible to analyse parent and metabolites. Quantification of parent test chemical and metabolites at steady state or at the end of the uptake phase provides important information. The samples should then be 'cleaned up' so that the parent test chemical can be quantified separately. If single metabolites exceed 10 % of total radioactivity in the analysed sample(s), the identification of these metabolites is recommended.
  
51. The overall recovery, and the recovery of test chemical in worms, soil, and if used, in traps containing absorbents to retain evaporated test chemical, should be recorded and reported.
  
52. Pooling of the individuals sampled from a given test vessel is acceptable for enchytraeid worms which are smaller than earthworms. If pooling involves the reduction of the number of replicates, this limits the statistical procedures which can be applied to the data. If a specific statistical procedure and power are required, then an adequate number of replicate test vessels should be included in the test to accommodate the desired pooling, procedure and power.
  
53. It is recommended that the BAF be expressed both as a function of total dry weight and, when required (i.e. for highly hydrophobic chemicals), as a function of the lipid content. Suitable methods should be used for determination of lipid content (some existing methods – e.g. (31) (58) – should be adapted for this purpose). These methods use a chloroform/methanol extraction technique. However, to avoid the use of chlorinated solvents, a modification of the Bligh and Dyer method (9) as described in (17) should be used. Since the various methods may not give identical values, it is important to give details of the method used. When possible, i.e. if sufficient worm tissue is available, the lipid analysis should ideally be made on the same sample or extract as the one used for analysis of the test chemical, since the lipids often have to be removed from the extract before it can be analysed chromatographically (49). Alternatively, control animals may be used to measure the lipid content, which can then be used to normalise BAF values. This latter approach reduces the contamination of equipment with the test chemical.

**▼ M4****DATA AND REPORTING****Treatment of results**

54. The uptake curve of the test chemical is obtained by plotting its concentration in/on the worms during the uptake phase against time on arithmetic scales. When the curve has reached a plateau, or steady state (see definitions in Appendix 1), the steady state bioaccumulation factor BAF<sub>ss</sub> is calculated from:

$$\frac{C_a \text{ at steady state or at end of uptake phase (mean)}}{C_s \text{ at steady state or at end of uptake phase (mean)}}$$

$C_a$  is the concentration of test chemical in the test organism

$C_s$  is the concentration of test chemical in the soil

55. When no steady state is reached, the BAF<sub>K</sub>, based on the rate constants, should be determined instead of BAF<sub>ss</sub>, as described below:

- Determine the accumulation factor (BAF<sub>K</sub>) as the ratio  $k_s/k_e$ .
- Uptake and elimination rates are preferably calculated simultaneously (see Equation 11 in Appendix 2)
- The elimination rate constant ( $k_e$ ) is usually determined from the elimination curve (i.e. a plot of the concentration of the test item in the worms during the elimination phase). The uptake rate constant  $k_s$  is then calculated given  $k_e$  and a value of  $C_a$  which is derived from the uptake curve – See Appendix 2 for a description of these methods. The preferred method for obtaining BAF<sub>K</sub> and the rate constants,  $k_s$ , and  $k_e$ , is to use non-linear parameter estimation methods on a computer. If the elimination is obviously not first-order, then more complex models should be employed.

**Test report**

56. The test report should include the following information:

*Test chemical:*

- Any available information on acute or long term toxicity (e.g. EC<sub>x</sub>, LC<sub>x</sub>, NOEC) of the test chemical towards soil-dwelling oligochaetes;
- purity, physical nature and, physicochemical properties e.g. log K<sub>ow</sub>, water solubility;
- chemical identification data; source of the test item, identity and concentration of any solvent used;
- if radiolabelled test chemical is used, the precise position of the labelled atoms, the specific radioactivity, and the radiochemical purity.

*Test species:*

- scientific name, strain, source, any pre-treatment, acclimation, age, size-range, etc..

**▼ M4***Test conditions:*

- test procedure used;
- type and characteristics of illumination used and photoperiod(s);
- test design (e.g. number and size of test vessels, soil mass and height of soil layer, number of replicates, number of worms per replicate, number of test concentrations, duration of uptake and elimination phases, sampling frequency);
- rationale for the choice of test vessel material;
- method of test item preparation and application as well as reasons for choosing a specific method;
- the nominal test concentrations, the means of the measured values and their standard deviations in the test vessels, and the method by which these values were obtained;
- source of the constituents of the artificial soil or – if natural media are used – origin of the soil, description of any pre-treatment, results of the controls (survival, biomass development, reproduction), soil characteristics (pH, total organic carbon content, particle size distribution (percent sand, silt, and clay), WHC<sub>max</sub>, percent water content at start and at end of the test, and any other measurements made);
- detailed information on the treatment of soil and worm samples, including details of preparation, storage, spiking procedures, extraction, and analytical procedures (and precision) for the test item in worms and soil, and lipid content (if measured), and recoveries of the test item.

*Results:*

- mortality of the control worms and the worms in each test vessel and any observed abnormal behaviour (e.g. soil avoidance, lack of reproduction in a bioaccumulation test with enchytraeids);
- the dry weight to wet weight ratio of the soil and the test organisms (useful for normalisation);
- the wet weights of the worms at each sampling time; for earthworms, the wet weights at start of the test, and at each sampling time before and after gut purging;
- the lipid content of the test organisms (if determined);
- curves, showing the uptake and elimination kinetics of the test chemical in the worms, and the time to steady state;
- $C_a$  and  $C_s$  (with standard deviation and range, if appropriate) for all sampling times ( $C_a$  expressed in  $\text{g kg}^{-1}$  wet and dry weight of whole body,  $C_s$  expressed in  $\text{g kg}^{-1}$  wet and dry weight of soil). If a biota-soil accumulation factor (BSAF) is required (e.g. for comparison of results from two or more tests performed with animals of differing lipid content),  $C_a$  may additionally be expressed as  $\text{g kg}^{-1}$  lipid content of the organism, and  $C_s$  may be expressed as  $\text{g kg}^{-1}$  organic carbon (OC) of the soil;
- BAF (expressed in  $\text{kg soil}\cdot\text{kg}^{-1}$  worm), soil uptake rate constant  $k_s$  (expressed in  $\text{g soil kg}^{-1}$  of worm  $\text{day}^{-1}$ ), and elimination rate constant  $k_e$  (expressed in  $\text{day}^{-1}$ ); BSAF (expressed in  $\text{kg soil OC kg}^{-1}$  worm lipid content) may be reported additionally;

**▼M4**

— if measured: percentages of parent chemical, metabolites, and bound residues (i.e. the percentage of test chemical that cannot be extracted with common extraction methods) detected in soil and test animals;

— methods used for the statistical analyses of data.

*Evaluation of results:*

— compliance of the results with the validity criteria as listed in paragraph 17;

— unexpected or unusual results, e.g. incomplete elimination of the test chemical from the test animals.

*LITERATURE:*

- (1) Amorim M (2000). Chronic and toxicokinetic behavior of Lindane ( $\gamma$ -HCH) in the Enchytraeid *Enchytraeus albidus*. Master thesis, University Coimbra.
- (2) ASTM (2000). Standard guide for the determination of the bioaccumulation of sediment-associated contaminants by benthic invertebrates. American Society for Testing and Materials, E 1688-00a.
- (3) ASTM International (2004). Standard guide for conducting laboratory soil toxicity or bioaccumulation tests with the Lumbricid earthworm *Eisenia fetida* and the Enchytraeid potworm *Enchytraeus albidus*. ASTM International, E1676-04: 26 pp.
- (4) Beek B, Boehling S, Bruckmann U, Franke C, Joehncke U, Studinger G (2000). The assessment of bioaccumulation. In Hutzinger, O. (editor), *The Handbook of Environmental Chemistry, Vol. 2 Part J* (Vol. editor: B. Beek): Bioaccumulation — New Aspects and Developments. Springer-Verlag Berlin Heidelberg: 235-276.
- (5) Belfroid A, Sikkenk M, Seinen W, Van Gestel C, Hermens J (1994). The toxicokinetic behavior of chlorobenzenes in earthworms (*Eisenia andrei*): Experiments in soil. *Environ. Toxicol. Chem.* 13: 93-99.
- (6) Belfroid A, Van Wezel A, Sikkenk M, Van Gestel C, Seinen W & Hermens J (1993). The toxicokinetic behavior of chlorobenzenes in earthworms (*Eisenia andrei*): Experiments in water. *Ecotox. Environ. Safety* 25: 154-165.
- (7) Belfroid A, Meiling J, Drenth H, Hermens J, Seinen W, Van Gestel C (1995). Dietary uptake of superlipophilic compounds by earthworms (*Eisenia andrei*). *Ecotox. Environ. Safety* 31: 185-191.
- (8) Bell AW (1958). The anatomy of *Enchytraeus albidus*, with a key to the species of the genus *Enchytraeus*. *Ann. Mus. Novitat.* 1902: 1-13.
- (9) Bligh EG and Dyer WJ (1959). A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* 37: 911-917.
- (10) Bouche M (1972). *Lombriciens de France. Ecologie et Systematique.* INRA, Annales de Zoologie-Ecologie animale, Paris, 671 p.
- (11) Bruns E, Egeler Ph, Moser T, Römbke J, Scheffczyk A, Spörlein P (2001a). Standardisierung und Validierung eines Bioakkumulationstests mit terrestrischen Oligochaeten. Report to the German Federal Environmental Agency (Umweltbundesamt Berlin), R & D No.: 29864416.
- (12) Bruns E, Egeler Ph, Römbke J, Scheffczyk A, Spörlein P (2001b). Bioaccumulation of lindane and hexachlorobenzene by the oligochaetes *Enchytraeus luxuriosus* and *Enchytraeus albidus* (Enchytraeidae, Oligochaeta, Annelida). *Hydrobiologia* 463: 185-196.
- (13) Conder JM and Lanno RP (2003). Lethal critical body residues as measures of Cd, Pb, and Zn bioavailability and toxicity in the earthworm *Eisenia fetida*. *J. Soils Sediments* 3: 13-20.

▼ **M4**

- (14) Connell DW and Markwell RD (1990). Bioaccumulation in the Soil to Earthworm System. *Chemosphere* 20: 91-100.
- (15) Didden WAM (1993). Ecology of Terrestrial Enchytraeidae. *Pedobiologia* 37: 2-29.
- (16) Didden W (2003). Oligochaeta, In: Bioindicators and biomonitors. Markert, B.A., Breure, A.M. & Zechmeister, H.G. (eds.). Elsevier Science Ltd, The Netherlands, pp. 555-576.
- (17) De Boer J, Smedes F, Wells D, Allan A (1999). Report on the QUASH interlaboratory study on the determination of total-lipid in fish and shellfish. Round 1 SBT-2, Exercise 1000, EU, Standards, Measurement and Testing Programme.
- (18) Dietrich DR, Schmid P, Zweifel U, Schlatter C, Jenni-Eiermann S, Bachmann H, Bühler U, Zbinden N (1995). Mortality of birds of prey following field application of granular carbofuran: A Case Study. *Arch. Environ. Contam. Toxicol.* 29: 140-145.
- (19) Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC (OJ L 396, 30.12.2006, p. 1).
- (20) Edwards CA and Bohlen PJ (1996). Biology and ecology of earthworms. Third Edition, Chapman & Hall, London, 426 pp.
- (21) OECD (2008), *Bioaccumulation in Sediment-dwelling Benthic Oligochaetes*, Test Guideline No 315, Guidelines for the testing of chemicals, OECD, Paris
- (22) Egeler Ph, Gilberg D, Scheffczyk A, Moser Th and Römbke J (2009). Validation of a Soil Bioaccumulation Test with Terrestrial Oligochaetes by an International Ring Test (Validierung einer Methode zur standardisierten Messung der Bioakkumulation mit terrestrischen Oligochaeten). Report to the Federal Environmental Agency (Umweltbundesamt Dessau-Rosslau), R & D No.: 20467458: 149 pp. Available for download at: <http://www.oecd.org/dataoecd/12/20/42552727.pdf>.
- (23) Elmgaard N and Jagers op Akkerhuis GAJM (2000). Safety factors in pesticide risk assessment, Differences in species sensitivity and acute-chronic relations. National Environmental Research Institute, NERI Technical Report 325: 57 pp.
- (24) Environment Canada (1995). Guidance document on measurement of toxicity test precision using control sediments spiked with a reference toxicant. Environmental Protection Series Report EPS 1/RM/30.
- (25) EPPO (2003). Environmental Risk Assessment scheme for plant protection products. Soil organisms and functions, EPPO (European Plant Protection Organization) Standards, Bull, OEPP/EPPO 33: 195-208.
- (26) Franke C (1996). How meaningful is the bioconcentration factor for risk assessment? *Chemosphere* 32: 1897-1905.
- (27) Franke C, Studinger G, Berger G, Böhling S, Bruckmann U, Cohors-Fresenborg D, Jöhncke U (1994). The assessment of bioaccumulation. *Chemosphere* 29: 1501-1514.
- (28) Füll C (1996). Bioakkumulation und Metabolismus von -1,2,3,4,5,6-Hexachlorcyclohexan (Lindan) und 2-(2,4-Dichlorphenoxy)-propionsäure (Dichlorprop) beim Regenwurm *Lumbricus rubellus* (Oligochaeta, Lumbricidae). Dissertation University Mainz, 156 pp.

▼ **M4**

- (29) Füll C, Schulte C, Kula C (2003). Bewertung der Auswirkungen von Pflanzenschutzmitteln auf Regenwürmer. UWSF — Z. Umweltchem, Ökotox. 15: 78-84.
- (30) Gabric A.J, Connell DW, Bell PRF (1990). A kinetic model for bioconcentration of lipophilic compounds by oligochaetes. Wat. Res. 24: 1225-1231.
- (31) Gardner WS, Frez WA, Cichocki EA, Parrish CC (1985). Micromethods for lipids in aquatic invertebrates. Limnology and Oceanography 30: 1099-1105.
- (32) Hawker DW and Connell DW (1988). Influence of partition coefficient of lipophilic compounds on bioconcentration kinetics with fish. Wat. Res. 22: 701-707.
- (33) Hund-Rinke K and Wiechering H (2000). Earthworm avoidance test for soil assessments: An alternative for acute and reproduction tests. J. Soils Sediments 1: 15-20.
- (34) Hund-Rinke K, Römbke J, Riepert F, Achazi R (2000). Beurteilung der Lebensraumfunktion von Böden mit Hilfe von Regenwurmtests. In: Toxikologische Beurteilung von Böden. Heiden, S., Erb, R., Dott, W. & Eisen-traeager, A. (eds.), Spektrum Verl., Heidelberg, 59-81.
- (35) ISO 11268-2 (1998) Soil Quality – Effects of pollutants on earthworms (*Eisenia fetida*). Part 2: Determination of effects on reproduction.
- (36) Jaenike J (1982). '*Eisenia foetida*' is two biological species. Megadrilogica 4: 6-8.
- (37) Jager T (1998). Mechanistic approach for estimating bioconcentration of organic chemicals in earthworms (Oligochaeta). Environ. Toxicol. Chem. 17: 2080-2090.
- (38) Jager T, Sanchez PA, Muijs B, van der Welde E, Posthuma L (2000). Toxicokinetics of polycyclic aromatic hydrocarbons in *Eisenia andrei* (Oligochaeta) using spiked soil. Environ. Toxicol. Chem. 19: 953-961.
- (39) Jager T, Baerselman R, Dijkman E, De Groot AC, Hogendoorn EA, DeJong A, Kruitbosch JAW, Peijnenburg W J G. M (2003a). Availability of polycyclic aromatic hydrocarbons to earthworms (*Eisenia andrei*, Oligochaeta) in field-polluted soils and soil-sediment mixtures. Environ. Toxicol. Chem. 22: 767-775.
- (40) Jager T, Fleuren RLJ, Hoogendoorn E, de Korte G (2003b). Elucidating the routes of exposure for organic chemicals in the earthworm, *Eisenia andrei* (Oligochaeta). Environ. Sci. Technol. 37: 3399-3404.
- (41) Janssen MPM, Bruins A, De Vries TH, Van Straalen NM (1991). Comparison of cadmium kinetics in four soil arthropod species. Arch. Environ. Contam. Toxicol. 20: 305-312.
- (42) Kasprzak K (1982). Review of enchytraeid community structure and function in agricultural ecosystems. Pedobiologia 23: 217-232.
- (43) Khalil AM (1990). Aufnahme und Metabolismus von <sup>14</sup>C-Hexachlorbenzol und <sup>14</sup>C-Pentachlornitrobenzol in Regenwürmern. Dissertation University München, 137 pp.
- (44) Landrum PF (1989). Bioavailability and toxicokinetics of polycyclic aromatic hydrocarbons sorbed to sediments for the amphipod *Pontoporeia hoyi*. Environ. Sci. Toxicol. 23: 588-595.
- (45) Marinussen MPJC, Van der Zee SEATM, De Haan FA (1997). Cu accumulation in *Lumbricus rubellus* under laboratory conditions compared with accumulation under field conditions. Ecotox. Environ. Safety 36: 17-26.
- (46) Mount DR, Dawson TD, Burkhard LP (1999). Implications of gut purging for tissue residues determined in bioaccumulation testing of sediment with *Lumbriculus variegates*. Environ. Toxicol. Chem. 18: 1244-1249.

▼ **M4**

- (47) Nendza M (1991). QSARs of bioaccumulation: Validity assessment of log  $K_{ow}$ /log BCF correlations, In: R. Nagel and R. Loskill (eds.): Bioaccumulation in aquatic systems, Contributions to the assessment, Proceedings of an international workshop, Berlin 1990, VCH, Weinheim.
- (48) Chapter C.8 of this Annex, Toxicity for Earthworms
- (49) Chapter C.13 of this Annex, Bioconcentration: flow-through fish test.
- (50) Chapter C.21 of this annex, Soil Microorganisms: Nitrogen Transformation Test.
- (51) OECD (2004a), Enchytraeid reproduction test, Test Guideline No 220, Guidelines for the testing of chemicals, OECD, Paris.
- (52) OECD (2004b), Earthworm reproduction test (*Eisenia fetida*/*Eisenia Andrei*), Test Guideline No 222, Guidelines for the testing of chemicals, OECD, Paris.
- (53) OECD (2008), Bioaccumulation in Sediment-dwelling Benthic Oligochaetes, Test Guideline No 315, Guidelines for the testing of chemicals, OECD, Paris.
- (54) Petersen H and Luxton M (1982). A comparative analysis of soil fauna populations and their role in decomposition processes. *Oikos* 39: 287-388.
- (55) Phillips DJH (1993). Bioaccumulation. In: Handbook of Ecotoxicology Vol. 1. Calow P. (ed.). Blackwell Scientific Publ., Oxford. 378-396.
- (56) Pflugmacher J (1992). Struktur-Aktivitätsbestimmungen (QSAR) zwischen der Konzentration von Pflanzenschutzmitteln und dem Octanol-Wasser-Koeffizienten UWSF- Z. Umweltchem. Ökotox. 4: 77-81.
- (57) Posthuma L, Weltje L, Anton-Sanchez FA (1996). Joint toxic effects of cadmium and pyrene on reproduction and growth of the earthworm *Eisenia fetida*. RIVM Report No 607506001, Bilthoven.
- (58) Randall RC, Lee II H, Ozretich RJ, Lake JL, Pruell RJ (1991). Evaluation of selected lipid methods for normalising pollutant bioaccumulation. *Environ.Toxicol. Chem.* 10: 1431-1436.
- (59) Römbke J, Egele, P, Füll C (1998). Literaturstudie über Bioakkumulationstests mit Oligochaeten im terrestrischen Medium. UBA-Texte 28/98, 84 S.
- (60) Römbke J and Moser Th (1999). Organisation and performance of an international ring-test for the validation of the Enchytraeid reproduction test. UBA-Texte 4/99: 373 pp.
- (61) Römbke J, Riepert F, Achazi R (2000). Enchytraeen als Testorganismen, In: Toxikologische Beurteilung von Böden. Heiden, S., Erb, R., Dott, W. & Eisentraeger, A. (eds.). Spektrum Verl., Heidelberg. 105-129.
- (62) Romijn CA,FM, Luttik R, Van De Meent D, Slooff W, Canton JH (1993). Presentation of a General Algorithm to Include Effect Assessment on Secondary Poisoning in the Derivation of Environmental Quality Criteria, Part 2: Terrestrial food chains. *Ecotox. Envir. Safety* 27: 107-127.
- (63) Sample BE, Suter DW, Beauchamp JJ, Efrogmson RA (1999). LITERATURE-derived bioaccumulation models for earthworms: Development and validation. *Environ. Toxicol. Chem.* 18: 2110-2120.
- (64) Schlosser H-J and Riepert F (1992). Entwicklung eines Prüfverfahrens für Chemikalien an Bodenraubmilben (Gamasina), Teil 2: Erste Ergebnisse mit Lindan und Kaliumdichromat in subletaler Dosierung. *Zool. Beitr. NF* 34: 413-433.
- (65) Schmelz R and Collado R (1999). *Enchytraeus luxuriosus* sp. nov., a new terrestrial oligochaete species (Enchytraeide, Clitellata, Annelida). *Carolinae* 57: 93-100.



## ▼M4

- (66) Sims R W and Gerard BM (1985). Earthworms, In: Kermack, D. M. & Barnes, R. S. K. (Hrsg.): Synopses of the British Fauna (New Series) No 31.171 S. London: E. J. Brill/Dr W. Backhuys.
- (67) Sousa JP, Loureiro S, Pieper S, Frost M, Kratz W, Nogueira AJA, Soares AMVM (2000). Soil and plant diet exposure routes and toxicokinetics of lindane in a terrestrial isopod. *Environ. Toxicol. Chem.* 19: 2557–2563.
- (68) Spacie A and Hamelink JL (1982). Alternative models for describing the bioconcentration of organics in fish. *Environ. Toxicol. Chem.* 1, 309-320.
- (69) Stephenson GL, Kaushik A, Kaushik NK, Solomon KR, Steele T, Scroggins RP (1998). Use of an avoidance-response test to assess the toxicity of contaminated soils to earthworms. In: *Advances in earthworm ecotoxicology*. S. Sheppard, J. Bembridge, M. Holmstrup, L. Posthuma (eds.). Setac Press, Pensacola, 67-81.
- (70) Sterenborg I, Vork NA, Verkade SK, Van Gestel CAM, Van Straalen NM (2003). Dietary zinc reduces uptake but not metallothionein binding and elimination of cadmium in the springtail *Orchesella cincta*. *Environ. Toxicol. Chemistry* 22: 1167-1171.
- (71) UBA (Umweltbundesamt) (1991). Bioakkumulation — Bewertungskonzept und Strategien im Gesetzesvollzug. UBA-Texte 42/91. Berlin.
- (72) US EPA (2000). Methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates. Second Edition, EPA 600/R-99/064, US, Environmental Protection Agency, Duluth, MN, March 2000.
- (73) Van Brummelen TC and Van Straalen NM (1996). Uptake and elimination of benzo(a)pyrene in the terrestrial isopod *Porcellio scaber*. *Arch. Environ. Contam. Toxicol.* 31: 277-285.
- (74) Van Gestel CAM. (1992). The influence of soil characteristics on the toxicity of chemicals for earthworms; a review, In: *Ecotoxicology of Earthworms* (Ed. Becker, H, Edwards, PJ, Greig-Smith, PW & Heimbach, F). Intercept Press, Andover (GB).
- (75) Van Gestel CA and Ma W-C (1990). An approach to quantitative structure-activity relationships (QSARs) in earthworm toxicity studies. *Chemosphere* 21: 1023-1033.
- (76) Van Straalen NM, Donker MH, Vijver MG, van Gestel CAM (2005). Bioavailability of contaminants estimated from uptake rates into soil invertebrates. *Environmental Pollution* 136: 409-417.
- (77) Venter JM and Reinecke AJ (1988). The life-cycle of the compost-worm *Eisenia fetida* (Oligochaeta). *South African J. Zool.* 23: 161-165.
- (78) Vijver MG, Vink JPM, Jager T, Wolterbeek HT, van Straalen NM, van Gestel CAM (2005). Biphasic elimination and uptake kinetics of Zn and Cd in the earthworm *Lumbricus rubellus* exposed to contaminated floodplain soil. *Soil Biol, Biochem.* 37: 1843-1851.
- (79) Widianarko B and Van Straalen NM (1996). Toxicokinetics-based survival analysis in bioassays using nonpersistent chemicals, *Environ. Toxicol. Chem.* 15: 402–406.

▼ M4*Appendix 1*

## DEFINITIONS

**Bioaccumulation** is the increase in concentration of the test chemical in or on an organism relative to the concentration of the test chemical in the surrounding medium. Bioaccumulation results from both bioconcentration and biomagnification processes (see below).

**Bioconcentration** is the increase in concentration of the test chemical in or on an organism, resulting from the uptake of the chemical exclusively from the surrounding medium (i.e. via the body surface and ingested soil), relative to the concentration of the test chemical in the surrounding medium.

**Biomagnification** is the increase in concentration of the test chemical in or on an organism, resulting mainly from uptake from contaminated food or prey, relative to the concentration of the test chemical in the food or prey. Biomagnification can lead to a transfer or accumulation of the test item within food webs.

The **elimination** of a test chemical is the loss of this chemical from the test organism tissue by active or passive processes that occurs independently of presence or absence of the test item in the surrounding medium.

The **bioaccumulation factor** (BAF) at any time during the uptake phase of this bioaccumulation test is the concentration of test chemical in/on the test organism ( $C_a$  in  $\text{g}\cdot\text{kg}^{-1}$  dry weight of worm) divided by the concentration of the chemical in the surrounding medium ( $C_s$  as  $\text{g}\cdot\text{kg}^{-1}$  of dry weight of soil); the BAF has the units of  $\text{kg soil}\cdot\text{kg}^{-1}$  worm.

The **steady state bioaccumulation factor** ( $\text{BAF}_{\text{ss}}$ ) is the BAF at steady state and does not change significantly over a prolonged period of time, the concentration of the test chemical in the surrounding medium ( $C_s$  as  $\text{g}\cdot\text{kg}^{-1}$  of dry weight of soil) being constant during this period of time.

**Bioaccumulation factors** calculated directly from the ratio of the soil uptake rate constant and the elimination rate constant ( $k_s$  and  $k_e$ , see below) are termed kinetic bioaccumulation factor ( $\text{BAF}_k$ ).

The **biota-soil accumulation factor** (BSAF) is the lipid-normalised concentration of the test chemical in/on the test organism divided by the organic carbon-normalised concentration of the test chemical in the soil at steady state.  $C_a$  is then expressed as  $\text{g}\cdot\text{kg}^{-1}$  lipid content of the organism, and  $C_s$  as  $\text{g}\cdot\text{kg}^{-1}$  organic content of the soil; the BSAF has the units of  $\text{kg OC}\cdot\text{kg}^{-1}$  lipid.

A **plateau** or **steady state** is defined as the equilibrium between the uptake and elimination processes that occur simultaneously during the exposure phase. The steady state is reached in the plot of BAF against time when the curve becomes parallel to the time axis and three successive analyses of BAF made on samples taken at intervals of at least two days are within 20 % of each other, and there are no statistically significant differences among the three sampling periods. For test chemicals which are taken up slowly, more appropriate intervals would be seven days (49).

The **organic carbon-water partitioning coefficient** ( $K_{\text{oc}}$ ) is the ratio of a chemical's concentration in/on the organic carbon fraction of a soil and the chemical's concentration in water at equilibrium.

The **octanol-water partitioning coefficient** ( $K_{\text{ow}}$ ) is the ratio of a chemical's solubility in n-octanol and water at equilibrium, also sometimes expressed as  $P_{\text{ow}}$ . The logarithm of  $K_{\text{ow}}$  ( $\log K_{\text{ow}}$ ) is used as an indication of a chemical's potential for bioaccumulation by aquatic organisms.

**▼ M4**

The **uptake or exposure phase** is the time during which the test organisms are exposed to the test chemical.

The **soil uptake rate constant** ( $k_s$ ) is the numerical value defining the rate of increase in the concentration of the test item in/on the test organism resulting from uptake from the soil phase.  $k_s$  is expressed in  $\text{g soil kg}^{-1}$  of worm  $\text{d}^{-1}$ .

The **elimination phase** is the time, following the transfer of the test organisms from a contaminated medium to a medium free of the test item, during which the elimination (or the net loss) of the chemical from the test organisms is studied.

The **elimination rate constant** ( $k_e$ ) is the numerical value defining the rate of reduction in the concentration of the test item in/on the test organism, following the transfer of the test organisms from a medium containing the test item to a chemical-free medium;  $k_e$  is expressed in  $\text{d}^{-1}$ .

**Test chemical:** Any substance or mixture tested using this Test Method.

▼ **M4***Appendix 2***Calculation of uptake and elimination parameters**

The main endpoint of a bioaccumulation test is the bioaccumulation factor, BAF. The measured BAF can be calculated by dividing the concentration in the test organism,  $C_a$ , by the concentration in the soil,  $C_s$ , at steady state. If the steady state is not reached during the uptake phase, the  $BAF_K$  is calculated from the rate constants instead of BAFs. However, it should be noted if the BAF is based on steady state concentrations or not.

The usual means for obtaining the kinetic bioaccumulation factor ( $BAF_K$ ), the soil uptake rate constant ( $k_s$ ) and the elimination rate constant ( $k_e$ ) is to use non-linear parameter estimation methods on a computer, e.g. based on the models described in (68). Given a set of sequential time concentration data and the model equations:

$$C_a = \frac{k_s}{k_e} \times C_s (1 - e^{-k_e t}) \quad 0 < t < t_c \quad \text{[equation 1]}$$

or

$$C_a = \frac{k_s}{k_e} \times C_s (e^{-k_e (t-t_c)} - e^{-k_e t}) \quad t > t_c \quad \text{[equation 2]}$$

where:

$C_a$  = concentration of chemical in worms [g kg<sup>-1</sup> wet or dry weight]

$k_s$  = uptake rate constant in tissue [g soil kg<sup>-1</sup> of worm d<sup>-1</sup>]

$C_s$  = concentration of chemical in soil [g kg<sup>-1</sup> of wet or dry weight]

$k_e$  = elimination rate constant [d<sup>-1</sup>]

$t_c$  = time at the end of the uptake phase,

these computer programs calculate values for  $BAF_K$ ,  $k_s$  and  $k_e$ .

When the background concentration in the non-exposed worms e.g. on day 0 differs significantly from zero (this may e.g. be the case for metals), this background concentration ( $C_{a,0}$ ) should be included in these equations, to make them read:

$$C_a = C_{a,0} + \frac{k_s}{k_e} \times C_s (1 - e^{-k_e t}) \quad 0 < t < t_c \quad \text{[equation 3]}$$

and

$$C_a = C_{a,0} + \frac{k_s}{k_e} \times C_s (e^{-k_e (t-t_c)} - e^{-k_e t}) \quad t > t_c \quad \text{[equation 4]}$$

In cases where a significant decrease of the test chemical concentration in the soil is observed over time during the uptake phase, the following models can be used e.g. (67) (79):

$$C_s = C_0 (e^{-k_0 t}) \quad \text{[equation 5]}$$

▼ **M4**

where:

$C_s$  = concentration of chemical in the soil [g kg-1 wet or dry weight]

$k_0$  = degradation rate constant in soil [d-1]

$C_0$  = initial concentration of chemical in soil [g kg-1 of wet or dry weight]

$$C_a = \frac{k_s}{k_e - k_0} \times (e^{-k_0 t} - e^{-k_e t}) \quad 0 < t < t_c \quad [\text{equation 6}]$$

$$C_a = \frac{k_s}{k_e - k_0} \times e^{-k_0 t_c} - e^{-k_e t_c} * e^{-k(t-t_c)} \quad t > t_c \quad [\text{equation 7}]$$

where:

$C_a$  = concentration of chemical in worms [g kg-1 wet or dry weight]

$k_s$  = uptake rate constant in tissue [g soil kg-1 of worm d-1]

$k_0$  = degradation rate constant in soil [d-1]

$k_e$  = elimination rate constant [d-1]

$t_c$  = time at the end of the uptake phase.

When steady state is reached during the uptake phase (i.e.  $t = \infty$ ), equation 1

$$C_a = \frac{k_s}{k_e} \times C_s (1 - e^{-k_e t}) \quad 0 < t < t_c \quad [\text{equation 1}]$$

may be reduced to:

$$C_a = \frac{k_s}{k_e} \times C_s$$

or

$$C_a/C_s = k_s/k_e = \text{BAF}_K \quad [\text{equation 8}]$$

Then  $k_s/k_e \times C_s$  is an approach to the concentration of the test item in the worm tissue at steady state ( $C_{a,ss}$ ).

The biota-soil accumulation factor (BSAF) can be calculated as follows:

$$\text{BSAF} = \text{BAF}_K * \frac{f_{oc}}{f_{lip}} \quad [\text{equation 9}]$$

where  $f_{oc}$  is the fraction of soil organic carbon, and  $f_{lip}$  is the fraction of worm lipid, both preferably determined on samples taken from the test, and based either on dry weight or on wet weight, respectively.

The elimination kinetics can be modelled using the data from the elimination phase and applying the following model equation and a computer-based non-linear parameter estimation method. If the data points plotted against time indicate a constant exponential decline of the test item concentration in the animals, a one-compartment model (equation 9) can be used to describe the time course of elimination.

$$C_a(t) = C_{a,ss} \times e^{-k_e t} \quad [\text{equation 10}]$$

**▼ M4**

Elimination processes sometimes appear to be biphasic, showing a rapid decline of  $C_a$  during the early phases, that changes to a slower loss of test items in the later phases of the elimination, e.g. (27) (68). The two phases can be interpreted by the assumption, that there are two different compartments in the organism, from which the test item is lost with different velocities. In these cases, specific LITERATURE should be studied e.g. (38) (39) (40) (78).

Using the model equations above, the kinetic parameters ( $k_s$  and  $k_e$ ) may also be calculated in one run by applying the first order kinetics model to all data from both the uptake and elimination phase simultaneously. For a description of a method that may allow for such a combined calculation of uptake and elimination rate constants, references (41), (73) and (70) may be consulted.

$$C_a = \left[ \frac{K_s}{K_e} \cdot C_s (1 - e^{-k_e t}) \times (m = 1) \right] + \left[ \frac{K_s}{k_e} \times C_s (e^{-K_e(t-t_0)} - e^{-K_e t}) \times (m = 2) \right] \quad \text{[equation 11]}$$

*Note:* When uptake and elimination parameters are estimated simultaneously from the combined uptake and the elimination data, 'm' as shown in equation 11 is a descriptor that allows the computer program to assign the equation's sub-terms to the data sets of the respective phase and to perform the evaluation correctly (m = 1 for uptake phase; m = 2 for elimination phase).

Nevertheless, these model equations should be used with caution, especially when changes in the test chemical's bioavailability, or (bio)degradation occur during the test (see e.g. (79)).

▼ **M4***Appendix 3*

## EXAMPLES OF SCHEDULES FOR SOIL BIOACCUMULATION TESTS

**Earthworm test**

- (a) Uptake phase with 8 sampling dates used for calculation of kinetics

Day	Activity
- 6	Conditioning of the prepared soil for 48 h;
- 4	Spiking of the soil fraction with the test chemical solution; evaporating of any solvent; mixing of the soil constituents; distributing the soil to the test vessels; equilibration at test conditions for 4 days (3 weeks for metal-spiked soil);
- 3 to - 1	Separation of the test organisms from the culture for acclimation; preparation and moisturising of the soil constituents;
0	Measuring temperature, and soil pH; removing soil samples from treated vessels and solvent controls for determination of test chemical concentration; addition of food ration; weighing and randomised distribution of the worms to the test vessels; retaining of sufficient subsamples of worms for determination of analytical background values, wet and dry weight, and lipid content; weighing of all test vessels to control soil moisture; controlling air supply, if closed test system is used;
1	Controlling air supply, recording worm behaviour and temperature; taking soil and worm samples for determination of test item concentration;
2	Same as day 1;
3	Controlling air supply, worm behaviour and temperature;
4	Same as day 1;
5-6	Same as day 3;
7	Same as day 1; addition of food ration; control soil moisture by re-weighing the test vessels and compensate evaporated water;
8-9	Same as day 3;
10	Same as day 1;
11-13	Same as day 3;
14	Same as day 1; addition of food ration; control soil moisture by re-weighing the test vessels and compensate evaporated water;
15-16	Same as day 3;
17	Same as day 1;
18-20	Same as day 3;

▼ **M4**

Day	Activity
21	Same as day 1; measuring temperature and soil pH; control soil moisture by re-weighing the test vessels; end of uptake phase; transfer worms from remaining exposed replicates to vessels containing clean soil for elimination phase (no gut-purging); sampling of soil and worms from solvent controls.
	Pre-exposure activities (equilibration phase) should be scheduled taking into account the properties of the test chemical.
	Activities described for day 3 should be performed daily (at least on workdays).

## (b) Elimination phase

Day	Activity
- 6	Preparation and moisturising of the soil constituents; conditioning of the prepared soil for 48 h;
- 4	Mixing of the soil constituents; distributing the soil to the test vessels; incubation at test conditions for 4 days;
0 (end of uptake phase)	Measuring temperature and soil pH; weighing and randomised distribution of the worms to the test vessels; addition of food ration; transfer worms from remaining exposed replicates to vessels containing clean soil; taking soil and worm samples after 4-6 h for determination of test chemical concentration;
1	Controlling air supply, recording worm behaviour and temperature; taking soil and worm samples for determination of test chemical concentration;
2	Same as day 1;
3	Controlling air supply, worm behaviour and temperature;
4	Same as day 1;
5-6	Same as day 3;
7	Same as day 1; addition of food ration; control soil moisture by re-weighing the test vessels and compensate evaporated water;
8-9	Same as day 3;
10	Same as day 1;
11-13	Same as day 3;
14	Same as day 1; addition of food ration; control soil moisture by re-weighing the test vessels and compensate evaporated water;
15-16	Same as day 3;
17	Same as day 1;



▼ **M4**

Day	Activity
18-20	Same as day 3;
21	Same as day 1; measuring temperature and soil pH; control soil moisture by re-weighing the test vessels; sampling of soil and worms from solvent controls.
	Preparation of the soil prior to start of elimination phase should be done in the same manner as before the uptake phase.
	Activities described for day 3 should be performed daily (at least on workdays).

**Enchytraeid test**

- (a) Uptake phase with 8 sampling dates used for calculation of kinetics

Day	Activity
- 6	Conditioning of the prepared soil for 48 h;
- 4	Spiking of the soil fraction with the test chemical solution; evaporating of any solvent; mixing of the soil constituents; distributing the soil to the test vessels; equilibration at test conditions for 4 days (3 weeks for metal-spiked soil);
- 3 to - 1	Separation of the test organisms from the culture for acclimation; preparation and moisturising of the soil constituents;
0	Measuring temperature, and soil pH; removing soil samples from treated vessels and solvent controls for determination of test chemical concentration; addition of food ration to soil; weighing and randomised distribution of the worms to the test vessels; retaining of sufficient subsamples of worms for determination of analytical background values, wet and dry weight, and lipid content; weighing of all test vessels to control soil moisture; controlling air supply, if closed test system is used;
1	Controlling air supply, recording worm behaviour and temperature; taking soil and worm samples for determination of test item concentration;
2	Same as day 1;
3	Controlling air supply, worm behaviour and temperature;
4	Same as day 1;
5-6	Same as day 3;
7	Same as day 1; addition of food ration to soil; control soil moisture by re-weighing the test vessels and compensate evaporated water;
9	Same as day 1;
10	Same as day 3;

▼ **M4**

Day	Activity
11	Same as day 1;
12-13	Same as day 3;
14	Same as day 1; addition of food ration to soil; measuring temperature and soil pH; control soil moisture by re-weighing the test vessels; end of uptake phase; transfer worms from remaining exposed replicates to vessels containing clean soil for elimination phase (no gut-purging); sampling of soil and worms from solvent controls.
	Pre-exposure activities (equilibration phase) should be scheduled taking into account the properties of the test chemical.
	Activities described for day 3 should be performed daily (at least on workdays).

▼ **M4***Appendix 4***Artificial soil – preparation and storage recommendations**

Since natural soils from a particular source may not be available throughout the year, and indigenous organisms as well as the presence of micro-pollutants can influence the test, an artificial substrate, the artificial soil according to Chapter C.8 of this Annex, Toxicity for Earthworms (48), is recommended for use in this test. Several test species can survive, grow, and reproduce in this soil, and maximum standardisation as well as intra- and interlaboratory comparability of test and culture conditions are provided.

Soil constituents:

Peat:	10 %	Sphagnum-peat, in accordance with the OECD Guideline 207 (48);
Quartz sand:	70 %	Industrial quartz sand (air dried); grain size: more than 50 % of the particles should be in the range of 50-200 µm, but all particles should be ≤ 2 mm;
Kaolinite clay:	20 %	Kaolinite content ≥ 30 %;
Calcium carbonate:	≤ 1 %	CaCO <sub>3</sub> , pulverised, chemically pure.

As an option, the organic carbon content of the artificial soil may be reduced, e.g. by lowering the peat content to 4-5 % of dry soil and increasing the sand content accordingly. By such a reduction in organic carbon content, the possibilities of adsorption of test chemical to the soil (organic carbon) may be decreased, and the availability of the test chemical to the worms may increase (74). It has been demonstrated that *Enchytraeus albidus* and *Eisenia fetida* can comply with the validity criteria on reproduction when tested in field soils with lower organic carbon content, e.g. 2,7 % (33), (61), and there is experience that this can also be achieved in artificial soil with 5 % peat.

**Preparation**

The dry constituents of the soil are mixed thoroughly (e.g. in a large-scale laboratory mixer). This should be done about one week before starting the test. The mixed dry soil constituents should be moistened with deionised water at least 48 h before application of the test item in order to equilibrate/stabilise the acidity. For the determination of pH a mixture of soil and 1 M KCl solution in a 1:5 ratio is used. If the pH value is not within the required range ( $6,0 \pm 0,5$ ), a sufficient amount of CaCO<sub>3</sub> is added to the soil, or a new batch of soil is prepared.

The maximum water holding capacity (WHC) of the artificial soil is determined according to ISO 11268-2 (35). At least two days before starting the test, the dry artificial soil is moistened by adding enough deionised or reconstituted water to obtain approximately half of the final water content. The final water content should be 40 % to 60 % of the maximum WHC. At the start of the test, the pre-moistened soil is divided into as many batches as the number of test concentrations and controls used for the test, and the moisture content is adjusted to 40-60 % of WHC<sub>max</sub> by using the solution of the test item and/or by adding deionised or reconstituted water. The moisture content is determined at the beginning and at the end of the test (at 105 °C). It should be optimal for the species' requirements (the moisture content can also be checked as follows: when the soil is gently squeezed in the hand, small drops of water should appear between the fingers).

**▼ M4****Storage**

The dry constituents of the artificial soil may be stored at room temperature until use. The prepared, pre-moistened soil may be stored in a cool place for up to three days prior to spiking; care should be taken to minimise evaporation of water. Soil spiked with the test item should be used immediately unless there is information indicating that the particular soil can be stored without affecting the toxicity and bioavailability of the test item. Samples of spiked soil may then be stored under the conditions recommended for the particular test item until analysis.

▼ **M4***Appendix 5***Species of terrestrial oligochaetes recommended for testing bioaccumulation from soil****Earthworms**

The recommended test species is *Eisenia fetida* (Savigny 1826), belonging to the family Lumbricidae. Since 1972 it is divided into two subspecies (*Eisenia fetida* and *Eisenia andrei* (10)). According to Jaenike (36), they are true, separate species. *Eisenia fetida* is easily recognised by its bright intersegmental yellow stripes whereas *Eisenia andrei* has a uniform, dark red colour. Originating probably from the region of the Black Sea, they are distributed worldwide today, especially in anthropogenically modified habitats like compost heaps. Both can be used for ecotoxicological as well as bioaccumulation tests.

*Eisenia fetida* and *Eisenia andrei* are commercially available, e.g. as fish bait. In comparison to other lumbricid earthworms, they have a short life-cycle, reaching maturity within ca. 2-3 months (at room temperature). Their optimum temperature is approximately at 20-24 °C. They prefer relatively moist substrates with a nearly neutral pH and a high content of organic material. Since these species have been widely used in standardised ecotoxicological tests for about 25 years, their culturing is well established (48) (77).

Both species can be bred in a wide range of animal wastes. The breeding medium recommended by ISO (35) is a 50:50 mixture of horse or cattle manure and peat. The medium should have a pH value of about 6 to 7 (regulated with calcium carbonate), a low ionic conductivity (less than 6 mS/cm or less than 0,5 % salt concentration) and should not be contaminated excessively with ammonia or animal urine. Also, a commercial gardening soil free of additives, or artificial soil according to OECD (48), or a 50:50 mixture of both can be used. The substrate should be moist but not too wet. Breeding boxes of 10 litre to 50 litre volume are suitable.

To obtain worms of standard age and mass, it is best to start the culture with cocoons. Therefore, adult worms are added to a breeding box containing fresh substrate to produce cocoons. Practical experience has shown that a population density of approximately 100 adult worms per kg substrate (wet weight) leads to good reproduction rates. After 28 days, the adult worms are removed. The earthworms hatched from the cocoons are used for testing when mature after at least 2 months but less than 12 months.

Worms of the species described above can be considered healthy if they move through the substrate, do not try to leave the substrate, and reproduce continuously. Very slow motioning or a yellow posterior end (in the case of *Eisenia fetida*) indicates substrate exhaustion. In this case, fresh substrate and/or a lower number of animals per box is recommended.

**Additional selected references**

Gerard BM (1964). Synopsis of the British fauna. No 6 Lumbricidae. Linnean Soc. London, 6: 1-58.

Graff O (1953). Die Regenwürmer Deutschlands. Schr. Forsch. Anst. Landwirtschaft. 7: 1-81.

Römbke J, Egeler P, Füll C (1997). Literaturstudie über Bioakkumulationstests mit Oligochaeten im terrestrischen Medium. Bericht für das UBA F + E 206 03 909, 86 S.

Rundgren S (1977). Seasonality of emergence in lumbricids in southern Sweden. Oikos 28: 49-55.

Satchell JE (1955). Some aspects of earthworm ecology. Soil Zoology (Kevan): 180-201.

**▼ M4**

Sims RW and Gerard BM (1985). A synopsis of the earthworms. Linnean Soc. London 31: 1-171.

Tomlin AD (1984). The earthworm bait market in North America. In: Earthworm Ecology — from Darwin to vermiculture. Satchell, J.E. (ed.), Chapman & Hall, London. 331-338 pp.

**Enchytraeids**

The recommended test species is *Enchytraeus albidus* Henle 1837 (white potworm). *Enchytraeus albidus* is one of the biggest (up to 15 mm) species of the annelid oligochaete family Enchytraeidae and it is worldwide distributed e.g. (8). *Enchytraeus albidus* is found in marine, limnic and terrestrial habitats, mainly in decaying organic matter (seaweed, compost) and rarely in meadows (42). This broad ecological tolerance and some morphological variations indicate that there might be different races for this species.

*Enchytraeus albidus* is commercially available, sold as food for fish. It should be checked whether the culture is contaminated by other, usually smaller species (60). If contamination occurs, all worms should be washed with water in a Petri dish. Large adult specimens of *Enchytraeus albidus* are then selected (by using a stereomicroscope) to start a new culture. All other worms are discarded. Its life cycle is short as maturity is reached between 33 days (at 18 °C) and 74 days (at 12 °C). Only cultures which have been kept in the laboratory for at least 5 weeks (one generation) without problems should be used for a test.

Other species of the *Enchytraeus* genus are also suitable, especially *Enchytraeus luxuriosus*. This species is a true soil inhabitant, which has been newly described in (65). If other species of *Enchytraeus* are used, they should be clearly identified and the rationale for the selection of the species should be reported.

*Enchytraeus crypticus* (Westheide & Graefe 1992) is a species belonging to the same group as *Enchytraeus luxuriosus*. It has not been found to exist with certainty in the field, having only been described from earthworm cultures and compost heaps (Römbke 2003). Its original ecological requirements are therefore not known. However, recent laboratory studies in various field soils have confirmed that this species has a broad tolerance towards soil properties like pH and texture (Jänsch et al. 2005). In recent years, this species has often been used in ecotoxicological studies because of the simplicity of its breeding and testing, e.g. Kuperman et al. 2003). However, it is small (3-12 mm; 7 mm on average (Westheide & Müller 1996), and this makes handling more difficult compared with *Enchytraeus albidus*. When using this species instead of *Enchytraeus albidus*, the size of the test vessel can but needs not to be smaller. In addition, it should be considered that this species reproduces very rapidly having a generation time of less than 20 days at 20 ± 2 °C (Achazi et al. 1999) and even quicker at higher temperatures.

Enchytraeids of the species *Enchytraeus albidus* (as well as other *Enchytraeus* species) can be bred in large plastic boxes (e.g. 30 × 60 × 10 cm or 20 × 12 × 8 cm which is suitable for culture of worms of small size) filled with a mixture of artificial soil and commercially available, uncontaminated garden soil free of additives. Compost material should be avoided since it could contain toxic chemicals like heavy metals. Fauna should be removed from the breeding soil before use by three times deep-freezing. Pure artificial soil can also be used but the reproduction rate could be slower compared to that obtained with mixed substrates. The substrate should have a pH of 6,0 ± 0,5. The culture is kept in an incubator at a temperature of 15 ± 2 °C without light. In any case, a temperature higher than 23 °C should be avoided. The artificial/natural soil moisture should be moist but not wet. When the soil is gently pressed by hand, only small drops of water should appear. In any case, anoxic conditions should be avoided (e.g. if a lid is used, the number of lid holes should be high enough to provide sufficient exchange of air). The breeding soil should be aerated by carefully mixing it once per week.

**▼M4**

The worms should be fed at least once per week ad libitum with rolled oats which are placed into a cavity on the soil surface and covered with soil. If food from the last feeding date remains in the container, the amount of food given should be adjusted accordingly. If fungi grow on the remaining food, it should be replaced by a new quantity of rolled oats. In order to stimulate reproduction, the rolled oats may be supplemented with commercially available, vitamin amended protein powder every two weeks. After three months, the animals are transferred to a freshly prepared culture or breeding substrate. The rolled oats, which have to be stored in sealed vessels, should be autoclaved or heated before use in order to avoid infections by flour mites (e.g. *Glyphagus* sp., Astigmata, Acarina) or predacious mites (e.g. *Hypoaspis (Cosmolaelaps) miles*, Gamasida, Acarina). After disinfecting, the food is ground up so that it can easily be strewn on the soil surface. Another possible food source is baker's yeast or the fish food TetraMin®.

In general, the culturing conditions are sufficient if worms do not try to leave the substrate, move quickly through the soil, exhibit a shiny outer surface without soil particles clinging to it, are more or less whitish coloured, and if worms of different ages are visible. Actually, worms can be considered healthy if they reproduce continuously.

**Additional selected references**

Achazi RK, Fröhlich E, Henneken M, Pilz C (1999). The effect of soil from former irrigation fields and of sewage sludge on dispersal activity and colonizing success of the annelid *Enchytraeus crypticus* (Enchytraeidae, Oligochaeta). Newsletter on Enchytraeidae 6: 117-126.

Jänsch S, Amorim MJB, Römbke J (2005). Identification of the ecological requirements of important terrestrial ecotoxicological test species. Environ. Reviews 13: 51-83.

Kuperman RG, Checkai RT, Simini M, Phillips CT, Kolakowski JE, Kurnas CW, Sunahara GI (2003). Survival and reproduction of *Enchytraeus crypticus* (Oligochaeta, Enchytraeidae) in a natural sandy loam soil amended with the nitro-heterocyclic explosives RDX and HMX. Pedobiologia 47: 651-656.

Römbke J (2003). Ecotoxicological laboratory tests with enchytraeids: A review. Pedobiologia 47: 607-616.

Westheide W and Graefe U (1992). Two new terrestrial *Enchytraeus* species (Oligochaeta, Annelida). J. Nat. Hist. 26: 479-488.

Westheide W and Müller MC (1996). Cinematographic documentation of enchytraeid morphology and reproductive biology. Hydrobiologia 334: 263-267.