ECETOC Document No. 36

Comments on OECD Draft Detailed Review Paper: Appraisal of Test Methods for Sex-Hormone Disrupting Chemicals

August 1997

Comments on OECD Draft Detailed Review Paper: Appraisal of Test Methods for Sex-Hormone Disrupting Chemicals

CONTENTS

SUMMARY

The Draft Detailed Review Paper (DDRP): Appraisal of Test Methods for Sex-Hormone Disrupting Chemicals is an OECD Environmental Health and Safety Publication which represents a monumental effort to review a large number of assays and forms the most comprehensive, well-written, and balanced overview of available test methods to date. Its focus is to list, describe and partly evaluate the potential methods that can be used to detect compounds which can act via oestrogenic/anti-oestrogenic and androgenic/anti-androgenic mechanisms. The DDRP clearly states that it does not cover other mechanisms of endocrine disruption (involving other hormonal systems). With a focus on these two endocrine mechanisms, there is an inherent potential risk of providing a false sense of security that enacting screening strategies based on these two endocrine classes will protect human health and the environment.

The DDRP endorses the definition of an endocrine disrupter[1] agreed at the European Workshop and correctly recognises that *in vivo* procedures are required to define potential hazard and ultimate risk. However, the DDRP does not address the fact that "endocrine disruption" is neither an endpoint nor an adverse effect *per se*, but one of many known mechanisms which may or may not result in an adverse effect. This has a major consequence since toxicological and ecotoxicological tests (especially OECD Guideline studies) have been designed to identify hazards and not the underlying mechanisms, e.g., endocrine disruption. Hazard identification through adequately designed *in vivo* studies is necessary for identification of inherent endocrine disruptive potential and along with dose-response data, is fundamental to meaningful risk assessment. Mechanistic data, on the other hand, may be important in defining appropriate experimental design (e.g. timing of dosage, period of observation of effects etc.) of the hazard identification studies.

A key factor in the interpretation of toxicological findings in the area of reproduction/developmental toxicity as well as in assessing endocrine disrupting potential is consideration of the role of general toxicity. It is so crucial for understanding and interpreting results, that it is difficult to think of a compound which would not cause endocrine disruption under some circumstances, namely at high doses in the range of overt (general) toxicity. Thus it would be advisable to include a special Section devoted to "Selection of Dose/Interpretation of Results", because of its importance to the process of obtaining valid and meaningful results.

[1] "An endocrine disrupter is an exogenous substance that causes adverse health effects in an intact organism, or its progeny, consequent to changes in endocrine function". This definition was agreed upon at the European Workshop on the Impact of Endocrine Disrupters on Human Health and Wildlife, Weybridge, UK, Dec 1996 (MRC Institute for Environment and Health, 1996).

Despite the general adequacy of current Guideline studies to detect endocrine disrupting potential, it has been proposed in the DDRP that additional endpoints should be added to the existing protocols. The addition of endpoints should be carefully considered however, since current observations may provide adequate sensitivity which may not be enhanced by the added measurement thus increasing complexity and cost without the benefit of improving detection of potential hazards.

It is recommended that the DDRP addresses separately the specific protocols or tools which define mechanisms from those which assess hazards. *In vitro* assays may contribute to ultimate understanding of mechanism of action, but the potential application of *in vitro* assays needs to be put into context with the limitations noted in the DDRP and with the following: i) effects on reproductive organs are not always mediated by endocrine interaction; and ii) it is questionable if any interaction seen *in vitro* is a surrogate of an *in vivo* hazard. In this respect, *in vitro* assays do not seem suitable for hazard identification. Further, thorough validation of any assay is vital prior to its implementation as a Guideline.

The main focus of testing substances with endocrine-disruptive potential in wildlife organisms is the impairment of reproduction and development ultimately affecting the stability of populations. Therefore, the evaluation of testing procedures should address these effects, while the identification of mechanisms is of lesser importance. Existing, validated and standardised (regulatory) methods and potential enhancements should be evaluated for their suitability to address reproductive and developmental endpoints. Realistic ways of extending these methods to all stages of *in vivo* testing, short-term screening tests to long-term multi-generation studies should be a primary direction of the DDRP.

In conclusion:

- endocrine disruption is a mechanism (not an endpoint or an adverse effect, *per se*);
- the need for protocols or tools for mechanistic studies, in addition to those assessing hazards, warrants further discussion;
- *in vitro* assays are not suitable for hazard identification;
- at present only *in vivo* studies adequately reflect the great variety of subtle interactions and feed-back mechanisms of the endocrine system;
- introducing new procedures or new parameters in standard test protocols will need careful selection and a thorough validation process;
- addition of a special Section devoted to "Selection of Dose/Interpretation of Results", is strongly recommended because of high-dose phenomena and dose-response issues.

KEY TO DDRP PAGINATION

The specific page numbers of *Draft Detailed Review Paper: Appraisal of Test Methods for Sex-Hormone Disrupting Chemicals* differ between the electronic version OECD web page and the OECD printed copy supplied to the member countries representatives and selected reviewers. The citation of a specific page number in these ECETOC comments on the OECD draft Detailed Review Paper are based on the printed distributed copy. The key is presented in the following Table.

Table: Key to DDRP Pagination

GENERAL COMMENTS

1. The Draft Detailed Review Paper (DDRP) evaluates the potential methods that can be used to detect oestrogenic/anti-oestrogenic and androgenic/anti-androgenic compounds. The review of *in vivo* and *in vitro* mammalian tests is clearly the most comprehensive, well-written, and balanced overview of available test methods to date for its stated goal. However, the DDRP does not cover other mechanisms of endocrine disruption (involving other hormonal systems), as clearly stated in the document. The issue of 'environmental oestrogens' originally focused on chemicals which mimic the action of the natural hormone oestrogen. However, the concern is now encompassing effects on the whole endocrine system and the collective term 'endocrine disrupter' is in general use (by scientists, journalists, environmental organisations, and the general public). Currently, the most favoured definition of an endocrine disrupter is "an exogenous substance that causes adverse health effects in an intact organism, or its progeny, consequent to changes in endocrine function". This definition was agreed upon at the European Workshop on the Impact of Endocrine Disrupters on Human Health and Wildlife, Weybridge, U.K, Dec 1996 (MRC Institute for Environment and Health, 1996). Whilst the DDRP covers comprehensively the detection of oestrogenic/androgenic agonistic and antagonistic actions, it does not cover action mediated by other endocrine mechanisms.

   By focusing solely on these two endocrine mechanisms, there are at least two possible consequences. The first is provision of a false sense of security that enacting screening strategies on these two endocrine classes will protect human health and the environment from unintended exposure to endocrine active compounds (EACs). The second is that it will be necessary to expand the number of endpoints in the future to identify not only hazards, but also mechanisms of EACs. Hence, there is a potential for a proliferation of new/revised test methods. This could lead to a complex mixture of studies which cannot be conducted in a cost-effective manner. It should be borne in mind that toxicological tests (especially OECD protocols) have been designed to identify hazards (without the necessity to describe the underlying mechanisms such as, endocrine disruption); hence the emphasis (of the DDRP) should be concentrated on establishing tests which can detect biologically relevant effects caused by endocrine disruption. On the other hand, tests which describe mechanisms may be important in establishing the appropriate experimental design of the studies of effects (e.g. timing of dosage, period of observation of effects).

   It is strongly recommended that any testing strategies proposed should include reference to, and recognition of, these additional endocrine-disrupting mechanisms. An explanation of how these issues will be addressed and ultimately incorporated into the OECD process should be included in the DDRP.

There are many mechanisms by which a compound can act to enhance or attenuate hormone action and ultimately affect reproduction as well as numerous other functions. As indicated in the DDRP, *in vivo* tests are considered necessary to establish potential hazards for hormonally-active compounds using *in vitro* tests to prioritise and help determine mechanism of action. At present, several screening procedures have been proposed (Reel *et al*, 1996; Shelby *et al*, 1996; Carney *et al*, 1997). Another example being developed under the direction of Chemical Manufacturers' Association and Chlorine Chemistry Council is to validate a Tier I comprehensive testing scheme which looks at many types of mechanisms using 15 different types of endocrine-active control substances (O'Connor *et al*, 1996; Cook *et al*, 1997). It is strongly encouraged that the results and potential for implementation of this type of approach be one of the considerations before propagation of new/revised Guidelines.

2. Discussions are in progress on the best approach to detect substances having potential for endocrine activity and on how to assess potential hazards both for humans and wildlife. It is recognized that *in vitro* assays provide a valuable contribution to an ultimate understanding of mechanism of action. Further these assays require a limited amount of time and money. However, at present the application of *in vitro* assays needs to be put in context with the following:

- The currently known *in vitro* assays for oestrogenicity are capable of indicating interaction between the receptor and the chemical. Even though they are well established in pharmacological research for detection of hormonally-effective drugs, they serve, even in the pharmaceutical area, only as prioritisation tools. A more firmly based decision is made on the results of whole-animal experiments.

- As stated in the DDRP, such an interaction is only one (of many known) mechanism which can cause changes in the endocrine system. In addition effects on reproductive organs are not necessarily always mediated by endocrine interaction.

- It is questionable if any interaction seen *in vitro* is a surrogate for an *in vivo* hazard. There are examples of substances showing low potency of receptor interaction but lacking adverse effects on endocrine-sensitive endpoints in appropriate *in vivo* assays. In this respect, *in vitro* assays are not necessarily suitable for hazard identification.

- The DDRP focuses on several further limitations of *in vitro* assays (Section 6.3, page 104 of the print version) including the following:

  - the interaction with the receptor construct or response elements may not mimic *in vivo* modes of action;

  - *in vitro* assays may be unable to distinguish agonists from antagonists;

- existing *in vitro* assays lack satisfactory metabolic activity (some yeast-based assays are not metabolically inert impeding interpretation of the results);
- some *in vitro* assays show limited chemical uptake;
- absorption, metabolism, excretion or bioaccumulation may play critical roles in determining activity.

In vivo assays, especially studies with repeated exposure, are not compromised by such limitations. At present, only such *in vivo* studies adequately reflect the subtle interactions and feed-back mechanisms of the endocrine system. Conducting repeat-exposure *in vivo* studies at different dose levels not only affords the opportunity to identify hazards but also to characterise the dose response. These are prerequisites for a meaningful risk assessment. Applying the current risk assessment paradigm on the results of such studies is the best known approach for adequate protection of humans and wildlife (Stevens *et al,* 1997a; Crisp *et al*, 1997).

NOTE: It should be recognised that currently there are questions regarding the availability of the technology for transfection assays in general. Specifically, Ligand Pharmaceuticals holds a patent on the technology which could make general implementation difficult. Prior to the availability of any Guidelines, these issues need to be resolved.

3. There are two large Sections of the DDRP which cover the review of the non-regulatory individual test methods - these are Chapter 5 "Critical Assessment of Non-Regulatory Test Methodologies" and Chapter 8 "Annexes". In Chapter 5, non-regulatory methods are reviewed, followed by a tabulated listing of methods and Sections in which certain methods are recommended. This structure is difficult to read and the basis for selecting and recommending certain methods or assays is not explained in sufficient detail. It is also confusing as some key references, which would influence the critique, are mentioned in the Annexes but not in the critical review. The tables list a number of procedures without any further qualification or comment on their usefulness. It would make it significantly clearer to draw conclusions from the vast amount of information available if these Chapters were combined. For example, for each test system, the details given in Chapter 8 could be followed by the critical review and recommendations (as given in Chapter 5). Further, the Annexes (Sections 8.1.5 and 8.2.6) are not comprehensive.

4. The terms "human relevant" and "wildlife relevant" (page 40, Section 5.1) are confusing and almost suggest that there should be Sections on "human non-relevant" etc. which of course is incorrect.

In addition, there is inconsistency in Section names which are termed "Systems of Relevance to Assessment of Human Toxic Potential" etc. later in the document (page 129, Section 8.1.1).

The authors' rationale for classifying a test method as relevant is not explained, and it would appear that none of the tests is relevant to anything but potential hazard assessment. In addition, we are not really dealing with "toxic potential" when referring to endocrine disruption. Identification and regulation of a genuine chemical hazard should depend on the observation of an adverse toxicological effect and not just on observation of mechanistic changes. For example, small perturbations in a hormone level in a study where no effects are observed in reproduction or chronic toxicity studies must be of questionable significance. Where possible the authors should clearly indicate the criteria upon which they base the statement relevant to human or wildlife.

5. It is suggested that a special Section be devoted to Selection of Dose/Interpretation of Results because of its importance to the process of obtaining valid and meaningful results (Foran *et al*, 1997) .

It is evident that *in vitro* and *in vivo* screening assays are being used without regard to appropriate or logical doses. There are examples in the literature of assays being used to measure the effects of oestradiol and xenobiotics at concentrations ranging over 8 or more orders of magnitude (Soto *et al,* 1995, Jobling and Sumpter, 1993). There is, however no effort to determine if the assay or the endpoint respond appropriately to high concentrations of chemicals or that the defined mechanism is functional at these high concentrations. Further, there is no effort to determine realistic concentrations over which the assay is valuable in identiying the potential hazard of a test material. In addition, Ren *et al* (1996) indicated that mRNA for vitellogenin in fish was induced by dimethyl formamide and ethylene glycol. It has been clearly demonstrated that ethylene glycol does not bind to the oestrogen receptor *in vitro*. The confounding effects produced by these solvents must be considered in selecting doses.

6. For this discussion, it needs to be stated that as more endpoints are added to standard tests, more random/spurious findings will be obtained. Interpretation and conclusions based on these findings must be treated with caution. It is likely that, except in cases of obvious and extreme effect, observation of related effects will be required to justify the conclusion of endocrine disruption. Therefore, these random/spurious findings will result in a large number of equivocal tests. Decisions on how these will be resolved are critical prior to embarking on a panoply of additions to standard test procedures.

7. A key factor in the interpretation of toxicological findings in the area of reproduction/ developmental toxicity is consideration of the role of general toxicity. The potential for this to be the explanation of effects (rather than a subtle or specific effect on the endocrine system) is real and should be brought out more in the DDRP. There is a mention of toxicity on page 99, paragraph 2, but its key role in attributing correct interpretation to findings should be stressed in other Sections that deal with the *in vivo* models.

8. The DDRP has avoided defining which *in vivo* endpoints definitely establish the sex-hormone disruption of a substance (Stevens *et al*, 1997b). These endpoints are essential for validating possible screening approaches. The DDRP presents numerous possible tests. However, it appears in the discussion that the establishment of sensitivity, relevance and relative importance of possible marker endpoints has been added as an afterthought. These features must be addressed with major emphasis, particularly after the unfortunate episode concerning synergy (McLachlan, 1997). The same lack of definition or identification applies to the need for a reference set of chemicals and dose-response profiles.

9. Parts of the environmental and wildlife discussions need to be considered for re-drafting as it is often difficult to follow what is being proposed by the authors. It is recommended that the draft be edited to improve significantly the ease with which it can be read. Also, the Executive Summary does not give sufficient attention to the wildlife issue given the scope of the full document and therefore should be amended to correct this imbalance. The definition given on page 15 should also be moved to the Executive Summary.

   There should be a stronger emphasis in the Executive Summary and Introduction on the ecological context of sex-hormone disrupting chemicals in wildlife. Although it is referenced on page 7 as an important consideration, this needs to be more prominent. It should be pointed out that the environmental hazard assessment for such hormone disrupters should be treated in the same way as any other agent that can cause adverse reproductive effects, regardless of the mechanism involved. As is true of other areas of wildlife (eco) toxicology, this Introduction should act as a signpost in ensuring that mechanistic studies are best targeted towards providing information that helps in ecosystem management, rather than such mechanistic data being seen as the final endpoint in their own right for environmental risk assessment.

   An additional concern is that much of the relevant ecotoxicology literature is not cited in the DDRP. There are many publications on reproductive toxicity in diverse wildlife species that should be considered in such a review. Specifically, a good deal of research work published by the US EPA over many years and also the standard methods for fish and invertebrates published by the US EPA and ASTM has not been included. Moreover scant attention has been given to

many important papers published on the comparative endocrinology of wildlife species used in the OECD testing scheme. For the papers that have been included, many appear as uncritical listings which do not make best use of the information. The authors are advised to refer to the report from the recent SETAC/OECD meeting on *Endocrine Modulators in Wildlife - Assessment and Testing* (EMWAT workshop), as well as the ECETOC Document 33 "Compendium of Test Methods (1996).

The authors of the DDRP should be wary of specifying tests using individual wildlife species that are not part of the OECD scheme, e.g., mosquito fish (Table 4 and 5). Such a recommendation does not reflect the current "state of the science" in either wildlife endocrine research or ecotoxicology. These tables should be supplemented with footnotes to indicate that, in principle, other species can be used for the same purposes.

10. In Chapter 5, non-regulatory methods are reviewed, followed by a tabulated listing of methods and a Section in which certain methods are recommended. This structure is very difficult to read and the basis for selecting and recommending certain methods or assays is not explained in sufficient detail. The tables list a number of procedures without any further qualification or comment on their usefulness. The different Sections of this Chapter should be consolidated.

Note: ECETOC together with EMSG has developed a research programme addressing existing data gaps related to specific testing approaches for wildlife organisms. ECETOC is happy to cooperate with OECD in providing further literature references and to update the Organisation about the on-going research activities being conducted by the chemical industry.

SPECIFIC COMMENTS

In this Section, detailed comments are listed under the chapter headings (italicised) employed in the DDRP.

1. EXECUTIVE SUMMARY

1.1 MODIFICATIONS TO EXISTING TEST METHODS

Page 5, Section 1.1: Although it is agreed that the most promising approach seems to be to make full use of the existing study Guidelines with inclusion of appropriate enhancements as necessary, it must be pointed out that the merit of adding some of the possible enhancements has yet to be established.

As the sensitivities of these endpoints need to be established, it would appear inappropriate to add these endpoints in test systems without first establishing (relative) sensitivity and then proposing possible extension of existing test-Guidelines.

Further, the modification to existing test methods is presented as if applicable for all (or most) reviewed OECD tests. It is very doubtful that it is desirable to include these modifications in all tests.

Since the added value of some of the modifications have not been clearly established perhaps the modification should only be made to a limited number of the most relevant OECD tests and then only on a trial basis. Because a base set of information is required in Europe, the best approach is to have the proposed extra parameters included in only one of the required tests, most logically a 28-day type study such as a 'modified OECD 407' Guideline.

1.2 NON-REGULATORY TEST MODELS PROPOSED FOR FURTHER DEVELOPMENT/ ADOPTION

Page 6, Section 1.2, third bullet: As written, the castrated rat model appears to be focused on detecting androgens. It should clearly be designed to detect androgen-receptor antagonists. If the castrated model is to be implanted with testosterone implants O'Connor *et al* (1997) recommend using silastic tubing rather than pellets.

Page 6, Section 1.2: Strongly agree that at present it is not possible to recommend adoption of any of the *in vitro* assays as a regulatory acceptable model because of the various limitations and difficulties inherent in current designs.