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

12-Sep-2006

English - Or. English

ENV/JM/MONO(2006)27
Unclassified

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

SERIES ON TESTING AND ASSESSMENT
Number 60

REPORT OF THE INITIAL WORK TOWARDS THE VALIDATION OF THE 21-DAY FISH SCREENING ASSAY FOR THE DETECTION OF ENDOCRINE ACTIVE SUBSTANCES (PHASE 1A)

JT03213435

English - Or. English

OECD Environment Health and Safety Publications

Series on Testing and Assessment

No. 60

REPORT OF THE INITIAL WORK TOWARDS THE VALIDATION OF THE 21-DAY FISH SCREENING ASSAY FOR THE DETECTION OF ENDOCRINE ACTIVE SUBSTANCES

(PHASE 1A)

Environment Directorate
ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Paris

September 2006

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, revised 2006)*

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, *Guidance 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, *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 approval)*

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 (2004)*

No 29, *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)

No. 33, *Harmonised Integrated Classification System for Human Health and Environmental Hazards of Chemical Substances and Mixtures (2001)*

No. 34, *Guidance Document on the Development, Validation and Regulatory Acceptance of New and Updated Internationally Acceptable Test Methods in Hazard Assessment (in preparation)*

No. 35, *Guidance notes for analysis and evaluation of chronic toxicity and carcinogenicity studies (2002)*

No. 36, *Report of the OECD/UNEP Workshop on the use of Multimedia Models for estimating overall Environmental Persistence and long range Transport in the context of PBTS/POPS Assessment (2002)*

No. 37, *Detailed Review Document on Classification Systems for Substances Which Pose an Aspiration Hazard (2002)*

No. 38, *Detailed Background Review of the Uterotrophic Assay Summary of the Available Literature in Support of the Project of the OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA) to Standardise and Validate the Uterotrophic Assay (2003)*

No. 39, *Guidance Document on Acute Inhalation Toxicity Testing (in preparation)*

No. 40, *Detailed Review Document on Classification in OECD Member Countries of Substances and Mixtures Which Cause Respiratory Tract Irritation and Corrosion (2003)*

No. 41, *Detailed Review Document on Classification in OECD Member Countries of Substances and Mixtures which in Contact with Water Release Toxic Gases (2003)*

No. 42, *Guidance Document on Reporting Summary Information on Environmental, Occupational and Consumer Exposure (2003)*

No. 43, *Draft Guidance Document on Reproductive Toxicity Testing and Assessment (in preparation)*

No. 44, *Description of Selected Key Generic Terms Used in Chemical Hazard/Risk Assessment (2003)*

No. 45, *Guidance Document on the Use of Multimedia Models for Estimating Overall Environmental Persistence and Long-range Transport (2004)*

No. 46, *Detailed Review Paper on Amphibian Metamorphosis*

Assay for the Detection of Thyroid Active Substances (2004)

No. 47, *Detailed Review Paper on Fish Screening Assays for the Detection of Endocrine Active Substances (2004)*

No. 48, New Chemical Assessment Comparisons and Implications for Work sharing (2004)

No. 49, Report from the Expert Group on (Q)SARs on the Principles for the Validation of (Q)SARs (2004)

No. 50*, Report of the OECD/IPCS Workshop on Toxicogenomics (2005)*

No. 51*, Approaches to Exposure Assessment in OECD Member Countries: Report from the Policy Dialogue on Exposure Assessment in June 2005 (2006)*

No. 52*, Comparison of emission estimation methods used in Pollutant Release and Transfer Registers (PRTRs) and Emission Scenario Documents (ESDs): Case study of pulp and paper and textile sectors (2006)*

No. 53*, Guidance Document on Simulated Freshwater Lentic Field Tests (Outdoor Microcosms and Mesocosms) (2006)*

No. 54*, Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to Application (2006)*

No. 55*, Detailed Review Paper on Aquatic Arthropods in Life Cycle Toxicity Tests with an Emphasis on Developmental, Reproductive and Endocrine Disruptive Effects (2006)*

No. 56, *Guidance Document on the Breakdown of Organic Matter in Litter Bags (2006)*

No. 57, *Detailed Review Paper on Thyroid Hormone Disruption Assays (2006)*

No. 58, *Report on the Regulatory Uses and Applications in OECD Member Countries of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models in the Assessment of New and Existing Chemicals (2006)*

No. 59, *Report of the Validation of the Updated Test Guideline 407: Repeat Dose 28-Day Oral Toxicity Study in Laboratory Rats (2006)*

No. 60, *Report of the Initial Work Towards the Validation of the 21-Day Fish Screening Assay for the Detection of Endocrine Active Substances (Phase 1A) (2006)*

No. 61, *Report of the Validation of the 21-Day Fish Screening Assay fort he Detection of Endocrine Active Substances (Phase 1B) (2006)*

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 30 industrialised countries in North America, Europe and the Asia and Pacific region, 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 working 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 working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in ten different series: **Testing and Assessment**; **Good Laboratory Practice and Compliance Monitoring**; **Pesticides and Biocides**; **Risk Management**; **Harmonisation of Regulatory Oversight in Biotechnology**; **Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; and the Safety of Manufactured Nanomaterials.** More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (http://www.oecd.org/ehs/).

This publication was produced within the framework of the Inter-Organisation Programme for the Sound Management of Chemicals (IOMC).

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 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. The participating organisations are FAO, ILO, OECD, UNEP, UNIDO, UNITAR and WHO. The World Bank and UNDP are observers. 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.

OECD Environment Directorate,
Environment, Health and Safety Division
2 rue André-Pascal
75775 Paris Cedex 16
France

Fax: (33-1) 44 30 61 80
E-mail: ehscont@oecd.org

FOREWORD

This document is the report of the OECD Phase 1A study to validate the 21-day Fish Screening Assay for the detection of some endocrine active substances. The Phase 1A was an investigation of the feasibility of a protocol using adult male and female fish for the detection of strong estrogen and androgen; to that end, three core endpoints were measured and subsequently evaluated in this report. Three laboratories agreed to take part in this study and performed the assay with two or three of the fish species proposed. Each laboratory used the two test chemicals: 17beta-estradiol and 17beta-trenbolone. The test substances were distributed from a central chemical repository, situated at the United States Environmental Protection Agency. A total of twenty experiments took place between January and May 2003 for this Phase 1A.

The proposed studies, the model protocol, the test substances and the concentrations were agreed by the Fish Drafting Group and approved by the Validation Management Group for Ecotoxicity Testing (VMG-eco). Each of the three laboratories submitted the raw data to the lead laboratory for analysis and to the OECD Secretariat for archiving. Dr. Masanori Seki from the lead laboratory (CERI, Japan) prepared the initial draft of the report, and checked the accuracy of the data with the participating laboratories. Comments and input to the draft report were also contributed from the participating laboratories and from members of the VMG.

The VMG approved the draft report in February 2004. The Task Force on Endocrine Disrupters Testing and Assessment (EDTA) then endorsed the Phase 1A report, together with the Phase 1B report which constitutes a separate document, in January 2006. At its 18th Meeting, the Working Group of the National Coordinators of the Test Guidelines Programme agreed to the submission of the report to the Joint Meeting with a view to its declassification. This document is published on the responsibility of the Joint Meeting of the Chemicals Committee and Working Party on Chemicals, Pesticides and Biotechnology.

Contact for further details

Environment, Health and Safety Division
Environment Directorate
Organisation for Economic Cooperation and Development
2, rue André-Pascal
75775 Paris Cédex 16, France

Tel : +33 1 45 24 1674 or 9843
Fax : +33 1 45 24 1675
Email : env.edcontact@oecd.org

TABLE OF CONTENTS

FOREWARD 11
GLOSSARY 14
SUMMARY 15
INTRODUCTION 17
PURPOSE OF THE FISH SCREENING ASSAY AND OBJECTIVES OF PHASE 1 OF THE VALIDATION 17
Purpose and Major Characteristics of the Fish Screening Assay 17
Objective of the validation study 18
Fish Species 18
Positive control chemicals and Sampling Times in Phase 1A 18
Summary of Key Goals of Phase 1A 18
ACTION PLAN AND ORGANISATION OF THE VALIDATION PROJECT 19
METHOD 20
Life-stage 20
Preferred species 20
Endpoints studied 20
Route of exposure 20
Test concentrations 20
Housing of male and female fish 21
Test conditions 21
Analysis of data 21
RESULTS 22
Measured test concentrations 22
Statistical approach followed 22
17 β- estradiol study 25
Vitellogenin (VTG) 25
Gonado-somatic index (GSI) 28
Secondary sex characteristics (number of papillary processes in medaka, number and index of nuptial tubercles in fathead minnow) 29
Gonad histology 29
17 β-trenbolone study 32
Vitellogenin (VTG) 32
Gonado-somatic index (GSI) 35
Secondary sex characteristics (number of papillary processes in medaka, number and index of nuptial tubercles in fathead minnow) 36
Gonad histology 37
Overall findings 40
DISCUSSION 42
RECOMMENDATIONS 47
ACKNOWLEDGEMENT 48
ANNEXURES 49
ANNEX 1: PROTOCOL FOR PHASE 1A OF THE VALIDATION 50
REFERENCES 58

LIST OF FIGURES

Figure 1: Statistic flow-chart 1
Figure 2: Statistic flow-chart 2
Figure 3a: VTG in males fathead minnow at d-14
Figure 3b: VTG in females fathead minnow at d-14
Figure 3c: VTG in males fathead minnow at d-21
Figure 3d: VTG in females fathead minnow at d-21
Figure 4a: VTG in males medaka at d-14
Figure 4b: VTG in femals medaka at d-14
Figure 4c: VTG in males medaka at d-21
Figure 4d: VTG in females medaka at d-21
Figure 5a: VTG in males zebrafish at d-14
Figure 5b: VTG in females zebrafish at d-14
Figure 5c: VTG in males zebrafish at d-21
Figure 5d: VTG in females zebrafish at d-21
Figure 6a: GSI in males fathead minnow at d-21
Figure 6b: GSI in females fathead minnow at d-21
Figure 6c: GSI in males medaka at d-21
Figure 6d: GSI in females medaka at d-21
Figure 6e: GSI in males zebrafish at d-21
Figure 6f: GSI in females zebrafish at d-21
Figure 7a: VTG in males fathead minnow at d-14
Figure 7b: VTG in females fathead minnow at d-14
Figure 7c: VTG in males fathead minnow at d-21
Figure 7d: VTG in females fathead minnow at d-21
Figure 8a: VTG in males medaka at d-14
Figure 8b: VTG in females medaka at d-14
Figure 8c: VTG in males medaka at d-21
Figure 8d: VTG in females medaka at d-21
Figure 9a: VTG in males zebrafish d-14
Figure 9b: VTG in females zebrafish at d-14
Figure 9c: VTG in males zebrafish at d-21
Figure 9d: VTG in females zebrafish at d-21
Figure 10a: GSI in males fathead minnow at d-21
Figure 10b: GSI in females fathead minnow at d-21
Figure 10c: GSI in males medaka at d-21
Figure 10d: GSI in females medaka at d-21
Figure 10e: GSI in males zebrafish at d-21
Figure 10f: GSI in females zebrafish at d-21
Figure 11a: Number of nuptial tubercles in females fathead minnow after 14 days exposure
Figure 11b: Number of nuptial tubercles in females fathead minnow after 21 days exposure
Figure 12a: Number of papillary processes in females medaka after 14 days exposure
Figure 12b: Number of papillary processes in females medaka after 21 days exposure

LIST OF TABLES

Table 1: List of participating laboratories
Table 2: Measured test concentrations of 17-β estradiol
Table 3: Measured test concentrations of 17-β trenbolone
Table 4: Summary of the study with 17 β-estradiol
Table 5: Summary of the study with 17 β-trenbolone
Table 6: Summary of the evaluation of the gonad histopathology after 17 β-estradiol exposure in all laboratories and for all fish species
Table 7: Summary of the evaluation of the gonad histopathology after 17 β-trenbolone exposure in all laboratories for all three fish species.

GLOSSARY

Relevance: description of the relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect and species of interest.

Reproducibility: the agreement among results obtained from testing the same substance using the same test protocol. (see Reliability).

Reliability: extent of reproducibility of results from a test over time within and among laboratories, when performed using the same protocol**.**

Robust(ness): the insensitivity of a test to departure from the specified test conditions. The ability of a test to provide similar results over a range of test conditions under which the test may be used in different laboratories.

Transferability: the ability of a test method or procedure to be accurately and reliably performed in independent, competent laboratories.

SUMMARY

i) This report provides the results from an OECD inter-laboratory study conducted in 2003 to examine the relevance and transferability of a standardized OECD protocol for the fish screening assay. This work is considered the first step in a process to validate at the international level the fish screening assay for the detection of endocrine active substances.

ii) The need to develop and validate the fish screening assay originates from the concerns that environmental levels of chemicals may be causing adverse effects in both humans and wildlife due to the interaction of these chemicals with the endocrine system. Several cases were reported where high level exposures have resulted in effects in wildlife, and in particular in fish. In 1997, the OECD concluded that existing test methods were insufficient to identify such effects. As part of the OECD Test Guidelines Programme a *Special Activity on the Testing and Assessment of Endocrine Disrupters* was initiated to revise existing, and develop new OECD Test Guidelines for the testing of potential endocrine disrupters. A Task Force on Endocrine Disrupters Testing and Assessment was subsequently established to provide a focal point within OECD to consider and recommend priorities for the development of testing methods for endocrine disrupters.

iii) Although the need for that assay was identified earlier in the scope of the EDTA Task force, existing Test Guidelines were not suitable for that purpose. Therefore, international fish experts met in two occasions in London in 1998 and in Tokyo in 2000 to review the state-of -the-art in the area of fish screening and fish testing in relation to endocrine disrupters. The outcome of these Expert Consultations was an agreement on three promising core endpoints that should be part of the future fish screening assay. These endpoints were the vitellogenin level, the gross morphology (including secondary sex characteristics and the gonado-somatic index) and gonad histology. However, no protocol was yet existing that encompassed these core endpoints. It is only in 2001, during the First Meeting of the Validation Management Group for Ecotoxicity Testing (VMG-eco) that proposals were made for candidate protocols.

iv) A Fish Drafting Group was established to follow-up on the elaboration of a protocol and was responsible of preparing a proposal to the VMG-eco for validation of the test method selected. One major constraint in the task was that the protocol should fit the three fish species commonly used in regulatory testing: fathead minnow, medaka and zebrafish. In December 2002, the Secretariat transmitted the proposal to the VMG-eco. The protocol proposed exposes adult fish for a period of 14 and 21 days. Fish are sampled at 14 days and at 21 days where core endpoints are measured. Males and females are analysed separately.

v) In Phase 1A, two test compounds were used: 17 β-estradiol as an estrogen agonist (CAS. No. 50-28-2) and 17 β-trenbolone as an androgen agonist (CAS No. 10161-33-8). In order to have a sufficient number of repeats of tests by fish species across laboratories, all laboratories were required to use at least two fish species and both test substances. In the end, zebrafish was tested in all four laboratories, and fathead minnow and medaka in three laboratories. Data were collected from one laboratory in Germany and three laboratories in Japan.

vi) At 14 days and 21 days after exposure, almost all laboratories were successful in detecting an increase in vitellogenin production in males following 17 β-estradiol exposure in all species. There were differences in the lowest observed effect concentration among the laboratories, but in average, all laboratories were able to detect a difference at 32 ng/l at 14-days.

vii) At 14 days and 21 days after exposure, almost all laboratories were successful in detecting a decrease in vitellogenin production in females following 17 β-trenbolone exposure in all species. There were differences in the lowest observed effect concentration (50 ng/l in fathead minnow and 500 ng/l in

zebrafish and medaka at 14-days).

viii) The gonado-somatic index poorly responded to both trenbolone and estradiol in both sexes at 14 and 21 days after exposure. Whenever a significant level was observed, the trends were not always consistent across laboratories.

ix) Secondary sex characteristics, which are responsive to endocrine active substances, are measurable in fathead minnow and medaka. Following trenbolone exposure, the number of papillary processes in medaka increased significantly after trenbolone exposure in females at both 14 and 21 days. Nuptial tubercles number in female fathead minnow was slightly affected after trenbolone exposure.

x) Histology evaluation demonstrated that gonads were responsive to both androgenic and estrogenic exposures in most of the laboratories (LAB 4, LAB 1 and LAB 3). However, one laboratory (LAB 2) never found any change, except for hermaphroditic gonad, induced by either of the chemical. Evaluation criteria were not harmonized at this early stage of endpoint standardisation. Reproducibility of findings was not checked *sensu stricto*, but observations tended to show a serious lack of consistency.

xi) The feasibility of the assay in detecting changes caused by 17-βestradiol, an estrogen agonist and 17-β trenbolone, an androgen agonist was successfully demonstrated through respectively an increase in males and a decrease in females in vitellogenin production.

xii) A number of protocol refinements and additional guidance will be necessary to improve the performance of the assay before starting Phase 1B, where a few compounds with different modes of action will be added. These refinements include standardization of vitellogenin measurement (through blood or plasma and liver sampling procedures and harmonization of measurement method); allotment of fish in test vessels in a biological relevant way; standard operating procedures for secondary sex characteristics and gonad fixation and sectioning; improvement of the guidance for the evaluation of gonad histology with the inclusion of evaluation criteria; and the definition of an appropriate sample size or replication design.

xiii) Overall, it can be concluded from Phase 1A that the current assay is relevant and reproducible for the detection of strong estrogenic and androgenic compounds. Separate analysis of females and males for vitellogenin measurement as indicators of estrogenic exposure (and maybe other modes of action), secondary sex characteristics in medaka and fathead minnow as indicators of androgenic exposure and probably also gonad histology are relevant in this assay.

INTRODUCTION

1. A conceptual framework was developed already at the first meeting of the EDTA Task Force in 1998. This framework was aiming at the identification of short-term screening assays and long-term tests of increasing complexity to gather information on potential endocrine disrupters. The range of assays and tests spans from (quantitative) structure activity relationships, *in vitro* assays (e.g. receptor binding assays), short-term *in vivo* assays to longer-term assays involving different life-stages of the animal.

2. This framework was amended several times and it is considered as a flexible toolbox where, depending on the needs, tools can be picked up at different levels to provide the information needed. The fish screening assay, once validated will be added to the framework as a short-term *in vivo* assay.

3. The fish screening assay is based on the principle that the egg-yolk precursor protein, called vitellogenin, normally produced by female fish, is also produced by male exposed to estrogenic substances. In addition, other sex hormone sensitive parameters are incorporated in the assay (secondary sex characteristics, the gonado-somatic index and gonad histology) as potential indicators of exposure to estrogen (ant-) agonists and androgen (ant-) agonists. A separate analysis of males and females is possible in adults and enables sex dependant features to be interpreted for each sex.

4. The aim of the validation of the fish screening assay is to develop a robust, relevant and reliable test method for the detection of chemicals that have the potential to act as (anti-)androgens or (anti-) estrogens and interfere with males and females sex hormone systems. It is also the purpose of the validation to understand the possibilities for application of the assay, and any limitations to its use.

PURPOSE OF THE FISH SCREENING ASSAY AND OBJECTIVES OF PHASE 1 OF THE VALIDATION

Purpose and Major Characteristics of the Fish Screening Assay

5. The *in vivo* fish screening assay is intended to detect chemicals that affect androgenic or estrogenic activity in adult fish exposed during a limited part of their life-cycle. The assay is not designed to identify specific mechanisms of hormonal disruption (such as hormone receptor agonistic or antagonistic effects, or disturbance of hormone synthesis or metabolism). The use of adult fish allows simultaneous measurements of core endpoints in each individual, thus reducing the total number of animals needed in the assay.

6. The fish screening assay addresses three core biological endpoints: i) gross morphology, including secondary sexual characteristics and gonado-somatic index (GSI); ii) plasma or liver vitellogenin levels; and iii) histopathology of gonads.

7. With respect to vitellogenin determinations, measurements are species-specific and could be done either in plasma or in liver homogenates. Works conducted in the United States on the fathead minnow, the medaka and the zebrafish, and in France on the zebrafish were completed at the same time as Phase 1A of the Validation. All results, including Phase 1A of the Validation, were discussed in a meeting of fish experts in October 2003 resulting in recommendations to use homologous ELISA methods (homologous VTG standard and antibodies) with demonstrated detection capabilities in the order of ng/ml. Additional guidance on vitellogenin measurement, including standard operating procedures for sample preparation and details of the method used, will be provided for Phase 1B of the Validation.. In Phase 1A, laboratories had the choice of the ELISA methods they used for all three species. In addition, it was recognized that standard operating procedures would facilitate a harmonized evaluation of secondary sex characteristics in medaka and fathead minnow in future validation exercise. One of the most challenging endpoint in the

assay is the evaluation of gonadal histopathology. Experience gained in Phase 1A showed that a guidance document building consensus on the evaluation of gonad histology (gonad staging and on severity scoring systems) is highly needed. A draft of the guidance document was available on time for the start of Phase 1B. This helped to reduce a subjective interpretation, leading to observer to observer differences.

Objective of the validation study

8. The objective of the validation study is to experimentally establish the relevance of the assay, when conducted in a choice of three generally used OECD fish species, to detect weak and strong estrogenic and androgenic effects of chemicals and to assess the reliability (reproducibility) of the assay by comparing test results obtained by a variety of laboratories in geographical diverse locations. Robustness, relevance and reliability have been defined in the draft OECD Guidance Document no. 34, based on the Solna Workshop Report on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. Definitions can also be found in the Glossary of this Report.

9. The specific objective of the Phase 1A of the validation study is to experimentally assess the feasibility of a test design, developed by an international group of experts that is built of components which are considered sensitive to endocrine disruption. The aim of phase 1A is to obtain initial information on: i) the relevance of the endpoints used in terms of their sensitivity and (species) specificity and ii) the reproducibility of test results in a limited number of laboratories. To that end, Phase 1 has been divided in Phase 1A and Phase 1B. In Phase 1A, two known endocrine disrupting chemicals (a strong estrogen and a strong androgen) were tested to check the feasibility, relevance of endpoints and to enable necessary refinements to the protocol before Phase 1B. In Phase 1B, three weakly active endocrine disrupting chemicals are used to check whether the assay is able to detect weak compounds as well. A Phase 2 will include negative substances for an evaluation of the specificity of the assay (i.e. the possibility to have false positive outcomes).

Fish Species

10. Three species are used in the validation study: fathead minnow (*Pimephales promelas*), Japanese medaka (*Oryzias latipes*) and zebrafish (*Danio rerio*).

Positive control chemicals and Sampling Times in Phase 1A

11. Phase 1A includes the use of two positive test substances: the estrogen 17β-estradiol (CAS number 50-28-2) and the androgen 17β-trenbolone (CAS number 10161-33-8).

12. In Phase 1A of the validation, two sampling points are defined (at day-14 and at day-21 after exposure via the water), in recognition that certain effects may take more than 14 days of exposure to be detectable (e.g. change in gonad histology).

Summary of Key Goals of Phase 1A

- Demonstrate the overall feasibility of the fish endocrine screening assay test design;
- Obtain initial information on the relevance of endpoints and in particular their ability to respond to 17β-estradiol and 17β-trenbolone exposures in the context of this screening assay.
- Collect a set of data for two reference test substances;
- Obtain initial information on possible differences in species sensitivity to androgen and estrogen

agonist activity;

- Obtain information on adequate exposure durations to observe significant changes of endpoints in the three fish species;

- Collect sufficient information to enable protocol optimization and other refinements, especially as regards gonad histology procedures; VTG measurement methods and optimization of the number of fish used.

ACTION PLAN AND ORGANISATION OF THE VALIDATION PROJECT

13. Upon agreement by the Fish Drafting Group on a preferred protocol for the fish screening assay, expressions of interest were sought from laboratories to participate in the practical validation work. They were explained that Phase 1 was divided in two steps, the first step (Phase 1A) with only a limited number of laboratory (3 to 4) for an evaluation of the feasibility of the protocol and the second step (Phase 1B) with possibly a few more laboratories once the feasibility and transferability of the protocol would be ascertained. A selection had to be made for Phase 1A and three Japanese laboratories and a German laboratory conducted the assay. Table 1 presents the list of laboratories.

14. It was required from the laboratories that each of them uses at least two of the three fish species. This was important in order to ensure comparability of data across laboratories. One laboratory was identified as the lead co-ordinating laboratory for the validation study; it was LAB 1 in Japan. LAB 1 co-ordinated the practical details of the protocol, the exchange of technical advice, the collection of results, the statistical analysis and the preparation of a preliminary report. The assay was conducted under full GLP compliance in LAB 4, but not in the other three laboratories.

15. The OECD Secretariat co-ordinated the overall project, assisted in the preparation of Excel spreadsheets, in the co-ordination of the supply of test compounds to the laboratories and in the preparation of the report. Due to restrictions in import and export of the androgenic substance, 17 β-trenbolone, it was not possible for the laboratories to all use substance from the same lot number. The estrogen 17 β-estradiol was supplied from a single lot by Battelle in the United States who acted as a chemical repository.

16. Only LAB 4 in Germany wrote detailed experimental procedures, based on the common agreed OECD protocol, in compliance with GLP, and submitted it to the OECD Secretariat prior to the start of the study**.**

17. The experimental work started shortly after approval of the study design by the VMG-eco in March 2003. The in-life part of Phase 1A work was completed at the end of June 2003. Provisional results were made available to the 2nd meeting of the VMG-eco on 2nd May 2003, but the quantity of information was still limited at that time to enable a comprehensive discussion..

Laboratory	Country
LAB 1	Japan
LAB 2	Japan
LAB 3	Japan
LAB 4	Germany

Table 1: list of laboratories

METHOD

18. Historically, there is no internationally agreed test method to detect endocrine active chemicals in fish. Three proposals were made in 2001 at the First Meeting of the VMG-eco. One method including reproduction was developed in the United States on fathead minnow. The second method was the non-spawning fish screening assay, developed on fathead minnow in Europe by a co-operation of six industry and university laboratories. The third one was the 40-d Zebrafish assay, developed by the Nordic countries, and which will further be developed into a fish sexual development test. Extensive discussions took place to compare the methods, which all have advantages and limitations, if considered under specific angles.

19. The non-spawning fish assay was thought to be the closest to what is normally understood as a screening assay with respect to exposure duration and measurable endpoints. Optimisation was thought to be possible to obtain a short, relevant and reliable assay that can fit all three common fish species.

Life-stage

20. In the non-spawning fish screening assay, adult fish are used, as it allows a separation of males from females and subsequent analysis of sex dependant features for males and females in all three species. Furthermore, due to the larger size of animals at the adult stage compared with juveniles, it becomes possible to measure different endpoints in a single animal, thus reducing the total number of animals needed.

Preferred species

21. The fathead minnow, the Japanese medaka and the zebrafish are commonly used for regulatory purposes in the United States, Japan and Europe respectively. Although it would have been interesting to also test the stickleback for example, this option was not taken due to the additional burden of adding a species in the validation exercise, and also because of the limited regulatory use of this species at this time.

22. LAB 2 and LAB 1 used all three species. LAB 3 used medaka and zebrafish; and LAB 4 used fathead minnow and zebrafish.

Endpoints studied

23. Vitellogenin, gross morphology (including sex characteristics and gonado-somatic index) and gonad histology were the core endpoints of the study. Laboratories used ELISA methods to measure the vitellogenin level in fish. Different ELISA methods were selected by laboratories and may have generated variability in the different responses. All were homologous methods, except one method (ELISA for fathead minnow used in LAB 4).

Route of exposure

24. Chemicals were placed in the water under flow though conditions and no carrier solvent was used for either of the two substances. A long stirring period was necessary to achieve the desired test concentration for 17-β trenbolone.

Test concentrations

25. For the estrogen, 17 β-estradiol, the nominal test concentrations were 10, 32 and 100 ng/l. For the androgen, 17 β-trenbolone, the nominal test concentrations were 50, 500 and 5000 ng/l.

26. The controls received dilution water (reconstituted test water in LAB 4 and dechlorinated tap water in the other laboratories) alone. A control was set for the respective test, however, only in case of LAB 2 the same control was used for both 17 β-estradiol and 17 β-trenbolone in all species.

Housing of male and female fish

27. In the assays using medaka, two laboratories (LAB 1 and LAB 3) separated the males from the females by a mesh and one laboratory (LAB 2) placed males and females in distinct test chambers to compare both separation methods. In zebrafish and fathead minnow assays, males and females were separated by a mesh barrier in all laboratories.

Test conditions

28. Each laboratory recorded the test conditions, including the weekly measurements of test concentrations, the dissolved oxygen level, pH, water hardness and temperature, the diet, the frequency of food distribution, the occurrence of dead fish, external abnormalities and abnormal behaviour, including secondary sex characteristics.

Analysis of data

29. Participating laboratories recorded the raw experimental data on standardized Excel spreadsheets (Annex I) developed specifically for this validation exercise. All Excel spreadsheets completed with laboratory results were sent for data analysis to the lead laboratory and were also submitted to the OECD Secretariat for the records.

30. Additionally, LAB 4 submitted a comprehensive evaluation of data using a scoring system for gonad histology for both fathead minnow and zebrafish. This was not analysed in details for reproducibility assessment.

31. The lead laboratory hired a consultant to perform the statistical analysis. An analysis of variance (ANOVA) was applied to vitellogenin, the gonado-somatic index and quantifiable secondary sex characteristics. Statistical analysis was not performed on the evaluation of gonad histology, due to the lack of an agreed scoring system between all four laboratories.

32. The experimental data, except for the index of dorsal nape pad and that of nuptial tubercles which are secondary sexual characteristics in fathead minnow, were checked for homogeneity of variances across treatments by Levene's test (Figures 1 and 2). When no homogeneity was observed in the data, a logarithmic transformation was performed and to check for homogeneity of variances across treatments again. When the assumptions were met (with or without transformation), the data were subjected to one-way analysis of variance (ANOVA) followed by Dunnett's multiple comparison test. When no homogeneity was observed even in the transformed data, the nonparametric Kruskal-Wallis test was used, followed by the Mann-Whitney U test with Bonferroni's adjustment. The VTG values lower than the determination limit were transformed to half the value of the determination limit for each analysis. Data on secondary sexual characteristics of fathead minnow (index of dorsal nape pad, index of nuptial tubercles) were subjected to Kruskal-Wallis test followed by the Mann-Whitney U test with Bonferroni's adjustment without treatment by Levene's test. Differences were considered to be significant at $p<0.05$ in all tests, however, Bonferroni's p value was used in Mann-Whitney U test. All statistical analyses were performed by JMP ver.4.05J produced by SAS Institute Japan.

RESULTS

Measured test concentrations

Means of measured concentrations (% nominal) in tanks 1 and 2				
Nominal values		10 ng/l	32 ng/l	100 ng/l
LAB 1	Medaka	8.9 (89)	28.2 (88)	85 (85)
LAB 2		10.1 (101)	32.8 (102)	104.3 (104)
LAB 3		8.9 (89)	28.3 (88)	93.6 (94)
LAB 1	Fathead minnow	8.76 (88)	28.6 (89)	85.9 (86)
LAB 2		9 (90)	27 (84)	83.7 (84)
LAB 4		6.2 (62)	20 (64)	63.5 (63)
LAB 4	Zebrafish	86.5 (87)	22.4 (70)	86.5 (87)
LAB 1		10 (100)	24.4 (76)	85.8 (86)
LAB 2		8.9 (89)	34.1 (106)	107.7 (108)
LAB 3		9.1 (91)	28.6 (89)	91.1 (91)

Table 2: measured test concentrations of 17-β estradiol.
Shaded cell: low measured concentration (outside the 80-120% nominal concentration)

Means of measured concentrations (% nominal) in tanks 1 and 2				
Nominal values		50 ng/l	500 ng/l	5000 ng/l
LAB 1	Medaka	34.7 (79)	365 (74)	4012 (80)
LAB 2		42.4 (85)	476 (95)	4854 (97)
LAB 3		45 (90)	540 (108)	5699 (114)
LAB 1	Fathead minnow	40.8 (82)	400 (80)	4057 (81)
LAB 2		51.9 (104)	459 (92)	4363 (87)
LAB 4		40.3 (80)	441 (88)	4580 (96)
LAB 4	Zebrafish	46 (92)	457 (91)	5172 (103)
LAB 1		35 (71)	350 (70)	3439 (69)
LAB 2		44.5 (89)	482 (96)	5177 (103)
LAB 3		53.5 (107)	548 (110)	5354 (107)

Table 3: measured test concentrations of 17-β trenbolone**.**
Shaded cells: low measured concentrations (outside the 80-120% nominal concentration)

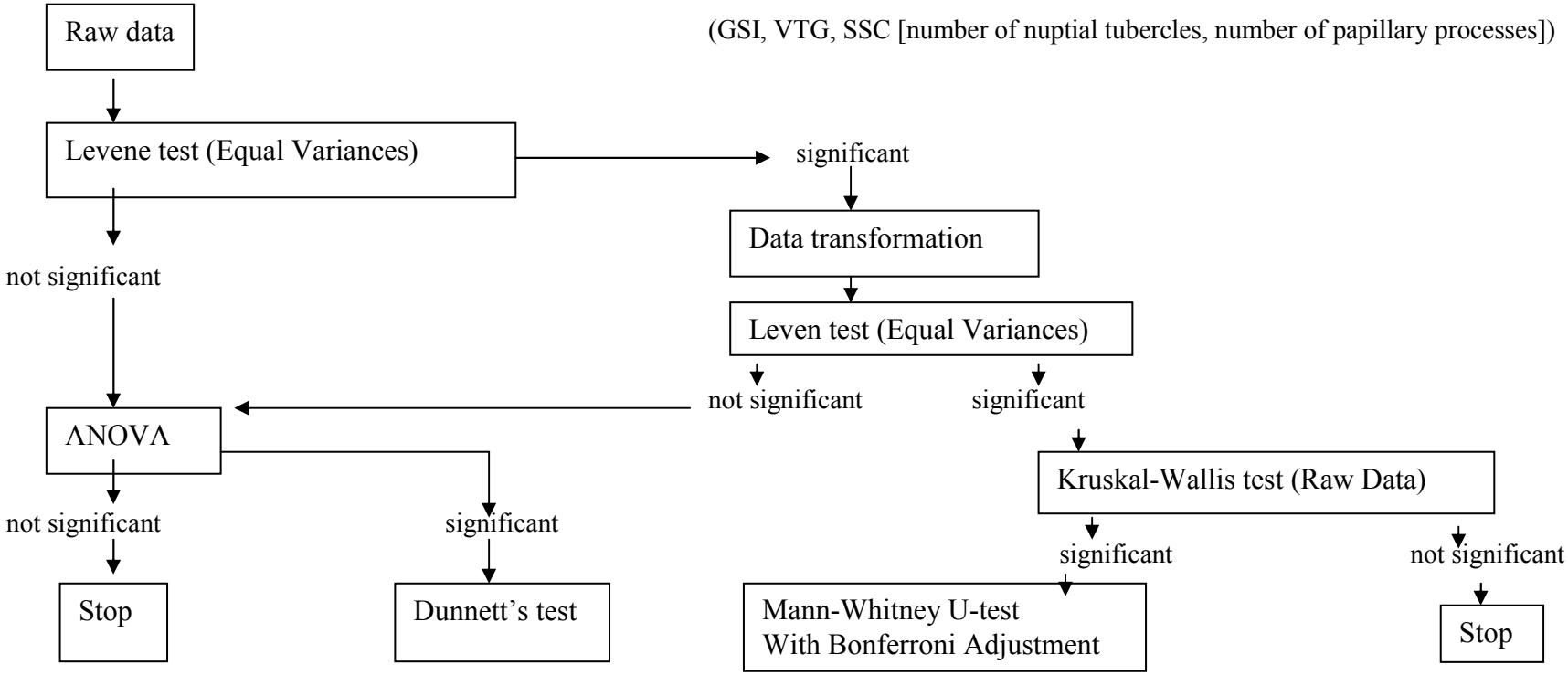
33. In general, measured concentrations ranged between 70 and 110% of nominal contrations. There were a few low measured concentrations, where measured concentrations were outside the range 80%-120% of nominal concentration.

Statistical approach followed

34. Figures 1 and 2 show the decision pathways followed to perform the statistical analysis on all quantitative data and the tests conducted to determine statistical significance.

35. The presence of outliers can be illustrated by the following examples: in a few cases (less than 5) a female fish was encountered among the males. Although not initially found out, it was discovered during the dissection, and in the very high vitellogenin level was found among low values. Only in this case, the data of concern was excluded in the statistical analysis as outliers.

Figure 1: Statistic Flow chart 1

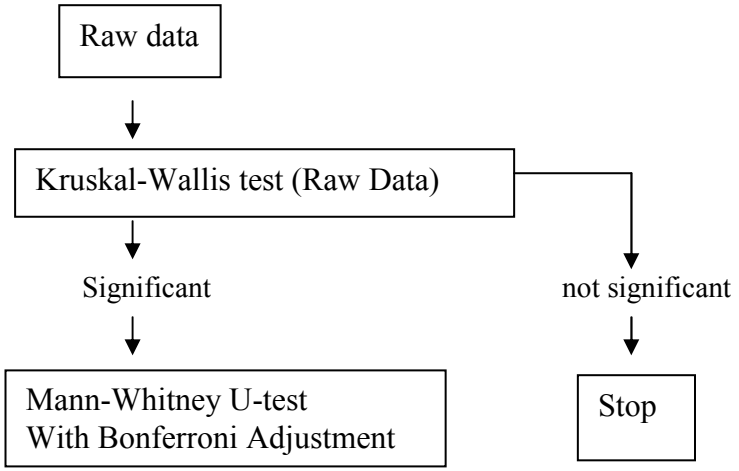


SSC: secondary sexual characteristics
Significance level: 0.05

Data Transform (Log; logarithmic transformation)
Standard deviation increased according to the increase in mean with log normal distribution (Shibata 1995).
Because a lot of data sets of VTG were heterogeneous variance among control group and groups treated with some concentrations of the hormone, data transformation was required in order to perform the ANOVA. The standard deviations increased in proportion to means, then it was presumed that the data of VTG followed log normal distribution, and logarithmic transformation of the data were carried out.
Shibata Yoshisada. 1995. Probability and probability distributions. In: Miyahara Hideo and Tango Toshizo (eds) Medicine statistics handbook. Asakura Shoten, Tokyo (In Japanese)

Figure 2: Statistic Flow chart 2

(SSC [index of dorsal nape pad, index of nuptial tubercles])



SSC: secondary sexual characteristics
Significance level: 0.05

17 β- estradiol study

Vitellogenin (VTG)

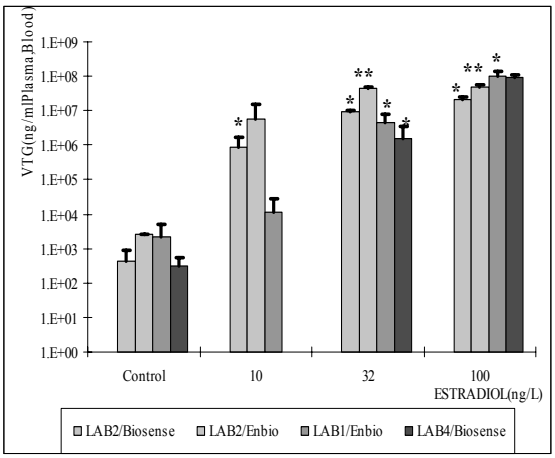


Figure 3a: VTG in males fathead minnow at d-14
LAB 4 DAY14 Control male Sample No.2 was lost during centrifugation.
LAB 4 DAY14 10ng/l male all Samples were lost during centrifugation.
LAB 4 DAY14 Control male Sample No.5: The data was excluded (no plasma available)
LAB 4 DAY14 100ng/l male Sample No.5: the fish died.

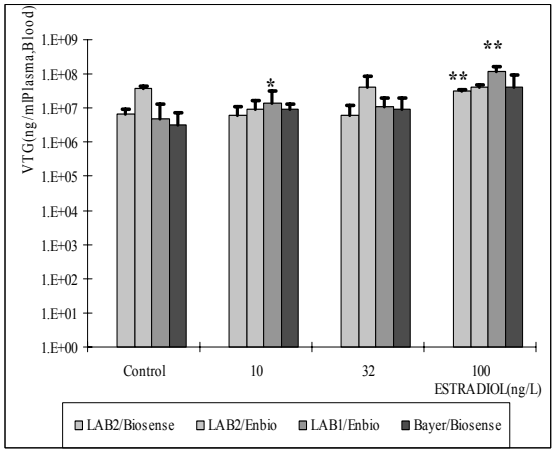


Figure 3b: VTG in females fathead minnow at d-14
LAB 4 DAY14 10ng/l female Sample No.2: The data was excluded (blood was yellowish).

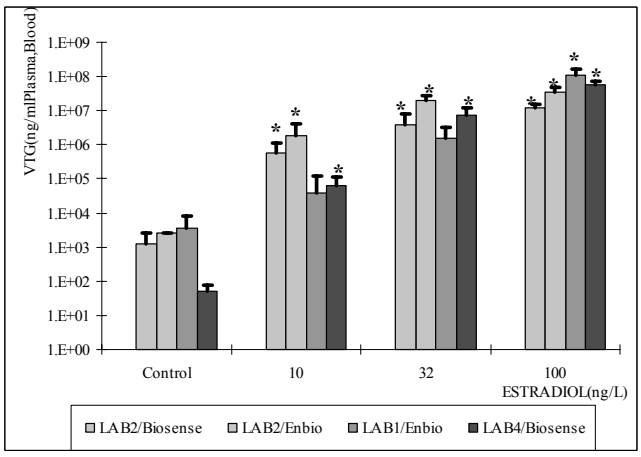


Figure 3c: VTG in males fathead minnow at d-21
LAB 1 DAY21 10ng/l male Sample No.5:The fish died

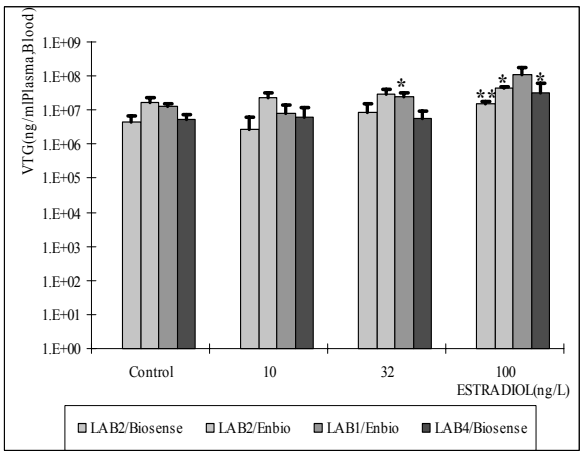


Figure 3d: VTG in females fathead minnow at d-21

LAB 2,LAB 1 measured Blood. LAB 4 measured Plasma.
significant level 5%;*,1%;**

36. Means and standard deviations are plotted on a logarithmic scale. Figures 3a and 3c show that significant changes in vitellogenin level after 17 β- estradiol exposure can be observed in males fathead minnow where basal levels are usually very low. All laboratories were able to detect significant levels at 14 days of exposure at 32 ng/l. One laboratory could detect a significant increase in VTG at 10 ng/l after 14 days.

37. The ELISA methods used by LAB 2 and LAB 1 were specific to fathead minnow (using homologous antibodies) and the other ELISA methods, used in LAB 4 and LAB 2 were using carp-raised antibodies. Overall, results are homogenous.

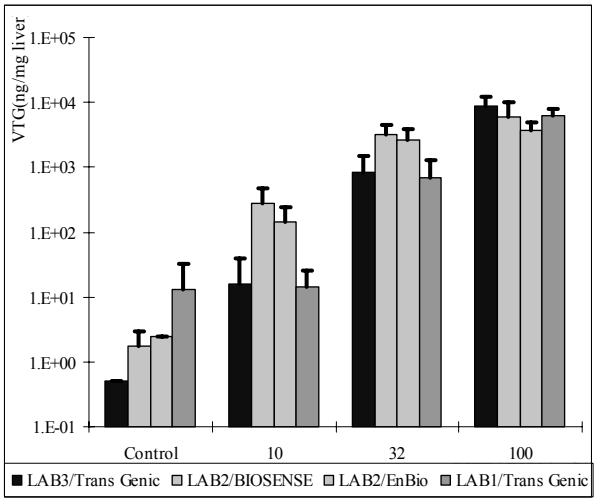


Figure 4a: VTG in males medaka at d-14

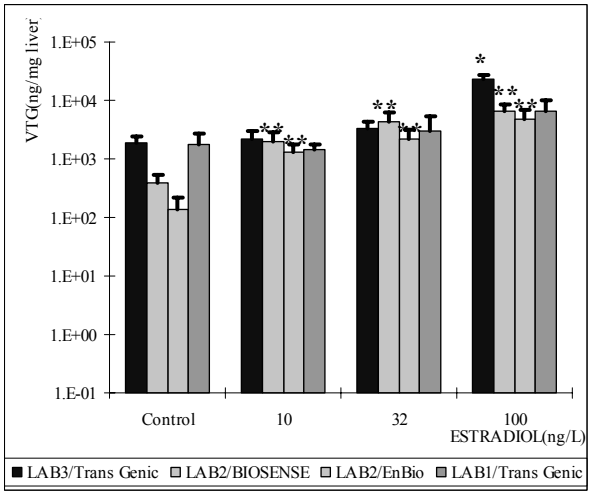


Figure 4b: VTG in femals medaka at d-14

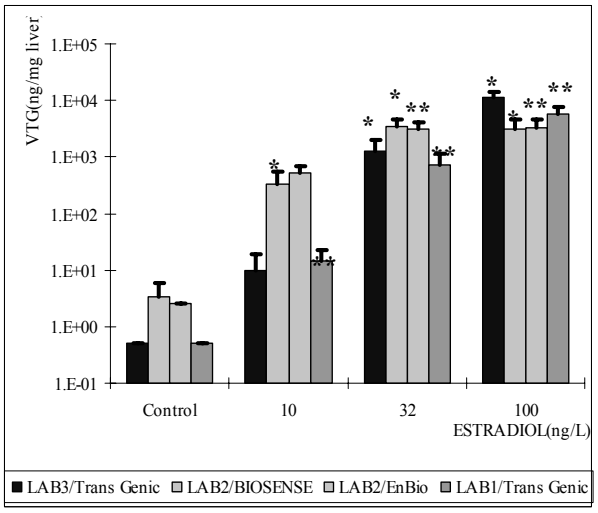


Figure 4c: VTG in males medaka at d-21
LAB 2 DAY21 10ng/l male Sample/No5: The fish died The Data is excluded.
significant level 5%;*,1%;**

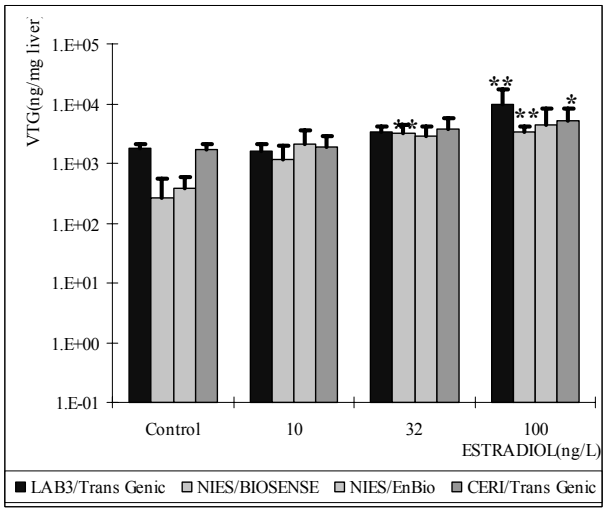


Figure 4d: VTG in females medaka at d-21
LAB 2 DAY21 32ng/l female Sample/No5: The fish died. The Data is excluded

38. Again, Figures 4a and 4c on males medaka show that in all laboratories significant changes in vitellogenin levels could be achieved at 14 days after exposure at 32 ng/l of estradiol. At 14 days, it seems that already at 10 ng/l, the level of significance is close to be achieved, but this assumption was not verified at 21 days.

39. The ELISA methods were specific to medaka, but those used by LAB 1 and LAB 3 were using polyclonal antibodies and the one used by LAB 2 was using monoclonal antibodies. Overall, results are homogenous.

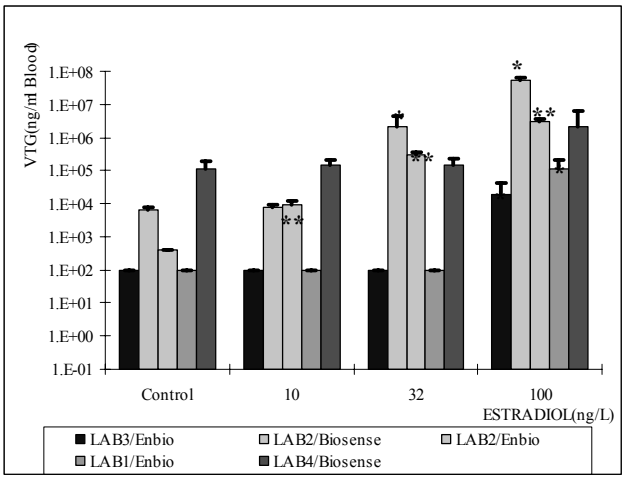


Figure 5a: VTG in males zebrafish at d-14
LAB 4 DAY14 Control male Sample/No1 was excluded because female was mistaken for male.

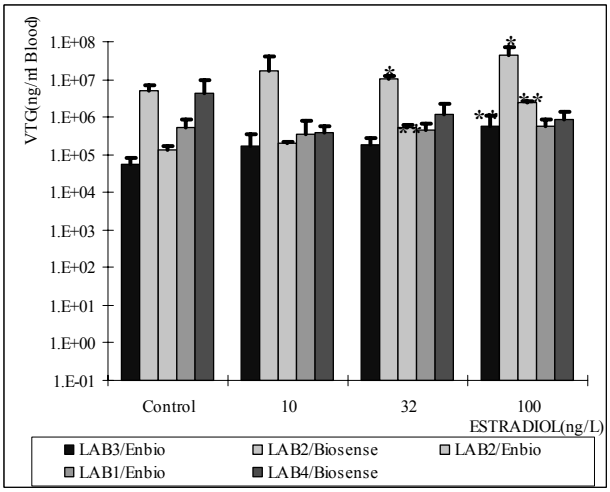


Figure 5b: VTG in females zebrafish at d-14
LAB 2 DAY14 10ng/l female Sample/ No5: The fish died.

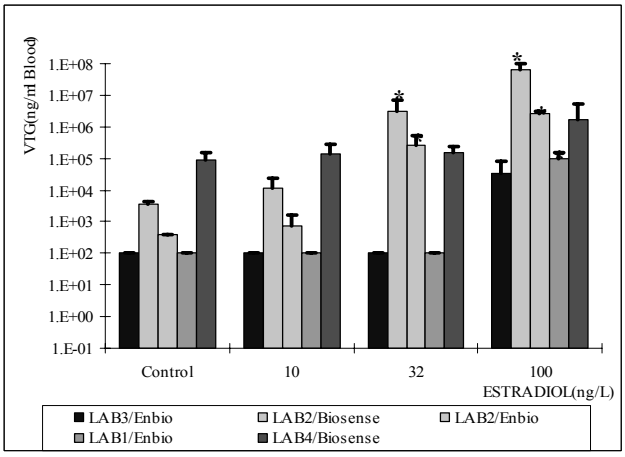


Figure 5c: VTG in males zebrafish at d-21
LAB 3 DAY21 10ng/l male Sample/ No5: The fish died

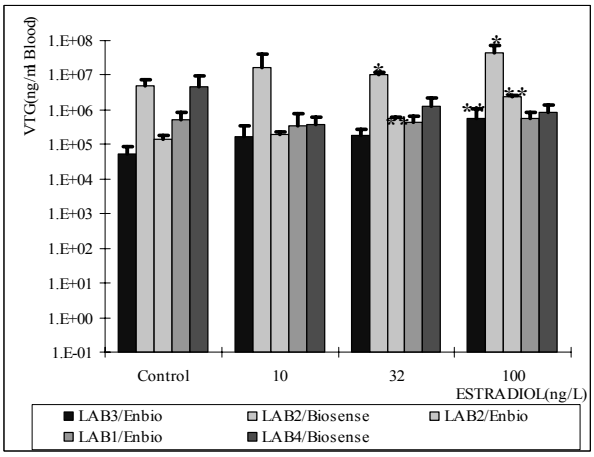


Figure 5d: VTG in females zebrafish at d-21
LAB 1 DAY21 100ng/l female Sample/No4-5; the fish died
LAB 2 DAY21 10ng/l female Sample/ No5: The fish died. These Data are excluded

significant level 5%;*,1%;**

40. Conclusions previously made on the other two species appear less clearly in zebrafish. One laboratory especially (LAB 4) found very high basal levels in (non exposed) male zebrafish. The method they used was the same as in LAB 2; however LAB 2 was able to detect significance level at 32 ng/l after 14 days of exposure. One female fish was found among the males in the control at day 14, but its VTG level was not accounted for in the calculation of the mean. Two laboratories (LAB 3 and LAB 1) found a significant increase in males at 100 ng/l after 14 days, and the other laboratory (LAB 2) found a significant increase in males at 10 ng/l after 14 days. Responses appear less obvious after 21 days than after 14 days of exposure.

41. Both ELISA methods used were specific to zebrafish (homologous standard and monoclonal antibodies). LAB 2 used both ELISA methods to quantify vitellogenin in plasma samples.

Gonado-somatic index (GSI)

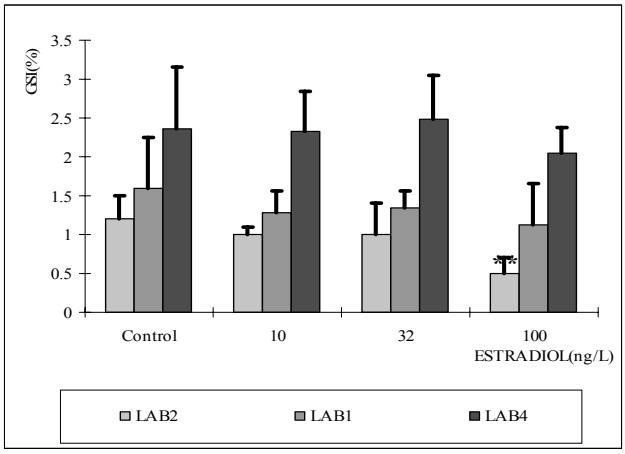


Figure 6a: GSI in males fathead minnow at d-21
LAB 1 DAY21 10ng/l male Sample/No5:The fish died.

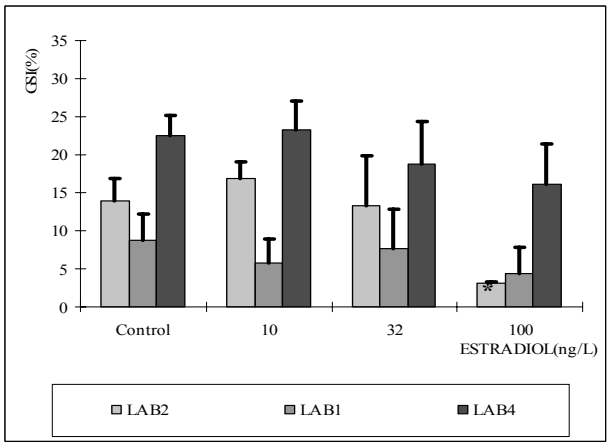


Figure 6b: GSI in females fathead minnow at d-21

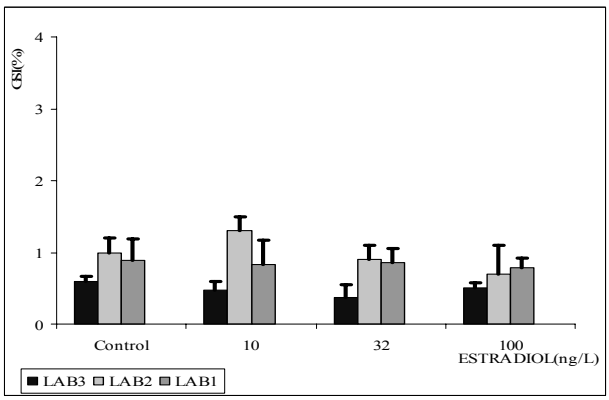


Figure 6c: GSI in males medaka at d-21
LAB 2 DAY21 10ng/l male Sample/No5: The fish died

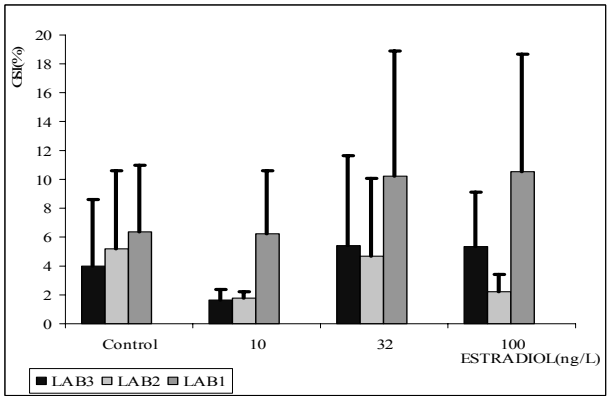


Figure 6d: GSI in females medaka at d-21
LAB 2 DAY21 32ng/l female Sample/No5: The fish died.

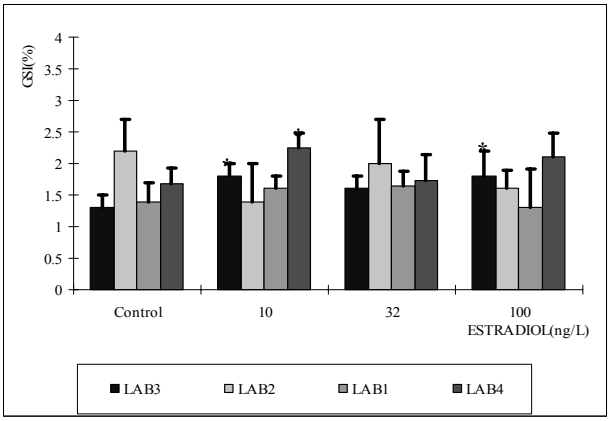


Figure 6e: GSI in males zebrafish at d-21
LAB 3 DAY21 10ng/l male Sample/No5: The fish died

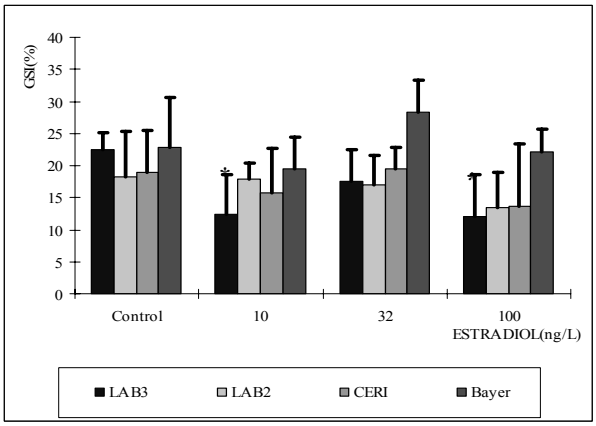


Figure 6f: GSI in females zebrafish at d-21
LAB 2 DAY21 10ng/l female Sample/No5, and
LAB 1 DAY21 100ng/l female Sample No.4 & No.5:
The fish died

significant level 5%;*,1%;**

42. All figures show the data collected on the gonado-somatic index (GSI)obtained after 21-days of exposure. There is no major difference in GSI measurements between 14 days and 21 days. There is no trend in the responses in the selected dose-range.

Secondary sex characteristics (number of papillary processes in medaka, number and index of nuptial tubercles in fathead minnow)

43. These measurements will not be shown on graphs here, but can be found in Annex IV (Collection of all graphs). The reason is that these indicators are mainly sensitive to androgens. Therefore, estradiol exposure moderately reduced the number of papillary processes on the anal fin of medaka in one laboratory only ($p<0.05$). Nothing was observed on fathead minnow.

Gonad histology

44. One laboratory (LAB 4) made 2 slides per gonad with a distance of 200µm and made similar observations for both sections. In other laboratories, differences were sometimes observed between slides (or sections) of the same gonad. In general, participating laboratories indicated that inter-individual differences are more commonly observed than differences between sections of the same individual..

45. Only LAB 4 evaluated gonad histology using a detailed scoring system; detailed findings are not presented here but constitute part of the full GLP draft reports. The scoring system is based on PathData©system where four grades are used for abnormalities (grade 1: minimal/very few/very small; grade 2: slight/few/small; grade 3: moderate/moderate number/moderate size; grade 4: marked/many/large). In oviducts, the following features were scored when observed: cellular debris, yolk plug, mature oocytes and inflammation. In the ovaries, the following features were scored when observed: atretic degenerating oocytes, increased immature oocytes, inflammation, epithelium hyperplasia, epithelium yolk accumulation. In the testes, the following features were scored when observed: size of sperm cysts, size of lumina, apoptosis, prominent light cells, reduction in number of sperm cysts, inflammation and spermatid clusters.

46. Additionally, fathead minnow and zebrafish control fish (20 males and 20 females) from LAB 4 were dissected for gonad histopathology with the view to collect historical control data. The following findings were recorded:

- In zebrafish, 20 male and 20 female "historical controls" were collected. In addition, 5 male and 5 female controls were collected for each compound and at each sampling date. In the testes, the size of spermatogenic cysts and lumina was assessed. Most animals were in a mid spermatogenic stage.

- In the ovaries, various stages of developing follicles/oocytes were observed up to the late vitellogenic stage. Minimal or slight atresia/degeneration of mature oocytes was observed in a high proportion of animals. In animals which had released mature eggs to the oviducts, a relative increase of the number of immature oocytes was recorded. Other findings in the ovaries were interstitial inflammation of undetermined cause and focal squamous metaplasia of the follicular epithelium.

- In animals in which the oviduct contained no egg material, the oviduct was very small and delicate. Therefore it is sometimes not contained in the sections. In a large number of animals however, the oviduct contained mature eggs, yolk plug and/or cellular debris. The oviduct wall was slightly thicker in these animals and sometimes contained some inflammatory cells.

- In fathead minnow, 20 male and 20 female "historical controls" were collected. In addition, 5 male and 5 female controls were collected for each compound at the start of the study, at the interim sacrifice after 2 weeks and at terminal sacrifice. In the testes, control animals mostly showed a mid spermatogenic stage of sexual maturation. In the lumina of the gonad, mature spermatozoa were present in large numbers. The germinal epithelium showed moderate thickness. In some areas a thin epithelium with scattered spermato-genesis (spermatogenic cysts with late phases of spermatid development) was seen indicating a progressing maturation. Late spermatogenic testes had a predominance of thin epithelium, more scattered spermatogenic cysts (often with later stages of spermatid development) and a large amount of spermatozoa. In late spermatogenic stages a relative prominence of light cells (spermatogonia and Sertoli cells) in the germinal epithelium was sometimes present along with a reduction of spermatogenic cysts. Early spermatogenic testes were characterised by narrow lumina containing few spermatozoa and thick germinal epithelium. The distinction between mid and late spermatogenic stages was vague since in the same testis areas of both stages were often present. Thus, the prevailence of one stage compared to the other was considered without further grading. In the lumina, foamy desquamated cells and/or granulomas were observed in some animals. No ovotestis was observed in control animals.

- The deferent duct was located adjacent to the testis and could be identified at trimming by its pigmentation. In animals with early spermatogenic testes which had not yet generated large numbers of mature spermatozoa, reduced spermatozoa/oligspermia was observed. Many animals showed small spherical bodies in the wall of the deferent duct some of which possessed concentrical layering or mineralization. Some of these structures were observed in the testes of a few animals. Often the epithelium of the deferent duct was vacuolated. Foamy desquamated cells, which represent presumably macrophages, granulomas and spherical structures, are most probably sequels of a parasitic infestation with pleistophora species (spp.) which however in males did not reach the same extent and relevance as in females.

- In the ovaries, various stages of developing follicles were observed up to late vitellogenic follicles/oocytes. The number of degenerating/atretic follicles/oocytes was mostly rated as minimal or slight. In most animals, follicular degeneration was observed associated with inflammatory changes caused by microsporidian parasitic infestation of the oocytes. The parasite was identified as *Pleistophora* spp. (Warner 1972) but the species could not be definitely classified. The parasites were contained mostly in late vitellogenic follicles and could eventually lead to follicle destruction. The spores the lead to a granulamatous inflammation of the interstitium. The parasite is known to parasitize various fish species including fathead minnows (Nagel and Hoffman 1977). No intersex formation was observed in control animals. In the oviducts, cell debris, yolk plug (degenerated egg material) or mature oocytes were observed at varying incidences. As the deferent ducts, the oviducts could be identified grossly by their pigmentation.

47. Table 4 provides a summary of the findings for 17β-estradiol study. Summaries of observations are qualitative because no common and precise scoring of severity for each species existed prior to Phase 1A of the Validation that could be uniformly used by all laboratories. Generally, after 17-β estradiol exposure an increase proportion of spermatogonia was observed after 14 days already (confirmed after 21 days of exposure) in fathead minnow and zebrafish (mentioned by LAB 4 and LAB 1 laboratories). The occurrence of testis-ova was noted, although observed in a low number of fish and also in control zebrafish.

Fish species	Laboratory	Testis after 17 β-estradiol exposure	Ovary after 17 β-estradiol exposure
Fathead minnow	LAB 4	- prominence of light cells (spermatogonia/Sertoli cells) in the germinal epithelium in all dose groups after 2 and 3 weeks of treatment without clear dose-response relationship - apoptosis in the germinal epithelium at the high dose level in one animal after 2 weeks of treatment and in 3 animals after 3 weeks of treatment	- no induced changes
	LAB 1	- increased proportion of spermatogonia at 100 ng/L (1 of 5 fish after 2 weeks and 3 of 5 fish after 3 weeks), but similar increase is attained in one of the controls - decreased proportion of spermatocytes at 100 ng/L (1 of 5 fish after 2 weeks and 3 of 5 fish after 3 weeks)	- no noticeable abnormalities compared with the controls - considerable increase of atretic oocytes in control and treatment groups after 2 and 3 weeks
	LAB 2	- no induced changes	- no induced changes
Medaka	LAB 1	- presence of testis-ova at 32 ng/L (1 of 5 fish after 2 and 3 weeks) and at 100 ng/L (1 of 5 fish after 2 weeks and 3 of 5 fish after 3 weeks) - relative decrease of spermatogenesis at 100 ng/L (1 of 5 fish after 3 weeks	- no noticeable abnormalities compared with the controls
	LAB 3	- no clear trend but high frequency of abnormalities at 100 ng/L [degenerated /reduced germ cell cyst (3 of 5 fish at 100 ng/L after 3 weeks)] - presence of testis-ova (1 of 5 fish at 32 ng/L after 2 weeks	- trend of increase in abortive oocyte (ateretic follicles) and degenerating ovarian mature oocytes dose-dependently after 2 and 3 weeks - decrease in degenerating of ovulated and remained mature oocytes in ovarian cavity dose-dependently after 2 and 3weeks
	LAB 2	- presence of testis-ova (1 of 5 fish at 32 ng/L after 3 weeks)	- no induced changes
Zebrafish	LAB 4	- increased incidence of prominent light cells (spermatogonia and/or Sertoli cells) after 3 weeks of treatment at 32 ng/L and above	- no induced changes
	LAB 1	- no noticeable abnormalities except increased proportion of spermatogonia at 100 ng/L (1 of 5 after 3 weeks) - presence of testis-ova in control and treatment groups	- no noticeable abnormalities compared with the controls
	LAB 3	- no abnormalities with increasing the exposure concentrations - presence of testis-ova in control and treatment groups	- no abnormalities with increasing the exposure concentrations
	LAB 2	- no induced changes	- no induced changes

Table 4: Summary of the evaluation of the gonad histopathology after 17 β-estradiol exposure in all laboratories for all three fish species.

17 β-trenbolone study

Vitellogenin (VTG)

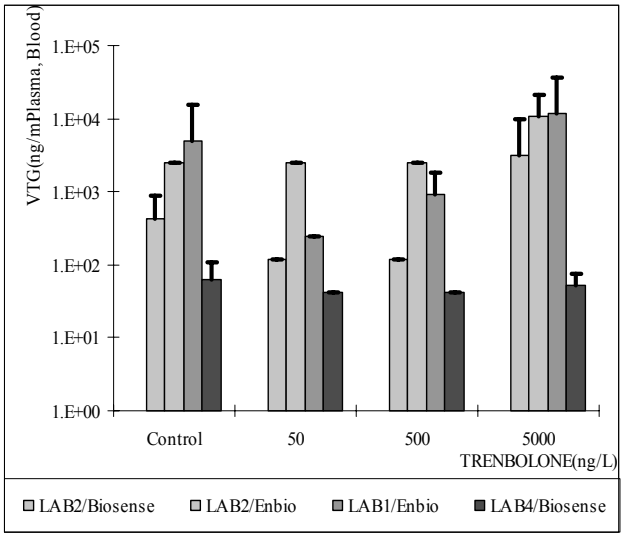


Figure 7a: VTG in males fathead minnow at d-14

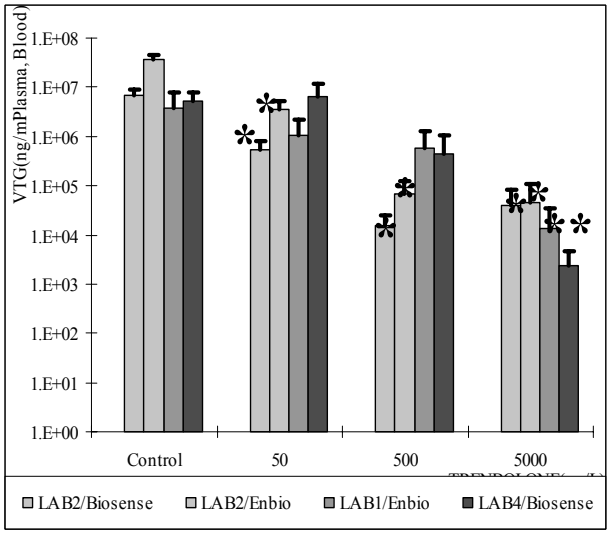


Figure 7b: VTG in females fathead minnow at d-14
LAB 4 DAY14 5000ng/l female Sample/No1,No5 : The fish died.

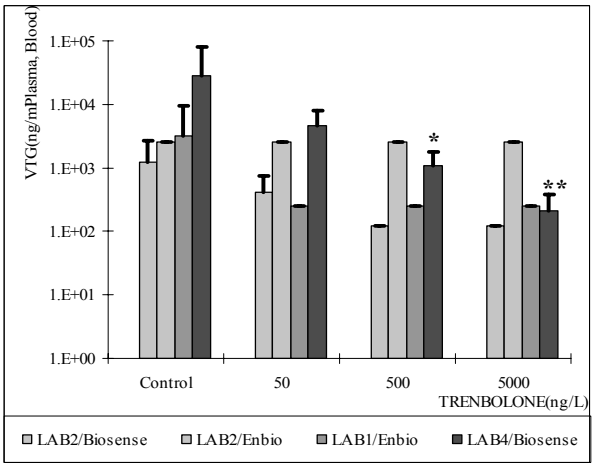


Figure 7c: VTG in males fathead minnow at d-21
LAB 4 DAY21 50ng/l male Sample/No1 :
The sample was lost during centrifugation.

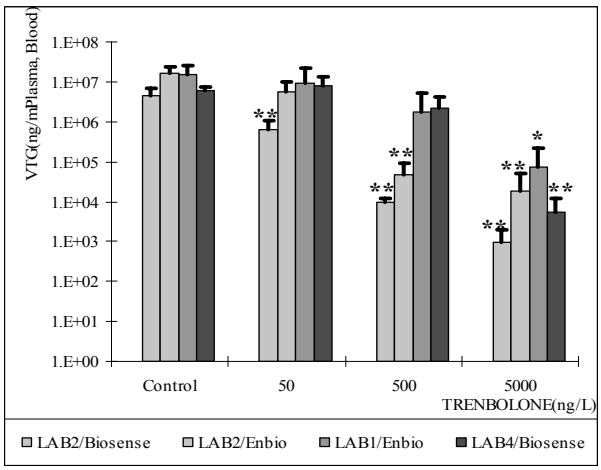


Figure 7d: VTG in females fathead minnow at d-21
LAB 2 DAY21 5000ng/l female Sample/No5 :The fish died.
LAB 1 DAY21 500ng/l female Sample No.3: The data was excluded because male was mistake for female.
LAB 4 DAY21 5000ng/l female
Sample No.5: The fish died

LAB 2,LAB 1 measured Blood. LAB 4 measured Plasma
significant level 5%;*,1%;**

48. All laboratories were able to detect a statistically significant decrease after 21 days of exposure at 5000 ng/l.

49. One laboratory (LAB 2) used two different methods (a carp-based ELISA and a fathead minnow-based ELISA) and was able to detect a significant decrease at a concentration as low as 50 ng/l. Both

ELISA methods were also used, separately by LAB 4 and LAB 1, but these laboratories only detected a significant decrease at the highest concentration (Figure 7b) at 21 days and 14 days after exposure respectively

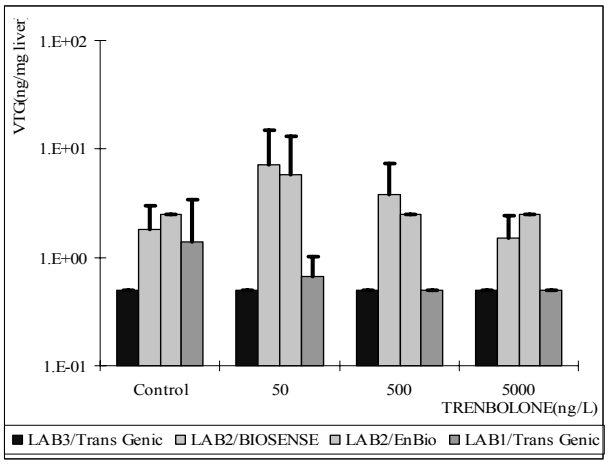


Figure 8a: VTG in males medaka at d-14

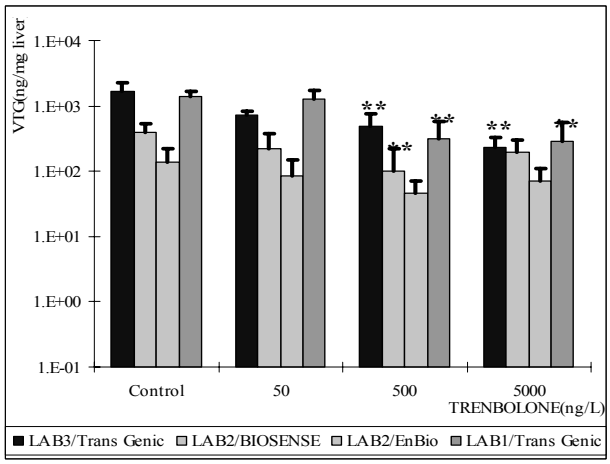


Figure 8b: VTG in females medaka at d-14

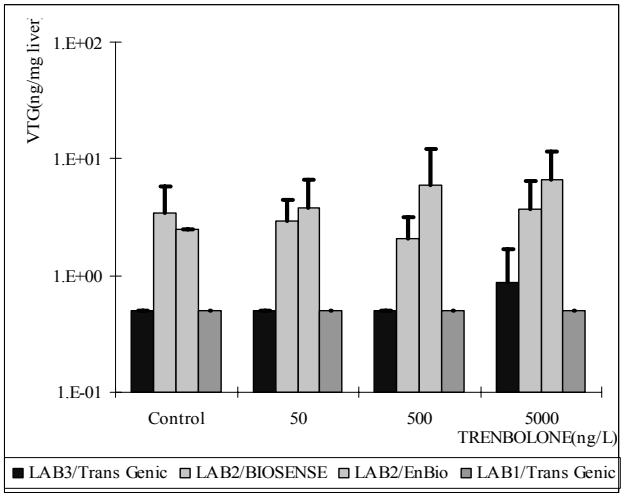


Figure 8c: VTG in males medaka at d-21
LAB 2 DAY21 500ng/l male Sample/No4,No5: the fish Died
LAB 2 DAY21 5000ng/l male Sample/No5: the fish died
LAB 3 DAY21 500ng/l male Sample/No5: The fish dead
significant level 5%;*, 1%;**

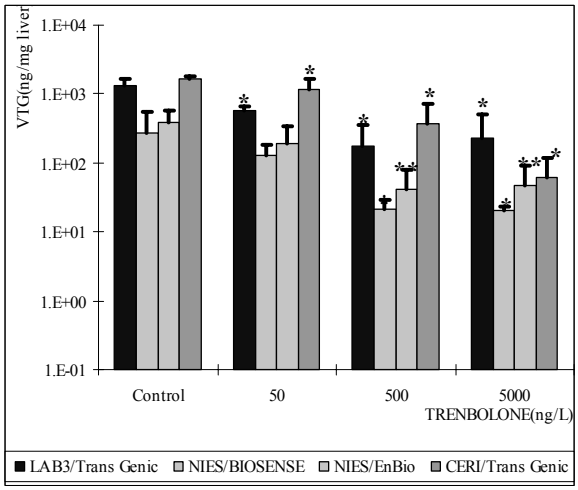


Figure 8d: VTG in females medaka at d-21
LAB 2 DAY21 50ng/l female Sample/No5: the fish died

50. A significant decrease was detected by all laboratories at 500 ng/l after 21 days of exposure. Already at 14 days, the trend appeared clearly ($p<0.01$), but one method did not detect it.

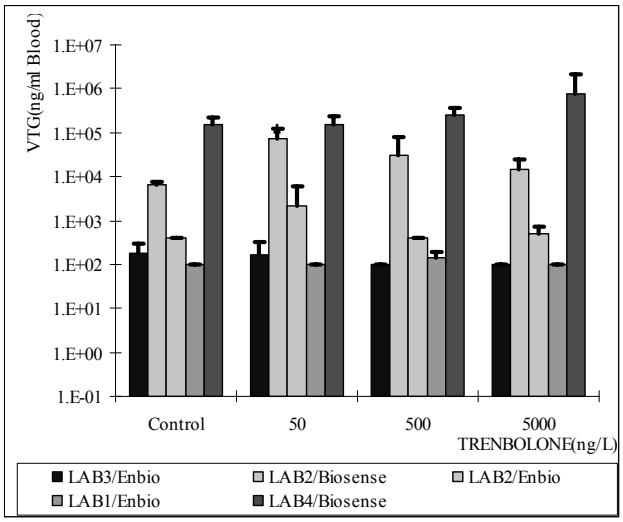


Figure 9a: VTG in males zebrafish d-14
LAB 4 DAY14 500ng/l male Sample/No1:
The data was excluded because female was mistaken for male.
LAB 1 DAY14 50ng/l male Sample No.2: The data was excluded because female was mistaken for male.

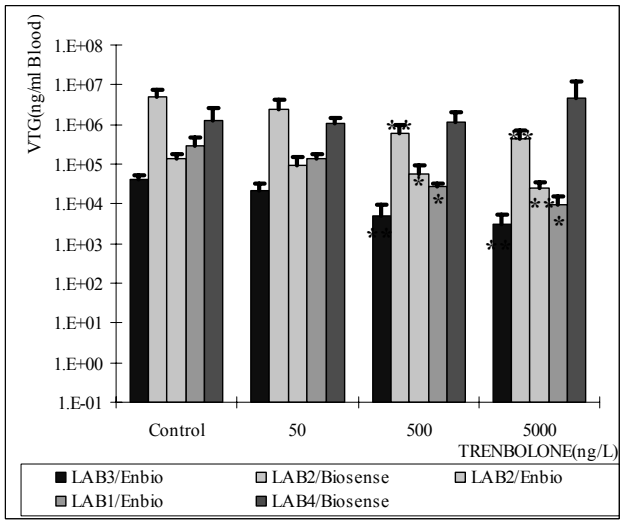


Figure 9b: VTG in females zebrafish at d-14

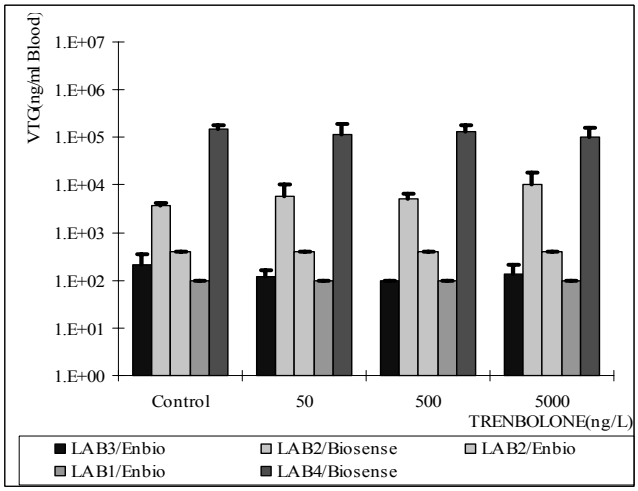


Figure 9c: VTG in males zebrafish at d-21
LAB 1 DAY21 500ng/l male Sample/No4,No5: the fish died
LAB 4 DAY21 500ng/l male No.5:
The fish died

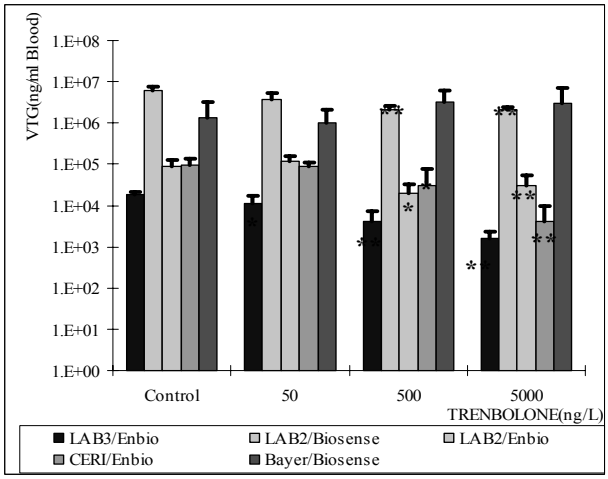


Figure 9d: VTG in females zebrafish at d-21
LAB 3 DAY21 500ng/ female Sample/No5:The fish died
LAB 2 DAY21 50ng/l female Sample/No5:The fish died

significant level 5%;*, 1%;**

51. At 500 ng/l, and after 14 days of exposure to 17 β-trenbolone, all laboratories except one were able to detect a statistically significant decrease (for two laboratories, $p<0.01$) in vitellogenin levels in females zebrafish. One laboratory (LAB 4) using the same method as LAB 2 (1st yellow stick) never detected a decrease in vitellogenin levels; all VTG concentrations were in general very high when measured with this method in both laboratories.

Gonado-somatic index (GSI)

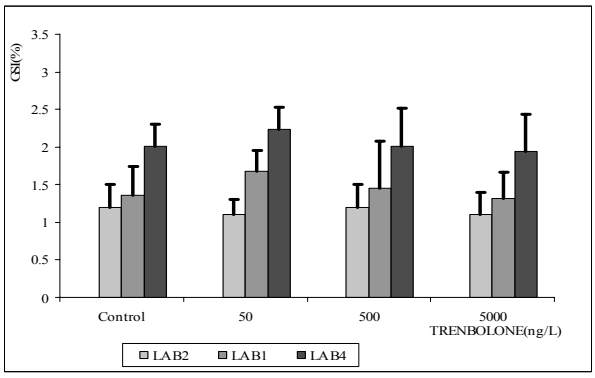


Figure 10a: GSI in males fathead minnow at d-21

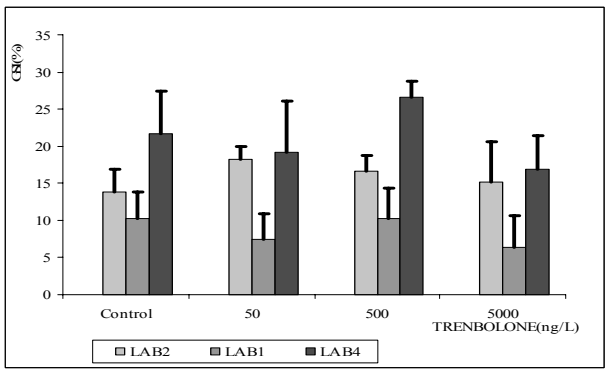


Figure 10b: GSI in females fathead minnow at d-21
LAB 4 DAY21 5000ng/l female Sample/No5:the fish died
LAB 2 DAY21 5000ng/l female Sample/No5 : the fishdied
LAB 1 DAY21 500ng/l female Sample No.3: The data was excluded because male was mistake for female.

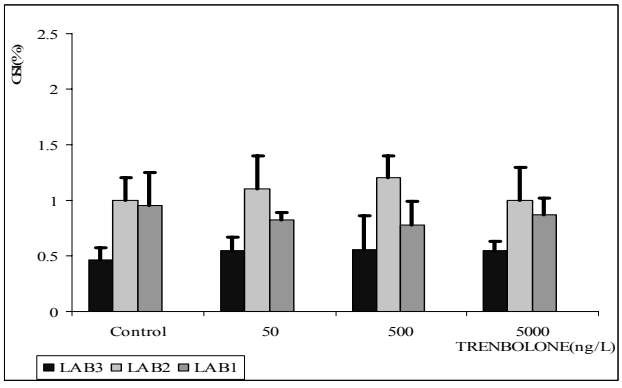


Figure 10c: GSI in males medaka at d-21
LAB 3 DAY21 500ng/l male Sample/No5,
LAB 2 DAY21 500ng/l male Sample/No4,No5, and
LAB 2 DAY21 5000ng/l male Sample/No5: The fish died

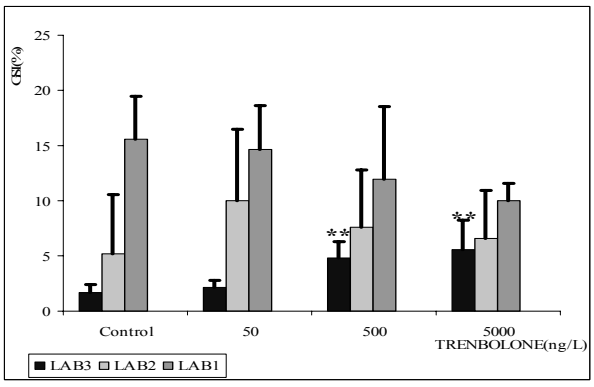


Figure 10d: GSI in females medaka at d-21
LAB 2 DAY21 50ng/l female Sample No.5:The fish died

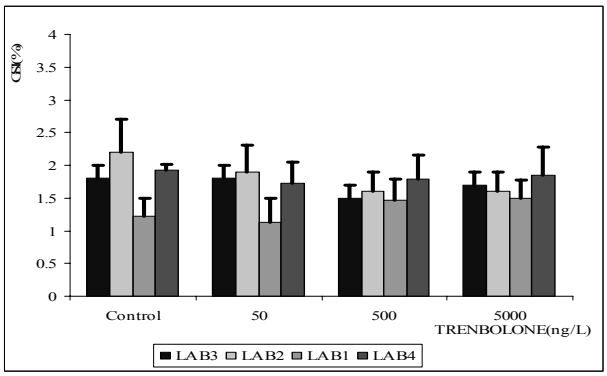


Figure 10e: GSI in males zebrafish at d-21
LAB 4 DAY21 500ng/l male Sample/No5,
LAB 1 DAY21 500ng/l male Sample/No4,No5 and
LAB 2 DAY21 50ng/l female Sample/,No5:
The fish died

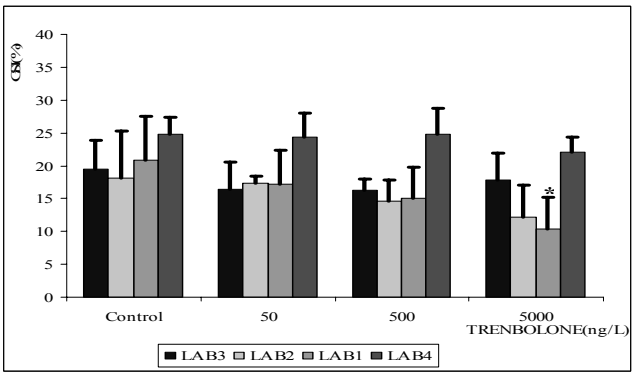


Figure 10f: GSI in females zebrafish at d-21
LAB 3 DAY21 500ng/l female Sample/No5: The fish
LAB 1 DAY21 500ng/l male Sample/No4,No5:
The fish died

significant level 5%;*,1%;**

52. As was already observed in the 17β-estradiol study, the gonad-somatic index poorly responded to 17β-trenbolone exposure as well, although some statistically significant results were obtained in female medaka in one laboratory (Figure 10d).

Secondary sex characteristics (number of papillary processes in medaka, number and index of nuptial tubercles in fathead minnow)

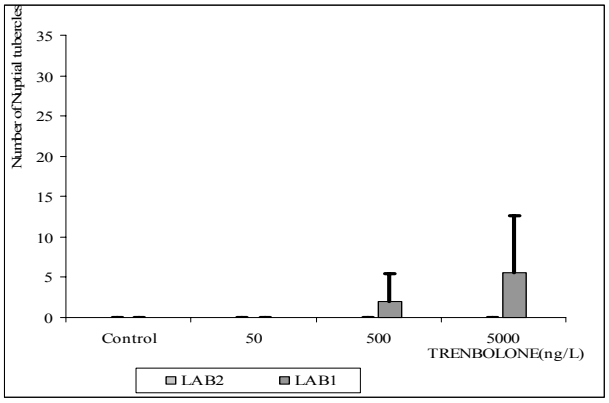


Figure 11a: Number of nuptial tubercles in females fathead minnow after 14 days exposure

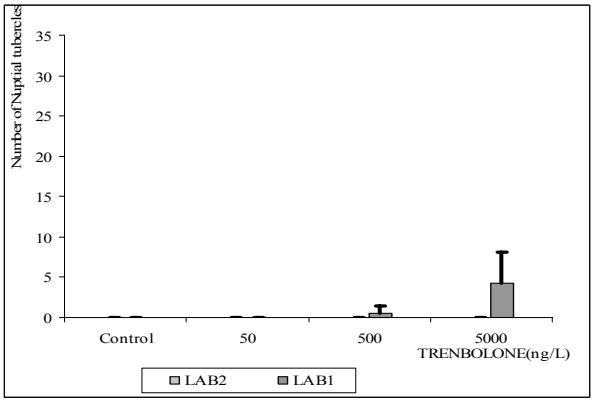


Figure 11b: Number of nuptial tubercles in females fathead minnow after 21 days exposure
LAB 2 DAY21 5000ng/l female Sample No.5: The fish died
LAB 1 DAY21 500ng/l female Sample No.3: The data was excluded because male was mistake for female

53. Figures 11 a and 11b are the most representative of the responses observed on secondary sex characteristics in fathead minnow after 17 β-trenbolone exposure. They represent the number of nuptial tubercles present on the face of the fish. Nuptial tubercles are normally seen in males but can also be observed and counted in females exposed to an androgenic substance. Other parameters measured (i.e. index of nuptial tubercles, index of dorsal nape pad) did not respond in a clear manner in males or females. Only in one laboratory, a significant increase of the index of dorsal nape pad was observed after 14 days in males, but this trend was not confirmed at 21 days.

54. Laboratories expressed difficulties in measuring the secondary sex characteristics in fathead minnow, in particular the height of the dorsal nape pad.

55. All other results not shown here can be found in Annex III (Phase 1A figures).

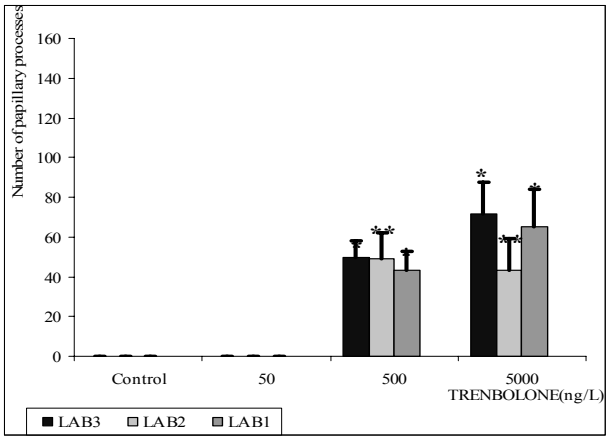


Figure 12a: Number of papillary processes in females medaka after 14 days exposure

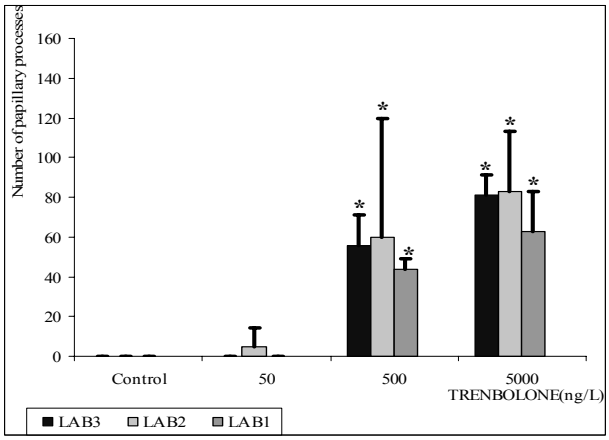


Figure 12b: Number of papillary processes in females medaka after 21 days exposure
LAB 2 DAY21 50ng/l female Sample No.5:The fish died

56. The number of papillary processes on the anal fin of medaka is the secondary sex characteristic measured for that species. It is normally seen in males but can also be observed and counted in females exposed to an androgenic substance. Figures 12a and 12b demonstrate that the endpoint respond clearly already after 14 days of exposure and statistical significance was detected at 500 ng/l in all three laboratories which used the medaka.

Gonad histology

57. Table 5 provides a summary of the findings for the 17β-trenbolone study. Historical data on controls are described earlier in this report, under the 17β-estradiol section. As regards chemical-related abnormalities, degeneration of vitellogenic oocytes could be observed in all species, but not in all laboratories. One laboratory never observed any treatment related change.

58. Some changes were observed at 500 ng/l after 14 days of exposure already. However, abnormalities observed in exposed groups were sometimes also observed in control groups (e.g. follicular atresia in zebrafish).

59. One laboratory (LAB 4) found parasite infection with *Pleistophora spp*. in the oocytes of fathead minnow. This infection confounded interpretation of exposure-related effects.

Fish species	Laboratory	Testis after 17 β-trenbolone exposure	Ovary after 17 β-trenbolone exposure
Fathead minnow	LAB 4	- advanced maturation (smaller spermatogenic cysts and fewer spermatogenic cysts or late spermatogenic stage with a high amount of spermatozoa) after 2 weeks of treatment at 5000 ng/L and after 3 weeks of treatment at 50 ng/L and above	- atresia/degeneration of late vitellogenic follicles at 5000 ng/L after 3 weeks of treatment and relative increase of mid vitellogenic follicles at 500 ng/L and above after 2 and 3 weeks of treatment - effects obscured by parallel infection with *Pleistophora* spp. in the oocytes of the majority of animals - incipient spermatogenesis in one female at 5000 ng/L after 3 weeks of treatment
	LAB 1	- no noticeable abnormalities compared with the controls	- increased proportion of vitellogenic oocytes at 500 ng/L and above after 2 and 3 weeks - considerable increase of atretic oocytes in control and treatment groups after 2 and 3 weeks
	LAB 2	- no induced changes, except one hermaphroditic gonad after 3 weeks of treatment at 50 ng/L	- no induced changes
Medaka	LAB 1	- presence of testis-ova (1 of 5 fish at 50 ng/L after 3 weeks)	- relative increase of vitellogenic oocytes at 500 ng/L (2 of 5 fish after 2 weeks and 4 of 5 fish after 3 weeks) and at 5000 ng/L (5 fish after 2 and 3 weeks) - increase of atretic oocytes at 500 ng/L L (3 of 5 fish after 2 weeks and 4 of 5 fish after 3 weeks) and at 5000 ng/L (4 of 5 fish after 2 weeks and 5 fish after 3 weeks)
	LAB 3	- no abnormalities with increasing the exposure concentrations	- increase of ateretic follicle and degenerated ovarian mature oocytes dose-dependently after 2 and 3 weeks - decrease in degenerating of ovulated and remained oocytes in ovarian cavity dose-dependently after 2 and 3 weeks
	LAB 2	- no induced changes	- no induced changes
Zebrafish	LAB 4	- no induced changes	- increased grade of follicular atresia/degeneration after 2 weeks at 5000 ng/L and after 3 weeks at 500 ng/L and above - the incidence of this finding is relatively high in the affected groups but similar incidences are attained in some of the control groups

	LAB 1	- no noticeable abnormalities compared with the controls - presence of testis-ova in control and treatment groups	- increased proportion of vitellogenic oocytes at 500 ng/L (1 of 5 fish after 3 weeks) and at 5000 ng/L (2 of 5 fish after 2 weeks and 3 weeks) - increase of atretic oocytes at 50 ng/L (1 of 5 fish after 3 weeks) and at 5000 ng/L (2 of 5 animals after 3 weeks)
	LAB 3	- no abnormalities with increasing the exposure concentrations - presence of testis-ova in control and treatment groups	- increase of degenerated ovarian pre-mature/mature oocytes after 2 weeks at 500 ng/L and above and after 3 weeks at treatment groups - decrease in degenerating of ovulated and remained mature oocytes in ovarian cavity dose-dependently after 2 weeks
	LAB 2	- no induced changes	- no induced changes

Table 5: Summary of the evaluation of the gonad histopathology after 17 β-trenbolone exposure in all laboratories for all three fish species.

Overall findings

60. In general, participating laboratories found that a number of refinements were needed following Phase 1A to standardize the protocol and provide additional guidance where highlighted. Detailed feedback was collected from participating laboratories through a questionnaire with open-ended questions. All answers can be found in Annex II.

61. In summary, the comments received related to:

- Additional guidance needed on the measurement of secondary sex characteristics (counting and sizing) especially for the dorsal nape pad in fathead minnow which seemed difficult to quantify; laboratories asked whether a semi-quantitative analysis (e.g. in 3 to 4 classes) would be sufficient instead;
- More instructions needed on the selection of the stock population. What does "reproductively mature fish" mean exactly?
- Additional guidance on the evaluation of gonad histology is needed; more information is needed on the frequency of abnormalities in gonads of non-exposed fish (testis-ova in all species);
- Guarantee through preliminary checking of the absence of parasitism in gonads, especially of fathead minnow, is necessary, as it influences the evaluation of gonad histology;
- Time for *in situ* fixation of the gonads was too short (10 minutes); instructions need to be revised;
- The number of slides (5) per gonad seemed to be too high as little variation was observed from one section of the gonads to the next;
- Revision of the spreadsheet for the recording of gonad histology evaluation is necessary;
- Weight of the fish: a lower limit is understandable, but the higher limit (+/- 20%) was felt to be too restrictive;
- The mesh barrier was inefficient in preventing medaka (and zebrafish to a lesser extent) from spawning; after discussion at the meeting of fish experts in October it was recommended that in future validation exercise, male and female fish should be held together in the same vessel to maintain biologically relevant conditions where females are not prevented from spawning;
- The number of animals seemed to be appropriate, when no mortality occurred; but in case 2 fish out of 5 were dead, the determination of statistical significance (e.g for VTG) was not possible anymore (in the case where the Mann-Whitney *U* test with Bonferroni's adjustment was used);
- Additional guidance needed on how to prepare the blood samples for VTG measurement; how to deal with the sediments; how to avoid haemolysis; freeze-thaw of samples is not appropriate.

62. Mortality was recorded: 14 fish for 17β-trenbolone and 6 fish for 17β-estradiol; all species and all treatment groups (except the control) were affected. The total number of fish used in the Phase 1A was1800.

63. Some laboratories believed that the lower limit of body weight was useful as an indication of reproductive maturity but that a higher limit was too restrictive; however, these boundaries for the variation of the weight are useful in ensuring that the GSI does not vary substantially. Body weights were generally maintained throughout the study in each laboratory. Substantial differences in body weight across laboratories were observed for zebrafish and fathead minnow.

64. There was no treatment related change in body weight in any study or laboratory (Annex III)

65. Blood sampling for the determination of vitellogenin level did not pose any particular problem, although there was initial concern that the small size of some species may hamper the collection of a sufficient blood volume. Only in the start-up phase, a few blood samples got lost during centrifugation in

LAB 4, but this was independent from biological constraints. For medaka, the liver was used for the measurement of vitellogenin in all three laboratories.

66. Statistical significance was most of the time achieved for VTG levels which decreased significantly in females after trenbolone exposure in all three species, and increased significantly in males after estradiol exposure in all three species. For zebrafish, VTG measurements were heterogeneous, trends observed for other fish species were not equally reproduced in all four laboratories using zebrafish. Different ELISA methods were used which may have produced differences in absolute values of VTG concentration. However, this did not hamper detection of statistical significance after the same exposure duration at the medium concentration in both 17β-estradiol and 17β-trenbolone (with exceptions described in the detailed statistical analysis).

67. Evaluation of gross morphology included an evaluation of the secondary sex characteristics and the measurement of the gonado-somatic index (GSI).

68. Secondary sex characteristics exist in fathead minnow and medaka, where they respond to endocrine active substances. In fathead minnow, the number of nuptial tubercles is a male characteristic which can also be measured in females exposed to an androgenic substance; an attempt was made to measure the height of the dorsal fat pad in females exposed to trenbolone but laboratories were not successful in doing so, in the absence of clear guidance. In medaka the number of papillary processes of the anal fin was measured. The number of papillary processes in female medaka exposed to trenbolone increased significantly after 14-d exposure in all labs testing the species.

69. The GSI poorly responded in both males and females after androgenic and estrogenic exposures.

70. In gonad histology, consistent results were not always achieved in all four laboratories. Rather than differences in the species or in the sexes of fish, much discrepancy was found between evaluations themselves. In general, except in one laboratory, findings were consistent with the exposure scenario: changes were observed in testis after estrogenic exposure and in ovaries after androgenic exposure. However, one laboratory did not find any changes in any three species tested following chemical exposure, because evaluation only focused on the occurrence of testis-ova.

71. This highlights the need for a round robin evaluation of the histopathological assessment of gonadal sections in fish exposed during Phase 1A and future validation exercise. A guidance document will be prepared before starting the next validation exercise.

72. Summary of the overall results obtained in the three fish species in 17β-estradiol and 17β-trenbolone can be found in Table 6 and Table 7.

Summary of the study with 17β-Estradiol and 17β-Trenbolone

Test species	Sex	Core endpoints			
		Gross morphology		VTG	Gonad histology
		GSI	2nd. sex characters		
fathead minnow	male	+	+	+++	x
	female	+	–	+	NC
medaka	male	–	+	+++	xx
	female	–	–	++	x
zebrafish	male	+	NR	+	x
	female	+	NR	+	NC

Table 6: 17β-Estradiol

Test species	Sex	Core endpoints			
		Gross morphology		VTG	Gonad histology
		GSI	2nd. sex characters		
fathead minnow	male	–	–	+	x
	female	–	+*	++	x
medaka	male	–	–	–	NC
	female	+	+++	++	x
zebrafish	male	–	NR	–	NC
	female	+	NR	+	x

Table 7: 17β-Trenbolone

Gross morphology and VTG

+++ Statistically significant difference was found in all studies and the response was observed in a concentration-dependent manner (with sampling at 21 days).

++ Statistically significant difference was found in all studies (with sampling at 21 days).

+ Statistically significant difference was found in few studies (with sampling at 21 days).

– Statistically significant difference was not found in any studies (with sampling at 21 days).

NR No results were obtained because zebrafish have no secondary sex characteristics (with sampling at 21 days).

* Statistically significant difference was not found, however nuptial tubercles appeared in two or three studies.

Gonad histology

xx Changes were observed in all studies (with sampling at 14 and/or 21 days).

x Changes were observed in a few studies (with sampling at 14 and/or 21 days).

NC No changes were observed in all studies (with sampling at 14 and 21 days).

DISCUSSION

Overall feasibility

73. Overall, it can be concluded from Phase 1A studies that the current assay is feasible and transferable from one laboratory to another for the detection of the strong estrogenic compound 17β-estradiol and the strong androgenic compound 17β-trenbolone in all three species, according to the

definition given in the draft OECD Guidance Document no. 34 (see Glossary). In particular, the analysis of male VTG (increase) is relevant for the detection of the strong estrogen; while the measurement of female secondary sex characteristics in medaka and fathead minnow is relevant for the detection of the androgen 17β-trenbolone. The measurement of VTG decrease in female is useful in supporting the detection of androgen, but relevance is not warranted because of the underlying mechanism of action is not so well understood between VTG and androgen. Regarding the other endpoints: changes did not follow a dose-response for the GSI and not always consistent or reproducible across laboratories for gonad histology. Secondary sex characteristics in fathead minnow and medaka and gonad histology for all three species warrant further standardization before relevance and reliability can be ascertained. Detailed discussion of each endpoint, test fish and test design are described below.

Endpoints

Vitellogenn measurement (VTG)

74. Statistical significance was most of the time achieved for VTG levels which increased significantly in males after estradiol exposure in all three species and decreased significantly in females after trenbolone exposure in all three species. These results indicate that VTG was adequate for strong estrogen and responsive to androgen 17β-trenbolone. However, issues of reproducibility emerged in particular for zebrafish due to large variations in absolute values of VTG across laboratories (although similar methods were used) and also due to high basal levels in males in experiments conducted in one laboratory (LAB 4), resulting in no significant difference compared with the treatment groups. The ELISA method used was the same as in LAB 2. VTG levels in control male zebrafish were low in LAB 2 and therefore LAB 2 was able to detect significant increase at 32 ng/l 17β-estradiol after 14 days of exposure. Although no clear explanation could be found for this high level in control males, some possible explanation are *1)* difference in the sensitivity between zebrafish strains, *2)* age of fish, and *3)* differences in sampling procedures and handling of samples. In the first case, the same strains should be used in future validation exercises, avoiding strains which are not historically well known for this type of study. In the second case, tight instructions on fish age of fish is needed in future. In the third case, tighter standard operating procedures for sampling and handling are needed and will be implemented in Phase 1B. LAB 4 confirmed that diet and water could be excluded from causing high basal levels in males. In one of the comparative VTG study conducted in the United States (USEPA Comparative Evaluation of Vitellogenin Methods on medaka and zebrafish, 2003), it was hypothesized that time between sampling and analysis can cause proteolysis of VTG in smaller proteins which can then be recognized by the antibodies as many VTG proteins. However, LAB 4 added protease inhibitor (aprotinin) to blood sample, suggesting that proteolysis of VTG was not affecting VTG measurement.

75. Another problem was the differences in absolute values of VTG concentrations between kits. One laboratory (LAB 2) compared two ELISA methods for each species using the same biological samples for both estradiol and trenbolone exposures. Differences of up to two orders of magnitude in VTG levels were found between two different ELISA methods. This outcome suggests that in future validation exercise, calibration of methods will be needed through the use of a common VTG standard for each fish species. It is also recommended that methods capable of detecting VTG levels as low as ng/ml plasma and for which detailed standard operating procedures are available, should be used in the next step of the validation.

76. In zebrafish and fathead minnow, VTG concentrations were measured using blood (LAB 1 and LAB 3) and plasma (LAB 4 and LAB 2). For comparison and evaluation of the result, it is desirable that all laboratories use the unified method, described in the detailed OECD protocol.

77. In medaka, all three laboratories used liver for VTG measurement. Homogenous results were obtained in both estradiol and trenbolone studies. No particular problem was observed for VTG

measurement. These results indicated that liver is an appropriate tissue for VTG measurement in medaka. However, for transparency and reproducibility purpose, it is necessary to develop detailed standard operating procedures for liver sampling, pre-treatment and VTG measurement for the next step in the validation.

78. It was not warranted that VTG always decreases in females following androgenic exposure. Ankley et al. (2001) reported that the putative androgen methyltestosterone increased the VTG levels in adult female (and male) fathead minnows. The authors suggested that this associated with its conversion to the estrogen receptor agonist. Further comparison with published literature is needed to confirm that VTG is a relevant endpoint for androgens in fish screening assay.

Gonado-somatic index (GSI)

79. The GSI poorly responded in both males and females after androgenic and estrogenic exposures. Low sensitivity of GSI in above reference chemicals might be due to variation in measurement of gonad weight, i.e., the difficulties associated with the measurement of this endpoint, such as the in situ fixation procedure (e.g. time of fixation), the excision of the gonad and removal of the fat around it. Extension of the time for in situ fixation (i.e. full fixation of the gonad) may give us much time in excision of the gonad and perfect removal of the fat around it, resulting in decreasing variation in measurement of gonad weight. Another reason might be variation despite the fact that selected fish were same age. The selection of fish with similar reproductive status (e.g. selection of the fish with spawning) may be one solution to minimise variation of GSI. In some of the Phase 1A exposures, spawning was observed in medaka and zebrafish assays, while no females laid eggs in fathead minnow and medaka (when males and females were separated by distinct chambers). It is unclear to what extent spawning affected the GSI.

80. In any case, it appeared that the GSI was a not a good indicator of estrogenic and androgenic effects in either males or females for any of the three species under present test design.

Secondary sex characteristics

81. Secondary sex characteristics were observed in fathead minnow and medaka. This parameter poorly responded to the estrogen in both males and females. In medaka assays with trenbolone, induction of papillary processes on the anal fin of female medaka was clearly observed already after 14 days of exposure; statistical significance was detected at 500 ng/l in all three laboratories. These results indicated that papillary processes of medaka are relavant and reproducible for androgen detection. In fathead minnow assays with trenbolone, the induction of nuptial tubercles in the exposed female groups was observed in two of three studies, no change was observed in one laboratory. These results indicate that secondary sex characteristics are adequate for strong androgen in fathead minnow and medaka. The Japanese laboratories (LAB 1 and LAB 2) were not familiar with fathead minnow, and this little experience might have affect the measurement of nuptial tubercles.

82. The secondary sex characteristics are also useful for the selection of male and female. In general, secondary sex characteristics are the clearest in the active reproductive stage. The use of fish at active reproductive stage can avoid misidentification of sex in selection of the test fish. Special care is needed for selecting properly males and females zebrafish.

83. Although quantitative measurements of papillary processes in medaka and nuptial tubercles in fathead minnow seemed to be technically possible, a semi-quantification method may be envisaged as experience is gained in routine use of the assay.

84. The difficulty of measurement is an issue in this parameter, as laboratories commented (in fathead

minnow in particular). Therefore, technical lead laboratory for each fish species will be required to provide detailed guidance for evaluation in future validation exercise to improve the reproducibility of findings.

Gonad histology

85. The evaluation of gonad histology is probably the endpoint that was most challenging in the assay. In general, except for one laboratory in LAB 2, where the evaluation was focused only on the occurrence of testis-ova), findings were related to the exposure scenario: some sorts of changes were observed in testis after estrogenic exposure and in ovaries after androgenic exposure. This demonstrates a certain level of relevance for this endpoint. However, consistency and reproducibility are real concerns Inconsistencies are possibly due to the lack of clearly defined standards regarding abnormality and lesion evaluation. Therefore, reevaluation of gonad histology by histopathologists (including the histopathologists who are unrelated to Phase 1A laboratories) using the slides obtained from Phase 1A would be necessary to define clear and easy standards regarding abnormality and lesion evaluation. This will be attempted through a round-robin evaluation where histopathologists will meet after having exchanged digital photographs of lesions. Hopefully, this exercise can result in an initial and objective evaluation of the relevance and sensitivity of gonad histopathology; one of the expectations is that the outcome will result in a better guidance on how to score the lesions observed in all three species.

86. Testis-ova was observed in 32ng/L and/or 100ng/L in 17β-estradiol medaka studies (no incidence in the controls). However, on zebrafish, incidence of testis-ova in control male was observed in 2 of 4 laboratories. Since zebrafish is hermaphrodite at juvenile stage, finding of testis-ova is not an appropriate index for estrogenic chemicals in this species. In addition, no testis-ova was observed in fathead minnow in any of the three laboratories. Thus, induction of testis-ova seems to be specific to medaka.

87. In trenbolone exposures, atresia/degeneration of vitellogenic oocytes could be observed in some laboratories at 500 ng/l and higher after 14 days of exposure already. However, incidences observed in treatment groups were sometimes also observed in control groups. The fact that females could not (normally) spawn may explain the occurrence of incidences in treatment groups and controls and also could compromise utility of gonadal histology endpoints in the non-spawning design.

88. One laboratory commented that five sections per gonad for the examination were unnecessary because sections of the same gonad do not differ much; this was based on their own findings. In other laboratories which used five sections per gonad, differences were sometimes observed within the gonads in all species. Although parasite infection (*Pleistophora* spp.) was observed in the gonad of fathead minnow in one laboratory (LAB 4), it was not clear that this infection affected the gonad.

89. In Phase 1A design, ovaries poorly responded to 17β-estradiol. The published studies in fathead minnow and medaka have consistently shown alterations in ovarian histology (25)(14)(38)(39). This suggests that the ovaries of non-spawning animals are comparatively less responsive to the effects of strong estrogens than reproductive active fish.

Test fish

90. In Phase 1A, instructions for body weight at the start of the assay were not systematically followed. Some of the laboratories proposed the revision of the OECD protocol on the body weight, age and reproductive status for the test fish. One idea is that males and females that are fully mature and at reproductive status are selected for the test fish in Phase 1B; body weight and age are additional guidance.

Test design

Exposure duration and test chemical

91. Regarding VTG concentration, effects of estradiol and trenbolone were detectable after 14 days of exposure. In secondary sex characteristics, masculinization of females was observed in medaka and fathead minnow after 14 days of trenbolone exposure. The results of gonad histology indicated that some of changes were found after 14 days of exposure in estradiol and trenbolone. The GSI poorly responded both after 14 and 21 days of exposure in both estradiol and trenbolone. These results suggest that the effects of strong estrogen and androgen are generally detectable after 14 days of exposure. However, the minimum duration of exposure for the reliable detection of weak estrogen, anti-androgen or aromatase inhibitor is not clear.

Separation of males and females and spawning issue

92. Spawning was noted for medaka and zebrafish even when males and females were separated by a mesh barrier within the test vessel. This suggests that separation of males and females by a mesh barrier is not efficient to suppress oviposition in these two species. Furthermore, in several instances, such separation proved to be biologically irrelevant (e.g. abnormal histological responses were noted in control animals 2 to 3 weeks after test initiation in fathead minnow and zebrafish). Therefore, exposure of males and females in the same vessel should be proposed for future validation exercise.

Statistical analysis

93. In Phase 1A, statistical analysis with two-side test was performed because the initial hypothesis was that either an increase or a decrease in endpoints measured could be considered as damageable outcomes.

94. The present OECD protocol is not suitable for non-parametric analysis. However, such analysis is adequate and yields good power in some cases and certain test design. For instance, non-parametric analysis seems to be adequate when no homogeneity is observed in the transformed data. In Phase 1A, the VTG values, lower than the determination limit, were transformed to half the value of the determination limit for each analysis (agreement among four participating laboratories). There seems to be no problem using half the value of the determination limit for statistical analysis, and description of the method in the OECD protocol may be necessary (based on the discussion in FDG). When half the values of the determination limit are used in statistical analysis, non-parametric analysis may be adequate. However, there is a demerit in use of non-parametric analysis, i.e., statistically significant difference is not observed in this analysis when the group size is small. In fact, partial mortality was observed in some exposure studies in Phase 1A, and no significant difference was detected in a few endpoints of these studies because of small sample size (e.g. statistically significant difference is never found in Mann-Whitney U test with Bonferroni's adjustment if the group size is 3 in either of two data sets). Therefore, an evaluation of optimum group size and the number of replicate vessels will be needed, when more data are available to characterize the size of the response following diverse types of chemical exposures and response scenarios.

95. ANOVA/Kruskal-Wallis test followed by Dunnett's multiple comparison test/Mann-Whitney U test with Bonferroni's adjustment were performed in Phase 1A. However, the performance of ANOVA/Kruskal-Wallis test is arguable because of less severity in comparison to the tests like Dunnett's- or Williams-Test. The choice of the most appropriate statistical test(s) needs further consideration.

GLP

96. One laboratory (LAB 4) provided a detailed study plan before initiating experimental work, following GLP requirements and detailed reports for histopathological findings. In Phase 1B, laboratories will be encouraged to provide their experimental protocol in advance, and to perform assays under GLP to ensure that proper methodologies are used and agreed upon before starting experimental work, and to ensure that data are checked for quality assurance prior to submission to the lead laboratory.

Summary

97. It is concluded from Phase 1A study that the present OECD protocol is applicable to detect the effects of strong estrogen 17-β estradiol and androgen 17-β trenbolone. Analysis of male VTG was relevant and reproducible for the strong estrogen 17β-estradiol and measurement of secondary sex characteristics in medaka and fathead minnow females was responsive and almost reproducible in all laboratories, following the 17β-trenbolone exposure. GSI was poorly responsive to both 17β-estradiol and 17β-trenbolone. In gonad histology, reproducibility could not be verified due to the lack of standardization of criteria for evaluation. Further guidance on the secondary sex characteristics and gonad histology is necessary for Phase 1B.

RECOMMENDATIONS

Selecting test fish

98. Additional guidance from the technical lead laboratory for each fish species is required for breeding condition, state of maturation, method for selecting male and female etc. to prepare a Phase 1B OECD protocol. Detailed descriptions regarding the above should be included in the OECD protocol. In order to minimize error when selecting males and females, an idea is to use male and female fish with clear secondary sex characteristics and/or in reproductive status. Although there were no technical lead laboratories for fathead minnow and zebrafish in Phase 1A, it is recommended that a technical lead laboratory for each fish species collects the background data on each endpoint in breeding and acclimation group before Phase 1B, and delivers guidance to other participants as necessary.

99. It is recommended that fish are maintained in clean dilution water in laboratory at least for the generation preceding the assay. It is also recommended that histopathological examination of viscera (including gonad) be performed using a sub-sample of fish before the exposure, especially if the fish is purchased from a commercial breeder. This will prevent from using fish hosting parasites.

VTG

100. In Phase 1B, all laboratories should measure VTG concentrations using a uniform method for sampling (blood and liver) and pretreatment. For medaka, liver is recommended for use (LAB 1 will prepare a guidance of the concrete procedure). For fathead minnow and zebrafish, it may be necessary to determine and agree on the use of the most appropriate tissue-source for VTG measurement, and ensure that all laboratories are using the same source and measurement method.

GSI

101. The Fish Drafting Group concluded that GSI would not be measured as an endpoint in Phase 1B due to its lack of responsiveness in Phase 1A.

Secondary sex characteristics

102. The reproducibility of the endpoint of secondary sex characteristics should be checked again in Phase1B. Technical lead laboratory for fathead minnow and medaka should prepare additional guidance for the measurements of secondary sex characteristics. Although quantitative evaluation is desirable for the number of papillary processes on anal fin in medaka and nuptial tubercles in fathead minnow, semi-quantitative method might also be appropriate for the estimation in consideration of the measurement time.

Gonad histology

103. It is recommended that the gonad be removed after complete fixation with prolongation of in situ fixation. As for the problems of gonad histology, it is desirable to judge based on the results of round-robin evaluation. After the criteria of diagnosis are created in the evaluation, the applicability of the endpoint of gonad histology should be discussed again in Phase 1B or before Phase 1B if possible.

Others

104. The exposure duration should be discussed based on the results of Phase 1B using weak estrogen, anti-androgen, aromatase inhibitor and so on. Therefore, the duration of the assay in Phase 1B should be the same as in Phase 1A, i.e. 21-d.

105. The final evaluation for the applicability of each core endpoint should be validated based on the results of Phase 1B using weak estrogen, anti-androgen, aromatase inhibitor. Therefore, the three core endpoints gross morphology (secondary sex characteristics), VTG and gonad histology should be used again in Phase 1B. Only for gonad histology, the number of sections for observation should be discussed beforehand.

106. Exposure of males and females in the same tank will be recommended in Phase 1B, to increase the biological relevance of the assay.

107. Following Phase 1B, a detailed statistical analysis is warranted to advise on the power properties of the assay and any action required to increase the power to an acceptable level, considering the data properties and response profiles, the group size and the replication.

108. In order to assure the quality of the data, all participating laboratories should prepare their test protocol, and if possible, it is recommended that the assays be conducted under GLP principle.

109. In Phase 1A, there were no technical lead laboratories for fathead minnow and zebrafish. A technical lead laboratory for these species will be designated for Phase 1B.

ACKNOWLEDGEMENT

118. The dedicated efforts and work of the lead laboratory (LAB1) and especially the participating laboratories which generated the data are acknowledged. Also many thanks to Dr. Yukio Kawashima from Japan for his statistical and computational support.

ANNEXURES

ANNEX I

OECD Protocol for Phase 1A of the Validation of the Fish Screening Assay

ANNEX II

Comments from participating laboratories on Phase 1A of the Validation
(***Available from the OECD Secretariat upon request***)

ANNEX III

Figures showing test results for 17 β-estradiol and 17 β-trenbolone studies
(***Available from the OECD Secretariat upon request***)

ANNEX IV

Summary statistical tables
(***Available from the OECD Secretariat upon request***)

ANNEX I

PROTOCOL FOR PHASE 1A OF THE VALIDATION

Non-Spawning Fish Screening Assay for Endocrine Disrupters

INTRODUCTION

1. This protocol describes an *in vivo* screening assay for identifying endocrine active chemicals in sexually dimorphic fish. The non-spawning fish assay normally exposes fish to chemicals for up to 21 days. The *in vivo* non-spawning fish screening assay is intended to detect chemicals that affect androgenic or estrogenic activity in fish exposed during a limited part of their life-cycle, not including reproduction. The assay is not designed to identify specific mechanisms of hormonal disruption (such as hormone receptor agonistic or antagonistic effects, or disturbance of hormone synthesis or metabolism). The concept for this protocol is derived from work on the fathead minnow *(Pimephales promelas)* (1)(2)(3)(37), the Japanese medaka (*Oryzias latipes*) (4) (5) (6) and the zebrafish (*Danio rerio*) (7) (8) (9) (10). There is limited evidence that the protocol could also be used for the three spined sticleback (*Gasterosteus aculatus*) (11).

2. This protocol addresses a maximum 21-day chemical exposure period (with no *in situ* pre-exposure period) and measures three core biomarker endpoints as indicators of endocrine disrupter activity, namely: i) gross morphology (eg secondary sexual characteristics and gonadal somatic index (GSI)); ii) vitellogenin levels and iii) gonadal histology (OECD Fish Expert Consultation, Tokyo, 2000). To address all core endpoints, sexually dimorphic adult fish are normally used (pre-breeding animals should be used wherever possible).

PRINCIPLE OF THE ASSAY

3. Adult fish as test organisms. Overviews of the relevant bioassay conditions are provided in Appendix 1 to this protocol. The assay is normally initiated with fish sampled from populations that are in spawning condition. Where necessary (with expert technical judgment) non-spawning adult fish that are almost ready to breed may also be selected from pre-spawning populations, however, senescent animals should not be used. The assay is conducted using three chemical exposure concentrations, as well as a water control, with two vessels per treatment (one vessel containing 10 males and the other vessel containing 10 females). To be environmentally relevant, chemical delivery of 17-β estradiol and 17-β trenbolone should be via an aqueous route without carrier solvents. The exposure is conducted for 21-days. On sampling after 14 and 21-days, 10 fish (5 males and 5 females) per treatment are killed humanely and blood samples are collected for determination of vitellogenin (note - as an alternative to plasma collection, liver tissues may be sampled for VTG analysis). The gonads are also fixed *in situ*, promptly wet weighed for calculation of the gonadal-somatic index (GSI) and then processed for subsequent histological analyses (see Appendix 3).

INFORMATION ON THE TEST SUBSTANCE

4. Characterisation of the test substances is the responsibility of the original chemical supplier and those managing the chemical repository. It is not the responsibility of the lead or participating laboratories. Test substances from the same batch will be supplied to each participating laboratory by the chemical repository. The test substances 17β-oestradiol and 17β-trenbolone will be characterised by name, supplier,

batch number, structural formula, purity, appearance, stability in water and light, pK_a and P_{ow}, storage conditions, expiry date.

TEST ACCEPTANCE CRITERIA

5. For the test results to be acceptable the following conditions apply:

- the mortality in the control(s) does not exceed 10 per cent at the end of the exposure period;
- the dissolved oxygen concentration has been at least 60 per cent of the air saturation value (ASV) throughout the exposure period.
- the water temperature did not differ by more than ± 1 °C between test vessels at any one time during the exposure period and was maintained within a range of 2°C within the temperature ranges specified for the test species (Appendix 1).

DESCRIPTION OF THE METHOD

Apparatus

6. Normal laboratory equipment and especially the following:

(a) oxygen and pH meters;
(b) equipment for determination of water hardness and alkalinity;
(c) adequate apparatus for temperature control and preferably continuous monitoring;
(d) tanks made of chemically inert material and of a suitable capacity in relation to the recommended loading and stocking density (see Appendix 1);
(e) suitably accurate balance (i.e. accurate to ± 0.5%).

Water

7. Any water in which the test species shows suitable long-term survival and growth may be used as test water. It should be of constant quality during the period of the test. The pH of the water should be within the range 6.5 to 8.5, but during a given test it should be within a range of ± 0.5 pH units. Dilution water hardness should be above 140 mg/l (as $CaCO_3$) is recommended. In order to ensure that the dilution water will not unduly influence the test result (for example by complexion of test substance), samples should be taken at intervals for analysis. Measurements of heavy metals (e.g. Cu, Pb, Zn, Hg, Cd, Ni), major anions and cations (e.g. Ca, Mg, Na, K, Cl, SO_4), pesticides (e.g. total organophosphorus and total organochlorine pesticides), total organic carbon and suspended solids should be made, for example, every three months where a dilution water is known to be relatively constant in quality. If water quality has been demonstrated to be constant over at least one year, determinations can be less frequent and intervals extended (e.g. every six months). Some chemical characteristics of acceptable dilution water are listed in Appendix 2.

Test solutions

8. Test solutions of the chosen concentrations are prepared by dilution of a stock solution. The stock solution should preferably be prepared by simply mixing or agitating the test substance in the dilution water by using mechanical means (e.g. stirring or ultrasonication). Saturation columns (solubility columns) can be used for achieving a suitable concentrated stock solution. The use of solvents will not be necessary in Phase 1 with 17β-oestradiol and 17β-trenbolone.

9. A flow-through test system will be used. Such a system continually dispenses and dilutes a stock solution of the test substance (e.g. metering pump, proportional diluter, saturator system) in order to deliver a series of concentrations to the test chambers. The flow rates of stock solutions and dilution water should be checked at intervals, preferably daily, during the test and should not vary by more than 10% throughout the test. Care should be taken to avoid the use of low-grade plastic tubing or other materials that may contain biologically active substances.

Selection of species

10. The exposure phase will be started with sexually dimorphic adult fish from a laboratory supply of reproductively mature animals (ideally, actively breeding fish). Where necessary (with expert technical judgment) non-spawning adult fish that are almost ready to breed may also be selected from pre-spawning populations, however, senescent animals should not be used. Selected species for this assay are the fathead minnow (*Pimephales promelas*), Japanese medaka (*Oryzias latipes*) and zebrafish (*Danio rerio*).

11. Test fish shall be selected from a laboratory population of a single stock, preferably from the same spawning, which has been acclimated for at least two weeks prior to the test under conditions of water quality and illumination similar to those used in the test (note, this acclimation period is not an *in situ* pre-exposure period). Fish should be fed *ad libitum* throughout the holding period and during the exposure phase. Note-feed should not be fed within 12 hours of necropsy.

12. Following a 48-hour settling-in period, mortalities are recorded and the following criteria applied:

- mortalities of greater than 10% of population in seven days: reject the entire batch;
- mortalities of between 5% and 10% of population: acclimation for seven additional days; if more than 5% mortality during second seven days, reject the entire batch;
- mortalities of less than 5% of population in seven days: accept the batch.

13. Fish should not receive treatment for disease in the two week acclimation period preceding the test, or during the exposure period.

TEST DESIGN

14. Three concentrations of the chemical are used per experiment. The data may be analyzed in order to define the Lowest Observed Effect Concentration or No-Observed Effect Concentration based on endocrine biomarker endpoints (the $LOEC_{biomarker}$ and $NOEC_{biomarker}$, respectively). Calculation of these statistical parameters will be useful in order to establish whether any further longer term testing for adverse effects (namely, survival, development, growth and reproduction) is required for the chemical.

15. At initiation of the experiment on day-0, 5 males and 5 females from the non-exposed population are sampled for the measurement of the three core endpoints. After 14 days of exposure, 5 males and 5 females from each treatment level and from the control are sampled again for the measurement of the three endpoints. Finally, at termination of the assay after 21 days of exposure, the 5 remaining males and 5 remaining females are killed for the measurement of the three endpoints (see Appendix 3 for diagram of test design).

16. For the purpose of acquiring adequate background data, the lead laboratories for each species will perform gonad histology on an additional group of 40 fish (20 male and 20 female) from the same culture as the fish used for the assay. This additional task is necessary to collect information on the normal status of gonads and on the natural variability in gonad histology in each species. Other laboratories (apart from the technical leads) may also carry out this task if they wish.

PROCEDURE

Selection and weighing of test fish

17. It is important to minimise variation in weight of the fish at the beginning of the non-spawning fish assay. Suitable size ranges for the different species recommended for use in this test are given in Appendix 1 to this protocol. For the whole batch of fish used in the test, the range in individual weights at the start of the test should be kept to within ± 20% of the arithmetic mean weight. It is recommended to weigh a subsample of fish before the test in order to estimate the mean weight.

Conditions of exposure

Duration

18. The test duration is up to 21 days (with sampling at days 0, 14 and 21).

Loading rates and stocking densities

19. It is important that the loading rate and stocking density (for definitions, see Appendix 4) is appropriate for the test species used (see Appendix 1). If the stocking density is too high, then overcrowding stress will occur leading to reduced growth rates and possibly to disease. If it is too low, territorial behavior may be induced which could also affect growth. In any case, the loading rate should be low enough in order that a dissolved oxygen concentration of at least 60% ASV can be maintained without aeration.

Feeding

20. The fish should be fed *ad libitum* with an appropriate food (Appendix 1) at a sufficient rate to maintain body condition (2). Care should be taken to avoid microbial growth and water turbidity. The daily ration may be divided into two or three equal portions for multiple feeds per day, separated by at least three hours between each feed. The ration is based on the initial total fish weight for each test vessel. Food should be withheld from the fish for 12 hours prior to the day of sampling (day 0, day 14 and day 21) to aid in histology processing of small fish.

21. Fish foods should be evaluated for the presence of contaminants such as organochlorine pesticides, polycyclic aromatic hydrocarbons (PAHs) and polychlorinated biphenyls (PCBs).

22. Uneaten food and fecal material should be removed from the test vessels each day by carefully cleaning the bottom of each tank using suction.

Light and temperature

23. The photoperiod and water temperature should be appropriate for the test species (see Appendix 1).

Test concentrations and controls

24. Concentrations of the test substances in Phase 1 will be as follows:

- for 17β-oestradiol: 10, 32, 100 ng/l (+ water control), and
- for 17β-trenbolone 50, 500 and 5000ng/l (+ water control).

Frequency of analytical determinations and measurements

25. During the test, the concentrations of the test substance are determined at regular intervals, as follows: the flow rates of diluent and toxicant stock solution should be checked at intervals, at least twice per week, and should not vary by more than 15% throughout the test. It is recommended that the actual test chemical concentrations be measured in all vessels at the start of the test and at weekly intervals thereafter.

26. Results should always be based on measured concentrations and included in reporting.

27. Samples may need to be filtered (e.g. using a 0.45 μm pore size) or centrifuged. Centrifugation is the recommended procedure. However, if the test material does not adsorb to filters, filtration may also be acceptable.

28. During the test, dissolved oxygen, pH and temperature should be measured in all test vessels at least once per week. Total hardness, alkalinity and salinity should be measured in the controls and one vessel at the highest concentration at least once per week. Temperature should preferably be monitored continuously in at least one test vessel.

Observations

29. A number of general (eg survival) and core biological responses are assessed over the course of the non-spawning fish assay. Collection of these endpoints and their utility are described below:

Survival

30. Fish should be examined daily during the test period and any external abnormalities (such as hemorrhage, discoloration) noted. Any mortality should be recorded and the dead fish removed as soon as possible. Dead fish should not be replaced in either the control or treatment vessels.

Behaviour

31. Any abnormal behavior (relative to controls) should be noted; this might include signs of general toxicity including hyperventilation, uncoordinated swimming, loss of equilibrium, and atypical quiescence or feeding. Such behavioral observations may provide useful qualitative information to inform potential future fish testing requirements (for example, territorial aggressiveness in normal males or masculinized females has been observed in fathead minnows under androgenic exposure).

Appearance

32. Secondary sexual characteristics are under endocrine control; therefore observations of physical appearance of the fish should be made over the course of the test, and at conclusion of the study. Experience to date with fathead minnows suggests some endocrine active chemicals may initially induce changes in the following external characteristics: body color (light or dark), coloration patterns (presence of vertical bands), body shape (head and pectoral region), and specialized secondary sex characteristics (size of dorsal nape pad, number of nuptial tubercles in males; ovipositor size in females). Notably, chemicals with certain modes of action may cause abnormal occurrence of secondary sex characteristic in animals of the opposite sex; for example, androgen receptor agonists, such as methyltestosterone and dihydrotestosterone, can cause female fathead minnows to develop pronounced nuptial tubercules (3)(12) (13). It also has been reported that estrogen receptor agonists can decrease nuptial tubercle numbers and size of the dorsal nape pad in adult males (14) (15). Such gross morphological observations may provide useful qualitative information to inform potential future fish testing requirements.

33. Because some aspects of appearance (primarily color) can change quickly with handling, it is important that qualitative observations be made prior to removal of animals from the test system. This type of qualitative assessment might be enhanced through the use of photographs or videotape. Other endpoints, such as the number and size of nuptial tubercles, can be quantified directly or in preserved specimens (3).

Humane killing of fish

34. At day 0, day 14 and at day 21 at conclusion of the exposure (see also Appendix 3), the fish should be anaesthetized with MS-222 (100 mg per L buffered with 200 mg $NaHCO_3$ per L), individually weighed as wet weights (blotted dry) and blood collected from the caudal artery/vein with a heparinised microhematocrit capillary tubule. Depending upon size of the fish, collectable blood volumes generally range from 20 to 80 µl per individual for fathead minnows (1) (16) and 5-15 µl per individual for zebrafish (17). Plasma is separated from the blood via centrifugation (3 min; 15,000 g; room temperature), and stored with protease inhibitors at -80^{o}C, until analyzed for vitellogenin. Alternatively, liver could be used in vitellogenin analysis, depending on the fish species studied; in such case weighing of the fish should be made prior to removal of the liver.

Relative Gonad Weight

35. After sampling the blood, fish should be weighed, the gonads fixed *in situ* (see paragraph 39), then removed and weighed (to the nearest 0.001 g) for determination of the gonado-somatic index (GSI=100 × gonad wt/body wt). For example, typical GSI values for adult fathead minnows range from 7 to 14% for females, and 1 to 2% for males (16)(18). Many chemicals that reduce fecundity also will reduce the GSI in one or both sexes (3)(19). After removal of the fixed gonads, the remainder of the carcass of the fish may be discarded. Optionally, fish could be preserved for further measurements (e.g., frozen for chemical analysis).

Gonad Histology

36. The first step of gonad histological analysis is necropsy and rapid fixation of the gonad to prevent autolysis and cellular deterioration. Immediately after humane killing of an individual fish (plus length and weight measurements and collection of fresh tissues – eg plasma for VTG analysis – the abdomen should be opened and the paired gonads fixed *in situ*. Suitable fixatives are 4% formaldehyde or 1% glutaraldehyde (Bouin's fixative may also be used), all of which are ideal for maintaining the structural integrity of the gonad. After a short period of a few minutes *in situ* fixation, the gonads are carefully removed and wet weighed (for calculation of GSI) then transferred to a suitable labeled container containing fresh fixative (as guidance, the fixative: tissue volume ratio should be ≥10:1). It should be noted that in rare cases, some fish naturally lack a fully developed pair of gonads and only one gonad may be present. All such observations must be recorded.

37. The fixed gonads should be processed and embedded in paraffin wax. Using standard histological techniques (20), microtome cut histological sections should be taken along the longitudinal axis of the gonads in a serial-step fashion (for example, at 10-20 µm intervals). Two longitudinal sections per gonad should be collected from each fish.

38. After staining with haematoxylin and eosin, histological sections should be evaluated by an experienced pathologist without prior knowledge of the treatment regime associated with specific samples. In general, evaluation of the testis is based on the amount of germinal epithelium present, and the degree of spermatogenic activity. The ovary is evaluated based upon relative numbers of perinucleolar, cortical alveolar, and vitellogenic oocytes (21). Classification schemes potentially useful for assessing fish gonads are summarized elsewhere (3)(22). Recent studies have documented a variety of alterations in fathead

minnow gonadal histology associated with exposure to endocrine active chemicals with oestrogenic or androgenic properties (3) (14) (23) (24) (25) (26).

39. The gonadal histology observations will be interpreted based on the expert judgement of a professional pathologist. It is not necessary to evaluate the histological sections using extensive image-analysis techniques with associated statistical analyses; however, in some cases such techniques may provide useful supplemental information. The pathologist's interpretation of the material should be fully documented in the study report.

40. In addition, as indicated in paragraph 16, a group of 40 fish from the same culture (20 males and 20 females), will be sampled by the technical lead laboratories for each fish species, to perform gonad histology. In this way, additional information will be collected on the normal status of gonads and the natural variability in gonad histology in each species.

Vitellogenin (VTG)

41. VTG is a phospholipoglycoprotein precursor to egg yolk protein that normally occurs in sexually-active females of all oviparous species; the production of VTG is controlled by interaction of oestrogens with the estrogen receptor. Significantly, males maintain the capacity to produce VTG in response to stimulation with estrogen receptor agonists; as such, induction of VTG in males and immature females has been successfully exploited as a biomarker specific for oestrogenic compounds in a variety of OECD fish species (1)(2) (3) (28) (29)(30) (31) (32) (33)

COMMENT:
Once the results are available from the US EPA's work on VTG assays, this text may need to be extended to address the specific ELISA method(s), the source of reagents, the type of antibodies used, the intra- and inter-assay CVs and the supporting references.

42. The measurement of VTG should be based upon a quantitative and validated method. Information should be available upon the intra-assay and inter-assay variability of the method used in a given laboratory. For example, using an enzyme-linked immunoabsorbant assay (ELISA) raised to carp (*Cyprinus carpio*) VTG and applied to fathead minnow VTG, coefficients of variation for these values were below 10%, while the inter-laboratory coefficient of variation for juvenile fathead minnow VTG was 34% (34). Similarly, Fenske *et al* (35) reported a competitive ELISA for adult zebrafish plasma VTG, with intra- and inter-assay coefficients of variation of 7.5% and 4.9%, respectively. Further examples of VTG measurements include: fathead minnow (3) (30) (36); medaka (32) (33); zebrafish (15) (17).

43. Different methods are available to assess vitellogenin production in different fish species considered in this assay; a measurement technique that is both relatively sensitive and specific is determination of protein concentrations in plasma via enzyme-linked immunoabsorbant assay (ELISA). For ELISA, polyclonal fathead minnow vitellogenin antibody and purified vitellogenin protein also from the fathead minnow are utilized (28) (29). Polyclonal and/or monoclonal vitellogenin antibodies prepared using protein from other fish species may cross-react with fathead minnow vitellogenin and, hence, also could be useful for assessing this endpoint (30)(36). In either case of homologous immunoassay (antibodies to fathead minnow vitellogenin) or heterologous immunoassay (antibodies raised to vitellogenin from a related fish species) validation of a quantitative method is essential to obtain meaningful results. One challenge to the implementation of the fish endocrine screening assay for routine use will be standardization of the measurement of vitellogenin; ideally, both a standard (perhaps monoclonal) antibody and protein (vitellogenin) standards should be broadly available.

Optional endpoint: measurement of gonadal steroids

44. Laboratories that want to measure sex steroids concentration on plasma can do so. Standard operating procedures are available upon request from the lead laboratory.

DATA AND REPORTING

Treatment of results

45. Specific guidance on statistical analyses of biological responses is not given here. To guide on the type of data that should be provided, a proposal for a data template can be found in Appendix 5. All laboratories will use the same template for data collection. It is the responsibility of the lead laboratory to ensure that the template adequately captures all necessary data, and that all laboratories are using it.

Evaluation of Biomarker Responses by Analysis of Variance (ANOVA)

46. To identify potential endocrine activity by a chemical, biomarker responses (including GSI, VTG) are compared between treatments versus controls groups using analysis of variance (ANOVA). If the required assumptions for parametric methods are not met - non-normal distribution (e.g. Shapiro-Wilk's test) or heterogeneous variance (Bartlett's test or Levene's test), consideration should be given to transforming the data to homogenize variances prior to performing the ANOVA, or to carrying out a weighted ANOVA. All biological response data should be analyzed and reported separately by sex (either females or males) where possible.

Evaluation of Biomarker Responses by Regression Analysis

47. To identify potential endocrine activity by a chemical, biomarker responses (including GSI, VTG) across the different treatment concentrations should be examined based on regression analysis (based on three exposure concentrations per chemical). All biological response data should be analyzed and reported separately by sex (either females or males) where possible. Direct observation of concentration-related trends in biological responses can also usefully inform the outcome of such a study for the intrinsic endocrine activity of a test chemical.

Interpretation of results

48. The results should be interpreted with caution where measured toxicant concentrations in test solutions occur at levels near the detection limit of the analytical method.

49. Different options are available for data analysis; these are comprised primarily of routine hypothesis testing procedures, based on ANOVA, and are described elsewhere. For the results of the non-spawning fish assay to be acceptable, the following performance criteria should be met:

Test report

50. The test report must include the following information:

Test substance: physical nature and relevant physical-chemical properties, chemical identification data including purity and analytical method for quantification of the test substance where appropriate.

Test species: at a minimum scientific name, supplier and any pretreatment.

Test conditions: test procedure used (test-type, loading rate, stocking density, etc.); method of preparation of stock solutions and flow-rate; the nominal test concentrations, the means of the measured values and standard deviations in the test vessels and the method by which these were attained and evidence that the measurements refer to the concentrations of the test substance in true solution; dilution water characteristics (including pH, hardness, alkalinity, temperature, dissolved oxygen concentration, residual chlorine levels, total organic carbon, suspended solids, salinity of the test medium (if measured) and any other measurements made); water quality within test vessels: pH, hardness, temperature and dissolved oxygen concentration; detailed information on feeding (e.g. type of food(s), source, amount given and frequency and analyses for relevant contaminants (eg PCBs, PAHs and organochlorine pesticides);

Results: evidence that controls met the validity criterion for survival, and data on mortalities occurring in any of the test concentrations; statistical analytical techniques used, statistics based on fish, treatment of data and justification of techniques used; tabulated data on biological observations of gross morphology (including GSI) and vitellogenin; detailed report on gonadal histology; results of the statistical analysis preferably in tabular and graphical form; incidence of any unusual reactions by the fish and any visible effects produced by the test substance.

REFERENCES

1. Panter GH, Thompson RS, Sumpter JP. 1998a. Adverse reproductive effects in male fathead minnows (*Pimephales promelas*) exposed to environmentally relevant concentrations of the natural oestrogens, oestradiol and oestrone. *Aquat Toxicol* 42:243-253.

2. Panter GH, Campbell PM, Hutchinson TH, Kime DE, Länge R, Lye C, McCahon P, Sumpter JP, Tattersfield LJ, Tyler CR. 1998b. Effects of endocrine disruptors in the fathead minnow (*Pimephales promelas*). *Abstracts*, 19th Annual Meeting, Society of Environmental Toxicology and Chemistry, Charlotte, NC, USA.

3. Ankley, G.T., Jensen, K.M., Kahl, M.D., Korte, J.J., Makynen, E.A. (2001) Description and evaluation of a short-term reproduction test with the fathead minnow (*Pimephales promelas*). *Environmental Toxicology & Chemistry*, 20, 1276-1290.

4. Gray MA and Metcalfe CD (1999). Toxicity of 4-*tert*-octylphenol to early life stages of Japanese medaka (*Oryzias latipes*) *Aquat Toxicol* 46(2):149-154

5. Groonen S, Denslow N, Manning S, Barnes S, Barnes D and Brouwer M (1999) Serum vitellogenin levels and reproductive impairment of male Japanese medaka (*Oryzias latipes*) exposed to 4-tert-octylphenol. *Environ Hlth Perspect* 10: 385-390

6. Shioda T and Wakabayashi M (2000) Effect of certain chemicals on the reproduction of medaka (*Oryzias latipes*) *Chemosphere* 40(3): 239-243

7. Bresch H, Beck H, Ehlermann D, Schlaszus H, Urbanek M (1990) A long-term test comprising reproduction and growth of zebrafish (*Danio rerio*) with 4-chloroaniline. *Arch Environ Contam Toxicol* 19: 419-427

8. Andersen, L., Bengtsson, B.E., Björk, M., Gessbo, Å., Holbech, H., Hylland, K., Norrgren, L., Pedersen, K.L., Lundgren, A., Petersen, G.I., Steinholz, Å. and Örn, S. (2000) Zebrafish for testing

endocrine disrupting chemicals. TemaNord 2000:555. ISBN:92-893-0480-4.

9. Andersen, L., Gessbo, Å., Holbech, H. Norrgren, L., Petersen, GI. (2001). Effects of exposure during early development to 17α-ethinylestradiol on the sexual differentiation and induction of vitellogenin in zebrafish (*Danio rerio*). *Aquatic Toxicology*, submitted.

10. Van den Belt K, Verheyen R and Witters H (2001) Reproductiv effects of ethynylestrdadiol and 4t-octylphenol on the Zebrafish (Danio rerio) *Arch Environ Contam Toxicol* 19: 419-427

11. Katsiadaki I, Scott AP, Mayer I. The potential of the three-spined stickleback (Gasterosteus aculeatus L.) as a combined biomarker for oestrogens and androgens in European waters. Mar Environ Res 2002 Sep- Dec;54(3-5):725-8

12. Smith, RJF. 1974. Effects of 17α-methyltestosterone on the dorsal pad and tubercles of fathead minnows (*Pimephales promelas*). *Can J Zool* 52:1031-1038.

13. Panter GH, Hutchinson TH, Hurd KS, Sherren A, Stanley RD, Tyler CR. (2003). Successful detection of (anti-) androgens and aromatase inhibitors using a sub-adult fathead minnow non-spawning assay. CEFIC-EMSG research report reference AQ006. CEFIC, Brussels, Belgium.

14. Miles-Richardson SR, Kramer VJ, Fitzgerald SD, Render JA, Yamini B, Barbee SJ, Giesy JP. 1999b. Effects of waterborne exposure of 17â-estradiol on secondary sex characteristics and gonads of fathead minnows (*Pimephales promelas*). *Aquat Toxicol* 47:129-145.

15. Holbech H, Andersen L, Petersen GI, Korsgaard B, Pedersen KL, Bjerregaard P. 2001. Development of an ELISA for vitellogenin in whole body homogenate of zebrafish (*Danio rerio*). Comp. Biochem. Physiol. C. 130: 119-131.

16. Jensen KM, Korte JJ, Kahl MD, Pasha MS, Ankley GT. 2001. Aspects of basic reproductive biology and endocrinology in the fathead minnow (*Pimephales promelas*). *Comp Biochem Physiol* 28C(1): 127-141

17. Van den Belt K, Wester PW, Van der Ven LT, Verheyen R, Witters H. Effects of ethynilestradiol on the reproductive physiology in zebrafish (*Danio rerio*): time dependency and reversibility. *Env Toxicol Chem*, 21, 4, April 2002, pp. 767-775

18. Panter GH, Hutchinson TH, LängeR, Whale G, Sumpter JP, Tyler CR. 2001. Quantifying gonadal development in sexually maturing fathead minnows (*Pimephales promelas*). CEFIC-EMSG research report reference AQ004. CEFIC, Brussels, Belgium.
19. Makynen EA, Kahl MD, Jensen KM, Tietge JE, Wells KL, Van Der Kraak G, Ankley GT. 2000. Effects of the mammalian anti-androgen vinclozolin on development and reproduction of the fathead minnow (*Pimephales promelas*). *Aquat Toxicol* 48:461-475.

20. Roberts RJ (1978). Fish Pathology. Baillere Tindall, London. ISBN 0 7020 0674.

21. West G (1990) Methods of assessing ovarian development in fish: a review. *Aust J Mar Freshwater Res* 41: 199-222

22. Wester PW and Vos JG (1994) Toxicological pathology in laboratory fish: an evaluation with two species & various environmental contaminants. *Ecotoxicology* 3: 21-44

23. Miles-Richardson SR, Pierens SL, Nichols KM, Kramer VJ, Snyder EM, Snyder SA, Render JA, Fitzgerald SD, Giesy JP. 1999a. Effects of waterborne exposure to 4-nonylphenol and nonylphenol ethoxylate on secondary sex characteristics and gonads of fathead minnows (*Pimephales promelas*). *Environ Res* 80:122-137.

24. Harries JE, Runnalls T, Hill E, Harris C, Maddix S, Sumpter JP, Tyler CR. 2000. Development and validation of a reproductive performance test for endocrine disrupting chemicals using pair-breeding fathead minnows (*Pimephales promelas*). *Environ Sci Technol* (in press).

25. Kramer VJ, Miles-Richardson S, Pierens SL, Giesy JP. 1998. Reproductive impairment and induction of alkaline-labile phosphate, a biomarker of estrogen exposure in fathead minnows (*Pimephales promelas*) exposed to waterborne 17â-estradiol. *Aquat Toxicol* 40:335-360.

26. Panter GH, Hutchinson TH, Länge R, Whale G, Sumpter JP, van Aerle R, Tyler CR. 2002. Quantifying histological development in the gonads of sexually maturing fathead minnow: phase 2. CEFIC-EMSG research report reference AQ005. CEFIC, Brussels, Belgium.

27. Thorpe KL, Hutchinson TH, Hetheridge MJ, Scholze M, Sumpter JP, Tyler CR., 2001. Assessing the biological potency of binary mixtures of environmental estrogens using vitellogenin induction in juvenile rainbow trout (Oncorhynchus mykiss). Environ Sci Technol 15;35(12):2476-81

28. Parks LG, Cheek AO, Denslow ND, Heppell SA, McLachlan JA, LeBlanc GA, Sullivan CV. 1999. Fathead minnow (*Pimephales promelas*) vitellogenin: Purification, characterization, and quantitative immunoassay for the detection of estrogenic compounds. *Comp Biochem Physiol* 132C:113-125.

29. Korte JJ, Kahl MD, Jensen KM, Pasha MS, Parks LG, LeBlanc GA, Ankley GT. 2000. Fathead minnow vitellogenin: Complementary DNA sequence and messenger RNA and protein expression after 17â-estradiol treatment. *Environ Toxicol Chem* 19:972-981.

30. Tyler CR, Vandereerden B, Jobling S, Panter G, Sumpter JP. 1996. Measurement of vitellogenin, a biomarker for exposure to oestrogenic chemicals, in a wide variety of cyprinid fish. *Journal of Comp Physiol* 166B: 418-426.

31. Rose J, Holbech H, Lindholst C, Noerum U, Povlsen A, Korsgaard B, Bjerregaard P. 2002. Vitellogenin induction by 17β-estradiol and 17α-ethinylestradiol in male zebrafish (*Danio rerio*). Comp. Biochem. Physiol. C. 131: 531-539.

32. Nishi K, Chikae M, Hatano Y, Mizukami H, Yamashita M, Sakakibara R, Tamiya E, 2002. Development and application of a monoclonal antibody-based sandwich ELISA for quantification of Japanese medaka (Oryzias latipes) vitellogenin. Comp.biochem. Phys. C 132 (2): 161-169.

33. Yokota H, Morita H, Nakano N, Kang IJ, Tadokoro H, Oshima Y, Honjo T, Kobayashi K. 2001. Development of an ELISA for determination of the Hepatic vitellogenin in Medaka (*Oryzias latipes*). *Jpn J Environ Toxicol* 4:87–98.

34. Panter GH, Tyler CR, Maddix S, Campbell PM, Hutchinson TH, Länge R, Lye C, Sumpter JP, 1999. Application of an ELISA to quantify vitellogenin concentrations in fathead minnows (*Pimephales promelas*) exposed to endocrine disrupting chemicals. CEFIC-EMSG research report reference AQ001. CEFIC, Brussels, Belgium.

35. Fenske M., van Aerle, R.B., Brack, S.C., Tyler, C.R., Segner, H., (2001). Development and validation of a homologous zebrafish (*Danio rerio* Hamilton- Buchanan) vitellogenin enzyme-linked immunosorbent assay (ELISA) and its application for studies on estrogenic chemicals. Comp. Biochem. Phys. C 129 (3): 217-232.

36. Tyler CR, Aerle RV, Hutchinson TH, Maddix S, Trip H. 1999. An in vivo testing system for endocrine disruptors in fish early life stages using induction of vitellogenin. *Environ Toxicol Chem* 18:337-347.

37. Ankley, G.T., K.M. Jensen, E.A. Makynen, M.D. Kahl, J.J. Korte, M.W. Hornung, T.R. Henry, J.S. Denny, R.L. Leino, V.S. Wilson, M.C. Cardon, P.C. Hartig and L.E. Gray. 2003. Effects of the androgenic growth promoter 17-B-trenbolone on fecundity and reproductive endocrinology of the fathead minnow (Pimephales promelas). Environ. Toxicol. Chem. 22, 1350-1360.

38. Lange R, Hutchinson TH, Croudace CP, Siegmund F, Schweinfurth H, Hampe P, Panter GH, Sumpter JP. 2001. Effects of the synthetic estrogen 17 alpha-ethinylestradiol on the life-cycle of the fathead minnow (*Pimephales promelas*). *Environ Toxicol Chem* 20: 1216-1227.

39. Seki M, Yokota H, Matsubara H, Tsuruda Y, Maeda M, Tadokoro H, Kobayashi K. 2002. Effect of ethinylestradiol on the reproduction and induction of vitellogenin and testis-ova in medaka (*Oryzias latipes*). *Environ Toxicol Chem* 21: 1692-1698.