Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

14-Aug-2001

English text only

ENV/JM/MONO(2001)6
Unclassified

ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND
THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

Cancels & replaces the same document of 14 August 2001

OECD SERIES ON TESTING AND ASSESSMENT
Number 33

HARMONISED INTEGRATED CLASSIFICATION SYSTEM FOR HUMAN HEALTH AND ENVIRONMENTAL HAZARDS OF CHEMICAL SUBSTANCES AND MIXTURES

JT00111570

English text only

HARMONISED INTEGRATED CLASSIFICATION SYSTEM FOR HUMAN HEALTH AND ENVIRONMENTAL HAZARDS OF CHEMICAL SUBSTANCES AND MIXTURES

As endorsed by:

- **the 28th Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, November 1998 (Part I; Part II: Chapters 1-7, 10);**
- **the 31st Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology, November 2000 (Part II: Chapters 8-9; Part III: Chapters 1, 2, 4-8);**
- **the 32nd Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology, June 2001 (Part III: Chapters 3, 9, Annex 2-3).**

Also published in the Series on Testing and Assessment:

No. 1, *Guidance Document for the Development of OECD Guidelines for Testing of Chemicals (1993; reformatted 1995)*

No. 2, *Detailed Review Paper on Biodegradability Testing (1995)*

No. 3, *Guidance Document for Aquatic Effects Assessment (1995)*

No. 4, *Report of the OECD Workshop on Environmental Hazard/Risk Assessment (1995)*

No. 5, *Report of the SETAC/OECD Workshop on Avian Toxicity Testing (1996)*

No. 6, *Report of the Final Ring-test of the Daphnia magna Reproduction Test (1997)*

No. 7, G*uidance Document on Direct Phototransformation of Chemicals in Water (1997)*

No. 8, *Report of the OECD Workshop on Sharing Information about New Industrial Chemicals Assessment (1997)*

No. 9, *Guidance Document for the Conduct of Studies of Occupational Exposure to Pesticides During Agricultural Application (1997)*

No. 10, *Report of the OECD Workshop on Statistical Analysis of Aquatic Toxicity Data (1998)*

No. 11, *Detailed Review Paper on Aquatic Testing Methods for Pesticides and industrial Chemicals (1998)*

No. 12, *Detailed Review Document on Classification Systems for Germ Cell Mutagenicity in OECD Member Countries (1998)*

No. 13, *Detailed Review Document on Classification Systems for Sensitising Substances in OECD Member Countries 1998)*

No. 14, *Detailed Review Document on Classification Systems for Eye Irritation/Corrosion in OECD Member Countries (1998)*

No. 15, *Detailed Review Document on Classification Systems for Reproductive Toxicity in OECD Member Countries (1998)*

No. 16, *Detailed Review Document on Classification Systems for Skin Irritation/Corrosion in OECD Member Countries(1998)*

No. 17, *Environmental Exposure Assessment Strategies for Existing Industrial Chemicals in OECD Member Countries (1999)*

No. 18, *Report of the OECD Workshop on Improving the Use of Monitoring Data in the Exposure Assessment of Industrial Chemicals (2000)*

No. 19, *Draft Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals used in Safety Evaluation (1999)*

No. 20, *Revised Draft Guidance Document for Neurotoxicity Testing (in preparation)*

No. 21, *Detailed Review Paper: Appraisal of Test Methods For Sex Hormone Disrupting Chemicals (2000)*

No. 22, *Guidance Document for the Performance of Out-door Monolith Lysimeter Studies (2000)*

No. 23, *Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures (2000)*

No. 24, *Guidance Document on Acute Oral Toxicity Testing(2001)*

No. 25, *Detailed Review Document on Hazard Classification Systems for Specifics Target Organ Systemic Toxicity Repeated Exposure in OECD Member Countries (2001)*

No. 26, *Revised Analysis of Responses Received from Member Countries to the Questionnaire on Regulatory Acute Toxicity Data Needs (2001)*

No 27, *Guidance Document On The Use Of The Harmonised System For The Classification Of Chemicals Which Are Hazardous For The Aquatic Environment (2001)*

No 28, *Guidance Document for the Conduct of Skin Absorption Studies (in preparation)*

No 29, *Draft Guidance Document on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media (2001)*

No 30, *Detailed Review Document on Hazard Classification Systems for Mixtures (2001)*

No 31, *Detailed Review Paper on Non-Genotoxic Carcinogens Detection: The Performance of In-Vitro Cell Transformation Assays(draft)*

No. 32, *Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies (2000)*

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 29 industrialised countries in North America, Europe and the Pacific, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised Committees and subsidiary groups composed of Member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's Workshops and other meetings. Committees and subsidiary groups are served by the OECD Secretariat, located in Paris, France, which is organised into Directorates and Divisions.

The work of the OECD related to chemical safety is carried out in the **Environment, Health and Safety Programme**. As part of its work on chemical testing, the OECD has issued several Council Decisions and Recommendations (the former legally binding on Member countries), as well as numerous Guidance Documents and technical reports. The best known of these publications, the **OECD Test Guidelines**, is a collection of methods used to assess the hazards of chemicals and of chemical preparations. These methods cover tests for physical and chemical properties, effects on human health and wildlife, and accumulation and degradation in the environment. The OECD Test Guidelines are recognised world-wide as the standard reference tool for chemical testing.

More information about the Environment, Health and Safety Programme and its publications (including the Test Guidelines) is available on the OECD's World Wide Web site (see page 8).

The Environment, Health and Safety Programme co-operates closely with other international organisations. This document was produced within the framework of the Inter-Organisation Programme for the Sound Management of Chemicals (IOMC).

The Inter-Organization Programme for the Sound Management of Chemicals (IOMC) was established in 1995 by UNEP, ILO, FAO, WHO, UNIDO and the OECD (the Participating Organisations), following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. UNITAR joined the IOMC in 1997 to become the seventh Participating Organisation. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

TABLE OF CONTENTS

PART 1:

GENERAL INTRODUCTION TO THE HARMONISED INTEGRATED HAZARD CLASSIFICATION SYSTEM

Chapter 1.1:
INTRODUCTION

1. The production and use of chemicals is fundamental in the economic development of all countries and, at the same time, it may pose a risk to the health and well-being of all people and the environment if not managed in a responsible manner. The primary objective of hazard classification and communication systems is to provide information to protect human health and the environment.

2. One essential step leading to the safe use of chemicals is the identification of the specific hazards and the organisation of that information so that it can be conveyed to users of chemicals in a form that is easy to understand. Measures can then be taken to avoid or manage potential risks in circumstances where exposure may occur. This is the fundamental rationale behind the hazard classification and labelling of chemicals. It has traditionally led at the national level to sector-specific regulations (transport, industry, environment, health, agriculture, consumer products, occupational health). Because of differences in use and exposure, hazard classification systems usually vary between sectors. In some cases, there is little or no consistency within sectors between different countries.

3. In 1952, the International Labor Office (ILO) began a study of the classification and labelling of dangerous substances which led in 1989 to a Resolution considering the harmonisation of systems of classification and labelling for the use of hazardous chemicals at work.

4. In 1953, the UN Economic and Social Council created the UN Committee of Experts on the Transport of Dangerous Goods (UNCETDG) charged with developing recommendations addressed to governments and international organisations concerned with the regulation of the transportation of dangerous goods; amongst other aspects, these recommendations cover the principles of classification and definitions of the categories of dangerous goods. In 1956, the UNCETDG first published its UN Recommendations on Transport of Dangerous Goods (UNRTDG) which were recently modified (1999) for the eleventh time. The UNRTDG are now included in the transport legislation of many UN states and they are used by the International Maritime Organisation (IMO), the International Civil Aviation Organisation (ICAO) and other international bodies covering transport modes. Thus land-sea-air transport is the only sector where harmonisation of hazard classification and labelling has been to a large degree achieved.

5. The UN Conference on Environment and Development (UNCED) in 1992 identified the harmonisation of classification and labelling of chemicals as one of six action programs in Chapter XIX of UNCED Agenda 21. Its objective was: “a globally harmonised hazard classification and compatible labelling system (GHS) including material safety data sheets and easily understandable symbols, should be available, if feasible, by the year 2000.” It was recognised that, while a harmonised classification system might be feasible, harmonised labelling may or may not be appropriate or possible across all sectors, but that compatibility of labelling systems might be achievable.

6. UNCED identified the International Program on Chemical Safety (IPCS) as the nucleus for international co-operation on Chapter XIX activities. Under the umbrella of IPCS a Co-ordinating Group for the Harmonisation of Chemical Classification Systems (CG/HCCS) was

established to promote and oversee the work to develop a GHS. Later, the oversight of the work of the CG/HCCS was provided by the broader Inter Organisational Programme for the Sound Management of Chemicals - IOMC. As expressed in the CG/HCCS Terms of Reference, the goals of international harmonisation are to:

- enhance the protection of people and the environment by providing an internationally comprehensible system for hazard communication;
- provide a recognised framework for those countries without an existing system;
- reduce the need for testing and evaluation of chemicals;
- facilitate international trade in chemicals whose hazards have been properly assessed and identified on an international basis.

Chapter 1.2:
ORGANISATIONAL CONTEXT FOR DEVELOPMENT OF THE GHS

7. The first priority of the CG/HCCS was the development of a harmonised classification system defining the hazards of various endpoints of concern. The Organisation for Economic Co-operation and Development (OECD) was identified as the Focal Point for work on human health and environmental hazards, ILO/UNCETDG as the Focal Point for work on physical hazards, and ILO as the Focal Point for work on Hazard Communication. The CG/HCCS would integrate the harmonised classification scheme with a harmonised hazard communication system to give an overall Globally Harmonised Classification and labelling System (GHS).

The OECD Advisory Group on Harmonisation of Classification and Labelling (AG-HCL)

8. The AG-HCL was formally established in 1994 by the Joint Meeting of the OECD Chemicals Group and Management Committee to develop proposals for a harmonised classification system for the hazards of chemicals to human health and the environment. It based its work on the initial efforts of an OECD Clearing House (1991-1993) on the Acute Human Toxicity and on the Acute Aquatic Toxicity of chemicals.

9. In its work the AG-HCL followed a set of general principles developed by the IOMC-GG/HCCS for the work on harmonisation of the hazard classification of chemicals, that specifically:

a) the level of protection offered to workers, consumers, the general public and the environment should not be reduced as a result of harmonising the classification and labelling systems;

b) the hazard classification process refers only to the hazards arising from the intrinsic properties of chemical elements and compounds, and mixtures thereof, whether natural or synthetic;

c) harmonisation means establishing a common and coherent basis for chemical hazard classification and communication, from which the appropriate elements relevant to means of transport, consumer, worker and environment protection can be selected;

d) the scope of harmonisation includes both hazard classification criteria and hazard communication tools, e.g. labelling and chemical safety data sheets;

e) changes in all existing systems will be required to achieve a single globally harmonised system; transitional measures should be included in the process of moving to the new system;

f) the involvement of concerned international organisations of employers, workers, consumers, and other relevant organisations in the process of harmonisation should be ensured;

g) the comprehension of chemical hazard information, by the target audience, e.g. workers, consumers and the general public, should be addressed;

h) test data already generated for the classification of chemicals under the existing systems, should be accepted when reclassifying these chemicals under the harmonised system;

i) a new harmonised classification system may require adaptation of existing methods for testing of chemicals;

j) in relation to chemical hazard communication and the safety and health of workers, consumers and the public in general should be ensured while protecting confidential business information, as prescribed by the competent authorities.

10. The work of the AG-HCL was generally of three related kinds:

a) Comparison of the major classification systems, identification of similar or identical elements and, for the elements which were dissimilar, development of a consensus on a compromise;

b) Examination of the scientific basis for the criteria which define the end-point of concern, gaining expert consensus on the test methods, data interpretation and level of concern, and then seeking consensus on the criteria. For some end-points, the existing schemes had no criteria and the relevant criteria were developed by the AG-HCL;

c) Where there was a decision-tree approach (e.g. irritation) or where there were dependent criteria in the classification scheme (acute aquatic toxicity), development of consensus on the process or the scheme for using the criteria.

11. The AG-HCL proceeded stepwise in developing its harmonised classification criteria. For each end-point the following steps were undertaken:

Step 1:

A thorough analysis of existing classification systems, including the scientific basis for the system and its criteria, its rationale and explanation of the mode of use. A Step 1 document was prepared for a number of endpoints, as appropriate, and amended as necessary after discussion by AG-HCL.

Step 2:

A proposal for a harmonised classification system and criteria for each category was developed. A Step 2 document was prepared and amended as necessary after discussion by AG-HCL.

Step 3:

(a) AG-HCL reached consensus on the revised Step 2 proposal; or

(b) After attempts at consensus building failed, the specific non-consensus items

were identified as alternatives in a revised Step 2 proposal.

Step 4:

Final proposal was submitted to the OECD Joint Meeting for approval and subsequently to the IOMC CG-HCCS for global implementation.

12. As experience with the use of the system is accumulated, and as new scientific information emerges, the test methods, the interpretation of the test data and the harmonised criteria *per se* may have to be updated. Thus, international work will continue to be needed in the future and, depending on the nature of the future international instrument for the implementation of the GHS, decisions will have to be made on the mechanism for carrying out the updating work in the future.

Chapter 1.3:
GENERAL CONSIDERATIONS

Scope of the Harmonised Classification System

13. The work on harmonisation of hazard classification and labelling focuses on a harmonised system for all chemicals and mixtures of chemicals. The application of the components of the system may vary by type of product or stage of the life cycle.

14. The classification system applies to pure chemical substances, their dilute solutions and to mixtures of chemical substances. However, since special considerations are needed to classify mixtures, a separate OECD Expert Group on Classification Criteria for Mixtures has addressed harmonisation in this area.

15. One objective of the harmonised hazard classification system is for it to be simple and transparent with a clear distinction between categories in order to allow for "self classification" as far as possible. For many end-points the criteria are semi-quantitative or qualitative and expert judgement is required to interpret the data for classification purposes. Furthermore, for some end-points, e.g. eye irritation, a decision tree approach is given as an example.

Presentation of Criteria

16. The current criteria for specific endpoints are presented as a series of chapters in this paper. These chapters include a number of sections all of which are relevant to classification decisions. Some chapters also have an Appendix which, unless clearly indicated to the contrary, are not part of the criteria and should be regarded as background information only. For one endpoint (hazardous for the aquatic environment) a separate Guidance Document is considered essential for a good understanding and use of the system.

Test Methods and Test Data Quality

17. The classification of a chemical substance depends both on the criteria and on the reliability of the test methods underpinning the criteria. In some cases the classification is determined by a pass or fail of a specific test, e.g. the ready biodegradation test, while in other cases, interpretations are made from dose/response curves and observations during testing. In all cases, the test conditions need to be standardised so that the results are reproducible with a given chemical substance and the standardised test yields "valid" data for defining the end-point of concern. In this context, validation is the process by which the reliability and the relevance of a procedure are established for a particular purpose.

18. Tests that determine hazardous properties which are conducted according to internationally recognised scientific principles can be used for purposes of a hazard determination for health and environmental hazards. The GHS criteria for determining health and environmental hazards should be test method neutral, allowing different approaches as long as they are scientifically sound and validated according to international procedures and criteria already referred to in existing systems for the endpoint of concern and produce mutually acceptable data.

Previously Classified Chemicals

19. One of the general principles established by the IOMC-CG-HCCS states that test data already generated for the classification of chemicals under the existing systems should be accepted when classifying these chemicals under the harmonised system thereby avoiding duplicative testing and the unnecessary use of test animals. This policy has important implications in those cases where the criteria in the GHS are different from those in an existing system. In some cases, it may be difficult to determine the quality of existing data from older studies. In such cases, expert judgement will needed.

Substances Posing Special Problems

20. The effect of a substance on biological and environmental systems is influenced, *inter alia*, by the physico chemical properties of the substance and the way in which it is biologically available. Some groups of substances present special problems in this respect, for example some polymers and metals.

Animal Welfare

21. The welfare of experimental animals is a concern. This ethical concern includes not only the alleviation of stress and suffering but also, in some countries, the use and consumption *per se* of test animals. Where possible and appropriate, tests and experiments that do not require the use of live animals are preferred to those using sentient live experimental animals. To that end, for certain end-points (skin and eye irritation/corrosion) testing schemes starting with non-animal observation/measurements are included as part of the classification system. For other endpoints such as acute toxicity, alternative animal tests, using fewer animals or causing less suffering are internationally accepted and should be preferred to the conventional LD50 test.

Evidence From Humans

22. For classification purposes, reliable epidemiological data and experience on the effects of chemicals on humans (e.g. occupational data, data from accident data bases) should be taken into account in the evaluation of human health hazards of a chemical. Testing on humans solely for hazard identification purposes is generally not acceptable.

Weight of Evidence

23. For some hazard endpoints, classification results directly when the data satisfy the criteria. For others, classification of a chemical is made on the basis of the total weight of evidence. This means that all available information bearing on the determination of toxicity is considered together, including the results of valid in vitro tests, relevant animal data, and human experience such as epidemiological and clinical studies and well-documented case reports and observations.

24. The quality and consistency of the data are important. Evaluation of substances related to the material under study should be included, as should site of action and mechanism or mode of action study results. Both positive and negative results are assembled together in a single weight of evidence determination.

25. Positive effects which are consistent with the criteria for classification in each chapter, whether seen in humans or animals, will normally justify classification. Where evidence is available from both sources and there is a conflict between the findings, the quality and reliability of the

evidence from both sources must be assessed in order to resolve the question for classification. Generally, data of good quality and reliability in humans will have precedence over other data. However, even well-designed and conducted epidemiological studies may lack sufficient numbers of subjects to detect relatively rare but still significant effects, or to assess potentially confounding factors. Positive results from well-conducted animal studies are not necessarily negated by the lack of positive human experience but require an assessment of the robustness and quality of both the human and animal data relative to the expected frequency of occurrence of effects and the impact of potentially confounding factors.

26. Route of exposure, mechanistic information and metabolism studies are pertinent to determining the relevance of an effect in humans. When such information raises doubt about relevance in humans, a lower classification may be warranted. When it is clear that the mechanism or mode of action is not relevant to humans, the substance should not be classified.

27. Both positive and negative results are assembled together in the weight of evidence determination. However, a single positive study performed according to good scientific principles and with statistically and biologically significant positive results may justify classification.

Chapter 1.4:
BUILDING BLOCK APPROACH

28. At various times during the development of harmonised classification criteria, concerns have arisen concerning the way a harmonised classification system might be used and whether it would meet the needs of its various end-users.

29. One of the consequences of the application of the classification system is expressed in the IOMC CG/HCCS General Principle (c):

> "harmonisation means establishing a common and coherent basis for chemical hazard classification and communication, from which the appropriate elements relevant to means of transport, consumer, worker and environment protection can be selected."

30. In the following chapters, sufficient sub-categories have been included under some endpoints to accommodate the fundamental needs of the existing systems. The application of the classification scheme may vary according to the circumstances, type of product and stage of the life cycle of the chemical.

31. It is essential that the cut-offs be recognised as a fundamental basis for the harmonised classification system. The use of different cut-offs for any use of the classification system would be contrary to harmonisation.

PART 2:

HARMONISED HAZARD CLASSIFICATION SYSTEM FOR CHEMICAL SUBSTANCES

Chapter 2.1:

HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICALS WHICH CAUSE ACUTE TOXICITY

PURPOSE, BASIS AND APPLICABILITY

32. The purpose of this document is to present a harmonised system of classification for acute toxicity by the oral, dermal, and inhalation routes to be used internationally.

33. The basis for the harmonised criteria are those which are currently in use in OECD countries as well as those recommended by the United National Committee of Experts on the Transport of Dangerous Goods (UNCETDG). Elements from these sources have been integrated so as a to maintain a high level of protection under a globally harmonised system of classification.

34. The classification scheme included elements that will be used by all authorities as well as other categories that will be applied only by some (e.g. transport).

CLASSIFICATION CLASSES

35. Chemicals can be allocated to one of five toxicity categories based on acute toxicity by the oral, dermal or inhalation route according to the numeric criteria expressed as (approximate) LD50 (oral, dermal) or LC50 (inhalation) values are shown in the table below. Explanatory notes are shown in italics following the table.

Table 1: Acute toxicity hazard categories and (approximate) LD50/LC50 values defining the respective categories.

	Category 1	Category 2	Category 3	Category 4	Category 5
Oral (mg/kg)	5	50	300	2000	5000 See detailed criteria
Dermal (mg/kg)	50	200	1000	2000	
Gases (ppm) see: Note a	100	500	2500	5000	
Vapours (mg/l) see: Note a Note b Note c	0.5	2.0	10	20	
Dusts and Mists (mg/l) see: Note a Note d	0.05	0.5	1.0	5	

Notes:

a: Inhalation cut-off values in the table are based on 4 hour testing exposures. Conversion of existing inhalation toxicity data which has been generated according to 1 hour exposures should be by dividing by a factor of 2 for gases and vapours and 4 for dusts and mists.

b: It is recognised that saturated vapour concentration may be used as an additional element by some regulatory systems to provide for specific health and safety protection. (e.g. UN Recommendations for the Transport of Dangerous Goods).

c: For some chemicals the test atmosphere will not just be a vapour but will consist of a mixture of liquid and vapour phases. For other chemicals the test atmosphere may consist of a vapour which is near the gaseous phase. In these latter cases, classification should be based on ppm as follows: Category 1 (100 ppm), Category 2 (500 ppm), Category 3 (2500 ppm), Category 4 (5000 ppm). Work in the OECD Test Guidelines Programme should be undertaken to better define the terms "dusts", "mists" and "vapours" in relation to inhalation toxicity testing.

d: The values for dusts and mists should be reviewed to adapt to any future changes to OECD Test Guidelines with respect to technical limitation in generating, maintaining and measuring dust and mist concentrations in respirable form.

CRITERIA FOR CATEGORY 5

36. Criteria for Category 5 are intended to enable the identification of 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 LD50 in the range of 2000-5000 mg/kg or equivalent doses for other routes.

37. The specific criteria for Category 5 are:

a) The substance is classified in this category if reliable evidence is already available that indicates the LD50 or (LC50) to be in the range of Category 5 values or other animal studies or toxic effects in humans indicate a concern for human health or an acute nature.

b) The substance is classified in this category, through extrapolation, estimation or measurement of data, if assignment to a more hazardous category 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, inhalation, or dermal routes; 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 other animal studies.

38. Recognising the need to protect animal welfare, testing in animals in Category 5 ranges is discouraged and should only be considered when there is a strong likelihood that results of such a test would have a direct relevance for protecting human health.

RATIONALE FOR THE PROPOSED SYSTEM

General considerations

39. The harmonised classification system for acute toxicity has been developed in such a way as to accommodate the needs of existing systems. A basic principle set by the IOMC CG/HCCS is that "harmonisation means establishing a common and coherent basis for chemical hazard classification and communication from which the appropriate elements relevant to means of transport, consumer, worker and environment protection can be selected." To that end, five categories have been included in the acute toxicity scheme.

40. The preferred test species for evaluation of acute toxicity by the oral and inhalation routes is the rat, while the rat or rabbit are preferred for evaluation of acute dermal toxicity. As noted by the CG/HCCS, "Test data already generated for the classification of chemicals under existing systems should be accepted when reclassifying these chemicals under the harmonised system." When experimental data for acute toxicity are available in several animal species, scientific judgement should be used in selecting the most appropriate LD50 value from among valid, well-performed tests.

41. Category 1, the highest toxicity category, has cut off values of 5 mg/kg by the oral route, 50 mg/kg by the dermal route, 100 ppm for gases or gaseous vapours, 0.5 mg/l for vapours, and 0.05 mg/l for dusts and mists. These toxicity values are currently used primarily by the transport sector for classification for packing groups.

42. Category 5 is for chemicals which are of relatively low acute toxicity but which, under certain circumstances, may pose a hazard to especially vulnerable populations. Criteria for identifying substances in Category 5 are provided in addition to the table. These substances are anticipated to have an oral or dermal LD50 value in the range 2000 - 5000 mg/kg or equivalent doses for other routes of exposure. In light of animal welfare considerations, testing in animals in Category 5 ranges is discouraged and should only be considered when there is a strong likelihood that results of such testing would have a direct relevance for protecting human health.

Special considerations for inhalation toxicity

43. Values for inhalation toxicity are based on 4 hour tests in laboratory animals. When experimental values are taken from tests using a 1 hour exposure, they can be converted to a 4 hour equivalent by dividing the 1 hour value by a factor of 2 for gases and vapours and 4 for dusts and mists.

44. Units for inhalation toxicity are a function of the form of the inhaled material. Values for dusts and mists are expressed in mg/l. Values for gases are expressed in ppm. Acknowledging the difficulties in testing vapours, some of which consist of mixtures of liquid and vapours phases, the table provides values in units of mg/l. However, for those vapours which are near the gaseous phase, classification should be based on ppm. As inhalation test methods are updated, the OECD and other test guideline programs will need to define vapours in relation to mists for greater clarity.

45. Vapour inhalation values are intended for use in classification of acute hazard for all sectors. It is also recognised that the saturated vapour concentration of a chemical is used by the transport sector as an additional element in classifying chemicals for packing groups.

46. Of particular importance is the use of well articulated values in the high toxicity categories for dusts and mists. Inhaled particles between 1 and 4 microns mean mass aerodynamic diameter

(MMAD) will deposit in all regions of the rat respiratory tract. This particle size range corresponds to a maximum dose of about 2 mg/l. In order to achieve applicability of animal experiments to human exposure, dusts and mists would ideally be tested in this range in rats. The cut off values in the table for dusts and mists allow clear distinctions to be made for materials with a wide range of toxicities measured under varying test conditions. The values for dusts and mists should be reviewed in the future to adapt to any future changes in OECD or other test guidelines with respect to technical limitations in generating, maintaining, and measuring dust and mist concentrations in respirable form.

Chapter 2.2:

HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICALS WHICH CAUSE SKIN IRRITATION/CORROSION

EXECUTIVE SUMMARY

47. From a comparison of existing dermal irritation/corrosion classification procedures currently in use, a harmonised system was formulated. It includes an evaluation strategy of existing information and specific testing for dermal effects. In developing potential harmonised positions for dermal irritation/corrosion testing, two objectives have been kept in mind: to define criteria for both corrosion and irritation classification that are in the range of sensitivity of existing systems and to have the possibility of subdividing effects into different subcategories for those authorities that need them.

48. A single category is adopted for skin corrosion. Authorities wanting to have up to three subcategories may subdivide the single corrosive category. These subcategories are modelled after those currently in use in the United Nations transport authority.

49. A single category is adopted for skin irritation. The classification procedure draws upon those currently employed by the European Union (EU). Erythema/eschar and oedema are graded separately; an animal's mean score from readings over the first three days after exposure must meet a defined level to be positive; and at least 2 of 3 tested animals must be positive for the test to be positive. Positive responses can also be obtained using other, less common criteria. The proportion of test substances expected to be positive by the proposed irritant category is within the range of positives among existing classification systems; it is somewhat higher than that of some of the current classification systems but below those of other systems. Authorities wanting to have two hazard categories can use both irritant and mild irritant categories.

PURPOSE, BASIS AND APPLICABILITY

50. The purpose of the document is to present a harmonised system of classification for skin irritation and corrosion that can be agreed upon and utilised internationally.

51. The harmonised classification system grew out of the major systems that are currently employed. It is based on concepts already in effect and does not deviate significantly from those currently in use.

52. The harmonised system for classification of skin irritation and corrosion include elements that are harmonised and will be used by all authorities as well as other categories that will be applied by only some authorities (e.g., transport, pesticides).

CLASSIFICATION CATEGORIES AND CRITERIA

53. The harmonised system includes guidance for the use of initial considerations, that is those data elements that are evaluated before animal testing for dermal corrosion and irritation is undertaken. It also includes hazard categories for corrosion and irritation.

Initial Considerations

54. Several factors should be considered in determining the corrosion and irritation potential of chemicals before testing is undertaken. Existing human experience and data including from single or repeated exposure and animal observations and data should be the first line of analysis, as it gives information directly referable to effects on the skin. In some cases enough information may be available from structurally related compounds to make classification decisions. Likewise, pH extremes like $\leq$ 2 and $\geq$ 11.5, may indicate dermal effects, especially when buffering capacity is known, although the correlation is not perfect. Generally, such agents are expected to produce significant effects on the skin. It also stands to reason that if a chemical is highly toxic by the dermal route, a dermal irritation/corrosion study may not be practicable since the amount of test substance to be applied would considerably exceed the toxic dose and, consequently, would result in the death of the animals. When observations are made of dermal irritation/corrosion in acute toxicity studies and are observed up through the limit dose, additional testing would not be needed, provided that the dilutions used and species tested are equivalent. *In vitro* alternatives that have been validated and accepted may also be used to help make classification decisions.

55. All the above information that is available on a chemical should be used in determining the need for *in vivo* dermal irritation testing. Although information might be gained from the evaluation of single parameters within a tier (e.g., caustic alkalies with extreme pH should be considered as dermal corrosives), there is merit in considering the totality of existing information and making an overall weight of evidence determination. This is especially true when there is information available on some but not all parameters. Generally, primary emphasis should be placed upon existing human experience and data, followed by animal experience and testing data, followed by other sources of information, but case-by-case determinations are necessary.

56. A tiered approach to the evaluation of initial information should be considered, where applicable (Figure 1), recognising that all elements may not be relevant in certain cases.

Corrosion

57. A single harmonised corrosion category is adopted using the results of animal testing. A corrosive is a test material that produces destruction of skin tissue, namely, visible necrosis through the epidermis and into the dermis) in $\geq$ 1 of 3 tested animals after exposure up to a 4 hour duration. 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 discern questionable lesions.

Figure 1. Tiered testing and evaluation of dermal corrosion and irritation potential (see also the "Testing and evaluation strategy for eye irritation/corrosion")

Step	Parameter	Finding	Conclusion
1a	Existing human or animal experience [g)]	→ Corrosive	→ Classify as corrosive [a)]
	↓ Not corrosive or no data		
1b	Existing human or animal experience [g)]	→ Irritant	→ Classify as irritant [a)]
	↓ Not irritant or no data		
1c	Existing human or animal experience	→ Not corrosive or irritant	→ No further testing
	↓ No data		
2a	Structure-activity relationships or structure-property relationships [b)]	→ Corrosive	→ Classify as corrosive [a)]
	↓ Not corrosive or no data		
2b	Structure-activity relationships or structure-property relationships [b)]	→ Irritant	→ Classify as irritant [a)]
	↓ Not irritating or no data		
3	pH with buffering [c)]	→ pH $\leq$ 2 or $\geq$11.5	→ Classify as corrosive [a)]
	↓ Not pH extreme or no data		
4	Existing dermal data in animals indicate no need for animal testing [d)]	→ Yes	→ Possibly no further testing may be deemed corrosive/ irritant
	↓ No indication or no data		

Figure 1. Tiered testing and evaluation of dermal corrosion and irritation potential (see also the "Testing and evaluation strategy for eye irritation/corrosion")

Step	Parameter	Finding	Conclusion
5	Valid and accepted in vitro dermal corrosion test [e)]	→ Positive response	→ Classify as corrosive [a)]
	↓ Negative response or no data		
6	Valid and accepted in vitro dermal irritation test [f)]	→ Positive response	→ Classify as irritant [a)]
	↓ Negative response or no data		
7	*In vivo* dermal corrosion test (1 animal)	→ Corrosive response	→ Classify as corrosive [a)]
	↓ Negative response		
8	*In vivo* dermal irritation test (3 animals total) [h)]	→ Irritant response	→ Classify as irritant [a)]
	↓ Negative response	→ No further testing	→ Classify as irritant [a)]
9	When it is ethical to perform human patch testing [g)]	→ Irritant response	→ Classify as irritant [a)]
	Not as above	→ Non-irritant response	→ No further testing

a. Classify in the harmonised category, below.
b. Structure-activity and structure-property relationships are presented separately but would be conducted in parallel.
c. Measurement of pH alone may be adequate, but assessment of acid or alkali reserve is preferable; methods are needed to assess buffering capacity.
d. Pre-existing animal data should be carefully reviewed to determine if in vivo dermal corrosion/irritation testing is needed. As examples, testing may not be needed when a test material has not produced any dermal irritation in an acute dermal toxicity test at the limit dose, or produces very toxic effects in an acute dermal toxicity test. In the latter case, the material would be classified as being very hazardous by the dermal route for acute toxicity; it

is moot whether the material is also irritating or corrosive on the skin. It should be kept in mind in evaluating acute dermal toxicity information that the reporting of dermal lesions may be incomplete, testing and observations may be made on a species other than the rabbit, and species may differ in sensitivity in their responses.

e. Currently there are not yet internationally accepted validated in vitro methods of dermal corrosion, but a validation study on several methods has been completed.
f. Presently there are not yet validated and internationally accepted in vitro test methods for dermal irritation.
g. This evidence could be derived from single or repeated exposures. There is no internationally accepted test method for human dermal irritation testing.
h. Testing is usually conducted in 3 animals, one coming from the negative corrosion test.

58. For those authorities wanting more than one designation of corrosivity, up to three subcategories are adopted which divide up responses in the corrosive category (Category 1, see Table 2): **subcategory 1A** --where responses are noted following up to 3 minutes exposure and up to 1 hour observation; **subcategory 1B** --where responses are described following exposure between 3 minutes and 1 hour and observations up to 14 day; and **subcategory 1C** --where responses occur after exposures between 1 hour and 4 hours and observations up to 14 days.

Table 2. Skin corrosive category and subcategories [a)]

Corrosive category (category 1)	**Potential corrosive subclasses**	**Corrosive in ≥ 1 of 3 animals**	
(applies to authorities not using subcategories)	(only applies to some authorities)	**Exposure**	**observation**
corrosive	corrosive subcategory 1A	≤ 3 minutes	≤ 1 hour
	corrosive subcategory 1B	> 3 minutes -- ≤ 1 hour	≤ 14 days
	corrosive subcategory 1C	> 1 hour -- ≤ 4 hours	≤ 14 days

a). In case human data are considered, the use of human data is discussed in Part 1, Chapter 1.3: "General Considerations".

Irritation

59. A single irritant category is adopted that (a) is centrist in sensitivity among existing classifications, (b) recognises that some test materials may lead to effects which persist throughout the length of the test, and (c) acknowledges that animal responses in a test may be quite variable. The current EU 3-animal classification system is modified to generate the proposed position. An additional mild irritant category is available for those authorities that want to have more than one dermal irritant category.

60. Reversibility of dermal lesions is another consideration in evaluating irritant responses. When inflammation persists to the end of the observation period in 2 or more test animals, taking into consideration alopecia (limited area), hyperkeratosis, hyperplasia and scaling, then a material should be considered to be an irritant.

61. Animal irritant responses within a test can be quite variable, as they are with corrosion. A separate irritant criterion should be added to accommodate cases when there is a significant irritant response but less than the mean score criterion for a positive test. For example, a test material might be designated as an irritant if 1 of 3 tested animals shows a very elevated mean score throughout the study, including lesions persisting at the end of an observation period of normally 14 days. Other responses could also fulfil this criterion. However, the responses should be ascertained as being the result of chemical exposure. Addition of this criterion increases the sensitivity of the classification system beyond that of the current EU system.

62. To counterbalance the increases in sensitivity of a designation of an irritant position and to make room for a mild irritant category, the endpoint mean score for a positive animal response is raised from $\geq$ 2.0 under the current EU system to $\geq$ 2.3. From a training set of data, the proportion of positive tests for the total data base decreases from 0.59 for the current EU system to 0.34. The exact proportion of positive test materials in the proposed system is not known, but it would definitely be higher than 0.34 and, thus, closer to the proportion of positives in the current EU system. In addition, the proportion of positives will vary considerably with the composition of materials being tested. From the training set, about 0.34 of the chemicals are in the mild irritant category, and the total is the sum of the proportion of irritants and mild irritants, or 0.68 of the chemicals.

63. A single **irritant** category (Category 2) is adopted using the results of animal testing. Authorities (e.g., pesticides) also have available a less severe **mild irritant** category (Category 3). Several criteria distinguish the two categories (Table 3). They mainly differ in the severity of dermal reactions. The major criterion for the irritant category is that at least 2 tested animals have a mean score of $\geq$ 2.3 - $\leq$ 4.0. For the mild irritant category, the mean score cut-offs are $\geq$ 1.5 - < 2.3 for at least 2 tested animals. Test materials in the irritant category would be excluded from being placed in the mild irritant category.

Table 3. Skin irritant category and subclass[a]

Classes	**Criteria**
Irritant (Category 2) (applies to all authorities)	(1) Mean value of $\geq$ 2.3 - < 4.0 for erythema/eschar or for oedema in at least 2 of 3 tested animals from gradings at 24, 48 and 72 hours after patch removal or, if reactions are delayed, from grades on 3 consecutive days after the onset of dermal reactions, or (2) Inflammation that persists to the end of the observation period normally 14 days in at least 2 animals, particularly taking into account alopecia (limited area), hyperkeratosis, hyperplasia, and scaling, or (3) In some cases where there is pronounced variability of response among animals, with very definite positive effects related to chemical exposure in a single animal but less than the criteria above.
Mild irritant (Category 3) (applies to only some authorities)	Mean value of $\geq$ 1.5 - < 2.3 for erythema/eschar or for oedema from gradings in at least 2 of 3 tested animals from grades at 24, 48 and 72 hours or, if reactions are delayed, from grades on 3 consecutive days after the onset of dermal reactions (when not included in the irritant category above).

a. In case human data are considered, the use of human data is discussed in Part 1, Chapter 1.3: "General Considerations".

Chapter 2.3:

HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICALS WHICH CAUSE EYE IRRITATION/CORROSION

EXECUTIVE SUMMARY

64. In the following harmonised system for eye irritation/corrosion hazard classification the collection of test guidelines and classification schemes worked out by the EC, the tier scheme of the U.S. regulators, the experiences of the German regulators based on the EU chemicals notification procedure and the outcome of the "OECD Workshop on Harmonisation of Validation Criteria for Alternative Tests / Harmonisation and Acceptance Criteria for Alternative Toxicological Test Methods" in Solna, Sweden (22nd -24th January, 1996) have been considered.

65. Also reflected are eye irritation/corrosion classification schemes for chemicals which are in force in the member countries of the Organisation for Economic Co-operation and Development, OECD (6), in the European Union, EU and the Canadian Pest Management Regulatory Agency and the Canadian workplace system, WHMIS. Within the transport sectors of the United Nations, UN, only dermal corrosivity is taken into account; eye corrosivity or eye irritating properties are not included within the “Orange Book” of the UN.

66. The harmonised system includes an evaluation strategy of existing information and specific testing for eye effects. In developing harmonised positions for eye irritation/corrosion testing, three objectives have been kept in mind:

- to define criteria for both serious damage to eyes and eye irritation that are in the range of sensitivity of existing systems,
- to have the option of subdividing effects in two parts for those authorities that need them, and
- to avoid testing for local effects on eyes with skin corrosive substances.

67. A single harmonised hazard group is defined for the classification of serious damage to eyes. Serious damage to eyes is defined as severe irreversible effects on the eye including not only corrosive effects like destruction of cornea or conjunctivae but also persistent indication of serious impairment of sight.

68. A single harmonised hazard group is defined for the classification of eye irritation that reverses within an appropriate observation time. The proposed harmonised classification of reversible eye irritation draws upon procedures currently employed by the European Union (EU) and by regulatory authorities in the United States of America (USA) and in Canada. Classified are local effects detected in a Draize test with rabbits that reverse within 21 days after instillation of the substance into the eye. Effects on the cornea, effects on the iris and conjunctival erythema and oedema are graded separately; an animal’s mean score from readings over the first three days after instillation must meet a defined level to be positive, and at least 2 of 3 tested animals must be positive for the test to be positive. The proportion of test substances expected to be positive by the proposed harmonised system is somewhat higher than that of the current EU system but less than that of the current US and Canadian systems. Authorities wanting to distinguish between mild and

moderate eye irritants have the option to use a subcategorisation that considers the differences within the current classification systems.

PURPOSE, BASIS AND APPLICABILITY

69. The purpose of the document is to present a harmonised system of hazard classification for eye irritation, destruction of eye tissues and other serious damage to tissues and function of eyes that can be agreed upon and utilised by OECD Member countries.

70. A tiered testing and evaluation scheme is presented that combines pre-existing information on local corrosivity and on eye irritation (including data relating to historical human or animal experience) as well as considerations on structure-activity relationships (SAR) or structure-property relationships (SPR) and the output of validated *in vitro* tests in order to avoid unnecessary animal testing.

71. The harmonised hazard classification system grew out of the currently employed systems within the OECD Member countries. It is based on concepts already in effect and melds together a position that does not deviate significantly from those currently in use.

72. The proposals for classification of eye irritation and serious damage to the eye include elements that are harmonised and will be used by all authorities as well as optional subcategories that will be applied by only some authorities (e.g., authorities classifying pesticides).

CLASSIFICATION CATEGORIES AND CRITERIA

73. The harmonised system includes guidance for the use of initial considerations, that is those data elements that are evaluated before animal testing for eye damaging effects is undertaken. It also includes hazard categories for local lesions on the eyes.

Initial considerations / tier testing and evaluation strategy

74. Before there is any *in vivo* dermal or eye irritation/corrosion testing all existing information on a test material should be reviewed. Preliminary decisions can often be made from them as to whether an agent is corrosive. If a test material can be classified, no testing is required. A highly recommended way of evaluating existing information on agents or of approaching new uninvestigated substances, is to utilise a tier testing strategy for eye irritation/corrosion.

75. Several factors should be considered in determining the eye damage or irritation potential of chemicals before testing is undertaken. Accumulated human and animal experience should be the first line of analysis, as it gives information directly referable to effects on the eye. In some cases enough information may be available from structurally related compounds to make hazard decisions. Likewise, pH extremes like ≤ 2 and ≥ 11.5, may indicate corrosive effects, especially when buffering capacity is known. Such agents are expected to produce significant effects on the eyes. Possible skin corrosion has to be evaluated prior to consideration of eye irritation/corrosion in order to avoid testing for local effects on eyes with skin corrosive substances. *In vitro* alternatives that have been validated and accepted may be used to make classification decisions.

76. All the above information that is available on a chemical should be used in determining the need for *in vivo* eye irritation testing. Although information might be gained from the evaluation of single parameters within a tier (e.g., caustic alkalies with extreme pH should be considered as local corrosives), there is merit in considering the totality of existing information and making an overall

weight of evidence determination. This is especially true when there is information available on some but not all parameters. Generally, primary emphasis should be placed upon expert judgement considering human experience with the substance, followed by the outcome of skin irritation testing and of well validated alternative methods. Animal testing with corrosive substances should be avoided whenever possible.

77. A tiered approach to the evaluation of initial information should be considered, where applicable recognising that all elements may not be relevant in certain cases. The tiered approach explained in Figure 2 was developed with contributions from (inter)national centres and committees for the testing and validation of alternatives to animal testing during a workshop in Solna, Sweden.

Figure 2: Testing and evaluation strategy for eye irritation/corrosion (see also: "Testing and evaluation strategy for skin irritation/corrosion")

Step	Parameter	Findings	Conclusions
1a	Data relating to historical human or animal experience	→ Severe damage to eyes	→ Category 1
		→ Eye irritant	Category 2
	↓ No or don't know		
1b	Data relating to historical human or animal experience	→ Skin corrosive	→ No evaluation of effects on eyes; deemed to be Category 1
	↓ No or don't know		
1c	Data relating to historical human or animal experience	→ Skin irritant	→ No evaluation of effects on eyes; deemed to be Category 2
	↓ No or don't know		
2a	SAR/SPR	→ Severe damage to eyes	→ Category 1
	↓ No or don't know		
2b	SAR/SPR	→ Eye irritant	→ No evaluation of effects on eyes; deemed to be Category 2
	↓		

Figure 2 (cont.): Testing and evaluation strategy for eye irritation/corrosion (see also: "Testing and evaluation strategy for skin irritation/corrosion")

Step	Parameter	Findings	Conclusions
	No or don't know ↓		
2c	SAR/SPR →	Skin corrosive →	No evaluation of effects on eyes; deemed to be Category 1
	↓ No or don't know ↓		
3a	pH/acid or alkaline reserve →	pH ≥ 11.5 or pH ≤ 2 (considering acid or alkaline reserve) →	Category 1
	↓		
3b	2 < pH < 11.5 (no buffering potential) ↓		
4	Other information indicating the material is a dermal corrosive →	Yes →	No evaluation of effects on eyes; deemed to be Category 1
	↓ No ↓		
5	Is a valid *in vitro* test available to assess severe damage to eyes →	No →	Go to step 6
5a	*In vitro* test for severe eye irritation →	Severe damage to eyes →	Category 1
	↓ Not a severe eye irritant ↓		
6	Is a valid *in vitro* test for eye irritation available ↓ No →	but *in vitro* test for severe eye irritancy was negative →	Go to step 8
		in the absence of any *in vitro* test →	Go to Step 7

Figure 2 (cont.): Testing and evaluation strategy for eye irritation/corrosion (see also: "Testing and evaluation strategy for skin irritation/corrosion")

Step	Parameter	Findings	Conclusions
	Yes ↓		
6a	*In vitro* eye irritation test	→ Eye irritant	→ Category 2
	↓ No indication of eye irritant properties ↓		
7	Experimentally assess skin corrosion potential (see Testing Strategy for Skin Irritation/Corrosion)	→ Skin corrosive	→ No evaluation of effects on eyes, deemed to be Category 1
	↓ Not corrosive	→ Serious damage to eyes	→ Category 1
	↓		
8	1 rabbit eye test ↓		
	No serious damage	→ Eye irritant	→ Category 2
	↓		
9	1 or 2 further rabbits Not an eye irritant		

Notes to the testing and evaluation strategy for eye irritation / corrosion

78. Step 1a/b: Data relating to historical human or animal experience: Pre-existing information on eye irritation and skin corrosion are shown separately because evaluation of skin corrosion has to be considered if there is no information on local effects on eyes. Analysis of pre-existing experience with the chemical may identify both corrosion and irritation potential for both dermal and ocular effects: i) Step 1a - reliable determination of eye irritancy basing on human or animal experience - depends on expert judgement: In most cases human experience is based on accidental events and thus, the local effects detected after an accident have to be compared with classification criteria created for evaluation of animal test data. ii) Step 1b - evaluation of data on skin corrosivity - skin corrosive substances should not be instilled into the eyes of animals; such substances should be considered as corrosive to the eyes as well. **(Category 1)**

79. Step 2a/b: SAR (Structure Activity Relationships) / SPR (Structure Property Relationships) for eye irritation and skin corrosion are shown separately but in reality would

probably be done in parallel. This stage should be completed using validated and accepted SAR/SPR approaches. The SAR/SPR analysis may identify both corrosion and irritation potential for both dermal and ocular effects: i) Step 2a - reliable determination of eye irritancy only by theoretical evaluations - in most cases it will only be appropriate for substances that are homologous to agents with very well known properties. ii) Step 2c - theoretical evaluation of skin corrosivity - skin corrosive substances should not be instilled into the eyes of animals; such substances should be considered as corrosive to the eyes as well. **(Category 1)**

80. Step 3: pH extremes like <2 and >11.5 may indicate strong local effects, especially in combination with assessment of acid or alkaline reserve, substances exhibiting such physico-chemical properties should be considered as corrosive to eyes. **(Category 1)**

81. Step 4: All attainable information should be used, including probable human experience. But this information should be restricted to that which pre-exists (e.g. the results of a dermal LD50 test or historical information on dermal corrosion).

82. Step 5: These must be alternative methods for the assessment of severe eye irritation/corrosion or serious damage to eyes (e.g., irreversible corneal opacity) which have been validated in accordance with internationally agreed principles and criteria (see "General Considerations" of the General Introduction to the Harmonised Integrated Hazard Classification System).

83. Step 6: At present this step seems not be achievable in the near future. Validated alternative methods for the reliable assessment of (reversible) eye irritation need to be worked out.

84. Step 7: In the absence of any other relevant information, it is essential to obtain this via an internationally recognised corrosion/irritation test before proceeding to a rabbit eye irritation test. This must be conducted in a staged manner. If possible, this should be achieved using a validated, accepted in vitro skin corrosivity assay. If this is not available, then the assessment should be completed using animal tests (see the skin irritation/corrosion strategy).

85. Step 8: Staged assessment of eye irritation in vivo. If in a limit test with one rabbit serious damage to eyes/severe eye irritation/corrosion is detected no further testing is needed.

86. Step 9: Only two animals may be employed for irritation testing (including the one used for evaluation of possible severe effects) if these two animals give concordant clearly irritant or clearly non-irritant responses. In the case of different or borderline responses a third animal is needed. Depending on the result of this three-animal test, classification may be required or not.

87. Where data needed for such a testing strategy cannot be required, the proposed tier testing approach demonstrates a good guidance how to organise existing information on a test material and to make a weight-of-evidence decision about hazard assessment and hazard classification - ideally without conducting new animal tests.

Irreversible effects on the eye / serious damage to eyes

88. A single harmonised hazard category is adopted for substances that have the potential to damage the eyes seriously. This hazard category - Category 1 (irreversible effects on the eye) - includes the criteria listed below. These observations include animals with grade 4 cornea lesions and other severe reactions (e.g., destruction of cornea) observed at any time during the test, as well as persistent corneal opacity, discoloration of the cornea by a dye substance, adhesion, pannus, and

interference with the function of the iris or other effects that impair sight. In this context, persistent lesions are considered those which are not fully reversible within an observation period of normally 21 days. Hazard classification: Category 1 also contains substances fulfilling the criteria of corneal opacity ≥ 3 or iritis > 1.5 detected in a Draize eye test with rabbits, because severe lesions like these usually do not reverse within a 21 days observation period.

IRREVERSIBLE EYE EFFECTS CLASSES

An eye irritant Category 1 (irreversible effects on the eye) is a test material that produces:

- at least in one animal effects on the cornea, iris or conjunctiva that are not expected to reverse or have not fully reversed within an observation period of normally 21 days

  and/or

- at least in 2 of 3 tested animals a positive response of:

  corneal opacity ≥ 3 and/or
  iritis > 1.5

  calculated as the mean scores following grading at 24, 48 and 72 hours after installation of the test material.

89. The use of human data is discussed under "General Considerations" in the introductory chapters of the Harmonised Integrated Hazard Classification System for Human Health and Environmental Effects of Chemicals.

Reversible effects on the eye

90. A single category is adopted for substances that have the potential to induce reversible eye irritation. This single hazard category provides the option to identify within the category a sub-category for substances inducing eye irritant effects reversing within an observation time of 7 days.

91. Those authorities desiring one single category for classification of "eye irritation" may use the overall harmonised Category 2 (irritating to eyes): others may want to distinguish between Category 2A (irritating to the eyes) and Category 2B (mildly irritating to eyes).

REVERSIBLE EYE EFFECTS CLASSES

An eye irritant Category 2A (irritating to eyes) is a test material that produces:

- at least in 2 of 3 tested animals a positive response of:

  corneal opacity ≥ 1 and/or
  iritis ≥ 1, and/or
  conjunctival redness≥ 2

  conjunctival oedema (chemosis) ≥ 2

  calculated as the mean scores following grading at 24, 48 and 72 hours after installation of the test material, and

- which fully reverses within an observation period of normally 21 days

Within this category an eye irritant is considered **mildly irritating to eyes (Category 2B)** when the effects listed above are fully reversible within 7 days of observation.

92. For those chemicals where there is pronounced variability among animal responses, this information may be taken into account in determining the classification.

Chapter 2.4:

HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICALS WHICH CAUSE RESPIRATORY OR SKIN SENSITISATION [1)]

PURPOSE, BASIS AND APPLICABILITY

93. The purpose of the harmonised criteria for classification of respiratory and dermal sensitisers is to give a common ground, which could be used internationally, for the hazard classification of sensitising properties of chemicals.

94. The basis for the harmonised criteria are those criteria which are currently in use in the OECD countries. Elements from these were integrated so as to maintain a high level of protection and to form harmonised criteria which could be agreed upon.

95. The criteria should be applicable on the hazard classification of chemicals irrespective of their end use.

I. RESPIRATORY SENSITISERS

Definitions

96. A respiratory sensitiser is a substance that will induce hypersensitivity of the airways following inhalation of the substance.

Classification Criteria

97. Substances shall be classified as respiratory sensitisers in accordance with the criteria given below:

- if there is evidence in humans that the substance can induce specific respiratory hypersensitivity and/or
- where there are positive results from an appropriate animal test.

1. There has been considerable discussion about what to convey about sensitisation effects to those exposed, and at what point it should be conveyed. While the current cut-off for mixtures is 1%, it appears that the major systems all believe information should be conveyed below that level. This may be appropriate both to warn those already sensitised, as well as to warn those who may become sensitised. This issue was not clear during the initial deliberations on the criteria for mixtures containing sensitisers, and thus has not been adequately discussed nor options explored.

   Before the system becomes implemented, this issue should be revisited by the ECOSOC Subcommittee on the GHS as one of its first priorities. It should be noted that the sensitisation criteria for substances will also have to be re-opened to consider this issue and the inclusion of new information and evolving testing approaches that addresses the question of strong sensitisers versus those that are weaker. Appropriate hazard communication should be considered along with the discussions on the criteria and the availability of an appropriate test method.

RATIONALE FOR THE SYSTEM

Human evidence

98. Evidence that a substance can induce specific respiratory hypersensitivity will normally be based on human experience. In this context, hypersensitivity is normally seen as asthma, but other hypersensitivity reactions such as rhinitis/conjunctivitis and alveolitis are also considered. The condition will have the clinical character of an allergic reaction. However, immunological mechanisms do not have to be demonstrated.

99. When considering the human evidence, it is necessary for a decision on classification to take into account in addition to the evidence from the cases:

- the size of the population exposed
- the extent of exposure.

100. The evidence referred to above could be

- clinical history and data from appropriate lung function tests related to exposure to the substance, confirmed by other supportive evidence which may include:
  - in vivo immunological test (e.g. skin prick test)
  - in vitro immunological test (e.g. serological analysis)
  - studies that may indicate other specific hypersensitivity reactions where immunological mechanisms of action have not been proven, e.g. repeated low-level irritation, pharmacologically mediated effects
  - a chemical structure related to substances known to cause respiratory hypersensitivity
- data from positive bronchial challenge tests with the substance conducted according to accepted guidelines for the determination of a specific hypersensitivity reaction.

101. Clinical history should include both medical and occupational history to determine a relationship between exposure to a specific substance and development of respiratory hypersensitivity. Relevant information includes aggravating factors both in the home and workplace, the onset and progress of the disease, family history and medical history of the patient in question. The medical history should also include a note of other allergic or airway disorders from childhood, and smoking history.

102. The results of positive bronchial challenge tests are considered to provide sufficient evidence for classification on their own. It is however recognised that in practice many of the examinations listed above will already have been carried out.

Animal studies

103. Data from appropriate animal studies which may be indicative of the potential of a substance to cause sensitisation by inhalation in humans may include:

- measurements of IgE and other specific immunological parameters, for example in mice
- specific pulmonary responses in guinea pigs.

EXPLANATORY NOTES

104. The mechanisms by which substances induce symptoms of asthma are not yet fully known. For preventative reasons these substances are considered as respiratory sensitisers. However, if on the basis of the evidence mentioned in paragraph 100, it can be demonstrated that these substances induce symptoms of asthma by irritation only in people with bronchial hyperreactivity, they should not be considered as respiratory sensitisers.

105. At present recognised animal models for the testing of respiratory hypersensitivity are not available. Under certain circumstances, animal testing may be used, e.g. a modification of the guinea pig maximisation test for determination of relative allergenicity of proteins. However, these tests still need further validation.

106. Some substances causing respiratory sensitisation may in addition cause immunological contact urticaria and therefore should be considered for classification as a contact sensitisers (see part II).

II. CONTACT SENSITISERS

Definitions

107. A contact sensitiser is a substance that will induce an allergic response following skin contact.

Classification Criteria

108. Substances shall be classified as contact sensitisers in accordance with the criteria given below:

- if there is evidence in humans that the substance can induce sensitisation by skin contact in a substantial number of persons, or
- where there are positive results from an appropriate animal test.

RATIONALE FOR THE SYSTEM

109. For classification of a substance evidence should include any or all of the following:

- Positive data from patch testing, normally obtained in more than one dermatology clinic.
- Epidemiological studies showing allergic contact dermatitis caused by the substance. Situations in which a high proportion of those exposed exhibit characteristic symptoms are to be looked at with special concern, even if the number of cases is small.
- Positive data from appropriate animal studies.
- Positive data from experimental studies in man. (see Part 1, Chapter 1.3, paragraph 22).

- Well documented episodes of allergic contact dermatitis, normally obtained in more than one dermatology clinic.

110. Positive effects seen in either humans or animals will normally justify classification. Evidence from animal studies is usually much more reliable than evidence from human exposure. However, in cases where evidence is available from both sources, and there is conflict between the results, the quality and reliability of the evidence from both sources must be assessed in order to resolve the question of classification on a case-by-case basis. Normally, human data are not generated in controlled experiments with volunteers for the purpose of hazard classification but rather as part of risk assessment to confirm lack of effects seen in animal tests. Consequently, positive human data on contact sensitisation are usually derived from case-control or other, less defined studies. Evaluation of human data must therefore be carried out with caution as the frequency of cases reflect, in addition to the inherent properties of the substances, factors such as the exposure situation, bioavailability, individual predisposition and preventive measures taken. Negative human data should not normally be used to negate positive results from animal studies.

111. If none of the above mentioned conditions are met the substance need not be classified as a contact sensitiser. However, a combination of two or more indicators of contact sensitisation as listed below may alter the decision. This shall be considered on a case-by-case basis.

- Isolated episodes of allergic contact dermatitis.
- Epidemiological studies of limited power, e.g. where chance, bias or confounders have not been ruled out fully with reasonable confidence.
- Data from animal tests, performed according to existing guidelines, which do not meet the criteria given in the section on animal studies but are sufficiently close to the limit to be considered significant.
- Positive data from non-standard methods.
- Positive results from close structural analogues.

EXPLANATORY NOTES

Immunological Contact Urticaria

112. Substances meeting the criteria for classification as respiratory sensitisers may in addition cause immunological contact urticaria. Consideration should be given to classify these substances also as contact sensitisers. Substances which cause immunological contact urticaria without meeting the criteria for respiratory sensitisers should also be considered for classification as contact sensitisers.

113. There is no recognised animal model available to identify substances which cause immunological contact urticaria. Therefore, classification will normally be based on human evidence which will be similar to that for skin sensitisation.

Animal Studies

114. When an adjuvant type test method for skin sensitisation is used, a response of at least 30% of the animals is considered as positive. For a non-adjuvant test method a response of at least 15% of the animals is considered positive. Test methods for skin sensitisation are described in the

OECD Guideline 406 (the Guinea Pig Maximisation test and the Buehler guinea pig test). Other methods may be used provided that they are well-validated and scientific justification is given.

115. The mouse ear swelling test, MEST, and the local lymph node assay, LLNA, appear to be reliable screening tests to detect moderate to strong sensitisers. The LLNA or the MEST can be used as a first stage in the assessment of skin sensitisation potential. In case of a positive result in either assay it may not be necessary to conduct a further guinea pig test.

116. When evaluating animal data, produced by testing according to the OECD or equivalent Guidelines for skin sensitisation, the rate of sensitised animals may be considered. This rate reflects the sensitising capacity of a substance in relation to its mildly irritating dose. This dose may vary between substances. A more appropriate evaluation of the sensitising capacity of a substance could be carried out if the dose-response relationship was known for the substance. This is an area that needs further development.

117. There are substances that are extremely sensitising at low doses where others require high doses and long time of exposure for sensitisation. For the purpose of hazard classification it may be preferable to distinguish between strong and moderate sensitisers. However, at present animal or other test systems to subcategorise sensitisers have not been validated and accepted. Therefore, subcategorisation should not yet be considered as part of the harmonised classification system. (See Background Information).

APPENDIX: BACKGROUND INFORMATION

118. Categorisation of sensitisers accounting for differences in sensitising capacity among substances would be a useful concept to develop. It may be appropriate to allocate both respiratory and dermal sensitisers to, for example, one of the following categories:

Category 1, Strong Sensitiser:

A strong sensitiser would be indicated by

- a high frequency of occurrence and/or severity of occurrence within an exposed population or
- a probability of occurrence of a high sensitisation rate in humans based on animal or other tests.

Category 2, Sensitiser:

A low to moderate sensitiser would be indicated by

- a low or moderate frequency or severity of occurrence within an exposed population or
- a probability of occurrence of a low to moderate sensitisation rate in humans based on animal or other tests.

119. Some authorities currently categorise strong sensitisers. However, at present, animal or other test systems to subcategorise sensitisers as indicated above, have not been validated and accepted. Work is going on to develop such models for the potency evaluation of contact allergens.

Chapter 2.5:

HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICALS WHICH CAUSE MUTATIONS IN GERM CELLS

PURPOSE, BASIS AND APPLICABILITY

120. The purpose of the harmonised scheme for the classification of chemicals which may cause heritable mutations in germ cells in humans is to provide a common ground which could be used internationally for the classification of mutagens. All tests conducted according to validated and internationally accepted test guidelines are acceptable for the purpose of classifying substances.

121. To arrive at that classification scheme, test results are considered from experiments determining mutagenic and/or genotoxic effects in germ and/or somatic cells of exposed animals. Mutagenic and/or genotoxic effects determined in *in vitro* tests may also be considered.

122. The system is hazard based, classifying chemicals on the basis of their intrinsic ability to induce mutations in germ cells. The scheme is, therefore, not meant for the (quantitative) risk assessment of chemical substances.

DEFINITIONS

123. The classification system is primarily concerned with chemicals which may cause mutations in the germ cells of humans and these mutations can be transmitted to the progeny. However, mutagenicity/genotoxicity tests *in vitro* and in mammalian somatic cells *in vivo* will also be considered in the sub-divisions of the classification system.

124. In the present context, commonly found definitions of the terms mutagenic, mutagen, mutations and genotoxic are used, and a mutation is defined here as a permanent change in the amount or structure of the genetic material in a cell.

125. The term “mutation” applies both for heritable genetic changes that may be manifested at the phenotypic level, and for the underlying DNA modifications when known (including, for example, specific base pair changes and chromosomal translocations). The term “mutagenic” and “mutagen” will be used for agents giving rise to an increased occurrence of mutations in populations of cells and/or organisms.

126. The more general terms “genotoxic” and “genotoxicity” apply to agents or processes which alter the structure, information content, or segregation of DNA, including those which cause DNA damage by interfering with normal replication processes, or which in a non-physiological manner (temporarily) alter its replication. Genotoxicity test results are usually taken as indicators for mutagenic effects.

CLASSIFICATION CATEGORIES AND CRITERIA

127. The classification system comprises two different categories of germ cell mutagens to accommodate the weight of evidence available. The two-category system is described in the following.

CATEGORY 1:

CHEMICALS KNOWN TO INDUCE HERITABLE MUTATIONS OR TO BE REGARDED AS IF THEY INDUCE HERITABLE MUTATIONS IN THE GERM CELLS OF HUMANS.

CATEGORY 1A: Chemicals known to induce heritable mutations in germ cells of humans

Criteria: Positive evidence from human epidemiological studies**.**

CATEGORY 1B: Chemicals which should be regarded as if they induce heritable mutations in the germ cells of humans.

Criteria:

- Positive result(s) from in vivo heritable germ cell mutagenicity tests in mammals; or
- Positive result(s) from in vivo somatic cell mutagenicity tests in mammals, in combination with some evidence that the substance has potential to cause mutations to germ cells. This supporting evidence may, for example, be derived from mutagenicity/genotoxic tests in germ cells in vivo, or by demonstrating the ability of the substance or its metabolite(s) to interact with the genetic material of germ cells; or
- Positive results from tests showing mutagenic effects in the germ cells of humans, without demonstration of transmission to progeny; for example, an increase in the frequency of aneuploidy in sperm cells of exposed people.

CATEGORY 2:

CHEMICALS WHICH CAUSE CONCERN FOR MAN OWING TO THE POSSIBILITY THAT THEY MAY INDUCE HERITABLE MUTATIONS IN THE GERM CELLS OF HUMANS.

Positive evidence obtained from experiments in mammals and/or in some cases from in vitro experiments, obtained from:

- Somatic cell mutagenicity tests in vivo, in mammals; or
- Other in vivo somatic cell genotoxicity tests which are to be supported by positive results from in vitro mutagenicity assays

Nota Bene:

- Chemicals which are positive in in vitro mammalian mutagenicity assays, and which also show chemical structure activity relationship to known germ cell mutagens, should be considered for classification as category 2 mutagens.

RATIONALE FOR THE PROPOSED SYSTEM

128. Classification for heritable effects in human germ cells is made on the basis of well conducted, sufficiently validated tests, preferably as described in OECD Test Guidelines. Evaluation of the test results should be done using expert judgement and all the available evidence should be weighed for classification.

129. Examples of *in vivo* heritable germ cell mutagenicity tests are:

Rodent dominant lethal mutation test (OECD 478)
Mouse heritable translocation assay (OECD 485)
Mouse specific locus test

130. Examples of in vivo somatic cell mutagenicity tests are:

Mammalian bone marrow micronucleus test (OECD 474)
Mammalian bone marrow chromosome aberration test (OECD 475)
Mouse spot test (OECD 484)
Mammalian erythrocyte micronucleus test (OECD 474)

131. Examples of mutagenicity/genotoxicity tests in germ cells are:

A) Mutagenicity tests:
Mammalian spermatogonial chromosome aberration test (OECD 483)
Spermatid micronucleus assay

B) Genotoxicity tests:
Sister chromatid exchange analysis in spermatogonia
Unscheduled DNA synthesis test (UDS) in testicular cells

132. Examples of genotoxicity tests in somatic cells are:

Liver Unscheduled DNA Synthesis (UDS) *in vivo* (OECD 486)
Mammalian bone marrow sister chromatid exchanges (SCE)

133. Examples of in vitro mutagenicity tests are:

In vitro mammalian chromosome aberration test (OECD 473)
In vitro mammalian cell gene mutation test (OECD 476)
Bacterial reverse mutation tests (OECD 471)

134. The classification of individual substances should be based on the total weight of evidence available, using expert judgement. In those instances where a single well-conducted test is used for classification, it should provide clear and unambiguously positive results. If new, well validated, tests arise these may also be used in the total weight of evidence to be considered. The relevance of the route of exposure used in the study of the chemical compared to the route of human exposure should also be taken into account.

EXPLANATORY NOTES

135. It becomes increasingly clear that the process of chemical-induced tumorigenesis in man and animals involves (an accumulation of) genetic changes in proto-oncogenes and/or tumour suppresser genes of somatic cells. Therefore, the demonstration of mutagenic properties of chemicals in somatic and/or germ cells of mammals *in vivo* may have implications for the potential classification of these chemicals as carcinogens (cf. chapter "Harmonised System for the Classification of Chemicals Which Cause Cancer").

Chapter 2.6:

HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICALS WHICH CAUSE CANCER

PURPOSE, BASIS AND APPLICABILITY

136. The purpose of the harmonised system for the classification of chemicals which may cause cancer is to provide common ground which could be used internationally for the classification of carcinogenic substances.

137. The scheme is applicable to the classification of all chemicals. This chapter deals only with chemical substances. The application to classification of preparations/products/mixtures is described in Chapter 3.6.

DEFINITIONS

138. The term "carcinogen" denotes a chemical substance or a mixture of chemical substances which induce cancer or increase its incidence. Substances which have induced benign and malignant tumours in well performed experimental studies on animals are considered also to be presumed or suspected human carcinogens unless there is strong evidence that the mechanism of tumour formation is not relevant for humans.

139. Classification of a chemical as posing a carcinogenic hazard is based on the inherent properties of the substance and does not provide information on the level of the human cancer risk which the use of the chemical may represent.

CLASSIFICATION CATEGORIES AND CRITERIA

140. For the purpose of classification for carcinogenicity, chemical substances are allocated to one of two categories based on strength of evidence and additional considerations (weight of evidence). In certain instances route specific classification may be warranted.

CATEGORY 1: KNOWN OR PRESUMED HUMAN CARCINOGENS

The placing of a chemical in Category 1 is done on the basis of epidemiological and/or animal data. An individual chemical may be further distinguished:

CATEGORY 1A: KNOWN to have carcinogenic potential for humans; the placing of a chemical is largely based on human evidence.

CATEGORY 1B: PRESUMED to have carcinogenic potential for humans; the placing of a chemical is largely based on animal evidence.

Based on strength of evidence together with additional considerations, such evidence may be derived from human studies that establish a causal relationship between human exposure to a chemical and the development of cancer (known human carcinogen). Alternatively, evidence may be derived from animal experiments for which there is sufficient evidence to demonstrate animal carcinogenicity (presumed human carcinogen). In addition, on a case by case basis, scientific judgement may warrant a decision of presumed human carcinogenicity derived from studies showing limited evidence of carcinogenicity in humans together with limited evidence of carcinogenicity in experimental animals.

Classification: Category 1 (A and B) Carcinogen

CATEGORY 2: SUSPECTED HUMAN CARCINOGENS

The placing of a chemical in Category 2 is done on the basis of evidence obtained from human and/or animal studies, but which is not sufficiently convincing to place the chemical in Category 1.

Based on strength of evidence together with additional considerations, such evidence may be from either limited evidence of carcinogenicity in human studies or from limited evidence of carcinogenicity in animal studies.

Classification: Category 2 Carcinogen

RATIONALE FOR THE PROPOSED SYSTEM

141. **Classification as Carcinogen** is made on the basis of evidence from reliable and acceptable methods, and is intended to be used for chemicals which have an intrinsic property to produce such toxic effects. The evaluations should be based on all existing data, peer-reviewed published studies and additional data accepted by regulatory agencies.

142. **Carcinogen classification** is a one-step, criterion-based process that involves two interrelated determinations: evaluations of strength of evidence and consideration of all other relevant information to place chemicals with human cancer potential into hazard categories.

143. **Strength of evidence** involves the enumeration of tumours in human and animal studies and determination of their level of statistical significance. Sufficient human evidence demonstrates

causality between human exposure and the development of cancer, whereas sufficient evidence in animals shows a causal relationship between the agent and an increased incidence of tumours. Limited evidence in humans is demonstrated by a positive association between exposure and cancer, but a causal relationship cannot be stated. Limited evidence in animals is provided when data suggest a carcinogenic effect, but are less than sufficient. The terms "sufficient" and "limited" are used here as they have been defined by the International Agency for Research on Cancer (IARC) and are cited in the Background Information for this document.

144. **Additional considerations** (weight of evidence). Beyond the determination of the strength of evidence for carcinogenicity, a number of other factors should be considered that influence the overall likelihood that an agent may pose a carcinogenic hazard in humans. The full list of factors that influence this determination is very lengthy, but some of the important ones are considered here.

145. The factors can be viewed as either increasing or decreasing the level of concern for human carcinogenicity. The relative emphasis accorded to each factor depends upon the amount and coherence of evidence bearing on each. Generally there is a requirement for more complete information to decrease than to increase the level of concern. Additional considerations should be used in evaluating the tumour findings and the other factors in a case-by-case manner.

146. Some important factors which may be taken into consideration, when assessing the overall level of concern are:

- Tumour type and background incidence.
- Multisite responses.
- Progression of lesions to malignancy.
- Reduced tumour latency.

Additional factors on which the evaluation may increase or decrease the level of concern include:

- Whether responses are in single or both sexes.
- Whether responses are in a single species or several species.
- Structural similarity or not to a chemical(s) for which there is good evidence of carcinogenicity.
- Routes of exposure.
- Comparison of absorption, distribution, metabolism and excretion between test animals and humans.
- The possibility of a confounding effect of excessive toxicity at test doses.
- Mode of action and its relevance for humans, such as mutagenicity, cytotoxicity with growth stimulation, mitogenesis, immunosuppression.

147. **Mutagenicity.** It is recognised that genetic events are central in the overall process of cancer development. Therefore evidence of mutagenic activity *in vivo* may indicate that a chemical has a potential for carcinogenic effects.

EXPLANATORY NOTES

148. The following additional considerations apply to classification of chemicals into either Category 1 or Category 2. A chemical that has not been tested for carcinogenicity may in certain instances be classified in Category 1 or Category 2 based on tumour data from a structural analogue

together with substantial support from consideration of other important factors such as formation of common significant metabolites, e.g. for benzidine congener dyes.

149. The classification should take into consideration whether or not the chemical is absorbed by a given route(s); or whether there are only local tumours at the site of administration for the tested route(s), and adequate testing by other major route(s) show lack of carcinogenicity.

150. It is important that whatever is known of the physico-chemical, toxicokinetic and toxicodynamic properties of the substances, as well as any available relevant information on chemical analogues, i.e. structure activity relationship, is taken into consideration when undertaking classification.

151. It is realised that some regulatory authorities may need flexibility beyond that developed in the hazard classification scheme. For inclusion into Safety Data Sheets positive results in any carcinogenicity study performed according to good scientific principles with statistically significant results may be considered.

152. Guidance on the importance of the different factors mentioned in paragraph 146 has to be elaborated in order to indicate their effects or level of concern.

153. The relative hazard potential of a chemical is a function of its intrinsic potency. There is great variability in potency among chemicals, and it may be important to account for these potency differences. The work that remains to be done is to examine methods for potency estimation. Carcinogenic potency as used here does not preclude risk assessment. (See Background Information below).

154. The proceedings of the recent WHO/IPCS working group to harmonise risk assessment for carcinogenicity points to a number of scientific questions arising for classification of chemicals e.g. mouse liver tumours, peroxisome proliferation, receptor-mediated reactions, chemicals which are carcinogenic only at toxic doses and which do not demonstrate mutagenicity. Accordingly, there is a need to articulate the principles necessary to resolve these scientific issues which have led to diverging classifications in the past. Once these issues are resolved, there would be a firm foundation for classification of a number of chemical carcinogens.

155. Data already generated for classifying chemicals under existing systems should be acceptable when reviewing these chemicals with regard to classification under the harmonised system. Further testing should not (normally) be necessary.

APPENDIX : BACKGROUND INFORMATION

I. Evaluation of the Strength of Evidence for Carcinogenicity Arising from Human and Experimental Data Adopted by the International Agency for Research on Cancer (IARC)

Carcinogenicity in humans

156. The evidence relevant to carcinogenicity from studies in humans is classified into one of the following categories:

- **Sufficient evidence of carcinogenicity:** The Working Group considers that a causal relationship has been established between exposure to the agent, mixture or exposure circumstance and human cancer. That is, a positive relationship has been observed between exposure and cancer in studies in which chance, bias and confounding could be ruled out with reasonable confidence.

- **Limited evidence of carcinogenicity:** A positive association has been observed between exposure to the agent, mixture or exposure circumstance and cancer for which a causal interpretation is considered by the Working Group to be credible, but chance, bias or confounding could not be ruled out with reasonable confidence.

157. In some instances the above categories may be used to classify the degree of evidence related to carcinogenicity in specific organs or tissues.

Carcinogenicity in experimental animals

158. The evidence relevant to carcinogenicity in experimental animals is classified into one of the following categories:

- **Sufficient evidence of carcinogenicity:** The Working Group considers that a causal relationship has been established between the agent and an increased incidence of malignant neoplasms or of an appropriate combination of benign and malignant neoplasms in (a) two or more species of animals or (b) in two or more independent studies in one species carried out at different times or in different laboratories or under different protocols.
- Exceptionally, a single study in one species might be considered to provide sufficient evidence of carcinogenicity when malignant neoplasms occur to an unusual degree with regard to incidence, site, type of tumour or age at onset.
- **Limited evidence of carcinogenicity:** The data suggest a carcinogenic effect but are limited for making a definitive evaluation because, e.g., (a) the evidence of carcinogenicity is restricted to a single experiment; or (b) there are unresolved questions regarding the adequacy of the design, conduct or interpretation of the study; or (c) the agent or mixture increases the incidence only of benign neoplasms or lesions of uncertain neoplastic potential, or of certain neoplasms which may occur spontaneously in high incidences in certain strains.

II. Considerations of Potency for Labelling Limits

159. The considerations as laid out below were excerpted from the Report of the Meeting of the Working Group on Harmonisation of Classification and Labelling of Carcinogens, Washington, DC, 17-18 October 1995.

Purpose

160. The purpose of establishing a potency scheme to be used for labelling of substances, preparations (mixtures) and contaminants is to provide for practical minimum levels of carcinogens in substances for which labelling would be required. It will result in labelling highly potent materials more strictly and less potent materials less strictly. A further purpose is to eliminate unnecessary labelling. In addition, use of a potency scheme may encourage risk reduction through purification of chemical substances or reformulating preparations.

Background

161. A large number of chemicals have been classified as carcinogenic and placed into various categories for labelling or other regulatory purpose. Chemicals that have been identified as carcinogenic may also occur as components of preparations (mixtures), impurities or additives. Gold and co-authors (Environ Health Perspect 79: 259, 1989) calculated doses from animal testing which result in tumours in half the dosed animals (TD50 values span a range of more than eight

orders of magnitude). Most classification systems do not take into account the wide range of potencies of these chemicals.

162. Carcinogens are in some countries divided into three potency groups: high, medium and low. Potency is in these instances determined using dose-response data in the observed dosing range for laboratory animals. Additional indicators of potency such as tumour site and species specificity, or species differences in toxicokinetics may also be used. Such potency groups are used to set upper limits for the classification of substances as carcinogens and for the purpose of initiating labelling. They have also been used for the classification and determination of labelling provisions for preparations (mixtures) of carcinogenic chemicals.

163. Some countries have implemented a scheme where 0.1% is used as a default limit value for labelling of substances and preparations (mixtures) as carcinogens with sufficient data for carcinogenicity. In these countries chemicals with medium carcinogenic potency are labelled if they occur in chemical substances at or above this level. Many carcinogenic compounds fall into the medium range. Carcinogens with high potency might be classified and labelled at lower levels and carcinogens with low potency could be classified and labelled only when they occur at higher levels. Some countries use 1% as a default limit value for low potency carcinogens and for carcinogens with more limited data.

164. Some regulatory authorities do not have the obligation to perform potency determinations. If a chemical carcinogen is a candidate for a potency rating outside of the default range, such chemicals should be referred to an international group for its determination.

Observations

165. The Working Group agreed that it would be useful to explore further the concept of using potency to make labelling decisions. Initial thoughts of the Working Group are presented here.

166. Potency ranking of carcinogens should not be determined or refined more precisely than by ten-fold factors in light of differences in species response, tumour types and the limits of standardisation of test protocols. In light of these points, a scheme for classification and labelling purposes which separates carcinogens into potency groupings serves the practical purposes listed above.

167. The use of potency for establishing limits does not preclude the ability of authorities to perform quantitative risk assessments of exposures to carcinogenic substances for regulatory purposes.

168. Potency determinations should be based on well performed studies which are peer reviewed, performed according to good laboratory practices, or are deemed acceptable by regulatory authorities.

Chapter 2.7:

HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICALS WHICH CAUSE REPRODUCTIVE TOXICITY

PURPOSE, BASIS, AND APPLICABILITY

169. The purpose of the harmonised system for the classification of chemicals which may cause an adverse effect on reproduction in humans is to provide a common ground which could be used internationally for the classification of reproductive toxicants.

170. The system is hazard based, classifying chemicals on the basis of intrinsic ability to produce an adverse effect on reproductive function or capacity, and/or on development of the offspring. The present system involves consideration of any substance-related adverse effect on reproduction seen in humans, or observed in appropriate tests conducted in experimental animals.

171. The Explanatory Notes provide essential guidance and should be regarded as an integral part of the Classification System.

REPRODUCTIVE TOXICITY: DEFINITIONS

172. Reproductive toxicity includes adverse effects on sexual function and fertility in adult males and females, as well as developmental toxicity in the offspring. The definitions presented below are adapted from those agreed at the IPCS/OECD Workshop for the Harmonisation of Risk Assessment for Reproductive and Developmental Toxicity, Carshalton, UK, 17-21 October, 1994. For classification purposes, the known induction of genetically-based inheritable effects in the offspring is addressed elsewhere, since in the present classification system it is considered more appropriate to address such effects under the separate end-point of germ-cell mutagenicity.

173. In this classification system, reproductive toxicity is subdivided under two main headings:

a) Adverse effects on reproductive ability or capacity

174. Any effect of chemicals that would interfere with reproductive ability or capacity. This may include, but not be limited to, alterations to the female and male reproductive system, adverse effects on onset of puberty, gamete production and transport, reproductive cycle normality, sexual behaviour, fertility, parturition, premature reproductive senescence, or modifications in other functions that are dependent on the integrity of the reproductive systems.

175. Adverse effects on or via lactation can also be included in reproductive toxicity, but for classification purposes, such effects are treated separately (see paragraph 183). This is because it is desirable to be able to classify chemicals specifically for adverse effect on lactation so that a specific hazard warning about this effect can be provided for lactating mothers.

b) Adverse effects on development of the offspring

176. Taken in its widest sense, developmental toxicity includes any effect which interferes with normal development of the conceptus, either before or after birth, and resulting from exposure of either parent prior to conception, or exposure of the developing offspring during prenatal development, or postnatally, to the time of sexual maturation.

177. However, it is considered that classification under the heading of developmental toxicity is primarily intended to provide hazard warning for pregnant women and men and women of reproductive capacity. Therefore, for pragmatic purposes of classification, developmental toxicity essentially means adverse effects induced during pregnancy, or as a result of parental exposure. These effects can be manifested at any point in the life span of the organism. The major manifestations of developmental toxicity include (1) death of the developing organism, (2) structural abnormality, (3) altered growth, and (4) functional deficiency.

CLASSIFICATION

Weight of Evidence

178. Classification as a reproductive toxicant is made on the basis of an assessment of the total weight of evidence. This means that all available information that bears on the determination of reproductive toxicity is considered together. Included are such information as epidemiological studies and case reports in humans and specific reproduction studies along with sub-chronic, chronic and special study results in animals that provide relevant information regarding toxicity to reproductive and related endocrine organs. Evaluation of substances chemically related to the material under study may also be included, particularly when information on the material is scarce. The weight given to the available evidence will be influenced by factors such as the quality of the studies, consistency of results, nature and severity of effects, level of statistical significance for intergroup differences, number of endpoints affected, relevance of route of administration to humans and freedom from bias. Both positive and negative results are assembled together into a weight of evidence determination. However, a single, positive study performed according to good scientific principles and with statistically or biologically significant positive results may justify classification (see also paragraph 180).

179. Toxicokinetic studies in animals and humans, site of action and mechanism or mode of action study results may provide relevant information, which could reduce or increase concerns about the hazard to human health. If it can be conclusively demonstrated that the clearly identified mechanism or mode of action has no relevance for humans or when the toxicokinetic differences are so marked that it is certain that the hazardous property will not be expressed in humans then a substance which produces an adverse effect on reproduction in experimental animals should not be classified.

180. In some reproductive toxicity studies in experimental animals the only effects recorded may be considered of low or minimal toxicological significance and classification may not necessarily be the outcome. These include for example small changes in semen parameters or in the incidence of spontaneous defects in the foetus, small changes in the proportions of common foetal variants such as are observed in skeletal examinations, or in foetal weights, or small differences in postnatal developmental assessments.

181. Data from animal studies ideally should provide clear evidence of specific reproductive toxicity in the absence of other, systemic, toxic effects. However, if developmental toxicity occurs together with other toxic effects in the dam, the potential influence of the generalised adverse effects should be assessed to the extent possible. The preferred approach is to consider adverse effects in the embryo/foetus first, and then evaluate maternal toxicity, along with any other factors which are likely to have influenced these effects, as part of the weight of evidence. In general, developmental effects that are observed at maternal toxic doses should not be automatically discounted. Discounting developmental effects that are observed at maternal toxic doses can only be done on a case-by-case basis when a causal relationship is established or refuted.

182. If appropriate information is available it is important to try to determine whether developmental toxicity is due to a specific maternally mediated mechanism or to a non-specific secondary mechanism, like maternal stress and the disruption of homeostasis. Generally, the presence of maternal toxicity should not be used to negate findings of embryo/foetal effects, unless it can be clearly demonstrated that the effects are secondary non-specific effects. This is especially the case when the effects in the offspring are significant, e.g. irreversible effects such as structural malformations. In some situations it is reasonable to assume that reproductive toxicity is due to a secondary consequence of maternal toxicity and discount the effects, for example if the chemical is so toxic that dams fail to thrive and there is severe inanition; they are incapable of nursing pups; or they are prostrate or dying.

Hazard classes

183. For the purpose of classification for reproductive toxicity, chemical substances are allocated to one of two categories. Effects on reproductive ability or capacity, and on development, are considered as separate issues.

CATEGORY 1:

KNOWN OR PRESUMED HUMAN REPRODUCTIVE OR DEVELOPMENTAL TOXICANT

This Category includes substances which are known to have produced an adverse effect on reproductive ability or capacity or on development in humans or for which there is evidence from animal studies, possibly supplemented with other information, to provide a strong presumption that the substance has the capacity to interfere with reproduction in humans. For regulatory purposes, a substance can be further distinguished on the basis of whether the evidence for classification is primarily from human data (Category 1A) or from animal data (Category 1B).

CATEGORY 1A: KNOWN to have produced an adverse effect on reproductive ability or capacity or on development in humans. The placing of the substance in this category is largely based on evidence from humans.

CATEGORY 1B: PRESUMED to produce an adverse effect on reproductive ability or capacity or on development in humans. The placing of the substance in this category is largely based on evidence from experimental animals. Data from animal studies should provide clear evidence of specific reproductive toxicity in the absence of other toxic effects, or if occurring together with other toxic effects the adverse effect on reproduction is considered not to be a secondary non-specific consequence of other toxic effects. However, when there is mechanistic information that raises doubt about

the relevance of the effect for humans, classification in Category 2 may be more appropriate.

CATEGORY 2:

SUSPECTED HUMAN REPRODUCTIVE OR DEVELOPMENTAL TOXICANT

This Category includes substances for which there is some evidence from humans or experimental animals, - possibly supplemented with other information - of an adverse effect on reproductive ability or capacity, or on development, in the absence of other toxic effects, or if occurring together with other toxic effects the adverse effect on reproduction is considered not to be a secondary non-specific consequence of the other toxic effects, and where the evidence is not sufficiently convincing to place the substance in Category 1. For instance, deficiencies in the study may make the quality of evidence less convincing, and in view of this Category 2 could be the more appropriate classification.

EFFECTS ON OR VIA LACTATION

Effects on or via lactation are allocated to a separate single category. It is appreciated that for many substances there is no information on the potential to cause adverse effects on the offspring via lactation. However, for substances which are absorbed by women and have been shown to interfere with lactation or which may be present (including metabolites) in breast milk in amounts sufficient to cause concern for the health of a breastfed child, should be classified to indicate this property hazardous to breastfed babies. This classification can be assigned on the basis of:

(a) absorption, metabolism, distribution and excretion studies that would indicate the likelihood the substance would be present in potentially toxic levels in breast milk; and/or

(b) results of one or two generation studies in animals which provide clear evidence of adverse effect in the offspring due to transfer in the milk or adverse effect on the quality of the milk; and/or

(c) human evidence indicating a hazard to babies during the lactation period.

BASIS OF CLASSIFICATION

184. Classification is made on the basis of the appropriate criteria, outlined above, and an assessment of the total weight of evidence. Classification as a reproductive or developmental toxicant is intended to be used for chemicals which have an intrinsic, specific property to produce an adverse effect on reproduction or development and chemicals should not be so classified if such an effect is produced solely as a non-specific secondary consequence of other toxic effects.

185. In the evaluation of toxic effects on the developing offspring, it is important to consider the possible influence of maternal toxicity.

186. For human evidence to provide the primary basis for a Category 1A classification there must be reliable evidence of adverse effect on reproduction in humans. Evidence used for classification should ideally be from well conducted epidemiological studies which include the use of appropriate controls, balanced assessment, and due consideration of bias or confounding factors. Less rigorous data from studies in humans should be supplemented with adequate data from studies in experimental animals and classification in Category 1B should be considered.

187. Data already generated for classifying chemicals under existing systems should be acceptable when reviewing these chemicals with regard to classification under the harmonised system. Further testing should not normally be necessary.

EXPLANATORY NOTES

Maternal toxicity

188. Development of the offspring throughout gestation and during the early post-natal stages can be influenced by toxic effects in the mother either through non-specific mechanisms related to stress and the disruption of maternal homeostasis, or by specific maternally-mediated mechanisms. So, in the interpretation of the developmental outcome to decide classification for developmental effects it is important to consider the possible influence of maternal toxicity. This is a complex issue because of uncertainties surrounding the relationship between maternal toxicity and developmental outcome. Expert judgement and a weight of evidence approach, using all available studies, should be used to determine the degree of influence that should be attributed to maternal toxicity when interpreting the criteria for classification for developmental effects. The adverse effects in the embryo/foetus should be first considered, and then maternal toxicity, along with any other factors which are likely to have influenced these effects, as weight of evidence, to help reach a conclusion about classification.

189. Based on pragmatic observation, it is believed, that maternal toxicity may, depending on severity, influence development via non-specific secondary mechanisms, producing effects such as depressed foetal weight, retarded ossification, and possibly resorptions and certain malformations in some strains of certain species. However, the limited number of studies which have investigated the relationship between developmental effects and general maternal toxicity have failed to demonstrate a consistent, reproducible relationship across species. Developmental effects which occur even in the presence of maternal toxicity are considered to be evidence of developmental toxicity, unless it can be unequivocally demonstrated on a case by case basis that the developmental effects are secondary to maternal toxicity. Moreover, classification should be considered where there is significant toxic effect in the offspring, e.g. irreversible effects such as structural malformations, embryo/foetal lethality, significant post-natal functional deficiencies.

190. Classification should not automatically be discounted for chemicals that produce developmental toxicity only in association with maternal toxicity, even if a specific maternally-mediated mechanism has been demonstrated. In such a case, classification in Category 2 may be considered more appropriate than Category 1. However, when a chemical is so toxic that maternal death or severe inanition results, or the dams are prostrate and incapable of nursing the pups, it may be reasonable to assume that developmental toxicity is produced solely as a secondary consequence of maternal toxicity and discount the developmental effects. Classification may not necessarily be the outcome in the case of minor developmental changes e.g. small reduction in foetal/pup body weight, retardation of ossification when seen in association with maternal toxicity.

191. Some of the end points used to assess maternal toxicity are provided below. Data on these end points, if available, needs to be evaluated in light of their statistical or biological significance and dose response relationship.

> Maternal Mortality: An increased incidence of mortality among the treated dams over the controls should be considered evidence of maternal toxicity if the increase occurs in a dose-related manner and can be attributed to the systemic toxicity of the test material.

Maternal mortality greater than 10% is considered excessive and the data for that dose level should not normally be considered for further evaluation.

Mating Index (no. animals with seminal plugs or sperm/no. mated x 100)[1]

Fertility Index (no. animals with implants/no. of matings x 100)[1]

Gestation Length (if allowed to deliver)

Body Weight and Body Weight Change: Consideration of the maternal body weight change and/or adjusted (corrected) maternal body weight should be included in the evaluation of maternal toxicity whenever such data are available. The calculation of a adjusted (corrected) mean maternal body weight change, which is the difference between the initial and terminal body weight minus the gravid uterine weight (or alternatively, the sum of the weights of the foetuses), may indicate whether the effect is maternal or intrauterine. In rabbits, the body weight gain may not be useful indicators of maternal toxicity because of normal fluctuations in body weight during pregnancy.

Food and Water Consumption (if relevant): The observation of a significant decrease in the average food or water consumption in treated dams compared to the control group may be useful in evaluating maternal toxicity, particularly when the test material is administered in the diet or drinking water. Changes in food or water consumption should be evaluated in conjunction with maternal body weights when determining if the effects noted are reflective of maternal toxicity or more simply, unpalatability of the test material in feed or water.

Clinical evaluations (including clinical signs, markers, haematology and clinical chemistry studies): The observation of increased incidence of significant clinical signs of toxicity in treated dams relative to the control group may be useful in evaluating maternal toxicity. If this is to be used as the basis for the assessment of maternal toxicity, the types, incidence, degree and duration of clinical signs should be reported in the study. Examples of frank clinical signs of maternal intoxication include: coma, prostration, hyperactivity, loss of righting reflex, ataxia, or laboured breathing.

Post-mortem data: Increased incidence and/or severity of post-mortem findings may be indicative of maternal toxicity. This can include gross or microscopic pathological findings or organ weight data, e.g., absolute organ weight, organ-to-body weight ratio, or organ-to-brain weight ratio. When supported by findings of adverse histopathological effects in the affected organ(s), the observation of a significant change in the average weight of suspected target organ(s) of treated dams, compared to those in the control group, may be considered evidence of maternal toxicity.

Potency and cut-off doses

192. In the present scheme, the relative potency of a chemical to produce a toxic effect on reproduction is not included in the criteria for reaching a conclusion regarding classification. Nevertheless, during the development of this scheme it was suggested that cut-off dose levels should be included, in order to provide some means of assessing and categorising the potency of chemicals for the ability to produce an adverse effect on reproduction. This concept has not been readily accepted by all member countries because of concerns that any specified cut-off level may be exceeded by human exposure levels in certain situations, e.g. inhalation of volatile solvents, the

1 . It is recognised that this index can also be affected by the male.

level may be inadequate in cases where humans are more sensitive than the animal model, and because of disagreements about whether or not potency is a component of hazard.

193. There has been interest in this concept to further consider it as a future development of the classification scheme.

Limit dose

194. Member countries appear to be in agreement about the concept of a limit dose, above which the production of an adverse effect may be considered to be outside the criteria which lead to classification. However, there is disagreement between members regarding the inclusion within the criteria of a specified dose as a limit dose. Some Test Guidelines specify a limit dose, other Test Guidelines qualify the limit dose with a statement that higher doses may be necessary if anticipated human exposure is sufficiently high that an adequate margin of exposure would not be achieved. Also, due to species differences in toxicokinetics, establishing a specific limit dose may not be adequate for situations where humans are more sensitive than the animal model.

195. In principle, adverse effects on reproduction seen only at very high dose levels in animal studies (for example doses that induce prostration, severe inappetence, excessive mortality) would not normally lead to classification, unless other information is available, e.g. toxicokinetics information indicating that humans may be more susceptible than animals, to suggest that classification is appropriate. Please also refer to the section on Maternal Toxicity for further guidance in this area.

196. However, specification of the actual ‘limit dose' will depend upon the test method that has been employed to provide the test results, e.g. in the OECD Test Guideline for repeated dose toxicity studies by the oral route, an upper dose of 1000 mg/kg unless expected human response indicates the need for a higher dose level, has been recommended as a limit dose.

Animal and experimental data

197. A number of internationally accepted test methods are available; these include methods for developmental toxicity testing (e.g., OECD Test Guideline 414, ICH Guideline S5A, 1993), methods for peri- and post-natal toxicity testing (e.g. ICH S5B, 1995) and methods for one or two-generation toxicity testing (e.g. OECD Test Guidelines 415, 416).

198. Results obtained from Screening Tests (e.g. OECD Guidelines 421 - Reproduction/Developmental Toxicity Screening Test, and 422 - Combined Repeated Dose Toxicity Study with Reproduction/Development Toxicity Screening Test) can also be used to justify classification, although it is recognised that the quality of this evidence is less reliable than that obtained from full studies.

199. Adverse effects or changes, seen in short- or long-term repeated dose toxicity studies, which are judged likely to impair reproductive ability or capacity and which occur in the absence of significant generalised toxicity, may be used as a basis for classification, e.g. histopathological changes in the gonads.

200. Evidence from in vitro assays, or non-mammalian tests, and from analogous substances using structure-activity relationship (SAR), can contribute to the procedure for classification. In all cases of this nature, expert judgement must be used to assess the adequacy of the data. Inadequate data should not be used as a primary support for classification.

201. It is preferable that animal studies are conducted using appropriate routes of administration which relate to the potential route of human exposure. However, in practice, reproductive toxicity studies are commonly conducted using the oral route, and such studies will normally be suitable for evaluating the hazardous properties of the substance with respect to reproductive toxicity. However, if it can be conclusively demonstrated that the clearly identified mechanism or mode of action has no relevance for humans or when the toxicokinetic differences are so marked that it is certain that the hazardous property will not be expressed in humans then a substance which produces an adverse effect on reproduction in experimental animals should not be classified.

202. Studies involving routes of administration such as intravenous or intraperitoneal injection, which may result in exposure of the reproductive organs to unrealistically high levels of the test substance, or elicit local damage to the reproductive organs, e.g. by irritation, must be interpreted with extreme caution and on their own would not normally be the basis for classification.

Chapter 2.8:

HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICALS WHICH CAUSE SPECIFIC TARGET ORGAN ORIENTED SYSTEMIC TOXICITY FOLLOWING A SINGLE EXPOSURE

PURPOSE, BASIS AND APPLICABILITY

203. The purpose of this document is to provide a means of classifying substances that produce specific, non lethal target organ/systemic toxicity arising from a single exposure. All significant health effects that can impair function, both reversible and irreversible, immediate and/or delayed are included.

204. Specific target organ/systemic toxicity following a repeated exposure is classified elsewhere in the GHS as a separate chapter, and therefore, is excluded from the present chapter. Other specific toxic effects, such as acute lethality/toxicity, eye and skin corrosivity/irritation, skin and respiratory sensitisation, carcinogenicity, mutagenicity and reproductive toxicity are assessed separately in the GHS and consequently are not included here.

205. Specific target organ/systemic toxicity can occur by any route that is relevant for humans, i.e., principally oral, dermal or inhalation.

DEFINITIONS

206. Classification identifies the chemical substance as being a specific target organ/systemic toxicant and, as such, it may present a potential for adverse health impact to people who are exposed to it.

207. Classification depends upon the availability of reliable evidence that a single exposure to the substance has produced a consistent and identifiable toxic effect in humans, or, in experimental animals, toxicologically significant changes which have affected the function or morphology of a tissue/organ, or has produced serious changes to the biochemistry or haematology of the organism and these changes are relevant for human health. It is recognised that human data will be the primary source of evidence for this end point.

208. Assessment should take into consideration not only significant changes in a single organ or biological system but also generalised changes of a less severe nature involving several organs.

CLASSIFICATION

209. Substances are classified for immediate or delayed effects separately by the use of expert judgement on the basis of the weight of all evidence available, including the use of recommended guidance values (see paragraphs 219-223). Then substances are placed in one of two categories, depending upon the nature and severity of the effect(s) observed.

CATEGORY 1:

SUBSTANCES THAT HAVE PRODUCED SIGNIFICANT TOXICITY IN HUMANS, OR THAT, ON THE BASIS OF EVIDENCE FROM STUDIES IN EXPERIMENTAL ANIMALS CAN BE PRESUMED TO HAVE THE POTENTIAL TO PRODUCE SIGNIFICANT TOXICITY IN HUMANS FOLLOWING SINGLE EXPOSURE

Placing a substance in Category 1 is done on the basis of:

- reliable and good quality evidence from human cases or epidemiological studies;
  or,
- observations from appropriate studies in experimental animals in which significant and/or severe toxic effects of relevance to human health were produced at generally low exposure concentrations. Guidance dose/concentration values are provided below (see paragraphs 219-223) to be used as part of weight-of-evidence evaluation.

CATEGORY 2:

SUBSTANCES THAT, ON THE BASIS OF EVIDENCE FROM STUDIES IN EXPERIMENTAL ANIMALS CAN BE PRESUMED TO HAVE THE POTENTIAL TO BE HARMFUL TO HUMAN HEALTH FOLLOWING SINGLE EXPOSURE

Placing a substance in Category 2 is done on the basis of observations from appropriate studies in experimental animals in which significant toxic effects, of relevance to human health, were produced at generally moderate exposure concentrations. Guidance dose/concentration values are provided below (see paragraphs 219-223) in order to help in classification.

In exceptional cases, human evidence can also be used to place a substance in Category 2 (see paragraph 214).

For both categories the classified substance may be named for specific target organ/system that has been primarily affected, or as a general systemic toxicant. Attempts should be made to determine the primary target organ of toxicity and classify for that purpose, e.g. hepatoxicants, neurotoxicants. One should carefully evaluate the data and, where possible, not include secondary effects, e.g., a hepatotoxin can secondarily produce effects of the nervous or gastro-intestinal systems.

210. The classified substance should be named for the relevant route of exposure.

Criteria

211. Classification is determined by expert judgement, on the basis of the weight of all evidence available including the guidance presented below.

212. Weight of evidence of all data, including human incidents, epidemiology, and studies conducted in experimental animals, is used to substantiate specific target organ/systemic toxic effects that merit classification.

213. The information required to evaluate specific target organ/systemic toxicity comes either from single exposure in humans, e.g., exposure at home, in the workplace or environmentally, or from studies conducted in experimental animals. The standard animal studies in rats or mice that provide this information are acute toxicity studies which can include clinical observations and detailed macroscopic and microscopic examination to enable the toxic effects on target tissues/organs to be identified. Results of acute toxicity studies conducted in other species may also provide relevant information.

214. In exceptional cases, based on expert judgement, it may be appropriate to place certain substances with human evidence of target organ/systemic toxicity in Category 2: (1) when the weight of human evidence is not sufficiently convincing to warrant Category 1 classification, and/or (2) based on the nature and severity of effects. Dose/concentration levels in humans should not be considered in the classification and any available evidence from animal studies should be consistent with the Category 2 classification. In other words, if there are also animal data available on the chemical that warrant Category 1 classification, the chemical should be classified as Category 1.

Effects Considered To Support Classification

215. Evidence associating single exposure to the substance with a consistent and identifiable toxic effect.

216. It is recognised that evidence from human experience/incidents is usually restricted to an adverse health consequence often with uncertainty about exposure conditions, and may not provide the scientific detail that can be obtained from well-conducted studies in experimental animals.

217. Evidence from appropriate studies in experimental animals can furnish much more detail, in the form of clinical observations, and macroscopic and microscopic pathological examination - and this can often reveal hazards that may not be life-threatening but could indicate functional impairment. Consequently all available evidence, and relevance to human health, must be taken into consideration in the classification process. Examples of relevant toxic effects in humans and/or animals are provided below:

- Morbidity resulting from single exposure.
- Significant functional changes in the central or peripheral nervous systems or other organ systems, including signs of central nervous system depression and effects on special senses (e.g., sight, hearing and sense of smell).
- Any consistent and significant adverse change in clinical biochemistry, haematology, or urinalysis parameters.
- Significant organ damage that may be noted at necropsy and/or subsequently seen or confirmed at microscopic examination.
- Multifocal or diffuse necrosis, fibrosis or granuloma formation in vital organs with regenerative capacity.
- Morphological changes that are potentially reversible but provide clear evidence of marked organ dysfunction.
- Evidence of appreciable cell death (including cell degeneration and reduced cell number) in vital organs incapable of regeneration.

Effects Considered Not To Support Classification**:**

218. It is recognised that effects may be seen that would not justify classification. Examples of such effects in humans and/or animals are provided below:

- Clinical observations or small changes in bodyweight gain, food consumption or water intake that may have some toxicological importance but that do not, by themselves, indicate "significant" toxicity.

- Small changes in clinical biochemistry, haematology or urinalysis parameters and/or transient effects, when such changes or effects are of doubtful or minimal toxicological importance.
- Changes in organ weights with no evidence or organ dysfunction.
- Adaptive responses that are not considered toxicologically relevant.
- Substance-induced species-specific mechanisms of toxicity, i.e. demonstrated with reasonable certainty to be not relevant for human health, should not justify classification.
- Where there are only local effects, at the site of administration for the routes tested, and especially when adequate testing by other principal routes show lack of specific target organ/systemic toxicity.

Guidance values to assist with classification based on the results obtained from studies conducted in experimental animals

219. In order to help reach a decision about whether a substance should be classified or not, and to what degree it would be classified (Category 1 vs. Category 2), dose/concentration 'guidance values' are provided for consideration of the dose/concentration which has been shown to produce significant health effects. The principal argument for proposing such guidance values is that all chemicals are potentially toxic and there has to be a reasonable dose/concentration above which a degree of toxic effect is acknowledged.

220. Thus, in animal studies, when significant toxic effects are observed, that would indicate classification, consideration of the dose/concentration at which these effects were seen, in relation to the suggested guidance values, can provide useful information to help assess the need to classify (since the toxic effects are a consequence of the hazardous property(ies) and also the dose/concentration).

221. The guidance value ranges proposed for single-dose exposure which has produced a significant non-lethal toxic effect are those applicable to acute toxicity testing, as indicated in Table 4 below:

Table 4: Guidance value ranges for single-dose exposures

		Guidance value ranges for :	
Route of exposure	**Units**	**Category 1 classification**	**Category 2 classification**
Oral (rat)	mg/kg bw	$c \leq 300$	$2000 \geq c > 300$
Dermal (rat or rabbit)	mg/kg bw	$c \leq 1000$	$2000 \geq c > 1000$
Inhalation (rat) gas	ppm	$c \leq 2500$	$5000 \geq c > 2500$
Inhalation (rat) vapour	mg/1	$c \leq 10$	$20 \geq c > 10$
Inhalation (rat) dust/mist/fume	mg/l/4h	$c \leq 1.0$	$5.0 \geq c > 1.0$

222. It is important to recognise that the guidance values and ranges mentioned in paragraph 221 above are intended only for guidance purposes, i.e., to be used as part of the weight of evidence approach, and to assist with decision about classification. They are not intended as strict demarcation values.

223. Thus it is feasible that a specific profile of toxicity is seen to occur at a dose/concentration below the guidance value, eg. <2000 mg/kg bw by the oral route, however the nature of the effect may result in the decision not to classify. Conversely, a specific profile of toxicity may be seen in animal studies occurring at or above a guidance value, eg. ≥2000 mg/kg bw by the oral route, and in addition there is supplementary information from other sources, e.g. other single dose studies, or human case experience, which supports a conclusion that, in view of the weight of evidence, classification would be the prudent action to take.

RATIONALE FOR THE PROPOSED SYSTEM

224. When a chemical is characterised only by use of animal data (typical of new chemicals, but also true for many existing chemicals), the classification process would include reference to dose/concentration guidance values as one of the elements that contribute to the weight of evidence approach.

225. When well-substantiated human data are available showing a specific target organ/systemic toxic effect that can be reliably attributed to single exposure to a chemical substance, the substance may be classified. Positive human data, regardless of probable dose, predominates over animal data. Thus, if a chemical is unclassified because specific target organ/systemic toxicity observed was considered not relevant or significant to humans, if subsequent human incident data become available showing a specific target organ/systemic toxic effect, the substance should be classified.

226. A chemical that has not been tested for specific target organ/systemic toxicity may in certain instances, where appropriate, be classified on the basis of data from a validated structure activity relationship and expert judgement-based extrapolation from a structural analogue that has previously been classified together with substantial support from consideration of other important factors such as formation of common significant metabolites.

227. It is recognised that saturated vapour concentration may be used as an additional element by some regulatory systems to provide for specific health and safety protection.

Chapter 2.9:

HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICALS WHICH CAUSE SPECIFIC TARGET ORGAN ORIENTED SYSTEMIC TOXICITY FOLLOWING REPEATED EXPOSURE

PURPOSE, BASIS AND APPLICABILITY

228. The purpose of this document is to provide a means of classifying substances that produce specific target organ/systemic toxicity arising from repeated exposure that is not specifically addressed elsewhere in the harmonised classification system (GHS). All significant health effects that can impair function, both reversible and irreversible, following repeated or long-term exposure, are included. Other specific toxic effects, such as acute lethality/toxicity, eye and skin corrosivity/irritation, skin and respiratory sensitisation, carcinogenicity, mutagenicity and reproductive toxicity are assessed separately in the GHS and consequently are not included in this chapter.

229. Non-lethal toxic effects observed after a single-event exposure are classified elsewhere in the GHS as a separate chapter and, therefore, are excluded from the present chapter.

230. Specific target organ/systemic toxicity can occur by any route that is relevant for humans, i.e., principally oral, dermal or inhalation.

DEFINITIONS

231. Classification identifies the chemical substance as being a specific target organ/systemic toxicant and, as such, it may present a potential for adverse health impact to people who are exposed to it.

232. Classification depends upon the availability of reliable evidence that repeated exposure to the substance has produced a consistent and identifiable toxic effect in humans, or, in experimental animals, toxicologically significant changes which have affected the function or morphology of a tissue/organ, or has produced serious changes to the biochemistry or haematology of the organism and these changes are relevant for human health.

233. Assessment of specific target organ/systemic toxicity should take into consideration not only significant changes in a single organ or biological system but also generalised changes of a less severe nature involving several organs.

CLASSIFICATION

234. Substances are classified as specific target organ/systemic toxicant by expert judgement on the basis of the weight of all evidence available, including the use of recommended guidance values which take into account the duration of exposure and the dose/concentration which produced the

effect(s), (see paragraphs 244-252), and are placed in one of two categories, depending upon the nature and severity of the effect(s) observed.

CATEGORY 1:

SUBSTANCES THAT HAVE PRODUCED SIGNIFICANT TOXICITY IN HUMANS, OR THAT, ON THE BASIS OF EVIDENCE FROM STUDIES IN EXPERIMENTAL ANIMALS CAN BE PRESUMED TO HAVE THE POTENTIAL TO PRODUCE SIGNIFICANT TOXICITY IN HUMANS FOLLOWING REPEATED EXPOSURE.

Placing a substance in Category 1 is done on the basis of:

- reliable and good quality evidence from human cases or epidemiological studies; or,
- observations from appropriate studies in experimental animals in which significant and/or severe toxic effects, of relevance to human health, were produced at generally low exposure concentrations. Guidance dose/concentration values are provided below (see paragraphs 244-252) to be used as part of weight-of- evidence evaluation.

CATEGORY 2:

SUBSTANCES THAT, ON THE BASIS OF EVIDENCE FROM STUDIES IN EXPERIMENTAL ANIMALS CAN BE PRESUMED TO HAVE THE POTENTIAL TO BE HARMFUL TO HUMAN HEALTH FOLLOWING REPEATED EXPOSURE.

Placing a substance in Category 2 is done on the basis of observations from appropriate studies in experimental animals in which significant toxic effects, of relevance to human health, were produced at generally moderate exposure concentrations. Guidance dose/concentration values are provided below (see paragraphs 244-252) in order to help in classification.

In exceptional cases human evidence can also be used to place a substance in Category 2 (see paragraph 239).

The classified substance may be named for the specific target organ/system that has been primarily affected, or generally as a general systemic toxicant. Attempts should be made to determine the primary target organ of toxicity and classify for that purpose, e.g., hepatotoxicants, neurotoxicants. One should carefully evaluate the data and, where possible, not include secondary effects, e.g. hepatotoxin can secondarily produce effects of the nervous or gastro-intestinal systems.

235. The classified substance should be named for the relevant route of exposure.

Criteria

236. Classification is determined by expert judgement, on the basis of the weight of all evidence available including the guidance presented below.

237. Weight of evidence of all data, including human incidents, epidemiology, and studies conducted in experimental animals, is used to substantiate specific target organ/systemic toxic effects that merit classification. This taps the considerable body of industrial toxicology data

collected over the years. Evaluation should be based on all existing data, including peer-reviewed published studies and additional data acceptable to regulatory agencies.

238. The information required to evaluate specific target organ/systemic toxicity comes either from repeated exposure in humans, e.g., exposure at home, in the workplace or environmentally, or from studies conducted in experimental animals. The standard animal studies in rats or mice that provide this information are 28 day, 90 day or lifetime studies (up to 2 years) that include haematological, clinicochemical and detailed macroscopic and microscopic examination to enable the toxic effects on target tissues/organs to be identified. Data from repeat dose studies performed in other species may also be used. Other long-term exposure studies, eg. for carcinogenicity, neurotoxicity or reproductive toxicity, may also provide evidence of specific target organ/systemic toxicity that could be used in the assessment of classification.

239. In exceptional cases, based on expert judgement, it may be appropriate to place certain substances with human evidence of target organ/systemic toxicity in Category 2: (1) when the weight of human evidence is not sufficiently convincing to warrant Category 1 classification, and/or (2) based on the nature and severity of effects. Dose/concentration levels in humans should not be considered in the classification and any available evidence from animal studies should be consistent with the Category 2 classification. In other words, if there are also animal data available on the chemical that warrant Category 1 classification, the chemical should be classified as Class1.

Effects Considered To Support Classification**:**

240. Reliable evidence associating repeated exposure to the substance with a consistent and identifiable toxic effect.

241. It is recognised that evidence from human experience/incidents is usually restricted to an adverse health consequence, often with uncertainty about exposure conditions, and may not provide the scientific detail that can be obtained from well-conducted studies in experimental animals.

242. Evidence from appropriate studies in experimental animals can furnish much more detail, in the form of clinical observations, haematology, clinical chemistry, and macroscopic and microscopic pathological examination - and this can often reveal hazards that may not be life-threatening but could indicate functional impairment. Consequently all available evidence, and relevance to human health, must be taken into consideration in the classification process. Examples of relevant toxic effects in humans and/or animals are provided below:

- Morbidity or death resulting from repeated or long-term exposure. Morbidity or death may result from repeated exposure, even to relatively low doses/concentrations, due to bioaccumulation of the substance or its metabolites, or accumulation of effect owing to the ability of the de-toxification process becoming overwhelmed by repeated exposure to the substance or its metabolites.
- Significant functional changes in the central or peripheral nervous systems or other organ systems, including signs of central nervous system depression and effects on special senses (e.g., sight, hearing and sense of smell).
- Any consistent and significant adverse change in clinical biochemistry, haematology, or urinalysis parameters.
- Significant organ damage that may be noted at necropsy and/or subsequently seen or confirmed at microscopic examination.

- Multifocal or diffuse necrosis, fibrosis or granuloma formation in vital organs with regenerative capacity.
- Morphological changes that are potentially reversible but provide clear evidence of marked organ dysfunction (e.g., severe fatty change in the liver).
- Evidence of appreciable cell death (including cell degeneration and reduced cell number) in vital organs incapable of regeneration.

Effects Considered Not To Support Classification:

243. It is recognised that effects may be seen that would not justify classification. Examples of such effects in humans and/or animals are provided below:

- Clinical observations or small changes in bodyweight gain, food consumption or water intake that may have some toxicological importance but that do not, by themselves, indicate “significant" toxicity.
- Small changes in clinical biochemistry, haematology or urinalysis parameters and /or transient effects, when such changes or effects are of doubtful or minimal toxicological importance.
- Changes in organ weights with no evidence or organ dysfunction.
- Adaptive responses that are not considered toxicologically relevant.
- Substance-induced species-specific mechanisms of toxicity, i.e. demonstrated with reasonable certainty to be not relevant for human health, should not justify classification.

Guidance values to assist with classification based on the results obtained from studies conducted in experimental animals

244. In studies conducted in experimental animals, reliance on observation of effects alone, without reference to the duration of experimental exposure and dose/concentration, omits a fundamental concept of toxicology, i.e., all substances are potentially toxic, and what determines the toxicity is a function of the dose/concentration and the duration of exposure. In most studies conducted in experimental animals the test guidelines use an upper limit dose value.

245. In order to help reach a decision about whether a substance should be classified or not, and to what degree it would be classified (Category 1 vs. Category 2), dose/concentration ‘guidance values’ are provided for consideration of the dose/concentration which has been shown to produce significant health effects. The principal argument for proposing such guidance values is that all chemicals are potentially toxic and there has to be a reasonable dose/concentration above which a degree of toxic effect is acknowledged. Also, repeated-dose studies conducted in experimental animals are designed to produce toxicity at the highest dose used in order to optimise the test objective - and so most studies will reveal some toxic effect at least at this highest dose. What is therefore to be decided is not only what effects have been produced, but also at what dose/concentration they were produced and how relevant is that for humans.

246. Thus, in animal studies, when significant toxic effects are observed, that would indicate classification, consideration of the duration of experimental exposure and the dose/concentration at which these effects were seen, in relation to the suggested guidance values, can provide useful information to help assess the need to classify (since the toxic effects are a consequence of the hazardous property(ies) and also the duration of exposure and the dose/concentration).

247. The decision to classify at all can be influenced by reference to the dose/concentration guidance values at or below which a significant toxic effect has been observed.

248. The guidance values proposed refer basically to effects seen in a standard 90-day toxicity study conducted in rats. They can be used as a basis to extrapolate equivalent guidance values for toxicity studies of greater or lesser duration, using dose/exposure time extrapolation similar to Haber's rule for inhalation, which states essentially that the effective dose is directly proportional to the exposure concentration and the duration of exposure. The assessment should be done on a case-by-case basis; e.g., for a 28-day study the guidance values below would be increased by a factor of three.

249. Thus for Category 1 classification, significant toxic effects observed in a 90-day repeated-dose study conducted in experimental animals and seen to occur at or below the (suggested) guidance values as indicated in Table 5 below would justify classification:

Table 5: Guidance values to assist in Category 1 classification

Route of exposure	Units	Guidance values (dose/concentration)
Oral (rat)	mg/kg bw/d	10
Dermal(rat or rabbit)	mg/kg bw/d	20
Inhalation (rat)gas	ppm/6h/d	50
Inhalation (rat)vapour	mg/litre/6h/d	0.2
Inhalation (rat) dust/mist/fume	mg/litre/6h/d	0.02

250. For Category 2 classification, significant toxic effects observed in a 90-day repeated-dose study conducted in experimental animals and seen to occur within the (suggested) guidance value ranges as indicated in Table 6 below would justify classification:

Table 6: Guidance values to assist in Category 2 classification

Route of Exposure	**Units**	**Guidance Value Ranges:** (dose/concentration)
Oral (rat)	mg/kg bw/d	10-100
Dermal (rat or rabbit)	mg/kg bw/d	20-200
Inhalation (rat) gas	ppm/6h/d	50-250
Inhalation (rat)vapour	mg/litre/6h/d	0.2-1.0
Inhalation (rat) dust/mist/fume	mg/litre/6h/d	0.02-0.2

251. It is important to recognise that the guidance values and ranges mentioned in paragraphs 249 and 250 are intended only for guidance purposes, i.e., to be used as part of the weight of evidence approach, and to assist with decisions about classification. They are not intended as strict demarcation values.

252. Thus it is feasible that a specific profile of toxicity is seen to occur in repeat-dose animal studies at a dose/concentration below the guidance value, eg. <100 mg/kg bw/day by the oral route, however the nature of the effect, e.g., nephrotoxicity seen only in male rats of a particular strain known to be susceptible to this effect may result in the decision not to classify. Conversely, a specific profile of toxicity may be seen in animal studies occurring at or above a guidance value, eg. ≥100 mg/kg bw/day by the oral route, and in addition there is supplementary information from other sources, e.g., other long-term administration studies, or human case experience, which supports a conclusion that, in view of the weight of evidence, classification would be the prudent action to take.

RATIONALE FOR THE PROPOSED SYSTEM

253. When a chemical is characterised only by use of animal data (typical of new chemicals, but also true for many existing chemicals), the classification process would include reference to dose/concentration guidance values as one of the elements that contribute to the weight of evidence approach.

254. When well-substantiated human data are available showing a specific target organ/systemic toxic effect that can be reliably attributed to repeated or prolonged exposure to a chemical substance, the substance may be classified. Positive human data, regardless of probable dose, predominates over animal data. Thus, if a chemical is unclassified because no specific target organ/systemic toxicity was seen at or below the proposed dose/concentration guidance value for animal testing, if subsequent human incident data become available showing a specific target organ/systemic toxic effect, the substance should be classified.

255. A chemical that has not been tested for specific target organ/systemic toxicity may in certain instances and, where appropriate, be classified on the basis of data from a validated structure activity relationship and expert judgement-based extrapolation from a structural analogue that has previously been classified together with substantial support from consideration of other important factors such as formation of common significant metabolites.

256. It is recognised that saturated vapour concentration may be used as an additional element by some regulatory systems to provide for specific health and safety protection.

Chapter 2.10:

HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICALS WHICH ARE HAZARDOUS FOR THE AQUATIC ENVIRONMENT

PURPOSE, BASIS AND APPLICABILITY

257. The harmonised system for classifying chemical substances for the hazards they present to the aquatic environment is based on a consideration of the existing systems listed below. The aquatic environment may be considered in terms of the aquatic organisms that live in the water, and the aquatic ecosystem of which they are part. To that extent, the proposal does not address aquatic pollutants for which there may be a need to consider effects beyond the aquatic environment such as the impacts on human health etc. The basis, therefore, of the identification of hazard is the aquatic toxicity of the substance, although this may be modified by further information on the degradation and bioaccumulation behaviour.

258. The proposed system is intended specifically for use with chemical substances and is not intended at this stage to cover preparations or other mixtures such as formulated pesticides. Its application to mixtures is described in Part 3, Chapter 3.9. While the scheme is intended to apply to all substances, it is recognised that for some substances, e.g. metals, poorly soluble substances etc., special guidance will be necessary.

259. A Guidance Document has been prepared to cover issues such as data interpretation and the application of the criteria defined below to such groups of substances. Considering the complexity of this endpoint and the breadth of the application of the system, the Guidance Document is considered an important element in the operation of the harmonised scheme (see Annex 2 of this document).

260. Consideration has been given to existing classification systems as currently in use, including the EU Supply and Use Scheme, the revised GESAMP (Group of Experts on the Scientific Aspects of Marine Environmental Protection) hazard evaluation procedure, IMO Scheme for Marine Pollutant, the European Road and Rail Transport Scheme (RID/ADR), the Canadian and US Pesticide systems and the US Land Transport Scheme. The harmonised scheme is considered suitable for use for packaged goods in both supply and use and multimodal transport schemes, and elements of it may be used for bulk land transport and bulk marine transport under MARPOL 73/78 Annex II insofar as this uses aquatic toxicity.

DEFINITIONS AND DATA REQUIREMENTS

261. The basic elements for use within the harmonised system are:

- acute aquatic toxicity;
- potential for or actual bioaccumulation;
- degradation (biotic or abiotic) for organic chemicals; and
- chronic aquatic toxicity.

262. While data from internationally harmonised test methods are preferred, in practice, data from national methods may also be used where they are considered as equivalent. In general, it has

been agreed that freshwater and marine species toxicity data can be considered as equivalent data and are preferably to be derived using OECD Test Guidelines or equivalent according to the principles of GLP. Where such data are not available classification should be based on the best available data.

Acute toxicity

263. Acute aquatic toxicity would normally be determined using a fish 96 hour LC_{50} (OECD Test Guideline 203 or equivalent), a crustacea species 48 hour EC_{50} (OECD Test Guideline 202 or equivalent) and/or an algal species 72 or 96 hour EC_{50} (OECD Test Guideline 201 or equivalent). These species are considered as surrogate for all aquatic organisms and data on other species such as Lemna may also be considered if the test methodology is suitable.

Bioaccumulation potential

264. The potential for bioaccumulation would normally be determined by using the octanol/water partition coefficient, usually reported as a log Kow determined by OECD Test Guideline 107 or 117. While this represents a potential to bioaccumulate, an experimentally determined Bioconcentration Factor (BCF) provides a better measure and should be used in preference when available. A BCF should be determined according to OECD Test Guideline 305.

Rapid degradability

265. Environmental degradation may be biotic or abiotic (e.g. hydrolysis) and the criteria used reflect this fact (Annex I). Ready biodegradation can most easily be defined using the OECD biodegradability tests OECD Test Guideline 301 (A - F). A pass level in these tests can be considered as indicative of rapid degradation in most environments. These are freshwater tests and thus the use of the results from OECD Test Guideline 306 which is more suitable for marine environments has also been included. Where such data are not available, a BOD(5 days)/COD ratio >0.5 is considered as indicative of rapid degradation.

266. Abiotic degradation such as hydrolysis, primary degradation, both abiotic and biotic, degradation in non-aquatic media and proven rapid degradation in the environment may all be considered in defining rapid degradability. Special guidance on data interpretation will be provided in the Guidance Document.

Chronic toxicity

267. Chronic toxicity data are less available than acute data and the range of testing procedures less standardised. Data generated according to the OECD Test Guidelines 210 (Fish Early Life Stage), or 211 (Daphnia Reproduction) and 201 (Algal Growth Inhibition) can be accepted. Other validated and internationally accepted tests could also be used. The NOECs or other equivalent L(E)Cx should be used.

CLASSIFICATION CATEGORIES AND CRITERIA

268. Substances classified under the following criteria will be categorised as ‘hazardous to the aquatic environment’. These criteria describe in detail the classification categories detailed diagrammatically in Appendix 2 to this chapter.

Acute toxicity

Category: Acute I

Acute toxicity:

96 hr LC_{50} (for fish)	≤1 mg/L and/or
48 hr EC_{50} (for crustacea)	≤1 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	≤1 mg/L.

Category: Acute I may be subdivided for some regulatory systems to include a lower band at $L(E)C_{50}$ ≤0.1 mg/L.

Category: Acute II

Acute toxicity:

96 hr LC_{50} (for fish)	>1 - ≤10 mg/L and/or
48 hr EC_{50} (for crustacea)	>1 - ≤10 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	>1 - ≤10 mg/L.

Category: Acute III

Acute toxicity:

96 hr LC_{50} (for fish)	>10 - ≤100 mg/L and/or
48 hr EC_{50} (for crustacea)	>10 - ≤100 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	>10 - ≤100 mg/L.

Some regulatory systems may extend this range beyond an $L(E)C_{50}$ of 100 mg/L through the introduction of another category.

Chronic toxicity

Category: Chronic I

Acute toxicity:

96 hr LC_{50} (for fish)	≤1 mg/L and/or
48 hr EC_{50} (for crustacea)	≤1 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	≤1 mg/L

and the substance is not rapidly degradable and/or the log Kow ≥ 4 (unless the experimentally determined BCF <500).

Category: Chronic II

Acute toxicity

96 hr LC_{50} (for fish)	>1 to ≤10 mg/L and/or
48 hr EC_{50} (for crustacea)	>1 to ≤10 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	>1 to ≤10 mg/L

and the substance is not rapidly degradable and/or the log Kow ≥4 (unless the experimentally determined BCF <500), unless the chronic toxicity NOECs are > 1 mg/L.

Category: Chronic III

Acute toxicity:

96 hr LC_{50} (for fish)	>10 to ≤100 mg/L and/or
48 hr EC_{50} (for crustacea)	>10 to ≤100 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	>10 to ≤100 mg/L

and the substance is not rapidly degradable and/or the log Kow ≥4 (unless the experimentally determined BCF <500) unless the chronic toxicity NOECs are >1 mg/L.

Category: Chronic IV

Poorly soluble substances for which no acute toxicity is recorded at levels up to the water solubility, and which are not rapidly degradable and have a log Kow ≥ 4, indicating a potential to bioaccumulate, will be classified in this category unless other scientific evidence exists showing classification to be unnecessary. Such evidence would include an experimentally determined BCF <500, or a chronic toxicity NOECs >1 mg/L, or evidence of rapid degradation in the environment.

RATIONALE FOR THE SYSTEM

269. The system for classification recognises that the core intrinsic hazard to aquatic organisms is represented by both the acute and chronic toxicity of a substance, the relative importance of which is determined by the specific regulatory system in operation. Distinction can be made between the acute hazard and the chronic hazard and therefore separate hazard categories are defined for both properties representing a gradation in the level of hazard identified. The lowest of the available toxicity values will normally be used to define the appropriate hazard category(ies). There may be circumstances, however, when a weight of evidence approach may be used. Acute toxicity data are the most readily available and the tests used are the most standardised. For that reason, these data form the core of the classification system.

270. Acute toxicity represents a key property in defining the hazard where transport of large quantities of a substance may give rise to short-term dangers arising from accidents or major spillages. Hazards categories up to $L(E)C_{50}$ values of 100 mg/L are thus defined although categories up to 1000 mg/L may be used in certain regulatory frameworks. The Acute: Category I may be further sub-divided to include an additional category for acute toxicity $L(E)C_{50} \leq 0.1$ mg/L in certain regulatory systems such as that defined by MARPOL 73/78 Annex II. It is anticipated that their use would be restricted to regulatory systems concerning bulk transport.

271. For packaged substances it is considered that the principal hazard is defined by chronic toxicity, although acute toxicity at $L(E)C_{50}$ levels ≤1 mg/L are also considered hazardous. Levels of substances up to 1 mg/L are considered as possible in the aquatic environment following normal use and disposal. At toxicity levels above this, it is considered that the short-term toxicity itself does not describe the principle hazard, which arises from low concentrations causing effects over a longer time scale. Thus, a number of hazard categories are defined which are based on levels of chronic aquatic toxicity. Chronic toxicity data are not available for many substances, however, and it is necessary to use the available data on acute toxicity to estimate this property. The intrinsic properties of a lack of rapid degradability and/or a potential to bioconcentrate in combination with acute toxicity may be used to assign a substance to a chronic hazard category. Where chronic toxicity is available showing NOECs >1 mg/L, this would indicate that no classification in a chronic hazard category would be necessary. Equally, for substances with an $L(E)C_{50}$ >100 mg/L, the toxicity is considered as insufficient to warrant classification in most regulatory systems.

272. While the current system will continue to rely on the use of acute toxicity data in combination with a lack of rapid degradation and/or a potential to bioaccumulate as the basis for classification for assigning a chronic hazard category, it is recognised that actual chronic toxicity data would form a better basis for classification where these data are available. It is thus the intention that the scheme should be further developed to accommodate such data. It is anticipated that in such a further development, the available chronic toxicity data would be used to classify in the chronic hazard in preference to that derived from their acute toxicity in combination with a lack of rapid degradation and/or a potential to bioaccumulate.

273. Recognition is given to the classification goals of MARPOL 73/78 Annex II which covers the transport of bulk quantities in ships tanks, which are aimed at regulating operational discharges from ships and assigning of suitable ship types. They go beyond that of protecting aquatic ecosystems, although that clearly is included. Additional hazard categories may thus be used which take account of factors such as physico-chemical properties and mammalian toxicity.

EXPLANATORY NOTES

274. The organisms fish, crustacea and algae are tested as surrogate species covering a range of trophic levels and taxa, and the test methods are highly standardised. Data on other organisms may also be considered, however, provided they represent equivalent species and test endpoints. The algal growth inhibition test is a chronic test but the EC_{50} is treated as an acute value for classification purposes. This EC_{50} should normally be based on growth rate inhibition. If only the EC_{50} based on reduction in biomass is available, or it is not indicated which EC_{50} is reported, this value may be used in the same way.

275. Aquatic toxicity testing by its nature, involves the dissolution of the substance under test in the water media used and the maintenance of a stable bioavailable exposure concentration over the course of the test. Some substances are difficult to test under standard procedures and thus special guidance has been developed on data interpretation for these substances and how the data should be used when applying the classification criteria (Annex 3 to this document).

276. It is the bioaccumulation of substances within the aquatic organisms that can give rise to toxic effects over longer time scales even when actual water concentrations are low. The potential to bioaccumulate is determined by the partitioning between n-octanol and water. The relationship between the partition coefficient of an organic substance and its bioconcentration as measured by the BCF in fish has considerable scientific literature support. Using a cut-off value of log K(o/w) ≥ 4 is intended to identify only those substances with a real potential to bioconcentrate. In recognition that the log P(o/w) is only an imperfect surrogate for a measured BCF, such a measured value would always take precedence. A BCF in fish of <500 is considered as indicative of a low level of bioconcentration.

277. Substances that rapidly degrade can be quickly removed from the environment. While effects can occur, particularly in the event of a spillage or accident, they will be localised and of short duration. The absence of rapid degradation in the environment can mean that a substance in the water has the potential to exert toxicity over a wide temporal and spatial scale. One way of demonstrating rapid degradation utilises the biodegradation screening tests designed to determine whether a substance is 'readily biodegradable'. Thus a substance which passes this screening test is one that is likely to biodegrade 'rapidly' in the aquatic environment, and is thus unlikely to be persistent. However, a fail in the screening test does not necessarily mean that the substance will not degrade rapidly in the environment. Thus a further criterion was added which would allow the use of data to show that the substance did actually degrade biotically or abiotically in the aquatic environment by >70% in 28 days. Thus, if degradation could be demonstrated under environmentally realistic conditions, then the definition of 'rapid degradability' would have been met. Many degradation data are available in the form of degradation half-lives and these can also be used in defining rapid degradation. Details regarding the interpretation of these data is further elaborated in the Guidance Document (Annex 3). Some tests measure the ultimate biodegradation of the substance, i.e. full mineralisation is achieved. Primary biodegradation would not normally qualify in the assessment of rapid degradability unless it can be demonstrated that the degradation products do not fulfil the criteria for classification as hazardous to the aquatic environment.

278. It must be recognised that environmental degradation may be biotic or abiotic (e.g. hydrolysis) and the criteria used reflect this fact. Equally, it must be recognised that failing the ready biodegradability criteria in the OECD tests does not mean that the substance will not be degraded rapidly in the real environment. Thus where such rapid degradation can be shown, the substance should be considered as rapidly degradable. Hydrolysis can be considered if the hydrolysis products do not fulfil the criteria for classification as hazardous to the aquatic environment. A specific definition of rapid degradability is included as Appendix 1. Other evidence of rapid degradation in the environment may also be considered and may be of particular importance where the substances are inhibitory to microbial activity at the concentration levels used in standard testing. The range of available data and guidance on its interpretation are provided in the Guidance Document (Annex 2).

279. For inorganic compounds and metals, the concept of degradability as applied to organic compounds has limited or no meaning. Rather the substance may be transformed by normal environmental processes to either increase or decrease the bioavailability of the toxic species. Equally the use of bioaccumulation data should be treated with care. Specific guidance is provided in Annex 2 on how these data for such materials may be used in meeting the requirements of the classification criteria.

280. Poorly soluble inorganic compounds and metals may be acutely or chronically toxic in the aquatic environment depending on the intrinsic toxicity of the bioavailable inorganic species and the rate and amount of this species which may enter solution. A protocol for testing these poorly soluble materials is being developed and is included in Annex 3.

281. The system also introduces as 'safety net' classification (Category: Chronic IV) for use when the data available do not allow classification under the formal criteria but there are nevertheless some grounds for concern. The precise criteria are not defined with one exception. For poorly water soluble organic substances for which no toxicity has been demonstrated, classification can occur if the substance is both not rapidly degraded and has a potential to bioaccumulate. It is considered that for such poorly soluble substances, the toxicity may not have been adequately assessed in the short-term test due to the low exposure levels and potentially slow uptake into the organism. The need for this classification can be negated by demonstrating the absence of long-term effects, i.e. a long-term NOECs > water solubility or 1 mg/L, or rapid degradation in the environment.

282. While experimentally derived test data are preferred, where no experimental data are available, validated Quantitative Structure Activity Relationships (QSARs) for aquatic toxicity and log Kow may be used in the classification process. Such validated QSARs may be used without modification to the agreed criteria, if restricted to chemicals for which their mode of action and applicability are well characterised. Validity may be judged according to the criteria established within the USEPA/EU/Japan Collaborative Project. Reliable calculated toxicity and log Kow values should be valuable in the safety net context. QSARs for predicting ready biodegradation are not yet sufficiently accurate to predict rapid degradation.

APPENDIX 1 to Chapter 2.10:

RAPID DEGRADABILITY

Substances are considered rapidly degradable in the environment if the following criteria hold true:

a) if in 28-day ready biodegradation studies, the following levels of degradation are achieved;

- tests based on dissolved organic carbon: 70%

- tests based on oxygen depletion or carbon dioxide generation: 60% of theoretical maxima

These levels of biodegradation must be achieved within 10 days of the start of degradation which point is taken as the time when 10% of the substance has been degraded.

or

b) if, in those cases where only BOD and COD data are available, when the ratio of BOD5/COD is ≥0.5

or

c) if other convincing scientific evidence is available to demonstrate that the substance can be degraded (biotically and/or abiotically) in the aquatic environment to a level >70% within a 28 day period.

APPENDIX 2 to Chapter 2.10:

Classification Scheme for Substances Hazardous to the Aquatic Environment

Toxicity		**Degradability (note 3)**	**Bioaccumulation (note 4)**	**Classification categories**	
Acute (note 1)	**Chronic (note 2)**			**Acute**	**Chronic**
Box 1 value ≤ 1.00		**Box 5**	**Box 6**	**Category:** **Acute I** **Box 1**	**Category:** **Chronic I** **Boxes 1+5+6** **Boxes 1+5** **Boxes 1+6**
Box 2 1.00 < value ≤ 10.0		lack of rapid degradability	BCF ≥ 500 or, if absent log Kow ≥ 4	**Category:** **Acute II** **Box 2**	**Category:** **Chronic II** **Boxes 2+5+6** **Boxes 2+5** **Boxes 2+6** **Unless Box 7**
Box 3 10.0 < value ≤ 100				**Category:** **Acute III** **Box 3**	**Category:** **Chronic III** **Boxes 3+5+6** **Boxes 3+5** **Boxes 3+6** **Unless Box 7**
Box 4 No acute toxicity **(note 5)**	**Box 7** value > 1.00				**Category:** **Chronic IV** **Boxes 4+5+6** **Unless Box 7**

Notes to the table:

Note 1a. Acute toxicity band based on L(E)C-50 values in mg/L for fish, crustacea and/or algae or other aquatic plants (or QSAR estimation if no experimental data).

Note 1b. Where the algal toxicity ErC-50 [= EC-50 (growth rate)] falls more than 100 times below the next most sensitive species and results in a classification based solely on this effect, consideration should be given to whether this toxicity is representative of the toxicity to aquatic plants. Where it can be shown that this is not the case, professional judgement should be used in deciding if classification should be applied. Classification should be based on the ErC-50. In circumstances where the basis of the EC-50 is not specified and no ErC-50 is recorded, classification should be based on the lowest EC-50 available.

Note 2a. Chronic toxicity band based on NOEC values in mg/L for fish or crustacea or other recognised measures for long-term toxicity.

Note 2b. It is the intention that the system be further developed to include chronic toxicity data.

Note 3. Lack of rapid degradability is based on either a lack of Ready Biodegradability or other evidence of lack of rapid degradation.

Note 4. Potential to bioaccumulate, based on an experimentally derived BCF ≥ 500 or, if absent, a log Kow ≥ 4 provided log Kow is an appropriate descriptor for the bioaccumulation potential of the substance. Measured log Kow values take precedence over estimated values and measured BCF values take precedence over log Kow values.

Note 5. "No acute toxicity" is taken to mean that the L(E)C-50 is above the water solubility. Also for poorly soluble substances, (w.s. < 1.00 mg/L), where there is evidence that the acute test would not have provided a true measure of the intrinsic toxicity.

PART 3:

HARMONISED HAZARD CLASSIFICATION CRITERIA FOR MIXTURES

Chapter 3.1:
GENERAL INTRODUCTION AND CONSIDERATIONS

INTRODUCTION

283. Part 2 of this document describes the harmonised classification criteria for chemical substances for specific health and environmental endpoints, viz., acute toxicity, skin and eye irritation/corrosion, contact and respiratory sensitisers, germ cell mutagenicity, carcinogenicity, reproductive toxicity, specific target organ toxicity, and aquatic hazards in the environment.

284. The development of these criteria for substances was part of the overall process to meet the objective defined, as one of six action programs, under Chapter XIX of the UN Conference on Environment and Development (UNCED) Agenda 21, namely: a globally harmonised hazard classification and compatible labelling system (GHS) including material safety data sheets and easily understood symbols. Part 1 of this document provides a description of the organisation and processes involved in the development of the GHS and the role of OECD, and should be consulted for further details.

285. OECD had formed an Advisory Group on Harmonisation of Classification and Labelling (AG-HCL) to pursue the development of the criteria for substances in the Integrated Document. An OECD Expert Group was subsequently formed to pursue the development of hazard classification criteria for chemical mixtures. The Expert Group on Classification Criteria for Chemical Mixtures followed similar processes to those established under the AG-HCL to achieve consensus on criteria for mixtures, including the development of documents in a stepwise manner as summarised below:

Step 1:

A thorough analysis of existing classification systems, including the scientific basis for the system and its criteria, its rationale and explanation of the mode of use.

Approaches analysis:

Many complex issues were identified that would require some resolution before a Step 2 document could be developed. Therefore, an analysis of these issues was carried out to identify critical issues together with some approaches to resolution, as an intermediate step in the process.

Step 2:

A proposal for a harmonised classification system and criteria for each endpoint was developed.

Step 3:

(a) The Expert Group on Classification Criteria for Chemical Mixtures reached consensus on a Step 2 proposal; or

(b) Any specific non-consensus items were identified as alternatives.

Step 4:

The final proposal and any non-consensus items were reviewed by the OECD AG-HCL and approved by the OECD Joint Meeting and subsequently submitted to the IOMC CG-HCCS for global implementation.

286. As experience with the use of the system is accumulated, and as new scientific information emerges, the test methods, the interpretation of the test data and the harmonised criteria *per se* may have to be updated. Thus, international work will continue to be needed in the future and, depending on the nature of the future international instrument for the implementation of the GHS, decisions will have to be made on the mechanism for carrying out the updating work in the future.

GENERAL CONSIDERATIONS

Scope of the Harmonised Classification System

287. The work on harmonisation of hazard classification and labelling focuses on a harmonised system for all chemicals and mixtures of chemicals. The application of the ingredients of the system may vary by type of product or stage of the life cycle. The classification system applies to pure chemical substances, and to mixtures of chemical substances.

288. One objective of the harmonised classification system is for it to be simple and transparent with a clear distinction between categories in order to allow for self classification as far as possible. For many endpoints the criteria are semi-quantitative or qualitative and expert judgement is required to interpret the data for classification purposes. Furthermore, for some endpoints, e.g., eye irritation, a decision tree approach is given as an example.

289. Articles as defined in the US OSHA Hazard Communication Standard (29 CFR 1910.1200), or by similar definition, are outside the scope of this document.

Presentation of Criteria

290. The GHS itself does not include requirements for testing chemicals. Therefore, there is no requirement under the GHS to generate test data for any endpoint. It is recognised that some parts of regulatory systems do require data to be generated (e.g., pesticides), but these requirements are not related specifically to the GHS.. The criteria established for classifying a mixture will allow the use of available data for the mixture itself and /or similar mixtures and /or data for ingredients of the mixture.

291. The classification criteria are presented in chapters, each of which is for a specific endpoint or a group of closely related endpoints. These chapters are based on the criteria for substances presented in the Integrated Document. The recommended process of classification for all endpoints is in the following sequence:

(1) Where test data are available for the complete mixture, the classification of the mixture will always be based on that data.
(2) Where test data are not available for the mixture itself, then the bridging principles should be considered to see whether they permit classification of the mixture.
(3) If (1) test data are not available for the mixture itself, and (2), the available information is not sufficient to allow application of the bridging principles then the agreed method(s) described in each chapter for estimating the hazards based on the information known will be applied to classify the mixture.

Test Methods and Test Data Quality[1]

292. The classification of a mixture, when it has been tested for a specific endpoint, depends both on the criteria for that endpoint and on the reliability of the test methods. In some cases the classification is determined by a pass or fail of a specific test, while in other cases, interpretations are made from dose / response curves and observations during testing. In all cases, the test conditions need to be standardised so that the results are reproducible with a given mixture and the standardised test yields valid data for defining the endpoint of concern. In this context, validation is the process by which the reliability and the relevance of a procedure are established for a particular purpose.

293. Tests that determine hazardous properties that are conducted according to internationally recognised scientific principles can be used for purposes of a hazard determination for health and environmental hazards. The GHS criteria for determining health and environmental hazards should be test method neutral, allowing different approaches as long as they are scientifically sound and validated according to international procedures and criteria already referred to in existing systems for the endpoint of concern and produce mutually acceptable data.

Previously Classified Chemicals

294. One of the general principles established by the IOMC-CG-HCCS states that test data already generated for the classification of chemicals under the existing systems should be accepted when classifying these chemicals under the harmonised system thereby avoiding duplicative testing and the unnecessary use of test animals. This policy has important implications in those cases where the criteria in the GHS are different from those in the existing system. In some cases, it may be difficult to determine the quality of existing data from older studies. In such cases, expert judgement will be needed.

Substances / Mixtures Posing Special Problems

295. The effect of a mixture on biological and environmental systems is influenced, *inter alia*, by the physico chemical properties of the mixture and / or the ingredient substances in the mixture and the way in which ingredient substances are biologically available. Some groups of substances may present special problems in this respect, for example, some polymers and metals. A mixture need not be classified when it can be shown by conclusive experimental data from internationally acceptable test methods that the mixture is not biologically available. Similarly, the result of such bioavailability data on ingredients of a mixture should be used in conjunction with the harmonised classification criteria when classifying these mixtures.

Animal Welfare

296. The welfare of experimental animals is a concern. This ethical concern includes not only the alleviation of stress and suffering but also, in some countries, the use and consumption *per se* of test animals. Where possible and appropriate, tests and experiments that do not require the use of live animals are preferred to those using sentient live experimental animals. To that end, for certain endpoints (e.g., skin and eye irritation/corrosion) testing schemes starting with non-animal observations/measurements are included as part of the classification system. For other endpoints

[1] Paragraphs 292-306 are similar or identical to paragraphs 17-31 of Part 1 of this document. They are repeated here in case Part 3 is used as a stand-alone document.

such as acute toxicity, alternative animal tests, using fewer animals or causing less suffering are internationally accepted and should be preferred to the conventional LD50 test.

Expert Judgement

297. The approach to classifying mixtures includes the application of expert judgement in a number of areas in order to ensure existing information can be used for as many mixtures as possible to provide protection for human health and the environment.

Evidence from Humans

298. For classification purposes, reliable epidemiological data and experience on the effects of chemicals on humans (e.g., occupational data, data from accident data bases) should be taken into account in the evaluation of human health hazards of a chemical. Testing on humans solely for hazard identification purposes is generally not acceptable.

Weight of Evidence

299. For some hazard endpoints, classification results directly when the data satisfy the criteria. For others, classification of a substance or mixture is made on the basis of the total weight of evidence. This means that all available information bearing on the determination of toxicity is considered together, including the results of valid in vitro tests, relevant animal data, and human experience such as epidemiological and clinical studies and well-documented case reports and observations.

300. The quality and consistency of the data are important. Evaluation of substances or mixtures related to the material under study should be included, as should site of action and mechanism or mode of action study results. Both positive and negative results are assembled together in a single weight of evidence determination.

301. Positive effects which are consistent with the criteria for classification in each chapter, whether seen in humans or animals, will normally justify classification. Where evidence is available from both sources and there is a conflict between the findings, the quality and reliability of the evidence from both sources must be assessed in order to resolve the question for classification. Generally, data of good quality and reliability in humans will have precedence over other data. However, even well-designed and conducted epidemiological studies may lack sufficient numbers of subjects to detect relatively rare but still significant effects, or to assess potentially confounding factors. Positive results from well-conducted animal studies are not necessarily negated by the lack of positive human experience but require an assessment of the robustness and quality of both the human and animal data relative to the expected frequency of occurrence of effects and the impact of potentially confounding factors.

302. Route of exposure, mechanistic information and metabolism studies are pertinent to determining the relevance of an effect in humans. When such information raises doubt about relevance in humans, a lower classification may be warranted. When it is clear that the mechanism or mode of action is not relevant to humans, the substance or mixture should not be classified.

303. Both positive and negative results are assembled together in the weight of evidence determination. However, a single positive study performed according to good scientific principles and with statistically and biologically significant positive results may justify classification.

BUILDING BLOCK APPROACH

304. At various times during the development of harmonised classification criteria, concerns have arisen concerning the way a harmonised classification system might be used and whether it would meet the needs of its various end-users.

305. One of the consequences of the application of the classification system is expressed in the IOMC CG/HCCS General Principle (c):

> "Harmonisation means establishing a common and coherent basis for chemical hazard classification and communication, from which the appropriate elements relevant to means of transport, consumer, worker and environment protection can be selected."

The application of the classification scheme may vary according to the circumstances, type of product and stage of the life cycle of the chemical.

306. It is essential that the types and levels of hazards be recognised as a fundamental basis for the harmonised classification system. For hazard classification the use of categories and subcategories other than those specified in the GHS would be contrary to harmonisation.

DEFINITIONS

307. In order to ensure that everyone understands the provisions for classifying mixtures, definitions of certain terms are required. These definitions are for the purpose of evaluating or determining the hazards of a product for classification and labelling, and are not intended to be applied in other situations such as inventory reporting. The intent of the definitions as drawn is to ensure that 1) all products within the scope of the Globally Harmonised System are evaluated to determine their hazards, and are subsequently classified according to the GHS criteria as appropriate; and 2) the evaluation is based on the actual product involved, i.e., on a stable product. If a reaction occurs during manufacture and a new product evolves, a new hazard evaluation and classification must take place to apply the GHS to the new product.

308. The following have been accepted as "working definitions":

> Substance: Chemical elements and their compounds in the natural state or obtained by any production process, including any additive necessary to preserve the stability of the product and any impurities deriving from the process used, but excluding any solvent which may be separated without affecting the stability of the substance or changing its composition.
>
> Guidance on the use of hazard classification of a substance: Where impurities, additives or individual constituents of a substance have been identified and are themselves classified, they shall be taken into account during classification if they exceed the cut-off value/concentration limit for a given endpoint.
>
> Mixture: Mixtures or solutions composed of two or more substances in which they do not react.
>
> Alloy: An alloy is a metallic material, homogeneous on a macroscopic scale, consisting of two or more elements so combined that they cannot be readily separated by mechanical means. Alloys are considered to be mixtures for the purpose of classification under the GHS.

309. It is recognised, as a practical matter, that some substances may react slowly with atmospheric gases, e.g., oxygen, carbon dioxide, water vapour, to form different substances; or they may react very slowly with other ingredient substances of a mixture to form different substances; or they may self-polymerise to form oligomers or polymers. However, the concentrations of different substances produced by such reactions are typically considered to be sufficiently low that they do not affect the hazard classification of the mixture.

310. It is recognised that consistency must be maintained between the definitions used for substances and mixtures.

Definition of "Classification"

311. It is proposed to use the term hazard classification in the GHS, as opposed to classification, to indicate that only the intrinsic hazardous properties of substances or mixtures are considered.

312. Hazard classification incorporates only 3 steps, viz.,

- identification of relevant data regarding the hazards of a substance or mixture
- subsequent review of those data to ascertain the hazards associated with the substance or mixture, and
- a decision on whether the substance or mixture will be classified as a hazardous substance or mixture and the degree of hazard, where appropriate, by comparison of the data with agreed hazard classification criteria.

313. As noted by the IOMC Co-ordinating Group, it is recognised that once a chemical is classified, the likelihood of adverse effects may be considered in deciding what informational or other steps should be taken for a given product or use setting (Ref: GHS Scope Clarification in Document - IOMC/CG13/99.2 dated 11.08.98).

The Use Of Cut-off Values/Concentration Limits

314. When classifying an untested mixture through the hazards of its ingredients, generic cut-off values or concentration limits for the classified ingredients of the mixture are used for several endpoints in the GHS. While the adopted cut-off values/concentration limits adequately identify the hazard for most mixtures, there may be some that contain hazardous ingredients in smaller concentrations than the harmonised cut-off value/concentration limit that still pose an identifiable hazard. There may also be cases where the harmonised cut-off value/concentration limit is considerably lower than could be expected on the basis of an established non-hazardous level for an ingredient.

315. Normally, the generic cut-off values/concentration limits adopted in the GHS shall be applied uniformly in all jurisdictions and for all sectors. However, if the classifier has information that the hazard of an ingredient will be evident below the generic cut-off/concentration limits, the mixture containing that ingredient must be classified accordingly.

316. On occasion, conclusive data may show that the hazard of an ingredient will not be evident when present at a level above the generic GHS cut-off/concentration limit(s). In these cases the mixture could be classified according to that data. The data should exclude the possibility that the

ingredient would behave in the mixture in a manner that would increase the hazard over that of the pure substance. Furthermore, the mixture should not contain ingredients that would affect that determination.

317. Adequate documentation supporting the change in a generic cut-off/ concentration limit(s) should be retained and made available for review on request.

Synergistic or Antagonistic Effects

318. When performing an assessment in accordance with the GHS requirements, the evaluator must take into account all available information about the potential occurrence of synergistic effects among the ingredients of the mixture. Lowering classification of a mixture to a less hazardous category on the basis of antagonistic effects may be done only if the determination is supported by sufficient data.

Endpoint Chapters

319. Regarding the content of endpoint chapters: The classification criteria for substances given in the Integrated Document will not be repeated in these chapters unless it is necessary in order to clarify the criteria for mixtures.

Chapter 3.2:

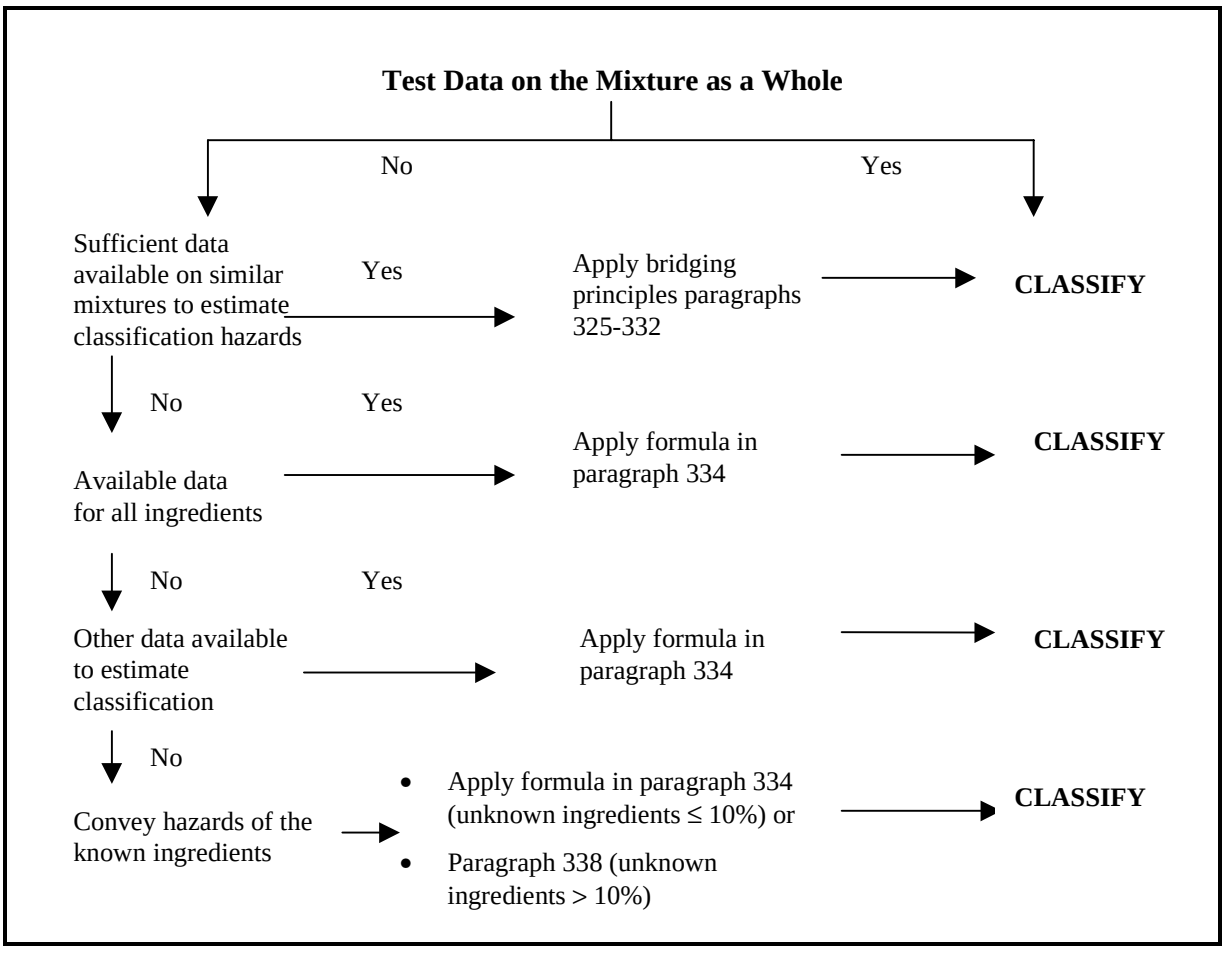
HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICAL MIXTURES WHICH CAUSE ACUTE TOXICITY

GENERAL CONSIDERATIONS

320. The harmonised criteria for the acute toxicity of substances are described in Part 2, Chapter 2.1 in this Document. The criteria for substances classify acute toxicity by use of lethal dose data (tested or derived). For mixtures, it is necessary to obtain or derive information that allows the criteria to be applied to the mixture for the purpose of classification.

321. The approach to classification for acute toxicity is tiered, and is dependent upon the amount of information available for the mixture itself and for its ingredients. The flow chart of Figure 3 below outlines the process to be followed:

Figure 3: Tiered approach to classification of mixtures for acute toxicity



322. Classification of mixtures for acute toxicity can be carried out for each route of exposure, but is only needed for one route of exposure as long as this route is followed (estimated or tested) for all ingredients. If the acute toxicity is determined for more than one route of exposure, the more severe hazard level will be used for classification. All available information should be considered and all relevant routes of exposure should be identified for hazard communication.

323. In order to make use of all available data for purposes of classifying the hazards of the mixtures, certain assumptions have been made and are applied where appropriate in the tiered approach:

a) The "relevant ingredients" of a mixture are those which are present in concentrations of 1% (w/w for solids, liquids, dusts, mists and vapours and v/v for gases) or greater, unless there is a presumption that an ingredient present at a concentration of less than 1% can still be relevant for classifying the mixture for acute toxicity.[1]

b) The acute toxicity estimate (ATE) for an ingredient in a mixture is derived using:

- The LD_{50}/LC_{50} where available,
- The appropriate conversion value from Table 7 that relates to the results of a range test for an ingredient, or
- The appropriate conversion value from Table 7 that relates to a classification for the ingredient.

c) Where a classified mixture is used as an ingredient of another mixture, the actual or derived acute toxicity estimate (ATE) for that mixture may be used when calculating the classification of the new mixture using the formulas in paragraph 334 - 338.

CLASSIFICATION OF MIXTURES WHERE ACUTE TOXICITY TEST DATA ARE AVAILABLE FOR THE COMPLETE MIXTURE.

324. Where the mixture itself has been tested to determine its acute toxicity, it will be classified according to the criteria that have been agreed for substances. In situations where such test data for the mixture are not available, the procedures presented below should be followed.

CLASSIFICATION OF MIXTURES WHERE ACUTE TOXICITY TEST DATA ARE NOT AVAILABLE FOR THE COMPLETE MIXTURE.

Bridging Principles

325. Where the mixture itself has not been tested to determine its acute toxicity, but there are sufficient data on the individual ingredients and similar tested mixtures to adequately characterise the hazards of the mixture, these data will be used in accordance with the following agreed bridging rules. This ensures that the classification process uses the available data to the greatest extent possible in characterising the hazards of the mixture without the necessity for additional testing in animals.

[1] this is particularly relevant in the case of ingredients classified in Category 1 and Category 2.

Dilution

326. If a mixture is diluted with a substance that has an equivalent or lower toxicity classification than the least toxic original ingredient, and which is not expected to affect the toxicity of other ingredients, then the new mixture may be classified as equivalent to the original mixture. Alternatively, the formula explained in paragraph 334 could be applied.

327. If a mixture is diluted with water or other totally non-toxic material, the toxicity of the mixture can be calculated from test data on the undiluted mixture. For example, if a mixture with an LD50 of 1000 mg/kg were diluted with an equal volume of water, the LD50 of the diluted mixture would be 2000 mg/kg.

Batching

328. The toxicity of one production batch of a complex mixture can be assumed to be substantially equivalent to that of another production batch of the same commercial product, and produced by or under the control of the same manufacturer, unless there is reason to believe there is significant variation such that the toxicity of the batch has changed. If the latter occurs, new classification is necessary.

Concentration Of Highly Toxic Mixtures

329. If a mixture is classified in Category 1, and the concentration of the ingredients of the mixture that are in Category 1 is increased, the new mixture should be classified in Category 1 without additional testing.

Interpolation Within One Toxicity Category

330. For three mixtures with identical ingredients, where A and B are in the same toxicity category and mixture C has toxicologically active ingredients with concentrations intermediate to those in mixtures A and B, then mixture C is assumed to be in the same toxicity category as A and B.

Substantially Similar Mixtures

331. Given the following:

a). Two mixtures: (i) A + B
(ii) C + B
b). The concentration of ingredient B is essentially the same in both mixtures.
c). The concentration of ingredient A in mixture (i) equals that of ingredient C in mixture (ii).
d). Data on toxicity for A and C are available and substantially equivalent, i.e. they are in the same hazard category and are not expected to affect the toxicity of B.

If mixture (i) is already classified by testing, mixture (ii) can be assigned the same hazard category.

Aerosols

332. An aerosol form of a mixture may be classified in the same hazard category as the tested, non aerosolised form of the mixture for oral and dermal toxicity provided the added propellant does not affect the toxicity of the mixture on spraying. Classification of aerosolised mixtures for inhalation toxicity should be considered separately.

CLASSIFICATION OF MIXTURES BASED ON INGREDIENTS OF THE MIXTURE (ADDITIVITY FORMULA).

Data Available For All Ingredients

333. In order to ensure that classification of the mixture is accurate, and that the calculation need only be performed once for all systems, sectors, and categories, the acute toxicity estimate (ATE) of ingredients should be considered as follows:

- Include ingredients with a known acute toxicity, which fall into any of the GHS acute toxicity categories.
- Ignore ingredients that are presumed not acutely toxic (e.g., water, sugar).
- Ignore ingredients if the oral limit test does not show acute toxicity at 2,000 mg/kg/body weight.

Ingredients that fall within the scope of this paragraph are considered to be ingredients with a known acute toxicity estimate (ATE).

334. The ATE of the mixture is determined by calculation from the ATE values for all relevant ingredients according to the following formula below for Oral, Dermal or Inhalation Toxicity:

$$\frac{100}{ATE_{mix}} = \sum_{\eta} \frac{C_i}{ATE_i}$$

where:

C_i= concentration of ingredient i

n ingredients and i is running from 1 to n

ATE_i = Acute Toxicity Estimate of ingredient i

Data Are Not Available For One Or More Ingredients Of The Mixture.

335. Where an ATE is not available for an individual ingredient of the mixture, but available information such as listed below can provide a derived conversion value, the formula in paragraph 334 may be applied.

This may include evaluation of:

(a) Extrapolation between oral, dermal and inhalation acute toxicity estimates[1]. Such an evaluation could require appropriate pharmacodynamic and pharmacokinetic data;

[1]. For ingredients with acute toxicity estimates available for other than the most appropriate exposure route, values may be extrapolated from the available exposure route to the most relevant route. Dermal and inhalatory route data are not always required for ingredients. However, in case data requirements for specific ingredients include acute toxicity estimates for the dermal and inhalatory route, the values to be used in the formula need to be from the required exposure route.

(b) Evidence from human exposure that indicates toxic effects but does not provide lethal dose data;

(c) Evidence from any other toxicity tests/assays available on the substance that indicates toxic acute effects but does not necessarily provide lethal dose data; or

(d) Data from closely analogous substances using structure/activity relationships.

336. This approach generally requires substantial supplemental technical information, and a highly trained and experienced expert, to reliably estimate acute toxicity. If such information is not available, proceed to the provisions of paragraph 337.

337. In the event that an ingredient without any useable information at all is used in a mixture at a concentration of 1% or greater, it is concluded that the mixture cannot be attributed a definitive acute toxicity estimate. In this situation the mixture should be classified based on the known ingredients only, with the additional statement that x percent of the mixture consists of ingredient(s) of unknown toxicity.

338. If the total concentration of the ingredient(s) with unknown acute toxicity is ≤ 10% then the formula presented in paragraph 334 should be used. If the total concentration of the ingredient(s) with unknown toxicity is >10%, the formula presented in paragraph 334 should be corrected to adjust for the total percentage of the unknown ingredient(s) as follows:

$$\frac{100 - (\sum C_{\text{unknown}} \text{ if } >10\% \ )}{ATE_{mix}} = \sum_{\eta} \frac{Ci}{ATEi}$$

Table 7: Conversion from the experimentally obtained acute toxicity range estimates or a classification to point estimates for the respective routes of exposure.

	Classification or experimentally obtained acute toxicity range estimate (see note 1)	**Conversion value (note 2)**
Oral (mg/kg)	0 < Category 1 ≤ 5 5 < Category 2 ≤ 50 50 < Category 3 ≤ 300 300 < Category 4 ≤ 2000 *2000 < Category 5 ≤ 5000*	0.5 5 100 500 2500
Dermal (mg/kg)	0 < Category 1 ≤ 50 50 < Category 2 ≤ 200 200 < Category 3 ≤ 1000 1000 < Category 4 ≤ 2000 *2000 < Category 5 ≤ 5000*	5 50 300 1100 2500
Gases (ppm)	0 < Class1 ≤ 100 100 < Category 2 ≤ 500 500 < Category 3 ≤ 2500 2500 < Category 4 ≤ 5000 *Category 5*	10 100 700 3000
Vapours (mg/l)	0 < Category 1 ≤ 0.5 0.5 < Category 2 ≤ 2.0 2.0 < Category 3 ≤ 10.0 10.0 < Category 4 ≤ 20.0 *Category 5*	0.05 0.5 3 11
Dust/mist (mg/l)	0 < Category 1 ≤ 0.05 0.05 < Category 2 ≤ 0.5 0.5 < Category 3 ≤ 1.0 1.0 < Category 4 ≤ 5.0 *Category 5*	0.005 0.05 0.5 1.5

Note1: Category 5 is for mixtures which are of relatively low acute toxicity but which under certain circumstances may pose a hazard to vulnerable populations. These mixtures are anticipated to have an oral or dermal LD_{50} value in the range of 2000-5000mg/kg or equivalent dose for other routes of exposure. In light of animal welfare considerations, testing in animals in Category 5 ranges is discouraged and should only be considered when there is a strong likelihood that results of such testing would have a direct relevance for protecting human health.

Note2: These values are designed to be used in the calculation of the ATE for a mixture based on its components and do not represent test results. The values are conservatively set at the lower end of the range of Categories 1 and 2, and at a point approximately $1/10^{th}$ from the lower end of the range for Categories 3 – 5.

Chapter 3.3:

HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICAL MIXTURES WHICH CAUSE SKIN AND EYE CORROSION/IRRITATION

GENERAL CONSIDERATION

339. The harmonised criteria for the skin and eye irritation / corrosion of substances are described in Part 2, Chapter 2.2 and 2.3 of this document.

CLASSIFICATION OF MIXTURES WHEN DATA ARE AVAILABLE FOR THE COMPLETE MIXTURE.

340. The mixture will be classified using the criteria for substances, and taking into account the testing and evaluation strategies to develop data for these endpoints.

341. Unlike other endpoints, there are alternative tests available for skin corrosivity of certain categories of chemicals that can give an accurate result for classification purposes, as well as being simple and relatively inexpensive to perform. When considering testing of the mixture manufacturers are encouraged to use a tiered weight of evidence strategy as included in the criteria for classification of substances for eye and skin corrosion and irritation to help ensure an accurate classification, as well as avoid unnecessary animal testing. A mixture is considered corrosive (Skin Category 1, Eye Category 1) if it has a pH of 2 or less or 11.5 or greater. If consideration of alkali/acid reserve suggests the substance or preparation may not be corrosive despite the low or high pH value, then further testing needs to be carried out to confirm this, preferably by use of an appropriate validated *in vitro* test.

CLASSIFICATION OF MIXTURES WHEN DATA ARE NOT AVAILABLE FOR THE COMPLETE MIXTURE.

Bridging Principles

342. Where the mixture itself has not been tested to determine its skin and eye irritation/corrosion, but there are sufficient data on the individual ingredients and similar tested mixtures to adequately characterise the hazards of the mixture, these data will be used in accordance with the following agreed bridging rules. This ensures that the classification process uses the available data to the greatest extent possible in characterising the hazards of the mixture without the necessity for additional testing in animals.

Dilution

343. Skin: If a mixture is diluted with a diluent which has an equivalent or lower corrosivity/irritancy classification than the least corrosive/irritant original ingredient and which is not expected to affect the corrosivity/irritancy of other ingredients, then the new mixture may be classified as equivalent to the original mixture. Alternatively, the method explained in paragraphs 350 - 355 could be applied.

344. Eye: If a mixture is diluted with a diluent which has an equivalent or lower corrosivity/irritancy classification than the least corrosive/irritant original ingredient and which is not expected to affect the corrosivity/irritancy of other ingredients, then the new mixture may be classified as equivalent to the original mixture. Alternatively, the method explained in paragraphs 350 - 355 could be applied.

Batching

345. The irritation/corrosion potential of one production batch of a complex mixture can be assumed to be substantially equivalent to that of another production batch of the same commercial product and produced by or under the control of the same manufacturer, unless there is reason to believe there is significant variation such that the toxicity of the batch has changed. If the latter occurs, new classification is necessary.

Concentration of Mixtures of the Highest Corrosion / Irritation Category

346. If a tested mixture classified in the highest subcategory for corrosion is concentrated, a more concentrated mixture should be classified in the highest corrosion subcategory without additional testing. If a tested mixture classified in the highest category for skin/eye irritation is concentrated and does not contain corrosive ingredients, a more concentrated mixture should be classified in the highest irritation category without additional testing.

Interpolation within One Toxicity Category

347. If mixtures A and B are in the same irritation/corrosion toxicity category and mixture C is made in which the toxicologically active ingredients have concentrations intermediate to those in mixtures A and B, then mixture C is assumed to be in the same irritation/corrosion category as A and B. Note that the identity of the ingredients is the same in all three mixtures.

Substantially Similar Mixtures

348. Given the following:

a). Two mixtures (i.) A +B
(ii.) C + B

b). The concentration of ingredient B is essentially the same in both mixtures.

c). The concentration of ingredient A in mixture (i) equals that of ingredient C in mixture (ii).

d). Data on irritation/corrosion for A and C are available and substantially equivalent, i.e., they are in the same hazard category and are not expected to affect the toxicity of B.

If mixture (i) is already classified by testing, mixture (ii) can be assigned in the same category.

Aerosols

349. An aerosol form of a mixture may be classified in the same hazard category as the tested non-aerosolised form of mixture provided that the added propellant does not affect the irritation or corrosive properties of the mixture upon spraying[1].

1. Bridging rules apply for the intrinsic hazard classification of aerosols, however, the need to evaluate the potential for "mechanical" eye damage from the physical force of the spray is recognised.

CLASSIFICATION OF MIXTURES WHEN DATA ARE AVAILABLE FOR ALL INGREDIENTS OR ONLY FOR SOME INGREDIENTS OF THE MIXTURE.

350. In order to make use of all available data for purposes of classifying the skin and eye irritation/corrosion hazards of the mixtures, the following assumption has been made and is applied where appropriate in the tiered approach:

> The "relevant ingredients" of a mixture are those which are present in concentrations of 1% (w/w for solids, liquids, dusts, mists and vapours and v/v for gases) or greater, unless there is a presumption (e.g., in the case of corrosive ingredients) that an ingredient present at a concentration of less than 1% can still be relevant for classifying the mixture for skin and eye irritation/corrosion.

351. In general, the approach to classification of mixtures as irritant or corrosive to skin and/or eye when data are available on the components, but not on the mixture as a whole, is based on the theory of additivity, such that each corrosive or irritant component contributes to the overall irritant or corrosive properties of the mixture in proportion to its potency and concentration. A weighting factor of 10 is used for corrosive components when they are present at a concentration below the concentration limit for classification with Category 1, but are at a concentration that will contribute to the classification of the mixture as an irritant. The mixture is classified as corrosive or irritant when the sum of the concentrations of such components exceeds a threshold concentration limit.

352. Tables 8 and 9 below provide the concentration limits to be used to determine if the mixture is considered to be an irritant or a corrosive for skin and eye respectively.

353. Particular care must be taken when classifying certain types of chemicals such as acids and bases, inorganic salts, aldehydes, phenols, and surfactants. The approach explained in paragraphs 351 and 352 might not work given that many of such substances are corrosive or irritant at concentrations < 1%. For mixtures containing strong acids or bases the pH should be used as classification criteria (see paragraph 341) since pH will be a better indicator of corrosion than the concentration limits of Tables 8 and 9. In the case of mixtures containing corrosive or irritant ingredients that cannot be classified based on the additivity approach applied in Tables 8 and 9 due to chemical characteristics that make this approach unworkable, a mixture will be classified as Skin Category 1 and Eye Category 1 if it contains ≥ 1% of a corrosive ingredient and as Skin Category 2/3 and Eye Category 2 when it contains ≥ 3% of an irritant ingredient. Classification of mixtures with ingredients for which the approach in Tables 8 and 9 does not apply is summarised in Table 10 below.

354. On occasion, reliable data may show that the skin corrosion/irritation or the reversible/irreversible eye effects of an ingredient will not be evident when present at a level above the generic concentration cut-off levels mentioned in Tables 8-10. In these cases the mixture could be classified according to that data (see also paragraph 316). On occasion, when it is expected that the skin corrosion/irritation or the reversible/irreversible eye effects of an ingredient will not be evident when present at a level above the generic concentration cut-off levels mentioned in Tables 8-10, testing of the mixture may be considered. In those cases the tiered weight of evidence strategy should be applied as referred to in paragraph 341 and explained in detail in the chapter on classification of substances for skin and eye hazards.

355. If there is data showing that (an) ingredient(s) may be corrosive or irritant at a concentration of < 1% (corrosive) or < 3% (irritant), the mixture should be classified accordingly (see also paragraph 314).

Table 8 : Concentration of ingredients of a mixture classified as skin category 1, 2 or 3 that would trigger classification of the mixture as <u>hazardous to</u> <u>skin</u> (category 1, 2 or 3).

Sum of ingredients classified as:	Concentration triggering classification of a mixture as:		
	Skin		
	Corrosive	Irritant	
	Category 1 (see note below)	Category 2	Category 3
Skin Category 1	≥5%	≥1% but < 5%	
Skin Category 2		≥10%	≥1% but < 10%
Skin Category 3			≥10%
(10 x Skin Category 1) + Skin Category 2		≥10%	≥1% but <10%
(10 x Skin Category 1) + Skin Category 2+Skin Category 3			≥10%

Note to Table 8 : Only some authorities will use the subcategories of Skin Category 1 (corrosive). In these cases, the sum of all ingredients of a mixture classified as Skin Category 1A, 1B or 1C respectively, should each be ≥ 5% in order to classify the mixture as either Skin Category 1A, 1B or 1C. In case the sum of the Skin Category 1A ingredients is < 5% but the sum of Skin Category ingredients 1A+1B is ≥ 5%, the mixture should be classified as Skin Category 1B. Similarly, in case the sum of Skin Category 1A+1B is < 5% but the sum of Category 1A+1B+1C is ≥ 5% the mixture would be classified as Category 1C.

Table 9: Concentration of ingredients of a mixture classified as skin category 1 and/or eye category 1 or 2 that would trigger classification of the mixtures as <u>hazardous to the eye</u> (category 1 or 2).

Sum of Ingredients Classified as:	Concentration triggering classification of a mixture as:	
	Eye	
	Irreversible	Reversible
	Category 1	Category 2
Eye or Skin Category 1	≥ 3%	≥1% but < 3%
Eye Category 2/2A		≥10%
(10 x Eye Category 1) + Eye Category 2/2A		≥10%
Skin Category 1 + Eye Category 1	≥ 3%	≥1% but <3%
10 x (Skin Category 1 + Eye Category 1) + Eye Category 2/2A		≥10%

Table 10: Concentration of ingredients of a mixture for which the additivity approach does not apply, that would trigger classification of the mixture as hazardous to skin or the eye.

Ingredient:	Concentration:	Mixture classified as:	
		Skin	Eye
Acid with pH ≤ 2	≥ 1%	Category 1	Category 1
Base with pH ≥11.5	≥ 1%	Category 1	Category 1
Other corrosive (Category 1) ingredients for which additivity does not apply	≥ 1%	Category 1	Category 1
Other irritant (Category 2) ingredients for which additivity does not apply, including acids and bases	≥ 3%	Category 2	Category 2

Chapter 3.4:

HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICAL MIXTURES WHICH CAUSE RESPIRATORY OR SKIN SENSITISATION [2]

GENERAL CONSIDERATIONS

356. The harmonised criteria for respiratory and skin sensitisation of substances are described in Part 2, Chapter 2.4 of this document.

CLASSIFICATION OF MIXTURES WHEN DATA ARE AVAILABLE FOR THE COMPLETE MIXTURE.

357. When reliable and good quality evidence from human experience or appropriate studies in experimental animals, as described in the criteria for substances, is available for the mixture, then the mixture can be classified by weight of evidence evaluation of these data. Care should be exercised in evaluating data on mixtures, that the dose used does not render the results inconclusive.

CLASSIFICATION OF MIXTURES WHEN DATA ARE NOT AVAILABLE FOR THE COMPLETE MIXTURE.

Bridging Principles

358. Where the mixture itself has not been tested to determine its sensitising properties, but there are sufficient data on the individual ingredients and similar tested mixtures to adequately characterise the hazards of the mixture, these data will be used in accordance with the following agreed bridging rules. This ensures that the classification process uses the available data to the greatest extent possible in characterising the hazards of the mixture without the necessity for additional testing in animals.

Dilution

359. If a mixture is diluted with a diluent which is not a sensitiser and which is not expected to affect the sensitisation of other ingredients, then the new mixture may be classified as equivalent to the original mixture.

[2] There has been considerable discussion about what to convey about sensitisation effects to those exposed, and at what point it should be conveyed. While the current cut-off for mixtures is 1%, it appears that the major systems all believe information should be conveyed below that level. This may be appropriate both to warn those already sensitised, as well as to warn those who may become sensitised. This issue was not clear during the initial deliberations on the criteria for mixtures containing sensitisers, and thus has not been adequately discussed nor options explored.

Before the system becomes implemented, this issue should be revisited by the ECOSOC Subcommittee on the GHS as one of its first priorities. It should be noted that the sensitisation criteria for substances will also have to be re-opened to consider this issue and the inclusion of new information and evolving testing approaches that addresses the question of strong sensitisers versus those that are weaker. Appropriate hazard communication should be considered along with the discussions on the criteria and the availability of an appropriate test method.

Batching

360. The sensitising properties of one production batch of a complex mixture can be assumed to be substantially equivalent to that of another production batch of the same commercial product and produced by or under the control of the same manufacturer, unless there is reason to believe there is significant variation such that the sensitisation of the batch has changed. If the latter occurs, new classification is necessary.

Substantially Similar Mixtures

361. Given the following:

a). Two mixtures: (i.) A + B
(ii.) C + B
b). The concentration of ingredient B is essentially the same in both mixtures.
c). The concentration of ingredient A in mixture (i) equals that of ingredient C in mixture (ii).
d). Ingredient B is a sensitiser and Ingredients A and C are not sensitisers.
e). A and C are not expected to affect the sensitisation of B.

If mixture (i) is already classified by testing, mixture (ii) can be assigned the same hazard category.

Aerosols

362. An aerosol form of the mixture may be classified in the same hazard category as the tested non-aerosolised form of the mixture provided that the added propellant does not affect the sensitising properties of the mixture upon spraying.

CLASSIFICATION OF MIXTURES WHEN DATA ARE AVAILABLE FOR ALL INGREDIENTS OR ONLY FOR SOME INGREDIENTS OF THE MIXTURE.

363. The mixture will be classified as a respiratory or skin sensitiser when at least one ingredient has been classified as a respiratory or skin sensitiser and is present at or above the appropriate cut-off value / concentration limit for the specific endpoint as mentioned in Table 11 below for solid/liquid and gas respectively.

Table 11: Cut-off values/concentration limits of ingredients of a mixture classified as either skin sensitisers or respiratory sensitisers, that would trigger classification of the mixture.

Ingredient classified as:	Cut-off/concentration limits triggering classification of a mixture as:	
	Skin sensitiser	Respiratory sensitisers
Skin sensitiser	≥1.0% w/w	≥1.0%v/v
Respiratory sensitiser	≥1.0% w/w	≥0.2% v/v

Chapter 3.5:

HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICAL MIXTURES WHICH CAUSE GERM CELL MUTAGENICITY

GENERAL CONSIDERATIONS

364. The harmonised criteria for germ cell mutagenicity of substances are described in Part 2, Chapter 2.5 of this document.

CLASSIFICATION OF MIXTURES WHEN DATA ARE AVAILABLE FOR THE COMPLETE MIXTURE.

365. Classification of mixtures will be based on the available test data on the individual constituents of the mixture using cut-off values/concentration limits for the components of the mixture. The classification may be modified on a case-by-case basis based on the available test data for the mixture as a whole. In such cases, the test results for the mixture as a whole must be shown to be conclusive taking into account dose and other factors such as duration, observations and analysis (e.g., statistical analysis, test sensitivity) of germ cell mutagenicity test systems. Adequate documentation supporting the classification should be retained and made available for review upon request.

CLASSIFICATION OF MIXTURES WHEN DATA ARE NOT AVAILABLE FOR THE COMPLETE MIXTURE.

Bridging Principles

366. Where the mixture itself has not been tested to determine its germ cell mutagenicity hazard, but there are sufficient data on the individual ingredients and similar tested mixtures to adequately characterise the hazards of the mixture, these data will be used in accordance with the following agreed bridging rules. This ensures that the classification process uses the available data to the greatest extent possible in characterising the hazards of the mixture without the necessity for additional testing in animals.

Dilution

367. If a mixture is diluted with a diluent which is not expected to affect the germ cell mutagenicity of other ingredients, then the new mixture may be classified as equivalent to the original mixture.

Batching

368. The germ cell mutagenic potential of one production batch of a complex mixture can be assumed to be substantially equivalent to that of another production batch of the same commercial product produced by and under the control of the same manufacturer unless there is reason to believe there is significant variation in composition such that the germ cell mutagenic potential of the batch has changed. If the latter occurs, a new classification is necessary.

Substantially similar mixtures

369. Given the following:

a). Two mixtures: i.) A + B
ii.) C + B
b). The concentration of mutagen Ingredient B is the same in both mixtures.
c). The concentration of ingredient A in mixture (i) equals that of ingredient C in mixture (ii).
d). Data on toxicity for A and C are available and substantially equivalent, i.e. they are not expected to affect the germ cell mutagenicity of B.

If mixture (i) is already classified by testing, mixture (ii) can be assigned the same category.

CLASSIFICATION OF MIXTURES WHEN DATA ARE AVAILABLE FOR ALL INGREDIENTS OR ONLY FOR SOME INGREDIENTS OF THE MIXTURE.

370. The mixture will be classified as a mutagen when at least one ingredient has been classified as a Category 1 or Category 2 mutagen and is present at or above the appropriate cut-off value/concentration limit as mentioned in Table 12 below for Category 1 and 2 respectively.

Table 12: Cut-off values/concentration limits of ingredients of a mixture classified as germ cell mutagens that would trigger classification of the mixture.

Ingredient classified as:	Cut-off/concentration limits triggering classification of a mixture as:	
	Category 1 mutagen	Category 2 mutagen
Category 1 mutagen	≥ 0.1 %	-
Category 2 mutagen	-	≥ 1.0%

Note: The cut-off values/concentration limits in the table above apply to solids and liquids (w/w units) as well as gases (v/v units).

Chapter 3.6:

HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICAL MIXTURES WHICH CAUSE CARCINOGENICITY

GENERAL CONSIDERATIONS

371. The harmonised criteria for carcinogenicity of substances are described Part 2, Chapter 2.6 of this document.

CLASSIFICATION OF MIXTURES WHEN DATA ARE AVAILABLE FOR THE COMPLETE MIXTURE.

372. Classification of mixtures will be based on the available test data on the individual constituents of the mixture using cut-off values/concentration limits for the components of the mixture. The classification may be modified on a case-by case basis based on the available test data for the mixture as a whole. In such cases, the test results for the mixture as a whole must be shown to be conclusive taking into account dose and other factors such as duration, observations and analysis (e.g., statistical analysis, test sensitivity) of carcinogenicity test systems. Adequate documentation supporting the classification should be retained and made available for review upon request.

CLASSIFICATION OF MIXTURES WHEN DATA ARE NOT AVAILABLE FOR THE COMPLETE MIXTURE.

Bridging Principles

373. Where the mixture itself has not been tested to determine its carcinogenic hazard, but there are sufficient data on the individual ingredients and similar tested mixtures to adequately characterise the hazards of the mixture, these data will be used in accordance with the following agreed bridging rules. This ensures that the classification process uses the available data to the greatest extent possible in characterising the hazards of the mixture without the necessity for additional testing in animals.

Dilution

374. If a mixture is diluted with a diluent which is not expected to affect the carcinogenicity of other ingredients, then the new mixture may be classified as equivalent to the original mixture.

Batching

375. The carcinogenic potential of one production batch of a complex mixture can be assumed to be substantially equivalent to that of another production batch of the same commercial product produced by and under the control of the same manufacturer unless there is reason to believe there is significant variation in composition such that the carcinogenic potential of the batch has changed. If the latter occurs, a new classification is necessary.

Substantially similar mixtures

376. Given the following:

a). Two mixtures: i.) A + B
ii.) C + B
b). The concentration of carcinogen ingredient B is the same in both mixtures.
c). The concentration of ingredient A in mixture i equals that of ingredient C in mixture ii.
d). Data on toxicity for A and C are available and substantially equivalent, i.e. they are not expected to affect the carcinogenicity of B.

If mixture (i) is already classified by testing, mixture (ii) can be assigned the same category.

CLASSIFICATION OF MIXTURES WHEN DATA ARE AVAILABLE FOR ALL COMPONENTS OR ONLY FOR SOME COMPONENTS OF THE MIXTURE.

377. The mixture will be classified as a carcinogen when at least one ingredient has been classified as a Category 1 or Category 2 carcinogen and is present at or above the appropriate cut-off value/concentration limit as mentioned in Table 13 below for Category 1 and 2 respectively.

Table 13: Cut-off values/concentration limits of ingredients of a mixture classified as carcinogen that would trigger classification of the mixture[1].

Ingredient classified as:	Cut-off/concentration limits triggering classification of a mixture as:	
	Category 1 carcinogen	Category 2 carcinogen
Category 1 carcinogen	≥ 0.1 %	
Category 2 carcinogen	-	≥ 0.1% (note1)
		≥ 1.0% (note 2)

Note 1: If a Category 2 carcinogen ingredient is present in the mixture at a concentration between 0.1% and 1%, every regulatory authority would require information on the SDS for a product. However, a label warning would be optional. Some authorities will choose to label when the ingredient is present in the mixture between 0.1% and 1%, whereas others would normally not require a label in this case.

Note 2: If a Category 2 carcinogen ingredient is present in the mixture at a concentration of $\geq$ 1%, both an SDS and a label would generally be expected.

[1] This compromise classification scheme involves consideration of differences in hazard communication practices in existing systems. Although it is recognised that this may result in a lack of harmonisation for some mixtures, the OECD Expert Group is recommending to the ILO Hazard Communication Work Group that this compromise be accepted as a way to move the process forward. It is expected that the number of affected mixtures will be small; the differences will be limited to label warnings; and the situation will evolve over time to a more harmonised approach. All of these hazard communication recommendations are subject to review by the ILO Work Group, and may be affected by that group's determinations regarding the possibility of using risk considerations in labelling in the consumer sector.

Chapter 3.7:

HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICAL MIXTURES WHICH CAUSE REPRODUCTIVE TOXICITY

GENERAL CONSIDERATION

378. The harmonised criteria for reproductive toxicity of substances are described in Part 2, Chapter 2.7 of this document.

CLASSIFICATION OF MIXTURES WHEN DATA ARE AVAILABLE FOR THE COMPLETE MIXTURE.

379. Classification of mixtures will be based on the available test data on the individual constituents of the mixture using cut-off values/concentration limits for the components of the mixture. The classification may be modified on a case-by case basis based on the available test data for the mixture as a whole. In such cases, the test results for the mixture as a whole must be shown to be conclusive taking into account dose and other factors such as duration, observations and analysis (e.g., statistical analysis, test sensitivity) of reproduction test systems. Adequate documentation supporting the classification should be retained and made available for review upon request.

CLASSIFICATION OF MIXTURES WHEN DATA ARE NOT AVAILABLE FOR THE COMPLETE MIXTURE.

Bridging Principles

380. Where the mixture itself has not been tested to determine its reproductive toxicity, but there are sufficient data on the individual ingredients and similar tested mixtures to adequately characterise the hazards of the mixture, these data will be used in accordance with the following agreed bridging rules. This ensures that the classification process uses the available data to the greatest extent possible in characterising the hazards of the mixture without the necessity for additional testing in animals.

Dilution

381. If a mixture is diluted with a diluent which is not expected to affect the reproductive toxicity of other ingredients, then the new mixture may be classified as equivalent to the original mixture.

Batching

382. The reproductive toxicity potential of one production batch of a complex mixture can be assumed to be substantially equivalent to that of another production batch of the same commercial product produced by and under the control of the same manufacturer unless there is reason to believe there is significant variation in composition such that the reproductive toxicity potential of the batch has changed. If the latter occurs, a new classification is necessary.

Substantially similar mixtures

383. Given the following:

a). Two mixtures: i.) A + B
ii.) C + B
b). The concentration of Ingredient B, toxic to reproduction, is the same in both mixtures.
c). The concentration of ingredient A in mixture i equals that of ingredient C in mixture ii.
d). Data on toxicity for A and C are available and substantially equivalent, i.e. they are not expected to affect the reproductive toxicity of B.

If mixture (i) is already classified by testing, mixture (ii) can be assigned the same category.

CLASSIFICATION OF MIXTURES WHEN DATA ARE AVAILABLE FOR ALL COMPONENTS OR ONLY FOR SOME COMPONENTS OF THE MIXTURE.

384. The mixture will be classified as a reproductive toxin when at least one ingredient has been classified as a Category 1 or Category 2 reproductive toxicant and is present at or above the appropriate cut-off value/concentration limit as mentioned in Table 14 below for Category 1 and 2 respectively.

Table 14 : Cut-off values/concentration limits of ingredients of a mixture classified as reproductive toxicants that would trigger classification of the mixture.[1]

Ingredient classified as:	Cut-off/concentration limits triggering classification of a mixture as:	
	Category 1 reproductive toxicant	Category 2 reproductive toxicant
Category 1 reproductive toxicant	≥ 0.1 % (note 1) --- ≥ 0.3 % (note 2)	
Category 2 reproductive toxicant		≥ 0.1 % (note 3) --- ≥ 3.0 % (note 4)

Note 1: If a Category 1 reproductive toxicant is present in the mixture as an ingredient at a concentration between 0.1% and 0.3%, every regulatory authority would require information on the SDS for a product. However, a label warning would be optional. Some authorities will choose to label when the ingredient is present in the mixture between 0.1% and 0.3%, whereas others would normally not require a label in this case.

[1] This compromise classification scheme involves consideration of differences in hazard communication practices in existing systems. Although it is recognised that this may result in a lack of harmonisation for some mixtures, the OECD Expert Group is recommending to the ILO Hazard Communication Work Group that this compromise be accepted as a way to move the process forward. It is expected that the number of affected mixtures will be small; the differences will be limited to label warnings; and the situation will evolve over time to a more harmonised approach. All of these hazard communication recommendations are subject to review by the ILO Work Group, and may be affected by that group's determinations regarding the possibility of using risk considerations in labelling in the consumer sector.

Note 2: If a Category 1 reproductive toxicant reproductive toxicant is present in the mixture as an ingredient at a concentration of $\geq$ 0.3%, both an SDS and a label would generally be expected.

Note 3: If a Category 2 reproductive toxicant is present in the mixture as an ingredient at a concentration between 0.1% and 3.0%, every regulatory authority would require information on the SDS for a product. However, a label warning would be optional. Some authorities will choose to label when the ingredient is present in the mixture between 0.1% and 3.0%, whereas others would normally not require a label in this case.

Note 4: If a Category 2 reproductive toxicant is present in the mixture as an ingredient at a concentration of $\geq$ 3.0%, both an SDS and a label would generally be expected.

Chapter 3.8:

HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICAL MIXTURES WHICH CAUSE SPECIFIC TARGET ORGAN SYSTEMIC TOXICITY

GENERAL CONSIDERATION

385. The harmonised criteria for the classification of chemical substances for specific target organ/systemic toxicity, following single or repeated/prolonged exposure, are described in Part 2, Chapters 2.8 and 2.9 of this document. Mixtures are classified using the same criteria as for substances, or alternatively as described below. As with substances, mixtures may be classified for target organ/systemic toxicity following single exposure, repeated exposure, or both.

CLASSIFICATION OF MIXTURES WHEN RELIABLE EVIDENCE OR TEST DATA ARE AVAILABLE FOR THE COMPLETE MIXTURE.

386. When reliable and good quality evidence from human experience or appropriate studies in experimental animals, as described in the criteria for substances, is available for the mixture, then the mixture can be classified by weight of evidence evaluation of this data. Care should be exercised in evaluating data on mixtures, that the dose, duration, observation or analysis, do not render the results inconclusive.

CLASSIFICATION OF MIXTURES WHEN DATA ARE NOT AVAILABLE FOR THE COMPLETE MIXTURE.

Bridging Principles

387. Where the mixture itself has not been tested to determine its target organ/systemic toxicity, but there are sufficient data on the individual ingredients and similar tested mixtures to adequately characterise the hazards of the mixture, these data can be used in accordance with the following bridging principles. This ensures that the classification process uses the available data to the greatest extent possible in characterising the hazards of the mixture without the necessity of additional testing in animals.

Dilution

388. If a mixture is diluted with a diluent which has the same or a lower toxicity classification as the least toxic original ingredient and which is not expected to affect the toxicity of other ingredients, then the new mixture may be classified as equivalent to the original mixture.

Batching

389. The toxicity of one production batch of a complex mixture can be assumed to be substantially equivalent to that of another production batch of the same commercial product and produced by or under the control of the same manufacturer, unless there is reason to believe there is significant variation such that the toxicity of the batch has changed. If the latter occurs, new classification is necessary.

Concentration of Highly Toxic Mixtures

390. If in a mixture of Category 1, the concentration of a toxic ingredient is increased, the concentrated mixture should be classified in Category 1 without additional testing.

Interpolation within One Toxicity Category

391. If mixtures A and B are classified in the same toxicity category and mixture C is made in which the toxicologically active ingredients have concentrations intermediate to those in mixtures A and B, then mixture C is assumed to be in the same toxicity category as A and B. Note that the identity of the ingredients should be the same in all three mixtures.

Substantially Similar Mixtures

392. Given the following:

a). Two mixtures: (i) A + B
(ii) C + B
b). The concentration of ingredient B is essentially the same in both mixtures.
c). The concentration of ingredient A in mixture (i) equals that of ingredient C in mixture (ii)
d). Data on toxicity for A and C are available and substantially equivalent, i.e. they are in the same hazard category and are not expected to affect the toxicity of B.

If mixture (i) is already classified by testing, mixture (ii) can assigned the same category.

Aerosols

393. An aerosol form of a mixture may be classified in the same hazard category as the tested, non-aerosolised form of the mixture for oral and dermal toxicity provided the added propellant does not affect the toxicity of the mixture on spraying. Classification of aerosolised mixtures for inhalation toxicity should be considered separately.

CLASSIFICATION OF MIXTURES WHEN DATA ARE AVAILABLE FOR ALL INGREDIENTS OR ONLY FOR SOME INGREDIENTS OF THE MIXTURE.

394. Where there is no reliable evidence or test data for the specific mixture itself, and the bridging principles cannot be used to enable classification, then classification of the mixture is based on the classification of the ingredient substances. In this case, the mixture will be classified as a target organ/systemic toxicant (specific organ specified), following single exposure, repeat exposure, or both when at least one ingredient has been classified as a Category 1 or Category 2 target organ/systemic toxicant and is present at or above the appropriate cut-off value/concentration limit as mentioned in Table 15 below for Category 1 and 2 respectively.

Table 15: Cut-off values/concentration limits of ingredients of a mixture classified as a Target Organ/ Systemic Toxicant that would trigger classification of the mixture.[1]

Ingredient classified as:	Cut-off/concentration limits triggering classification of a mixture as:	
	Category 1 Target Organ Systemic Toxicant (TOST)	Category 2 Target Organ Systemic Toxicant (TOST)
Category 1 (TOST) Target Organ Systemic Toxicant	≥ 1.0 % (note 1)	1.0≤ ingredient < 10% (note 3)
	≥ 10 % (note 2)	1.0≤ ingredient < 10% (note 3)
Category 2 (TOST) Target Organ Systemic Toxicant		≥ 1.0 % (note 4)
		≥ 10 % (note 5)

Note 1: If a Category 1 target organ/systemic toxicant is present in the mixture as an ingredient at a concentration between 1.0% and 10%, every regulatory authority would require information on the SDS for a product. However, a label warning would be optional. Some authorities will choose to label when the ingredient is present in the mixture between 1.0% and 10%, whereas others would normally not require a label in this case.

Note 2: If a Category 1 target organ/systemic toxicant is present in the mixture as an ingredient at a concentration of ≥ 10%, both an SDS and a label would generally be expected.

Note 3: If a Category 1 target organ/systemic toxicant is present in the mixture as an ingredient at a concentration between 1.0% and 10%, some authorities classify this mixture as a Category 2 target organ/systemic toxicant, whereas others would not.

Note 4: If a Category 2 target organ/systemic toxicant is present in the mixture as an ingredient at a concentration between 1.0% and 10%, every regulatory authority would require information on the SDS for a product. However, a label warning would be optional. Some authorities will choose to label when the ingredient is present in the mixture between 1.0% and 10%, whereas others would normally not require a label in this case.

Note 5: If a Category 2 target organ/systemic toxicant is present in the mixture as an ingredient at a concentration of ≥ 10%, both an SDS and a label would generally be expected.

395. These cut-off values and consequent classifications should be applied equally and appropriately to both single- and repeated-dose target organ toxicants.

[1] This compromise classification scheme involves consideration of differences in hazard communication practices in existing systems. Although it is recognised that this may result in a lack of harmonisation for some mixtures, the OECD Expert Group is recommending to the ILO Hazard Communication Work Group that this compromise be accepted as a way to move the process forward. It is expected that the number of affected mixtures will be small; the differences will be limited to label warnings; and the situation will evolve over time to a more harmonised approach. All of these hazard communication recommendations are subject to review by the ILO Work Group, and may be affected by that group's determinations regarding the possibility of using risk considerations in labelling in the consumer sector.

396. Mixtures should be classified for either or both single- and repeated-dose toxicity independently.

397. Care should be exercised when toxicants affecting more than one organ system are combined that the potentiation or synergistic interactions are considered, because certain substances can cause target organ toxicity at <1% concentration when other ingredients in the mixture are known to potentiate its toxic effect.

CHAPTER 3.9

HARMONISED SYSTEM FOR THE CLASSIFICATION OF THE CHEMICAL MIXTURES WHICH ARE HAZARDOUS FOR THE AQUATIC ENVIRONMENT

GENERAL CONSIDERATIONS

398. The harmonised criteria for the classification of substances as "hazardous for the aquatic environment" are described in Part 2 , Chapter 2.10 of this document and were already endorsed by the 28th Joint Meeting of the Chemicals Committee and the Working Party on Chemicals in November 1998. The harmonised classification system for substances consists of three acute classification categories and four chronic classification categories. The acute and the chronic classification categories are applied independently. The criteria for classification of a substance in acute categories I to III are defined on the basis of the acute toxicity data only (EC_{50} or LC_{50}). The criteria for classification of a substance into chronic categories combine two types of information, i.e. acute toxicity data and environmental fate data (degradability and bioaccumulation data). For assignment of mixtures to chronic categories, degradation and bioaccumulation properties are derived from tests on components.

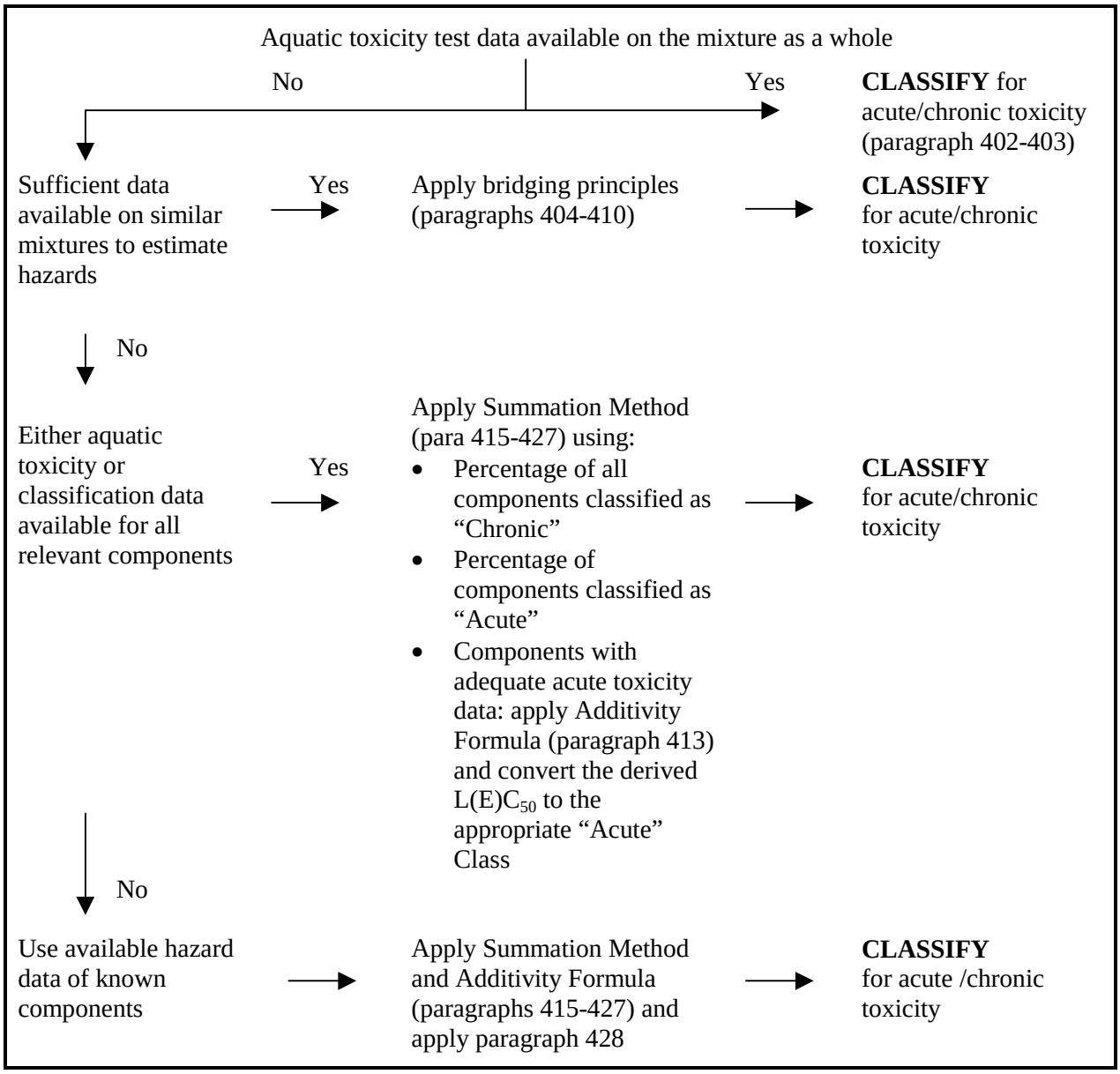
399. The classification system for mixtures covers all classification categories which are used for substances meaning acute categories I to III and chronic categories I to IV.

400. In order to make use of all available data for purposes of classifying the aquatic environmental hazards of the mixture, the following assumption has been made and is applied where appropriate.

> The "relevant components" of a mixture are those which are present in a concentration of 1% (w/w) or greater, unless there is a presumption (e.g. in the case of highly toxic components) that a component present at less than 1% can still be relevant for classifying the mixture for aquatic environmental hazards.

401. The approach for classification of aquatic environmental hazards is tiered, and is dependent upon the type of information available for the mixture itself and for its components. Elements of the tiered approach include: i) classification based on tested mixtures; ii) classification based on bridging principles, iii) the use of "summation of classifed components" and /or an "additivity formula". Figure 4 outlines the process to be followed.

Figure 4: Tiered Approach to Classification of Mixtures for Acute and Chronic Aquatic Environmental Hazards



CLASSIFICATION OF MIXTURES WHEN AQUATIC (TOXICITY) TEST DATA ARE AVAILABLE FOR THE COMPLETE MIXTURE.

402. When the mixture as a whole has been tested to determine its aquatic toxicity, it can be classified according to the criteria that have been agreed for substances, but only for acute toxicity. The classification should be based on the data from: fish, crustacea and algae/plants. Classification of mixtures by using LC_{50} or EC_{50} data for the mixture as a whole is not possible for chronic categories since both toxicity data and environmental fate data are needed, and there are no degradability and bioaccumulation data for mixtures as a whole. It is not possible to apply the criteria for chronic classification because the data from degradability and bio-accumulation tests of mixtures cannot be interpreted; they are meaningful only for single substances.

403. When there is acute toxicity test data (LC_{50} or EC_{50}) available for the mixture as a whole, this data as well as information with respect to the classification of components for chronic toxicity should be used to complete the classification for tested mixtures as follows. When chronic (long term) toxicity data (NOEC) is also available, this should be used as well.

- <u>L(E)C_{50} (LC_{50} or EC_{50}) of the tested mixture ≤ 100mg/L and NOEC of the tested mixture ≤ 1.0 mg/L or unknown:</u>
  → Classify mixture as Acute I, II or III
  → Apply Summation of Classified Components approach (see paragraphs 423-428) for chronic classification (Chronic I, II, III, IV or no need of chronic classification).
- <u>L(E)C_{50} of the tested mixture ≤ 100mg/L and NOEC of the tested mixture > 1.0 mg/L</u>:
  → Classify mixture as Acute I, II or III
  → Apply Summation of Classified Components approach (see paragraphs 423-428) for classification as Chronic I. If the mixture is not classified as Chronic I, then there is no need for chronic classification.
- <u>L(E)C_{50} of the tested mixture >100mg/L, or above the water solubility, and NOEC of the tested mixture ≤ 1.0mg/L or unknown</u>:
  → No need to classify for acute toxicity
  → Apply Summation of Classified Components approach (see paragraphs 423-428) for chronic classification (Chronic IV or no need for chronic classification).
- <u>L(E)C_{50} of the tested mixture >100mg/L, or above the water solubility, and NOEC of the tested mixture > 1.0 mg/L</u>
  → No need to classify for acute or chronic toxicity

CLASSIFICATION OF MIXTURES WHEN AQUATIC TEST DATA ARE NOT AVAILABLE FOR THE COMPLETE MIXTURE.

Bridging Principles

404. Where the mixture itself has not been tested to determine its aquatic environmental hazard, but there are sufficient data on the individual components and similar tested mixtures to adequately characterise the hazards of the mixture, this data will be used in accordance with the following

agreed bridging rules. This ensures that the classification process uses the available data to the greatest extent possible in characterising the hazards of the mixture without the necessity for additional testing in animals.

Dilution

405. If a mixture is formed by diluting another classified mixture or a substance with a diluent which has an equivalent or lower aquatic hazard classification than the least toxic original component and which is not expected to affect the aquatic hazards of other components, then the mixture may be classified as equivalent to the original mixture or substance.

406. If a mixture is formed by diluting another classified mixture or a substance with water or other totally non-toxic material, the toxicity of the mixture can be calculated from the original mixture or substance.

Batching

407. The aquatic hazard classification of one production batch of a complex mixture can be assumed to be substantially equivalent to that of another production batch of the same commercial product and produced by or under the control of the same manufacturer, unless there is reason to believe there is significant variation such that the aquatic hazard classification of the batch has changed. If the latter occurs, new classification is necessary.

Concentration of Mixtures which are classified with the most severe classification categories (Chronic I and Acute I)

408. If a mixture is classified as chronic I and/or acute I, and components of the mixture which are classified as chronic I and/or acute I are further concentrated, the more concentrated mixture should be classified with the same classification category as the original mixture without additional testing.

Interpolation within One Toxicity Category

409. If mixtures A and B are in the same classification category and mixture C is made in which the toxicologically active components have concentrations intermediate to those in mixtures A and B, then mixture C is assumed to be in the same category as A and B. Note that the identity of the components is the same in all three mixtures.

Substantially similar mixtures

410. Given the following:

a). Two mixtures: i.) A + B
ii.) C + B
b). The concentration of component B is the same in both mixtures.
c). The concentration of component A in mixture (i) equals that of component C in mixture (ii).
d). Classification for A and C are available and are the same, i.e. they are in the same hazard category and are not expected to affect the aquatic toxicity of B.

Then there is no need to test mixture (ii). If mixture (i) is already characterised by testing, mixture (ii) can be classified the same hazard category.

CLASSIFICATION OF MIXTURES BASED ON AQUATIC TEST DATA OR AVAILABLE CLASSIFICATION OF COMPONENTS.

411. The classification of a mixture is based on summation of the classification of its components. The percentage of components classified as "Acute" or "Chronic" will feed straight in to the summation method. Details of the summation method are described in paragraphs 416-428.

412. Mixtures can be made of a combination of both components that are classified (as Acute I, II, III and/or Chronic I, II, III, IV) and those for which adequate test data is available. When adequate toxicity data is available for more than one component in the mixture, the combined toxicity of those components may be calculated using the following additivity formula, and the calculated toxicity may be used to assign that portion of the mixture an acute toxicity category which is then subsequently used in applying the summation method.

$$\frac{\sum Ci}{L(E)C_{50m}} = \sum_{\eta} \frac{Ci}{L(E)C_{50i}}$$

where:

C_i = concentration of component i (weight percentage)
$L(E)C_{50i}$ = (mg/L) LC_{50} or EC_{50} for component i
η = number of components
$L(E)\ C_{50m}$ = $L(E)C_{50}$ of the part of the mixture with test data

413. When applying the additivity formula for part of the mixture, it is preferable to calculate the toxicity of this part of the mixture using for each substance toxicity values that relate to the same species (i.e.; fish, daphnia or algae) and then to use the highest toxicity (lowest value) obtained (viz., use the most sensitive of the three species). However, when toxicity data for each component are not available in the same species, the toxicity value of each component should be selected in the same manner that toxicity values are selected for the classification of substances, i.e. the higher toxicity (from the most sensitive test organism) is used. The calculated acute toxicity may then be used to classify this part of the mixture as Acute I, II or III using the same criteria described in the Harmonised Integrated System for pure substances.

414. If a mixture is classified in more than one way, the method yielding the more conservative result should be used.

Summation Method

Rationale

415. In case of the substance classification categories Acute I/Chronic I to Acute III/Chronic III, the underlying toxicity criteria differ by a factor of 10 in moving from one category to another. Substances with a classification in a high toxicity band may therefore contribute to the classification of a mixture in a lower band. The calculation of these classification categories therefore needs to consider the contribution of all substances classified Acute I/Chronic I to Acute III/Chronic III together.

416. When a mixture contains components classified as Acute Category I, attention should be paid to the fact that such components, when their acute toxicity is well below 1 mg/L (see also

paragraph 314), contribute to the toxicity of the mixture even at a low concentration. Active ingredients in pesticides often possess such high aquatic toxicity but also some other substances like organometallic compounds. Under these circumstances the application of the normal cut-off values/concentration limits may lead to an "underclassification" of the mixture. Therefore, multiplying factors should be applied to account for highly toxic components, as described in paragraph 427.

Classification Procedure

417. In general a more severe classification for mixtures overrides a less severe classification, e.g. a classification with Chronic I overrides a classification with Chronic II. As a consequence the classification procedure is already completed if the results of the classification is Chronic I. A more severe classification than chronic I is not possible therefore it is not necessary to undergo the further classification procedure.

Classification for the Acute Categories I, II and III

418. First all components classified as Acute I are considered. If the sum of these components is greater than 25% the whole mixture is classified as Category Acute I. If the result of the calculation is a classification of the mixture as Category Acute I, the classification process is completed.

419. In cases where the mixture is not classified as Acute I, classification of the mixture as Acute II is considered. A mixture is classified as Acute II if ten times the sum of all components classified as Acute I plus the sum of all components classified as Acute II is greater than 25%. If the result of the calculation is classification of the mixture as Category Acute II, the classification process is completed.

420. In cases where the mixture is not classified either as Acute I or Acute II, classification of the mixture as Acute III is considered. A mixture is classified as Acute III if 100 times the sum of all components classified as Acute I plus 10 times the sum of all components classified as Acute II plus the sum of all components classified as Acute III is greater than 25%.

421. The classification of mixtures for acute hazards based on this summation of classified components, is summarised in Table 16 below.

Table 16: Classification of a mixture for acute hazards, based on summation of classified components.

Sum of components classified as:		Mixture is classified as:
Acute I x M[1)]	>25%	Acute I
(M x 10 x Acute I) +Acute II	>25%	Acute II
(M x 100 x Acute I)+ (10 x Acute II) + Acute III	>25%	Acute III

1) for explanation of the M factor, see paragraph 427

Classification for the Chronic Categories I, II, III and IV

422. First all components classified as Chronic I are considered. If the sum of these components is greater than 25% the mixture is classified as Category Chronic I. If the result of the calculation is a classification of the mixture as Category Chronic I the classification procedure is completed.

423. In cases where the mixture is not classified as Chronic I, classification of the mixture as Chronic II is considered. A mixture is classified as Chronic II if 10 times the sum of all components classified as Chronic I plus the sum of all components classified as Chronic II is greater than 25%. If the result of the calculation is classification of the mixture as Chronic II, the classification process is completed.

424. In cases where the mixture is not classified either as Chronic I or Chronic II, classification of the mixture as Chronic III is considered. A mixture is classified as Chronic III if 100 times the sum of all components classified as Chronic I plus 10 times the sum of all components classified with Chronic II plus the sum of all components classified as Chronic III is greater than 25%.

425. If the mixture is still not classified in either Category Chronic I, II or III, classification of the mixture as Chronic IV should be considered. A mixture is classified as Chronic IV if the sum of the percentages of components classified as Chronic I, II, III and IV is greater than 25%.

426. The classification of mixtures for chronic hazards, based on this summation of classified components, is summarised in Table 17 below.

Table 17: Classification of a mixture for chronic hazards, based on summation of classified components.

Sum of components classified as:		Mixture is classified as:
Chronic I x M[1)]	>25%	Chronic I
(M x 10 x Chronic I)+Chronic II	>25%	Chronic II
(M x 100 x Chronic I)+(10x Chronic II)+Chronic III	>25%	Chronic III
Chronic I + Chronic II + Chronic III +Chronic IV	> 25%	Chronic IV

1) for explanation of the M factor, see paragraph 427

Mixtures with highly toxic components

427. Acute Category I components with toxicities well below 1 mg/L may influence the toxicity of the mixture and should be given increased weight in applying the summation of classification approach. When a mixture contains components classified as Acute or Chronic Category I, the tiered approach described in paragraphs 418-426 should be applied using a weighted sum by multiplying the concentrations of each Acute Category I components by a factor, instead of merely adding up the percentages. This means that the concentration of "Acute I"in the left column of Table 16 and the concentration of "Chronic I" in the left column of Table 17 are multiplied by the appropriate multiplying factor. The multiplying factors to be applied to these components are defined using the toxicity value, as summarised in Table 18 below. Therefore, in order to classify a mixture containing Acute/Chronic I components, the classifier needs to be informed of the value of the M factor in order to apply the summation method. Alternatively, the additivity formula (paragraph 412) may be used when toxicity data are available for all highly toxic components in the mixture and there is convincing evidence that all other components, including those for which specific acute toxicity data are not available, are of low or no toxicity and do not significantly contribute to the environmental hazard of the mixture.

Table 18: Multiplying factors for highly toxic components of mixtures

$L(E)C_{50}$ value	Multiplying factor (M)
$0.1 < L(E)C_{50} \leq 1$	1
$0.01 < L(E)C_{50} \leq 0.1$	10
$0.001 < L(E)C_{50} \leq 0.01$	100
$0.0001 < L(E)C_{50} \leq 0.001$	1000
$0.00001 < L(E)C_{50} \leq 0.0001$	10000
(continue in factor 10 intervals)	

CLASSIFICATION OF MIXTURES WITH COMPONENTS WITHOUT ANY USEABLE INFORMATION.

428. In the event that no useable information on acute and/or chronic aquatic toxicity is available for one or more relevant components, it is concluded that the mixture cannot be attributed (a) definitive hazard category(ies). In this situation the mixture should be classified based on the known components only, with the additional statement that: *"x percent of the mixture consists of components(s) of unknown hazards to the aquatic environment"*.

ANNEX 1

SCHEMATIC PRESENTATION OF THE HARMONISED INTEGRATED HAZARD CLASSIFICATION SYSTEM FOR CHEMICAL SUBSTANCES

ANNEX 1

SCHEMATIC PRESENTATION OF THE INTEGRATED CLASSIFICATION SYSTEM FOR HUMAN HEALTH AND ENVIRONMENTAL HAZARDS OF CHEMICAL SUBSTANCES

For the convenience and comparison of the various endpoints, the scheme and criteria for classifying each hazard are presented in the following diagram. The criteria have been drastically abridged and the end-point chapters must be consulted for the specific details to avoid misunderstanding.

ENDPOINT	HAZARD CATEGORIES AND CRITERIA				
ACUTE TOXICITY	Category 1	Category 2	Category 3	Category 4	Category 5
Oral (mg/kg)	5	50	300	2 000	5 000 (or equivalent doses for other routes) Criteria: • Indication of significant effect in human • Any mortality at Category 4 • Significant clinical signs at Category 4 • Indications from other studies
Dermal (mg/kg)	50	200	1 000	2 000	
Inhalation note 1					
gas (ppm)	100	500	2 500	5 000	
vapour (mg/L) note 2,3	0.5	2.0	10	20	
dust/mists (mg/L/4 hrs) note 4	0.05	0.5	1.0	5	

Note 1: Inhalation cut-off values are based on 4 hour testing exposures. Conversion of existing inhalation toxicity data which has been generated according to 1 hour exposures should be by dividing by a factor of 2 for gases and vapours and 4 for dusts and mists.

Note 2: Saturated vapour concentration may be used as an additional element to provide for specific health and safety.

Note 3: For some chemicals the test atmosphere will not just be a vapour but will consist of a mixture of liquid and vapour phases. For other chemicals the test atmosphere may consist of a vapour which is near the gaseous phase. In these latter cases, classification should be based on ppm as follows: Category 1 (100 ppm), Category 2 (500 ppm), Category 3 (2500 ppm), Category 4 (5000 ppm).

Note 4: The values for dusts and mists should be reviewed to adapt to any future changes to OECD Test Guidelines with respect to technical limitation in generating, maintaining and measuring dust and mist concentrations in respirable form.

ENDPOINT	HAZARD CATEGORIES AND CRITERIA

DERMAL IRRITATION/ CORROSION	Category 1			Category 2:	Category 3:
	Destruction of dermal tissue: visible necrosis in at least one animal			- Reversible adverse effects in dermal tissue	- Reversible adverse effects in dermal tissue
	Subcategory 1A Exposure ≤ 3 minutes Observation ≤ 1 hour	Subcategory 1B Exposure ≤ 1 hour Observation ≤ 14 days	Subcategory 1C Exposure ≤ 4 hours Observation ≤ 14 days	- Mean Draize score in 2 of 3 animals: 2.3 ≤erythema/eschar/ edema < 4.0, or - persistent inflammation	- Mean Draize score in 2 of 3 animals: 1.5 ≤ erythema/ eschar/ edema < 2.3

EYE IRRITATION/ CORROSION	Category 1 - Irreversible damage to cornea, iris, conjunctiva 21 days after exposure in at least one animal - mean Draize score in 2 of 3 animals: corneal opacity ≥ 3, iritis >1.5	Category 2 - reversible adverse effects on cornea, iris, conjuctiva - mean Draize score in 2 of 3 animals: - corneal opacity: ≥1,iritis: ≥1,redness ≥ 2,chemosis: ≥2	
		Subcategory 2A: reversible in 21 days	Subcategory 2B: reversible in 7 days

RESPIRATORY SENSITISATION	Category 1: - evidence of specific respiratory hypersensitivity, or - positive results from animal test

DERMAL SENSITISATION	Category 1: - evidence in humans of sensitisation by skin contact, or - positive results from animal tests

ENDPOINT	HAZARD CATEGORIES AND CRITERIA			
GERM CELL MUTAGENICITY	Category 1 known to produce heritable mutations in human germ cells		Category 2:	
	Subcategory 1A positive evidence from epidemiological studies	Subcategory 1B positive results in: - *in vivo* heritable germ cell tests in mammals - human germ cell tests - *in vivo* somatic mutagenicity tests, combined with some evidence of germ cell mutagenicity	- may induce heritable mutations in human germ cells - positive evidence from tests in mammals and somatic cell tests - *in vivo* somatic genotoxicity supported by *in vitro* mutagenicity	
CARCINOGENICITY	Category 1: Known or presumed carcinogen		Category 2: - suspected carcinogen - limited evidence of human or animal carcinogenicity	
	Subcategory 1A: known human carcinogen based on human evidence	Subcategory 1B: presumed human carcinogen based on demonstrated animal carcinogenicity		
REPRODUCTIVE TOXICITY	Category 1: known or presumed human reproductive or developmental toxicant		Category 2: suspected human reproductive or developmental toxicant	Additional Category effects on or via lactation
	Category 1A: known	Category 1B: presumed		

ENDPOINT	HAZARD CATEGORIES AND CRITERIA	
SPECIFIC TARGET ORGAN SYSTEMIC TOXICITY: **<u>SINGLE</u> EXPOSURE**	CATEGORY 1	CATEGORY 2
	Presumed to have the potential to produce significant toxicity • Reliable evidence from humans • Observations from animal studies • Expert judgement based on weight of evidence including the following guidance values of dose levels showing the effect: - oral ≤ 300 mg/kg/bw - dermal ≤ 1000 mg/kg/bw - inhalation (gas) ≤ 2500 ppm - inhalation (vapour) ≤ 10 mg/L - inhalation (dust/mist) ≤ 1.0 mg/L	Presumed to have the potential to be harmful • Observations from animal studies • Expect judgement based on weight of evidence including the following guidance values of dose level showing the effects - oral 2000 ≥c > 300 mg/L - dermal 2000 ≥c >1000 mg/L - inhalation (gas) 5000 ≥c >2500 ppm - inhalation (vapour) 20 ≥c > 10 mg/L - inhalation (dust/mist) 5 ≥c > 1.0 mg/L
SPECIFIC TARGET ORGAN SYSTEMIC TOXICITY: **<u>REPEATED</u> EXPOSURE**	CATEGORY 1	CATEGORY 2
	Presumed to have the potential to produce significant toxicity • Reliable evidence from humans • Observations from animal studies • Expert judgement based on weight of evidence including the following guidance values of dose levels showing the effect: - oral ≤ 10 mg/kg/bw - dermal ≤ 20 mg/kg/bw - inhalation (gas) ≤ 50 ppm - inhalation (vapour) ≤ 0.2 mg/L - inhalation (dust/mist) ≤ 0.02 mg/L	Presumed to have the potential to be harmful • Observations from animal studies • Expect judgement based on weight of evidence including the following guidance values of dose level showing the effects - oral 100 ≥c > 10 mg/L - dermal 200 ≥c > 20 mg/L - inhalation (gas) 250 ≥c > 50 ppm - inhalation (vapour) 110 ≥c > 0.2 mg/L - inhalation (dust/mist) 0.2 ≥c > 0.02 mg/L

ENDPOINT	HAZARD CATEGORIES AND CRITERIA

AQUATIC TOXICITY	Acute Category 1: acute toxicity ≤ 1.00mg/L	Acute Category 2: acute toxicity > 1.00 but ≤ 10.0mg/L	Acute Category 3: acute toxicity > 10.0 but ≤ 100mg/L	
	Chronic Category 1: acute toxicity ≤ 1.00mg/L and lack of rapid degradability and log Kow ≥ 4 unless BCF < 500	Chronic Category 2: acute toxicity > 1.00 but ≤ 10.0mg/L and lack of rapid degradability and log Kow ≥ 4 unless BCF < 500 and unless chronic toxicity > 1 mg/L	Chronic Category 3: acute toxicity > 10.0 but ≤ 100mg/L and lack of rapid degradability and log Kow ≥ 4 unless BCF < 500 and unless chronic toxicity > 1mg/L	Chronic Category 4: acute toxicity > 100 mg/L and lack of rapid degradability and log Kow ≥ 4 unless BCF < 500 and unless chronic toxicity > 1mg/L

ANNEX 2

OECD GUIDANCE DOCUMENT No. 27
GUIDANCE DOCUMENT ON THE USE OF THE HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICALS WHICH ARE HAZARDOUS FOR THE AQUATIC ENVIRONMENT

OECD Environment, Health and Safety Publications

Series on Testing and Assessment

No. 27

GUIDANCE DOCUMENT ON THE USE OF THE HARMONISED SYSTEM FOR THE CLASSIFICATION OF CHEMICALS WHICH ARE HAZARDOUS FOR THE AQUATIC ENVIRONMENT

Environment Directorate

Organisation for Economic Co-operation and Development

April 2001

Glossary of important terms used in the Guidance Document [1)]

Substance [2)]	Chemical elements and their compounds in the natural state or obtained by any production process, including any additive necessary to preserve the stability of the product and any impurities deriving from the process used, but excluding any solvent which may be separated without affecting the stability of the substances or changing its composition.
Mixture [2)]	Mixtures or solutions composed of two or more substances in which they do not react.
Multi-component substances or Complex substances [3)]	Mixtures comprising a complex mix of individual substances with different solubilities and physico-chemical properties. In most cases, they can be characterised as a homologous series of substances with a certain range of carbon chain length/number or degree of substitution. These materials are frequently referred to as "complex mixtures". But, in this Guidance Document, these are referred to as "multi-component substances".
Geometric mean of the effect concentrations	Antilog of the mean of the log-transformed effect concentrations.
Availability	Availability is the extent to which a substance becomes a soluble or disaggregate species. For metals availability is the extent to which the metal ion portion of a metal (M^O) compound can disaggregate from the rest of compound (molecule).
Bioavailability	Extent to which a substance is taken up by an organism, and distributed to an area within the organism. It is dependent upon: physicochemical properties of the substance; anatomy and physiology of the organism; pharmacokinetics; and route of exposure. Availability is not a prerequisite for bioavailability.
Acute toxicity	Intrinsic property of a substance to be injurious to an organism in a short-term exposure to that substance.
Chronic Toxicity	Potential or actual properties of a substance to cause adverse effects to aquatic organisms during exposures which are determined in relation to the life-cycle of the organism.
Degradation	Decomposition of organic molecules to smaller molecules and eventually to carbon dioxide, water and salts.
Bioaccumulation	Net result of uptake, transformation, and elimination of a substance in an organism due to all routes of exposure (i.e., via air, water, sediment/soil, and food).
Bioconcentration	Net result of uptake, transformation, and elimination of a substance in an organism due to waterborne exposure.

Note 1. All terms and their description should be considered as working definitions for the purpose of this Guidance Document only.

Note 2. The definition is cited from a paper (ENV/JM/HCL(99)11), entitled "Step 2 proposal for Harmonised Classification Criteria for Mixtures" and therefore considered as a provisional definition.

Note 3. Consideration is given to the consistency with the definition of "multi-component substances" (or "complex substances") in Draft Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures.

1. INTRODUCTION

1. As part of a wider international effort on the global harmonisation of hazard classification systems, agreement was reached in technical working groups on a set of criteria that would form the basis of a global scheme for identifying substances hazardous to the aquatic environment. Such a scheme forms part of an international agreement on hazard classification of substances. The criteria were endorsed by the Joint Meeting of the OECD in November 1998 and form part of the Globally Harmonised Classification System (GHS) which is expected to be implemented under ECOSOC in 2001 (see Appendix). In developing the criteria, it was agreed that the detail needed to properly define the hazard to the environment resulted in a complex system for which some suitable guidance would be necessary. The harmonised proposal makes a number of references to a Guidance Document in the detailed explanation of the scheme. The purpose of this document is therefore twofold:

- to provide a description of and guidance to how the system will work
- to provide a guidance to the interpretation of data for use in applying the classification criteria

2. The hazard classification scheme has been developed with the object of identifying those chemical substances that present, through the intrinsic properties they possess, a danger to the aquatic environment. In this context, the aquatic environment is taken as the aquatic ecosystem in freshwater and marine, and the organisms that live in it. For most substances, the majority of data available addresses this environmental compartment. The definition is limited in scope in that it does not, as yet, include aquatic sediments, nor higher organisms at the top end of the aquatic food-chain, although these may to some extent be covered by the criteria selected.

3. Although limited in scope, it is widely accepted that this compartment is both vulnerable, in that it is the final receiving environment for many harmful substances, and the organisms that live there are sensitive. It is also complex since any system that seeks to identify hazards to the environment must seek to define those effects in terms of wider effects on ecosystems rather than on individuals within a species or population. As will be described in detail in the subsequent chapters, a limited set of specific properties of chemical substances have been selected through which the hazard can be best described: aquatic toxicity; lack of degradability; and potential or actual bioaccumulation. The rationale for the selection of these data as the means to define the aquatic hazard will be described in more detail in Chapter 2.

4. The application of the criteria is also limited, at this stage, to chemical substances. The term substances covers a wide range of chemicals, many of which pose difficult challenges to a classification system based on rigid criteria. The following chapters will thus provide some guidance as to how these challenges can be dealt with based both on experience in use and clear scientific rationale. A substance, in this context, is defined in the Step 2 Proposal for Harmonised Classification Criteria for Mixtures (ENV/JM/HCL(99)11) as “chemical elements and their compounds in the natural state or obtained by any production process, including any additive necessary to preserve the stability of the product and any impurities deriving from the process used, but excluding any solvent which may be separated without affecting the stability of the substance or changing its composition”. While the harmonised criteria apply most easily to the classification of individual substances of defined structure, some materials that fall under this definition are frequently referred to as “complex mixtures”. In most cases they can be characterised as a homologous series of substances with a certain range of carbon chain length/number or degree of substitution. Special methodologies have been developed for testing which provides data for evaluating the intrinsic hazard to aquatic organisms, bioaccumulation and degradation. More

specific guidance is provided in the separate chapters on these properties. For the purpose of this Guidance Document, these materials will be referred to as "complex substances" or "multi-component substances".

5. While aspects of the criteria can potentially be applied to chemical mixtures, the interpretation of test data is often complex and ambiguous and it is possible that another method of classification, such as a calculation based on the component substances may be preferred. The basis of a harmonised approach to the classification of mixtures is still under discussion and thus, while the criteria should form the basis of future decision making, it is not felt that they can or should be applied directly to mixtures at this time.

6. Each of these properties (i.e., aquatic toxicity, degradability, bioaccumulation) can present a complex interpretational problem, even for experts. While internationally agreed testing guidelines exist and should be used for any and all new data produced, many data usable in classification will not have been generated according to such standard tests. Even where standard tests have been used, some substances, such as complex substances, hydrolytically unstable substances, polymers etc, present difficult interpretational problems when the results have to be used within the classification scheme. Thus data are available for a wide variety of both standard and non-standard test organisms, both marine and freshwater, of varying duration and utilising a variety of endpoints. Degradation data may be biotic or abiotic and can vary in environmental relevance. The potential to bioaccumulate can, for many organic chemicals, be indicated by the octanol-water partition coefficient. It can however be affected by many other factors and these will also need to be taken into account.

7. It is clearly the objective of a globally harmonised system that, having agreed on a common set of criteria, a common data-set should also be used so that once classified, the classification is globally accepted. For this to occur, there must first be a common understanding of the type of data that can be used in applying the criteria, both in type and quality, and subsequently a common interpretation of the data when measured against the criteria. For that reason, it has been felt necessary to develop a transparent guidance document that would seek to expand and explain the criteria in such a way that a common understanding of their rationale and a common approach to data interpretation may be achieved. This is of particular importance since any harmonised system applied to the "universe of chemicals" will rely heavily on self-classification by manufacturers and suppliers, classifications that must be accepted across national boundaries without always receiving regulatory scrutiny. This guidance document, therefore, seeks to inform the reader, in a number of key areas, and as a result lead to classification in a consistent manner, thus ensuring a truly harmonised and self-operating system.

8. Firstly, it will provide a detailed description of the criteria, a rationale for the criteria selected, and an overview of how the scheme will work in practice (Chapter 2). This chapter will address the common sources of data, the need to apply ~~a~~ quality criteria, how to classify when the data-set is incomplete or when a large data-set leads to an ambiguous classification, and other commonly encountered classification problems.

9. Secondly, the guidance will provide detailed expert advice on the interpretation of data derived from the available databases, including how to use non-standard data, and specific quality criteria that may apply for individual properties. The problems of data interpretation for "difficult substances", those substances for which standard testing methods either do not apply or give difficult interpretational problems, will be described and advice provided on suitable solutions. The emphasis will be on data interpretation rather than testing since the system will, as far as possible, rely on the best available existing data and data required for regulatory purposes. The three core

properties, aquatic toxicity (Chapter 3), degradability (Chapter 4) and bioaccumulation (Chapter 5) are treated separately.

10. The range of interpretational problems can be extensive and as a result such interpretation will always rely on the ability and expertise of the individuals responsible for classification. However, it is possible to identify some commonly occurring difficulties and provide guidance that distils accepted expert judgement that can act as an aid to achieving a reliable and consistent result. Such difficulties can fall into a number of overlapping issues:

a) The difficulty in applying the current test procedures to a number of types of substance.
b) The difficulty in interpreting the data derived both from these "difficult to test" substances and from other substances.
c) The difficulty in interpretation of diverse data-sets derived from a wide variety of sources.

11. For many organic substances, the testing and interpretation of data present no problems when applying both the relevant OECD Guideline and the classification criteria. There are a number of typical interpretational problems, however, that can be characterised by the type of substance being studied. These are commonly called "difficult substances":

- poorly soluble substances: these substances are difficult to test because they present problems in solution preparation, and in concentration maintenance and verification during aquatic toxicity testing. In addition, many available data for such substances have been produced using "solutions" in excess of the water solubility resulting in major interpretational problems in defining the true $L(E)C_{50}$ for the purposes of classification. Interpretation of the partitioning behaviour can also be problematic where the poor solubility in water and octanol may be compounded by insufficient sensitivity in the analytical method. Water solubility may be difficult to determine and is frequently recorded as simply being less than the detection limit, creating problems in interpreting both aquatic toxicity and bioaccumulation studies. In biodegradation studies, poor solubility may result in low bioavailability and thus lower than expected biodegradation rates. The specific test method or the choice of procedures used can thus be of key importance.

- unstable substances: substance that degrade (or react) rapidly in the test system again present both testing and interpretational problems. It will be necessary to determine whether the correct methodology has been used, whether it is the substance or the degradation/reaction product that has been tested, and whether the data produced is relevant to the classification of the parent substance.

- volatile substances: such substances that can clearly present testing problems when used in open systems should be evaluated to ensure adequate maintenance of exposure concentrations. Loss of test material during biodegradation testing is inevitable in certain methods and will lead to misinterpretation of the results.

- complex or multi-component substances: such substances, for example, hydrocarbon mixtures, frequently cannot be dissolved into a homogeneous solution, and the multiple components make monitoring impossible. Consideration therefore needs to be given to using the data derived from the testing of water accommodated fractions (WAFs) for aquatic toxicity, and the utilisation of such data in the classification scheme. Biodegradation, bioaccumulation, partitioning behaviour and water solubility

all present problems of interpretation, where each component of the mixture may behave differently.

- polymers: such substances frequently have a wide range of molecular masses, with only a fraction being water soluble. Special methods are available to determine the water soluble fraction and these data will need to be used in interpreting the test data against the classification criteria.

- inorganic compounds and metals: such substances, which can interact with the media, can produce a range of aquatic toxicities dependant on such factors as pH, water hardness etc. Difficult interpretational problems also arise from the testing of essential elements that are beneficial at certain levels. For metals and inorganic metal compounds, the concept of degradability as applied to organic compounds has limited or no meaning. Equally the use of bioaccumulation data should be treated with care.

- surface active substances: such substances can form emulsions in which the bioavailablity is difficult to ascertain, even with careful solution preparation. Micelle formation can result in an overestimation of the bioavailable fraction even when "solutions" are apparently formed. This presents significant problems of interpretation in each of the water solubility, partition coefficient, bioaccumulation and aquatic toxicity studies.

- ionizable substances: such substances can change the extent of ionization according to the level of counter ions in the media. Acids and bases, for example, will show radically different partitioning behaviour depending on the pH.

- coloured substances: such substance can cause problems in the algal/aquatic plant testing because of the blocking of incident light.

- impurities: some substances can contain impurities that can change in % and in chemical nature between production batches. Interpretational problems can arise where either or both the toxicity and water solubility of the impurities are greater than the parent substance, thus potentially influencing the toxicity data in a significant way.

12. These represent some of the problems encountered in establishing the adequacy of data, interpreting the data and applying that data to the classification scheme. Detailed guidance on how to deal with these problems, as well as other issues related will be presented in the following Chapters. The interpretation of data on aquatic toxicity will be covered in Chapter 3. This chapter will deal with the specific interpretational problems encountered for the above "difficult substances", including providing some advice on when and how such data can be used within the classification scheme. Also covered will be a general description of the test data used and the testing methodologies suitable for producing such data.

13. A wide range of degradation data are available that must be interpreted according to the criteria for rapid degradability. Guidance is thus needed on how to use these data obtained by employing non-standard test methods, including the use of half-lives where these are available, of primary degradation, of soil degradation rates and their suitability for extrapolation to aquatic degradation and of environmental degradation rates. A short description of estimation techniques for evaluating degradability in relation to the classification criteria is also included. This guidance will be provided in Chapter 4.

14. Methods by which the potential to bioaccumulate can be determined will be described in Chapter 5. This chapter will describe the relationship between the partition coefficient criteria and the bioconcentration factor (BCF), provide guidance on the interpretation of existing data, how to estimate the partition coefficient by the use of QSARs when no experimental data are available and in particular deal with the specific problems identified above for difficult substances. The problems encountered when dealing with substances of high molecular mass will also be covered.

15. A chapter is also included which covers general issues concerning the use of QSARs within the system, when and how they may be used, for each of the three properties of concern. As a general approach, it is widely accepted that experimental data should be used rather than QSAR data when such data are available. The use of QSARs will thus be limited to such times when no reliable data are available. Not all substances are suitable for the application of QSAR estimations, however, and the guidance in Chapter 6 will address this issue.

16. Finally, a chapter is devoted to the special problems associated with the classification of metals and their compounds. Clearly, for these compounds, a number of the specific criteria such as biodegradability and octanol-water partition coefficient cannot be applied although the principle of lack of destruction via degradation, and bioaccumulation remain important concepts. Thus it is necessary to adopt a different approach. Metals and metal compounds can undergo interactions with the media which affect the solubility of the metal ion, partitioning from the water column, and the species of metal ion that exists in the water column. In the water column, it is generally the dissolved metal ions which are of concern for toxicity. The interaction of the substance with the media may either increase or decrease the level of ions and hence toxicity. It is thus necessary to consider whether metal ions are likely to be formed from the substance and dissolve in the water, and if so whether they are formed rapidly enough to cause concern. A scheme for interpreting the results from this type of study is presented in Chapter 7.

17. While the Guidance Document provides useful advice on how to apply the criteria to a wide variety of situations, it remains a guidance only. It cannot hope to cover all situations that arise in classification. It should therefore be seen as a living document that in part describes the fundamental principles of the system, e.g., hazard based rather than risk based, and the fixed criteria. It must also, in part, be a repository for the accumulated experience in using the scheme to include the interpretations which allow the apparently fixed criteria to be applied in a wide variety of non-standard situations.

2. THE HARMONIZED CLASSIFICATION SCHEME

2.1 SCOPE

18. The criteria were developed taking into account existing systems for hazard classification, such as EU- Supply and Use System, the Canadian and US Pesticide systems, GESAMP hazard evaluation procedure, IMO Scheme for Marine Pollutant, the European Road and Rail Transport Scheme (RID/ADR), and the US Land Transport. These systems include supply and subsequent use of chemicals, the sea transport of chemical substances as well as transport of chemical substances by road and rail. The harmonised criteria are therefore intended to identify hazardous chemicals in a common way for use throughout all these systems. To address the needs for all different sectors (transport and supply and use) it was necessary to create two different classification categories, one acute category, consisting of three categories and one chronic category, consisting of 4 categories. The acute classification category makes provision for two acute hazard categories (acute II and III) not normally used when considering packaged goods. For substances transported in bulk, there are a number of regulatory decisions that can uniquely arise because of the bulk quantities being considered. For these situations, for example where decisions are required on the ship type to be used, consideration of all acute classification categories as well as the chronic classification categories are considered important. The following paragraphs describe in detail the criteria to be used in defining each of these hazard categories.

2.2 CLASSIFICATION CATEGORIES AND CRITERIA

19. The hazard categories have been defined, according to the criteria set out below.

2.2.1 Acute toxicity

Category: Acute I

Acute toxicity:

96 hr LC_{50} (for fish)	≤1 mg/L and/or
48 hr EC_{50} (for crustacea)	≤1 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	≤1 mg/L.

Category: Acute I may be subdivided for some regulatory systems to include a lower band at $L(E)C_{50}$ ≤0.1 mg/L.

Category: Acute II

Acute toxicity:

96 hr LC_{50} (for fish)	>1 - ≤10 mg/L and/or
48 hr EC_{50} (for crustacea)	>1 - ≤10 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	>1 - ≤10 mg/L.

Category: Acute III

Acute toxicity:

96 hr LC_{50} (for fish)	>10 - ≤100 mg/L and/or
48 hr EC_{50} (for crustacea)	>10 - ≤100 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	>10 - ≤100 mg/L.

Some regulatory systems may extend this range beyond an $L(E)C_{50}$ of 100 mg/L through the introduction of another category.

2.2.2 Chronic toxicity

Category: Chronic I

Acute toxicity:

96 hr LC_{50} (for fish)	≤1 mg/L and/or
48 hr EC_{50} (for crustacea)	≤1 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	≤1 mg/L

and the substance is not rapidly degradable and/or the log Kow ≥ 4 (unless the experimentally determined BCF <500).

Category: Chronic II

Acute toxicity

96 hr LC_{50} (for fish)	>1 to ≤10 mg/L and/or
48 hr EC_{50} (for crustacea)	>1 to ≤10 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	>1 to ≤10 mg/L

and the substance is not rapidly degradable and/or the log Kow ≥4 (unless the experimentally determined BCF <500), unless the chronic toxicity NOECs are > 1 mg/L.

Category: Chronic III

Acute toxicity:

96 hr LC_{50} (for fish)	>10 to ≤100 mg/L and/or
48 hr EC_{50} (for crustacea)	>10 to ≤100 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	>10 to ≤100 mg/L

and the substance is not rapidly degradable and/or the log Kow ≥4 (unless the experimentally determined BCF <500) unless the chronic toxicity NOECs are >1 mg/L.

Category: Chronic IV

Poorly soluble substances for which no acute toxicity is recorded at levels up to the water solubility, and which are not rapidly degradable and have a log Kow ≥ 4, indicating a potential to bioaccumulate, will be classified in this category unless other scientific evidence exists showing classification to be unnecessary. Such evidence would include an experimentally determined BCF <500, or a chronic toxicity NOECs >1 mg/L, or evidence of rapid degradation in the environment.

2.3 RATIONALE

20. The harmonised system for classification recognises that the intrinsic hazard to aquatic organisms is represented by both the acute and chronic or longer-term toxicity of a substance, the relative importance of which is determined by the specific regulatory regimes in operation. Distinction can be made between the acute hazard and the chronic hazard and therefore hazard categories are defined for both properties representing a gradation in the level of hazard identified. Clearly the hazard identified by Chronic I is more severe than Chronic II. Since the acute hazard and chronic hazard represent distinct types of hazard, they are not comparable in terms of their relative severity. Both hazard classed should be applied independently for the classification of substances to establish a basis for all regulatory systems.

21. The principal hazard bands defined by the criteria relate largely to the potential for chronic hazard. This reflects the overriding concern with respect to chemicals in the environment, namely that the effects caused are usually sub-lethal, e.g., effects on reproduction, and caused by longer-term exposure. While recognising that the chronic hazard represents the principal concern,

particularly for packaged goods where environmental release would be limited in scope, it must also be recognised that chronic toxicity data are expensive to generate and generally not readily available for most substances. On the other hand, acute toxicity data are frequently readily available, or can be generated to highly standardised protocols. It is this acute toxicity which has therefore been used as the core property in defining both the acute and the chronic hazard. Nevertheless, it has been recognised that, where chronic toxicity data are available, it should be possible to use these in defining the appropriate hazard band. The development of specific criteria using such data is thus a high priority in the future development of the scheme.

22. While recognising that acute toxicity itself is not a sufficiently accurate predictor of chronic toxicity to be used solely and directly for establishing hazard, it is considered that, in combination with either a potential to bioaccumulate (i.e., a log K_{ow} ≥4 unless BCF <500) or potential longer-term exposure (i.e., lack of rapid degradation) it can be used as a suitable surrogate for classification purposes. Substances that show acute toxicity and also bioaccumulate to a significant degree will normally show chronic toxicity at a significantly lower concentration. Precise acute: chronic ratios are difficult to predict and thus the surrogate data are generally precautionary. Equally substances that do not rapidly degrade have a higher potential for giving rise to longer term exposures which again may result in long-term toxicity being realised. Thus, for example, Category Chronic I should be assigned if either of the following criteria are met:

i) L(E)C_{50} for any appropriate aquatic species ≤1 mg/l and a potential to bioaccumulate (log Kow ≥4 unless BCF <500).

ii) L(E)C_{50} for any appropriate aquatic species ≤1 mg/l and a lack of rapid degradation.

23. The precise definitions of acute toxicity of an appropriate species, lack of rapid degradation and potential to bioaccumulate are detailed in Chapters 3, 4 and 5 respectively.

24. For some poorly soluble substances, which are normally considered as those having a water solubility < 1 mg/l, no acute toxicity is expressed in toxicity tests performed at the solubility limit. If for such a substance, however, the BCF ≥ 500, or if absent, the log K_{ow} ≥ 4 (indicating a bioaccumulating potential) and the substance is also not rapidly degradable, a safety net classification is applied, Chronic Category IV. For these types of substance the exposure duration in short term tests may well be too short for a steady state concentration of the substance to be reached in the test organisms. Thus, even though no acute toxicity has been measured in a short term (acute) test, it remains a real possibility that such non-rapidly degradable and bioaccumulative substances may exert chronic effects, particularly since such low degradability may lead to an extended exposure period in the aquatic environment.

25. In defining acute aquatic toxicity, it is not possible to test all species present in an aquatic ecosystem. Representative species are therefore chosen which cover a range of trophic levels and taxonomic groupings. The taxa chosen, fish, crustacea and aquatic plants that represent the "base-set" in most hazard profiles, represent a minimum data-set for a fully valid description of hazard. The lowest of the available toxicity values will normally be used to define the hazard category. Given the wide range of species in the environment, the three tested can only be a poor surrogate and the lowest value is therefore taken for cautious reasons to define the hazard band. In doing so, it is recognised that the distribution of species sensitivity can be several orders of magnitude wide and that there will thus be both more and less sensitive species in the environment. Thus, when data are limited, the use of the most sensitive species tested gives a cautious but acceptable definition of the hazard. There are some circumstances where it may not be appropriate to use the lowest toxicity value as the basis for classification. This will usually only arise where it is possible to define the

sensitivity distribution with more accuracy than would normally be possible, such as when large data-sets are available. Such large data-sets should be evaluated with due caution.

2.4 APPLICATION

26. Generally speaking, in deciding whether a substance should be classified, a search of appropriate databases and other sources of data should be made for the following data elements:

- water solubility
- octanol/water partition coefficient (log K_{ow})
- fish bioconcentration factor (BCF)
- acute aquatic toxicity (L(E)C_{50}s)
- chronic aquatic toxicity (NOECs)
- available degradation (and specifically evidence of ready biodegradability)
- stability data, in water

The water solubility and stability data, although not used directly in the criteria, are nevertheless important since they are a valuable help in the data interpretation of the other properties (see para 11).

27. To classify, a review should first be made of the available aquatic toxicity data. It will be necessary to consider all the available data and select those which meet the necessary quality criteria for classification. If there are no data available that meet the quality criteria required by the internationally standardised methods, it will be necessary to examine any available data to determine whether a classification can be made. If the data indicate that the acute aquatic toxicity L(E)C_{50} >100 mg/l for soluble substances, then the substance is not classified as hazardous. There are a number of cases where no effects are observed in the test and the aquatic toxicity is thus recorded as a >water solubility value, i.e., there is no acute toxicity within the range of the water solubility in the test media. Where this is the case, and the water solubility in the test media is ≥1 mg/l, again, no classification need be applied.

28. Where the lowest aquatic toxicity data are below 100 mg/l, it is necessary to first decide which hazard band the toxicity falls in, and then to determine whether the chronic and/or the acute category should be applied. This can simply be achieved by examining the available data on the partition coefficient, log K_{ow} and the available data on degradation. If either the log K_{ow}≥4 or the substance cannot be considered as rapidly degradable, then the appropriate chronic hazard category and the corresponding acute category are applied independently. It should be noted that, although the log K_{ow} is the most readily available indication of a potential to bioaccumulate, an experimentally derived BCF is preferred. Where this is available, this should be used rather than the partition coefficient. In these circumstances, a BCF ≥500 would indicate bioaccumulation sufficient to classify in the appropriate chronic hazard category. If the substance is both rapidly degradable and has a low potential to bioaccumulate (BCF <500 or, if absent log K_{ow} <4) then it should not be assigned to a chronic hazard band, only the acute hazard bands need be applied (see para 18).

29. For poorly soluble substances, generally speaking, those with a water solubility in the test media of <1 mg/l, for which no aquatic toxicity has been found, should be further examined to determine whether chronic category IV need be applied. Thus, if the substance is both not rapidly degradable and has a potential to bioaccumulate (BCF ≥500 or, if absent log K_{ow} ≥4), the chronic category IV should be applied.

2.5 DATA AVAILABILITY

30. The data used to classify a substance can be drawn from data required for regulatory purposes as well as the relevant literature, although a number of internationally recognised data-bases exist which can act as a good starting point. Such databases vary widely in quality and comprehensiveness and it is unlikely that any one database will hold all he information necessary for classification to be made. Some databases specialise in aquatic toxicity and others in environmental fate. There is an obligation on the chemical supplier to make the necessary searches and checks to determine the extent and quality of the data available and to use it in assigning the appropriate hazard band.

2.6 DATA QUALITY

31. The precise use of the available data will be described in the relevant chapter but, as a general rule, data generated to standard international guidelines and to GLP is to be preferred over other types of data. Equally, however, it is important to appreciate that classification can be made based on the best available data. Thus if no data is available which conforms to the quality standard detailed above, classification can still be made provided the data used is not considered invalid. To assist this process, a quality scoring guide has been developed and used extensively in a number of fora and generally conforms to the following categories:

1. Data derived from official data sources that have been validated by regulatory authorities, such as EU Water Quality Monographs, USEPA Water Quality Criteria. These data can be considered as valid for classification purposes. No assumption should be made that these are the only data available, however, and due regard should be given to the date of the relevant report. Newly available data may not have been considered.

2. Data derived from recognised international guidelines (e.g., OECD Guidelines) or national guidelines of equivalent quality. Subject to the data interpretation issues raised in the following chapters, these data can be used for classification.

3. Data derived from testing which, while not strictly according to a guideline detailed above, follows accepted scientific principles and procedures and/or has been peer reviewed prior to publication. For such data, where all the experimental detail is not recorded, some judgement may be required to determine validity. Normally, such data may be used within the classification scheme.

4. Data derived from testing procedures which deviate significantly from standard guidelines and are considered as unreliable, should not be used in classification.

5. QSAR data. The circumstances of use and validity of QSAR data are discussed in the relevant chapters.

6. Data derived from secondary sources such as handbooks, reviews, citation, etc where the data quality cannot be directly evaluated. Such data should be examined where data from quality 1,2 and 3 are not available, to determine whether it can be used. Such data should have sufficient detail to allow quality to be assessed. In determining the acceptability of these data for the purposes of classification, due regard should be given to the difficulties in testing that may have affected data quality and the

significance of the reported result in terms of the level of hazard identified (see para 76).

32. Classification may also be made on incomplete toxicity data-sets, e.g., where data are not available on all three trophic levels. In these cases, the classification may be considered as 'provisional' and subject to further information becoming available. In general, all the data available will need to be considered prior to assigning a classification. Where good quality data are not available, lower quality data will need to be considered. In these circumstances, a judgement will need to be made regarding the true level of hazard. For example, where good quality data are available for a particular species or taxa, this should be used in preference to any lower quality data which might also be available for that species or taxa. However, good quality data may not always be available for all the basic data set trophic levels. It will be necessary to consider data of lower quality for those trophic levels for which good quality data are not available. Consideration of such data, however, will also need to consider the difficulties that may have affected the likelihood of achieving a valid result. For example, the test details and experimental design may be critical to the assessment of the usability of some data, such as that from hydrolytically unstable chemicals, while less so for other chemicals. Such difficulties are described further in Chapter 3.

33. Normally, the identification of hazard, and hence the classification will be based on information directly obtained from testing of the substance being considered. There are occasions, however, where this can create difficulties in the testing or the outcomes do not conform to common sense. For example, some chemicals, although stable in the bottle, will react rapidly (or slowly) in water giving rise to degradation products that may have different properties. Where such degradation is rapid, the available test data will frequently define the hazard of the degradation products since it will be these that have been tested. These data may be used to classify the parent substance in the normal way. However, where degradation is slower, it may be possible to test the parent substance and thus generate hazard data in the normal manner. The subsequent degradation may then be considered in determining whether an acute or chronic hazard category should apply. There may be occasions, however, when a substance so tested may degrade to give rise to a more hazardous product. In these circumstances, the classification of the parent should take due account of the hazard of the degradation product, and the rate at which it can be formed under normal environmental conditions.

3. AQUATIC TOXICITY

3.1 INTRODUCTION

34. The basis for the identification of hazard to the aquatic environment for a substance is the aquatic toxicity of that substance. Classification is predicated on having toxicity data for fish, crustacea, and algae/aquatic plant available. These taxa are generally accepted as representative of aquatic fauna and flora for hazard identification. Data on these particular taxa are more likely to be found because of this general acceptance by regulatory authorities and the chemical industry. Other information on the degradation and bioaccumulation behaviour is used to better delineate the aquatic hazard. This chapter describes the appropriate tests for ecotoxicity, provides some basic concepts in evaluating the data and using combinations of testing results for classification, summarises approaches for dealing with difficulty substances, and includes a brief discussion on interpretation of data quality.

3.2 DESCRIPTION OF TESTS

35. For classifying substances in the harmonized system, freshwater and marine species toxicity data can be considered as equivalent data. It should be noted that some types of substances, e.g.,

ionizable organic chemicals or organometallic substances may express different toxicities in freshwater and marine environments. Since the purpose of classification is to characterise hazard in the aquatic environment, the result showing the highest toxicity should be chosen.

36. The GHS criteria for determining health and environmental hazards should be test method neutral, allowing different approaches as long as they are scientifically sound and validated according to international procedures and criteria already referred to in existing systems for the endpoints of concern and produce mutually acceptable data. According to the proposed system (OECD 1998):

"Acute toxicity would normally be determined using a fish 96 hour LC50 (OECD Test Guideline 203 or equivalent), a crustacea species 48 hour EC50 (OECD Test Guideline 202 or equivalent) and/or an algal species 72 or 96 hour EC50 (OECD Test Guideline 201 or equivalent). These species are considered as surrogate for all aquatic organisms and data on other species such as the duckweed *Lemna* may also be considered if the test methodology is suitable. "

Chronic testing involves an exposure that is lingering or continues for a longer time; the term can signify periods from days to a year, or more depending on the reproductive cycle of the aquatic organism. Chronic tests can be done to assess certain endpoints relating to growth, survival, reproduction and development.

"Chronic toxicity data are less available than acute data and the range of testing procedures less standardised. Data generated according to the OECD Test Guidelines 210 (Fish Early Life Stage), 202 Part 2 or 211 (Daphnia Reproduction) and 201 (Algal Growth Inhibition) can be accepted. Other validated and internationally accepted tests could also be used. The NOECs or other equivalent L(E)Cx should be used."

37. It should be noted that several of the OECD guidelines cited as examples for classification are being revised or are being planned for updating. Such revisions may lead to minor modifications of test conditions. Therefore, the expert group that developed the harmonized criteria for classification intended some flexibility in test duration or even species used.

38. Guidelines for conducting acceptable tests with fish, crustacea, and algae can be found in many sources (OECD, 1999; EPA, 1996; ASTM, 1999; ISO EU). The OECD monograph No.11, Detailed Review Paper on Aquatic Toxicity Testing for Industrial Chemicals and Pesticides, is a good compilation of pelagic test methods and sources of testing guidance. This document is also a source of appropriate test methodologies.

3.2.1 Fish Tests

Acute testing

39. Acute tests are generally performed with young juveniles 0.1 - 5 g in size for a period of 96 hours. The observational endpoint in these tests is mortality. Fish larger than this range and/or durations shorter than 96 hours are generally less sensitive. However, for classification, they could be used if no acceptable data with the smaller fish for 96 hours are available or the results of these tests with different size fish or test durations would influence a more hazardous classification band. Tests consistent with OECD Test Guideline 203 (Fish 96 hour LC50) or equivalent should be used for classification.

Chronic testing

40. Chronic or long term tests with fish can be initiated with fertilised eggs, embryos, juveniles, or reproductively active adults. Tests consistent with OECD Test Guideline 210 (Fish Early Life Stage), the fish life-cycle test (US EPA 850.1500), or equivalent can be used in the classification scheme. Durations can vary widely depending on the test purpose (anywhere from 7 days to over 200 days). Observational endpoints can include hatching success, growth (length and weight changes), spawning success, and survival. Technically, the OECD 210 Guideline (Fish Early Life Stage) is not a "chronic" test, but a sub-chronic test on sensitive life stages. It is widely accepted as a predictor of chronic toxicity and is used as such for purposes of classification in the harmonized system. Fish early life stage toxicity data are much more available than fish life cycle or reproduction studies.

3.2.2 Crustacea Tests

Acute testing

41. Acute tests with crustacea generally begin with first instar juveniles. For daphnids, a test duration of 48 hours is used. For other crustacea, such as mysids or others, a duration of 96 hours is typical. The observational endpoint is mortality or immobilisation as a surrogate to mortality. Immobilisation is defined as unresponsive to gentle prodding. Tests consistent with OECD Test Guideline 202 Part 1 (Daphnia acute) or USA-EPA OPPTS 850.1035 (Mysid acute toxicity) or their equivalents should be used for classification.

Chronic testing

42. Chronic tests with crustacea also generally begin with first instar juveniles and continue through maturation and reproduction. For daphnids, 21 days is sufficient for maturation and the production of 3 broods. For mysids, 28 days is necessary. Observational endpoints include time to first brood, number of offspring produced per female, growth, and survival. It is recommended that tests consistent with OECD Test Guideline 202 Part 2 (Daphnia reproduction) or US-EPA 850.1350 (Mysid chronic) or their equivalents be used in the classification scheme.

3.2.3 Algae/Plant Tests

Tests in algae

43. Algae are cultured and exposed to the test substance in a nutrient-enriched medium. Tests consistent with OECD Test Guideline 201 (Algal growth inhibition) should be used. Standard test methods employ a cell density in the inoculum in order to ensure exponential growth through the test, usually 3 to 4 days duration.

44. The algal test is a short-term test and, although it provides both acute and chronic endpoints, only the acute EC50 is used for classification in the harmonized system. The preferred observational endpoint in this study is algal growth rate inhibition because it is not dependent on the test design, whereas biomass depends both on growth rate of the test species as well as test duration and other elements of test design. If the endpoint is reported only as reduction in biomass or is not specified, then this value may be interpreted as an equivalent endpoint.

Tests in aquatic macrophytes

45. The most commonly used vascular plants for aquatic toxicity tests are duckweeds (*Lemna gibba* and *Lemna minor*). The Lemna test is a short-term test and, although it provides both acute and sub-chronic endpoints, only the acute EC50 is used for classification in the harmonized system. The tests last for up to 14 days and are performed in nutrient enriched media similar to that used for algae, but may be increased in strength. The observational endpoint is based on change in the number of fronds produced. Tests consistent with OECD Test Guideline on Lemna (in preparation) and US-EPA 850.4400 (aquatic plant toxicity, Lemna) should be used.

3.3 AQUATIC TOXICITY CONCEPTS

46. This section addresses the use of acute and chronic toxicity data in classification, and special considerations for exposure regimes, algal toxicity testing, and use of QSARs. For a more detailed discussion of aquatic toxicity concepts, one can refer to Rand (1996).

3.3.1 Acute toxicity

47. Acute toxicity for purposes of classification refers to the intrinsic property of a substance to be injurious to an organism in a short-term exposure to that substance. Acute toxicity is generally expressed in terms of a concentration which is lethal to 50% of the test organisms (LC50), causes a measurable adverse effect to 50% of the test organisms (e.g., immobilisation of daphnids), or leads to a 50% reduction in test (treated) organism responses from control (untreated) organism responses (e.g., growth rate in algae).

48. Substances with an acute toxicity determined to be less than one part per million (1 mg/l) are generally recognised as being very toxic. The handling, use, or discharge into the environment of these substances poses a high degree of hazard and they are classified in chronic and/or acute band I. Decimal bands are accepted for categorising acute toxicity above this band. Substances with an acute toxicity measured from one to ten parts per million (1 - 10 mg/l) are classified in Category II for acute toxicity, from ten to one hundred parts per million (10 - 100 mg/l) are classified in Category III for acute toxicity, and those over one hundred parts per million are regarded as practically non-toxic.

3.3.2 Chronic toxicity

49. Chronic toxicity, for purposes of declassification, refers to the potential or actual properties of a substance to cause adverse effects to aquatic organisms during exposures which are determined in relation to the life-cycle of the organism. Such chronic effects usually include a range of sublethal endpoints and are generally expressed in terms of a No Observable Effect Concentration (NOEC), or an equivalent ECx. Observable endpoints typically include survival, growth and/or reproduction. Chronic toxicity exposure durations can vary widely depending on test endpoint measured and test species used.

50. Since chronic toxicity data are less common in certain sectors than acute data, for classification schemes, the potential for chronic toxicity is identified by appropriate combinations of acute toxicity, lack of degradability, and/or the potential or actual bioaccumulation. Where such data exist and show long-term NOECs > 1 mg/l, this can be taken into account when deciding whether the classification based on the acute data should be applied. In this context, the following general approach should be used. In order to remove a chronic classification, it must be demonstrated that the NOEC used would be suitable in removing the concern for all taxa which resulted in classification. This can often be achieved by showing a long-term NOEC >1 mg/l for the most sensitive species identified by the acute toxicity. Thus, if a classification has been applied based on a fish acute LC50, it would

generally not be possible to remove this classification using a long-term NOEC from an invertebrate toxicity test. In this case, the NOEC would normally need to be derived from a long-term fish test of the same species or one of equivalent or greater sensitivity. Equally, if classification has resulted from the acute toxicity to more than one taxa, it is likely that NOECs > 1 mg/l from each taxa will need to be demonstrated. In case of classification of a substance as chronic Category IV, it is sufficient to demonstrate that NOECs are greater than the water solubility of the substances under consideration.

51. Testing with algae/Lemna cannot be used for de-classifying chemicals because (1) the algae and Lemna tests are not long-term studies, (2) the acute to chronic ratio is generally narrow and (3) the endpoints are more consistent with the end points for other organisms.

However where classification is applied solely due to the acute toxicity ($L(E)C_{50}$) observed in single algae/aquatic plant tests, but there is evidence from a range of other algae tests that the chronic toxicity (NOECs) for this taxonomic group is above 1mg/l, this evidence could be used to consider declassification. At present this approach cannot be applied to aquatic plants since no standardised chronic toxicity tests have been developed.

52. The GHS is intended to contain a specific value of chronic toxicity below which substances would be classified as chronically toxic, but the criteria are not yet set.

3.3.3 Exposure regimes

53. Four types of exposure conditions are employed in both acute and chronic tests and in both freshwater and saltwater media: static, static-renewal (semi-static), recirculation, and flow-through. The choice for which test type to use usually depends on test substance characteristics, test duration, test species, and regulatory requirements.

3.3.4 Test media for algae

54. Algal tests are performed in nutrient-enriched media and use of one common constituent, EDTA, or other chelators, should be considered carefully. When testing the toxicity of organic chemicals, trace amounts of a chelator like EDTA are needed to complex micronutrients in the culture medium; if omitted, algal growth can be significantly reduced and compromise test utility. However, chelators can reduce the observed toxicity of metal test substances. Therefore, for metal compounds, it is desirable that data from tests with high concentration of chelators and/or tests with stoichiometrical excess of chelator relative to iron be critically evaluated. Free chelator may mask heavy metal toxicity considerably, in particular with strong chelators like EDTA. However, in the absence of available iron in the medium the growth of algae can become iron limited, and consequently data from tests with no or with reduced iron and EDTA should be treated with caution.

3.3.5 Use of QSARs

55. For purpose of classification, and in the absence of experimental data, QSARs can be relied upon to provide predictions of acute toxicity for fish, daphnia, and algae for non-electrolyte, non-electrophilic, and otherwise non-reactive substances (See Chapter 6 on Use of QSAR). Problems remain for substances such as organophosphates which operate by means of special mechanisms such as functional groups which interact with biological receptors, or which can form sulfhydryl bonds with cellular proteins. Reliable QSARs have been derived for chemicals acting by a basic narcosis mechanism. These chemicals are nonelectrolytes of low reactivity such as hydrocarbons, alcohols, ketones and certain aliphatic chlorinated hydrocarbons which produce their biological effects as a function of their partition coefficients. Every organic chemical can produce narcosis. However, if the

chemical is an electrolyte or contains specific functional groups leading to non-narcotic mechanisms as well, any calculations of toxicity based on partition coefficient alone would severely underestimate the toxicity. QSARs for acute aquatic toxicity of parent compounds cannot be used to predict the effects of toxic metabolites or degradates, when these arise after a longer time period than the duration of acute tests.

3.4 WEIGHT OF EVIDENCE

56. The best quality data should be used as the fundamental basis for classification. Classification should preferably be based on primary data sources. It is essential that test conditions be clearly and completely articulated.

57. Where multiple studies for a taxonomic group are available, a decision on what is the most sensitive and highest quality must be made. A judgement has to be made on a case by case basis whether a non-GLP study with a more sensitive observation is used in lieu of a GLP study. It would appear that results that indicate high toxicity from tests performed according to non-standard or non-GLP guidelines should be able to be used for classification, whereas studies, which demonstrate negligible toxicity, would require more careful consideration. Substances, which are difficult to test, may yield apparent results that are more or less severe than the true toxicity. Expert judgement would also be needed for classification in these cases.

58. Where more than one acceptable test is available for the same taxonomic group, the most sensitive (the one with the lowest L(E)C50 or NOEC) is generally used for classification. However, this must be dealt with on a case-by-case basis. When larger data sets (4 or more values) are available for the same species, the geometric mean of toxicity values may be used as the representative toxicity value for that species. In estimating a mean value, it is not advisable to combine tests of different species within a taxa group or in different life stages or tested under different conditions or duration.

3.5 DIFFICULT TO TEST SUBSTANCES

59. Valid aquatic toxicity tests require the dissolution of the test substance in the water media under the test conditions recommended by the guideline. In addition, a bioavailable exposure concentration should be maintained for the duration of the test. Some chemical substances are difficult to test in aquatic systems and guidance has been developed to assist in testing these materials (DoE 1996; ECETOC 1996; and US EPA 1996). OECD is in the process of finalising a Guidance Document on Aquatic Toxicity testing of Difficult Substances and Mixtures (OECD, 2000). This latter document is a good source of information on the types of substances that are difficult to test and the steps needed to ensure valid conclusions from tests with these materials.

60. Nevertheless, much test data exist that may have used testing methodologies which, while not in conformity with what might be considered best practice today, can still yield information suitable for application of the classification criteria. Such data require special guidance on interpretation, although ultimately, expert judgement must be used in determining data validity. Such difficult to test substances may be poorly soluble, volatile, or subject to rapid degradation due to such processes as phototransformation, hydrolysis, oxidation, or biotic degradation. When testing algae, coloured materials may interfere with the test endpoint by attenuating the light needed for cell growth. In a similar manner, substances tested as cloudy dispersions above solubility may give rise to false toxicity measurements. Loading of the water column with test material can be an issue for particulates or solids such as metals. Petroleum distillate fractions can also pose loading problems, as well as difficult interpretational problems when deciding on the appropriate concentrations for determining $L(E)C_{50}$

values. The draft Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures describes the more common properties of many types of substances which are likely to pose testing difficulties.

Stability: If test chemical concentrations are expected to fall below 80% of nominal, testing, in order to be valid, may require exposure regimes which provide for renewal of the test material. Semi-static or flow-through conditions are preferred. Special problems arise, therefore, with respect to testing on algae, where the standard guidelines generally include static tests to be conducted. While alternative exposure regimes are possible for crustacea and fish, these tests are frequently conducted on static conditions as included in the internationally agreed guidelines. In these tests, a certain level of degradation as well as other relevant factors has to be tolerated and appropriate account must be taken in calculations of toxic concentrations. Some approaches on how this can be dealt with are covered in para 64 and 65. Where degradation occurs, it is also important to consider the influence of the toxicity of the degradation products on the recorded toxicity in the test. Expert judgement will need to be exercised when deciding if the data can be used for classification.

Degradation: When a compound breaks down or degrades under test condition, expert judgement should be used in calculating toxicity for classification, including consideration of known or likely breakdown products. Concentrations of the parent material and all significant toxic degradates are desirable. If degradates are expected to be relatively non-toxic, renewable exposure regimes are desirable in order to ensure that levels of the parent compounds are maintained.

Saturation: For single component substances, classification should be based only on toxic responses observed in the soluble range, and not on total chemical loading above solubility. Frequently, data are available which indicate toxicity at levels in excess of water solubility and, while these data will often be regarded as not valid, some interpretation may be possible. These problems generally apply when testing poorly soluble substances, and guidance on how to interpret such data is included in para 66 and 67 (see also the Guidance Document on Aquatic Toxicity testing of Difficult Substances and Mixtures).

Perturbation of test media: Special provisions may be needed to ensure dissolution of difficult to test substances. Such measures should not lead to significant changes in the test media when such changes are likely to lead to an increase or decrease in the apparent toxicity and hence the classification level of the test substance.

Complex substances: Many substances covered by the classification scheme are in fact mixtures, for which measurement of exposure concentrations is difficult, and in some cases impossible. Substances such as petroleum distillate fractions, polymers, substances with significant levels of impurities, etc can pose special problems since the toxic concentration is difficult to define and impossible to verify. Typical testing procedures often rely on the formation of a Water Soluble Fraction (WSF) or Water Accommodated Fraction (WAF) and data are reported in terms of loading rates. These data may be used in applying the classification criteria.

61. For classification of organic compounds, it is desirable to have stabilised and analytically measured test concentrations. Although measured concentrations are preferred, classification may be based on nominal concentration studies when these are the only valid data available under certain circumstances. If the material is likely to substantially degrade or otherwise be lost from the water column, care must be taken in data interpretation and classification should be done taking the loss of the

toxicant during the test into account, if relevant and possible. Additionally, metals present their own set of difficulties and are discussed separately. Table 1 lists several properties of difficult to test substances and their relevance for classification.

62. In most difficult to test conditions, the actual test concentration is likely to be less than the nominal or expected test concentration. Where toxicities ($L(E)C_{50}$s) are estimated to be less than 1mg/l for a difficult to test substance, one can be fairly confident the classification in the Acute Category 1 (and Chronic I if appropriate) is warranted. However, if the estimated toxicity is greater than 1 mg/l, the estimated toxicity is likely to under-represent the toxicity. In these circumstances, expert judgement is needed to determine the acceptability of a test with a difficult to test substance for use in classification. Where the nature of the testing difficulty is believed to have a significant influence on the actual test concentration when toxicity is estimated to be greater than 1 mg/l and the test concentration is not measured, then the test should be used with due caution in classification.

63. The following paragraphs provide some detailed guidance on some of these interpretational problems. In doing so it should be remembered that this is guidance and hard and fast rules cannot be applied. The nature of many of the difficulties mean that expert judgement must always be applied both in determining whether there is sufficient information in a test for a judgement to be made on its validity, and also whether a toxicity level can be determined suitable for use in applying the classification criteria.

Unstable substances

64. While testing procedures should ideally have been adopted which minimised the impacts of instability in the test media, in practice, in certain tests, it can be almost impossible to maintain a concentration throughout the test. Common causes of such instability are oxidation, hydrolysis, photodegradation and biodegradation. While the latter forms of degradation can more readily be controlled, such controls are frequently absent in much existing testing. Nevertheless, for some testing, particularly acute and chronic fish toxicity testing, a choice of exposure regimes is available to help minimise losses due to instability, and this should be taken into account in deciding on the test data validity.

65. Where instability is a factor in determining the level of exposure during the test, an essential prerequisite for data interpretation is the existence of measured exposure concentrations at suitable time points throughout the test. In the absence of analytically measured concentrations at least at the start and end of test, no valid interpretation can be made and the test should be considered as invalid for classification purposes. Where measured data are available, a number of practical rules can be considered by way of guidance in interpretation:

- where measured data are available for the start and end of test (as is normal for the acute Daphnia and algal tests), the $L(E)C_{50,}$ for classification purposes, may be calculated based on the geometric mean of the start and end of test concentrations. Where the end of test concentrations are below the analytical detection limit, such concentrations shall be considered to be half that detection limit.
- where measured data are available at the start and end of media renewal periods (as may be available for the semi-static tests), the geometric mean for each renewal period should be calculated, and the mean exposure over the whole exposure period calculated from these data.
- where the toxicity can be attributed to a degradation breakdown product, and the concentrations of this are known, the $L(E)C_{50}$ for classification purposes, may be calculated based on the geometric mean of the degradation product concentration, back

calculated to the parent substance.

- similar principles may be applied to measured data in chronic toxicity testing.

Poorly soluble substances

66. These substances, usually taken to be those with a solubility in water of <1 mg/l, are frequently difficult to dissolve in the test media, and the dissolved concentrations will often prove difficult to measure at the low concentrations anticipated. For many substances, the true solubility in the test media will be unknown, and will often be recorded as < detection limit in purified water. Nevertheless such substances can show toxicity, and where no toxicity is found, judgement must be applied to whether the result can be considered valid for classification. Judgement should err on the side of caution and should not underestimate the hazard.

67. Ideally, tests using appropriate dissolution techniques and with accurately measured concentrations within the range of water solubility should be used. Where such test data are available, they should be used in preference to other data. It is normal, however, particularly when considering older data, to find such substances with toxicity levels recorded in excess of the water solubility, or where the dissolved levels are below the detection limit of the analytical method. Thus, in both circumstances, it is not possible to verify the actual exposure concentrations using measured data. Where these are the only data available on which to classify, some practical rules can be considered by way of general guidance:

- where the acute toxicity is recorded at levels in excess of the water solubility, the $L(E)C_{50}$ for classification purposes, may be considered to be equal to or below the measured water solubility. In such circumstances it is likely that Chronic I and/or Acute I categories should be applied. In making this decision, due attention should be paid to the possibility that the excess undissolved substance may have given rise to physical effects on the test organisms. Where this is considered the likely cause of the effects observed, the test should be considered as invalid for classification purposes.

- where no acute toxicity is recorded at levels in excess of the water solubility, the $L(E)C_{50}$ for classification purposes may be considered to be greater than the measured water solubility. In such circumstances, consideration should be given to whether the Chronic IV category should apply. In making a decision that the substance shows no acute toxicity, due account should be taken of the techniques used to achieve the maximum dissolved concentrations. Where these are not considered as adequate, the test should be considered as invalid for classification purposes.

- where the water solubility is below the detection limit of the analytical method for a substance, and acute toxicity is recorded, the $L(E)C_{50}$ for classification purposes, may be considered to be less than the analytical detection limit. Where no toxicity is observed, the $L(E)C_{50}$ for classification purposes, may be considered to be greater than the water solubility. Due consideration should also be given to the quality criteria mentioned above.

- where chronic toxicity data are available, the same general rules should apply. In principle, only data showing no effects at the water solubility limit, or greater than 1 mg/l need be considered. Again, where these data cannot be validated by consideration of measured concentrations, the techniques used to achieve the maximum dissolved concentrations must be considered as appropriate.

Other factors contributing to concentration loss

68. A number of other factors can also contribute to losses of concentration and, while some can be avoided by correct study design, interpretation of data where these factors have contributed may, from time to time, be necessary.

- sedimentation: this can occur during a test for a number of reasons. A common explanation is that the substance has not truly dissolved despite the apparent absence of particulates, and agglomeration occurs during the test leading to precipitation. In these circumstances, the $L(E)C_{50}$ for classification purposes, may be considered to be based on the end of test concentrations. Equally, precipitation can occur through reaction with the media. This is considered under instability above.

- adsorption: this can occur for substances of high adsorption characteristics such as high log K_{ow} substances. Where this occurs, the loss of concentration is usually rapid and exposure may best be characterised by the end of test concentrations.

- bioaccumulation: losses may occur through the bioaccumulation of a substance into the test organisms. This may be particularly important where the water solubility is low and log K_{ow} correspondingly high. The $L(E)C_{50}$ for classification purposes, may be calculated based on the geometric mean of the start and end of test concentrations.

Perturbation of the test media

69. Strong acids and bases may appear toxic because they may alter pH. Generally however changes of the pH in aquatic systems are normally prevented by buffer systems in the test medium. If no data are available on a salt, the salt should generally be classified in the same way as the anion or cation, i.e., as the ion that receives the most stringent classification. If the effect concentration is related to only one of the ions, the classification of the salt should take the molecular weight difference into consideration by correcting the effect concentration by multiplying with the ratio: $MW_{salt}/MW_{ion.}$

70. Polymers are typically not available in aquatic systems. Dispersible polymers and other high molecular mass materials can perturb the test system and interfere with uptake of oxygen, and give rise to mechanical or secondary effects. These factors need to be taken into account when considering data from these substances. Many polymers behave like complex substances, however, having a significant low molecular mass fraction which can leach from the bulk polymer. This is considered further below.

Complex substances

71. Complex substances are characterised by a range of chemical structures, frequently in a homologous series, but covering a wide range of water solubilities and other physico-chemical characteristics. On addition to water, an equilibrium will be reached between the dissolved and undissolved fractions which will be characteristic of the loading of the substance. For this reason, such complex substances are usually tested as a WSF or WAF, and the $L(E)C_{50}$ recorded based on the loading or nominal concentrations. Analytical support data are not normally available since the dissolved fraction will itself be a complex mixtures of components. The toxicity parameter is sometimes referred to as LL_{50}, related to the lethal loading level. This loading level from the WSF or WAF may be used directly in the classification criteria.

72. Polymers represent a special kind of complex substance, requiring consideration of the polymer type and their dissolution/dispersal behaviour. Polymers may dissolve as such without change, (true solubility related to particle size), be dispersible, or portions consisting of low molecular weight fractions may go into solution. In the latter case, in effect, the testing of a polymer is a test of the ability of low molecular mass material to leach from the bulk polymer, and whether this leachate is toxic. It can thus be considered in the same way as a complex mixture in that a loading of polymer can best characterise the resultant leachate, and hence the toxicity can be related to this loading.

Table 1. Classification of difficult test substances

Property	**Nature of difficulty**	**Relevance for Classification**
Poorly water soluble	Achieving/maintaining required exposure concentration. Analysing exposure.	When toxic responses are observed above apparent solubility, expert judgement is required to confirm whether effects are due to chemical toxicity or a physical effect; if no effects are observed, it should be demonstrated that full, saturated dissolution has been achieved.
Toxic at low concentrations	Achieving/maintaining required exposure concentration. Analysing exposure.	Classified based on toxicity < 1 mg/l
Volatile	Maintaining and measuring exposure concentration.	Classification should be based on reliable measurement of concentrations.
Photo-degradable	Maintaining exposure concentrations. Toxicity of breakdown products.	Classification requires expert judgement and should be based on measured concentrations. Toxicity of significant breakdown products should be characterised.
Hydrolytically unstable	Maintaining exposure concentrations. Toxicity of breakdown products. Comparison of degradation half-lives to the exposure regimen used in testing.	Classification requires expert judgement, should be based on measured concentrations, and needs to address the toxicity of significant breakdown products.
Oxidizable	Achieving, maintaining and measuring exposure concentration. Toxicity of modified chemical structures or breakdown products. Comparison of degradation half-lives to the exposure regimen used in testing.	Classification requires expert judgement, should be based on measured concentrations, and needs to address the toxicity of significant breakdown products.
Subject to corrosion/ transformation (this refers to metals /metal compounds)	Achieving, maintaining and measuring exposure concentration. Comparison of partitioning from the water column half-lives to the exposure regimen used in testing.	Classification requires expert judgement, should be based on measured concentrations, and needs to address the toxicity of significant breakdown products.
Biodegradable	Maintaining exposure concentrations. Toxicity of breakdown products. Comparison of degradation half-lives to the exposure regimen used in testing.	Classification requires expert judgement, should be based on measured concentrations, and needs to address the toxicity of significant breakdown products.
Adsorbing	Maintaining exposure concentrations. Analysing exposure. Toxicity mitigation due to reduced availability of test substance.	Classification should use measured concentration of available material.
Chelating	Distinguishing chelated and non-chelated fractions in media.	Classification should use measurement of concentration of bioavailable material
Coloured	Light attenuation (an algal problem).	Classification must distinguish toxic effects from reduced growth due to light attenuation.

Table 1. Classification of difficult test substances (continued)

Hydrophobic	Maintaining constant exposure concentrations.	Classification should use measured concentration
Ionised	Maintaining exposure concentrations. Toxicity of breakdown products. Comparison of degradation half-lives to the exposure regime used in testing.	Classification requires expert judgement, should be based on measured concentrations, and needs to address the toxicity of significant breakdown products.
Multi-component substances and preparations	Preparing representative test batches.	Considered same as complex mixture.

3.6 INTERPRETING DATA QUALITY

3.6.1 Standardisation

73. Many factors can influence the results of toxicity tests with aquatic organisms. These factors include characteristics of the test water, experimental design, chemical characteristics of the test material, and biological characteristics of the test organisms. Therefore, it is important in conducting aquatic toxicity tests to use standardised test procedures to reduce the influence of these sources of extraneous variability. The goal of test standardisation and international harmonisation of these standards is to reduce test variability and improve precision, reproducibility, and consistency of test results.

3.6.2 Data hierarchies

74. Classification should be based on primary data of good quality. Preference is given to data conforming to OECD Test Guidelines or equivalent and Good Laboratory Practices (GLP). While data from internationally harmonised test methods performed on standard test species are preferred, results of tests performed using widely recognised international or national methods or their equivalent may also be used, e.g., ISO or ASTM methods. Data from tests that appear to conform to accepted guidelines but which lacks provisions for GLP can be used in the absence of pertinent GLP data.

75. Pedersen et al (1995) provides a data quality-scoring system, which is compatible with many others in current use, including that, used by the US-EPA for its AQUIRE database. See also Mensink et al (1995) for discussions of data quality. The data quality scoring system described in Pedersen *et al.* includes a reliability ranking scheme, which can be a model for use with in classifying under the harmonised scheme. The first three levels of data described by Pedersen are for preferred data.

76. Data for classification under the harmonised scheme should come from primary sources. However, since many nations and regulatory authorities will perform classification using the globally harmonised scheme, classification should allow for use of reviews from national authorities and expert panels as long as the reviews are based on primary sources. Such reviews should include summaries of test conditions, which are sufficiently detailed for weight of evidence and classification decisions to be made. It may be possible to use the reviews, which were made by a well-recognised group such as GESAMP for which the primary data are accessible.

77. In the absence of empirical test data, validated Quantitative Structure Activity Relationships (QSARs) for aquatic toxicity may be used. Test data always take precedence over QSAR predictions, providing the test data are valid.

ANNEX 3.I

TEST GUIDELINES

78. Most of the guidelines mentioned are found in compilations from the organisation issuing them. The main references to these are:

- EC guidelines: European Commission (1996). Classification, Packaging and Labelling of Dangerous Substances in the European Union. Part 2 – Testing Methods. European Commission. 1997. ISBN92-828-0076-8. (Homepage: http://ecb.ei.jrc.it/testing-methods/);
- ISO guidelines: Available from the national standardisation organisations or ISO (Homepage: http://www.iso.ch/);
- OECD guidelines for the testing of chemicals. OECD, Paris, 1993 with regular updates (Homepage: http://www.oecd.org/ehs/test/testlist.htm);
- OPPTS guidelines: US-EPA homepage: http://www.epa.gov/opptsfrs/home/guidelin.htm;
- ASTM : ASTM's homepage: http://www.astm.org. Further search via “standards”.

OECD Test Guideline 201 (1984) Alga, Growth Inhibition Test

OECD Test Guideline 202 (1984) Daphnia sp. Acute Immobilisation Test and Reproduction Test

OECD Test Guideline 203 (1992) Fish, Acute Toxicity Test

OECD Test Guideline 204 (1984) Fish, Prolonged Toxicity Test: 14-Day Study

OECD Test Guideline 210 (1992) Fish, Early-Life Stage Toxicity Test

OECD Test Guideline 211 (1998) Daphnia magna Reproduction Test

OECD Test Guideline 212 (1998) Fish, Short-term Toxicity Test on Embryo and Sac-Fry Stages

OECD Test Guideline 215 (2000) Fish, Juvenile Growth Test

OECD Test Guideline 221 (in preparation) *Lemna* sp. Growth inhibition test

EC C.1: Acute Toxicity for Fish (1992)

EC C.2: Acute Toxicity for Daphnia (1992)

EC C.3: Algal Inhibition Test (1992)

EC C.14: Fish Juvenile Growth Test (2001)

EC C.15: Fish, Short-term Toxicity Test on Embryo and Sac-Fry Stages (2001)

EC C.20: Daphnia Magna Reproduction Test (2001)

OPPTS Testing Guidelines for Environmental Effects (850 Series Public Drafts)

850.1000 Special consideration for conducting aquatic laboratory studies (Adobe PDF)

850.1000 Special consideration for conducting aquatic laboratory studies (Text to HTML)

850.1010 Aquatic invertebrate acute toxicity, test, freshwater daphnids (Adobe PDF)

850.1010 Aquatic invertebrate acute toxicity, test, freshwater daphnids (Text to HTML)

850.1020 Gammarid acute toxicity test (Adobe PDF)

850.1020 Gammarid acute toxicity test (Text to HTML)

850.1035 Mysid acute toxicity test (Adobe PDF)

850.1035 Mysid acute toxicity test (Text to HTML)

850.1045 Penaeid acute toxicity test (Adobe PDF)

850.1045 Penaeid acute toxicity test (Text to HTML)

850.1075 Fish acute toxicity test, freshwater and marine (Adobe PDF)

850.1075 Fish acute toxicity test, freshwater and marine (Text to HTML)

850.1300 Daphnid chronic toxicity test (Adobe PDF)

850.1300 Daphnid chronic toxicity test (Text to HTML)

850.1350 Mysid chronic toxicity test (Adobe PDF)

850.1350 Mysid chronic toxicity test (Text to HTML)

850.1400 Fish early-life stage toxicity test (Adobe PDF)

850.1400 Fish early-life stage toxicity test (Text to HTML)

850.1500 Fish life cycle toxicity (Adobe PDF)

850.1500 Fish life cycle toxicity (Text to HTML)

850.1730 Fish BCF (Adobe PDF)

850.1730 Fish BCF (Text to HTML)

850.4400 Aquatic plant toxicity test using Lemna spp. Tiers I and II (Adobe PDF)

850.4400 Aquatic plant toxicity test using Lemna spp. Tiers I and II (Text to HTML)

850.4450 Aquatic plants field study, Tier III (Adobe PDF)

850.4450 Aquatic plants field study, Tier III (Text to HTML)

850.5400 Algal toxicity, Tiers I and II (Adobe PDF)

850.5400 Algal toxicity, Tiers I and II (Text to HTML)

*Note 1) :*This list of public drafts of environmental effects testing guidelines was taken from the homepage) of the U.S. Environmental Protection Agency on 19 September 2000. (http://www.epa.gov/OPPTS_Harmonized/850_Ecological_Effects_Test_Guidelines/Drafts) The list was last revised on 10 February 1997 by an automated conversion program. Further revisions may occur as the draft guidelines are updated.

ANNEX 3.II

REFERENCES

APHA 1992. Standard Methods for the Examination of Water and Wastewater, 18th edition. American Public Health Association, Washington, DC.

ASTM 1999. Annual Book of ASTM standards, Vol. 11.04. American Society for Testing and Materials, Philadelphia, PA.

DoE 1996. Guidance on the Aquatic Toxicity Testing of Difficult Substances. United Kingdom Department of the Environment, London.

ECETOC 1996. Aquatic Toxicity Testing of Sparingly Soluble, Volatile and Unstable Substances. ECETOC Monograph No. 26, ECETOC, Brussels.

Lewis, M. A. 1995. Algae and vascular plant tests. In: Rand, G. M. (ed.) 1995. Fundamentals of Aquatic Toxicology, Second Edition. Taylor & Francis, Washington, DC. pp. 135-169.

Mensink, B. J. W. G., M. Montforts, L. Wijkhuizen-Maslankiewicz, H. Tibosch, and J.B.H.J. Linders 1995. Manual for Summarising and Evaluating the Environmental Aspects of Pesticides. Report No. 679101022 RIVM, Bilthoven, The Netherlands.

OECD 1998. Harmonized Integrated Hazard Classification System for Human Health and Environmental Effects of Chemical Substances. OECD, Paris. http://www.oecd.org/ehs/Class/HCL6.htm

OECD 1999. Guidelines for Testing of Chemicals. Organisation for Economic Co-operation and Development, Paris.

OECD 2000. Revised Draft Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures, OECD, Paris.

Pedersen, F., H. Tyle, J. R. Niemeldi, B. Guttmann, L. Lander, and A. Wedebrand 1995. Environmental Hazard Classification - data collection and interpretation guide. TemaNord 1995:581.

US EPA 1996. Ecological Effects Test Guidelines - OPPTS 850.1000. Special Considerations for Conducting Aquatic Laboratory Studies. Public Draft, EPA 712-C-96-113. United States Environmental Protection Agency. http://www.epa.gov/docs/OPTS_Harmonized/

OECD Monograph 11, Detailed Review Paper on Aquatic Toxicity Testing for Industrial Chemicals and Pesticides.

Rand, Gary M., Fundamentals of Aquatic toxicology: Effects, Environmental Fate, and Risk Assessment.

4. DEGRADATION

4.1 INTRODUCTION

79. Degradability is one of the important intrinsic properties of chemical substances that determine their potential environmental hazard. Non-degradable substances will persist in the environment and may consequently have a potential for causing long-term adverse effects on biota. In contrast, degradable substances may be removed in the sewers, in sewage treatment plants or in the environment.

80. Classification of chemical substances is primarily based on their intrinsic properties. However, the degree of degradation depends not only on the intrinsic recalcitrance of the molecule, but also on the actual conditions in the receiving environmental compartment as e.g., redox potential, pH, presence of suitable micro-organisms, concentration of the substances and occurrence and concentration of other substrates. The interpretation of the degradation properties in an aquatic hazard classification context therefore requires detailed criteria that balance the intrinsic properties of the substance and the prevailing environmental conditions into a concluding statement on the potential for long-term adverse effects. The purpose of the present chapter is to present guidance for interpretation of data on degradability of organic substances. The guidance is based on an analysis of the above mentioned aspects regarding degradation in the aquatic environment. Based on the guidance a detailed decision scheme for use of existing degradation data for classification purposes is proposed. The types of degradation data included in this Guidance Document are ready biodegradability data, simulation data for transformation in water, aquatic sediment and soil, BOD_5/COD-data and techniques for estimation of rapid degradability in the aquatic environment. Also considered are anaerobic degradability, inherent biodegradability, sewage treatment plant simulation test data, abiotic transformation data such as hydrolysis and photolysis, removal process such as volatilisation and finally, data obtained from field investigations and monitoring studies.

81. The term degradation is defined in Glossary in this Guidance Document as the decomposition of organic molecules to smaller molecules and eventually to carbon dioxide, water and salts. For inorganic compounds and metals, the concept of degradability as applied to organic compounds has limited or no meaning. Rather the substance may be transformed by normal environmental processes to either increase or decrease the bioavailability of the toxic species. Therefore, the present chapter deals only with organic substances and organo-metals. Environmental partitioning from the water column is discussed in Chapter 7.

82. Data on degradation properties of a substance may be available from standardised tests or from other types of investigations, or they may be estimated from the structure of the molecules. The interpretation of such degradation data for classification purposes often requires detailed evaluation of the test data. Guidance is given in the present chapter and more details can be found in two annexes describing available methods (Annex 3) and factors influencing degradation in aquatic environments (Annex 4).

4.2 INTERPRETATION OF DEGRADABILITY DATA

4.2.1 Rapid degradability

83. Aquatic hazard classification of chemical substances is normally based on existing data on their environmental properties. Only seldom will test data be produced with the main purpose of facilitating a classification. Often a diverse range of test data is available that does not necessarily fits directly with the classification criteria. Consequently, guidance is needed on interpretation of existing test data in the context of the aquatic hazard classification. Based on the harmonised criteria, guidance for interpretation of degradation data is prepared below for the three types of data comprised by the expression "rapid degradation" in the aquatic environment (see para 8, 9, 20, 21 & 22 and the definition in Annex 1 of the "Harmonised system for the classification of chemicals which are hazardous for the aquatic environment" (OECD, 1998), which is attached to this Guidance Document as Appendix.

4.2.2 Ready biodegradability

84. Ready biodegradability is defined in the OECD Test Guidelines No. 301 (OECD 1992). All organic substances that degrade to a level higher than the pass level in a standard OECD ready biodegradability test or in a similar test should be considered readily biodegradable and consequently also rapidly degradable. Many literature test data, however, do not specify all of the conditions that should be evaluated to demonstrate whether or not the test fulfils the requirements of a ready biodegradability test. Expert judgement is therefore needed as regards the validity of the data before use for classification purposes. Before concluding on the ready biodegradability of a test substance, however, at least the following parameters should be considered.

Concentration of test substance

85. Relatively high concentrations of test substance are used in the OECD ready biodegradability tests (2-100 mg/L). Many substances may, however, be toxic to the inocula at such high concentrations causing a low degradation in the tests although the substances might be rapidly degradable at lower non-toxic concentrations. A toxicity test with micro-organisms (as e.g., the OECD Test Guideline 209 "Activated Sludge, Respiration Inhibition Test", the ISO 9509 nitrification inhibition test, or the ISO 11348 luminescent bacteria inhibition test) may demonstrate the toxicity of the test substance. When it is likely that inhibition is the reason for a substance being not readily degradable, results from a test employing lower non-toxic concentrations of the test substance should be used when available. Such test results could on a case by case basis be considered in relation to the classification criteria for rapid degradation, even though surface water degradation test data with environmentally realistic microbial biomass and non toxic realistic low concentration of the test substance in general are preferred, if available.

Time window

86. The harmonised criteria include a general requirement for all of the ready biodegradability tests on achievement of the pass level within 10 days. This is not in line with the OECD Test Guideline 301 in which the 10-days time window applies to the OECD ready biodegradability tests except to the MITI I test (OECD Test Guideline 301C). In the Closed Bottle test (OECD Test Guideline 301D), a 14-days window may be used instead when measurements have not been made after 10 days. Moreover, often only limited information is available in references of biodegradation tests. Thus, as a pragmatic approach the percentage of degradation reached after 28 days may be used directly for assessment of ready biodegradability when no information on the 10-days time

window is available. This should, however, only be accepted for existing test data and data from tests where the 10-days window does not apply.

4.2.3 BOD_5/COD

87. Information on the 5-day biochemical oxygen demand (BOD_5) will be used for classification purposes only when no other measured degradability data are available. Thus, priority is given to data from ready biodegradability tests and from simulation studies regarding degradability in the aquatic environment. The BOD_5 test is a traditional biodegradation test that is now replaced by the ready biodegradability tests. Therefore, this test should not be performed today for assessment of the ready biodegradability of substances. Older test data may, however, be used when no other degradability data are available. For substances where the chemical structure is known, the theoretical oxygen demand (ThOD) can be calculated and this value should be used instead of the chemical oxygen demand (COD).

4.2.4 Other convincing scientific evidence

88. Rapid degradation in the aquatic environment may be demonstrated by other data than referred to in criteria a) and b) in Annex I of the harmonised criteria (OECD 1998). These may be data on biotic and/or abiotic degradation. Data on primary degradation can only be used where it is demonstrated that the degradation products shall not be classified as hazardous to the aquatic environment, i.e., that they do not fulfil the classification criteria.

89. The fulfilment of criterion c) requires that the substance is degraded in the aquatic environment to a level of >70% within a 28-day period. If first-order kinetics are assumed, which is reasonable at the low substance concentrations prevailing in most aquatic environments, the degradation rate will be relatively constant for the 28-day period. Thus, the degradation requirement will be fulfilled with an average degradation rate constant, $k > -(\ln 0.3 - \ln 1)/28 = 0.043$ day^{-1}. This corresponds to a degradation half-life, $t_{½} < \ln 2/0.043 = 16$ days.

90. Moreover, as degradation processes are temperature dependent, this parameter should also be taken into account when assessing degradation in the environment. Data from studies employing environmentally realistic temperatures should be used for the evaluation. When data from studies performed at different temperatures need to be compared, the traditional Q10 approach could be used, i.e., that the degradation rate is halved when the temperature decreases by 10°C.

91. The evaluation of data on fulfilment of this criterion should be conducted on a case by case basis by expert judgement. However, guidance on the interpretation of various types of data that may be used for demonstrating a rapid degradation in the aquatic environment is given below. In general, only data from aquatic biodegradation simulation tests are considered directly applicable. However simulation test data from other environmental compartments could be considered as well, but such data require in general more scientific judgement before use.

Aquatic simulation tests

92. Aquatic simulation tests are tests conducted in laboratory, but simulating environmental conditions and employing natural samples as inoculum. Results of aquatic simulation tests may be used directly for classification purposes, when realistic environmental conditions in surface waters are simulated, i.e.,:

- substance concentration that is realistic for the general aquatic environment (often in the low µg/L range);
- inoculum from a relevant aquatic environment;
- realistic concentration of inoculum (10^3-10^6 cells/mL);
- realistic temperature (e.g., 5°C to 25°C); and
- ultimate degradation is determined (i.e., determination of the mineralisation rate or the individual degradation rates of the total biodegradation pathway).

93. Substances that under these conditions are degraded at least 70% within 28 days, i.e., with a half-life < 16 days are considered rapidly degradable.

Field investigations

94. Parallels to laboratory simulation tests are field investigations or mesocosm experiments. In such studies, fate and/or effects of chemicals in environments or environmental enclosures may be investigated. Fate data from such experiments might be used for assessing the potential for a rapid degradation. This may, however, often be difficult, as it requires that an ultimate degradation can be demonstrated. This may be documented by preparing mass balances showing that no non-degradable intermediates are formed, and which take the fractions into account that are removed from the aqueous system due to other processes such as sorption to sediment or volatilisation from the aquatic environment.

Monitoring data

95. Monitoring data may demonstrate the removal of contaminants from the aquatic environment. Such data are, however, very difficult to use for classification purposes. The following aspects should be considered before use:

- Is the removal a result of degradation, or is it a result of other processes such as dilution or distribution between compartments (sorption, volatilisation)?
- Is formation of non-degradable intermediates excluded?

Only when it can be demonstrated that removal as a result of ultimate degradation fulfils the criteria for rapid degradability, such data be considered for use for classification purposes. In general, monitoring data should only be used as supporting evidence for demonstration of either persistence in the aquatic environment or a rapid degradation.

Inherent biodegradability tests

96. Substances that are degraded more than 70% in tests for inherent biodegradability (OECD Test Guidelines 302) have the potential for ultimate biodegradation. However, because of the optimum conditions in these tests, the rapid biodegradability of inherently biodegradable substances in the environment cannot be assumed. The optimum conditions in inherent biodegradability tests stimulate adaptation of the micro-organisms thus increasing the biodegradation potential, compared to natural environments. Therefore, positive results in general should not be interpreted as evidence for rapid degradation in the environment (see Note 1).

Note 1: In relation to interpretation of degradation data equivalent with the harmonised OECD criteria for chronic Category IV, the standing EU working group for environmental hazard classification of substances is discussing whether certain types of data from inherent biodegradability tests may be

used in a case by case evaluation as a basis for not classifying substances otherwise fulfilling this classification criterion:
The inherent biodegradability tests concerned are the Zahn Wellens test (OECD TG 302 B) and the MITI II test (OECD TG 302 C). The conditions for use in this regard are:

a) The methods must not employ pre-exposed (pre-adapted) micro-organisms.
b) The time for adaptation within each test should be limited, the test endpoint should refer to the mineralisation only and the pass level and time for reaching these should be, respectively:

- MITI II pass level > 60 % within 14 days
- Zahn Wellens Test > 70 % within 7 days.

Sewage treatment plant simulation tests

97. Results from tests simulating the conditions in a sewage treatment plant (STP) (e.g., the OECD Test Guideline 303) cannot be used for assessing the degradation in the aquatic environment. The main reasons for this are that the microbial biomass in a STP is significantly different from the biomass in the environment, that there is a considerably different composition of substrates, and that the presence of rapidly mineralised organic matter in waste water facilitates degradation of the test substance by co-metabolism.

Soil and sediment degradation data

98. It has been argued that for many non-sorptive (non-lipophilic) substances more or less the same degradation rates are found in soil and in surface water. For lipophilic substances, a lower degradation rate may generally be expected in soil than in water due to partial immobilisation caused by sorption. Thus, when a substance has been shown to be degraded rapidly in a soil simulation study, it is most likely also rapidly degradable in the aquatic environment. It is therefore proposed that an experimentally determined rapid degradation in soil is sufficient documentation for a rapid degradation in surface waters when:

- no pre-exposure (pre-adaptation) of the soil micro-organisms has taken place, and
- an environmentally realistic concentration of substance is tested, and
- the substance is ultimately degraded within 28 days with a half-life <16 days corresponding to a degradation rate >0.043 day^{-1} .

99. The same argumentation is considered valid for data on degradation in sediment under aerobic conditions.

Anaerobic degradation data

100. Data regarding anaerobic degradation cannot be used in relation to deciding whether a substance should be regarded as rapidly degradable, because the aquatic environment is generally regarded as the aerobic compartment where the aquatic organisms, such as those employed for aquatic hazard classification, live.

Hydrolysis

101. Data on hydrolysis (e.g., OECD Test Guideline 111) might be considered for classification purposes only when the longest half-life $t_{½}$ determined within the pH range 4-9 is shorter than 16 days. However, hydrolysis is not an ultimate degradation and various intermediate degradation products may be formed, some of which may be only slowly degradable. Only when it can be

satisfactorily demonstrated that the hydrolysis products formed do not fulfil the criteria for classification as hazardous for the aquatic environment, data from hydrolysis studies could be considered.

102. When a substance is quickly hydrolysed (e.g., with $t_{½}$ < a few days), this process is a part of the degradation determined in biodegradation tests. Hydrolysis may be the initial transformation process in biodegradation.

Photochemical degradation

103. Information on photochemical degradation (e.g., OECD, 1997) is difficult to use for classification purposes. The actual degree of photochemical degradation in the aquatic environment depends on local conditions (e.g., water depth, suspended solids, turbidity) and the hazard of the degradation products is usually not known. Probably only seldom will enough information be available for a thorough evaluation based on photochemical degradation.

Estimation of degradation

104. Certain QSARs have been developed for prediction of an approximate hydrolysis half-life, which should only be considered when no experimental data are available. However, a hydrolysis half-life can only be used in relation to classification with great care, because hydrolysis does not concern ultimate degradability (see "Hydrolysis" of this Section). Furthermore the QSARs developed until now have a rather limited applicability and are only able to predict the potential for hydrolysis on a limited number of chemical categories. The QSAR program HYDROWIN (version 1.67, Syracuse Research Corporation) is for example only able to predict the potential for hydrolysis on less than $1/5^{th}$ of the existing EU substances which have a defined (precise) molecular structure (Niemelä, 2000).

105. In general, no quantitative estimation method (QSAR) for estimating the degree of biodegradability of organic substances is yet sufficiently accurate to predict rapid degradation. However, results from such methods may be used to predict that a substance is not rapidly degradable. For example, when in the Biodegradation Probability Program (e.g., BIOWIN version 3.67, Syracuse Research Corporation) the probability is < 0.5 estimated by the linear or non-linear methods, the substances should be regarded as not rapidly degradable (OECD, 1994; Pedersen *et al.*, 1995 & Langenberg *et al.*, 1996). Also other (Q)SAR methods may be used as well as expert judgement, for example, when degradation data for structurally analogue compounds are available, but such judgement should be conducted with great care. In general, a QSAR prediction that a substance is not rapidly degradable is considered a better documentation for a classification than application of a default classification, when no useful degradation data are available.

Volatilisation

106. Chemicals may be removed from some aquatic environments by volatilisation. The intrinsic potential for volatilisation is determined by the Henry's Law constant (H) of the substance. Volatilisation from the aquatic environment is highly dependent on the environmental conditions of the specific water body in question, such as the water depth, the gas exchange coefficients (depending on wind speed and water flow) and stratification of the water body. Because volatilisation only represents removal of a chemical from water phase, the Henry's Law constant can not be used for assessment of degradation in relation to aquatic hazard classification of substances. Substances that are gases at ambient temperature may however for example be considered further in this regard (see also Pedersen *et al.*, 1995).

4.2.5 No degradation data available

107. When no useful data on degradability are available - either experimentally determined or estimated data - the substance should be regarded as not rapidly degradable.

4.3 GENERAL INTERPRETATION PROBLEMS

4.3.1 Complex substances

108. The harmonised criteria for classification of chemicals as hazardous for the aquatic environment focus on single substances. A certain type of intrinsically complex substance are multi-component substances. They are typically of natural origin and need occasionally to be considered. This may be the case for chemicals that are produced or extracted from mineral oil or plant material. Such complex chemicals are normally considered as single substances in a regulatory context. In most cases they are defined as a homologous series of substances within a certain range of carbon chain length and/or degree of substitution. When this is the case, no major difference in degradability is foreseen and the degree of degradability can be established from tests of the complex chemical. One exception would be when a borderline degradation is found because in this case some of the individual substances may be rapidly degradable and other may be not rapidly degradable. This requires a more detailed assessment of the degradability of the individual components in the complex substance. When not-rapidly-degradable components constitute a significant part of the complex substance (e.g., more than 20%, or for a hazardous component, an even lower content), the substance should be regarded as not rapidly degradable.

4.3.2 Availability of the substance

109. Degradation of organic substances in the environment takes place mostly in the aquatic compartments or in aquatic phases in soil or sediment. Hydrolysis, of course, requires the presence of water. The activity of micro-organisms depends on the presence of water. Moreover, biodegradation requires that the micro-organisms are directly in contact with the substance. Dissolution of the substance in the water phase that surrounds the micro-organisms is therefore the most direct way for contact between the bacteria and fungi and the substrate.

110. The present standard methods for investigating degradability of chemical substances are developed for readily soluble test compounds. However, many organic substances are only slightly soluble in water. As the standard tests require 2-100 mg/L of the test substance, sufficient availability may not be reached for substances with a low water solubility. Tests with continuous mixing and/or an increased exposure time, or tests with a special design where concentrations of the test substance lower than the water solubility have been employed, may be available on slightly soluble compounds.

4.3.3 Test duration less than 28 days

111. Sometimes degradation is reported for tests terminated before the 28 days period specified in the standards (e.g., the MITI, 1992). These data are of course directly applicable when a degradation greater than or equal to the pass level is obtained. When a lower degradation level is reached, the results need to be interpreted with caution. One possibility is that the duration of the test was too short and that the chemical structure would probably have been degraded in a 28-day biodegradability test. If substantial degradation occurs within a short time period, the situation may be compared with the criterion $BOD_5/COD \geq 0.5$ or with the requirements on degradation within the

10-days time window. In these cases, a substance may be considered readily degradable (and hence rapidly degradable), if:

- the ultimate biodegradability exceeds 50% within 5 days; or
- the ultimate degradation rate constant in this period is greater than 0.1 day^{-1} corresponding to a half-life of 7 days.

112. These criteria are proposed in order to ensure that rapid mineralisation did occur, although the test was ended before 28 days and before the pass level was attained. Interpretation of test data that do not comply with the prescribed pass levels must be made with great caution. It is mandatory to consider whether a biodegradability below the pass level was due to a partial degradation of the substance and not a complete mineralisation. If partial degradation is the probable explanation for the observed biodegradability, the substance should be considered not readily biodegradable.

4.3.4 Primary biodegradation

113. In some tests, only the disappearance of the parent compound (i.e., primary degradation) is determined for example by following the degradation by specific or group specific chemical analyses of the test substance. Data on primary biodegradability may be used for demonstrating rapid degradability~~,~~ only when it can be satisfactorily demonstrated, that the degradation products formed do not fulfil the criteria for classification as hazardous to the aquatic environment.

4.3.5 Conflicting results from screening tests

114. The situation where more degradation data are available for the same substance introduces the possibility of conflicting results. In general, conflicting results for a substance which has been tested several times with an appropriate biodegradability test could be interpreted by a "weight of evidence approach". This implies that if both positive (i.e., higher degradation than the pass level) and negative results have been obtained for a substance in ready biodegradability tests, then the data of the highest quality and the best documentation should be used for determining the ready biodegradability of the substance. However, positive results in ready biodegradability tests could be considered valid, irrespective of negative results, when the scientific quality is good and the test conditions are well documented, i.e., guideline criteria are fulfilled, including the use of non-pre-exposed (non-adapted) inoculum. None of the various screening tests are suitable for the testing of all types of substances, and results obtained by the use of a test procedure which is not suitable for the specific substance should be evaluated carefully before a decision on the use is taken.

115. Thus, there are a number of factors that may explain conflicting biodegradability data from screening tests:

- inoculum;
- toxicity of test substance;
- test conditions;
- solubility of the test substance; and
- volatilisation of the test substance.

116. The suitability of the inoculum for degrading the test substance depends on the presence and amount of competent degraders. When the inoculum is obtained from an environment that has previously been exposed to the test substance, the inoculum may be adapted as evidenced by a degradation capacity, which is greater than that of an inoculum from a non-exposed environment. As far as possible the inoculum must be sampled from an unexposed environment, but for

substances that are used ubiquitously in high volumes and released widespread or more or less continuously, this may be difficult or impossible. When conflicting results are obtained, the origin of the inoculum should be checked in order to clarify whether or not differences in the adaptation of the microbial community may be the reason.

117. As mentioned above, many substances may be toxic or inhibitory to the inoculum at the relatively high concentrations tested in ready biodegradability tests. Especially in the Modified MITI (I) test (OECD Test Guideline 301C) and the Manometric Respirometry test (OECD Test Guideline 301F) high concentrations (100 mg/L) are prescribed. The lowest test substance concentrations are prescribed in the Closed Bottle test (OECD Test Guideline 301D) where 2-10 mg/L is used. The possibility of toxic effects may be evaluated by including a toxicity control in the ready biodegradability test or by comparing the test concentration with toxicity test data on micro-organisms, e.g., the respiration inhibition tests (OECD Test Guideline 209), the nitrification inhibition test (ISO 9509) or, if other microbial toxicity tests are not available, the bioluminescence inhibition test (ISO 11348). When conflicting results are found, this may be caused by toxicity of the test substance. If the substance is not inhibitory at environmentally realistic concentrations, the greatest degradation measured in screening tests may be used as a basis for classification. If simulation test data are available in such cases, consideration of these data may be especially important, because a low non inhibitory concentration of the substance may have been employed, thus giving a more reliable indication of the biodegradation half-life of the substance under environmentally realistic conditions.

118. When the solubility of the test substance is lower than the concentrations employed in a test, this parameter may be the limiting factor for the actual degradation measured. In these cases, results from tests employing the lowest concentrations of test substance should prevail, i.e., often the Closed Bottle test (OECD Test Guideline 301D). In general, the DOC Die-Away test (OECD Test Guideline 301A) and the Modified OECD Screening test (OECD Test Guideline 301E) are not suitable for testing the biodegradability of poorly soluble substances (e.g., OECD Test Guideline 301).

119. Volatile substances should only be tested in closed systems as the Closed Bottle test (OECD Test Guideline 301D), the MITI I test (OECD Test Guideline 301C) and the Manometric Respirometry test (OECD Test Guideline 301F). Results from other tests should be evaluated carefully and only considered if it can be demonstrated, e.g., by mass balance estimates, that the removal of the test substance is not a result of volatilisation.

4.3.6 Variation in simulation test data

120. A number of simulation test data may be available for certain high priority chemicals. Often such data provide a range of half lives in environmental media such as soil, sediment and/or surface water. The observed differences in half-lives from simulation tests performed on the same substance may reflect differences in test conditions, all of which may be environmentally relevant. A suitable half life in the higher end of the observed range of half lives from such investigations should be selected for classification by employing a weight of evidence approach and taking the realism and relevance of the employed tests into account in relation to environmental conditions. In general, simulation test data of surface water are preferred relative to aquatic sediment or soil simulation test data in relation to the evaluation of rapid degradability in the aquatic environment.

4.4 Decision scheme

121. The following decision scheme may be used as a general guidance to facilitate decisions in relation to rapid degradability in the aquatic environment and classification of chemicals hazardous to the aquatic environment.

122. A substance is considered to be not rapidly degradable unless at least one of the following is fulfilled:

1) the substance is demonstrated to be readily biodegradable in a 28-day test for ready biodegradability. The pass level of the test (70% DOC removal or 60% theoretical oxygen demand) must be achieved within 10 days from the onset of biodegradation, if it is possible to evaluate this according to the available test data. If this is not possible, then the pass level should be evaluated within a 14 days time window if possible, or after the end of the test; or

2) the substance is demonstrated to be ultimately degraded in a surface water simulation test [1] with a half-life of <16 days (corresponding to a degradation of >70% within 28 days); or

3) the substance is demonstrated to be primarily degraded (biotically or abiotically) in the aquatic environment with a half-life <16 days (corresponding to a degradation of >70% within 28 days) and it can be demonstrated that the degradation products do not fulfil the criteria for classification as hazardous to the aquatic environment; or

When these data are not available rapid degradation may be demonstrated if either of the following criteria are justified:

4) the substance is demonstrated to be ultimately degraded in an aquatic sediment or soil simulation test [1] with a half-life of < 16 days (corresponding to a degradation of > 70% within 28 days); or

5) in those cases where only BOD_5 and COD data are available, the ratio of BOD_5/COD is greater than or equal to 0.5. The same criterion applies to ready biodegradability tests of a shorter duration than 28 days, if the half-life furthermore is < 7 days.

Note 1. Simulations tests should reflect realistic environmental conditions such as low concentration of the chemical, realistic temperature and employment of ambient microbial biomass not pre-exposed to the chemical.

123. If none of the above types of data are available then the substance is considered as not rapidly degradable. This decision may be supported by fulfilment of at least one of the following criteria:

1. the substance is not inherently degradable in an inherent biodegradability test; or

2. the substances is predicted to be slowly biodegradable by scientifically valid QSARs, e.g., for the Biodegradation Probability Program, the score for rapid degradation (linear or non-linear model) < 0.5; or

3. the substance is considered to be not rapidly degradable based on indirect evidence, as e.g., knowledge from structurally similar substances; or

4. no other data regarding degradability are available.

ANNEX 4.I

DETERMINATION OF DEGRADABILITY OF ORGANIC SUBSTANCES

124. Organic substances may be degraded by abiotic or biotic processes or by a combination of these. A number of standard procedures or tests for determination of the degradability are available. The general principles of some of these are described below. It is by no way the intention to present a comprehensive review of degradability test methods, but only to place the methods in the context of aquatic hazard classification.

1. ABIOTIC DEGRADABILITY

125. Abiotic degradation comprises chemical transformation and photochemical transformation. Usually abiotic transformations will yield other organic compounds but will not cause a full mineralisation (Schwarzenbach *et al*., 1993). Chemical transformation is defined as transformation that happens without light and without the mediation of organisms whereas photochemical transformations require light.

126. Examples of relevant chemical transformation processes in aqueous environment are hydrolysis, nucleophilic substitution, elimination, oxidation and reduction reactions (Schwarzenbach *et al.*, 1993). Of these, hydrolysis is often considered the most important and it is the only chemical transformation process for which international test guidelines are generally available. The tests for abiotic degradation of chemicals are generally in the form of determination of transformation rates under standardised conditions.

2. HYDROLYSIS

127. Hydrolysis is the reaction of the nucleophiles H_2O or OH^- with a chemical where a (leaving) group of the chemical is exchanged with an OH group. Many compounds, especially acid derivatives, are susceptible to hydrolysis. Hydrolysis can both be abiotic and biotic, but in regard to testing only abiotic hydrolysis is considered. Hydrolysis can take place by different mechanisms at different pHs, neutral, acid- or base-catalysed hydrolysis, and hydrolysis rates may be very dependent on pH.

128. Currently two guidelines for evaluating abiotic hydrolysis are generally available, the OECD Test Guideline 111 Hydrolysis as a function of pH (corresponding to OPPTS 835.2110) and OPPTS 835.2130 Hydrolysis as a function of pH and temperature. In OECD Test Guideline 111, the overall hydrolysis rate at different pHs in pure buffered water is determined. The test is divided in two, a preliminary test that is performed for chemicals with unknown hydrolysis rates and a more detailed test that is performed for chemicals that are known to be hydrolytically unstable and for chemicals for which the preliminary test shows fast hydrolysis. In the preliminary test the concentration of the chemical in buffered solutions at pHs in the range normally found in the environment (pHs of 4, 7 and 9) at 50^{o}C is measured after 5 days. If the concentration of the chemical has decreased less than 10 % it is considered hydrolytically stable, otherwise the detailed test may be performed. In the detailed test, the overall hydrolysis rate is determined at three pHs (4, 7 and 9) by measuring the concentration of the chemical as a function of time. The hydrolysis rate is determined at different temperatures so that interpolations or extrapolations to environmentally relevant temperatures can be made. The OPPTS 835.2130 test is almost identical in design to the OECD Test Guideline 111, the difference mainly being in the treatment of data.

129. It should be noted that apart from hydrolysis the hydrolysis rate constants determined by the tests include all other abiotic transformations that may occur without light under the given test conditions. Good agreement has been found between hydrolysis rates in natural and in pure waters (OPPTS 835.2110).

3. PHOTOLYSIS

130. At present, there is no OECD guideline on aqueous photodegradation, but a guidance document, concerning aquatic direct photolysis, is available (OECD, 1997). The Guidance Document is supposed to form the basis for a scheduled guideline. According to the definitions set out in this Guidance Document, phototransformation of compounds in water can be in the form of primary or secondary phototransformation, where the primary phototransformation (photolysis) can be divided further into direct and indirect photolysis. Direct phototransformation (photolysis) is the case where the chemical absorbs light and as a direct result hereof undergoes transformation. Indirect phototransformation is the case where other excited species transfer energy, electrons or H-atoms to the chemical and thereby induces a transformation (sensitised photolysis). Secondary phototransformation is the case where chemical reactions occur between the chemical and reactive short lived species like hydroxy radicals, peroxy radicals or singlet oxygen that are formed in the presence of light by reactions of excited species like excited humic or fulvic acids or nitrate.

131. The only currently available guidelines on phototransformation of chemicals in water are therefore OPPTS 835.2210 *Direct photolysis rate in water by sunlight* and OPPTS 835.5270 *Indirect photolysis screening test*. The OPPTS 835.2210 test uses a tiered approach. In Tier 1 the maximum direct photolysis rate constant (minimum half-life) is calculated from a measured molar absorptivity. In Tier 2 there are two phases. In Phase 1 the chemical is photolysed with sunlight and an approximate rate constant is obtained. In Phase 2, a more accurate rate constant is determined by using an actinometer that quantifies the intensity of the light that the chemical has actually been exposed to. From the parameters measured, the actual direct photodegradation rate at different temperatures and for different latitudes can be calculated. This degradation rate will only apply to the uppermost layer of a water body, e.g., the first 50 cm or less and only when the water is pure and air saturated which may clearly not be the case in environment. However, the results can be extended over other environmental conditions by the use of a computer programme incorporating attenuation in natural waters and other relevant factors.

132. The OPPTS 835.5270 screening test concerns indirect photolysis of chemicals in waters that contain humic substances. The principle of the test is that in natural waters exposed to natural sunlight a measured phototransformation rate will include both direct and indirect phototransformation, whereas only direct phototransformation will take place in pure water. Therefore, the difference between the direct photodegradation rate in pure water and the total photodegradation in natural water is the sum of indirect photolysis and secondary photodegradation according to the definitions set out in the OECD Guidance Document. In the practical application of the test, commercial humic substances are used to make up a synthetic humic water, which mimics a natural water. It should be noted that the indirect phototransformation rate determined is only valid for the season and latitude for which it is determined and it is not possible to transfer the results to other latitudes and seasons.

4. BIOTIC DEGRADABILITY

133. Only a brief overview of the test methods is given below. For more information, the comprehensive OECD Detailed Review Paper on Biodegradability Testing (OECD, 1995) should be consulted.

5. READY BIODEGRADABILITY

134. Standard tests for determination of the ready biodegradability of organic substances are developed by a number of organisations including OECD (OECD Test Guidelines 301A-F), EU (C.4 tests), OPPTS (835.3110) and ISO (9408, 9439, 10707).

135. The ready biodegradability tests are stringent tests, which provide limited opportunity for biodegradation and acclimatisation to occur. The basic test conditions ensuring these specifications are:

- high concentration of test substance (2-100 mg/L);
- the test substance is the sole carbon and energy source;
- low to medium concentration of inoculum (10^4-10^8 cells/mL);
- no pre-adaptation of inoculum is allowed;
- 28 days test period with a 10-days time window (except for the MITI I method (OECD Test Guideline 301C)) for degradation to take place;
- test temperature < 25°C; and
- pass levels of 70% (DOC removal) or 60% (O_2 demand or CO_2 evolution) demonstrating complete mineralisation (as the remaining carbon of the test substance is assumed to be built into the growing biomass).

136. It is assumed that a positive result in one of the ready biodegradability tests demonstrates that the substance will degrade rapidly in the environment (OECD Test Guidelines).

137. Also the traditional BOD_5 tests (e.g., the EU C.5 test) may demonstrate whether a substance is readily biodegradable. In this test, the relative biochemical oxygen demand in a period of 5 days is compared to the theoretical oxygen demand (ThOD) or, when this is not available, the chemical oxygen demand (COD). The test is completed within five days and consequently, the pass level defined in the proposed hazard classification criteria at 50% is lower than in the ready biodegradability tests.

138. The screening test for biodegradability in seawater (OECD Test Guideline 306) may be seen as seawater parallel to the ready biodegradability tests. Substances that reach the pass level in OECD Test Guideline 306 (i.e., >70% DOC removal or >60 theoretical oxygen demand) may be regarded as readily biodegradable, since the degradation potential is normally lower in seawater than in the freshwater degradation tests.

6. INHERENT BIODEGRADABILITY

139. Tests for inherent biodegradability are designed to assess whether a substance has any potential for biodegradation. Examples of such tests are the OECD Test Guidelines 302A-C tests, the EU C.9 and C.12 tests, and the ASTM E 1625-94 test.

140. The basic test conditions favouring an assessment of the inherent biodegradation potential are:

- a prolonged exposure of the test substance to the inoculum allowing adaptation within the test period;
- a high concentration of micro-organisms;
- a favourable substance/biomass ratio.

141. A positive result in an inherent test indicates that the test substance will not persist indefinitely in the environment, however a rapid and complete biodegradation can not be assumed. A result demonstrating more than 70% mineralisation indicates a potential for ultimate biodegradation, a degradation of more than 20% indicates inherent, primary biodegradation, and a result of less than 20% indicates that the substance is persistent. Thus, a negative result means that non-biodegradability (persistence) should be assumed (OECD Test Guidelines).

142. In many inherent biodegradability tests only the disappearance of the test substance is measured. Such a result only demonstrates a primary biodegradability and not a total mineralisation. Thus, more or less persistent degradation products may have been formed. Primary biodegradation of a substance is no indication of ultimate degradability in the environment.

143. The OECD inherent biodegradation tests are very different in their approach and especially, the MITI II test (OECD Test Guideline 302C) employs a concentration of inoculum that is only three times higher than in the corresponding MITI I ready biodegradability test (OECD Test Guideline 301C). Also the Zahn-Wellens test (OECD Test Guideline 302B) is a relatively "weak" inherent test. However, although the degradation potential in these tests is not very much stronger than in the ready biodegradability tests, the results can not be extrapolated to conditions in the ready biodegradability tests and in the aquatic environment.

7. AQUATIC SIMULATION TESTS

144. A simulation test attempts to simulate biodegradation in a specific aquatic environment. As examples of a standard test for simulation of degradation in the aquatic environment may be mentioned the ISO/DS14592 Shake flask batch test with surface water or surface water/sediment suspensions (Nyholm and Toräng, 1999), the ASTM E 1279-89(95) test on biodegradation by a shake-flask die-away method and the similar OPPTS 835.3170 test. Such test methods are often referred to as river die-away tests.

145. The features of the tests that ensures simulation of the conditions in the aquatic environment are:

- use of a natural water (and sediment) sample as inoculum; and
- low concentration of test substance (1-100 µg/L) ensuring first-order degradation kinetics.

146. The use of a radiolabelled test compound is recommended as this facilitates the determination of the ultimate degradation. If only the removal of the test substance by chemical analysis is determined, only the primary degradability is determined. From observation of the degradation kinetics, the rate constant for the degradation can be derived. Due to the low concentration of the test substance, first-order degradation kinetics are assumed to prevail.

147. The test may also be conducted with natural sediment simulating the conditions in the sediment compartment. Moreover, by sterilising the samples, the abiotic degradation under the test conditions can be determined.

8. STP SIMULATION TESTS

148. Tests are also available for simulating the degradability in a sewage treatment plant (STP), e.g., the OECD Test Guideline 303A Coupled Unit test, ISO 11733 Activated sludge simulation test, and the EU C.10 test. Recently, a new simulation test employing low concentrations of organic pollutants has been proposed (Nyholm et. al., 1996).

9. ANAEROBIC DEGRADABILITY

149. Test methods for anaerobic biodegradability determine the intrinsic potential of the test substance to undergo biodegradation under anaerobic conditions. Examples of such tests are the ISO 11734:1995(E) test, the ASTM E 1196-92 test and the OPPTS 835.3400 test.

150. The potential for anaerobic degradation is determined during a period of up to eight weeks and with the test conditions indicated below:

- performance of the test in sealed vessels in the absence of O_2 (initially in a pure N_2 atmosphere);
- use of digested sludge;
- a test temperature of 35°C; and
- determination of head-space gas pressure (CO_2 and CH_4 formation).

151. The ultimate degradation is determined by determining the gas production. However, also primary degradation may be determined by measuring the remaining parent substance.

10. DEGRADATION IN SOIL AND SEDIMENT

152. Many chemical substances end up in the soil or sediment compartments and an assessment of their degradability in these environments may therefore be of importance. Among standard methods may be mentioned the OECD Test Guideline 304A test on inherent biodegradability in soil, which corresponds to the OPPTS 835.3300 test.

153. The special test characteristics ensuring the determination of the inherent degradability in soil are:

- natural soil samples are used without additional inoculation;
- radiolabelled test substance is used; and
- evolution of radiolabelled CO_2 is determined.

154. A standard method for determining the biodegradation in sediment is the OPPTS 835.3180 Sediment/water microcosm biodegradation test. Microcosms containing sediment and water are collected from test sites and test compounds are introduced into the system. Disappearance of the parent compound (i.e., primary biodegradation) and, if feasible, appearance of metabolites or measurements of ultimate biodegradation may be made.

155. Currently, two new OECD guidelines are being drafted on aerobic and anaerobic transformation in soil (OECD Test Guideline, 1999a) and in aquatic sediment systems (OECD Test Guideline 1999b), respectively. The experiments are performed to determine the rate of transformation of the test substance and the nature and rates of formation and decline of transformation products under environmentally realistic conditions including a realistic

concentration of the test substance. Either complete mineralisation or primary degradability may be determined depending on the analytical method employed for determining the transformation of the test substance.

11. METHODS FOR ESTIMATING BIODEGRADABILITY

156. In recent years, possibilities for estimating environmental properties of chemical substances have been developed and, among these, also methods for predicting the biodegradability potential of organic substances (e.g., the Syracuse Research Corporation's Biodegradability Probability Program, BIOWIN). Reviews of methods have been performed by OECD (1993) and by Langenberg *et al.* (1996). They show that group contribution methods seem to be the most successful methods. Of these, the Biodegradation Probability Program (BIOWIN) seems to have the broadest application. It gives a qualitative estimate of the probability of slow or fast biodegradation in the presence of a mixed population of environmental micro-organisms. The applicability of this program has been evaluated by the US EPA/EC Joint Project on the Evaluation of (Q)SARs (OECD, 1994), and by Pedersen *et al.* (1995). The latter is briefly referred below.

157. A validation set of experimentally determined biodegradation data was selected among the data from MITI (1992), but excluding substances for which no precise degradation data were available and substances already used for development of the programme. The validation set then consisted of 304 substances. The biodegradability of these substances were estimated by use of the programme's non-linear estimation module (the most reliable) and the results compared with the measured data. 162 substances were predicted to degrade "fast", but only 41 (25%) were actually readily degradable in the MITI I test. 142 substances were predicted to degrade "slowly", which was confirmed by 138 (97%) substances being not readily degradable in the MITI I test. Thus, it was concluded that the programme may be used for classification purposes only when no experimental degradation data can be obtained, and when the programme predicts a substance to be degraded "slowly". In this case, the substance can be regarded as not rapidly degradable.

158. The same conclusion was reached in the US EPA/EC Joint Project on the Evaluation of (Q)SARs by use of experimental and QSAR data on new substances notified in the EU. The evaluation was based on an analysis of QSAR predictions on 115 new substances also tested experimentally in ready biodegradability tests. Only 9 of the substances included in this analysis were readily biodegradable. The employed QSAR methodology is not fully specified in the final report of the Joint US EPA/EC project (OECD, 1994), but it is likely that the majority of predictions were made by using methods which later have been integrated in the Biodegradation Probability Program.

159. Also in the EU TGD (EC, 1996) it is recommended that estimated biodegradability by use of the Biodegradation Probability Program is used only in a conservative way, i.e., when the programme predicts fast biodegradation, this result should not be taken into consideration, whereas predictions of slow biodegradation may be considered (EC, 1996).

160. Thus, the use of results of the Biodegradability Probability Program in a conservative way may fulfil the needs for evaluating biodegradability of some of the large number of substances for which no experimental degradation data are available.

ANNEX 4.II

FACTORS INFLUENCING DEGRADABILITY IN THE AUQATIC ENVIRONMENT

161. The OECD classification criteria are considering the hazards to the aquatic environment only. However, the hazard classification is primarily based on data prepared by conduction of tests under laboratory conditions that only seldom are similar to the conditions in the environment. Thus, the interpretation of laboratory test data for prediction of the hazards in the aquatic environment should be considered.

162. Interpretation of test results on biodegradability of organic substances has been considered in the OECD Detailed Review Paper on Biodegradability Testing (OECD, 1995).

163. The conditions in the environment are typically very different from the conditions in the standardised test systems, which make the extrapolation of degradation data from laboratory tests to the environment difficult. Among the differences, the following have significant influence on the degradability:

- Organism related factors (presence of competent micro-organisms);
- Substrate related factors (concentration of the substance and presence of other substrates); and
- Environment related factors (physico-chemical conditions, presence of nutrients, bioavailability of the substance).

164. These aspects will be discussed further below.

1. PRESENCE OF COMPETENT MICRO-ORGANISMS

165. Biodegradation in the aquatic environment is dependent on the presence of competent micro-organisms in sufficient numbers. The natural microbial communities consist of a very diverse biomass and when a 'new' substance is introduced in a sufficiently high concentration, the biomass may be adapted to degrade this substance. Frequently, the adaptation of the microbial population is caused by the growth of specific degraders that by nature are competent to degrade the substance. However, also other processes as enzyme induction, exchange of genetic material and development of tolerance to toxicity may be involved.

166. Adaptation takes place during a "lag" phase, which is the time period from the onset of the exposure until a significant degradation begins. It seems obvious that the length of the lag phase will depend on the initial presence of competent degraders. This will again depend on the history of the microbial community, i.e., whether the community formerly has been exposed to the substance. This means that when a xenobiotic substance has been used and emitted ubiquitously in a number of years, the likelihood of finding competent degraders will increase. This will especially be the case in environments receiving emissions as e.g., biological wastewater treatment plants. Often more consistent degradation results are found in tests where inocula from polluted waters are used compared to tests with inocula from unpolluted water (OECD, 1995; Nyholm and Ingerslev, 1997).

167. A number of factors determine whether the potential for adaptation in the aquatic environment is comparable with the potential in laboratory tests. Among other things adaptation depends on:

- initial number of competent degraders in the biomass (fraction and number);

- presence of surfaces for attachment;
- concentration and availability of substrate; and
- presence of other substrates.

168. The length of the lag phase depends on the initial number of competent degraders and, for toxic substances, the survival and recovery of these. In standard ready biodegradability tests, the inoculum is sampled in sewage treatment plants. As the load with pollutants is normally higher than in the environment, both the fraction and the number of competent degraders may be higher than in the less polluted aquatic environment. It is, however, difficult to estimate how much longer the lag phase will be in the aquatic environment than in a laboratory test due to the likely lower initial number of competent degraders.

169. Over long periods of time, the initial concentration of competent degraders is not important as they will grow up when a suitable substrate is present in sufficient concentrations. However, if the degradability in a short period of time is of concern, the initial concentration of competent degrading micro-organisms should be considered (Scow, 1982).

170. The presence of flocs, aggregates and attached micro-organisms may also enhance adaptation by e.g., development of microbial niches with consortia of micro-organisms. This is of importance when considering the capability of adaptation in the diverse environments in sewage treatment plants or in sediment or soil. However, the total number of micro-organisms in ready biodegradability tests and in the aquatic environment are of the same orders of magnitude (10^4-10^8 cells/mL in ready biodegradability tests and 10^3-10^6 cells/mL or more in surface water (Scow, 1982). Thus, this factor is probably of minor importance.

171. When discussing the extrapolation to environmental conditions it may be valuable to discriminate between oligotrophic and eutrophic environments. Micro-organisms thriving under oligotrophic conditions are able to mineralise organic substrates at low concentrations (fractions of mg C/L), and they normally have a greater affinity for the substrate but lower growth rates and higher generation times than eutrophic organisms (OECD, 1995). Moreover, oligotrophs are unable to degrade chemicals in concentrations higher than 1 mg/L and may even be inhibited at high concentrations. Opposite to that, eutrophs require higher substrate concentrations before mineralisation begins and they thrive at higher concentrations than oligotrophs. Thus, the lower threshold limit for degradation in the aquatic environment will depend on whether the microbial population is an oligotroph or an eutroph population. It is, however, not clear whether oligotrophs and eutrophs are different species or whether there is only an oligotrophic and an eutrophic way of life (OECD, 1995). Most pollutants reach the aquatic environment directly through discharge of wastewater and consequently, these recipients are mostly eutrophic.

172. From the above discussion it may thus be concluded that the chance of presence of competent degraders is greatest in highly exposed environments, i.e., in environments continuously receiving substances (which more frequently occurs for high production volume chemicals than for low production volume chemicals). These environments are often eutrophic and therefore, the degradation may require relatively high concentrations of substances before onset. On the other hand, in pristine waters competent species may be lacking, especially species capable of degradation of chemicals only occasionally released as low production volume chemicals.

2. SUBSTRATE RELATED FACTORS

2.1 Concentration of test substance

173. In most laboratory tests, the test substance is applied in very high concentrations (2-100 mg/L) compared to the concentrations in the lower µg/L range that may be expected in the aquatic environment. In general, growth of micro-organisms is not supported when a substrate is present in concentrations below a threshold level of around 10 µg/L and at lower concentrations, even the energy requirement for maintenance is not met (OECD, 1995). The reason for this lower threshold level is possibly a lack of sufficient stimulus to initiate an enzymatic response (Scow, 1982). This means in general that the concentrations of many substances in the aquatic environment are at a level where they can only hardly be the primary substrate for degrading micro-organisms.

174. Moreover, the degradation kinetics depends on substance concentration (S_0) compared with the saturation constant (K_s) as described in the Monod equation. The saturation constant is the concentration of the substrate resulting in a specific growth rate of 50% of the maximum specific growth rate. At substrate concentrations much lower than the saturation constant, which is the normal situation in most of the aquatic environment, the degradation can be described by first order or logistic kinetics (OECD, 1995). When a low density of micro-organisms (lower than 10^3-10^5 cells/mL) prevails (e.g., in oligotrophic waters), the population grows at ever decreasing rates which is typical of logistic kinetics. At a higher density of micro-organisms (e.g., in eutrophic waters), the substrate concentration is not high enough to support growth of the cells and first order kinetics apply, i.e., the degradation rate is proportional with the substance concentration. In practice, it may be impossible to distinguish between the two types of degradation kinetics due to uncertainty of the data (OECD, 1995).

175. In conclusion, substances in low concentrations (i.e., below 10 µg/L) are probably not degraded as primary substrates in the aquatic environment. At higher concentrations, readily degradable substances will probably be degraded as primary substrates in the environment at a degradation rate more or less proportional with the concentration of the substance. The degradation of substances as secondary substrates is discussed below.

2.2 Presence of other substrates

176. In the standard tests, the test substance is applied as the sole substrate for the micro-organisms while in the environment, a large number of other substrates are present. In natural waters, concentrations of dissolved organic carbon are often found in the range 1-10 mg C/L, i.e., up to a factor 1000 higher than a pollutant. However, much of this organic carbon is relatively persistent with an increasing fraction of persistent matter the longer the distance from the shore.

177. Bacteria in natural waters are primarily nourishing on exudates from algae. These exudates are mineralised very quickly (within minutes) demonstrating that there is a high degradation potential in the natural micro-organism communities. Thus, as micro-organisms compete for the variety of substrates in natural waters, there is a selection pressure among micro-organisms resulting in growth of opportunistic species capable of nourishing on quickly mineralised substrates, while growth of more specialised species is suppressed. Experiences from isolation of bacteria capable of degrading various xenobiotics have demonstrated that these organisms are often growing relatively slowly and survive on complex carbon sources in competition with more rapidly growing bacteria. When competent micro-organisms are present in the environment, their numbers may increase if the specific xenobiotic substrate is continuously released and reach a concentration in the environment sufficient to support growth. However, most of the organic pollutants in the

aquatic environment are present in low concentrations and will only be degraded as secondary substrates not supporting growth.

178. On the other hand, the presence of quickly mineralised substrates in higher concentrations may facilitate an initial transformation of the xenobiotic molecule by co-metabolism. The co-metabolised substance may then be available for further degradation and mineralisation. Thus, the presence of other substrates may increase the possibilities for a substance to be degraded.

179. It may then be concluded that the presence of a variety of substrates in natural waters and among them quickly mineralised substrates, may on the one hand cause a selection pressure suppressing growth of micro-organisms competent of degrading micro-pollutants. On the other hand it may facilitate an increased degradation by an initial co-metabolism followed by a further mineralisation. The relative importance of these processes under natural conditions may vary depending on both the environmental conditions and the substance and no generalisation can yet be established.

3. ENVIRONMENT RELATED FACTORS

180. The environmental variables control the general microbial activity rather than specific degradation processes. However, the significance of the influence varies between different ecosystems and microbial species (Scow, 1982).

3.1 Redox potential

181. One of the most important environment related factors influencing the degradability is probably the presence of oxygen. The oxygen content and the related redox potential determines the presence of different types of micro-organisms in aquatic environments with aerobic organisms present in the water phase, in the upper layer of sediments and in parts of sewage treatment plants, and anaerobic organisms present in sediments and parts of sewage treatment plants. In most parts of the water phase, aerobic conditions are prevailing and the prediction of the biodegradability should be based on results from aerobic tests. However, in some aquatic environments the oxygen content may be very low in periods of the year due to eutrophication and the following decay of produced organic matter. In these periods, aerobic organisms will not be able to degrade the chemical, but anaerobic processes may take over if the chemical is degradable under anaerobic conditions.

3.2 Temperature

182. Another important parameter is the temperature. Most laboratory tests are performed at 20-25°C (standard aerobic ready biodegradability tests), but anaerobic tests may be performed at 35°C as this better mimics the conditions in a sludge reactor. Microbial activity is found in the environment at temperatures ranging from below 0°C to 100°C. However, optimum temperatures are probably in the range from 10°C to 30°C and roughly, the degradation rate doubles for every 10°C increase of temperature in this range (de Henau, 1993). Outside this optimum range the activity of the degraders is reduced drastically although some specialised species (termo- and psycrophilic bacteria) may thrive. When extrapolating from laboratory conditions, it should be considered that some aquatic environments are covered by ice in substantial periods of the year and that only minor or even no degradation can be expected during the winter season.

3.3 pH

183. Active micro-organisms are found in the entire pH range found in the environment. However, for bacteria as a group, slightly alkaline conditions favour the activity and the optimum pH range is 6-8. At a pH lower than 5, the metabolic activity in bacteria is significantly decreased. For fungi as a group, slightly acidic conditions favour the activity with an optimum pH range of 5-6 (Scow, 1982). Thus, an optimum for the degrading activity of micro-organisms will probably be within the pH range of 5-8, which is the range most often prevailing in the aquatic environment.

3.4 Presence of nutrients

184. The presence of inorganic nutrients (nitrogen and phosphorus) is often required for microbial growth. However, these are only seldom the activity limiting factors in the aquatic environment where growth of micro-organisms is often substrate limited. However, the presence of nutrient influences the growth of primary producers and then again the availability of readily mineralised exudates.

ANNEX 4.III

TEST GUIDELINES

185. Most of the guidelines mentioned are found in compilations from the organisation issuing them. The main references to these are:

- EC guidelines: European Commission (1996). Classification, Packaging and Labelling of Dangerous Substances in the European Union. Part 2 – Testing Methods. European Commission. 1997. ISBN92-828-0076-8. (Homepage: http://ecb.ei.jrc.it/testing-methods/);
- ISO guidelines: Available from the national standardisation organisations or ISO (Homepage: http://www.iso.ch/);
- OECD guidelines for the testing of chemicals. OECD. Paris. 1993 with regular updates (Homepage: http://www.oecd.org/ehs/test/testlist.htm);
- OPPTS guidelines: US-EPA's homepage: http://www.epa.gov/opptsfrs/home/guidelin.htm;
- ASTM : ASTM´s homepage: http://www.astm.org. Further search via "standards".

ASTM E 1196-92.

ASTM E 1279-89(95) Standard test method for biodegradation by a shake-flask die-away method.

ASTM E 1625-94 Standard test method for determining biodegradability of organic chemicals in semi-continuous activated sludge (SCAS).

EC C.4. A to F: Determination of ready biodegradability. Directive 67/548/EEC, AnnexV. (1992).

EC C.5. Degradation: biochemical oxygen demand. Directive 67/548/EEC, AnnexV. (1992).

EC C.7. Degradation: abiotic degradation: hydrolysis as a function of pH. Directive 67/548/EEC, AnnexV. (1992).

EC C.9. Biodegradation: Zahn-Wellens test. Directive 67/548/EEC, AnnexV. (1988).

EC C.10. Biodegradation: Activated sludge simulation tests. Directive 67/548/EEC, AnnexV. (1998).

EC C.11. Biodegradation: Activated sludge respiration inhibition test. Directive 67/548/EEC, AnnexV.(1988).

EC C.12. Biodegradation: Modified SCAS test. Directive 67/548/EEC, AnnexV. (1998).

ISO 9408 (1991). Water quality - Evaluation in an aqueous medium of the "ultimate" biodegradability of organic compounds - Method by determining the oxygen demand in a closed respirometer.

ISO 9439 (1990). Water quality - Evaluation in an aqueous medium of the "ultimate" biodegradability of organic compounds - Method by analysis of released carbon dioxide.

ISO 9509 (1996). Water quality - Method for assessing the inhibition of nitrification of activated sludge micro-organisms by chemicals and wastewaters.

ISO 9887 (1992). Water quality - Evaluation of the aerobic biodegradability of organic compounds in an aqueous medium - Semicontinuous activated sludge method (SCAS).

ISO 9888 (1991). Water quality - Evaluation of the aerobic biodegradability of organic compounds in an aqueous medium - Static test (Zahn-Wellens method).

ISO 10707 (1994). Water quality - Evaluation in an aqueous medium of the "ultimate" biodegradability of organic compounds - Method by analysis of biochemical oxygen demand (closed bottle test).

ISO 11348 (1997). Water quality - Determination of the inhibitory effect of water samples on the light emission of *Vibrio fischeri* (Luminescent bacteria test).

ISO 11733 (1994). Water quality - Evaluation of the elimination and biodegradability of organic compounds in an aqueous medium - Activated sludge simulation test.

ISO 11734 (1995). Water quality - Evaluation of the "ultimate" anaerobic biodegradability of organic compounds in digested sludge - Method by measurement of the biogas production.

ISO/DIS 14592 .(1999) Water quality - Evaluation of the aerobic biodegradability of organic compounds at low concentrations in water. Part 1: Shake flask batch test with surface water or surface water/sediment suspensions (22.11.1999).

OECD Test Guideline 111 (1981). Hydrolysis as a function of pH. OECD guidelines for testing of chemicals.

OECD Test Guideline 209 (1984). Activated sludge, respiration inhibition test. OECD guidelines for testing of chemicals.

OECD Test Guideline 301 (1992). Ready biodegradability. OECD guidelines for testing of chemicals.

OECD Test Guideline 302A (1981). Inherent biodegradability: Modified SCAS test. OECD guidelines for testing of chemicals.

OECD Test Guideline 302B (1992). Zahn-Wellens/EMPA test. OECD guidelines for testing of chemicals.

OECD Test Guideline 302C (1981). Inherent biodegradability: Modified MITI test (II). OECD guidelines for testing of chemicals.

OECD Test Guideline 303A (1981). Simulation test - aerobic sewage treatment: Coupled units test. OECD guidelines for testing of chemicals. Draft update available 1999.

OECD Test Guideline 304A (1981). Inherent biodegradability in soil. OECD guidelines for testing of chemicals.

OECD Test Guideline 306 (1992). Biodegradability in seawater. OECD guidelines for testing of chemicals.

OECD (1998~~b~~). Aerobic and anaerobic transformation in aquatic sediment systems. Draft proposal for a new guideline, December 1999.

OECD (1999). Aerobic and anaerobic transformation in soil. Final text of a draft proposal for a new guideline, October. 1999.

OECD (2000). Simulation test - Aerobic Transformation in Surface Water. Draft proposal for a new guideline, May 2000.

OPPTS 835.2110 Hydrolysis as a function of pH.

OPPTS 835.2130 Hydrolysis as a function of pH and temperature.

OPPTS 835.2210 Direct photolysis rate in water by sunlight.

OPPTS 835.3110 Ready biodegradability.

OPPTS 835.3170 Shake flask die-away test.

OPPTS 835.3180 Sediment/water microcosm biodegradability test.

OPPTS 835.3200 Zahn-Wellens/EMPA test.

OPPTS 835.3210 Modified SCAS test.

OPPTS 835.3300 Soil biodegradation.

OPPTS 835.3400 Anaerobic biodegradability of organic chemicals.

OPPTS 835.5270 Indirect photolysis screening test: Sunlight photolysis in waters containing dissolved humic substances.

ANNEX 4.IV

REFERENCES

Boesten J.J.T.I. & A.M.A. van der Linden (1991). Modeling the influence of sorption and transformation on pesticide leaching and persistence. *J. Environ. Qual.* 20, 425-435.

Boethling R.S., P.H. Howard, J.A. Beauman & M.E. Larosche (1995). Factors for intermedia extrapolation in biodegradability assessment. *Chemosphere* 30(4), 741-752.

de Henau H. (1993). Biodegradation. In: P. Calow. Handbook of Ecotoxicology, vol. I. Blackwell Scientific Publications, London. Chapter 18, pp. 355-377.

EC (1996). Technical guidance documents in support of the Commission Directive 93/67/EEC on risk assessment for new notified substances and the Commission Regulation (EC) No. 1488/94 on risk assessment for existing substances. European Commission, Ispra.

ECETOC (1998): QSARs in the Assessment of the Environmental Fate and Effects of Chemicals, Technical report No. 74. Brussels, June 1998.

Federle T.W., S.D. Gasior & B.A. Nuck (1997). Extrapolating mineralisation rates from the ready CO_2 screening test to activated sludge, river water, and soil. *Environmental Toxicology and Chemistry* 16, 127-134.

Langenberg J.H., W.J.G.M. Peijnenburg & E. Rorije (1996). On the usefulness and reliability of existing QSBRs for risk assessment and priority setting. *SAR and QSAR in Environmental Research* 5, 1-16.

Loonen H., F. Lindgren, B. Hansen & W. Karcher (1996). Prediction of biodegradability from chemical structure. In: Peijnenburg W.J.G.M. & J. Damborsky (eds.). Biodegradability Prediction. Kluwer Academic Publishers.

MITI (1992). Biodegradation and bioaccumulation data on existing data based on the CSCL Japan. Japan chemical industry, Ecology-toxicology & information center. ISBN 4-89074-101-1.

Niemelä J (2000). Personal communication to OECD Environment Directorate, 20 March 2000.

Nyholm N., U.T. Berg & F. Ingerslev (1996). Activated sludge biodegradability simulation test. Danish EPA, Environmental Report No. 337.

Nyholm N. & F. Ingerslev (1997). Kinetic biodegradation tests with low test substance concentrations: Shake flask test with surface water and short term rate measurement in activated sludge. In: Hales S.G. (ed.). Biodegradation Kinetics: Generation and use of data for regulatory decision making. From the SETAC-Europe Workshop. Port- Sunlight. September 1996. pp. 101-115. SETAC-Europe, Brussels.

Nyholm N. & L. Toräng (1999). Report of 1998/1999 Ring-test: Shalke flask batch test with surface water or surface water / sediment suspensions. ISO/CD 14592-1 Water Quality- Evaluation of the aerobic biodegradability of organic compounds at low concentrations, ISO/TC 147/ SC5/WG4 Biodegradability.

OECD (1993). Structure-Activity Relationships for Biodegradation. OECD Environment Monographs No. 68. Paris 1993.

OECD (1994): "US EPA/EC Joint Project on the Evaluation of (Quantitative) Structure Activity Relationships." OECD Environment Monograph No. 88. Paris.

OECD (1995). Detailed Review Paper on Biodegradability Testing. OECD Environmental Monograph No. 98. Paris.

OECD (1997). Guidance document on direct phototransformation of chemical in water. OECD/GD(97)21. Paris.

OECD (1998). Harmonized integrated hazard classification system for human health and environmental effects of chemical substances. Paris. http://www.oecd.org/ehs/Class/HCL6.htm.

Pedersen F., H. Tyle, J. R. Niemelä, B. Guttmann. L. Lander & A. Wedebrand (1995). Environmental Hazard Classification - data collection and interpretation guide for substances to be evaluated for classification as dangerous for the environment. Nordic Council of Ministers. 2nd edition. TemaNord 1995:581, 166 pp.

Schwarzenbach R.P., P.M. Gschwend & D.M. Imboden (1993). Environmental organic chemistry 1st ed. John Wiley & Sons, Inc. New York.

Scow K.M. (1982). Rate of biodegradation. In: Lyman W.J., W.F. Reehl & D.H. Rosenblatt (1982): Handbook of Chemical Property Estimation Methods Environmental Behaviour of Organic Compounds. American Chemical Society. Washington DC (ISBN 0-8412-1761-0). Chapter 9.

Struijs J. & R. van den Berg (1995). Standardized biodegradability tests: Extrapolation to aerobic environments. *Wat. Res.* 29(1), 255-262.

Syracuse Research Corporation. Biodegradation Probability Program (BIOWIN). Syracuse. N.Y. http://esc.syrres.com/~esc1/biodeg.htm.

Westermann P., B.K. Ahring & R.A. Mah (1989). Temperature compensation in *Methanosarcina barkeri* by modulation of hydrogen and acetate affinity. *Applied and Environmental Microbiology* 55(5), 1262-1266.

5. BIOACCUMULATION

5.1 INTRODUCTION

186. Bioaccumulation is one of the important intrinsic properties of chemical substances that determine the potential environmental hazard. Bioaccumulation of a substance into an organism is not a hazard in itself, but bioconcentration and bioaccumulation will result in a body burden, which may or may not lead to toxic effects. In the harmonised integrated hazard classification system for human health and environmental effects of chemical substances (OECD, 1998), the wording "potential for bioaccumulation" is given. A distinction should, however, be drawn between bioconcentration and bioaccumulation. Here bioconcentration is defined as the net result of uptake, transformation, and elimination of a substance in an organism due to waterborne exposure, whereas bioaccumulation includes all routes of exposure (i.e., via air, water, sediment/soil, and food). Finally, biomagnification is defined as accumulation and transfer of substances via the food chain, resulting in an increase of internal concentrations in organisms on higher levels of the trophic chain (European Commission, 1996). For most organic chemicals uptake from water (bioconcentration) is believed to be the predominant route of uptake. Only for very hydrophobic substances does uptake from food becomes important. Also, the harmonised classification criteria use the bioconcentration factor (or the octanol/water partition coefficient) as the measure of the potential for bioaccumulation. For these reasons, the present guidance document only considers bioconcentration and does not discuss uptake via food or other routes.

187. Classification of a chemical substance is primarily based on its intrinsic properties. However, the degree of bioconcentration also depends on factors such as the degree of bioavailability, the physiology of test organism, maintenance of constant exposure concentration, exposure duration, metabolism inside the body of the target organism and excretion from the body. The interpretation of the bioconcentration potential in a chemical classification context therefore requires an evaluation of the intrinsic properties of the substance, as well as of the experimental conditions under which bioconcentration factor (BCF) has been determined. Based on the guide, a decision scheme for application of bioconcentration data or log K_{ow} data for classification purposes has been developed. The emphasis of the present chapter is organic substances and organo-metals. Bioaccumulation of metals is also discussed in Chapter 7.

188. Data on bioconcentration properties of a substance may be available from standardised tests or may be estimated from the structure of the molecule. The interpretation of such bioconcentration data for classification purposes often requires detailed evaluation of test data. In order to facilitate this evaluation two additional annexes are enclosed. These annexes describe available methods (Annex 5.I) and factors influencing the bioconcentration potential (Annex 5.II). Finally, a list of standardised experimental methods for determination of bioconcentration and K_{ow} are attached (Annex 5.III) together with a list of references (Annex 5.IV).

5.2 INTERPRETATION OF BIOCONCENTRATION DATA

189. Environmental hazard classification of a chemical substance is normally based on existing data on its environmental properties. Test data will only seldom be produced with the main purpose of facilitating a classification. Often a diverse range of test data is available which does not necessarily match the classification criteria. Consequently, guidance is needed on interpretation of existing test data in the context of hazard classification.

190. Bioconcentration of an organic substance can be experimentally determined in bioconcentration experiments, during which BCF is measured as the concentration in the organism

relative to the concentration in water under steady-state conditions and/or estimated from the uptake rate constant (k_1) and the elimination rate constant (k_2) (OECD 305, 1996). In general, the potential of an organic substance to bioconcentrate is primarily related to the lipophilicity of the substance. A measure of lipophilicity is the n-octanol-water partition coefficient (K_{ow}) which, for lipophilic non-ionic organic substances, undergoing minimal metabolism or biotransformation within the organism, is correlated with the bioconcentration factor. Therefore, K_{ow} is often used for estimating the bioconcentration of organic substances, based on the empirical relationship between log BCF and log K_{ow}. For most organic substances, estimation methods are available for calculating the K_{ow}. Data on the bioconcentration properties of a substance may thus be (1) experimentally determined, (2) estimated from experimentally determined K_{ow}, or (3) estimated from K_{ow} values derived by use of Quantitative Structure Activity Relationships (QSARs). Guidance for interpretation of such data is given below together with guidance on assessment of chemical categories, which need special attention.

5.2.1 Bioconcentration factor (BCF)

191. The bioconcentration factor is defined as the ratio on a weight basis between the concentration of the chemical in biota and the concentration in the surrounding medium, here water, at steady state. BCF can thus be experimentally derived under steady-state conditions, on the basis of measured concentrations. However, BCF can also be calculated as the ratio between the first-order uptake and elimination rate constants; a method which does not require equilibrium conditions.

192. Different test guidelines for the experimental determination of bioconcentration in fish have been documented and adopted, the most generally applied being the OECD test guideline (OECD 305, 1996).

193. Experimentally derived BCF values of high quality are ultimately preferred for classification purposes as such data override surrogate data, e.g., K_{ow}.

194. High quality data are defined as data where the validity criteria for the test method applied are fulfilled and described, e.g., maintenance of constant exposure concentration; oxygen and temperature variations, and documentation that steady-state conditions have been reached, etc. The experiment will be regarded as a high-quality study, if a proper description is provided (e.g., by Good Laboratory Practice (GLP)) allowing verification that validity criteria are fulfilled. In addition, an appropriate analytical method must be used to quantify the chemical and its toxic metabolites in the water and fish tissue (see Annex 1 for further details).

195. BCF values of low or uncertain quality may give a false and too low BCF value; e.g., application of measured concentrations of the test substance in fish and water, but measured after a too short exposure period in which steady-state conditions have not been reached (cf. OECD 306, 1996, regarding estimation of time to equilibrium). Therefore, such data should be carefully evaluated before use and consideration should be given to using K_{ow} instead.

196. If there is no BCF value for fish species, high-quality data on the BCF value for other species may be used (e.g., BCF determined on blue mussel, oyster, scallop (ASTM E 1022-94)). Reported BCFs for microalgae should be used with caution.

197. For highly lipophilic substances, e.g., with log K_{ow} above 6, experimentally derived BCF values tend to decrease with increasing log K_{ow}. Conceptual explanations of this non-linearity mainly refer to either reduced membrane permeation kinetics or reduced biotic lipid solubility for

large molecules. A low bioavailability and uptake of these substances in the organism will thus occur. Other factors comprise experimental artefacts, such as equilibrium not being reached, reduced bioavailability due to sorption to organic matter in the aqueous phase, and analytical errors. Special care should thus be taken when evaluating experimental data on BCF for highly lipophilic substances as these data will have a much higher level of uncertainty than BCF values determined for less lipophilic substances.

BCF in different test species

198. BCF values used for classification are based on whole body measurements. As stated previously, the optimal data for classification are BCF values derived using the OECD 305 test method or internationally equivalent methods, which uses small fish. Due to the higher gill surface to weight ratio for smaller organisms than larger organisms, steady-state conditions will be reached sooner in smaller organisms than in larger ones. The size of the organisms (fish) used in bioconcentration studies is thus of considerable importance in relation to the time used in the uptake phase, when the reported BCF value is based solely on measured concentrations in fish and water at steady-state. Thus, if large fish, e.g., adult salmon, have been used in bioconcentration studies, it should be evaluated whether the uptake period was sufficiently long for steady state to be reached or to allow for a kinetic uptake rate constant to be determined precisely.

199. Furthermore, when using existing data for classification, it is possible that the BCF values could be derived from several different fish or other aquatic species (e.g., clams) and for different organs in the fish. Thus, to compare these data to each other and to the criteria, some common basis or normalisation will be required. It has been noted that there is a close relationship between the lipid content of a fish or an aquatic organism and the observed BCF value. Therefore, when comparing BCF values across different fish species or when converting BCF values for specific organs to whole body BCFs, the common approach is to express the BCF values on a common lipid content. If e.g., whole body BCF values or BCF values for specific organs are found in the literature, the first step is to calculate the BCF on a % lipid basis using the relative content of fat in the fish (cf. literature/test guideline for typical fat content of the test species) or the organ. In the second step the BCF for the whole body for a typical aquatic organism (i.e., small fish) is calculated assuming a common default lipid content. A default value of 5% is most commonly used (Pedersen *et al.*, 1995) as this represents the average lipid content of the small fish used in OECD 305 (1996).

200. Generally, the highest valid BCF value expressed on this common lipid basis is used to determine the wet weight based BCF-value in relation to the cut off value for BCF of 500 of the harmonised classification criteria.

Use of radiolabelled substances

201. The use of radiolabelled test substances can facilitate the analysis of water and fish samples. However, unless combined with a specific analytical method, the total radioactivity measurements potentially reflect the presence of the parent substance as well as possible metabolite(s) and possible metabolised carbon, which have been incorporated in the fish tissue in organic molecules. BCF values determined by use of radiolabelled test substances are therefore normally overestimated.

202. When using radiolabelled substances, the labelling is most often placed in the stable part of the molecule, for which reason the measured BCF value includes the BCF of the metabolites. For some substances it is the metabolite which is the most toxic and which has the highest bioconcentration potential. Measurements of the parent substance as well as the metabolites may

thus be important for the interpretation of the aquatic hazard (including the bioconcentration potential) of such substances.

203. In experiments where radiolabelled substances have been used, high radiolabel concentrations are often found in the gall bladder of fish. This is interpreted to be caused by biotransformation in the liver and subsequently by excretion of metabolites in the gall bladder (Comotto *et al.*, 1979; Wakabayashi *et al.*, 1987; Goodrich *et al.*, 1991; Toshima *et al.*, 1992). When fish do not eat, the content of the gall bladder is not emptied into the gut, and high concentrations of metabolites may build up in the gall bladder. The feeding regime may thus have a pronounced effect on the measured BCF. In the literature many studies are found where radiolabelled compounds are used, and where the fish are not fed. As a result high concentrations of radioactive material are found in the gall bladder. In these studies the bioconcentration may in most cases have been overestimated. Thus when evaluating experiments, in which radiolabelled compounds are used, it is essential to evaluate the feeding regime as well.

204. If the BCF in terms of radiolabelled residues is documented to be ≥ 1000, identification and quantification of degradation products, representing ≥ 10% of total residues in fish tissues at steady-state, are for e.g., pesticides strongly recommended in the OECD guideline No. 305 (1996). If no identification and quantification of metabolites are available, the assessment of bioconcentration should be based on the measured radiolabelled BCF value. If, for highly bioaccumulative substances (BCF ≥ 500), only BCFs based on the parent compound and on radiolabelled measurements are available, the latter should thus be used in relation to classification.

5.2.2 Octanol-water-partitioning coefficient (K_{ow})

205. For organic substances experimentally derived high-quality K_{ow} values, or values which are evaluated in reviews and assigned as the "recommended values", are preferred over other determinations of K_{ow}. When no experimental data of high quality are available, validated Quantitative Structure Activity Relationships (QSARs) for log K_{ow} may be used in the classification process. Such validated QSARs may be used without modification to the agreed criteria if they are restricted to chemicals for which their applicability is well characterised. For substances like strong acids and bases, substances which react with the eluent, or surface-active substances, a QSAR estimated value of K_{ow} or an estimate based on individual *n*-octanol and water solubilities should be provided instead of an analytical determination of K_{ow} (EEC A.8., 1992; OECD 117, 1989). Measurements should be taken on ionizable substances in their non-ionised form (free acid or free base) only by using an appropriate buffer with pH below pK for free acid or above the pK for free base.

Experimental determination of K_{ow}

206. For experimental determination of K_{ow} values, several different methods, Shake-flask, and HPLC, are described in standard guidelines, e.g., OECD Test Guideline 107 (1995); OECD Test Guideline 117 (1989); EEC A.8. (1992); EPA-OTS (1982); EPA-FIFRA (1982); ASTM (1993); the pH-metric method (OECD Test Guideline in preparation). The shake-flask method is recommended when the log K_{ow} value falls within the range from –2 to 4. The shake-flask method applies only to essential pure substances soluble in water and *n*-octanol. For highly lipophilic substances, which slowly dissolve in water, data obtained by employing a slow-stirring method are generally more reliable. Furthermore, the experimental difficulties, associated with the formation of microdroplets during the shake-flask experiment, can to some degree be overcome by a slow-stirring method where water, octanol, and test compound are equilibrated in a gently stirred reactor. With the slow-stirring method (OECD Test Guideline in preparation) a precise and accurate determination of K_{ow}

of compounds with log K_{ow} of up to 8.2 is allowed (OECD draft Guideline, 1998). As for the shake-flask method, the slow-stirring method applies only to essentially pure substances soluble in water and *n*-octanol. The HPLC method, which is performed on analytical columns, is recommended when the log K_{ow} value falls within the range 0 to 6. The HPLC method is less sensitive to the presence of impurities in the test compound compared to the shake-flask method. Another technique for measuring log K_{ow} is the generator column method (USEPA 1985).

207. As an experimental determination of the K_{ow} is not always possible, e.g., for very water-soluble substances, very lipophilic substances, and surfactants, a QSAR-derived K_{ow} may be used.

Use of QSARs for determination of log K_{ow}

208. When an estimated K_{ow} value is found, the estimation method has to be taken into account. Numerous QSARs have been and continue to be developed for the estimation of K_{ow}. Four commercially available PC programmes (CLOGP, LOGKOW (KOWWIN), AUTOLOGP, SPARC) are frequently used for risk assessment if no experimentally derived data are available. CLOGP, LOGKOW and AUTOLOGP are based upon the addition of group contributions, while SPARC is based upon a more fundamental chemical structure algorithm. Only SPARC can be employed in a general way for inorganic or organometallic compounds. Special methods are needed for estimating log K_{ow} for surface-active compounds, chelating compounds and mixtures. CLOGP is recommended in the US EPA/EC joint project on validation of QSAR estimation methods (US EPA/EC 1993). Pedersen *et al*. (1995) recommended the CLOGP and the LOGKOW programmes for classification purposes because of their reliability, commercial availability, and convenience of use. The following estimation methods are recommended for classification purposes (Table 1).

Table 1. Recommended QSARs for estimation of K_{ow}

MODEL	Log K_{ow} range	Substance utility
CLOGP	<0 - > 9[1]	The program calculates log K_{ow} for organic compounds containing C, H, N, O, Hal, P, and/or S.
LOGKOW (KOWWIN)	-4 - 8[2]	The program calculates log K_{ow} for organic compounds containing C, H, N, O, Hal, Si, P, Se, Li, Na, K, and/or Hg. Some surfactants (e.g., alcohol ethoxylates, dyestuffs, and dissociated substances may be predicted by the program as well.
AUTOLOGP	> 5	The programme calculates log K_{ow} for organic compounds containing C, H, N, O, Hal, P and S. Improvements are in progress in order to extend the applicability of AUTOLOGP.
SPARC	Provides improved results over KOWWIN and CLOGP for compounds with log K_{ow} > 5.	SPARC is a mechanistic model based on chemical thermodynamic principles rather than a deterministic model rooted in knowledge obtained from observational data. Therefore, SPARC differs from models that use QSARs (i.e., KOWWIN, CLOGP, AUTOLOGP) in that no measured log K_{ow} data are needed for a training set of chemicals. Only SPARC can be employed in a general way for inorganic or organometallic compounds.

1) A validation study performed by Niemelä, who compared experimental determined log K_{ow} values with estimated values, showed that the program precisely predicts the log K_{ow} for a great number of organic chemicals in the log K_{ow} range from below 0 to above 9 (n = 501, r2 = 0.967) (TemaNord 1995: 581).
2) Based on a scatter plot of estimated vs. experimental log K_{ow} (Syracuse Research Corporation, 1999), where 13058 compound have been tested, the LOGKOW is evaluated being valid for compounds with a log K_{ow} in the interval -4 - 8.

5.3 CHEMICAL CATEGORIES THAT NEED SPECIAL ATTENTION WITH RESPECT TO BCF AND K_{ow} VALUES

209. There are certain physico-chemical properties, which can make the determination of BCF or its measurement difficult. These may be substances, which do not bioconcentrate in a manner consistent with their other physico-chemical properties, e.g., steric hindrance or substances which make the use of descriptors inappropriate, e.g., surface activity, which makes both the measurement and use of log K_{ow} inappropriate.

5.3.1 Difficult substances

210. Some chemical substances are difficult to test in aquatic systems and guidance has been developed to assist in testing these materials (DoE, 1996; ECETOC 1996; and US EPA 1996). OECD is in the process of finalising a guidance document for the aquatic testing of difficult substances (OECD, 2000). This latter document is a good source of information, also for bioconcentration studies, on the types of substances that are difficult to test and the steps needed to ensure valid conclusions from tests with these substances. Difficult to test substances may be poorly soluble, volatile, or subject to rapid degradation due to such processes as phototransformation, hydrolysis, oxidation, or biotic degradation.

211. To bioconcentrate organic compounds, a substance needs to be soluble in lipids, present in the water, and available for transfer across the fish gills. Properties which alter this availability will thus change the actual bioconcentration of a substance, when compared with the prediction. For example, readily biodegradable substances may only be present in the aquatic compartment for short periods of time. Similarly, volatility, and hydrolysis will reduce the concentration and the time during which a substance is available for bioconcentration. A further important parameter, which may reduce the actual exposure concentration of a substance, is adsorption, either to particulate matter or to surfaces in general. There are a number of substances, which have shown to be rapidly transformed in the organism, thus leading to a lower BCF value than expected. Substances that form micelles or aggregates may bioconcentrate to a lower extent than would be predicted from simple physico-chemical properties. This is also the case for hydrophobic substances that are contained in micelles formed as a consequence of the use of dispersants. Therefore, the use of dispersants in bioaccumulation tests is discouraged.

212. In general, for difficult to test substances, measured BCF and K_{ow} values – based on the parent substance – are a prerequisite for the determination of the bioconcentration potential. Furthermore, proper documentation of the test concentration is a prerequisite for the validation of the given BCF value.

5.3.2 Poorly soluble and complex substances

213. Special attention should be paid to poorly soluble substances. Frequently the solubility of these substances is recorded as less than the detection limit, which creates problems in interpreting

the bioconcentration potential. For such substances the bioconcentration potential should be based on experimental determination of log K_{ow} or QSAR estimations of log K_{ow}.

214. When a multi-component substance is not fully soluble in water, it is important to attempt to identify the components of the mixture as far as practically possible and to examine the possibility of determining its bioaccumulation potential using available information on its components. When bioaccumulating components constitute a significant part of the complex substance (e.g., more than 20% or for hazardous components an even lower content), the complex substance should be regarded as being bioaccumulating.

5.3.3 High molecular weight substances

215. Above certain molecular dimensions, the potential of a substance to bioconcentrate decreases. This is possibly due to steric hindrance of the passage of the substance through gill membranes. It has been proposed that a cut-off limit of 700 for the molecular weight could be applied (e.g., European Commission, 1996). However, this cut-off has been subject to criticism and an alternative cut-off of 1000 has been proposed in relation to exclusion of consideration of substances with possible indirect aquatic effects (CSTEE, 1999). In general, bioconcentration of possible metabolites or environmental degradation products of large molecules should be considered. Data on bioconcentration of molecules with a high molecular weight should therefore be carefully evaluated and only be used if such data are considered to be fully valid in respect to both the parent compound and its possible metabolites and environmental degradation products.

5.3.4 Surface-active agents

216. Surfactants consist of a lipophilic (most often an alkyl chain) and a hydrophilic part (the polar headgroup). According to the charge of the headgroup, surfactants are subdivided into categories of anionic, cationic, non-ionic, or amphoteric surfactants. Due to the variety of different headgroups, surfactants are a structurally diverse category of compounds, which is defined by surface activity rather than by chemical structure. The bioaccumulation potential of surfactants should thus be considered in relation to the different subcategories (anionic, cationic, non-ionic, or amphoteric) instead of to the group as a whole. Surface-active substances may form emulsions, in which the bioavailability is difficult to ascertain. Micelle formation can result in a change of the bioavailable fraction even when the solutions are apparently formed, thus giving problems in interpretation of the bioaccumulation potential.

Experimentally derived bioconcentration factors

217. Measured BCF values on surfactants show that BCF may increase with increasing alkyl chain length and be dependant of the site of attachment of the head group, and other structural features.

Octanol-water-partition coefficient (K_{ow})

218. The octanol-water partition coefficient for surfactants can not be determined using the shake-flask or slow stirring method because of the formation of emulsions. In addition, the surfactant molecules will exist in the water phase almost exclusively as ions, whereas they will have to pair with a counter-ion in order to be dissolved in octanol. Therefore, experimental determination of K_{ow} does not characterise the partition of ionic surfactants (Tolls, 1998). On the other hand, it has been shown that the bioconcentration of anionic and non-ionic surfactants increases with increasing lipophilicity (Tolls, 1998). Tolls (1998) showed that for some surfactants, an estimated log K_{ow}

value using LOGKOW could represent the bioaccumulation potential; however, for other surfactants some 'correction' to the estimated log K_{ow} value using the method of Roberts (1989) was required. These results illustrate that the quality of the relationship between log K_{ow} estimates and bioconcentration depends on the category and specific type of surfactants involved. Therefore, the classification of the bioconcentration potential based on log K_{ow} values should be used with caution.

5.4 CONFLICTING DATA AND LACK OF DATA

5.4.1 Conflicting BCF data

219. In situations where multiple BCF data are available for the same substance, the possibility of conflicting results might arise. In general, conflicting results for a substance, which has been tested several times with an appropriate bioconcentration test, should be interpreted by a "weight of evidence approach". This implies that if experimental determined BCF data, both ≥ and < 500, have been obtained for a substance the data of the highest quality and with the best documentation should be used for determining the bioconcentration potential of the substance. If differences still remain, if e.g., high-quality BCF values for different fish species are available, generally the highest valid value should be used as the basis for classification.

220. When larger data sets (4 or more values) are available for the same species and life stage, the geometric mean of the BCF values may be used as the representative BCF value for that species.

5.4.2 Conflicting log K_{ow} data

221. The situations, where multiple log K_{ow} data are available for the same substance, the possibility of conflicting results might arise. If log K_{ow} data both ≥ and < 4 have been obtained for a substance, then the data of the highest quality and the best documentation should be used for determining the bioconcentration potential of the substance. If differences still exist, generally the highest valid value should take precedence. In such situation, QSAR estimated log K_{ow} could be used as a guidance.

5.4.3 Expert judgement

222. If no experimental BCF or log K_{ow} data or no predicted log K_{ow} data are available, the potential for bioconcentration in the aquatic environment may be assessed by expert judgement. This may be based on a comparison of the structure of the molecule with the structure of other substances for which experimental bioconcentration or log K_{ow} data or predicted K_{ow} are available.

5.5 DECISION SCHEME

223. Based on the above discussions and conclusions, a decision scheme has been elaborated which may facilitate decisions as to whether or not a substance has the potential for bioconcentration in aquatic species.

224. Experimentally derived BCF values of high quality are ultimately preferred for classification purposes. BCF values of low or uncertain quality should not be used for classification purposes if data on log K_{ow} are available because they may give a false and too low BCF value, e.g., due to a too short exposure period in which steady-state conditions have not been reached. If no

BCF is available for fish species, high quality data on the BCF for other species (e.g., mussels) may be used.

225. For organic substances, experimentally derived high quality K_{ow} values, or values which are evaluated in reviews and assigned as the "recommended values", are preferred. If no experimentally data of high quality are available validated Quantitative Structure Activity Relationships (QSARs) for log K_{ow} may be used in the classification process. Such validated QSARs may be used without modification in relation to the classification criteria, if restricted to chemicals for which their applicability is well characterised. For substances like strong acids and bases, metal complexes, and surface-active substances a QSAR estimated value of K_{ow} or an estimate based on individual *n*-octanol and water solubilities should be provided instead of an analytical determination of K_{ow}.

226. If data are available but not validated, expert judgement should be used.

227. Whether or not a substance has a potential for bioconcentration in aquatic organisms could thus be decided in accordance with the following scheme:

Valid/high quality experimentally determined BCF value → YES:
→BCF ≥ 500: *The substance has a potential for bioconcentration*
→BCF < 500: *The substance does not have a potential for bioconcentration*

Valid/high quality experimentally determined BCF value → NO:
→ Valid/high quality experimentally determined log K_{ow} value → YES:
→ log K_{ow} ≥ 4: *The substance has a potential for bioconcentration*
→ log K_{ow} < 4: *The substance does not have a potential for bioconcentration*

Valid/high quality experimentally determined BCF value → NO:
→ Valid/high quality experimentally determined log K_{ow} value → NO:
→ Use of validated QSAR for estimating a log K_{ow} value → YES:
→ log K_{ow} ≥ 4: *The substance has a potential for bioconcentration*
→ log K_{ow} < 4: *The substance does not have a potential for bioconcentration*

ANNEX 5.I

BASIC PRINCIPLES OF THE EXPERIMENTAL AND ESTIMATION METHODS FOR DETERMINATION OF BCF AND K_{ow} OF ORGANIC SUBSTANCES

1. BIOCONCENTRATION FACTOR (BCF)

228. The bioconcentration factor is defined as the ratio between the concentration of the chemical in biota and the concentration in the surrounding medium, here water, at steady state. BCF can be measured experimentally directly under steady-state conditions or calculated by the ratio of the first-order uptake and elimination rate constants, a method that does not require equilibrium conditions.

1.1 Appropriate methods for experimental determination of BCF

229. Different test guidelines for the experimental determination of bioconcentration in fish have been documented and adopted; the most generally applied being the OECD test guideline (OECD 305, 1996) and the ASTM standard guide (ASTM E 1022-94). OECD 305 (1996) was revised and replaced the previous version OECD 305A-E, (1981). Although flow-through test regimes are preferred (OECD 305, 1996), semi-static regimes are allowed (ASTM E 1022-94), provided that the validity criteria on mortality and maintenance of test conditions are fulfilled. For lipophilic substances ($\log K_{ow} > 3$), flow-through methods are preferred.

230. The principles of the OECD 305 and the ASTM guidelines are similar, but the experimental conditions described are different, especially concerning:

- method of test water supply (static, semi-static or flow through)
- the requirement for carrying out a depuration study
- the mathematical method for calculating BCF
- sampling frequency: Number of measurements in water and number of samples of fish
- requirement for measuring the lipid content of the fish
- the minimum duration of the uptake phase

231. In general, 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. A 28-day exposure phase is obligatory unless a steady state has been reached within this period. The time needed for reaching steady-state conditions may be set on the basis of $K_{ow} - k_2$ correlations (e.g., $\log k_2 = 1.47 - 0.41 \log K_{ow}$ (Spacie and Hamelink, 1982) or $\log k_2 = 1.69 - 0.53 \log K_{ow}$ (Gobas *et al.*, 1989)). The expected time (d) for e.g., 95% steady state may thus be calculated by: $-\ln(1-0.95)/k_2$, provided that the bioconcentration follows first order kinetics. During the depuration phase the fish are transferred to a medium free of the test substance. The concentration of the test substance in the fish is followed through both phases of the test. The BCF is expressed as a function of the total wet weight of the fish. As for many organic substances, there is a significant relationship between the potential for bioconcentration and the lipophilicity, and furthermore, there is a corresponding relationship between the lipid content of the test fish and the observed bioconcentration of such substances. Therefore, to reduce this source of variability in the test results for the substances with high lipophilicity, bioconcentration should be expressed in relation to the lipid content in addition to whole body weight (OECD 305 (1996), ECETOC (1995)). The guidelines mentioned are based on the assumption that bioconcentration may be approximated by a first-order process (one-compartment model) and thus that $BCF = k_1/k_2$ (k_1: first-order uptake rate, k_2: first-order depuration

rate, described by a log-linear approximation). If the depuration follows biphasic kinetics, i.e., two distinct depuration rates can be identified, the approximation k_1/k_2 may significantly underestimate BCF. If a second order kinetic has been indicated, BCF may be estimated from the relation: C_{Fish}/C_{Water}, provided that "steady-state" for the fish-water system has been reached.

232. Together with details of sample preparation and storage, an appropriate analytical method of known accuracy, precision, and sensitivity must be available for the quantification of the substance in the test solution and in the biological material. If these are lacking it is impossible to determine a true BCF. The use of radiolabelled test substance can facilitate the analysis of water and fish samples. However, unless combined with a specific analytical method, the total radioactivity measurements potentially reflect the presence of parent substance, possible metabolite(s), and possible metabolised carbon, which have been incorporated in the fish tissue in organic molecules. For the determination of a true BCF it is essential to clearly discriminate the parent substance from possible metabolites. 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 purified so that the parent compound can be analysed separately.

233. In the log K_{ow} range above 6, the measured BCF data tend to decrease with increasing log K_{ow}. Conceptual explanations of non-linearity mainly refer to either biotransformation, reduced membrane permeation kinetics or reduced biotic lipid solubility for large molecules. Other factors consider experimental artefacts, such as equilibrium not being reached, reduced bioavailability due to sorption to organic matter in the aqueous phase, and analytical errors. Moreover, care should be taken when evaluating experimental data on BCF for substances with log K_{ow} above 6, as these data will have a much higher level of uncertainty than BCF values determined for substances with log K_{ow} below 6.

2. LOG K_{ow}

234. The log *n*-octanol-water partition coefficient (log K_{ow}) is a measure of the lipophilicity of a substance. As such, log K_{ow} is a key parameter in the assessment of environmental fate. Many distribution processes are driven by log K_{ow}, e.g., sorption to soil and sediment and bioconcentration in organisms.

235. The basis for the relationship between bioconcentration and log K_{ow} is the analogy for the partition process between the lipid phase of fish and water and the partition process between n-octanol and water. The reason for using K_{ow} arises from the ability of octanol to act as a satisfactory surrogate for lipids in fish tissue. Highly significant relationships between log K_{ow} and the solubility of substances in cod liver oil and triolin exist (Niimi, 1991). Triolin is one of the most abundant triacylglycerols found in freshwater fish lipids (Henderson and Tocher, 1987).

236. The determination of the *n*-octanol-water partition coefficient (K_{ow}) is a requirement of the base data set to be submitted for notified new and priority existing substances within the EU. As the experimental determination of the K_{ow} is not always possible, e.g., for very water-soluble and for very lipophilic substances, a QSAR derived K_{ow} may be used. However, extreme caution should be exercised when using QSARs for substances where the experimental determination is not possible (as for e.g., surfactants).

2.1 Appropriate methods for experimental determination of K_{ow} values

237. For experimental determination of K_{ow} values, two different methods, Shake-flask and HPLC, have been described in standard guidelines e.g., OECD 107 (1995); OECD 117 (1983); EEC

A.8. (1992); EPA-OTS (1982); EPA-FIFRA (1982); ASTM (1993). Not only data obtained by the employment of the shake-flask or the HPLC method according to standard guidelines are recommended. For highly lipophilic substances, which are slowly soluble in water, data obtained by employing a slow-stirring method are generally more reliable (De Bruijn *et al*., 1989; Tolls and Sijm, 1993; OECD draft Guideline, 1998). The slow stirring method is currently being ringtested for development of a final OECD guideline.

Shake-flask method

238. The basic principle of the method is to measure the dissolution of the substance in two different phases, water and *n*-octanol. In order to determine the partition coefficient, equilibrium between all interacting components of the system must be achieved after which the concentration of the substances dissolved in the two phases is determined. The shake-flask method is applicable when the log K_{ow} value falls within the range from -2 to 4 (OECD 107, 1995). The shake-flask method applies only to essential pure substances soluble in water and *n*-octanol and should be performed at a constant temperature (±1°C) in the range 20-25°C.

HPLC method

239. 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 at different rates because of the different degrees of partitioning between the mobile aqueous phase and the stationary hydrocarbon phase. The HPLC method is not applicable to strong acids and bases, metals complexes, surface-active materials, or substances that react with the eluent. The HPLC method is applicable when the log K_{ow} value falls within the range 0 to 6 (OECD 117, 1989). The HPLC method is less sensitive to the presence of impurities in the test compound compared to the shake-flask method.

Slow stirring method

240. With the slow-stirring method a precise and accurate determination of K_{ow} of compounds with log K_{ow} up till 8.2 is allowed (De Bruijn *et al*., 1989). For highly lipophilic compounds the shake-flask method is prone to produce artefacts (formation of microdroplets), and with the HPLC method K_{ow} needs to be extrapolated beyond the calibration range to obtain estimates of K_{ow}.

241. In order to determine a partition coefficient, water, n-octanol, and test compound are equilibrated with each other after which the concentration of the test compound in the two phases is determined. The experimental difficulties associated with the formation of microdroplets during the shake-flask experiment can to some degree be overcome in the slow-stirring experiment as water, octanol, and the test compound are equilibrated in a gently stirred reactor. The stirring creates a more or less laminar flow between the octanol and the water, and exchange between the phases is enhanced without microdroplets being formed.

Generator Column Method

242. Another very versatile method for measuring log K_{ow} is the generator column method. In this method, a generator column method is used to partition the test substance between the octanol and water phases. The column is packed with a solid support and is saturated with a fixed concentration of the test substance in *n*-octanol. The test substance is eluted from the octanol - saturated generator column with water. The aqueous solution exiting the column represents the equilibrium concentration of the test substance that has partitioned from the octanol phase into the

water phase. The primary advantage of the generator column method over the shake flask method is that the former completely avoids the formation of micro-emulsions. Therefore, this method is particularly useful for measuring K_{ow} for substances values over 4.5 (Doucette and Andren, 1987 and 1988; Shiu *et al.*, 1988) as well as for substances having log K_{ow} values less than 4.5. A disadvantage of the generator column method is that it requires sophisticated equipment. A detailed description of the generator column method is presented in the "Toxic Substances Control Act Test Guidelines" (USEPA 1985).

2.2 Use of QSARs for determination of log K_{ow} (see also Chapter 6: Use of QSARs)

243. Numerous QSARs have been and continue to be developed for the estimation of K_{ow}. Commonly used methods are based on fragment constants. The fragmental approaches are based on a simple addition of the lipophilicity of the individual molecular fragments of a given molecule. Three commercially available PC programs are recommended in the European Commission's Technical Guidance Document (European Commission, 1996) for risk assessment, part III, if no experimentally derived data are available.

244. CLOGP (Daylight Chemical Information Systems, 1995) was initially developed for use in drug design. The model is based on the Hansch and Leo calculation procedure (Hansch and Leo, 1979). The program calculates log K_{ow} for organic compounds containing C, H, N, O, Hal, P, and/or S. Log K_{ow} for salts and for compounds with formal charges cannot be calculated (except for nitro compounds and nitrogen oxides). The calculation results of log K_{ow} for ionizable substances, like phenols, amines, and carboxylic acids, represent the neutral or unionised form and will be pH dependent. In general, the program results in clear estimates in the range of log Kow between 0 and 5 (European Commission, 1996, part III). However a validation study performed by Niemelä (1993), who compared experimental determined log K_{ow} values with estimated values, showed that the program precisely predicts the log K_{ow} for a great number of organic chemicals in the log K_{ow} range from below 0 to above 9 (n=501, r2=0.967). In a similar validation study on more than 7000 substances the results with the CLOGP-program (PC version 3.32, EPA version 1.2) were r2= 0.89, s.d.= 0.58, n= 7221. These validations show that the CLOGP-program may be used for estimating reliable log K_{ow} values when no experimental data are available. For chelating compounds and surfactants the CLOGP program is stated to be of limited reliability (OECD, 1993). However, as regards anionic surfactants (LAS) a correction method for estimating adjusted CLOGP values has been proposed (Roberts, 1989).

245. LOGKOW or KOWWIN (Syracuse Research Corporation) uses structural fragments and correction factors. The program calculates log K_{ow} for organic compounds containing the following atoms: C, H, N, O, Hal, Si, P, Se, Li, Na, K, and/or Hg. Log K_{ow} for compounds with formal charges (like nitrogenoxides and nitro compounds) can also be calculated. The calculation of log K_{ow} for ionizable substances, like phenols, amines and carboxylic acids, represent the neutral or unionised form, and the values will thus be pH dependent. Some surfactants (e.g., alcohol ethoxylates (Tolls, 1998), dyestuffs, and dissociated substances may be predicted by the LOGKOW program (Pedersen *et al*, 1995). In general, the program gives clear estimates in the range of log K_{ow} between 0 and 9 (TemaNord 1995:581). Like the CLOGP-program, LOGKOW has been validated (Table 2) and is recommended for classification purposes because of its reliability, commercial availability, and convenience of use.

246. AUTOLOGP (Devillers *et al.*, 1995) has been derived from a heterogeneous data set, comprising 800 organic chemicals collected from literature. The program calculates log K_{ow} values for organic chemicals containing C, H, N, O, Hal, P, and S. The log K_{ow} values of salts cannot be

calculated. Also the log K_{ow} of some compounds with formal charges cannot be calculated, with the exception of nitro compounds. The log K_{ow} values of ionizable chemicals like phenols, amines, and corboxylic acids can be calculated although pH-dependencies should be noted. Improvements are in progress in order to extend the applicability of AUTOLOGP. According to the presently available information, AUTOLOGP gives accurate values especially for highly lipophilic substances (log K_{ow} > 5) (European Commission, 1996).

247. SPARC. The SPARC model is still under development by EPA's Environmental Research Laboratory in Athens, Georgia, and is not yet public available. SPARC is a mechanistic model based on chemical thermodynamic principles rather than a deterministic model rooted in knowledge obtained from observational data. Therefore, SPARC differs from models that use QSARs (i.e., KOWWIN, LOGP) in that no measured log K_{ow} data are needed for a training set of chemicals. EPA does occasionally run the model for a list of CAS numbers, if requested. SPARC provides improved results over KOWWIN and CLOGP only for compounds with log K_{ow} values greater than 5. Only SPARC can be employed in a general way for inorganic or organometallic compounds.

248. In Table 2 an overview of log K_{ow} estimation methods based on fragmentation methodologies is presented. Also other methods for the estimation of log K_{ow} values exist, but they should only be used on a case by case basis and only with appropriate scientific justification.

Table 2 Overview of QSAR methods for estimation of log K_{ow} based on fragmentation methodologies (Howard and Meylan (1997)).

Method	Methodology	Statistics
CLOGP Hansch and Leo (1979), CLOGP Daylight (1995)	Fragments + correction factors	Total n=8942, r2=0,917 sd = 0,482 Validation: n=501 r2=0,967 Validation: n=7221 r2=0,89 sd = 0,58
LOGKOW (KOWWIN) Meylan and Howard (1995), SRC	140 fragments 260 correction factors	Calibration: n=2430, r2=0,981 sd = 0,219 me=0,161 Validation: n=8855 r2=0,95 sd = 0,427 me = 0,327
AUTOLOGP Devillers *et al*. (1995)	66 atomic and group contributions from Rekker and Manhold (1992)	Calibration: n=800, r2=0,96 sd = 0,387
SPARC Under development by EPA, Athens, Georgia.	Based upon fundamental chemical structure algorithm.	No measured log Kow data are needed for a training set of chemicals.
Rekker and De Kort (1979)	Fragments + correction factors	Calibration n=1054, r2=0,99 Validation: n=20 r2=0,917 sd = 0,53 me = 0,40
Niemi *et al.* (1992)	MCI	Calibration n=2039, r2=0,77 Validation: n=2039 r2=0,49
Klopman *et al* (1994)	98 fragments + correction factors	Calibration n=1663, r2=0,928 sd = 0,3817
Suzuki and Kudo (1990)	424 fragments	Total: n=1686 me = 0,35 Validation: n=221 me = 0,49
Ghose *et al.* (1988) ATOMLOGP	110 fragments	Calibration: n=830, r2=0,93 sd = 0,47 Validation: n=125 r2=0,87 sd = 0,52
Bodor and Huang (1992)	Molecule orbital	Calibration: n=302, r2=0,96 sd = 0,31 me=0,24 Validation: n=128 sd = 0,38
Broto *et al*. (1984) ProLogP	110 fragments	Calibration: n=1868, me=ca. 0,4

ANNEX 5.II

INFLUENCE OF EXTERNAL AND INTERNAL FACTORS ON THE BIOCONCENTRATION POTENTIAL OF ORGANIC SUBSTANCES

1. FACTORS INFLUENCING THE UPTAKE

249. The uptake rate for lipophilic compounds is mainly a function of the size of the organism (Sijm and Linde, 1995). External factors such as the molecular size, factors influencing the bioavailability, and different environmental factors are of great importance to the uptake rate as well.

1.1 Size of organism

250. Since larger fish have a relatively lower gill surface to weight ratio, a lower uptake rate constant (k_1) is to be expected for large fish compared to small fish (Sijm and Linde, 1995; Opperhuizen and Sijm, 1990). The uptake of substances in fish is further controlled by the water flow through the gills; the diffusion through aqueous diffusion layers at the gill epithelium; the permeation through the gill epithelium; the rate of blood flow through the gills, and the binding capacity of blood constituents (ECETOC, 1995).

1.2 Molecular size

251. Ionised substances do not readily penetrate membranes; as aqueous pH can influence the substance uptake. Loss of membrane permeability is expected for substances with a considerable cross-sectional area (Opperhuizen *et al*., 1985; Anliker *et al*., 1988) or long chain length (> 4.3 nm) (Opperhuizen, 1986). Loss of membrane permeability due to the size of the molecules will thus result in total loss of uptake. The effect of molecular weight on bioconcentration is due to an influence on the diffusion coefficient of the substance, which reduces the uptake rate constants (Gobas *et al*., 1986).

1.3 Availability

252. Before a substance is able to bioconcentrate in an organism it needs to be present in water and available for transfer across fish gills. Factors, which affect this availability under both natural and test conditions, will alter the actual bioconcentration in comparison to the estimated value for BCF. As fish are fed during bioconcentration studies, relatively high concentrations of dissolved and particulate organic matter may be expected, thus reducing the fraction of chemical that is actually available for direct uptake via the gills. McCarthy and Jimenez (1985) have shown that adsorption of lipophilic substances to dissolved humic materials reduces the availability of the substance, the more lipophilic the substance the larger reduction in availability (Schrap and Opperhuizen, 1990). Furthermore, adsorption to dissolved or particulate organic matter or surfaces in general may interfere during the measurement of BCF (and other physical-chemical properties) and thus make the determination of BCF or appropriate descriptors difficult. As bioconcentration in fish is directly correlated with the available fraction of the chemical in water, it is necessary for highly lipophilic substances to keep the available concentration of the test chemical within relatively narrow limits during the uptake period.

253. Substances, which are readily biodegradable, may only be present in the test water for a short period, and bioconcentration of these substances may thus be insignificant. Similarly,

volatility and hydrolysis will reduce the concentration and time in which the substance is available for bioconcentration.

1.4 Environmental factors

254. Environmental parameters influencing the physiology of the organism may also affect the uptake of substances. For instance, when the oxygen content of the water is lowered, fish have to pass more water over their gills in order to meet respiratory demands (McKim and Goeden, 1982). However, there may be species dependency as indicated by Opperhuizen and Schrap (1987). It has, furthermore, been shown that the temperature may have an influence on the uptake rate constant for lipophilic substances (Sijm *et al*. 1993), whereas other authors have not found any consistent effect of temperature changes (Black *et al*. 1991).

2 FACTORS INFLUENCING THE ELIMINATION RATE

255. The elimination rate is mainly a function of the size of the organism, the lipid content, the biotransformation process of the organism, and the lipophilicity of the test compound.

2.1 Size of organism

256. As for the uptake rate the elimination rate is dependent on the size of the organism. Due to the higher gill surface to weight ratio for small organisms (e.g., fish larvae) than that of large organisms, steady-state and thus "toxic dose equilibrium" has shown to be reached sooner in early life stages than in juvenile/adult stages of fish (Petersen and Kristensen, 1998). As the time needed to reach steady-state conditions is dependent on k_2, the size of fish used in bioconcentration studies has thus an important bearing on the time required for obtaining steady-state conditions.

2.2 Lipid content

257. Due to partitioning relationships, organisms with a high fat content tend to accumulate higher concentrations of lipophilic substances than lean organisms under steady-state conditions. Body burdens are therefore often higher for "fatty" fish such as eel, compared to "lean" fish such as cod. In addition, lipid "pools" may act as storage of highly lipophilic substances. Starvation or other physiological changes may change the lipid balance and release such substances and result in delayed impacts.

2.3 Metabolism

258. In general, metabolism or biotransformation leads to the conversion of the parent compound into more water-soluble metabolites. As a result, the more hydrophilic metabolites may be more easily excreted from the body than the parent compound. When the chemical structure of a compound is altered, many properties of the compound are altered as well. Consequently the metabolites will behave differently within the organism with respect to tissue distribution, bioaccumulation, persistence, and route and rate of excretion. Biotransformation may also alter the toxicity of a compound. This change in toxicity may either be beneficial or harmful to the organism. Biotransformation may prevent the concentration in the organism from becoming so high that a toxic response is expressed (detoxification). However, a metabolite may be formed which is more toxic than the parent compound (bioactivation) as known for e.g., benzo(a)pyrene.

259. Terrestrial organisms have a developed biotransformation system, which is generally better than that of organisms living in the aquatic environment. The reason for this difference may

be the fact that biotransformation of xenobiotics may be of minor importance in gill breathing organisms as they can relatively easily excrete the compound into the water (Van Den Berg *et al*. 1995). Concerning the biotransformation capacity in aquatic organisms the capacity for biotransformation of xenobiotics increases in general as follows: Molluscs < crustaceans < fish (Wofford *et al*., 1981).

3. LIPOPHILICITY OF SUBSTANCE

260. A negative linear correlation between k_2 (depuration constant) and log K_{ow} (or BCF) has been shown in fish by several authors (e.g., Spacie and Hamelink, 1982; Gobas *et al*., 1989; Petersen and Kristensen, 1998), whereas k_1 (uptake rate constant) is more or less independent of the lipophilicity of the substance (Connell, 1990). The resultant BCF will thus generally increase with increasing lipophilicity of the substances, i.e., log BCF and log K_{ow} correlate for substances which do not undergo extensive metabolism.

ANNEX 5.III

TEST GUIDELINES

261. Most of the guidelines mentioned are found in compilations from the organisation issuing them. The main references to these are:

- EC guidelines: European Commission (1996). Classification, Packaging and Labelling of Dangerous Substances in the European Union. Part 2 – Testing Methods. European Commission. 1997. ISBN92-828-0076-8. (Homepage: http://ecb.ei.jrc.it/testing-methods/);
- ISO guidelines: Available from the national standardisation organisations or ISO (Homepage: http://www.iso.ch/);
- OECD guidelines for the testing of chemicals. OECD. Paris. 1993 with regular updates (Homepage: http://www.oecd.org/ehs/test/testlist.htm);
- OPPTS guidelines: US-EPA's homepage: http://www.epa.gov/opptsfrs/home/guidelin.htm;
- ASTM : ASTM´s homepage: http://www.astm.org. Further search via "standards".

ASTM, 1993. ASTM Standards on Aquatic Toxicology and Hazard Evaluation. Sponsored by ASTM Committee E-47 on Biological Effects and Environmental Fate. American Society for Testing and Materials. 1916 Race Street, Philadelphia, PA 19103. ASTM PCN: 03-547093-16., ISBN 0-8032-1778-7.

ASTM E 1022-94. 1997. Standard Guide for Conducting Bioconcentration Tests with Fishes and Saltwater Bivalve Molluscs. American Society for Testing and Materials.

EC, 1992. EC A.8. Partition coefficient. Annex V (Directive 67/548/EEC). Methods for determination of physico-chemical properties, toxicity and ecotoxicity.

EC, 1998. EC.C.13 Bioconcentration: Flow-through Fish Test.

EPA-OTS, 1982. Guidelines and support documents for environmental effects testing. Chemical fate test guidelines and support documents. United States Environmental Protection Agency. Office of Pesticides and Toxic Substances, Washington, D.C. 20960. EPA 560/6-82-002. (August 1982 and updates), cf. also Code of Federal Regulations. Protection of the Environment Part 790 to End. Revised as of July 1, 1993. ONLINE information regarding the latest updates of these test guidelines: US National Technical Information System.

EPA-FIFRA, 1982. The Federal Insecticide, Fungicide and Rodenticide Act. Pesticide Assessment Guidelines, subdivision N: chemistry: Environmental fate, and subdivision E, J & L: Hazard Evaluation. Office of Pesticide Programs. US Environmental Protection Agency, Washington D.C. (1982 and updates). ONLINE information regarding the latest updates of these test guidelines: US National Technical Information System.

OECD Test Guideline 107, 1995. OECD Guidelines for testing of chemicals. Partition Coefficient (n-octanol/water): Shake Flask Method.

OECD Test Guideline 117, 1989. OECD Guideline for testing of chemicals. Partition Coefficient (n-octanol/water), High Performance Liquid Chromatography (HPLC) Method.

OECD Test Guideline 305, 1996. Bioconcentration: Flow-through Fish Test. OECD Guidelines for testing of Chemicals.

OECD Test Guidelines 305 A-E, 1981. Bioaccumulation. OECD Guidelines for testing of chemicals.

OECD draft Test Guideline, 1998. Partition Coefficient n-Octanol/Water P_{ow}. Slow-stirring method for highly hydrophobic chemicals. Draft proposal for an OECD Guideline for Testing of Chemicals.

ANNEX 5.IV

REFERENCES

Anliker, R., Moser, P., Poppinger, D. 1988. Bioaccumulation of dyestuffs and organic pigments in fish. Relationships to hydrophobicity and steric factors. Chem. 17(8):1631-1644.

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. Vol.1.pp.29-39.

Black, M.C., Millsap, D.S., McCarthy, J.F. 1991. Effects of acute temperature change on respiration and toxicant uptake by rainbow trout, *Salmo gairdneri* (Richardson). Physiol. Zool. 64:145-168.

Bodor, N., Huang, M.J. 1992. J. Pharm. Sci. 81:272-281.

Broto, P., Moreau, G., Vandycke, C. 1984. Eur. J. Med. Chem. 19:71-78.

Chiou, T. 1985. Partition coefficients of organic compounds in lipid-water systems and correlations with fish bioconcentration factors. Environ. Sci. Technol 19:57-62.

CLOGP. 1995. Daylight Chemical Information Systems, Inf. Sys. Inc. Irvine, Ca.

CSTEE (1999): DG XXIV Scientific Committee for Toxicity and Ecotoxicity and the Environment Opinion on revised proposal for a list of Priority substances in the context of the water framework directive (COMMs Procedure) prepared by the Frauenhofer-Institute, Germany,. Final report opinion adopted at the 11th CSTEE plenary meeting on 28th of September 1999.

Comotto, R.M., Kimerle, R.A., Swisher, R.D. 1979. Bioconcentration and metabolism of linear alkylbenzenesulfonate by Daphnids and Fathead minnows. L.L.Marking, R.A. Kimerle, Eds., Aquatic Toxicology (ASTM, 1979), vol. ASTM STP 667.

Connell, D.W., Hawker, D.W. 1988. Use of polynomial expressions to describe the bioconcentration of hydrophobic chemicals by fish. Ecotoxicol. Environ. Saf. 16:242-257.

Connell, D.W. 1990. Bioaccumulation of xenobiotic compounds, Florida: CRC Press, Inc. pp.1-213.

De Bruijn, J., Busser, F., Seinen, W. & Hermens, J. 1989. Determination of octanol/water partition coefficients with the “slow stirring” method. Environ. Toxicol. Chem. 8:499-512.

Devillers, J., Bintein, S., Domine, D. 1996. Comparison of BCF models based on log P. Chemosphere 33(6):1047-1065.

DoE, 1996. Guidance on the aquatic toxicity testing of difficult substance. Unites Kingdom Department of the Environment, London.

Doucette, W.J., Andren, A.W. 1987. Correlation of octanol/water partition coefficients and total molecular surface area for highly hydrophobic aromatic compounds. Environ. Sci. Technol., 21, pages 821-824.

Doucette, W.J., Andren, A.W. 1988. Estimation of octanol/water partition coefficients: evaluation of six methods for highly hydrophobic aromatic compounds. Chemosphere, 17, pages 345-359.

Driscoll, S.K., McElroy, A.E. 1996. Bioaccumulation and metabolism of benzo(a)pyrene in three species of polychaete worms. Environ. Toxicol. Chem. 15(8):1401-1410.

ECETOC, 1995. The role of bioaccumulation in environmental risk assessment: The aquatic environment and related food webs, Brussels, Belgium.

ECEOOC, 1996. Aquatic toxicity testing of sparingly soluble, volatile and unstable substances. ECETOC Monograph No. 26, ECETOC, Brussels.

European Commission, 1996. Technical Guidance Document in support of Commission Directive 93/96/EEC on Risk Assessment for new notified substances and Commission Regulation (EC) No 1488/94 on Risk Assessment for Existing Substances. Brussels

Ghose, A.K., Prottchet, A., Crippen, G.M. 1988. J. Computational Chem. 9:80-90.

Gobas, F.A.P.C., Opperhuizen, A., Hutzinger, O. 1986. Bioconcentration of hydrophobic chemicals in fish: Relationship with membrane permeation. Environ. Toxicol. Chem. 5:637-646.

Gobas, F.A.P.C., Clark, K.E., Shiu, W.Y., Mackay, D. 1989. Bioconcentration of polybrominated benzenes and biphenyls and related superhydrophobic chemicals in fish: Role of bioavailability and elimination into feces. Environ. Toxicol. Chem. 8:231-245.

Goodrich, M.S., Melancon, M.J., Davis, R.A., Lech J.J. 1991. The toxicity, bioaccumulation, metabolism, and elimination of dioctyl sodium sulfosuccinate DSS in rainbow trout (*Oncorhynchus mykiss*) Water Res. 25: 119-124.

Hansch, C., Leo, A. 1979. Substituent constants for correlation analysis in chemistry and biology. Wiley, New York, NY, 1979.

Henderson, R.J., Tocher, D.R. 1987. The lipid composition and biochemistry of freshwater fish. Prog. Lipid. Res. 26:281-347.

Howard, P.H. and Meyland, W.M., 1997. Prediction of physical properties transport and degradation for environmental fate and exposure assessments, QSAR in environmental science VII. Eds. Chen, F. and Schüürmann, G. pp. 185-205.

Kimerle, R.A., Swisher, R.D., Schroeder-Comotto, R.M. 1975. Surfactant structure and aquatic toxicity, Symposium on Structure-Activity correlations in Studies on Toxicity and Bioconcentration with Aquatic Organisms, Burlington, Ontario, Canada, pp. 22-35.

Klopman, G., Li, J.Y., Wang, S., Dimayuga, M. 1994. Computer automated log P calculations based on an extended group contribution approach. J. Chem. Inf. Comput. Sci. 34:752-781.

Knezovich, J.P., Lawton, M.P., Inoue, L.S. 1989. Bioaccumulation and tissue distribution of a quaternary ammonium surfactant in three aquatic species. Bull. Environ. Contam. Toxicol. 42:87-93.

Knezovich, J.P., Inoue, L.S. 1993. The influence of sediment and colloidal material on the bioavailability of a quaternary ammonium surfactant. Ecotoxicol. Environ. Safety. 26:253-264.

Kristensen, P. 1991. Bioconcentration in fish: Comparison of BCFs derived from OECD and ASTM testing methods; influence of particulate matter to the bioavailability of chemicals. Danish Water Quality Institute.

Mackay, D. 1982. Correlation of bioconcentration factors. Environ. Sci. Technol. 16:274-278.

McCarthy, J.F., Jimenez, B.D. 1985. Reduction in bioavailability to bluegills of polycyclic aromatic hydrocarbons bound to dissolved humic material. Environ. Toxicol. Chem. 4:511-521.

McKim, J.M., Goeden, H.M. 1982. A direct measure of the uptake efficiency of a xenobiotic chemical across the gill of brook trout (*Salvelinus fontinalis*) under normoxic and hypoxic conditions. Comp. Biochem. Physiol. 72C:65-74.

Meylan, W.M. and Howard, P.H., 1995. Atom/Fragment Contribution Methods for Estimating Octanol-Water Partition Coefficients. J.Pharm.Sci. 84, 83.

Niemelä, J.R. 1993. QTOXIN-program (ver 2.0). Danish Environmental Protection Agency.

Niemi, G.J., Basak, S.C., Veith, G.D., Grunwald, G. Environ. Toxicol. Chem. 11:893-900.

Niimi, A.J. 1991. Solubility of organic chemicals in octanol, triolin and cod liver oil and relationships between solubility and partition coefficients. Wat. Res. 25:1515-1521.

OECD, 1993. Application of structure activity relationships to the estimation of properties important in exposure assessment. OECD Environment Directorate. Environment Monograph No. 67.

OECD, 1998. Harmonized 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.

OECD, 2000. Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures, OECD, Paris.

Opperhuizen, A., Van der Velde, E.W., Gobas, F.A.P.C., Liem, A.K.D., Van der Steen, J.M.D., Hutzinger, O. 1985. Relationship between bioconcentration in fish and steric factors of hydrophobic chemicals. Chemosphere 14:1871-1896.

Opperhuizen, A. 1986. Bioconcentration of hydrophobic chemicals in fish. In: Poston T.M., Purdy, R. (eds), Aquatic Toxicology and Environmental Fate : Ninth Volume, ASTM STP 921. American Society for Testing and Materials, Philadelphia, PA, 304-315.

Opperhuizen, A., Schrap, S.M. 1987. Relationship between aqueous oxygen concentration and uptake and elimination rates during bioconcentration of hydrophobic chemicals in fish. Environ. Toxicol. Chemosphere 6:335-342.

Opperhuizen, A., Sijm, D.T.H.M. 1990. Bioaccumulation and biotransformation of polychlorinated dibenzo-p-dioxins and dibenzofurans in fish. Environ. Toxicol. Chem. 9:175-186.

Pedersen, F., Tyle, H., Niemelä, J.R., Guttmann, B., Lander,L. and Wedebrand, A., 1995. Environmental Hazard Classification – data collection and interpretation guide (2nd edition). TemaNord 1995:581.

Petersen, G.I., Kristensen, P. 1998. Bioaccumulation of lipophilic substances in fish early life stages. Environ. Toxicol. Chem. 17(7):1385-1395.

Rekker, R.F., de Kort, H.M. 1979. The hydrophobic fragmental constant: An extension to a 1000 data point set. Eur. J. Med. Chem. – Chim. Ther. 14:479-488.

Roberts, D.W. 1989. Aquatic toxicity of linear alkyl benzene sulphonates (LAS) – a QSAR analysis. Communicaciones Presentadas a las Jornadas del Comite Espanol de la Detergencia, 20 (1989) 35-43. Also in J.E. Turner, M.W. England, T.W. Schultz and N.J. Kwaak (eds.) QSAR 88. Proc. Third International Workshop on Qualitative Structure-Activity Relationships in Environmental Toxicology, 22-26 May 1988, Knoxville, Tennessee, pp. 91-98. Available from the National Technical Information Service, US Dept. of Commerce, Springfield, VA.

Schrap, S.M., Opperhuizen, A. 1990. Relationship between bioavailability and hydrophobicity: reduction of the uptake of organic chemicals by fish due to the sorption of particles. Environ. Toxicol. Chem. 9:715-724.

Shiu, WY, Doucette, W., Gobas, FAPC., Andren, A., Mackay, D. 1988. Physical-chemical properties of chlorinated dibenzo-p-dioxins. Environ. Sci. Technol. 22: pages 651-658.

Sijm, D.T.H.M., van der Linde, A. 1995. Size-dependent bioconcentration kinetics of hydrophobic organic chemicals in fish based on diffusive mass transfer and allometric relationships. Environ. Sci. Technol. 29:2769-2777.

Sijm, D.T.H.M., Pärt, P., Opperhuizen, A. 1993. The influence of temperature on the uptake rate constants of hydrophobic compounds determined by the isolated perfused gill of rainbow trout (*Oncorhynchus mykiss*). Aquat. Toxicol. 25:1-14.

Spacie, A., Hamelink, J.L. 1982. Alternative models for describing the bioconcentration of organics in fish. Environ. Toxicol. Chem. 1:309-320.

Suzuki, T., Kudo, Y.J. 1990. J. Computer-Aided Molecular Design 4:155-198.

Syracuse Research Corporation, 1999. http://esc_plaza.syrres.com/interkow/logkow.htm

Tas, J.W., Seinen, W., Opperhuizen, A. 1991. Lethal body burden of triphenyltin chloride in fish: Preliminary results. Comp. Biochem. Physiol. 100C(1/2):59-60.

Tolls J. & Sijm, D.T.H.M., 1993. Bioconcentration of surfactants, RITOX, the Netherlands (9. Nov. 1993). Procter and Gamble Report (ed.: M.Stalmans).

Tolls, J. 1998. Bioconcentration of surfactants. Ph.D. Thesis. Utrecht University, Utrecht, The Netherlands.

Toshima, S., Moriya, T. Yoshimura, K. 1992. Effects of polyoxyethylene (20) sorbitan monooleate on the acute toxicity of linear alkylbenzenesulfonate (C_{12}-LAS) to fish. Ecotoxicol. Environ. Safety 24: 26-36.

USEPA 1985. U.S. Environmental Protection Agency. Office of Toxic Substances. Toxic Substances Control Act Test Guidelines. 50 FR 39252.

US EPA/EC, 1993. US EPA/EC Joint Project on the Evaluation of (Quantitative) Structure Activity Relationships.

US EPA, 1996. Ecological effects test guidelines – OPPTS 850.1000. Special considerations for conducting aquatic laboratory studies. Public Draft, EPA712-C-96-113. United States Environmental Protection Agency. http:/www.epa.gov/docs/OPTS_harmonized/

Van Den Berg, M., Van De Meet, D., Peijnenburg, W.J.G.M., Sijm, D.T.H.M., Struijs, J., Tas, J.W. 1995. Transport, accumulation and transformation processes. In: Risk Assessment of Chemicals: An Introduction. van Leeuwen, C.J., Hermens, J.L.M. (eds). Dordrecht, NL. Kluwer Academic Publishers, 37-102.

Wakabayashi, M., Kikuchi, M., Sato, A. Yoshida, T. 1987. Bioconcentration of alcohol ethoxylates in carp (*Cyprinus carpio*), Ecotoxicol. Environ. Safety 13, 148-163.

Wofford, H.W., C.D. Wilsey, G.S. Neff, C.S. Giam & J.M. Neff (1981): Bioaccumulation and metabolism of phthalate esters by oysters, brown shrimp and sheepshead minnows. Ecotox.Environ.Safety 5:202-210, 1981.

6. USE OF QSAR

6.1 HISTORY

262. Quantitative Structure-Activity Relationships (QSAR) in aquatic toxicology can be traced to the work at the turn of the century of Overton in Zürich (Lipnick, 1986) and Meyer in Marburg (Lipnick, 1989a). They demonstrated that the potency of substances producing narcosis in tadpoles and small fish is in direct proportion to their partition coefficients measured between olive oil and water. Overton postulated in his 1901 monograph "Studien über die Narkose," that this correlation reflects toxicity taking place at a standard molar concentration or molar volume within some molecular site within the organism (Lipnick, 1991a). In addition, he concluded that this corresponds to the same concentration or volume for a various organisms, regardless of whether uptake is from water or via gaseous inhalation. This correlation became known in anaesthesia as the Meyer-Overton theory.

263. Corwin Hansch and co-workers at Pomona College proposed the use of n-octanol/water as a standard partitioning system, and found that these partition coefficients were an additive, constitutive property that can be directly estimated from chemical structure. In addition, they found that regression analysis could be used to derive QSAR models, providing a statistical analysis of the findings. Using this approach, in 1972 these workers reported 137 QSAR models in the form log (1/C) = A log K_{ow} + B, where K_{ow} is the n-octanol/water partition coefficient, and C is the molar concentration of a chemical yielding a standard biological response for the effect of simple non-electrolyte non-reactive organic compounds on whole animals, organs, cells, or even pure enzymes. Five of these equations, which relate to the toxicity of five simple monohydric alcohols to five species of fish, have almost identical slopes and intercepts that are in fact virtually the same as those found by Könemann in 1981, who appears to have been unaware of Hansch's earlier work. Könemann and others have demonstrated that such simple non-reactive non-electrolytes all act by a narcosis mechanism in an acute fish toxicity test, giving rise to minimum or baseline toxicity (Lipnick, 1989b).

6.2 EXPERIMENTAL ARTIFACTS CAUSING UNDERESTIMATION OF HAZARD

264. Other non-electrolytes can be more toxic than predicted by such a QSAR, but not less toxic, except as a result of a testing artefact. Such testing artefacts include data obtained for compounds such as hydrocarbons which tend to volatilise during the experiment, as well as very hydrophobic compounds for which the acute testing duration may be inadequate to achieve steady state equilibrium partitioning between the concentration in the aquatic phase (aquarium test solution), and the internal hydrophobic site of narcosis action. A QSAR plot of log K_{ow} vs log C for such simple non-reactive non-electrolytes exhibits a linear relationship so long as such equilibrium is established within the test duration. Beyond this point, a bilinear relationship is observed, with the most toxic chemical being the one with the highest log K_{ow} value for which such equilibrium is established (Lipnick, 1995).

265. Another testing problem is posed by water solubility cut-off. If the toxic concentration required to produce the effect is above the compound's water solubility, no effect will be observed even at water saturation. Compounds for which the predicted toxic concentration is close to water solubility will also show no effect if the test duration is insufficient to achieve equilibrium partitioning. A similar cut-off is observed for surfactants if toxicity is predicted at a concentration beyond the critical micelle concentration. Although such compounds may show no toxicity under these conditions when tested alone, their toxic contributions to mixtures are still present. For compounds with the same log K_{ow} value, differences in water solubility reflect differences in

enthalpy of fusion related to melting point. Melting point is a reflection of the degree of stability of the crystal lattice and is controlled by intermolecular hydrogen bonding, lack of conformational flexibility, and symmetry. The more highly symmetric a compound, the higher the melting point (Lipnick, 1990).

6.3 QSAR MODELLING ISSUES

266. Choosing an appropriate QSAR implies that the model will yield a reliable prediction for the toxicity or biological activity of an untested chemical. Generally speaking, reliability decreases with increasing complexity of chemical structure, unless a QSAR has been derived for a narrowly defined set of chemicals similar in structure to the candidate substance. QSAR models derived from narrowly defined categories of chemicals are commonly employed in the development of pharmaceuticals once a new lead compound is identified and there is a need to make minor structural modifications to optimise activity (and decrease toxicity). Overall, the objective is make estimates by interpolation rather than extrapolation.

267. For example, if 96-h LC50 test data for fathead minnow are available for ethanol, n-butanol, n-hexanol, and n-nonanol, we have some confidence in making a prediction for this endpoint for n-propanol and n-pentanol. In contrast, we would have less confidence in making such a prediction for methanol, which is an extrapolation, with fewer carbon atoms than any of the tested chemicals. In fact, the behaviour of the first member of such a homologous is typically the most anomalous, and should not be predicted using data from remaining members of the series. Even the toxicity of branched chain alcohols may be an unreasonable extrapolation, depending upon the endpoint in question. Such extrapolation becomes more unreliable to the extent that toxicity is related to production of metabolites for a particular endpoint, as opposed to the properties of the parent compound. Also, if toxicity is mediated by a specific receptor binding mechanism, dramatic effects may be observed with small changes in chemical structure.

268. What ultimately governs the validity of such predictions is the degree to which the compounds used to derive the QSAR for a specific biological endpoint, are acting by a common molecular mechanism. In many and perhaps most cases, a QSAR does not represent such a mechanistic model, but merely a correlative one. A truly valid mechanistic model must be derived from a series of chemicals all acting by a common molecular mechanism, and fit to an equation using one or more parameters that relate directly to one or more steps of the mechanism in question. Such parameters or properties are more generally known as molecular descriptors. It is also important to keep in mind that many such molecular descriptors in common use may not have a direct physical interpretation. For a correlative model, the statistical fit of the data are likely to be poorer than a mechanistic one given these limitations. Mechanisms are not necessarily completely understood, but enough information may be known to provide confidence in this approach. For correlative models, the predictive reliability increases with the narrowness with which each is defined, e.g., categories of electrophiles, such as acrylates, in which the degree of reactivity may be similar and toxicity can be estimated for a "new" chemical using a model based solely on the log K_{ow} parameter.

269. As an example, primary and secondary alcohols containing a double or triple bond that is conjugated with the hydroxyl function (i.e., allylic or propargylic) are more toxic than would be predicted for a QSAR for the corresponding saturated compounds. This behaviour has been ascribed to a proelectrophile mechanism involving metabolic activation by the ubiquitous enzyme alcohol dehydrogenase to the corresponding α,β-unsaturated aldehydes and ketones which can act as electrophiles via a Michael-type acceptor mechanism (Veith *et al.*, 1989). In the presence of an

alcohol dehydrogenase inhibitor, these compounds behave like other alcohols and do not show excess toxicity, consistent with the mechanistic hypothesis.

270. The situation quickly becomes more complex once one goes beyond such a homologous series of compounds. Consider, for example, simple benzene derivatives. A series of chlorobenzenes may be viewed as similar to a homologous series. Not much difference is likely in the toxicities of the three isomeric dichlorobenzenes, so that a QSAR for chlorobenzenes based upon test data for one of these isomers is likely to be adequate. What about the substitution of other functional groups on benzene ring? Unlike an aliphatic alcohol, addition of a hydroxyl functionality to a benzene ring produces a phenol which is no longer neutral, but an ionizable acidic compound, due to the resonance stabilisation of the resulting negative charge. For this reason, phenol does not act as a true narcotic agent. With the addition of electron withdrawing substituents to phenol (e.g., chlorine atoms), there is a shift to these compounds acting as uncouplers of oxidative phosphorylation (e.g., the herbicide dinoseb). Substitution of an aldehyde group leads to increased toxicity via an electrophile mechanism for such compounds react with amino groups, such as the lysine ε-amino group to produce a Schiff Base adduct. Similarly, a benzylic chloride acts as an electrophile to form covalent abducts with sulfhydryl groups. In tackling a prediction for an untested compound, the chemical reactivity of these and many other functional groups and their interaction with one another should be carefully studied, and attempts made to document these from the chemical literature (Lipnick, 1991b).

271. Given these limitations in using QSARs for making predictions, it is best employed as a means of establishing testing priorities, rather than as a means of substituting for testing, unless some mechanistic information is available on the untested compound itself. In fact, the inability to make a prediction along with known environmental release and exposure may in itself be adequate to trigger testing or the development of a new QSAR for a category of chemicals for which such decisions are needed. A QSAR model can be derived by statistical analysis, e.g., regression analysis, from such a data set. The most commonly employed molecular descriptor, log K_{ow}, may be tried as a first attempt.

272. By contrast, derivation of a mechanism based QSAR model requires an understanding or working hypothesis of molecular mechanism and what parameter or parameters would appropriately model these actions. It is important to keep in mind that this is different from a hypothesis regarding mode of action, which relates to biological/physiological response, but not molecular mechanism.

6.4 USE OF QSARs IN AQUATIC CLASSIFICATION

273. The following inherent properties of substances are relevant for classification purposes concerning the aquatic environment:

- partition coefficient n-octanol-water log Kow;
- bioconcentration factor BCF;
- degradability - abiotic and biodegradation;
- acute aquatic toxicity for fish, daphnia and algae;
- prolonged toxicity for fish and daphnia.

274. Test data always take precedence over QSAR predications, providing the test data are valid, with QSARs used for filling data gaps for purposes of classification. Since the available QSARs are of varying reliability and application range, different restrictions apply for the prediction of each of these endpoints. Nevertheless, if a tested compound belongs to a chemical category or

structure type (see above) for which there is some confidence in the predictive utility of the QSAR model, it is worthwhile to compare this prediction with the experimental data, as it is not unusual to use this approach to detect some of the experimental artefacts (volatilisation, insufficient test duration to achieve equilibrium, and water solubility cut-off) in the measured data, which would mostly result in classifying substances as lower than actual toxicity.

275. When two or more QSARs are applicable or appear to be applicable, it is useful to compare the predictions of these various models in the same way that predicted data should be compared with measured (as discussed above). If there is no discrepancy between these models, the result provides encouragement of the validity of the predictions. Of course, it may also mean that the models were all developed using data on similar compounds and statistical methods. On the other hand, if the predictions are quite different, this result needs to be examined further. There is always the possibility that none of the models used provides a valid prediction. As a first step, the structures and properties of the chemicals used to derive each of the predictive models should be examined to determine if any models are based upon chemicals similar in both of these respects to the one for which a prediction is needed. If one data set contains such an appropriate analogue used to derive the model, the measured value in the database for that compound vs model prediction should be tested. If the results fit well with the overall model, it is likely the most reliable one to use. Likewise, if none of the models contain test data for such an analogue, testing of the chemical in question is recommended.

276. The U.S. EPA has recently posted a draft document on its website "Development of Chemical Categories in the HPV Challenge Program," that proposes the use of chemical categories to "... voluntarily compile a Screening Information Data Set (SIDS) on all chemicals on the US HPV list ... [to provide] basic screening data needed for an initial assessment of the physicochemical properties, environmental fate, and human and environmental effects of chemicals" (US EPA, 1999). This list consists of "...about 2,800 HPV chemicals which were reported for the Toxic Substances Control Act's 1990 Inventory Update Rule (IUR)".

277. One approach being proposed "...where this is scientifically justifiable ... is to consider closely related chemicals as a group, or category, rather than test them as individual chemicals. In the category approach, not every chemical needs to be tested for every SIDS endpoint". Such limited testing could be justified providing that the "...final data set must allow one to assess the untested endpoints, ideally by *interpolation* [emphasis added here] between and among the category members." The process for defining such categories and in the development of such data are described in the proposal.

278. A second potentially less data intensive approach being considered (US EPA, 2000a) is "... applying SAR principles to a single chemical that is closely related to one or more better characterised chemicals ("analogs")." A third approach proposed consists of using "... a combination of the analogue and category approaches ... [for] individual chemicals ... [similar to that] used in ECOSAR (US EPA, 2000b), a SAR-based computer program that generates ecotoxicity values. ". The document also details the history of the use of SARs within the U.S. EPA new chemicals program, and how to go about collecting and analysing data for the sake of such SAR approaches.

279. The Nordic Council of Ministers issued a report (Pederson *et al.*, 1995) entitled "Environmental Hazard Classification," that includes information on data collection and interpretation, as well as a section (5.2.8) entitled "QSAR estimates of water solubility and acute aquatic toxicity". This section also discusses the estimation of physicochemical properties, including log K_{ow}. For the sake of classification purposes, estimation methods are recommended for

prediction of "minimum acute aquatic toxicity," for "...neutral, organic, non-reactive and non-ionizable compounds such as alcohols, ketones, ethers, alkyl, and aryl halides, and can also be used for aromatic hydrocarbons, halogenated aromatic and aliphatic hydrocarbons as well as sulphides and disulphides," as cited in an earlier OECD Guidance Document (OECD, 1995). The Nordic document also includes diskettes for a computerised application of some of these methods.

280. The European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) has published a report entitled "QSARs in the Assessment of the Environmental Fate and Effects of Chemicals," which describes the use of QSARs to "...check the validity of data or to fill data gaps for priority setting, risk assessment and classification" (ECETOC, 1998). QSARs are described for predicting environmental fate and aquatic toxicity. The report notes that "a consistent dataset for [an endpoint] covered ... for a well defined scope of chemical structures ("domain") [is needed] ... from which a training set is developed. The document also discusses the advantage of mechanism based models, the use of statistical analysis in the development of QSARs, and how to assess "outliers".

6.4.1 Partition coefficient n-octanol-water log Kow

281. Computerised methods such as CLOGP (US EPA, 1999), LOGKOW (US EPA, 2000a) and SPARC (US EPA, 2000b) are available to calculate log Kow directly from chemical structure. CLOGP and LOGKOW are based upon the addition of group contributions, while SPARC is based upon a more fundamental chemical structure algorithm. Caution should be used in using calculated values for compounds that can undergo hydrolysis in water or some other reaction, since these transformations need to be considered in the interpretation of aquatic toxicity test data for such reactive chemicals. Only SPARC can be employed in a general way for inorganic or organometallic compounds. Special methods are needed in making estimates of log Kow or aquatic toxicity for surface-active compounds, chelating compounds, and mixtures.

282. Log Kow values can be calculated for pentachlorophenol and similar compounds, both for the ionised and unionised (neutral) forms. These values can potentially be calculated for certain reactive molecules (e.g., benzotrichloride), but the reactivity and subsequent hydrolysis also need to be considered. Also, for such ionizable phenols, pKa is a second parameter. Specific models can be used to calculate log Kow values for organometallic compounds, but they need to be applied with caution since some of these compounds really exist in the form of ion pairs in water.

283. For compounds of extremely high lipophilicity, measurements up to about 6 to 6.5 can be made by shake flask, and can be extended up to about log Kow of 8 using the slow stirring approach (Bruijn *et al.*, 1989). Calculations are considered useful even in extrapolating beyond what can be measured by either of these methods. Of course, it should be kept in mind that if the QSAR models for toxicity, etc. are based on chemicals with lower log K_{ow} values, the prediction itself will also be an extrapolation; in fact, it is known that in the case of bioconcentration, the relationship with log K_{ow} becomes non-linear at higher values. For compounds with low log K_{ow} values, the group contribution can also be applied, but this is not very useful for hazard purposes since for such substances, particularly with negative log K_{ow} values, little if any partitioning can take place into lipophilic sites and as Overton reported, these substances produce toxicity through osmotic effects (Lipnick, 1986).

6.4.2 Bioconcentration factor BCF

284. If experimentally determined BCF values are available, these values should be used for classification. Bioconcentration measurements must be performed using pure samples at test concentrations within water solubility, and for an adequate test duration to achieve steady state equilibrium between the aqueous concentration and that in the fish tissue. Moreover, with bioconcentration tests of extended duration, the correlation with log K_{ow} levels off and ultimately decreases. Under environmental conditions, bioconcentration of highly lipophilic chemicals takes place by a combination of uptake from food and water, with the switch to food taking place at log Kow $\approx$ 6. Otherwise log K_{ow} values can be used with a QSAR model as a predictor of the bioaccumulation potential of organic compounds. Deviations from these QSARs tend to reflect differences in the extent to which the chemicals undergo metabolism in the fish. Thus, some chemicals, such as phthalate, can bioconcentrate significantly less than predicted for this reason. Also, caution should be applied in comparing predicted BCF values with those using radiolabeled compounds, where the tissue concentration thus detected may represent a mix of parent compound and metabolites or even covalently bound parent or metabolite.

285. Experimental log K_{ow} values are to be used preferentially. However, older shake flask values above 5.5 are not reliable and we are in many cases better off using some average of calculated values or having these remeasured using the slow stirring method (Bruijn *et al.*, 1989). If there is reasonable doubt about the accuracy of the measured data, calculated log Kow values shall be used.

6.4.3 Degradability - abiotic and biodegradation

286. QSARs for abiotic degradation in water phases are narrowly defined linear free energy relationships (LFERs) for specific categories of chemicals and mechanisms. For example, such LFERs are available for hydrolysis of benzylic chlorides with various substituents on the aromatic ring. Such narrowly defined LFER models tend to be very reliable if the needed parameters are available for the Substituent(s) in question. Photo degradation, i.e., reaction with UV produced reactive species, may be extrapolated from estimates for the air compartment. While these abiotic processes do not usually result in complete degradation of organic compounds, they are frequently significant starting points, and may be rate limiting. QSARs for calculating biodegradability are either compound specific (OECD, 1995) or group contribution models like the BIODEG program (Hansch and Leo, 1995; Meylan and Howard 1995; Hilal *et al.*, 1994; Howard *et al.*, 1992; Boethling *et al.*, 1994; Howard and Meylan 1992; Loonen *et al.*, 1999). While validated compound category specific models are very limited in their application range, the application range of group contribution models is potentially much broader, but limited to compounds containing the model substructures. Validation studies have suggested that the biodegradability predictions by currently available group contribution models may be used for prediction of "not ready biodegradability" (Pedersen *et al.*, 1995; Langenberg *et al.*, 1996; USEPA, 1993) – and thus in relation to aquatic hazard classification "not rapid degradability."

6.4.4 Acute aquatic toxicity for fish, daphnia and algae

287. The acute aquatic toxicity of non-reactive, non-electrolyte organic chemicals (baseline toxicity) can be predicted from their log K_{ow} value with a quite high level of confidence, provided the presence of electrophile, proelectrophile, or special mechanism functional groups (see above) were not detected. Problems remain for such specific toxicants, for which the appropriate QSAR has to be selected in a prospective manner: Since straightforward criteria for the identification of the relevant modes of action are still lacking, empirical expert judgement needs to be applied for selecting a suitable model. Thus, if an inappropriate QSAR is employed, the predictions may be in

error by several orders of magnitude, and in the case of baseline toxicity, will be predicted less toxic, rather than more.

6.4.5 Prolonged toxicity for fish and Daphnia

288. Calculated values for chronic toxicity to fish and Daphnia should not be used to overrule classification based on experimental acute toxicity data. Only a few validated models are available for calculating prolonged toxicity for fish and Daphnia. These models are based solely on log Kow correlations and are limited in their application to non-reactive, non-electrolyte organic compounds, and are not suitable for chemicals with specific modes of action under prolonged exposure conditions. The reliable estimation of chronic toxicity values depends on the correct discrimination between non-specific and specific chronic toxicity mechanisms; otherwise, the predicted toxicity can be wrong by orders of magnitude. It should be noted that although for many compounds, excess toxicity[3] in a chronic test correlates with excess toxicity in an acute test, this is not always the case.

[3] Excess toxicity, T_e = (Predicted baseline toxicity) / Observed toxicity

ANNEX 6.I

REFERENCES

Boethling, R.S., Howard, P.H., Meylan, W.M. Stiteler, W.M., Beauman, J.A., and Tirado, N. (1994). Group contribution method for predicting probability and rate of aerobic biodegradation. Envir. Sci. Technol., 28, 459-465.

De Bruijn, J, Busser, F., Seinen, W., and Hermens, J. (1989), Determination of octanol/water partition coefficients for hydrophobic organic chemicals with the "slow-stirring method," Environ. Toxicol. Chem.., 8, 499-512.

ECETOC (1998), QSARs in the Assessment of the Environmental Fate and Effects of Chemicals, Technical report No 74.

Hansch, C. and A. Leo (1995), *Exploring QSAR,* American Chemical Society.

Hilal, S. H., L. A. Carreira and S. W. Karickhoff (1994), *Quantitative Treatments of Solute/solvent Interactions, Theoretical and Computational Chemistry, Vol. 1,* 291-353, Elsevier Science.

Howard, P.H., Boethling, R.S, Stiteler, W.M., Meylan, W.M., Hueber, A.E., Beaumen, J.A. and Larosche, M.E. (1992). Predictive model for aerobic biodegradation developed from a file of evaluated biodegradation data. Envir. Toxicol. Chem. 11, 593-603.

Howard, P. And Meylan, W.M. (1992). Biodegradation Probability Program, Version 3, Syracuse Research Corp., NY.

Langenberg, J.H., Peijnenburg, W.J.G.M. and Rorije, E. (1996). On the usefulness and reliability of existing QSARs for risk assessment and priority setting. SAR QSAR Environ. Res., 5, 1-16.

R.L. Lipnick (1986). Charles Ernest Overton: Narcosis studies and a contribution to general pharmacology. *Trends Pharmacol. Sci*., 7, 161-164.

R.L. Lipnick (1989a). Hans Horst Meyer and the lipoid theory of narcosis, *Trends Pharmacol. Sci*., 10 (7) July, 265-269; Erratum: 11 (1) Jan (1990), p. 44.

R.L. Lipnick (1989b). Narcosis, electrophile, and proelectrophile toxicity mechanisms. Application of SAR and QSAR. *Environ. Toxicol. Chem*., 8, 1-12 .

R.L. Lipnick (1990). Narcosis: Fundamental and Baseline Toxicity Mechanism for Nonelectrolyte Organic Chemicals. In: W. Karcher and J. Devillers (eds.) *Practical Applications of Quantitative Structure-Activity Relationships (QSAR) in Environmental Chemistry and Toxicology*, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 129-144.

R.L. Lipnick (ed.) (1991a). *Charles Ernest Overton: Studies of Narcosis and a Contribution to General Pharmacology*, Chapman and Hall, London, and Wood Library-Museum of Anesthesiology.

R.L. Lipnick (1991b). Outliers: their origin and use in the classification of molecular mechanisms of toxicity, *Sci. Tot. Environ*., 109/110 131-153.

R.L. Lipnick (1995). Structure-Activity Relationships. In: Fundamentals of Aquatic Toxicology, 2nd edition, (G.R. Rand, ed.), Taylor & Francis, London, 609-655.

Loonen, H., Lindgren, F., Hansen, B., Karcher, W., Niemela, J., Hiromatsu, K., Takatsuki, M., Peijnenburg, W., Rorije, E., and Struijs, J. (1999). Prediction of biodegradability from chemical structure: modeling of ready biodegradation test data. Environ. Toxicol. Chem., 18, 1763-1768.

Meylan, W. M. and P. H. Howard (1995), *J. Pharm. Sci.*, 84, 83-92.

OECD (1993), Structure-Activity Relationships for Biodegradation. OECD Environment Monograph No. 68 OECD, Paris, France.

OECD (1995). Environment Monographs No. 92. Guidance Document for Aquatic Effects Assessment. OECD, Paris.

F. Pedersen, H. Tyle, J. R. Niemelä, B. Guttmann, L. Lander, and A. Wedebrand (1995), Environmental Hazard Classification: Data Collection and Interpretation Guide for Substances to be Evaluated for Classification as Dangerous for the Environment, 2nd Edition, TemaNord 1995:581, Nordic Council of Ministers, Copenhagen, January.

US EPA (1999) Development of Chemical Categories in the HPV Challenge Program, http://www.epa.gov/chemrtk/categuid.htm

US EPA (2000a), The Use of Structure-Activity Relationships (SAR) in the High Production Volume Chemicals Challenge Program,
http://www.epa.gov/chemrtk/sarfinl1.htm

US EPA (2000b), ECOSAR,
http://www.epa.gov/oppt/newchems/21ecosar.htm

US EPA/EC (1993): US EPA Joint Project on the Evaluation of (Quantitative) Structure Activity Relationships, Commission of European Communities, Final Report, July.

G.D. Veith, R.L. Lipnick, and C.L. Russom (1989). The toxicity of acetylenic alcohols to the fathead minnow, Pimephales promelas. Narcosis and proelectrophile activation. *Xenobiotica*, 19(5), 555-565.

7. CLASSIFICATION OF METALS AND METAL COMPOUNDS

7.1 INTRODUCTION

289. The harmonised system for classifying chemical substances is a hazard-based system, and the basis of the identification of hazard is the aquatic toxicity of the substances, and information on the degradation and bioaccumulation behaviour (OECD 1998). Since this document deals only with the hazards associated with a given substance when the substance is dissolved in the water column, exposure from this source is limited by the solubility of the substance in water and bioavailability of the substance in species in the aquatic environment. Thus, the hazard classification schemes for metals and metal compounds are limited to the hazards posed by metals and metal compounds when they are available (i.e., exist as dissolved metal ions, for example, as M^+ when present as $M\text{-}NO_3$), and do not take into account exposures to metals and metal compounds that are not dissolved in the water column but may still be bioavailable, such as metals in foods. This chapter does not take into account the non-metallic ion (e.g., CN-) of metal compounds which may be toxic or which may be organic and may pose bioaccumulation or persistence hazards. For such metal compounds the hazards of the non-metallic ions must also be considered.

290. The level of the metal ion which may be present in solution following the addition of the metal and/or its compounds, will largely be determined by two processes: the extent to which it can be dissolved, i.e., its water solubility, and the extent to which it can react with the media to transform to water soluble forms. The rate and extent at which this latter process, known as "transformation" for the purposes of this guidance, takes place can vary extensively between different compounds and the metal itself, and is an important factor in determining the appropriate hazard category. Where data on transformation are available, they should be taken into account in determining the classification. The Protocol for determining this rate is available as a separate Guidance Document (OECD, 2001).

291. Generally speaking, the rate at which a substance dissolves is not considered relevant to the determination of its intrinsic toxicity. However, for metals and many poorly soluble inorganic metal compounds, the difficulties in achieving dissolution through normal solubilisation techniques is so severe that the two processes of solubilisation and transformation become indistinguishable. Thus, where the compound is sufficiently poorly soluble that the levels dissolved following normal attempts at solubilisation do not exceed the available $L(E)C_{50}$, it is the rate and extent of transformation, which must be considered. The transformation will be affected by a number of factors, not least of which will be the properties of the media with respect to pH, water hardness, temperature etc. In addition to these properties, other factors such as the size and specific surface area of the particles which have been tested, the length of time over which exposure to the media takes place and, of course the mass or surface area loading of the substance in the media will all play a part in determining the level of dissolved metal ions in the water. Transformation data can generally, therefore, only be considered as reliable for the purposes of classification if conducted according to the standard Protocol referenced above.

292. This Protocol aims at standardising the principal variables such that the level of dissolved ion can be directly related to the loading of the substance added. It is this loading level which yields the level of metal ion equivalent to the available $L(E)C_{50}$ that can then be used to determine the hazard band appropriate for classification. The testing methodology is beyond the scope of this guidance but the strategy to be adopted in using the data from the testing protocol, and the data requirements needed to make that strategy work, will be described.

293. In considering the classification of metals and metal compounds, both readily and poorly soluble, recognition has to be paid to a number of factors. As defined in the Glossary of this document, the term "degradation" refers to the decomposition of organic molecules. For inorganic compounds and metals, clearly the concept of degradability, as it has been considered and used for organic substances, has limited or no meaning. Rather, the substance may be transformed by normal environmental processes to either increase or decrease the bioavailability of the toxic species. Equally, the log K_{ow} cannot be considered as a measure of the potential to accumulate. Nevertheless, the concepts that a substance, or a toxic metabolite/reaction product may not be rapidly lost from the environment and/or may bioaccumulate are as applicable to metals and metal compounds as they are to organic substances.

294. Speciation of the soluble form can be affected by pH, water hardness and other variables, and may yield particular forms of the metal ion which are more or less toxic. In addition, metal ions could be made non-available from the water column by a number of processes (e.g., mineralisation and partitioning). Sometimes these processes can be sufficiently rapid to be analogous to degradation in assessing chronic classification. However, partitioning of the metal ion from the water column to other environmental media does not necessarily mean that it is no longer bioavailable, nor does it mean that the metal has been made permanently unavailable.

295. Information pertaining to the extent of the partitioning of a metal ion from the water column, or the extent to which a metal has been or can be converted to a form that is less toxic or non-toxic is frequently not available over a sufficiently wide range of environmentally relevant conditions, and thus, a number of assumptions will need to be made as an aid in classification. These assumptions may be modified if available data show otherwise. In the first instance it should be assumed that the metal ions, once in the water, are not rapidly partitioned from the water column and thus these compounds do not meet the criteria. Underlying this is the assumption that, although speciation can occur, the species will remain available under environmentally relevant conditions. This may not always be the case, as described above, and any evidence available that would suggest changes to the bioavailability over the course of 28 days, should be carefully examined. The bioaccumulation of metals and inorganic metal compounds is a complex process and bioaccumulation data should be used with care. The application of bioaccumulation criteria will need to be considered on a case-by-case basis taking due account of all the available data.

296. A further assumption that can be made, which represents a cautious approach, is that, in the absence of any solubility data for a particular metal compound, either measured or calculated, the substance will be sufficiently soluble to cause toxicity at the level of the $L(E)C_{50}$, and thus may be classified in the same way as other soluble salts. Again, this is clearly not always the case, and it may be wise to generate appropriate solubility data.

297. This chapter deals with metals and metal compounds. Within the context of this Guidance Document, metals and metal compounds are characterised as follows, and therefore, organo-metals are outside the scope of this chapter:

(1) metals, M^0, in their elemental state are not soluble in water but may transform to yield the available form. This means that a metal in the elemental state may react with water or a dilute aqueous electrolyte to form soluble cationic or anionic products, and in the process the metal will oxidise, or transform, from the neutral or zero oxidation state to a higher one.

(2) in a simple metal compound, such as an oxide or sulphide, the metal already exists in the

oxidised state, so that further metal oxidation is unlikely to occur when the compound is introduced into an aqueous medium.

However, while oxidisation may not change, interaction with the media may yield more soluble forms. A sparingly soluble metal compound can be considered as one for which a solubility product can be calculated, and which will yield a small amount of the available form by dissolution. However, it should be recognised that the final solution concentration may be influenced by a number of factors, including the solubility product of some metal compounds precipitated during the transformation/dissolution test, e.g. aluminium hydroxide.

7.2 APPLICATION OF AQUATIC TOXICITY DATA AND SOLUBILITY DATA FOR CLASSIFICATION

7.2.1 Interpretation of aquatic toxicity data

298. Aquatic toxicity studies carried out according to a recognised protocol should normally be acceptable as valid for the purposes of classification. Chapter 3 should also be consulted for generic issues that are common to assessing any aquatic toxicity data point for the purposes of classification.

Metal complexation and speciation

299. The toxicity of a particular metal in solution, appears to depend primarily on (but is not strictly limited to) the level of dissolved free metal ions. Abiotic factors including alkalinity, ionic strength and pH can influence the toxicity of metals in two ways: by influencing the chemical speciation of the metal in water (and hence affecting the availability) and by influencing the uptake and binding of available metal by biological tissues.

300. Where speciation is important, it may be possible to model the concentrations of the different forms of the metal, including those that are likely to cause toxicity. Analysis methods for quantifying exposure concentrations, which are capable of distinguishing between the complexed and uncomplexed fractions of a test substance, may not always be available or economic.

301. Complexation of metals to organic and inorganic ligands in test media and natural environments can be estimated from metal speciation models. Speciation models for metals, including pH, hardness, DOC, and inorganic substances such as MINTEQ (Brown and Allison, 1987), WHAM (Tipping, 1994) and CHESS (Santore and Driscoll, 1995) can be used to calculate the uncomplexed and complexed fractions of the metal ions. Alternatively, the Biotic Ligand Model (BLM), allows for the calculation of the concentration of metal ion responsible for the toxic effect at the level of the organism. The BLM model has at present only been validated for a limited number of metals, organisms, and end-points (Santore and Di Toro, 1999). The models and formula used for the characterisation of metal complexation in the media should always be clearly reported, allowing for their translation back to natural environments (OECD, 2000).

7.2.2 Interpretation of solubility data

302. When considering the available data on solubility, their validity and applicability to the identification of the hazard of metal compounds should be assessed. In particular, a knowledge of the pH at which the data were generated should be known.

Assessment of existing data

303. Existing data will be in one of three forms. For some well-studied metals, there will be solubility products and/or solubility data for the various inorganic metal compounds. It is also possible that the pH relationship of the solubility will be known. However, for many metals or metal compounds, it is probable that the available information will be descriptive only, e.g., poorly soluble. Unfortunately there appears to be very little (consistent) guidance about the solubility ranges for such descriptive terms. Where these are the only information available it is probable that solubility data will need to be generated using the Transformation/Dissolution Protocol.

Screening test for assessing solubility of metal compounds

304. In the absence of solubility data, a simple "Screening Test" for assessing solubility, based on the high rate of loading for 24 h can be used for metal compounds as described in the Transformation/Dissolution Protocol. The function of the screening test is to identify those metal compounds which undergo either dissolution or rapid transformation such that they are indistinguishable from soluble forms and hence may be classified based on the dissolved ion concentration. Where data are available from the screening test detailed in the Transformation/Dissolution Protocol, the maximum solubility obtained over the tested pH range should be used. Where data are not available over the full pH range, a check should be made that this maximum solubility has been achieved by reference to suitable thermodynamic speciation models or other suitable methods (see paragraph 301). It should be noted that this test is only intended to be used for metal compounds.

Full test for assessing solubility of metals and metal compounds

305. The first step in this part of the study is, as with the screening test, an assessment of the pH(s) at which the study should be conducted. Normally, the Full Test should have been carried out at the pH that maximises the concentration of dissolved metal ions in solution. In such cases, the pH may be chosen following the same guidance as given for the screening test.

306. Based on the data from the Full Test, it is possible to generate a concentration of the metal ions in solution after 7 days for each of the three loadings (i.e., 1 mg/L as "low", 10 mg/L as "medium" and 100mg/L as "high") used in the test. If the purpose of the test is to assess the long-term hazard of the substance, then the test at the low loading may be extended to 28 days, at an appropriate pH.

7.2.3 Comparison of aquatic toxicity data and solubility data

307. A decision whether or not the substance be classified will be made by comparing aquatic toxicity data and solubility data. If the $L(E)C_{50}$ is exceeded, irrespective of whether the toxicity and dissolution data are at the same pH and if this is the only data available then the substance should be classified. If other solubility data are available to show that the dissolution concentration would not exceed the $L(E)C_{50}$ across the entire pH range then the substance should not be classified on its soluble form. This may involve the use of additional data either from ecotoxicological testing or from applicable bioavailability-effect models.

7.3 ASSESSMENT OF ENVIRONMENTAL TRANSFORMATION

308. Environmental transformation of one species of a metal to another species of the same does not constitute degradation as applied to organic compounds and may increase or decrease the availability and bioavailability of the toxic species. However as a result of naturally occurring geochemical processes metal ions can partition from the water column. Data on water column residence time, the processes involved at the water – sediment interface (i.e., deposition and re-mobilisation) are fairly extensive, but have not been integrated into a meaningful database. Nevertheless, using the principles and assumptions discussed above in Section 7.1, it may be possible to incorporate this approach into classification.

309. Such assessments are very difficult to give guidance for and will normally be addressed on a case by case approach. However, the following may be taken into account:

- Changes in speciation if they are to non-available forms, however, the potential for the reverse change to occur must also be considered;
- Changes to a metal compound which is considerably less soluble than that of the metal compound being considered.

Some caution is recommended, see paragraph 293 and 294.

7.4 BIOACCUMULATION

310. While log K_{ow} is a good predictor of BCF for certain types of organic compounds e.g., non-polar organic substances, it is of course irrelevant for inorganic substances such as inorganic metal compounds.

311. The mechanisms for uptake and depuration rates of metals are very complex and variable and there is at present no general model to describe this. Instead the bioaccumulation of metals according to the classification criteria should be evaluated on a case by case basis using expert judgement.

312. While BCFs are indicative of the potential for bioaccumulation there may be a number of complications in interpreting measured BCF values for metals and inorganic metal compounds. For some metals and inorganic metal compounds the relationship between water concentration and BCF in some aquatic organisms is inverse, and bioconcentration data should be used with care. This is particularly relevant for metals that are biologically essential. Metals that are biologically essential are actively regulated in organisms in which the metal is essential. Since nutritional requirement of the organisms can be higher than the environmental concentration, this active regulation can results in high BCFs and an inverse relationship between BCFs and the concentration of the metal in water. When environmental concentrations are low, high BCFs may be expected as a natural consequence of metal uptake to meet nutritional requirements and in these instances can be viewed as a normal phenomenon. Additionally, if internal concentration is regulated by the organism, then measured BCFs may decline as external concentration increases. When external concentrations are so high that they exceed a threshold level or overwhelm the regulatory mechanism, this can cause harm to the organism. Also, while a metal may be essential in a particular organism, it may not be essential in other organisms. Therefore, where the metal is not essential or when the bioconcentration of an essential metal is above nutritional levels special consideration should be given to the potential for bioconcentration and environmental concern.

7.5 APPLICATION OF CLASSIFICATION CRITERIA TO METALS AND METAL COMPOUNDS

7.5.1 Introduction to the classification strategy for metals and metal compounds

313. The schemes for the classification of metals and metal compounds are described below and summarised diagrammatically in Figure 1. There are several stages in these schemes where data are used for decision purposes. It is not the intention of the classification schemes to generate new data. In the absence of valid data, it will be necessary to use all available data and expert judgement.

In the following sections, the reference to the $L(E)C_{50}$ refers to the data point(s) that will be used to select the classification band for the metal or metal compound.

314. When considering $L(E)C_{50}$ data for metal compounds, it is important to ensure that the data point to be used as the justification for the classification is expressed in the weight of the molecule of the metal compound to be classified. This is known as correcting for molecular weight. Thus while most metal data is expressed in, for example, mg/L of the metal, this value will need to be adjusted to the corresponding weight of the metal compound. Thus:

$L(E)C_{50}$ metal compounds
= $L(E)C_{50}$ of metal x (Molecular Weight of metal compound/Atomic Weight of metal)

NOEC data may also need to be adjusted to the corresponding weight of the metal compounds.

7.5.2 Classification Strategy for Metals

315. Where the $L(E)C_{50}$ for the metal ions of concern is greater than 100mg/L, the metals need not be considered further in the classification scheme.

316. Where the $L(E)C_{50}$ for the metal ions of concern is less than or equal to 100mg/L, consideration must be given to the data available on the rate and extent to which these ions can be generated from the metal. Such data, to be valid and useable should have been generated using the Transformation/Dissolution Protocol.

317. Where such data are unavailable, i.e., there is no clear data of sufficient validity to show that the transformation to metal ions will not occur, the safety net classification (Chronic IV) should be applied since the known classifiable toxicity of these soluble forms is considered to produce sufficient concern.

318. Where data from dissolution protocol are available, then, the results should be used to aid classification according to the following rules:

<u>7 day Transformation Test</u>

319. If the dissolved metal ion concentration after a period of 7 days (or earlier) exceeds that of the $L(E)C_{50}$, then the default classification for the metals is replaced by the following classification:

i) If the dissolved metal ion concentration at the low loading rate is greater than or equal to the $L(E)C_{50}$, then classify Acute Category I. Classify also as Chronic Category I,

unless there is evidence of both rapid partitioning from the water column and no bioaccumulation;

ii) If the dissolved metal ion concentration at the medium loading rate is greater than or equal to the $L(E)C_{50}$, then classify Acute Category II. Classify also as Chronic Category II unless there is evidence of both rapid partitioning from the water column and no bioaccumulation;

iii) If the dissolved metal ion concentration at the high loading rate is greater than or equal to the $L(E)C_{50}$, then classify Acute Category III. Classify also as Chronic Category III unless there is evidence of both rapid partitioning from the water column and no bioaccumulation.

28 day Transformation Test

320. If the process described in paragraph 319 results in the classification of Chronic I, no further assessment is required, as the metal will be classified irrespective of any further information.

321. In all other cases, further data may have been generated through the dissolution/transformation test in order to show that the classification may be amended. If for substances classified Chronic II, III or IV, the dissolved metal ion concentration at the low loading rate after a total period of 28 days is less than or equal to the of the long-term NOECs, then the classification is removed.

7.5.3 Classification strategy for metal compounds

322. Where the $L(E)C_{50}$ for the metal ions of concern is greater than 100mg/L, the metal compounds need not be considered further in the classification scheme.

If solubility ≥ $L(E)C_{50}$, classify on the basis of soluble ion

323. All metal compounds with a water solubility (either measured e.g., through 24-hour Dissolution Screening test or estimated e.g., from the solubility product) greater or equal to the $L(E)C_{50}$ of the dissolved metal ion concentration are considered as readily soluble metal compounds. Care should be exercised for compounds whose solubility is close to the acute toxicity value as the conditions under which solubility is measured could differ significantly from those of the acute toxicity test. In these cases the results of the Dissolution Screening Test are preferred.

324. Readily soluble metal compounds are classified on the basis of the $L(E)C_{50}$ (corrected where necessary for molecular weight):

i) If the $L(E)C_{50}$ of the dissolved metal ion is less than or equal to 1 mg/L then classify Acute Category I. Classify also as Chronic I unless there is evidence of both rapid partitioning from the water column and no bioaccumulation;

ii) If the $L(E)C_{50}$ of the dissolved metal ion is greater than 1 mg/L but less than or equal to 10 mg/L then classify Acute Category II. Classify also as Chronic II unless there is evidence of both rapid partitioning from the water column and no bioaccumulation;

iii) If the $L(E)C_{50}$ of the dissolved metal ion is greater than 10 mg/L and less than or equal to 100 mg/L then classify Acute Category III, Classify also as Chronic

Category III unless there is evidence of both rapid partitioning from the water column and no bioaccumulation.

If solubility <L(E)C_{50}, classify default Chronic IV

325. In the context of the classification criteria, poorly soluble compounds of metals are defined as those with a known solubility (either measured e.g., through 24-hour Dissolution Screening test or estimated e.g., from the solubility product) less than the L(E)C_{50} of the soluble metal ion. In those cases when the soluble forms of the metal of poorly soluble metal compounds have a L(E)C_{50} less than or equal to 100 mg/L and the substance can be considered as poorly soluble the default safety net classification (Chronic IV) should be applied.

7 day Transformation Test

326. For poorly soluble metal compounds classified with the default safety net classification further information that may be available from the 7-day transformation/dissolution test can also be used. Such data should include transformation levels at low, medium and high loading levels.

327. If the dissolved metal ion concentration after a period of 7 days (or earlier) exceeds that of the L(E)C_{50}, then the default classification for the metals is replaced by the following classification:

i) If the dissolved metal ion concentration at the low loading rate is greater than or equal to the L(E)C_{50}, then classify Acute Category I. Classify also as Chronic Category I, unless there is evidence of both rapid partitioning from the water column and no bioaccumulation;
ii) If the dissolved metal ion concentration at the medium loading rate is greater than or equal to the L(E)C_{50}, then classify Acute Category II. Classify also as Chronic Category II unless there is evidence of both rapid partitioning from the water column and no bioaccumulation;
iii) If the dissolved metal ion concentration at the high loading rate is greater than or equal to the L(E)C_{50}, then classify Acute Category III. Classify also as Chronic Category III unless there is evidence of both rapid partitioning from the water column and no bioaccumulation.

28 day Transformation Test

328. If the process described in paragraph 327 results in the classification of Chronic I, no further assessment is required as the metal compound will be classified irrespective of any further information.

329. In all other cases, further data may have been generated through the dissolution/transformation test for 28 days in order to show that the classification may be amended. If for poorly soluble metal compounds classified as Chronic II, III or IV, the dissolved metal ion concentration at the low loading rate after a total period of 28 days is less than or equal to the long-term NOECs, then classification is removed.

7.5.4 Particle size and surface area

330. Particle size, or moreover surface area, is a crucial parameter in that any variation in the size or surface area tested may cause a significant change in the levels of metals ions released in a given

time-window. Thus, this particle size or surface area is fixed for the purposes of the transformation test, allowing the comparative classifications to be based solely on the loading level. Normally, the classification data generated would have used the smallest particle size marketed to determine the extent of transformation. There may be cases where data generated for a particular metal powder is not considered as suitable for classification of the massive forms. For example, where it can be shown that the tested powder is structurally a different material (e.g., different crystallographic structure) and/or it has been produced by a special process and cannot be generated from the massive metal, classification of the massive can be based on testing of a more representative particle size or surface area, if such data are available. The powder may be classified separately based on the data generated on the powder. However, in normal circumstances it is not anticipated that more than two classification proposals would be made for the same metal.

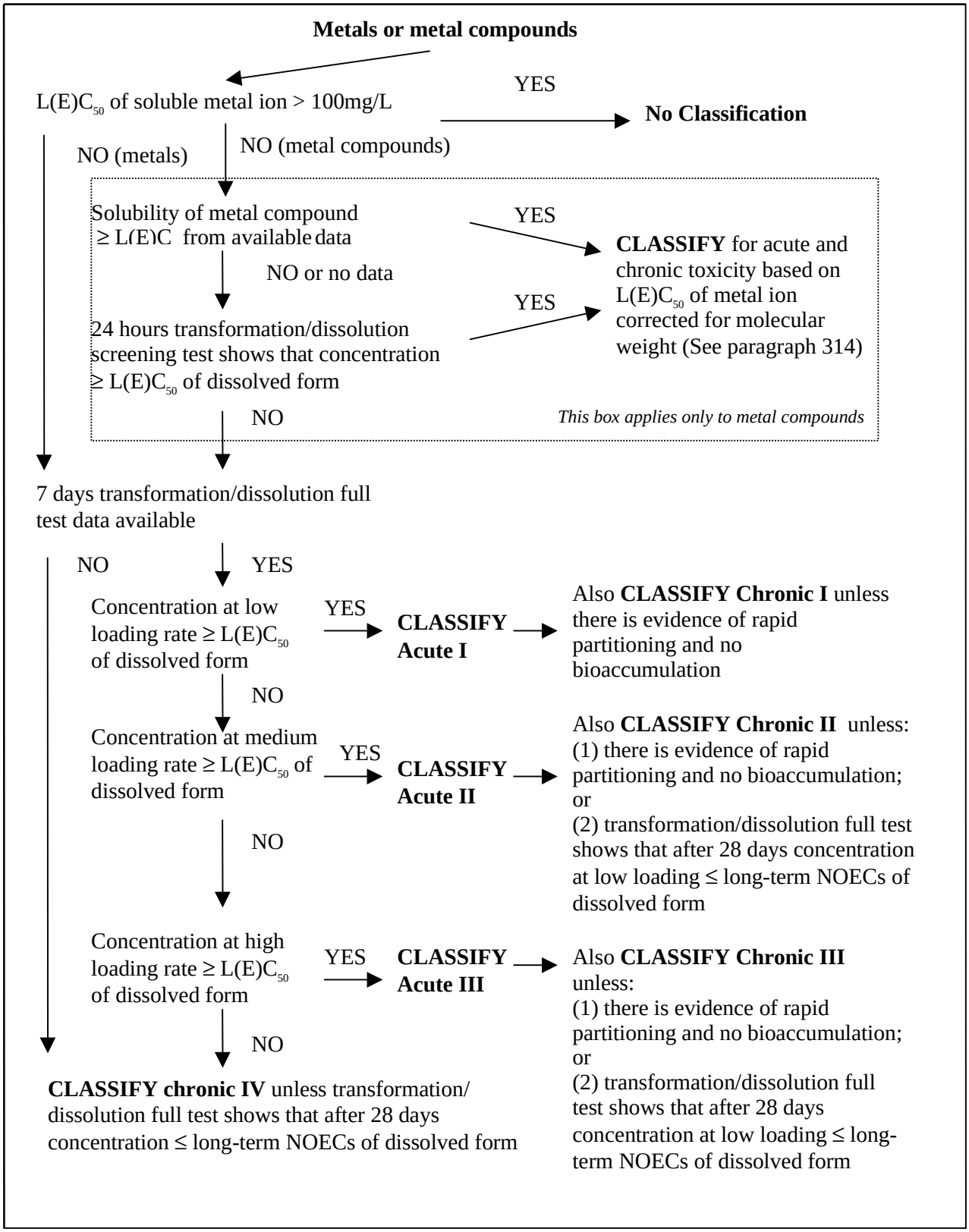
331. Metals with a particle size smaller than the default diameter value of 1 mm can be tested on a case-by-case basis. One example of this is where metal powders are produced by a different production technique or where the powders give rise to a higher dissolution (or reaction) rate than the massive form leading to a more stringent classification.

332. The particle sizes tested depend on the substance being assessed and are shown in the table below:

Type	Particle size	Comments
Metal compounds	Smallest representative size sold	Never larger than 1 mm
Metals – powders	Smallest representative size sold	May need to consider different sources if yielding different crystallographic / morphologic properties
Metals – massive	1 mm	Default value may be altered if sufficient justification

333. For some forms of metals, it may be possible, using the Transformation/Dissolution Protocol (OECD 2001), to obtain a correlation between the concentration of the metal ion after a specified time interval as a function of the surface area loadings of the forms tested. In such cases, it could then be possible to estimate the level of dissolved metal ion concentration of the metal with different particles, using the critical surface area approach as proposed by Skeaff *et. al.* (2000). That is, from this correlation and a linkage to the appropriate toxicity data, it may be possible to determine a critical surface area of the substance that delivers the $L(E)C_{50}$ to the medium and then to convert the critical surface area to the low, medium and high mass loadings used in hazard identification. While this approach is not normally used for classification it may provide useful information for labelling and downstream decisions.

FIGURE 1: Classification Strategy for metals and metal compounds



ANNEX 7.I

REFERENCES

Brown, D.S. and Allison, J.D. (1987). MINTEQA1 Equilibrium Metal Speciation Model: A user's manual. Athens, Georgia, USEPA Environmental Research Laboratory, Office of Research and Development.

OECD (1998). Harmonized Integrated Hazard Classification System for Human Health and Environmental Effects of Chemical Substances, http://www.oecd.org/ehs/Class/HCL6.htm

OECD (2000). Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures.

OECD (2001). Guidance Document on Transformation/Dissolution of Metals and Metals Compounds in Aqueous Media.

Santore, R.C. and Driscoll, C.T. (1995). The CHESS Model for Calculating Chemical Equilibria in Soils and Solutions, Chemical Equilibrium and Reaction Models. The Soil Society of America, American Society of Agronomy.

Santore, R.C. and Di Toro, D.M. et al (1999). A biotic ligand model of the acute toxicity of metals. II. Application to fish and daphnia exposure to copper. Environ. Tox. Chem. Submitted.

Skeaff, J., Delbeke, K., Van Assche, F. and Conard, B. (2000) A critical surface are concept for acute hazard classification of relatively insoluble metal-containing powders in aquatic environments. Environ. Tox. Chem. 19:1681-1691.

Tipping, E. (1994). WHAM – A computer equilibrium model and computer code for waters, sediments, and soils incorporating discrete site/electrostatic model of ion-binding by humic substances. Computers and Geoscience 20 (6): 073-1023.

APPENDIX

HARMONIZED SYSTEM FOR THE CLASSIFICATION OF CHEMICAL SUBSTANCES WHICH ARE HAZARDOUS FOR THE AQUATIC ENVIRONMENT

PURPOSE, BASIS AND APPLICABILITY

1. The harmonised system for classifying chemical substances for the hazards they present to the aquatic environment is based on a consideration of the existing systems listed below. The aquatic environment may be considered in terms of the aquatic organisms that live in the water, and the aquatic ecosystem of which they are part. To that extent, the proposal does not address aquatic pollutants for, which there may be a need to consider effects beyond the aquatic environment such as the impacts on human health etc. The basis, therefore, of the identification of hazard is the aquatic toxicity of the substance, although this may be modified by further information on the degradation and bioaccumulation behaviour.

2. The proposed system is intended specifically for use with chemical substances and is not intended at this stage to cover preparations or other mixtures such as formulated pesticides. Its application to mixtures is deferred to the OECD Working Group on Mixtures. While the scheme is intended to apply to all substances, it is recognised that for some substances, e.g. metals, poorly soluble substances etc., special guidance will be necessary. A Guidance Document will thus be prepared to cover issues such as data interpretation and the application of the criteria defined below to such groups of substances. Considering the complexity of this endpoint and the breadth of the application of the system, the Guidance Document is considered an important element in the operation of the harmonised scheme.

3. Consideration has been given to existing classification systems as currently in use, including the EU Supply and Use Scheme, the revised GESAMP hazard evaluation procedure, IMO Scheme for Marine Pollutant, the European Road and Rail Transport Scheme (RID/ADR), the Canadian and US Pesticide systems and the US Land Transport Scheme. The harmonised scheme is considered suitable for use for packaged goods in both supply and use and multimodal transport schemes, and elements of it may be used for bulk land transport and bulk marine transport under MARPOL 73/78 Annex II insofar as this uses aquatic toxicity.

DEFINITIONS AND DATA REQUIREMENTS

4. The basic elements for use within the harmonised system are:
- acute aquatic toxicity;
- potential for or actual bioaccumulation;
- degradation (biotic or abiotic) for organic chemicals; and
- chronic aquatic toxicity.

5. While data from internationally harmonised test methods are preferred, in practice, data from national methods may also be used where they are considered as equivalent. In general, it has been agreed that freshwater and marine species toxicity data can be considered as equivalent data and are preferably to be derived using OECD Test Guidelines or equivalent according to the principles of GLP. Where such data are not available classification should be based on the best available data.

Acute toxicity

6. Acute aquatic toxicity would normally be determined using a fish 96 hour LC_{50} (OECD Test Guideline 203 or equivalent), a crustacea species 48 hour EC_{50} (OECD Test Guideline 202 or equivalent) and/or an algal species 72 or 96 hour EC_{50} (OECD Test Guideline 201 or equivalent). These species are considered as surrogate for all aquatic organisms and data on other species such as Lemna may also be considered if the test methodology is suitable.

Bioaccumulation potential

7. The potential for bioaccumulation would normally be determined by using the octanol/water partition coefficient, usually reported as a log Kow determined by OECD Test Guideline 107 or 117. While this represents a potential to bioaccumulate, an experimentally determined Bioconcentration Factor (BCF) provides a better measure and should be used in preference when available. A BCF should be determined according to OECD Test Guideline 305.

Rapid degradability

8. Environmental degradation may be biotic or abiotic (e.g. hydrolysis) and the criteria used reflect this fact (Annex I). Ready biodegradation can most easily be defined using the OECD biodegradability tests OECD Test Guideline 301 (A - F). A pass level in these tests can be considered as indicative of rapid degradation in most environments. These are freshwater tests and thus the use of the results from OECD Test Guideline 306, which is more suitable for marine environments, has also been included. Where such data are not available, a BOD(5 days)/COD ratio >0.5 is considered as indicative of rapid degradation.

9. Abiotic degradation such as hydrolysis, primary degradation, both abiotic and biotic, degradation in non-aquatic media and proven rapid degradation in the environment may all be considered in defining rapid degradability. Special guidance on data interpretation will be provided in the Guidance Document.

Chronic toxicity

10. Chronic toxicity data are less available than acute data and the range of testing procedures less standardised. Data generated according to the OECD Test Guidelines 210 (Fish Early Life Stage), 202 Part 2 or 211 (Daphnia Reproduction) and 201 (Algal Growth Inhibition) can be accepted. Other validated and internationally accepted tests could also be used. The NOECs or other equivalent L(E)Cx should be used.

CLASSIFICATION CATEGORIES AND CRITERIA

11. Substances classified under the following criteria will be categorised as ‘hazardous to the aquatic environment’. These criteria describe in detail the classification categories detailed diagrammatically in Annex 2 to Appendix.

Acute toxicity

Category: Acute I

Acute toxicity:

96 hr LC_{50} (for fish)	≤1 mg/L and/or
48 hr EC_{50} (for crustacea)	≤1 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	≤1 mg/L.

Category: Acute I may be subdivided for some regulatory systems to include a lower band at $L(E)C_{50}$ ≤0.1 mg/L.

Category: Acute II

Acute toxicity:

96 hr LC_{50} (for fish)	>1 - ≤10 mg/L and/or
48 hr EC_{50} (for crustacea)	>1 - ≤10 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	>1 - ≤10 mg/L.

Category: Acute III

Acute toxicity:

96 hr LC_{50} (for fish)	>10 - ≤100 mg/L and/or
48 hr EC_{50} (for crustacea)	>10 - ≤100 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	>10 - ≤100 mg/L.

Some regulatory systems may extend this range beyond an $L(E)C_{50}$ of 100 mg/L through the introduction of another category.

Chronic toxicity

Category: Chronic I

Acute toxicity:

96 hr LC_{50} (for fish)	≤1 mg/L and/or
48 hr EC_{50} (for crustacea)	≤1 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	≤1 mg/L

and the substance is not rapidly degradable and/or the log Kow ≥ 4 (unless the experimentally determined BCF <500).

Category: Chronic II

Acute toxicity

96 hr LC_{50} (for fish)	>1 to ≤10 mg/L and/or
48 hr EC_{50} (for crustacea)	>1 to ≤10 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	>1 to ≤10 mg/L

and the substance is not rapidly degradable and/or the log Kow ≥4 (unless the experimentally determined BCF <500), unless the chronic toxicity NOECs are > 1 mg/L.

Category: Chronic III

Acute toxicity:

96 hr LC_{50} (for fish)	>10 to ≤100 mg/L and/or
48 hr EC_{50} (for crustacea)	>10 to ≤100 mg/L and/or
72 or 96hr ErC_{50} (for algae or other aquatic plants)	>10 to ≤100 mg/L

and the substance is not rapidly degradable and/or the log Kow ≥4 (unless the experimentally determined BCF <500) unless the chronic toxicity NOECs are >1 mg/L.

Category: Chronic IV
Poorly soluble substances for which no acute toxicity Is recorded at levels up to the water solubility, and which are not rapidly degradable and have a log Kow ≥ 4, indicating a potential to bioaccumulate, will be classified in this category unless other scientific evidence exists showing classification to be unnecessary. Such evidence would include an experimentally determined BCF <500, or a chronic toxicity NOECs >1 mg/L, or evidence of rapid degradation in the environment.

RATIONALE FOR THE SYSTEM

12. The system for classification recognises that the core intrinsic hazard to aquatic organisms is represented by both the acute and chronic toxicity of a substance, the relative importance of which is determined by the specific regulatory system in operation. Distinction can be made between the acute hazard and the chronic hazard and therefore separate hazard categories are defined for both properties representing a gradation in the level of hazard identified. The lowest of the available toxicity values will normally be used to define the appropriate hazard class(es). There may be circumstances, however, when a weight of evidence approach may be used. Acute toxicity data are the most readily available and the tests used are the most standardised. For that reason, these data form the core of the classification system.

13. Acute toxicity represents a key property in defining the hazard where transport of large quantities of a substance may give rise to short-term dangers arising from accidents or major spillages. Hazard categories up to $L(E)C_{50}$ values of 100 mg/L are thus defined although categories up to 1000 mg/L may be used in certain regulatory frameworks. The Acute: Category I may be further sub-divided to include an additional category for acute toxicity $L(E)C_{50}$ ≤0.1 mg/L in certain regulatory systems such as that defined by MARPOL 73/78 Annex II. It is anticipated that their use would be restricted to regulatory systems concerning bulk transport.

14. For packaged substances it is considered that the principal hazard is defined by chronic toxicity, although acute toxicity at $L(E)C_{50}$ levels ≤1 mg/L are also considered hazardous. Levels of substances up to 1 mg/L are considered as possible in the aquatic environment following normal use and disposal. At toxicity levels above this, it is considered that the short-term toxicity itself does not describe the principle hazard, which arises from low concentrations causing effects over a longer time scale. Thus, a number of hazard categories are defined which are based on levels of chronic aquatic toxicity. Chronic toxicity data are not available for many substances, however, and it is necessary to use the available data on acute toxicity to estimate this property. The intrinsic properties of a lack of rapid degradability and/or a potential to bioconcentrate in combination with acute toxicity may be used to assign a substance to a chronic hazard category. Where chronic toxicity is available showing NOECs >1 mg/L, this would indicate that no classification in a chronic hazard category would be necessary. Equally, for substances with an $L(E)C_{50}$ >100 mg/L, the toxicity is considered as insufficient to warrant classification in most regulatory systems.

15. While the current system will continue to rely on the use of acute toxicity data in combination with a lack of rapid degradation and/or a potential to bioaccumulate as the basis for classification for assigning a chronic hazard category, it is recognised that actual chronic toxicity data would form a better basis for classification where these data are available. It is thus the intention that the scheme should be further developed to accommodate such data. It is anticipated that in such a further development, the available chronic toxicity data would be used to classify in the chronic hazard in preference to that derived from their acute toxicity in combination with a lack of rapid degradation and/or a potential to bioaccumulate.

16. Recognition is given to the classification goals of MARPOL 73/78 Annex II that covers the transport of bulk quantities in ship tanks, which are aimed at regulating operational discharges from ships and assigning of suitable ship types. They go beyond that of protecting aquatic ecosystems, although that clearly is included. Additional hazard categories may thus be used which take account of factors such as physico-chemical properties and mammalian toxicity.

EXPLANATORY NOTES

17. The organisms fish, crustacea and algae are tested as surrogate species covering a range of trophic levels and taxa, and the test methods are highly standardised. Data on other organisms may also be considered, however, provided they represent equivalent species and test endpoints. The algal growth inhibition test is a chronic test but the EC_{50} is treated as an acute value for classification purposes. This EC_{50} should normally be based on growth rate inhibition. If only the EC_{50} based on reduction in biomass is available, or it is not indicated which EC_{50} is reported, this value may be used in the same way.

18. Aquatic toxicity testing by its nature, involves the dissolution of the substance under test in the water media used and the maintenance of a stable bioavailable exposure concentration over the course of the test. Some substances are difficult to test under standard procedures and thus special guidance will be developed on data interpretation for these substances and how the data should be used when applying the classification criteria.

19. It is the bioaccumulation of substances within the aquatic organisms that can give rise to toxic effects over longer time scales even when actual water concentrations are low. The potential to bioaccumulate is determined by the partitioning between n-octanol and water. The relationship between the partition coefficient of an organic substance and its bioconcentration as measured by the BCF in fish has considerable scientific literature support. Using a cut-off value of log P(o/w) $\geq$ 4 is intended to identify only those substances with a real potential to bioconcentrate. In recognition that the log P(o/w) is only an imperfect surrogate for a measured BCF, such a measured value would always take precedence. A BCF in fish of <500 is considered as indicative of a low level of bioconcentration.

20. Substances that rapidly degrade can be quickly removed from the environment. While effects can occur, particularly in the event of a spillage or accident, they will be localised and of short duration. The absence of rapid degradation in the environment can mean that a substance in the water has the potential to exert toxicity over a wide temporal and spatial scale. One way of demonstrating rapid degradation utilises the biodegradation screening tests designed to determine whether a substance is 'readily biodegradable'. Thus a substance, which passes this screening test, is one that is likely to biodegrade 'rapidly' in the aquatic environment, and is thus unlikely to be persistent. However, a fail in the screening test does not necessarily mean that the substance will not degrade rapidly in the environment. Thus a further criterion was added which would allow the use of data to show that the substance did actually degrade biotically or abiotically in the aquatic environment by >70% in 28 days. Thus, if degradation could be demonstrated under environmentally realistic conditions, then the definition of 'rapid degradability' would have been met. Many degradation data are available in the form of degradation half-lives and these can also be used in defining rapid degradation. Details regarding the interpretation of these data will be further elaborated in the Guidance Document. Some tests measure the ultimate biodegradation of the substance, i.e., full mineralisation is achieved. Primary biodegradation would not normally qualify in the assessment of rapid degradability unless it can be demonstrated that the degradation products do not fulfil the criteria for classification as hazardous to the aquatic environment.

21. It must be recognised that environmental degradation may be biotic or abiotic (e.g. hydrolysis) and the criteria used reflect this fact. Equally, it must be recognised that failing the ready biodegradability criteria in the OECD tests does not mean that the substance will not be degraded rapidly in the real environment. Thus where such rapid degradation can be shown, the substance should be considered as rapidly degradable. Hydrolysis can be considered if the hydrolysis products do not fulfil the criteria for classification as hazardous to the aquatic environment. A specific definition of rapid degradability is included as Annex 1. Other evidence of rapid degradation in the environment may also be considered and may be of particular importance where the substances are inhibitory to microbial activity at the concentration levels used in standard testing. The range of available data and guidance on its interpretation will be provided in the Guidance Document.

22. For inorganic compounds and metals, the concept of degradability as applied to organic compounds has limited or no meaning. Rather the substance may be transformed by normal environmental processes to either increase or decrease the bioavailability of the toxic species. Equally the use of bioaccumulation data should be treated with care. Specific guidance will be provided on how these data for such materials may be used in meeting the requirements of the classification criteria.

23. Poorly soluble inorganic compounds and metals may be acutely or chronically toxic in the aquatic environment depending on the intrinsic toxicity of the bioavailable inorganic species and the rate and amount of this species which may enter solution. A protocol for testing these poorly soluble materials is being developed and will be covered further in the special guidance.

24. The system also introduces as 'safety net' classification (Category: Chronic IV) for use when the data available does not allow classification under the formal criteria but there are nevertheless some grounds for concern. The precise criteria are not defined with one exception. For poorly water-soluble organic substances for which no toxicity has been demonstrated, classification can occur if the substance is both not rapidly degraded and has a potential to bioaccumulate. It is considered that for such poorly soluble substances, the toxicity may not have been adequately assessed in the short-term test due to the low exposure levels and potentially slow uptake into the organism. The need for this classification can be negated by demonstrating the absence of long-term effects, i.e., a long-term NOECs > water solubility or 1 mg/L, or rapid degradation in the environment.

25. While experimentally derived test data are preferred, where no experimental data are available, validated Quantitative Structure Activity Relationships (QSARs) for aquatic toxicity and log Kow may be used in the classification process. Such validated QSARs may be used without modification to the agreed criteria, if restricted to chemicals for which their mode of action and applicability are well characterised. Validity may be judged according to the criteria established within the USEPA/EU/Japan Collaborative Project. Reliable calculated toxicity and log Kow values should be valuable in the safety net context. QSARs for predicting ready biodegradation are not yet sufficiently accurate to predict rapid degradation.

ANNEX 1 to Appendix 2

RAPID DEGRADABILITY

Substances are considered rapidly degradable in the environment if the following criteria hold true:

a) if in 28-day ready biodegradation studies, the following levels of degradation are achieved;

- tests based on dissolved organic carbon: 70%
- tests based on oxygen depletion or carbon dioxide generation: 60% of theoretical maxima

These levels of biodegradation must be achieved within 10 days of the start of degradation which point is taken as the time when 10% of the substance has been degraded.

or

b) if, in those cases where only BOD and COD data are available, when the ratio of BOD5/COD is ≥0.5

or

c) if other convincing scientific evidence is available to demonstrate that the substance can be degraded (biotically and/or abiotically) in the aquatic environment to a level >70% within a 28 day period.

ANNEX 2 to Appendix 2

Classification Scheme for Substances Hazardous to the Aquatic Environment

<table>
<tr><th colspan="2">Toxicity</th><th rowspan="2">Degradability (note 3)</th><th rowspan="2">Bioaccumulation (note 4)</th><th colspan="2">Classification categories</th></tr>
<tr><th>Acute (note 1)</th><th>Chronic (note 2)</th><th>Acute</th><th>Chronic</th></tr>
<tr><td>Box 1
value ≤ 1.00</td><td rowspan="4"></td><td rowspan="4">Box 5

lack of rapid degradability</td><td rowspan="4">Box 6

BCF ≥ 500 or,
if absent
log Kow ≥ 4</td><td><u>Category:
Acute I</u>
Box 1</td><td><u>Category:
Chronic I</u>
Boxes 1+5+6
Boxes 1+5
Boxes 1+6</td></tr>
<tr><td>Box 2
1.00 < value
≤ 10.0</td><td><u>Category:
Acute II</u>
Box 2</td><td><u>Category:
Chronic II</u>
Boxes 2+5+6
Boxes 2+5
Boxes 2+6
Unless Box 7</td></tr>
<tr><td>Box 3
10.0 < value
≤ 100</td><td><u>Category:
Acute III</u>
Box 3</td><td><u>Category:
Chronic III</u>
Boxes 3+5+6
Boxes 3+5
Boxes 3+6
Unless Box 7</td></tr>
<tr><td>Box 4
No acute toxicity (note 5)</td><td>Box 7
value > 1.00</td><td><u>Category:
Chronic IV</u>
Boxes 4+5+6
Unless Box 7</td></tr>
</table>

<u>Notes to the table:</u>

Note 1a. Acute toxicity band based on L(E)C-50 values in mg/L for fish, crustacea and/or algae or other aquatic plants (or QSAR estimation if no experimental data)

Note 1b Where the algal toxicity ErC-50 [= EC-50 (growth rate)] falls more than 100 times below the next most sensitive species and results in a classification based solely on this effect, consideration should be given to whether this toxicity is representative of the toxicity to aquatic plants. Where it can be shown that this is not the case, professional judgement should be used in deciding if classification should be applied. Classification should be based on the ErC-50. In circumstances where the basis of the EC-50 is not specified and no ErC-50 is recorded, classification should be based on the lowest EC-50 available.

Note 2a. Chronic toxicity band based on NOEC values in mg/L for fish or crustacea or other recognised measures for long-term toxicity.

Note 2b. It is the intention that the system be further developed to include chronic toxicity data.

Note 3. Lack of rapid degradability is based on either a lack of Ready Biodegradability or other evidence of lack of rapid degradation.

Note 4. Potential to bioaccumulate, based on an experimentally derived BCF ≥ 500 or, if absent, a log Kow ≥ 4 provided log Kow is an appropriate descriptor for the bioaccumulation potential of the substance. Measured log Kow values take precedence over estimated values and measured BCF values take precedence over log Kow values.

Note 5. "No acute toxicity" is taken to mean that the L(E)C-50 is above the water solubility. Also for poorly soluble substances, (w.s. < 1.00 mg/L), where there is evidence that the acute test would not have provided a true measure of the intrinsic toxicity.

ANNEX 3

OECD GUIDANCE DOCUMENT No.29
GUIDANCE DOCUMENT ON TRANSFORMATION/DISSOLUTION OF METALS AND METAL COMPOUNDS IN AQUEOUS MEDIA

OECD Environment, Health and Safety Publications

Series on Testing and Assessment

No. 29

Draft Guidance Document on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media

Environment Directorate

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Paris

April 2001

FOREWORD

As part of a wider international effort on the global harmonisation of hazard classification systems, agreement was reached in the technical working groups on a set of criteria that would form the basis of a global scheme for classifying substances hazardous to the aquatic environment. Such scheme forms part of an international agreement on hazard classification of substances. The criteria were endorsed by the Joint Meeting of the OECD in November 1998 and form part of the Globally Harmonised Classification System (GHS) which will be implemented under ECOSOC in 2001. In developing the criteria, it was agreed that the detail needed to properly define the hazard to the environment resulted in a complex system for which some suitable guidance would be necessary. The harmonised proposal makes a number of references to a Guidance Document in the detailed explanation of the scheme. This Guidance document has been published in the Environment, Health and Safety Series on testing and Assessment as Document no 27.

In the Guidance Document a chapter (Chapter 7) is dedicated to the classification of metals and metal compounds. One of the major issues in this chapter is the bio-availability of metals and/or metal compounds. An OECD Workshop on Aquatic Toxicity Testing of Sparingly Soluble Metals, Inorganic Metal Compounds and Minerals" held in Ottawa in 1995 addressed this issue and concluded that a protocol on the transformation/dissolution of metals and metal compounds in aquatic media should be developed. The Metals Working Group took the lead in developing this protocol, until the group was merged with the Expert Group on Aquatic Environmental Hazards in March 2000. At the 6th Meeting of the newly formed Extended Expert Group on Aquatic Environmental Hazards it was agreed that the protocol which was then in its final stages of development should be prepared as a separate document.

This document is the outcome of the work undertaken by an ad-hoc Expert Group established under the Extended Expert Group.

The current protocol, as included in this Guidance Document is currently being considered for formal international validation. Therefore, it may be subject to changes depending on the outcome of the validation work and, therefore, will be revisited after completion of that exercise, if needed.

INTRODUCTION

1. This Test Guidance is designed to determine the rate and extent to which metals and sparingly soluble metal compounds can produce soluble available ionic and other metal-bearing species in aqueous media under a set of standard laboratory conditions representative of those generally occurring in the environment. Once determined, this information can be used to evaluate the short term and long term aquatic toxicity of the metal or sparingly soluble metal compound from which the soluble species came. This Test Guidance is the outcome of an international effort under the OECD to develop an approach for the toxicity testing and data interpretation of metals and sparingly soluble inorganic metal compounds (SSIMs) [ref to Ottawa workshop (1) and to Chapter 7 of the Guidance document]. As a result of recent meetings and discussions [references 1,2,3,4 + Chapter 7] held within the OECD and EU, the experimental work on several metals and metal compounds upon which this Test Guidance is based has been conducted and reported [references 5 to 11].

2. The evaluation of the short term and long term aquatic toxicity of metals and sparingly soluble metal compounds is to be accomplished by comparison of (a) the concentration of the metal ion in solution, produced during transformation or dissolution in a standard aqueous medium with (b) appropriate standard ecotoxicity data as determined with the soluble metal salt (acute and chronic values). This document gives guidance for performing the transformation/dissolution tests. The strategy to derive an environmental hazard classification using the results of the dissolution/transformation protocol is not within the scope of this Guidance document and can be found elsewhere (ref. to Chapter 7 of the Guidance document).

3. For this Test Guidance, the transformations of metals and sparingly soluble metal compounds are, within the context of the test, defined and characterised as follows :

(1) metals, M^0 , in their elemental state are not soluble in water but may transform to yield the available form. This means that a metal in the elemental state may react with the media to form soluble cationic or anionic products, and in the process the metal will oxidise, or transform, from the neutral or zero oxidation state to a higher one.

(2) in a simple metal compound, such as an oxide or sulphide, the metal already exists in an oxidised state, so that further metal oxidation is unlikely to occur when the compound is introduced into an aqueous medium. However, while oxidisation state may not change, interaction with the media may yield more soluble forms. A sparingly soluble metal compound can be considered as one for which a solubility product can be calculated, and which will yield small amount of the available form by dissolution. However, it should be recognised that the final solution concentration may be influenced by a number of factors, including the solubility product of some metal compounds precipitated during the transformation/dissolution test, e.g. aluminium hydroxide.

PRINCIPLES

4. This Test Guidance is intended to be a standard laboratory transformation/ dissolution protocol based on a simple experimental procedure of agitating various quantities of the test substance in a pH buffered aqueous medium, and sampling and analysing the solutions at specific time intervals to determine the concentrations of dissolved metal ions in the water. Two different types of tests are described in this document:

A. Screening transformation/dissolution test – sparingly soluble metal compounds

5. For sparingly soluble metal compounds, the maximum concentration of total dissolved metal can be determined by the solubility limit of the metal compound or from a screening transformation/dissolution test. The intent of the screening test, performed at a single loading, is to identify those compounds which undergo either dissolution or rapid transformation such that their ecotoxicity potential is indistinguishable from soluble forms.

6. Sparingly soluble metal compounds, having the smallest representative particle size on the market are introduced into the aqueous medium at a single loading of 100 mg/L. Such dissolution as will occur is achieved by agitation during a 24 hours period. After 24 hours agitation, the dissolved metal ion concentration is measured.

B. Full transformation/dissolution test - metals and sparingly soluble metal compounds

7. The full transformation/dissolution test is intended to determine level of the dissolution or transformation of metals and metal compounds after a certain time period at different loadings of the aqueous phase. Normally massive forms and/or powders are introduced into the aqueous medium at three different loadings: 1, 10 and 100 mg/L. A single loading of 100 mg/L may be used if a significant release of dissolved metal species is not anticipated. Transformation/dissolution is accomplished by standardised agitation, without causing abrasion of the particles. The short term transformation/dissolution endpoints are based on the dissolved metal ion concentrations obtained after a 7 days transformation/dissolution period. The long term transformation/dissolution endpoint is obtained during a 28 days transformation/dissolution test, using a single load of 1 mg/L.

8. As pH has a significant influence on transformation/dissolution both the screening test and the full test should in principle be carried out at a pH that maximises the concentration of the dissolved metal ions in solution. With reference to the conditions generally found in the environment a pH range of 6 to 8.5 must be used, except for the 28 day full test where the pH range of 5.5 to 8.5 should be used in order to take into consideration possible long term effects on acidic lakes.

9. As in addition the surface area of the particles in the test sample has an important influence on the rate and extent of transformation/dissolution, powders are tested at the smallest representative particle size as placed on the market, while massives are tested at a particle size representative of normal handling and use. A default diameter value of 1 mm should be used in absence of this information. For massive metals, this default may only be exceeded when sufficiently justified. The specific surface area should be determined in order to characterise and compare similar samples.

APPLICABILITY OF THE TEST

10. This test applies to all metals and sparingly soluble inorganic metal compounds. Exceptions, such as certain water reactive metals, should be justified.

INFORMATION ON THE TEST SUBSTANCE

11. Substances as placed on the market should be used in the transformation/dissolution tests. In order to allow for correct interpretation of the test results, it is important to obtain the following information on the test substance(s):

- substance name, formula and use on the market;
- physical-chemical method of preparation;
- identification of the batch used for testing;
- chemical characterisation: overall purity (%) and specific impurities (% or ppm);
- density (g/cm^3) or specific gravity;
- measured specific surface area (m^2/g)- measured by BET N_2 adsorption-desorption or equivalent technique;
- storage, expiration date;
- known solubility data and solubility products;
- hazard identification and safe handling precautions;
- Material Safety Data Sheets (MSDS) or equivalent;

DESCRIPTION OF THE TEST METHOD

Apparatus and reagents

12. The following apparatus and reagents are necessary for performing tests.

- Pre-cleaned and acid rinsed closed glass sample bottles (paragraph 13);
- transformation /dissolution medium (ISO 6341) (paragraph 14);
- test solution buffering facilities (paragraph 15);
- agitation equipment: orbital shaker, radial impeller, laboratory shaker or equivalent (paragraph 16);
- appropriate filters (e.g.0.2 µm Acrodisc) or centrifuge for solids-liquid separation (paragraph 18);
- means to control the temperature of the reaction vessels to + 2°C within the temperature range of 20°C to 25°C, such as a temperature controlled cabinet or a water bath;
- syringes and/or automatic pipettes;
- pH meter showing acceptable results within + 0.2 pH units;
- dissolved oxygen meter, with temperature reading capability;
- thermometer or thermocouple; and
- analytical equipment for metal analysis (e.g. atomic adsorption spectrometry, inductively coupled axial plasma spectrometry).

13. All glass test vessels must be carefully cleaned by standard laboratory practices, acid-cleaned (e.g. HCl) and subsequently rinsed with de-ionised water. The test vessel volume and configuration (one- or two-litre reaction kettles) should be sufficient to hold 1 or 2 L of aqueous medium without overflow during the agitation specified. If air buffering is used (tests carried out at pH 8), it is advised to increase the air buffering capacity of the medium by increasing the headspace/liquid ratio (e.g. 1 L medium in 2.8 L flasks).

14. A reconstituted standard water based on ISO 6341 should be used[4], as the standard transformation/dissolution medium. The medium should be sterilised by filtration (0.2 µm) before use in the tests. The chemical composition of the standard transformation/dissolution medium (for tests carried out at pH 8) is as follows:

$NaHCO_3$: 65.7 mg/L
KCl : 5.75 mg/L
$CaCl_2.2H_2O$: 294 mg/L
$MgSO_4.7H_2O$: 123 mg/L

For tests carried out at lower pH values, adjusted chemical compositions are given in paragraph 18.

15. The concentration of total organic carbon in the medium should not exceed 2.0mg/L.

16. In addition to the fresh water medium, the use of a standardised marine test medium may also be considered when the solubility or transformation of the metal compound is expected to be significantly affected by the high chloride content or other unique chemical characteristics of marine waters and when toxicity test data are available on marine species. When marine waters are considered, the chemical composition of the standard marine medium is as follows:

NaF:3mg/L
$SrCl_2 \cdot 6H_2O$:20mg/L
H_3BO_3:30mg/L
KBr:100mg/L
KCl:700mg/L
$CaCl_2 \cdot 2H2O$:1.47g/L
Na_2SO_4:4.0g/L
$MgCl_2 \cdot 6H2O$:10.78g/L
NaCl:23.5g/L
$Na_2SiO_3 \cdot 9H2O$:20mg/L
$NaHCO_3$:200mg/L

The salinity should be 34 $\pm$ 0.5g/kg and the pHshould be 8.0 $\pm$ 0.2. The reconstituted salt water should also be stripped of trace metals. (from ASTM E 729-96)

17. The transformation/dissolution tests are to be carried out at a pH that maximises the concentration of the dissolved metal ions in solution within the prescribed pH range. A pH-range of 6 to 8.5 must be used for the screening test and the 7 day full test, and a range of 5.5 to 8.5 for the 28 day full test (paragraph 8).

18. Buffering at pH 8 may be established by equilibrium with air, in which the concentration of CO_2 provides a natural buffering capacity sufficient to maintain the pH within an average of $\pm$ 0.2 pH units over a period of one week (reference 7). An increase in the headspace/liquid ratio can be used to improve the air buffering capacity of the medium.

[4] For hazard classification purposes the results of the dissolution/transformation protocol are compared with existing ecotoxicity data for metals and metal compounds. However, for purposes such as data validation, there might be cases where it may be appropriate to use the aqueous medium from a completed transformation test directly in an OECD 202 and 203 daphnia and fish ecotoxicity test. If the $CaCl_2.2H_2O$ and $MgSO_4.7H_2O$ concentrations of the transformation medium are reduced to one-fifth of the ISO 6341 medium, the completed transformation medium can also be used (upon the addition of micronutrients) in an OECD 201 algae ecotoxicity test.

19. For pH adjustment and buffering down to pH 7 and 6, Table 1 shows the recommended chemical compositions of the media, as well as the CO_2 concentrations in air to be passed through the headspace, and the calculated pH values under these conditions.

TABLE 1

Chemical composition of medium	$NaHCO_3$	6.5 mg/L	12.6 mg/L
	KCl	0.58 mg/L	2.32 mg/L
	$CaCl_2.2H_2O$	29.4 mg/L	117.6 mg/L
	$MgSO_4.7H_2O$	12.3 mg/L	49.2 mg/L
CO_2 concentration (balance is air) in test vessel		0.50%	0.10%
Calculated pH		6.09	7.07

Note: The pH values were calculated using the FACT (Facility for the Analysis of Chemical Thermodynamics) System (http://www.crct.polymtl.ca/fact/fact.htm)

20. Alternative equivalent buffering methods may be used if the influence of the applied buffer on the chemical speciation and transformation rate of the dissolved metal fraction would be minimal.

21. During the full transformation/dissolution tests, agitation should be used which is sufficient to maintain the flow of aqueous medium over the test substance while maintaining the integrity of the surface of the test substance and of any solid reaction product coatings formed during the test. For 1 L of aqueous medium, this may be accomplished by the use of :

- a radial impeller set at 200 r.p.m., with blades deployed 5 cm from the bottom of a 1 L reaction kettle. The radial impellers consist of two fixed polypropylene blades of dimensions 40 mm width x 15 mm height on a PVC-coated steel rod 8 mm diameter and 350 mm long; or
- a 1.0 to 3.0 L flask capped with a rubber stopper and placed on an orbital or laboratory shaker set at 100 r.p.m.

22. Other methods of gentle agitation may be used provided they meet the criteria of surface integrity and homogeneous solution.

23. The choice of solids-liquid separation method depends on whether adsorption of soluble metal ions on filters occurs and whether or not a suspension is generated by the agitation prescribed in paragraph 16, which will in turn depend on particle size distributions and particle density. For solids of density greater than approximately 6 g/cm^3 and particle size ranges as low as 50% < 8 µm, experience has shown that the gentle agitation methods prescribed in paragraph 16 are unlikely to result in suspensions. Hence, filtration of a sample through e.g. a 25 mm diameter 0.2 µm hydrophilic polyethersulphone membrane syringe filter (as an option, overlain by a 0.8 µm prefilter) will result in a solution essentially free of solids. However, in the event that suspensions occur, stopping the agitation to allow the suspension to settle for about 5 minutes prior to taking a solution sample may be useful.

Prerequisites

Analytical method

24. A suitable validated analytical method for the total dissolved metal analysis is essential to the study. The analytical detection limit should be lower than the appropriate chronic or long term value from the exotoxicity tests.

25. The following analytical validation aspects are at a minimum to be reported:

- detection and quantification limit of the analytical method;
- analytical linearity range within the applicable analytical range;
- a blank run consisting of transformation medium (this can be done during the tests);
- matrix effect of the transformation medium on the measurement of the dissolved metal ion;
- mass balance (%) after completion of the transformation test;
- reproducibility of the analysis;
- adsorptive properties of the soluble metal ions on the filters (if filtration is used for the separation of the soluble from the solid metal ion).

Determination of the appropriate pH of the dissolution medium

26. If no relevant literature data exist, a preliminary screening test may need to be carried out in order to ensure that the test is performed at a pH maximising transformation/dissolution within the pH range described in paragraph 8 and 16.

Reproducibility of transformation data

27. For a standard set-up of three replicate test vessels and two replicate samples per test vessel at each sampling time, it is reasonable to anticipate that for a constant loading of a substance, tested in a narrow particle size (e.g., 37 - 44 µm) and total surface area range, the within-vessel variation in transformation data should be less than 10% and the between-vessel variation should be less than 20 % [reference 5].

28. To estimate the reproducibility of the transformation test, some Guidance is given in the following. The results can be used to eventually improve on reproducibility by adjusting the final test set-up through varying the number of replica test vessels and/or replica samples or further screening of the particles. The preliminary tests also allow for a first evaluation of the transformation rate of the tested substance and can be used to establish the sampling frequency.

29. In preparing the transformation/dissolution medium, the pH of the medium should be adjusted to the desired pH (air buffering or CO_2 buffering) by agitation for about half an hour to bring the aqueous medium into equilibrium with the buffering atmosphere. At least three samples (e.g. 10 - 15 mL) are drawn from the test medium prior to addition of the substance, and the dissolved metal concentrations are measured as controls and background.

30. At least five test vessels, containing the metal or metal compound (e.g.100 mg solid/L medium), are agitated as described in paragraph 16 at a temperature $\pm$ 2 °C in the range 20 - 25°C, and triplicate samples are taken by syringe from each test vessel after 24 hours. The solid and solution are separated by membrane filter as described in paragraph 18, the solution is acidified with 1% HNO_3 and analysed for total dissolved metal concentration.

31. The within-test vessel and between-test vessel means and coefficients of variation of the measured dissolved metal concentrations are calculated.

Test performance

a. Dissolution screening test – sparingly soluble metal compounds

32. After dissolution medium is prepared, add the medium into at least three test vessels (number of test vessels depend on the reproducibility obtained during the preliminary test). After a half-hour of agitation to bring the aqueous medium into equilibrium with the atmosphere or buffering system (paragraph 15), the pH, temperature and dissolved O_2 concentrations of the medium are measured. Then at least two 10 - 15 mL samples are taken from the test medium (prior to addition of the solids) and the dissolved metal concentration measured as controls and background.

33. The metal compound is added to the test vessels at a loading of 100 mg/L and the test vessels are covered and agitated rapidly and vigorously. After the 24 hours agitation, the pH, temperature and dissolved O_2 concentrations are measured in each test vessel, and two to three solution samples are drawn by syringe from each test vessel and the solution is passed through a membrane filter as described in paragraph 18 above, acidified (e.g. 1 % HNO3) and analysed for total dissolved metal concentration.

b. Full test - metals and metal compounds

34. Repeat paragraph 32.

35. For 7 day test, substance loadings of 1, 10 and 100 mg/L, respectively, are added to the test vessels (number of which depends on the reproducibility as established in paragraphs 23- 26), containing the aqueous medium. The test vessels are closed and agitated as described in paragraph 16. If a 28 day test is to be conducted, the test with 1 mg/L loading may be extended to 28 days, provided that the same pH value is to be chosen for both 7 day and 28 day tests. However, since 7-day tests are only conducted at pH ranges of 6 and higher, separate 28-day tests are needed to cover the pH range between 5.5 and 6. It may also be useful to include a concurrent control test with no substance loaded (i.e. a blank test solution). At established time intervals (e.g. 2 hours, 6 hours, 1, 4 and 7 days), the temperature, pH and dissolved O_2 concentrations are measured in each test vessel, and at least two samples (e.g. 10 - 15 mL) are drawn by syringe from each test vessel. The solid and dissolved fractions are separated as per paragraph 18 above. The solutions are acidified (e.g. 1 % HNO_3) and analysed for dissolved metal concentration. After the first 24 hours, the solution volumes should be replenished with a volume of fresh dissolution medium equal to that already drawn. Repeat after subsequent samplings. The maximum total volume taken from the test solutions should not exceed 20% of the initial test solution volume. The test can be stopped when three subsequent total dissolved metal concentration data points vary no more than 15%. The maximum duration for the loadings of 10 and 100 mg/L is seven days (the short term test) and 28 days for the loading of 1 mg/L test medium (long term test).

Test Conditions

36. The transformation/dissolution tests should be done at a controlled ambient temperature $\pm$ 2 °C in the range 20 - 25°C.

37. The transformation/dissolution tests are to be carried out within the pH range described in paragraphs 8 and 16. The test solution pH should be recorded at each solution sampling interval. The pH can be expected to remain constant ($\pm$ 0.2 units) during most tests, although some short-term pH variations have been encountered at 100 mg/L loadings of reactive fine powders [7], due to the inherent properties of the substance in the finely divided state.

38. Above the aqueous medium, the head space provided by the reaction vessel should be adequate in most instances to maintain the dissolved oxygen concentration above 70% of its saturation in air, which is about 8.5 mg/L. However, in certain instances, reaction kinetics may be limited not by the availability of molecular oxygen in the head space above the solution but by the transfer of dissolved oxygen to, and removal of reaction product away from, the solid-solution interface. In this case, little can be done, other than await the restoration of equilibrium.

39. To reduce chemical and biological contamination as well as evaporation, the transformation/dissolution kinetics must be performed in closed vessels and in the dark, whenever possible.

TREATMENT OF THE RESULTS

Screening test

40. The mean dissolved metal concentrations at 24 hours are calculated (with confidence intervals).

Full test

a. Determination of the extent of transformation/dissolution

41. The dissolved metal concentrations, measured during the different short term (7 days) tests, are plotted versus time, and the transformation/dissolution kinetics may be determined, if possible. The following kinetic models could be used to describe the transformation/dissolution curves:

(1) Linear model :

$C_t = C_0 + kt$, mg/L
where :
C_0 = initial total dissolved metal concentration (mg/L) at time t = 0;
C_t = total dissolved metal concentration (mg/L) at time t;
k = linear rate constant, mg/L-days.

(2) First order model :

$C_t = A\ (1-e^{(-kt)})$, mg/L
where :
A = limiting dissolved metal concentration (mg/L) at apparent equilibrium = constant;
C_t = total dissolved metal concentration (mg/L) at time t;
k = first order rate constant, 1/days.

(3) Second order model :

$C_t = A\ (1-e^{(-at)}) + B\ (1-e^{(-bt)})$, mg/L
where :
C_t = total dissolved metal concentration (mg/L), at time t;
a = first order rate constant, 1/days;

b = second order rate constant, 1/days;
C = A + B = limiting dissolved metal concentration (mg/L).

(4) Reaction kinetic equation :

$C_t = a[1-e^{-bt} - (c/n)\{1 + (b\ e^{-nt} - n\ e^{-bt})/(n - b)\}]$, mg/L
where :
C_t = total dissolved metal concentration (mg/L) at time t;
a = regression coefficient (mg/L);
b,c,d = regression coefficients (1/days);
n = c+d.

Other reaction kinetic equations may also apply [7,8].

42. For each replicate vessel in the transformation test, these model parameters are to be estimated by regression analyses. The approach avoids possible problems of correlation between successive measurements of the same replicate. The mean values of the coefficients can be compared using standard analysis of variance if at least three replicate test vessel were used. The coefficient of determination, r^2, is estimated as a measure of the "goodness of fit" of the model.

43. The dissolved metal concentrations, measured from the 1 mg/L loading during the 28 day test, are plotted versus time and the transformation/dissolution kinetics determined, if possible, as described in paragraphs 40 and 41.

TEST REPORT

44. The test report should include (but is not limited to) the following information, also see paragraph 11 and 24:

- identification of the sponsor and testing facility;
- description of the tested substance;
- description of the reconstituted test medium and metal loadings;
- test medium buffering system used and validation of the pH used (as per paragraph 21)description of the analytical method;
- detailed descriptions of the test apparatus and procedure;
- preparation of the standard metal solution;
- results of the method validation;
- results from the analyses of metal concentrations, pH, temperature, oxygen;
- dates of tests and analyses at the various time intervals;
- mean dissolved metal concentration at different time intervals (with confidence intervals);
- transformation curves (total dissolved metal as a function of time);
- results from transformation/dissolution kinetics, if determined;
- estimated reaction kinetic quation, if determined;
- deviations from the study plan if any and reasons;
- any circumstances that may have affected the results; and
- reference to the records and raw data.

REFERENCES

1. "Draft Report of the OECD Workshop on Aquatic Toxicity Testing of Sparingly Soluble Metals, Inorganic Metal Compounds and Minerals", Sept. 5-8, 1995, Ottawa.

2. OECD Metals Working Group Meeting, Paris, June 18-19, 1996.

3. European Chemicals Bureau. Meeting on Testing Methods for Metals and Metal Compounds, Ispra, February 17-18, 1997.

4. OECD Metals Working Group Meeting, Paris, October 14-15, 1997.

5. LISEC[5] Staff, "Final report “transformation/dissolution of metals and sparingly soluble metal compounds in aqueous media - zinc", LISEC no. BO-015 (1997).

6. J.M. Skeaff[6] and D. Paktunc, "Development of a Protocol for Measuring the Rate and Extent of Transformations of Metals and Sparingly Soluble Metal Compounds in Aqueous Media. Phase I, Task 1: Study of Agitation Method." Final Report, January 1997. Mining and Mineral Sciences Laboratories Division Report 97-004(CR)/Contract No. 51545.

7. Jim Skeaff and Pierrette King, "Development of a Protocol For Measuring the Rate and Extent of Transformations of Metals and Sparingly Soluble Metal Compounds in Aqueous Media. Phase I, Tasks 3 and 4: Study of pH and of Particle Size/Surface Area.", Final Report, December 1997. Mining and Mineral Sciences Laboratories Division Report 97-071(CR)/Contract No. 51590.

8. Jim Skeaff and Pierrette King, Development of Data on the Reaction Kinetics of Nickel Metal and Nickel Oxide in Aqueous Media for Hazard Identification, Final Report, January 1998. Mining and Mineral Sciences Laboratories Division Report 97-089(CR)/Contract No. 51605.

9. LISEC Staff, "Final report “transformation/dissolution of metals and sparingly soluble metal compounds in aqueous media - zinc oxide", LISEC no. BO-016 (January, 1997).

10. LISEC Staff, "Final report “transformation/dissolution of metals and sparingly soluble metal compounds in aqueous media - cadmium", LISEC no. WE-14-002 (January, 1998).

11. LISEC Staff, "Final report “transformation/dissolution of metals and sparingly soluble metal compounds in aqueous media - cadmium oxide", LISEC no. WE-14-002 (January, 1998).

[5] LISEC, Craenevenne 140, 3600 Genk, Belgium

[6] CANMET, Natural Resources Canada, 555 Booth St., Ottawa, Canada K1A 0G1

BIBLIOGRAPHY

OECD Guideline for testing of chemicals, Paris (1984). Guideline 201 Alga, Growth Inhibition Test.

OECD Guideline for testing of chemicals, Paris (1984). Guideline 202 :Daphnia sp. Acute immobilisation test and Reproduction Test.

OECD Guideline for testing of chemicals, Paris (1992). Guideline 203 : Fish, Acute Toxicity Test.

OECD Guideline for testing of chemicals, Paris (1992). Guideline 204 : Fish, Prolonged Toxicity Test : 14- Day study.

OECD Guideline for testing of chemicals, Paris (1992). Guideline 210 : Fish, Early-Life Stage Toxicity Test.

International standard ISO 6341 (1989 (E)). Determination of the inhibition of the mobility of Daphnia magna Straus (Cladocera, Crustacea).