Centre on Endocrine Disrupters

Information/testing strategy for identification of substances with endocrine disrupting properties

Final report

June 30, 2013

DANISH CENTRE ON ENDOCRINE DISRUPTERS

Ulla Hass and Sofie Christiansen
Division of Toxicology and Risk Assessment, National Food Institute,
Technical University of Denmark

Poul Bjerregaard and Henrik Holbech
Institute of Biology, University of Southern Denmark

DTU Food
National Food Institute
DTU

UNIVERSITY OF SOUTHERN DENMARK

Table of Contents

1. Terms of reference and scope

This report has been prepared by the Danish Centre on Endocrine Disrupters (CeHoS) as a project contracted by the Danish Environmental Protection Agency. The Danish Centre on Endocrine Disrupters is an interdisciplinary scientific network without walls. The main purpose of the Centre is to build and gather new knowledge on endocrine disrupters (EDs) with the focus on providing information requested for the preventive work of the regulatory authorities. The Centre is financed by the Ministry of the Environment and the scientific work programme is followed by an international scientific advisory board.

The overall scope of this project is to provide a science based input to the ongoing work in EU with regard to endocrine disruptors, i.e. the development of criteria for identification, REACH review on EDs and the revised strategy for the future work on endocrine disruptors, focusing on adequate detection of substances with endocrine disrupting properties under various legislative frameworks, including REACH (EC No 1907/2006), the Plant Protection Products Regulation (PPPR) (EC No 1107/2009) and the Biocidal Products Regulation (BPR) (EC No 528/2012) .

2. Background and aim

In general, there is currently no specific information requirements or testing strategies with regard to endocrine disruption in REACH and other relevant legislation. However, in relation to biocides, and recently also to plant protection products, indications of endocrine disrupting properties of a substance trigger additional information/testing requirements. With regard to plant protection products a new regulation from 1 March 2013 (EU 283/2013 (active substances) + EU 284/2013 (products)) sets out the general data requirements from 1 January 2014. According to these, new test methods that include endocrine sensitive endpoints have been included with regard to human health and the environment. Similar data requirements and new test methods that include endocrine sensitive endpoints are expected in the coming guidance on how to fulfil data requirements for biocides (DK EPA, pers. com.). To a limited extent *in vivo* test methods that include endocrine sensitive endpoints have also been included in the standard information requirements for substances regulated under REACH.

In a recent report, it was concluded that there are major limitations as to the ability of the current testing requirements to adequately screen for endocrine disrupting properties and effect sizes of human relevance may be present at the NOAEL (Hass et al. 2013). In the report “State of the art assessment of endocrine disrupters” several limitations are also mentioned and it is stated that “even multigenerational assays suffer from limitations imparted by their design” (Kortenkamp et al. 2012).

The aim of this report is to contribute to development of information/testing strategies for adequate identification of EDs in relation to both human health and the environment.

The report is structured in 3 parts dealing with the following topics:

- Part 1 (chapter 3): Proposals for changes of the current standard information requirements in REACH and legislation concerning plant production products and biocides with the aim of increasing the ability for detection of EDs. The proposed changes include new information requirements, new test methods and modifications of existing methods included in the current standard information requirements. The scientific arguments behind the proposed changes in relation to the existing concern is provided and relevant specific proposals for explanatory note-text for the proposed information requirements is included.
- Part 2 (chapter 4): An information/testing strategy for use when there is concern for endocrine disruption. The strategy consist of several steps depending on the level of knowledge leading to the concern, e.g QSAR models, *in vitro* testing, *in vivo* testing, epidemiological studies or environment studies (field studies). The information/testing strategy may include new and revised test methods and may as such either supplement the existing information requirements or alternatively, be used for substance evaluations. This is specifically relevant, if the existing REACH standard information requirements in relation to endocrine distrupting properties will not be changed.
- Part 3 (chapter 5): A proposal for a ” test package” for already registered substances in REACH to detect substances with endocrine disrupting properties.

The current report is to a high degree based on the OECD guidance document No. 150 on standardised test guidelines for evaluating chemicals for endocrine disruption (OECD GD), OECD 2012, and the OECD Fish Toxicity Framework (OECD STA No. 171, OECD 2012c).

To set up a reference to definition and criteria for identification of EDs, the Danish proposal for identification of EDs, that was submitted to EU in May 2011, has been used (Danish EPA, 2011).

2.1 OECD Conceptual Framework and OECD Guidance Document No. 150

The OECD Conceptual Framework for testing and assessment of endocrine disrupting chemicals (OECD CF) is a five-level organised ‘toolbox’ listing assays considered to provide different types of information regarding the hazards of a substance with regard to its potential endocrine disrupting properties. It is not anticipated to be used as a tiered testing strategy. An overview of the OECD CF is shown in Figure 1.

Both the OECD CF and the OECD GD are limited to endocrine mechanisms and hazard assessment. Information on chemical exposure (e.g. on use, volume, fate, levels, duration and route of exposure) is not considered even though the framework mentions that such information may generally case-by-case be considered in relation to the needs for further testing.

The OECD GD covers only the same endocrine modalities as the OECD CF, i.e.:
• Oestrogen receptor mediated
• Androgen receptor mediated
• Thyroid hormone mediated
• Steroidogenesis interference

Although the assays covered by the OECD GD are applicable to most types of endocrine modalities which are currently reasonably well known, i.e. those operating via oestrogen/ androgen/ thyroid/ steroidogenesis (EATS) - modalities, it should be recognised that the assays may not be responsive to other more poorly-understood types of ED modes of action.

The general approach in the OECD GD is primarily to consider the possible results that might be obtained from each ED-responsive assay, and to provide guidance about how these results might be interpreted in the light of data that may or may not already be available from other *in vitro* or *in vivo* assays. The key questions addressed concern likely mechanisms or mode of endocrine action and any resulting apical effects that can be attributed to such action. Given the widely agreed definition of endocrine disrupting chemicals (WHO, 2002), the advice in the OECD GD only suggests that a chemical is an ED if an adverse *in vivo* effect can be plausibly linked to an endocrine mode of action. The OECD GD provides advice on the next step in testing (if any) which might be appropriate for a regulatory authority to take, if there is a need for further (more definitive or confirmatory) data, given the various data scenarios. It should be noted that it has only been possible to cover the most likely scenarios in the OECD GD. Advice on further testing which may be needed to assist in deciding whether a chemical is an ED, is generally limited to a single next step, and the OECD GD therefore does not present one single overall hazard testing strategy. This is all the more important given that the guidelines for testing for endocrine disruption are relatively new and that the field will probably develop further in the coming years. Thus, the OECD GD does not present a testing strategy as it is restricted to a single step when further testing is recommended or proposed for consideration.

2.2 Danish proposal for criteria for identification of EDs

The proposed Danish criteria for identification of EDs are shown in Table 1 (DK, EPA 2011). They are inspired by the criteria for hazard classification for developmental/reproductive toxicity and they also include specificity as an important decision element.

As can be seen in Table 1, the substances can be allocated to one of three categories. Appropriate grouping should always depend on an integrated assessment of all available data and their interrelationship using a weight of evidence approach. Individual datasets should be analysed case-by-case using expert judgement.

Both adverse *in vivo* effects and mode of action information is important for identification of a substance as an ED in category 1 and consequently, the information/testing strategies in the following sections focus on how to obtain sufficient data for both of these areas.

Figure 1 OECD Conceptual Framework for Testing and Assessment of Endocrine Disrupting Chemicals (as revised in 2012) as included in the OECD Guidance Document No. 150 (Annex 1.4)

	Mammalian and non-mammalian Toxicology	
Level 1 Existing Data and Non-Test Information	• Physical & chemical properties, e.g., MW reactivity, volatility, biodegradability • All available (eco)toxicological data from standardized or non-standardized tests. • Read across, chemical categories, QSARs and other *in silico* predictions, and ADME model predictions	
Level 2 *In vitro* assays providing data about selected endocrine mechanism(s) / pathways(s) (Mammalian and non-mammalian methods)	• Oestrogen or androgen receptor binding affinity • Oestrogen receptor transactivation (OECD TG 455 – OECD TG 457) • Androgen or thyroid transactivation (If/when TGs are available) • Steroidogenesis in vitro (OECD TG 456) • MCF-7 cell proliferation assays (ER ant/agonist) • Other assays as appropriate	
	Mammalian Toxicology	**Non-Mammalian Toxicology**
Level 3 *In vivo* assays providing data about selected endocrine mechanism(s) / pathway(s)[1]	• Uterotrophic assay (OECD TG 440) • Hershberger assay (OECD TG 441)	• Xenopus embryo thyroid signalling assay (When/if TG is available) • Amphibian metamorphosis assay (OECD TG 231) • Fish Reproductive Screening Assay (OECD TG 229) • Fish Screening Assay (OECD TG 230) • Androgenized female stickleback screen (GD 140)
Level 4 *In vivo* assays providing data on adverse effects on endocrine relevant endpoints [2]	• Repeated dose 28-day study (OECD TG 407) • Repeated dose 90-day study (OECD TG 408) • 1-generation reproduction toxicity study (OECD TG 415) • Male pubertal assay (see GD 150, Chapter C4.3)[3] • Female pubertal assay (see GD 150, Chapter C4.4)[3] • Intact adult male endocrine screening assay (see GD 150, Chapter Annex 2.5) • Prenatal developmental toxicity study (OECD TG 414) • Chronic toxicity and carcinogenicity studies (OECD TG 451-3) • Reproductive screening test (OECD TG 421 if enhanced) • Combined 28-day/reproductive screening assay (OECD TG 422 if enhanced) • Developmental neurotoxicity (OECD TG 426)	• Fish sexual development test (OECD TG 234) • Fish Reproduction Partial Lifecycle Test (when/If TG is Available) • Larval Amphibian Growth & Development Assay (when TG is available) • Avian Reproduction Assay (OECD TG 206) • Mollusc Partial Lifecycle Assays (when TG is available) [4] • Chironomid Toxicity Test (TG 218-219) [4] • Daphnia Reproduction Test (with male induction) (OECD TG 211) [4] • Earthworm Reproduction Test (OECD TG 222) [4] • Enchytraeid Reproduction Test (OECD TG 220) [4] • Sediment Water Lumbriculus Toxicity Test Using Spiked Sediment (OECD TG 225) [4] • Predatory mite reproduction test in soil (OECD TG 226) [4] • Collembolan Reproduction Test in Soil (TG OECD 232) [4]
Level 5 *In vivo* assays providing more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism [2]	• Extended one-generation reproductive toxicity study (OECD TG 443)[5] • 2-Generation reproduction toxicity study (OECD TG 416 most recent update)	• FLCTT (Fish LifeCycle Toxicity Test) (when TG is available) • Medaka Multigeneration Test (MMGT) (when TG is available) • Avian 2 generation reproductive toxicity assay (when TG is available) • Mysid Life Cycle Toxicity Test (when TG is available)[4] • Copepod Reproduction and Development Test (when TG is available)[4] • Sediment Water Chironomid Life Cycle Toxicity Test (OECD TG 233) [4] • Mollusc Full Lifecycle Assays (when TG is available) [4] • Daphnia Multigeneration Assay (if TG is available) [4]

Footnote and notes to the OECD Revised Conceptual Framework (Figure 1)

[1] Some assays may also provide some evidence of adverse effects.

[2] Effects can be sensitive to more than one mechanism and may be due to non-ED mechanisms.

[3] Depending on the guideline/protocol used, the fact that a substance may interact with a hormone system in these assays does not necessarily mean that when the substance is used it will cause adverse effects in humans or ecological systems.

[4] At present, the available invertebrate assays solely involve apical endpoints which are able to respond to some endocrine disrupters and some non-EDs. Those in Level 4 are partial lifecycle tests, while those in Level 5 are full- or multiple lifecycle tests.

[5] The Extended one-generation reproductive Toxicity Study (OECD TG 443) is preferable for detecting endocrine disruption because it provides an evaluation of a number of endocrine endpoints in the juvenile and adult F1, which are not included in the 2-generation study (OECD TG 416) adopted in 2001.

Note 1: Entering at all levels and exiting at all levels is possible and depends upon the nature of existing information and needs for testing and assessment.

Note 2: The assessment of each chemical should be made on a case by case basis, taking into account all available information.

Note 3: The framework should not be considered as all inclusive at the present time. At levels 2, 3, 4 and 5 it includes assays that are either available or for which validation is under way. With respect to the latter, these are provisionally included.

Table 1 Danish proposal for criteria for identification of EDs

Category 1- Endocrine disrupter

Substances are placed in category 1 when they are known to have produced ED adverse effects in humans or animal species living in the environment or when there is evidence from animal studies, possibly supplemented with other information, to provide a strong presumption that the substance has the capacity to cause ED effects in humans or animals living in the environment.

The animal studies shall provide clear evidence of ED effect in the absence of other toxic effects, or if occurring together with other toxic effects, the ED effects should be considered not to be a secondary non-specific consequence of other toxic effects. However, when there is e.g. mechanistic information that raises doubt about the relevance of the adverse effect for humans or the environment, Category 2a may be more appropriate.

Substances can be allocated to this group based on:

- Adverse *in vivo* effects where an ED mode of action is highly plausible
- ED mode of action *in vivo* that is clearly linked to adverse *in vivo* effects (by e.g. read-across)

Category 2a - Suspected ED

Substances are placed in category 2a when there is some evidence from humans or experimental animals, and where the evidence is not sufficiently convincing to place the substance in category 1. If for example limitations in the study (or studies) make the quality of evidence less convincing, category 2a could be more appropriate. Such effects should be observed in the absence of other toxic effects, or if occurring together with other toxic effects, the ED effect should be considered not to be a secondary non-specific consequence of other toxic effects.

Substances can be allocated to this category based on:

- Adverse effects *in vivo* where an ED mode of action is suspected
- ED mode of action *in vivo* that is suspected to be linked to adverse effects *in vivo*
- ED mode of action *in vitro* combined with toxicokinetic *in vivo* data (and relevant non test information such as read across, chemical categorisation and QSAR predictions)

Category 2b – Substances with indications of ED properties (indicated ED)

Substances are placed in category 2b when there is *in vitro*/*in silico* evidence indicating potential for endocrine disruption in intact organisms. Evidence could also be observed effects *in vivo* that could be ED-mediated.

3. Proposals for changes of the current standard information/data requirements for industrial chemicals, plant protection products and biocidal products to include endocrine disruption

The objective in this section is to give guidance on a stepwise approach to hazard identification with regard to endocrine disruption based on the current information requirements in REACH, PPPR and BPR. It is attempted to ensure that the information needs are met in the most efficient and humane manner so that animal usage and costs are minimized. The data requirements and proposals for changes are separated for human toxicity and environmental effects.

3.1 Human toxicity - REACH standard information requirements and proposals for changes to include endocrine disruption

The standard information requirements for substances for registration under REACH are differentiated according to tonnage level. Generally, testing requirements at a lower tonnage level apply also to the higher tonnage level, unless specific exemptions are clearly stated or because the higher tier standard information requirement supersedes that at the lower tonnage level. REACH does not include specific standard information requirements with regard to endocrine disruption. Relevant standard information requirements related to human health relevant for detection of endocrine disrupters are only those covered by testing for repeated dose toxicity, carcinogenicity and reproductive toxicity. The current information requirements for repeated dose toxicity and reproductive toxicity are summarized in the left column in Table 2. (In relation to carcinogenicity there are no standard information requirements in REACH unless triggered by suspicion of *in vivo* genotoxicity (i.e. hazard classification Mut cat 2) and furthermore triggered by wide dispersive use or evidence of frequent or long-term human exposure to the substance). It should be noted that interpretation of the actual testing requirements also depends on a weight-of-evidence evaluation of existing data and that it may therefore be different and less than the minimum requirements as presented here. Factors that could influence the testing requirements include concerns for endocrine disruption due to substance structural relationships with other chemicals, QSAR predictions, toxicokinetic data, the results of already available relevant toxicity – and ecotoxicity - studies, available data from humans exposed to the substance, and the use and human exposure patterns.

The proposed information/testing strategy for ED includes new standard information requirements, new test methods and modifications of existing methods included in the current standard information requirements. An overview of the proposed changes is shown in the right column Table 2. Further details are described in the following sections.

Table 2. Current REACH standard information requirements and proposed changes

Current information requirements	Proposals for changes with regard to ED (see text for details)
1 -10 tpa/ manufacturer or importer: Assessment of all available *in vitro* and *in vivo* data (including data related to tests for endocrine disruption), human data, data from valid (Q)SARs and data from structurally related substances (read-across approach)	Evaluate all existing and available data for alerts for ED[1] QSAR model predictions for ED MoA and ED effects *In vitro* assays for interaction with different ED modalities i.e. ER, AR and steroidogenesis interference
10-100 tpa/manufacturer or importer: Short-term repeated dose toxicity study (28 days), i.e. TG 407 Screening for reproductive/developmental toxicity (TG 421) or combined repeated dose/reproductive toxicity screening test (TG 422) The two-generation reproductive toxicity study shall be proposed by the registrant if there are indications of potential reproductive toxicity from a repeated dose toxicity study (e.g. histopathological effects on the gonads) or the substance has a close structural relationship with a known reproductive toxicant	 Include as mandatory the optional ED endpoints in OECD TG 407 Extend TG 421/422 until PND 14 and enhance with ED related endpoints: AGD, nipple retention, thyroid hormones Replace TG 416 with TG 443, i.e.: The extended one-generation reproductive toxicity study (TG 443) shall be proposed by the registrant if there are indications of potential endocrine toxicity from a repeated dose toxicity study, TG 421/422, or the substance has a close structural relationship with a known endocrine toxicant or there are clear SAR alerts or positive valid predictions from (Q)SAR models or *in vitro* testing
100 -1000 tpa /manufacturer or importer: 90-day repeated dose toxicity study (TG 408) Prenatal developmental toxicity study (TG 414) in one species, possibly a second species Two-generation reproductive toxicity study, having regard to the likely route of human exposure, if the 28-day or 90-day study indicates adverse effects on reproductive organs or tissues	 Include as mandatory all ED relevant endpoints in TG 408 similarly as in TG 407 inclusive the optional endpoints in TG 407 Enhance TG 414 with ED related endpoints: AGD, malformations of external reproductive organs Replace TG 416 with TG 443, i.e.: The Extended one-generation reproductive toxicity study (OECD TG 443), having regard to the likely route of human exposure, if existing data indicates adverse effects on the endocrine system
1000 tpa or more: 90-day repeated dose toxicity study (TG 408) Prenatal developmental toxicity study (TG 414) in one species, possibly a second species Two-generation reproduction toxicity study (TG 416)	 Include as mandatory all ED related endpoints in TG 408 similarly as in OECD TG 407 inclusive the optional endpoints in TG 407 Enhance TG 414 with related ED endpoints: AGD, malformations of external reproductive organs Replace two-generation study (TG 416) with extended one-generation reproductive toxicity study (TG 443)

[1] In general alerts for ED in this document cover endocrine activity and potential endocrine related effects.

3.1.1 Tonnage level 1-10 tpa/manufacturer or importer

According to the current standard information requirements, the first step is to gather all available test data on the substance to be registered as well as all other available and relevant information on the substance regardless of whether testing for a given endpoint is required or not at the specific tonnage level. The same should been done with regard to ED to be able to evaluate whether there are alerts from the existing toxicological database for ED related effects. If there are such alerts, their implications for further testing are discussed in section 3 and 4.

An alert is any factor, with the exclusion of convincing evidence for ED that is present in the existing toxicological/ecotoxicological databases, whether based on theoretical considerations or from experimental or observational data that raises concerns that a substance may be an ED. As part of the data review the following questions should be asked:

- Are there any alerts for ED?
- Are the current data sufficient/adequate for assessing whether the substance is an ED?
- If the data are insufficient, what study (or studies) is most appropriate? This decision must take account of both the standard tonnage related information requirements of REACH, any available ED relevant information, the nature of the alert(s) and Weight of Evidence.
- Is there any knowledge of the chemical, chemical groups or categories that would indicate special features to be included in the study design?

From a scientific perspective, it is not possible to generate an exhaustive list of ED alerts that would automatically trigger a particular study or have clearly defined implications for ED hazard identification. Instead, the alerts mentioned below should be viewed as a guide of indicators that would provide input to a Weight of Evidence analysis requiring expert judgement that leads to the most appropriate testing strategy.

Read-across to structurally or mechanistically similar substances (SAR)
The potential ED toxicity of a substance, for which no data are available may, in some cases, be evaluated by read-across from structurally or mechanistically related substances for which experimental data exists. The read-across approach is based on the principle that structurally and/or mechanistically related substances may have similar toxicological properties. Note that there are no formal criteria to identify structural alerts for ED related toxicity or for read-across to closely related substances, however, a range of ad hoc considerations and practices exist.
Based on structural similarities between different substances, the ED toxicity potential of one substance or a group of substances may be extended by read-across to a substance, for which there are no or limited data on this endpoint.
A mechanism of toxicity or mode of action identified for a substance and/or group of substances and causally related to adverse ED effects may be extended (read-across) to a substance for which a similar mechanism or mode of action has been identified, but where no or limited data on ED related toxic effects are available. In such cases, the substance under evaluation may reasonably be expected to exhibit the same pattern of ED toxicity.

The chemical category concept has been developed under the OECD HPV (High Production Volume) programme (OECD, 2004) as an approach to fill data gaps without the need for conduction of tests. A chemical category is a group of chemicals whose physico-chemical and toxicological properties are likely to be similar or follow a regular pattern as a result of structural similarity. In the category approach, not every substance needs to be tested for every endpoint.

Further testing may be required, on a case-by-case basis, to support a read-across proposal.

Human data on ED toxicity
Epidemiological studies, case reports and clinical data may provide sufficient hazard and dose-response evidence for evaluating a substance as an endocrine disrupter in category 1 and for risk assessment, including the identification of a N(L)OAEL. In such cases, there will normally not be a need to test the chemical. However, convincing human evidence of ED toxicity for a specific chemical is rarely available because it is often impossible to identify a population suitable for study that is exposed only to the chemical of interest. Human data may provide limited evidence of ED toxicity that indicates a need for further studies of the chemical; the test method selected should be based on the effect suspected.

When evidence of an ED hazard has been derived from animal studies it is unlikely that the absence of evidence of this hazard in an exposed human population will negate the concerns raised by the animal model. This is because there will usually be methodological and statistical limitations related to the human data. Extensive, high quality and preferable prospective, data are necessary to support a conclusion that there is no risk from exposure to the chemical.

Quantitative Structure-Activity Relationship (QSAR)
There are a number of potential targets/mechanisms associated with ED toxicity that, on the basis of current knowledge, cannot be adequately covered by a battery of QSAR models.
QSAR approaches are currently not formally validated by e.g. OECD and EU and this also applies for ED related toxicity. Consequently a negative result from QSAR models cannot be interpreted as demonstrating the absence of ED hazard unless there is other supporting evidence. However, a positive prediction from a validated QSAR model within the applicability domain of the model provides a trigger (alert) for further testing considerations.

At the present stage of QSAR development in OECD, QSAR models predicting mechanism would be used for prioritisation, ranking and hazard identification (OECD, GD 150). Some QSAR models for endocrine disruptive activity and reproductive toxicity effects are now becoming available. The output of these models can be applicable (with caution) to interpretation of the mechanisms or mode of action underlying *in vivo* results from studies in vertebrates. Furthermore, other QSAR methodologies such as categorization in the OECD (Q)SAR Toolbox can be used to identify groups of chemicals and structural alerts that are linked to *in vivo* effects, thereby elucidating possible key modes of action or mechanisms.

Additionally, QSAR models could be used as part of a Weight of Evidence approach, when considered alongside other data (*in vivo* ADME information and *in vitro* data), provided the applicability domain is appropriate. Also, QSARs can be used as supporting evidence when assessing the toxicological properties by read-across within a substance grouping approach, providing the applicability domain is appropriate.
In conclusion, the use of QSAR data are recommended to elucidate whether there are alert(s) for further testing for ED effects.

In vitro data
The design of *in vitro* assays for ED activity is challenging in view of the complexity of the endocrine system and the large number of potential targets/mechanisms associated with endocrine toxicity.

In vitro screens are relevant for effects in humans and vertebrate wildlife because many are based on highly conserved hormone receptors or interaction with key enzymes or other key molecules involved in the regulation of hormone levels in all vertebrates. Chemicals that bind to these receptors or otherwise interfere with key processes of hormone regulation have the potential to cause effects in *in vivo* studies of both mammals and non-mammalian vertebrate wildlife, assuming concentrations that reach the target are sufficiently high (e.g. dependent on ADME- absorption, distribution, metabolism, and excretion). A novel paper concludes that when comparing data from fish and rat assays a high concordance was seen with respect to identifying chemicals that impacted specific endocrine pathways of concern (Ankley and Gray, 2013). Although most chemicals were detected as positive in both the rat and fish assays, eliminating data from one class of vertebrate would weaken the battery. For example, the effects of competitive inhibitors of steroid hormone synthesis were far more obvious in the fish assay, whereas the activity of androgen receptor antagonists was clearer in mammalian assays (Ankley and Gray, 2013).

The validated *in vitro* assays for which guidance is provided in the OECD Guidance Document include:

- ER (oestrogen receptor) Binding Assay (US EPA OPPTS 890.1250)
- AR (androgen receptor) Binding Assay (US EPA OPPTS 890.1150)
- OECD TG 455: Stably Transfected Human ERα Transcriptional Activation Assay (ER STTA) (including guidance for the antagonism assay – not part of OECD TG)
- OECD TG 456: H295R Steroidogenesis Assay
- Aromatase Assay (US EPA OPPTS 890.1200)
- Stably Transfected Human AR
- Transactivation Assay (AR STTA) (in Appendix 2 as no guidelines is available)

It is noted that currently there are no validated *in vitro* methods available regarding thyroid hormone system related mechanisms, but a project in this regard has been proposed to the OECD Test Guideline Programme for 2013. This project proposal includes numerous potential *in vitro* methods but also concludes that for a range of thyroid hormone interference related mechanisms no *in vitro* test method has yet been developed or alternatively, the mechanism is not fully clarified.

It is proposed in the OECD GD that a battery of *in vitro* tests should be carried out as a single test will usually only provide information on one modality. Thus, the results from a combination of tests will increase weight of evidence. Assays for interaction with different ED modalities e.g. ER, AR and steroidogenesis interference, should be conducted at the same time so that all results can be considered together. Thyroid Receptor (TR) binding assays and other assays concerning mechanisms of thyroid disruption may also be conducted but as mentioned above these *in vitro* methods are not in common use and are not yet formally validated.

The *in vitro* assays do not include the use of a xenobiotic metabolising system but consideration should be given to the inclusion of this depending upon the circumstances e.g. if the metabolism of a chemical is unknown, although the methods for inclusion of xenobiotic metabolising systems are not yet validated (OECD GD). Alternatively, for a chemical with known metabolites, these could also be tested in the *in vitro* assays.

Positive *in vitro* test results indicate the possibility of endocrine disrupting effects *in vivo* and thus provide an alert for further testing.

Current *in vitro* tests covered by the OECD CF are largely based on mammalian systems, but their results can be used with caution to draw conclusions about possible EDs in other vertebrates, although potency and adverse consequences may differ (OECD GD).

Negative *in vitro* results alone cannot be used to exclude possible endocrine disrupting activity because of their inherent limitations, such as inability or unknown capacity to metabolically activate toxicants. In addition, chemicals can interfere with the endocrine system in other ways than through the receptor or by interaction with steroidogenesis. An example is effects on the hypothalamic-pituitary-gonadal axis (HPG) that can only be detected in whole animal studies. Each *in vitro* assay measures a certain mechanism and thus conclusions can be drawn only in the context of what the *in vitro* assay evaluates. However, negative *in vitro* effects should only be interpreted as a tentative indication of a lack of endocrine disruption activity (mechanism) for the modality in question, if it can be substantiated that the compound does not undergo metabolic activation e.g. by the use of ADME information.

In conclusion, *in vitro* assays for interaction with different ED modalities e.g. ER, AR and steroidogenesis interference, should be conducted to elucidate whether there are alert(s) for further testing for ED effects. *In vitro* assays concerning mechanisms of thyroid disruption may later, when validated, be considered for inclusion in the testing requirements.

3.1.2 Tonnage level 10-100 tpa/ manufacturer or importer

The relevant test methods for identification of endocrine disrupters at this tonnage levels are mentioned in relation to testing for repeated dose toxicity and reproductive toxicity and includes the OECD TG 407 (repeated dose 28-day toxicity study) and OECD TG 421/422 (the developmental / reproduction toxicity screening test/combined 28-day and reproductive screening assay. The ability for detection of EDs can for these tests be enhanced without increasing the number of experimental animals used.

It is also stated at this tonnage level that the prenatal developmental toxicity study (OECD TG 414) or the two-generation reproduction toxicity study (OECD TG 416) shall be proposed by the registrant if there are serious concerns about the potential for adverse effects on fertility or development based on indications of potential reproductive toxicity from a repeated dose toxicity study (e.g. histopathological effects on the gonads) or based on a close structural relationship with a known reproductive toxicant. Also concerns based on available valid QSAR-predictions and *in vitro* data are referred to. With regard to EDs this is proposed to be modified to: "The extended one-generation reproductive toxicity study (OECD TG 443) shall be proposed by the registrant if there are indications of endocrine related toxicity from a repeated dose toxicity study, OECD TG 421/422, or the substance has a close structural relationship with a known endocrine toxicant or there are clear alerts from QSAR models or *in vitro* testing". The main reason for this proposal, besides that it addresses ED specifically, is that currently there is no reference in REACH for triggering a higher tier reproductive toxicity study based on alerts from TG 421/422 or from valid QSAR predictions, structural close similarity to known reproductive toxicants or *in vitro* tests. Furthermore, triggering of TG 443 instead of TG 416 is based on the preference for the former study for identification of EDs according to OECD ED GD (2012). Due to the higher sensitivity and broader scope of TG 443 it is also generally acknowledged that TG 443 is superior to TG 416 for reproductive toxicity in general.

OECD TG 407: Repeated Dose 28-Day Oral Toxicity Study in Rodents
The animals in this assay are young adults with intact hypothalamus-pituitary-gonadal/thyroid axes and are therefore are a relevant model for human health. The OECD validation of the assay for endocrine endpoints showed that this assay is relatively insensitive and would only detect chemicals that were moderate and strong EDs for (anti)-oestrogenicity and (anti)-androgenicity (e.g. ethinylestradiol and flutamide). However, it did detect EDs that were weak and strong modulators of thyroid hormone-related effects (e.g. propylthiouracil and methyl testosterone). It may also detect steroidogenesis inhibition although only one (potent) chemical was used in the validation study (CGS 18320B) (OECD, 2006b). It should be noted that, as this assay is not primarily designed to detect endocrine disruption, a higher degree of systemic toxicity is typically induced than is the case with the other level 3 and 4 assays in the OECD CF. The possibly confounding effect of systemic toxicity on endocrine endpoints therefore needs to be considered case-by-case.

The results of OECD TG 407 can pragmatically be divided into "apical" and "indicators of hormonal activity" (OECD GD). "Apical" endpoints are weights of testes, epididymides, prostate

(and seminal vesicles with coagulating glands), ovary, uterus, histopathologic changes in testes, epididymides, prostate, seminal vesicles, coagulating glands, ovary, uterus/cervix, vagina, thyroid and oestrous cyclicity. “Indicators of hormonal activity” are hormones (T3, T4 and TSH). Both groups have similar biological importance, although the “indicators of hormonal activity” in the mammalian assays are serum hormones and are generally, but not always, more variable than “apical endpoints” (OECD GD).

Several of the apical endpoints and all of the indicators of hormonal activity are optional endpoints in this TG and therefore they may not be measured. These endpoints include weight of uterus and ovaries, oestrous cyclicity, histopathologic changes in mammary glands and pituitary and circulating levels of T3, T4, and TSH. **All of these optional endpoints in TG 407 are of relevance for identification of the effects of EDs and were included in the OECD validation. It is therefore recommended that they become mandatory in the information requirements for EDs.**

Also other ED endpoints not specified in the guideline e.g. reproductive hormones may be measured and if positive, they would contribute to the overall assessment of a positive result.

The apical endpoints for the detection of effects on male and female reproductive organs tended to be less sensitive than the indicators of hormonal activity in the validation of the OECD TG 407 and therefore such changes are generally more likely to be indicative of an ED although the results in entirety should be considered rather than single isolated changes. However, in the OECD validation studies for TG 407 this was not the case for the thyroid as changes in thyroid histopathology were always as sensitive, or more sensitive, than changes in thyroid hormone and TSH levels. However, the OECD validation is based on a limited number of substances and evidence for other substances suggests that for example T4 thyroid hormone levels may be decreased in the absence of thyroid histopathology (OECD, 2012). Also, “normal” T4 level are of major importance during brain development (Zoeller and Crofton, 2000; Zoeller 2007). Therefore assessment of thyroid hormones is evaluated as important in a testing strategy for EDs.

A positive result for indicators of hormonal activity alone in TG 407 should be considered with caution as it is possible that these endpoints may have detected weak effects that were not detected by the apical endpoints in this study but may then be detected in longer-term studies (OECD GD).

A negative result for the OECD TG 407 is taken to be absence of changes in both endocrine relevant indicators of hormonal activity and apical endpoints. In the absence of other pertinent lines of evidence, negative results in this test alone cannot be taken as evidence that the chemical is not an ED (OECD GD), because according to the validation studies of TG 407 this test guideline is only able to detect moderately and strongly acting EDs.

OECD TG 421 Reproduction/Developmental Toxicity Screening Test and OECD TG 422 Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test
These assays are designed to provide limited information about the effects of a chemical on the male and female reproductive systems including gonadal function, mating, conception, gestation, development of the conceptus and parturition. The OECD TGs have similar experimental schedules but OECD TG 422 includes a more detailed assessment of repeated dose toxicity and thus more endpoints.

These assays have not been validated for the detection of EDs, but include endpoints that are suitable for the determination of endocrine effects. In addition to reproduction/development, OECD TG 421 and 422 may both provide information about endocrine effects on adult parental male reproductive organs. OECD TG 422 will also include information about effects on the thyroid. The parental female reproductive organs are also examined but detection of endocrine effects in these organs may be obscured because of pregnancy. So although these tests include exposure during pregnancy, the endpoints related to fertility and gestation maintenance are measured in the parent generation. Thus they do not include exposure during the critical time windows of development for those endpoints.

Gross evaluation of anogenital distance (AGD) is generally used for sexing the offspring in reproductive toxicity studies, because anogenital distance is normally twice as long in males compared to females. Thus major effects on male sexual differentiation induced by potent anti-androgens such as flutamide may be detected as all offspring may display female-like anogenital distance (e.g. Hass et al. 2007). Measurement of AGD is a sensitive endpoint for anti-androgens and may also be affected by oestrogens. The AGD endpoint is included in the OECD TG 443 and can as such be evaluated as a validated endpoint. In addition, the OECD GD 43 and GD 151 states “A statistically significant change in AGD that cannot be explained by the size of the animal indicates effects of the exposure and should be used for setting the NOAEL”. AGD is measured at birth and therefore this endpoint can be included in TGs 421/422 without any modification of the overall test design. **Inclusion of measurements of AGD similarly as in TG 443 is therefore recommended in TG 421/422 as this will significantly improve the ability for detection of ED effects.**

Assessment of nipple retention in male offspring is also included in the OECD TG 443 and is regarded as a sensitive endpoint for anti-androgens. This endpoint has to be assessed at postnatal day 13/14 and consequently inclusion of this endpoint in OECD TG 421/422 requires that the observation period is extended from postnatal day 3 to postnatal 13/14. However, obtaining data for both AGD and nipple retention will provide an increased ability for evaluating the potential anti-androgenicity of a substance compared to having only data for AGD, especially in cases where equivocal AGD data is found. **Thus, for the OECD TGs 421/422 an extension of the testing period from postnatal day 3 to 13/14, i.e. 10-11 day, and assessment of nipple retention is recommended on postnatal day 13/14.**

Assessment of thyroid hormone levels is already proposed here to be included in the OECD TG 407, i.e. at the tonnage level where OECD TG 421/422 is also required. Inclusion of assessment of

thyroid hormone levels also in OECD TG 421/422 might, however, provide useful additional information, because pregnancy may affect the sensitivity for effects on thyroid hormones in the dams. Also, data on thyroid hormone levels in the pups can be useful for evaluating if there may be insufficient exposure via maternal milk and thus a need for direct dosing of pups in further studies. It is noted that measurement of blood thyroid hormone levels in the pups can be done when the study is terminated and hence invasive measurement on the pups are not required.
Thus, for the OECD TGs 421/422, it is recommended to include assessment of thyroid hormone levels in the dams during pregnancy and in the offspring at the termination of the study.

It is noted that a proposal for considering enhancement of TG 421/422 in line with the recommendations above based on an analysis of existing data has been proposed to the OECD TGP by DK and has recently been added in the OECD TGP work plan (pers.com, DK National OECD Coordinator Sofie Christiansen, 2013).

The ED related proposals for enhancements of both the OECD TG 407 and the OECD TG 421 should to be included in enhancement of TG 422, because TG 422 is a merged version of TG 407 and TG 421).

3.1.3 Tonnage level 100 -1000 tpa/manufacturer or importer

The relevant test methods for detection of endocrine disrupters at this tonnage levels are mentioned in relation to testing for repeated dose toxicity and reproductive toxicant and include the TG 408 (repeated dose 90-day toxicity study (oral)) and OECD TG 414 (prenatal developmental toxicity test). The ability for identification of EDs can be enhanced by inclusion of both the mandatory and optional ED sensitive endpoints included in the updated OECD TG 407 (see section 3.1.2) into OECD TG 408 and by inclusion of a few relevant ED sensitive endpoints into the OECD TG 414. For both OECD TGs this can be done without increasing the number of animals used.

Also, the two-generation reproduction toxicity study may be triggered at this tonnage level, having regard to the likely route of human exposure, if the TG 407 or TG 408 (28-day or 90-day repeated dose toxicity) studies indicates adverse effects on reproductive organs or tissues.

With regard to the strategy for EDs the text is proposed to be modified to:
“Extended one-generation reproductive toxicity study (TG OECD 443), having regard to the likely route of human exposure, if existing data indicates adverse effects on the endocrine system. “

OECD TG 408: Repeated dose 90-day toxicity study
Both the 28 and the 90-day studies (OECD TG 407 and OECD TG 408, respectively) are included in level 4 of the OECD Conceptual Framework, however, only OECD TG 407 has been updated and validated for the detection of endocrine disrupters. **To increase the ability for detection of ED effects in OECD TG 408 it is recommended to include as mandatory similar endpoints as in OECD TG 407, incl. the current optional ED related endpoints** (see section 3.1.2).

OECD TG 414: Prenatal developmental toxicity study
The prenatal development toxicity study has not been validated for the detection of endocrine disrupters. In this study, animals are exposed from implantation to two days before expected birth. Although the test includes exposure during pregnancy, there are no endpoints related to fertility and endpoints related to gestation maintenance are measured only in the parental generation. Thus, the study does not include exposure during critical time windows of development for those endpoints. The foetuses are inspected for gross anomalies. However, differences between humans and rodents should be born in mind, as some parts of sexual differentiation that take place during the third trimester of human pregnancy occurs after birth in the rat.

Gross evaluation of anogenital distance (AGD) is generally used for sexing the offspring in reproductive toxicity studies, because anogenital distance is normally twice as long in males compared to females. Thus major effects on male sexual differentiation induced by potent anti-androgens such as flutamide may be detected as all offspring may display female-like anogenital distance (e.g. Hass et al. 2007). Measurement of AGD is a sensitive endpoint for anti-androgens and may also be affected by oestrogens. The endpoint is included in the OECD TG 443 and can as such be evaluated as a validated endpoint. In the OECD TG 443 AGD is measured at birth, but several studies have shown that AGD can also be measured in foetuses on gestation day 21 and give valuable information on effects of endocrine disrupters on the sexual dimorphic development of the AGD (Borch et al. 2006; Saillenfait et al. 2009, 2011a, 2011b, 2013; Dean et al. 2012). Thus, this endpoint can be included in OECD TG 414 without any modification of the overall test design.

Measurements of AGD in pups are in this information/testing strategy also recommended in the Reproduction/Developmental Toxicity Screening Tests (OECD TG 421/422) at a lower tonnage level and these data may therefore be available at this tonnage level. Inclusion of AGD measurement also in OECD TG 414 is, however, very likely to lead to a higher sensitivity in this test guideline than in TG 421/422 due to the higher number of litters evaluated in the OECD TG 414 compared to OECD TGs 421/422 (20 and 8 litters per group, respectively).
Inclusion of measurements of AGD is therefore recommended in OECD TG 414 as this will significantly improve the ability for detection of ED effects.

3.1.4 Tonnage level more than 1000 tpa/ manufacturer or importer

The relevant test methods for detection of endocrine disrupters at this tonnage levels are mentioned in relation to testing for repeated dose toxicity and reproductive toxicity and includes the OECD TG 408 and OECD TG 416 (the two-generation reproduction toxicity study). The ability for identification of EDs can be enhanced by inclusion of the ED sensitive endpoints in OECD TG 407 (see section 3.1.2) into OECD TG 408, and by substituting the OECD TG 416 with TG 443. For the OECD TG 408 this can be done without increasing the number of experimental animals used. The substitution of OECD TG 416 with OECD TG 443 leads to around 40% reduction of the number of experimental animals used if F2 is excluded. Omission of the DIT (developmental immunotoxicity) and DNT (developmental neurotoxicity) cohorts is not recommended, because these types of

functional developmental effects may be affected by EDs. The DNT cohort may for example provide relevant information on effects on sexual dimorphic behaviour caused by effects on sex hormones or effects on the developing brain caused by effects on thyroid hormone levels.

A carcinogenicity study may be required if the substance has a widespread dispersive use or there is evidence of frequent or long-term human exposure and the substance is classified as mutagen category 3 or there is evidence from the repeated dose toxicity study(ies) that the substance is able to induce hyperplasia and/or pre-neoplastic lesions. The study is not designed for detection of EDs, but hormonally mediated cancers may be detected in this study.

OECD TG 416 Two-Generation Reproduction Toxicity Study vs. OECD TG 443 Extended One-Generation Reproduction Toxicity Study

The OECD two-generation reproduction toxicity study is an apical assay designed to provide general information concerning the effects of a chemical on the male and female reproductive systems including gonadal function, the oestrus cycle, mating, conception, gestation, parturition, lactation, weaning and growth and development of the offspring. The study is not specifically designed to detect EDs but include several endpoints relevant for the assessment of possible endocrine disruption during development.

The two-generation study (OECD TG 416) was revised in 2001 to include a more comprehensive range of endpoints. These endpoints include sexual maturation (vaginal opening and preputial separation) which are particularly sensitive to EDs. Two-generation studies conducted prior to the adoption of the revised OECD TG 416 in 2001 is therefore unlikely to provide as much data as studies conducted with the revised OECD TG 416, particularly with regard to endocrine disruption (OECD GD). Nonetheless, even results of two-generation reproduction studies conducted after the revision in 2001 should be interpreted with caution as some endpoints sensitive to endocrine disruption are not included in the current 2001 version of the two-generation reproduction study, such as nipple retention and measurement of thyroid hormones. Assessment of anogenital distance is included in the F2 pups if triggered by effect on sex ratio in the litters or on sexual maturation in F1. However, sex ratio in mammals is generally a very insensitive endpoint for effects of EDs and sexual maturation has lower sensitivity for detecting anti-androgenic effects than assessment of anogenital distance. Thus, for OECD TG 416 there are uncertainties with regard to the ability to adequately detect endocrine disrupters.

The new extended one-generation reproductive toxicity study (OECD TG 443) includes the above mentioned ED sensitive endpoints as well as assessment of neurodevelopment and immunotoxicity. This test also requires an increased number of pups to be examined and the sensitivity is therefore higher than the sensitivity for TG 416. Therefore, the use of the extended one-generation reproductive toxicity study (OECD TG 443) instead of TG 416 study would significantly enhance the ability for detection of endocrine disrupters.

It is therefore strongly recommended to replace the existing two-generation reproduction toxicity study (TG 416 from 2001) with the extended one-generation reproductive toxicity study (OECD TG 443 from 2011).

TG 408 Repeated dose 90-day toxicity study
Both the 28 and the 90-day studies (OECD TG 407 and OECD TG 408, respectively) are included in level 4 of the OECD Conceptual Framework, however, only OECD TG 407 has been updated and validated for the detection of endocrine disrupters. **To increase the ability for detection of ED effects in OECD TG 408 it is recommended to include similar endpoints as in OECD TG 407, incl. the optional endpoints** (see section 3.1.2).

3.2 Human toxicity – Plant Protection Products and Biocidal Products Regulations data requirements and proposals for changes to include endocrine disruption

The current requirements for toxicological information about active ingredients (plus safeners and synergists) of plant protection products (PPP) and biocides are rather similar to the requirements in REACH at the tonnage levels > 1000 tpa. This means that for the active substances there is adverse effect data from long-term repeated dose toxicity data in mammals (TG 408, 409, 451/452/453) as well as from higher tier reproductive toxicity studies (TG 414 and TG 416) as in REACH. However, for biocides and in the recently revised data requirements for the active ingredients of PPPs (EU 283/2013) the extended one-generation reproductive toxicity study (TG 443) is mentioned as an alternative approach to TG 416. If there is evidence that the active substance may have endocrine disrupting properties, additional information or specific studies shall be required to elucidate the mode/mechanism of action and to provide sufficient evidence for relevant adverse effects. With regard to active ingredients of PPPs the new regulation requires that several specified ED relevant endpoints shall be investigated. Studies required shall be designed on an individual basis and taking into account Union or internationally agreed guidelines, in the light of the particular parameters to be investigated and the objectives to be achieved. The Commission Communication 2013/C 95/01 list the following tests as relevant for identification of endocrine disrupting properties:

OECD TG 456 (steroidogenesis assay)
OECD TG 441 (Hershberger assay)
OECD TG 455 (ER STTA)
OECD 440 (Uterotrophic assay)
OCSPP guideline 890.1500 (male pubertal assay)
OCSPPP Guideline 890.1450 (female pubertal assay)
U.S. EPA (2007): 15-day intact adult male rat assay.

Similar requirements and test methods are also expected in the coming guidance on how to fulfil the data requirements for biocides (Danish EPA, pers. com).

For biocides and pesticides it is likely that there may also be ED relevant data available from published studies.

It is expected that the coming criteria for identification of a substance as an ED will require both mode of action data and adverse effect data, however, ED mode of action data are only triggered for

biocides and active ingredients of PPPs if there are indications of endocrine disrupting properties. Currently, ED mode of action data for plant protection products and biocides are therefore most likely only available to a limited extent. Also, adverse effect data of relevance for ED may be quite limited when the testing was performed before the last revisions of TG 407 in 2008 and TG 416 in 2001. The best data for adverse effect will be available if the TG 443 is used and it is therefore proposed, similarly as proposed for REACH, to require the TG 443 instead of the TG 416. It is also, similarly as proposed for REACH, proposed to update and enhance TG 407, 408 and 414 with ED relevant endpoints.

To obtain mode of action data in order to secure a formalised searching for ED alerts for all active ingredients of PPPs and biocides, it is recommended generally to use an "ED relevant standard information package" based on the proposed changes of REACH standard information requirements shown in Table 2. This includes to

- evaluate all data for alerts for ED
- investigate the substance in QSAR models for ED MoA and ED effects
- perform i*n vitro* assays for interaction with different ED modalities, i.e. ER, AR and steroidogenesis interference.

Depending on the results from this integrated analysis of all relevant information, the need for further *in vivo* studies should be considered. When there are alerts for ED further studies should be selected based on the OECD GD and the general ED information/testing strategy described in section 4. Generally, this will have to be a case-by-case evaluation to be performed by experts within the field.

Thus, the same recommendations for ED relevant enhancements addressed above for substances regulated under REACH as regards TG 407, TG 408, TG 414 and TG 416/443 also apply for plant protection products and biocides. In particular, replacement of TG 416 with TG 443 is strongly recommended. It should also be considered to include an ED relevant standard information package as part of the data requirements for active substances in plant protection products and biocidal products to secure a formalised searching for ED alerts.

3.3 Environmental effects – current information requirements in REACH, PPPR and BPR and proposals for changes to include endocrine disruption

3.3.1 REACH - current information requirements and proposal for changes to include endocrine disruption

As already mentioned, REACH does not include standard information requirements specifically targeting identification of endocrine disruption. The REACH standard information requirements with regard to effects on the environment which may be relevant concerning endocrine disrupting properties are shown in table 3 as well as the proposed changes of the requirements to cover also ED effects.

Only a limited number of adopted OECD TGs are available on species relevant in relation to environmental ED effects besides the toxicological tests employing rodent species or rabbits (relevant for mammalian wildlife). With the exception of birds (see footnote ** to table 3 below) the only taxonomic groups represented are fish and frogs which are only relevant in relation to aquatic toxicity.

For the time being, there is only one OECD validated test, TG 234, the Fish Sexual Development Test (FSDT) that definitively can identify specific endocrine disrupting adverse effects in non-mammalian vertebrates and only regarding interference with sex hormones (in particular oestrogens and androgens) and steroidogenesis.

Table 3. Requirements for ecotoxicological information for higher organisms and proposal for changes according to REACH annexes VII to X (other types of tests without any relevance for endocrine aspects, e.g. tests on degradation and adsorption, plants, microorganisms, etc. are not included). The proposed changes should be regarded as supplementary tests to standard information requirements unless the tests selected to cover endocrine effects also cover all endpoints described in the information requirements(s). See also Table 4 for overview of tests that may provide data for standard information requirements for REACH, PPPR and BPR.

Production volume	**Information requirement(s)**	**Proposals for changes with regard to ED (see text and Annex 1 for details)**
≥1 t	Assessment of all available *in vitro* and *in vivo* data including data related to endocrine activity and disruption: human data, test data, data from valid (Q)SAR´s and data from structurally related substances (grouping and read-across approach). Short term invertebrate toxicity (*Daphnia*)	Evaluate all existing and available data for alerts for ED, QSAR model predictions for ED MoA and ED related adverse effects. *In vitro* assays for interaction with different ED modalities i.e. ER, AR and steroidogenesis and if/when available thyroid hormone system interference.
≥ 10 t	Short term fish toxicity. Long-term fish toxicity study shall be considered if the substance is poorly water	Read across from all toxicological data generated from the proposed changes in Table 2. If ED alerts are identified the MoA and

	soluble, if needed for the chemicals safety assessment (i.e. if PEC > PNEC when PNEC based on short term data and available long term daphnia data)*	exposure scenario should be used to select appropriate tests with fish, amphibians or birds** (TG 229, TG 230, TG 231, TG 234, GD 148, FLCTT, MMGT, LAGDA, TG 206, ATGT).
≥ 100 t	Long term toxicity in invertebrates (*Daphnia*) Long term toxicity in fish Fish early life stage test Fish short term toxicity (embryo and sac fry)*** Fish juvenile growth test*** Bioaccumulation in aquatic species (fish) Short term terrestrial toxicity (invertebrates)	Read across from all toxicological data generated from the proposed changes in Table 2. If ED alerts are identified the MoA and exposure scenario should be used to select appropriate tests with fish, amphibians or birds (TG 229, TG 230, TG 231, TG 234, GD 148, FLCTT, MMGT, LAGDA, TG 206, ATGT). A FSDT (TG 234) is recommended (see also***).
≥ 1000 t	Long term toxicity to invertebrates Long term toxicity to sediment organisms Long term or reproductive toxicity to birds**	Read across from all toxicological data generated from the proposed changes in Table 2. If ED alerts arise from these information´s the MoA and exposure scenario should be used to select appropriate tests with fish, amphibians or birds (TG 229, TG 230, TG 231, TG 234, GD 148, FLCTT, MMGT, LAGDA, TG 206**, ATGT). A FSDT (TG 234) is recommended (see also ***).

TG 206: Avian Reproduction Test, TG 229: Fish Short Time Reproduction Assay (FSTRA) (equals EPA/600/R-01/067), TG 230: 21-day Fish Assay, TG 231: Amphibian Metamorphosis Assay, TG 234: Fish Sexual Development Test (FSDT), FLCTT: Fish Life cycle toxicity test (FLCTT) (US EPA OPPTS 850.1500), MMGT: Draft Medaka Multi-Generation Test, LAGDA: Draft Larval Amphibian Growth and Development Assay, ATGT: Draft Avian Two Generation Test, OECD GD 148: Guidance Document on the Androgenised Female Stickleback Screen.

* (Pers. Com. Danish EPA). In few cases long term fish toxicity testing may we warranted in relation to completing PBT assessments, namely in those cases where the P and the B criterion (but not vPvB criterion) is fulfilled and it is required to check if the T criterion is fulfilled as regards the chronic fish toxicity NOEC–value. In cases where also no chronic Daphnia NOECs are available, such Daphnia data will be required before long term fish toxicity test may be warranted. (Need for such long term fish NOEC value may for example be indicated for substances with confirmed PB –properties if the chronic Daphnia toxicity NOEC is just slightly above the T_{eco} criterion of < 0.01 mg/L)

** (Pers. Com. Danish EPA). According to the REACH endpoint specific guidance document it will only in very exceptional cases be relevant to consider toxicity testing on birds and this may very well also apply in relation to ED related long term toxicity testing, so this aspect is not discussed further here.

*** (Pers. Com. Danish EPA). These fish toxicity tests may be accepted if data are already available but if new testing is being proposed, a fish early life stage test is requested.

The current standard information requirements with regard to aquatic toxicity are driven by the need for a chemicals safety assessment, which includes a quantitative risk assessment (RCR= PEC/PNEC) and the PBT assessment. It is proposed that because EDs with emission potential just like vPvBs and PBTs also belong to the category of Substances of Very High Concern that the concept of chemical safety assessment is explicitly being broadened to also covering an ED assessment when relevant. Hence, if there is an alert for ED properties from screening of all

available toxicological and ecotoxicological information at any tonnage level above 10 tpa/ year, the registrant should - in line with the requirements and testing strategy for potential vPvB/PBTs - be required to conduct further ED targeted relevant testing on aquatic organisms (fish and amphibians) unless there is no/negligible exposure of the aquatic environment.

Alert factors for ED effects are described in section 3.1.1 including alerts from epidemiological studies, QSAR´s, *in vitro* and *in vivo* studies. As described in section 3.1.1 i*n vitro* screens are relevant for effects in vertebrate wildlife also. The validated *in vitro* assays for which guidance is provided in the OECD GD 150 (OECD 2012) are listed in section 3.1.1.

In conclusion, *in vitro* assays for interaction with different ED modalities e.g. ER, AR and steroidogenesis interference, should be conducted to elucidate whether there are alert(s) for further testing for ED effects. *In vitro* assays concerning mechanisms of thyroid disruption may later when validated be considered for inclusion in the testing requirements.

Any alert should trigger further testing, taking into account exposure scenario and tonnage level. If clear alerts of oestrogenic agonism, androgenic agonism or steroidogenesis inhibition appears from QSAR, chemical categorization, *in vitro* and/or toxicological *in vivo* studies, and for which aquatic exposure is possible, it is proposed to request the a FSDT (TG 234), where MoA (vitellogenin induction or reduction) may be related to adverse effect on phenotypic sex ratio. If the evidence of e.g. oestrogenic activity is weak, the Fish Short Time Reproduction Assay (TG 229) or the 21-day Fish Assay (TG 230) could alternatively be triggered in accordance with the OECD Fish Testing Strategy Guidance Document (OECD 2012, STA 171). These tests can be used as screening tests which can inform about *in vivo* oestrogenic/androgenic MoA before it is decided whether to conduct the longer and more thorough and potentially confirmatory FSDT. *In vivo* androgenic antagonism MoA may be investigated by use of the Androgenised Female Stickleback Screen (AFSS (OECD GD 148)). This screen does, however, not inform about adverse effects. A Fish Life Cycle Toxicity Test (FLCTT) including endocrine relevant endpoints, a draft Medaka Multi-Generation Test (MMGT) or a draft Avian Two Generation Test (ATGT) could inform about adverse androgenic antagonistic effects. However, none of these draft test TGs are currently fully validated and adopted. Hence their use as standard information requirement relevant tests under even an adapted REACH standard information requirement scheme, which includes requirements relating to ED assessment, seems presently not to be a realistic option. Thyroidal effects could be tested *in vivo* using the AMA (TG 231) and adversity confirmed by a FLCTT with endocrine relevant endpoints included or a MMGT or a LAGDA.

In conclusion, it should be noted that currently the FSDT (TG 234) is the only validated OECD TG on non-mammalian vertebrate species which can inform about both endocrine activity and consequential adverse effects, i.e. if positive it can be used for a definitive identification of endocrine disrupters.

The ecotoxicological tests with included ED relevant endpoints that are either validated or under validation are more detailed described in the Annex and are further discussed in the following tonnage sections 3.3.1.1-3.3.1.4.

3.3.1.1 Tonnage level 1-10 tpa/manufacturer or importer

The required short-term invertebrate toxicity with Daphnia does not include any ED relevant endpoints or information´s on ED activity. The steps described in section 3.1.1 of gathering all available test data on the substance to be registered as well as all other available and relevant information on the substance regardless of whether testing for a given endpoint is required or not at the specific tonnage level should be followed. The same should been done with regard to ED to be able to evaluate whether there are alerts from the existing toxicological database for ED related effects. The alerts should be viewed as a guide of indicators that would provide input to a Weight of Evidence analysis requiring expert judgement that leads to the most appropriate testing strategy. If there are such alerts, the expected MoA and exposure scenario should be included in the decision of which environmental *in vivo* tests that should be performed as described above and in section 4. **In conclusion, it is recommended to include a formalised standard information procedure in REACH, including conduction of in vitro assays, to secure a targeted searching for ED alerts.**

3.3.1.2 Tonnage level 10-100 tpa/ manufacturer or importer

The required short-term toxicity test on fish (e.g. TG 203) does not include any ED specific endpoints or information´s on ED activity.

The results from the available toxicological studies proposed in Table 2 should be included in the evaluation of the substance. **In conclusion, if there are indications of ED effects from these data or *in vitro* or QSAR information´s, one or more ecotoxicological test with ED relevant endpoints, defined by MoA and exposure scenario, should be performed on aquatic organisms as described above and in section 4.**

3.3.1.3 Tonnage level 100 -1000 tpa/manufacturer or importer

None of the usual tests to fulfil the required information (Table 3) include any informative ED relevant endpoints although growth in fish potentially could respond to thyroidal effects. As for the < 100 tpa level, the results from the available toxicological studies proposed in Table 2 should be included in the evaluation of the substance. If there are indications of ED effects from these data or *in vitro* or QSAR information's, one or more ecotoxicological test with ED relevant endpoints, defined by MoA and exposure scenario, should be performed on aquatic organisms. However, as mentioned previously the Fish Sexual Development Test (TG 234) is a Fish Early Life Stage Test targeted towards detection of ED relevant activity and resulting adverse effects and **it is recommended to require the Fish Sexual Development Test as the standard test when a Fish Early Life Stage Test is warranted and an alert of oestrogenic agonism, androgenic agonism or steroidogenesis inhibition appears from QSAR, *in vitro* and/or toxicological *in vivo* studies.**

3.3.1.4 Tonnage level more than 1000 tpa/ manufacturer or importer

The required TG 206 avian reproduction test (See ** below Table 3) could potentially inform about avian ED effects (e.g. growth may respond to some thyroid disrupters; % cracked eggs and egg shell thickness may respond to chemicals interfering with the control of shell deposition (OECD 2012)) but the guideline does not include ED specific endpoints. The test could though inform about adverse apical effects on development, growth or reproduction over the reproductive part of the avian lifecycle. A potential limitation of TG 206 is that the effects of test chemicals may not

become fully apparent during the test because the offspring are not directly dosed, and only receive bio accumulated material which may be passed from their mothers via the egg (OECD 2012).
For this tonnage level the results from the available toxicological studies proposed in Table 2 should be included in the evaluation of the substance. If there are indications of ED effects from these data or *in vitro* or QSAR information´s, one or more ecotoxicological test with ED relevant endpoints should be performed on aquatic organisms. As this tonnage level may lead to widespread environmental exposure to a chemical **it is recommended to include the FSDT (TG 234) as a standard information requirement for this tonnage level when an ED alert appears from QSAR, *in vitro* and/or toxicological *in vivo* studies.**

3.3.2 Plant Protection Products and Biocidal Products Regulations – data requirement and proposals for changes to include endocrine disruption

Until recently PPPR did not contain any data requirements specifically targeting identification of endocrine disruption. By the new regulation (EU 283/2013) the data requirements have been updated and a Commission communication (2013/C 95/01) describes the relevant test methods and Guidance Documents. This new regulation requires that information about potential endocrine disrupting properties of the active substance is considered and that further studies for investigation should be discussed with the competent authorities. The two quotes below directly describe the requirements for information on birds and other terrestrial vertebrates and aquatic organisms:

For birds and other terrestrial vertebrates (paragraph 8.1.5):
"Consideration shall be given to whether the active substance is a potential endocrine disruptor according to Union or internationally agreed guidelines. This may be done in consulting the mammalian toxicology section (see Section 5 in Commission Regulation (EU) No 283/2013 of 1 March 2013). In addition, other available information on toxicity profile and mode of action shall be taken into account. If as a result of this assessment, the active substance is identified as a potential endocrine disruptor, the type and conditions of the study to be performed shall be discussed with the national competent authorities".

For aquatic organisms (paragraph 8.2.3):
"Consideration shall be given to whether the active substance is a potential endocrine disruptor in aquatic non-target organisms according to Union or internationally agreed guidelines. In addition, other available information on toxicity profile and mode of action shall be taken into account. If as a result of this assessment, the active substance is identified as a potential endocrine disruptor, the type and conditions of the studies to be performed shall be discussed with the national competent authorities".

The Commission communication (2013/C 95/01) lists the following tests as relevant for identification of endocrine disrupting properties, see Annex 1 for further details:
OECD TG 229: Fish Short Time Reproduction Assay
OECD TG 230: 21-day Fish Assay
OECD TG 231: Amphibian Metamorphosis Assay
OECD TG 234: Fish Sexual Development Test

Both TG 229 and TG 230 could inform about *in vivo* MoA. TG 229 has been validated for oestrogenic antagonistic effects but lack the connection between MoA and adversity. As mentioned under section 3.3.1 there is currently only one OECD validated test, the Fish Sexual Development Test (TG 234) that definitively can show specific endocrine disrupting adverse effects and only regarding interference with sex hormones and steroidogenesis. As mentioned earlier a FLCTT with endocrine relevant endpoints included, a MMGT or an ATGT could inform about adversity. Thyroidal effects could be tested *in vivo* using the AMA (TG 231) and adversity confirmed by a FLCTT with endocrine relevant endpoints included or a MMGT or a LAGDA. TG 206 could potentially inform about avian thyroid effects but the guideline does not include thyroid specific endpoints. The draft ATGT could possibly inform about thyroid effect but further validation is needed**.**

In conclusion, currently the FSDT (TG 234) is the only validated OECD TG on non-mammalian vertebrate species which can inform about both endocrine activity and consequential adverse effects, i.e. if positive, it can be used for a definitive identification of endocrine disrupters. Therefore, this test should be considered for pesticides when there are alert(s) for ED effects (and in particular if indicating oestrogenicity) in aquatic vertebrates unless the effects are exclusively related to thyroidal effects. In that case, a TG 231, and if positive, a confirmatory LAGDA test should be requested (when the LAGDA test has been fully validated and adopted as an OECD TG or if it is possible before this time to agree to a testing protocol which may be used in a regulatory context case by case). If reproductive effects not related to oestrogenicity or to development are suspected, test requirements according to the draft OECD TGs MMGT or FLCTT with ED endpoints should be considered case by case, if possible.

The BPR does not contain specific data requirements with regard to endocrine disruption. However, by indication of ED, information concerning identification of endocrine activity may be required as additional data (ADS). ADS requirement is defined in Annex 2 paragraph 2 as: "*Data elements to be provided for a specific active substance shall be determined by considering each of the ADS data elements indicated in this Annex taking into account, inter alia, the physical and chemical properties of the substance, existing data, information which is part of the CDS and the types of products in which the active substance will be used and the exposure patterns related to these uses*." Hence, depending on use pattern and level of suspicion of ED, biocides active ingredients shall be investigated for ED but this is not specified further in the BPR itself.

However, guidance on how to fulfil the data requirements in relation to biocides are in preparation and it is expected that the listed test methods will resemble the new data requirements for the active ingredients of plant protection products (Danish EPA, pers.com.).

To provide an overview in relation to ecotoxicological information, the standard information requirements which may be relevant for endocrine properties according to PPPR and BPR (and in addition REACH) are summarized in Table 4. Most of the information requirements refer to results

generated by specific OECD test guidelines or their EU Test Method equivalents. Test methods validated by other organisations (e.g. US EPA) may also be used.

Table 4. Requirements for ecotoxicological information for higher organisms (other types of tests without relevance for endocrine aspects, e.g. tests on degradation and adsorption, plants, microorganisms, etc. are not included). For some of the information demands, the OECD/US EPA validated test methods that may provide the information are indicated in parenthesis.

Examples of tests that covers the Information demand			REACH		PPPR	BPR	Endocrine relevant
	≥1 t	≥10 t	≥100 t	≥1000 t			endpoints
Acute toxicity to daphnia (TG 202)	X				X	X	None
Acute fish toxicity (TG203)		X§			X[A]	X[B]	None
Acute toxicity for aquatic gastropod, insect and non-daphnian crustacean (e.g. TG 235)					X		None
Long term or chronic toxicity in invertebrates (Daphnia) (TG211)			X§		X	X§[C]	Some[F]
Chronic toxicity for aquatic gastropods and insects					X		Some[F]
Long term or chronic toxicity in fish			X§		X	X§	Some[F]
Fish early life stage toxicity test (TG210)			X§		X[#]	X§#	None
Fish, embryo and sac fry (TG212)*			X§			X§#	None
Fish Short Time Reproduction Assay (TG 229)			X§		X§		Yes
21-day Fish Assay (TG 230)			X§		X§		Yes
Amphibian Metamorphosis Assay (TG 231)			X§		X§		Yes
Fish Sexual Development Test (TG234)			X§		X§		Yes
Fish juvenile growth test (TG215)*			X§			X§#	None
Chronic toxicity in juvenile fish					X[#]		None
Life cycle test with fish (US EPA OPPTS 850.1500)					X[#]		Some[F]
Bioaccumulation in fish (TG305)			X§		X	X§	None
Acute earthworm toxicity (TG207)			X§		X	X	None
Long term or sublethal toxicity to invertebrates (TG211 or TG222)				X§	X		Some[F]
Acute toxicity to bees (TG213/214)					X	X§	None
Feeding test for bee larvae					X		None
Effects on two non-target arthropods (besides the bees) (TG226/232)					X		None
Long term toxicity to sediment organisms (TG218/219)				X§	X§	X§	Some[F]
Acute oral bird toxicity, 2 species (TG223)					X[D]	X[E]	None
5 or 8 days oral bird toxicity (feeding) (TG205)					X[G]	X§	None
Long term/sub-chronic or reproductive toxicity to birds (TG206)				X§	X[G]	X§	Some[F]

[A]: two species (rainbow trout & a warm water species). [B]: two species. [C]: reproduction. [D]: mallard and quail. [E]: two species with different feeding habits. [F]: Effects on reproduction are determined. [G]: Test Guideline specified in regulation. [H]: EPA/600/R-01/067 is almost equivalent to OECD TG 229. [§]: "If the chemicals safety assessment indicates the need". [#]: May replace each other pending expert assessment.
Some of the legal requirements for information concerning invertebrates have been interpreted as daphnia- and earthworms tests for aquatic and terrestrial invertebrates, respectively, in the table.
* It should be noted that currently there appears to be limited demand for OECD TG 204, TG 212, and TG 215; thus, the retention of these guidelines should be evaluated (Danish EPA, pers.comm.).

It is obvious from Table 4 and the discussion in section 3.3 that the ecotoxicological tests performed to meet the requirements for information under the present regulations are – with few exceptions - not designed to reveal endocrine disrupting effects. Some of the tests are able to show adverse outcomes on reproduction, but generally they do not inform about whether the adverse outcomes are caused by endocrine mechanisms/modes of action or by other modes of action e.g. related to general systemic toxicity.

Note that the list in Table 4 does not include the rodent/rabbit tests normally used for toxicity (human health) purposes even though such tests are relevant to consider in relation to EDs and their effects on wildlife mammalian species. As a starting point it can simply be concluded that if a substance causes ED related adverse effects in laboratory rodents or rabbits in toxicity studies such findings are ecotoxicologically relevant for mammalian wild-life or at least there should be very good specific reasons provided for not considering that. Hence the same recommendations for enhancements of the OECD TGs on mammalian species mentioned above for human health (section 3.1) also apply for environmental effects.

The new data requirements for environmental effects of active substances in PPPs (EU 283/2013) imply that it should be considered whether the active substance is an endocrine disrupter. There is, however, no specific data requirements with regard to endocrine disruption and the suggested tests to elucidate ED related effects are only triggered if the available information indicates a potential for endocrine disruption.

It is suggested that the strategy presented for changes of the standard information requirements in REACH as outlined in sections 3.3.1.1 – 3.3.1.4 should also be followed with regard to plant protection products and biocidal products. This would assure that their active ingredients in supplement to the current regulatory assessments also go through an ED related targeted assessment and testing strategy as the basis for considering approval or non-approval of applications for their use in biocidal products and plant protection products.

4. Testing strategy when there is a concern for ED

Several chemical substances are under suspicion for having endocrine disrupting properties. In order to provide guidance for further investigation of different kinds of indications of endocrine activity a proposal for an information/testing strategy is developed.

This section contains a mode of action based outline for a general ED testing strategy that can be used by industry and authorities. It can therefore be used to provide input to ED related targeted testing strategies for substance evaluation under REACH and for potentially contributing to enhanced testing strategies to be established concerning active substances of plant protection products and biocides. As mentioned before more detailed ED testing strategy considerations can be found in OECD GD Nos. 150 & 171.

The strategy for testing will depend on the observations that raise the alert for endocrine disrupting properties of a chemical. The alert may be based on either MoA data from QSAR, *in vitro* or *in vivo* studies or adverse effect data.

An overview of the ED testing strategy when there is concern for ED mode of action based on *in silico* or *in vitro* data is shown in figure 2, whereas figure 3 shows an overview of the strategy when the concern is based on *in vivo* data.

The objectives and scope of the OECD GD is to be a tool to support regulatory authorities by helping to interpret assay results and suggesting possible additional studies for reducing uncertainty (OECD, 2012). The present proposal for a testing strategy will primarily be based on the screens and tests handled in the GD. These tests are all taken from the “OECD Conceptual Framework for the Screening and Testing of Endocrine Disrupting Chemicals” (CF) which was modified and updated by the Endocrine Disrupters Testing and Assessment Advisory Group (EDTA AG) in 2011. The OECD GD states that lower level tests in the OECD CF should generally be conducted before higher level tests in order to avoid unnecessary animal usage unless it is apparent that a Level 5 test will be required anyway. In this context it should be noted that the OECD CF “is not a testing strategy to be followed linearly from level 1 through to level 5, although in cases where little or no information is available (i.e. for new chemicals) it could provide ideas about where to start testing. In principle, any test can be conducted at any time in the hazard assessment process, depending on the perceived need for information” (OECD, 2012).

At Level 5, the extended one-generation study (OECD TG 443) is the most sensitive reproductive toxicity assay for detecting endocrine disruption in mammals because it provides an evaluation of a number of endocrine endpoints in the juvenile and adult F1, which are not included in the two-generation study (OECD TG 416) adopted in 2001. With regard to level 5 for wildlife tests, no tests with endocrine specific endpoints (except for sex ratio in fish Lifecycle toxicity test (FLCTT)) have yet been adopted. However, many of the endpoints in these apical tests are nevertheless affected by EATS EDs. Of particular interest in the context of oestrogens, androgens and steroidogenesis disrupters are time to sexual maturity, fecundity as well as fertility.

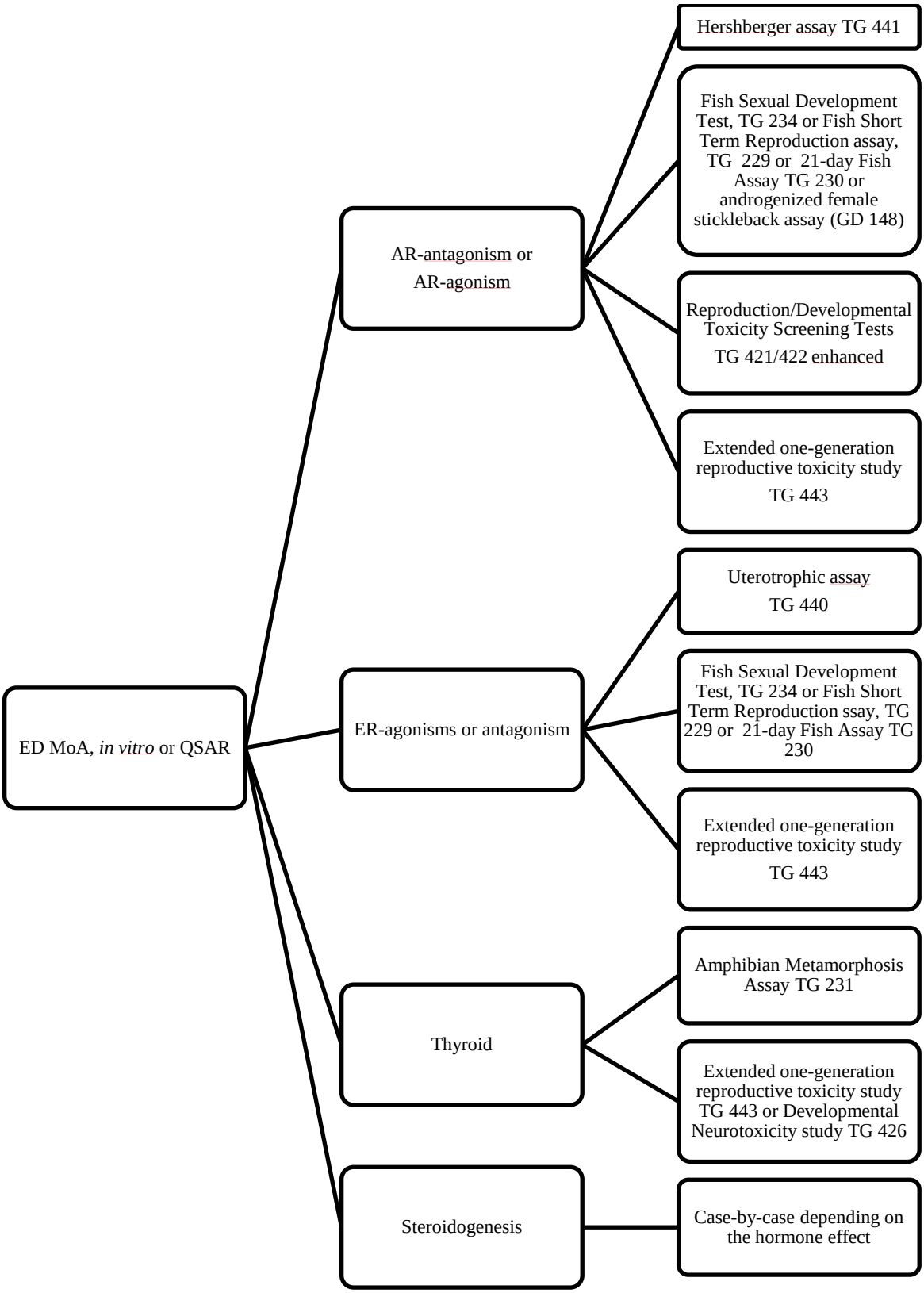

Figure 2 Overview of the ED testing strategy when there is concern based on mode of action data *in silico* or *in vitro*, see text for details.

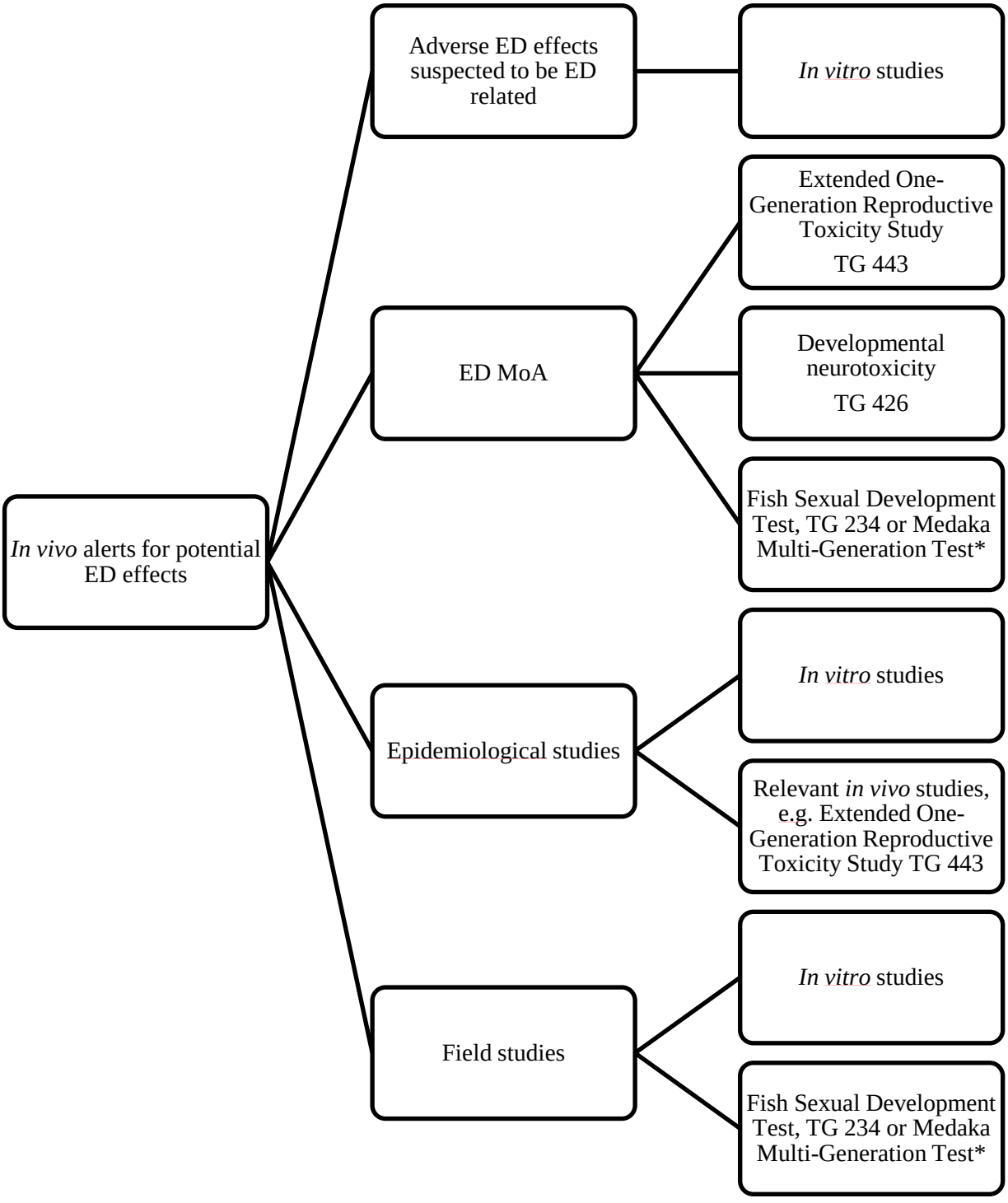

Figure 3 Overview of the ED testing strategy when there is concern based on *in vivo* data, see text for details.* Currently draft OECD TG

It should be recalled that due to the molecular similarities of endocrine systems and receptor homologies across the vertebrates, there may be some potential for using information from non-mammalian vertebrate test assays for assessing endocrine activity in mammals (and *vice versa*), and especially for extrapolation between various *in vitro* screens. This must be tempered with the knowledge that outcomes associated with a given endocrine modality can vary significantly across the vertebrates.

Indicators of *in vivo* endocrine activity might include 'read across' from *in vivo* results obtained with chemically related chemicals, or positive results from an *in vitro* screen for oestrogen or androgen receptor-mediated activity, or for effects on steroidogenesis (especially aromatase inhibition).

4.1 Alerts from *in silico* models

Alerts for endocrine disrupting properties from *in silico* modelling (*e.g.* quantitative structure–activity relationship (QSAR) predictions of endocrine activity) will typically be based on the structural properties of the chemical in relation to a specific receptor binding site. Therefore, the mechanism of action will be indicated and the obvious next step is to test the activity of the chemical in an *in vitro* test system containing the relevant hormone receptor if such a system is available. However, if the QSAR predictions are within the applicability domain, *in vitro* testing may be restricted to QSAR-predictions that are outside the applicability domain as the QSAR-models are based on in vitro test data. If no validated *in vitro* systems are available for the mechanism of action in focus it should be considered (based on physico-chemical properties of the chemical and the expected exposure scenario) to perform *in vivo* screening.

4.2 Alerts from *in vitro* assays

The *in vitro* screens (although at present based largely on mammalian receptors and/or enzymes) are generally capable of providing information applicable to both humans and vertebrate wildlife (OECD, 2010, OECD 2012).

Alert for endocrine disrupting properties raised from *in vitro* test systems are typically based on the ability of a chemical to bind to and activate or inhibit a specific hormone receptor or the ability of the chemical to interfere with the synthesis of a specific hormone (i.e. H295R Steroidogenesis Assay TG 456). The *in vitro* systems typically inform about the mechanism of action without informing about the mechanism of action *in vivo* or potential adverse effects. Therefore, the logical next steps will be to test the mechanism of action *in vivo* or the potential adverse effects of the chemical at the higher levels 3, 4 or 5 in OECD CF.
Alerts from *in vitro* testing will typically concern oestrogenic, anti-oestrogenic, androgenic, anti-androgenic, thyroid interfering mechanisms or interference with steroidogenesis and the choice of subsequent testing strategy will depend on the type of mechanism seen.

4.2.1 Anti-androgenicity in vitro, i.e. AR-antagonism

Mammals
The mechanism of action can be tested *in vivo* using the Hershberger assay (OECD TG 441) which has been validated for detection of anti-androgenic mode of action. An alternative model, i.e. the intact (uncastrated) weanling rat was also validated due to animal welfare concerns with the castration procedure but did not seem to consistently detect weak anti-androgenic chemicals at the doses tested.

Alternatively and depending on for example the tonnage level in relation to REACH, the substance can be tested using either the OECD TG 421/422 or TG 414 both enhanced with anti-androgenic endpoints (e.g. AGD) or the Extended One-Generation Reproductive Toxicity Study (TG 443).
In the validation process of OECD TG 407 assay (Repeated dose 28-day toxicity study) with endocrine endpoints, some studies failed to identify EDs that weakly affected androgen receptors as was demonstrated on the basis of data generated. In this validation only moderate EDs, such as nonylphenol and DDE, and strong EDs such as ethinylestradiol and flutamide (acting via ER and AR respectively) were detected. Thus OECD TG 407 cannot be regarded as a sensitive screening assay for endocrine activity and is therefore not recommended.

Avian
No avian tests have yet been validated with endocrine specific endpoints included. The OECD TG 206 Avian Reproduction Test (level 4) does, however, include endpoints as embryo viability, and hatchability, egg production, cracked eggs and eggshell thickness which could potentially be sensitive to EATS modalities. TG 206 could therefore be used to investigate whether a suspected ED result in adverse apical effects on development, growth or reproduction over the reproductive part of the avian lifecycle (OECD 2012). It should be noted that an avian two-generation test (ATGT) is under development including endpoints as sex ratio (phenotypic and/or genotypic), sex hormones, thyroid hormones, reproductive/thyroid organ weights, gonad histopathology and gross pathology, time to first egg laying, and sexual behaviour. When the ATGT is validated it could inform about anti-androgenic effects in avian species as Japanese quail (*Coturnix japonica*).

Amphibians
No amphibian tests have yet been developed to detect anti-androgenic modes of action. A Larval Amphibian Growth and development Assay (LAGDA) OECD draft proposal is currently under development and includes endpoints as vitellogenin, T4 and TSH, hormone titres, snout-vent length, body weight, thyroid and gonad histopathology, time to metamorphosis, nuptial pad development and phenotypic sex ratio. This proposed assay could inform about anti-androgenic effects in amphibians.

Fish
Anti-androgenic effects in fish can be tested *in vivo* using the Fish Short-Term Reproduction Assay (TG 229) with suppression of male secondary sex characteristics as an endpoint. The androgenized female stickleback screen is described in a Guidance Document (OECD GD 148) as a special

version of OECD TG 230 using the depression of the androgen dependant protein spiggin in androgenized female three-spined sticklebacks (*Gasterosteus aculeatus*) to confirm anti-androgenic MoA. The MoA has not been validated in the fish sexual development test TG 234 but a sex ratio skewed toward females is expected.

Invertebrates
As described in OECD GD and the revised CF (Footnote 4) at present the available invertebrate assays solely involve apical endpoints which are able to respond to some endocrine disrupters and some non-EDs. Those in level 4 are partial lifecycle tests, while those in level 5 are full- or multiple lifecycle tests. Invertebrate tests that could inform about ED effects on apical endpoints are at CF level 4: Chironomid toxicity test (TG 218-219), Daphnia reproduction test (with male induction) (OECD TG 211), Earthworm reproduction test (OECD TG 222, 2004), Enchytraeid reproduction test (OECD TG 220, 2004), Sediment water lumbriculus toxicity test using spiked sediment (OECD TG 225, 2007), Predatory mite reproduction test in soil (OECD TG 226, 2008), Collembolan reproduction test in soil (TG OECD 232, 2009) and at CF level 5: Sediment water chironomid life cycle toxicity test (OECD TG 233).

4.2.2 Androgenicity in vitro, i.e. AR-agonism

Mammals
Androgen mode of action can be studied in the Hershberger assay (TG 441). This assay is sensitive to both androgens and anti-androgens. An alternative model, i.e. the intact (uncastrated) weanling rat was also validated due to animal welfare concerns with the castration procedure seemed consistently to detect androgenic chemicals. Androgenic chemicals cause growth of the sex accessory tissues in this TG.

Studies using agonists are lacking in relation to TG 443. However, androgenic effects may show up in the following way: Increased AGD in male pups, changes in AGD in female pups, age at preputial separation (F1), genital abnormalities, weights of uterus, ovaries, testes, epididymides, prostate, seminal vesicles (including coagulating glands), histopathologic changes in the above organs and in mammary glands and changes in sperm parameters (P, F1).

Avian
No avian tests have yet been validated with endocrine specific endpoints included. The OECD TG 206: Avian Reproduction Test (level 4) does however include endpoints as embryo viability, and hatchability, egg production, cracked eggs and eggshell thickness which could potentially be sensitive to EATS modalities. TG 206 could therefore be used to investigate whether a suspected ED result in adverse apical effects on development, growth or reproduction over the reproductive part of the avian lifecycle (OECD 2012). It should be noted that an avian two-generation test (ATGT) is under development including endpoints as sex ratio (phenotypic and/or genotypic), sex hormones, thyroid hormones, reproductive/thyroid organ weights, gonad histopathology and gross pathology, time to first egg laying, and sexual behaviour. When the ATGT is validated it could inform about androgenic effects in avian species as Japanese quail (*Coturnix japonica*).

Amphibians
No amphibian tests have yet been developed to detect androgenic MoA. A Larval Amphibian Growth and development Assay (LAGDA) OECD draft proposal is currently under development and includes endpoints as vitellogenin, T4 and TSH, hormone titres, Snout-vent length, body weight, thyroid and gonad histopathology, time to metamorphosis, Nuptial pad development and phenotypic sex ratio. This proposal could inform about androgenic effects in amphibians.

Fish
The Fish Short-Term Reproduction Assay (TG 229) and the 21-day Fish Assay (TG 230) can inform about androgenic MoA if either Japanese Medaka or Fathead minnow is used as test species. The secondary sex characteristics and vitellogenin levels are the androgen sensitive endpoints. Also, the Fish Sexual Development Test (FSDT, TG 234) is well suited to demonstrate androgenic effects which are identified in sex ratios skewed towards males and lowered levels of vitellogenin in females. In the FSDT, the androgenic MoA can be linked to adverse effect at population level (sex ratio).

Invertebrates
As described in OECD GD the available invertebrate assays solely involve apical endpoints which are able to respond to some endocrine disrupters and some non-EDs. Those in level 4 are partial lifecycle tests, while those in level 5 are full- or multiple lifecycle tests.

4.2.3 Oestrogenicity in vitro, i.e. ER-agonism

Mammals
The *in vivo* mode of action can be studied in the Uterotrophic assay (OECD TG 440).
TG 443 can detect adverse effects related to oestrogenic mode of action and the effects may show up in the following way: Change in AGD in male and female pups, effect on oestrus cyclicity (P, F1), earlier age at vaginal opening (F1), possibly affected age at preputial separation (F1), genital abnormalities, changed weights of uterus, ovaries, testes, epididymides, prostate, seminal vesicles (including coagulating glands), histopathologic changes in the above mentioned organs and mammary glands and changes in sperm parameters (P, F1).

Avian
As already mentioned above no avian tests including endocrine specific endpoints have yet been validated. The OECD TG 206 Avian Reproduction Test (level 4) does however include endpoints as embryo viability, and hatchability, egg production, cracked eggs and eggshell thickness which could potentially be sensitive to oestrogenic effects. TG 206 could therefore be used to investigate whether a suspected oestrogen result in adverse apical effects on development, growth or reproduction over the reproductive part of the avian lifecycle (OECD 2012). It should be noted that an avian two-generation test (ATGT) is under development including endpoints as sex ratio (phenotypic and/or genotypic), sex hormones, thyroid hormones, reproductive/thyroid organ weights, gonad histopathology and gross pathology, time to first egg laying, and sexual behaviour.

When the ATGT is validated it could inform about oestrogenic effects in avian species as Japanese quail (*Coturnix japonica*).

Amphibians
No amphibian tests have yet been developed to detect oestrogenic MoA. A Larval Amphibian Growth and development Assay (LAGDA) OECD draft proposal is currently under development and includes endpoints as vitellogenin, T4 and TSH, hormone titres, Snout-vent length, body weight, thyroid and gonad histopathology, time to metamorphosis, Nuptial pad development and phenotypic sex ratio. This proposal could inform about oestrogenic effects in amphibians – especially with the endpoints vitellogenin and phenotypic sex ratio.

Fish
Since vitellogenin induction and skewing of sex ratios towards females are very sensitive markers of oestrogenic effects in fish, tests including fish are probably more efficient in revealing oestrogenic mode of action *in vivo* than tests with mammals. At OECD conceptual framework level 3, vitellogenin induction in male fish is used as a biomarker for oestrogenic effect in the Fish Short-Term Reproduction Assay (TG 229) and the 21-day Fish Assay (TG 230). The induction of vitellogenin confirms the oestrogenic mode of action but it is not in itself an adverse effect. Besides vitellogenin induction, egg production (fecundity) is determined in TG 229 and changes in the egg production may be an indicator of adverse effects. However, the robustness of the fecundity as an endpoint in the current design of TG 229 is under discussion because the statistical power is low and therefore, the risk of a false negative outcome is high. So if the purpose is to demonstrate adverse effects (as well as the mode of action via the vitellogenin induction) it may be considered to skip the level 3 tests and proceed directly to the Fish Sexual Development test (TG 234) at level 4.

Invertebrates
As described in OECD ED GD (OECD 2012) the available invertebrate assays solely involve apical endpoints which are able to respond to some endocrine disrupters and some non-EDs. Those in level 4 are partial lifecycle tests, while those in level 5 are full- or multiple lifecycle tests.

4.2.4 Anti-oestrogenicity in vitro

Mammals
Anti-oestrogenic mode of action in vivo may be studied in the Uterotrophic Assay (TG 440). The OECD validation of the Uterotrophic assay was, however, not considered adequate for anti-oestrogens as there were insufficient pure anti-oestrogens available (OECD GD). Thus, the OECD guideline for the Uterotrophic assay (TG 440) is for oestrogen agonists only. The use of the Uterotrophic assay as a test for anti-oestrogenenicity is, however, frequently used and it is described in an OECD guidance document (OECD, 2007a).

Studies using ER antagonists are lacking in relation to TG 443. However, anti-oestrogenic effects may show up in the following way: Dystocia (problems with giving birth), changes in AGD in male and female pups, age at vaginal opening (F1), effect on oestrus cyclicity (P, F1), genital

abnormalities, weights of uterus, ovaries, testes, epididymides, prostate, seminal vesicles (including coagulating glands), histopathologic changes in the above organs and in mammary glands and changes in sperm parameters (P, F1).

Avian
No avian tests have yet been validated with endocrine specific endpoints included. The OECD TG 206: Avian Reproduction Test (level 4) does however include endpoints as embryo viability, and hatchability, egg production, cracked eggs and eggshell thickness which could potentially be sensitive to anti-oestrogenic activity. TG 206 could therefore be used to investigate whether a suspected ED result in adverse apical effects on development, growth or reproduction over the reproductive part of the avian lifecycle (OECD 2012). It should be noted that an avian two-generation test (ATGT) is under development including endpoints as sex ratio (phenotypic and/or genotypic), sex hormones, thyroid hormones, reproductive/thyroid organ weights, gonad histopathology and gross pathology, time to first egg laying, and sexual behaviour. When the ATGT is validated it could inform about anti-oestrogenic effects in avian species as Japanese quail (*Coturnix japonica*).

Amphibians
No amphibian tests have yet been developed to detect anti-oestrogenic MoA. A Larval Amphibian Growth and development Assay (LAGDA) OECD draft proposal is currently under development and includes endpoints as vitellogenin, T4 and TSH, hormone titres, Snout-vent length, body weight, thyroid and gonad histopathology, time to metamorphosis, Nuptial pad development and phenotypic sex ratio. This proposal could inform about anti-oestrogenic effects in amphibians – especially with the endpoints vitellogenin and phenotypic sex ratio.

Fish
The anti-oestrogenic mode of action can be confirmed *in vivo* at OECD conceptual framework level 3 in TG 229 or TG 230 by determining reductions in vitellogenin concentrations in female fish. Adversity can be shown in level 3 in TG 229 if fecundity is reduced significantly (a negative result on fecundity cannot be used to argue against anti-oestrogenicity due to low power) and at OECD conceptual framework level 4 in TG234 with a decrease in female vitellogenin concentration and a decline in the ratio of female phenotypic sex would be expected. The MoA has, however, not yet been validated in TG234.

Invertebrates
As described in OECD ED GD (OECD 2012) the available invertebrate assays solely involve apical endpoints which are able to respond to some endocrine disrupters and some non-EDs. Those in level 4 are partial lifecycle tests, while those in level 5 are full- or multiple lifecycle tests.

4.2.5 Thyroid in vitro

In vitro screening assays for disruption of thyroid function have not been validated and the Detailed Review Paper on Thyroid Hormone Disruption Assays (OECD, 2006a) concluded that: “The

complicated nature of the thyroid system makes development of an *in vitro* battery of assays to detect thyroid disruption unlikely in the near future. The conclusion is based on two facts: the *in vitro* assays available need further development before they can be validated, and the number of *in vitro* assays required to encompass every potential point of disruption in the thyroid system would be too great for a manageable assay battery. Furthermore, *in vitro* assays alone would not detect interactions within the thyroid system in response to toxicants. However, recommendations were made on *in vitro* assays that could be developed and utilised for high throughput screens in the near future". No guidance has therefore been written in the OECD GD. There is, however, use of these assays in research and therefore data may be available and could be considered as "existing data" for evaluating whether there are *in vitro* alerts for thyroid activity. A few of these assays are described in the "Detailed Review Paper on the State of the Science on Novel *in Vitro* and *in Vivo* Screening and Testing Methods and Endpoints for Evaluating Endocrine Disruptors" (OECD 2012b). For example did Schmutzler et al. (2007) develop a novel *in vitro* Thyroid Peroxidase (TPO) inhibition assay based on human recombinant TPO (hrTPO) stably transfected into the human follicular thyroid carcinoma cell line and Cao et al. (2010) utilized a novel fluorescence displacement method to investigate displacement of T4 from the thyroid receptor caused by thyroid disrupting chemicals.

Mammals

Agonistic or antagonistic effects on the thyroid receptor are difficult to study in mammals, and no specific tests are included in OECD CF. Thus, there is a lack of assays that can be used for studying thyroid receptor modes of action *in vivo*. If the suspected thyroid mode of action, however, affects the production, metabolism or excretion of thyroid hormones such effects can be studied *in vivo* by inclusion of measurements of thyroid hormones (T4, TSH) in e.g. repeated dose toxicity and reproductive toxicity studies.

Neither OECD TG 416 nor TG 443 has been validated for effects related to substances with thyroid activities but, as also written in the OECD GD, it is reasonable to suppose that thyroid changes in those TGs would be: Increased thyroid weight, possible liver weight increase (in combination with other thyroid-related endpoints), or histopathologic changes in thyroid (follicular cell height increase & colloid area decrease). Also, effects on thyroid hormone levels during brain development can affect behaviour. Behavioural effects can be detected in TG 443 in the developmental neurotoxicity module and in the Developmental Neurotoxicity study (OECD TG 426). Studies of functional effects incl. behavioural effects are not included or described in TG 416, but inclusion of endpoints for such effects are actually proposed in the TG if they are not included in other studies. This proposal is, however, rarely followed, but TG 416 can detect behavioural effects similarly as the TG 443 if a developmental neurotoxicity module is added.

Avian

The OECD TG 206: Avian Reproduction Test (level 4) does not include endocrine specific endpoints but as described in OECD GD some of the endpoints are potentially affected by endocrine disrupters: Egg production, embryo viability, and hatchability could be affected by

steroidogenesis disrupters, oestrogens and androgens and growth by thyroid disrupters. Also percentage cracked eggs and egg shell thickness may respond to endocrine disrupters.

Amphibians
The Amphibian Metamorphosis Assay (AMA, TG 231) can be used for the *in vivo* confirmation of thyroid *in vivo* mode of action and effects. Signs of thyroid activity include advanced development, asynchronous development, and delayed development in absence of non-specific systemic toxicity and thyroid histopathology. This assay covers several different modes of action, including thyroid agonists and antagonists, as well as substances interfering with thyroid hormone synthesis and transport. There is disagreement about the implications of the different endpoints in this larval development screen. As referred in the OECD ED GD 150, some experts accept that changes in one of the thyroid-relevant apical endpoints (advanced development; asynchronous development; delayed development in absence of non-specific systemic toxicity) may on their own indicate thyroid activity, while others will only reach this conclusion if one of the apical endpoints is accompanied by significant thyroid histopathology such as moderate or severe follicular hypertrophy and/or hyperplasia (OECD, 2007c). Note that the AMA is subject to indirect thyroid effects such as those that result from cytochrome P450 induction (e.g. phenobarbital, the model compound for the latter effect, tests positive in the AMA). Therefore, interpretation of the AMA may be complicated (OECD, 2012). A Larval Amphibian Growth and development Assay (LAGDA) OECD draft proposal is currently under development and includes endpoints as vitellogenin, T4 and TSH, hormone titres, snout-vent length, body weight, thyroid and gonad histopathology, time to metamorphosis, nuptial pad development and phenotypic sex ratio. This proposal could inform about adverse thyroidal effects in amphibians.

Fish
No fish tests with thyroid specific endpoints have yet been approved. Apical endpoints as e.g. growth could though be affected by thyroidal effects in for example TG 210, TG 234 and US EPA OPPTS 850.1500 (Fish Full Life-cycle Toxicity Test).

Invertebrates
As described in OECD GD the available invertebrate assays solely involve apical endpoints which are able to respond to some endocrine disrupters and some non-EDs. Those in level 4 are partial lifecycle tests, while those in level 5 are full- or multiple lifecycle tests.

4.2.6 Steroidogenesis in vitro

Interferences with steroidogenesis may lead to different effects in various test animals (e.g. prochloraz shows anti-androgenic effects in rats and androgenic effects in zebrafish) and, consequently, this complication must be in mind when testing strategies are considered.

Mammals
Interference with steroidogenesis may lead to higher or lower levels of the endogenous hormones. If decreased level of androgens is found, the substance can be tested using either the OECD TG

421/422 enhanced with ED endpoints or TG 443. If increased level of oestrogen is found the *in vivo* mode of action can be studied in the Uterotrophic assay (OECD TG 440) using immature rodent, because these animals are not ovariectomized and have an intact HPG axis. Adverse effects of increased levels of oestrogens can be studied in TG 443.

Avian
No avian tests have yet been validated with endocrine specific endpoints included. The OECD TG 206: Avian Reproduction Test (level 4) does however include endpoints as embryo viability, and hatchability, egg production, cracked eggs and eggshell thickness which could potentially be sensitive to EATS modalities. TG 206 could therefore be used to investigate whether a suspected ED result in adverse apical effects on development, growth or reproduction over the reproductive part of the avian lifecycle (OECD 2012). It should be noted that an avian two-generation test (ATGT) is under development including endpoints as sex ratio (phenotypic and/or genotypic), sex hormones, thyroid hormones, reproductive/thyroid organ weights, gonad histopathology and gross pathology, time to first egg laying, and sexual behaviour. When the ATGT is validated it could inform about steroidogenic disrupting effects in avian species as Japanese quail (*Coturnix japonica*).

Amphibians
No amphibian tests have yet been developed to detect steroidogenic MoA. A Larval Amphibian Growth and development Assay (LAGDA) OECD draft proposal is currently under development and includes endpoints as vitellogenin, T4 and TSH, hormone titres, Snout-vent length, body weight, thyroid and gonad histopathology, time to metamorphosis, nuptial pad development and phenotypic sex ratio. This proposed assay could inform about steroidogenic effects in amphibians – especially with the endpoints vitellogenin and phenotypic sex ratio.

Fish
Effects on steroidogenesis could affect vitellogenin concentration which is included in TG 229, 230 and 234. Adverse effects are covered by TG 234 where sex ratio change is a sensitive endpoint on e.g. aromatase inhibition. Most probably, it is the ratio between T and E, which determines the sex ratio because both oestrogens and androgens can skew the ratio.

Invertebrates
As described in OECD GD the available invertebrate assays solely involve apical endpoints which are able to respond to some endocrine disrupters and some non-EDs. Those in level 4 are partial lifecycle tests, while those in level 5 are full- or multiple lifecycle tests.

4.3 Alerts from *in vivo* testing

4.3.1 Alert based on adverse effects in experimental studies
Alerts for endocrine disrupting properties from *in vivo* test systems may originate in many different ways, e.g. as the result of adverse effects suspected to be due to endocrine mode of action in repeated dose toxicity test, reproductive toxicity tests or ecotoxicity tests. Whatever the origin of the

signal is, the obvious next step is to test the chemical in *in vitro* systems attempting to select the *in vitro* systems with the modes of action expected to be of relevance for the observed adverse effects.

Depending on the results obtained in the *in vitro* test assays, the subsequent testing may proceed as indicated under 'Alert from *in vitro* assays' and read across information.

Negative results in the *in vitro* assays should be considered carefully as there can be cases where the adverse outcome pathway is not covered by the available *in vitro* assays. Further, as also stated in the OECD GD: "Negative *in vitro* results alone cannot be used to exclude possible endocrine disruption activity because of their inherent limitations, such as inability or unknown capacity to metabolically activate toxicants. In addition, chemicals can interfere with the endocrine system in other ways than through the receptor, such as effects on the hypothalamic-pituitary-gonadal axis (HPG) that can only be detected in whole animal studies. For example, chemicals can interfere with the hormonal feedback loops in the HPG axis which could only be revealed in intact animals *e.g*. by changes in hormone levels. Each *in vitro* assay measures a certain mechanism and thus, conclusions can be drawn only in the context of what the *in vitro* assay evaluates. However, negative *in vitro* effects should only be interpreted as a tentative indication of a lack of endocrine disruption activity for the modality in question, if it can be substantiated that the compound does not undergo metabolic activation *e.g*. by the use of ADME information."

4.3.2 Alert based on in vivo mode of action data

In OECD TG 407, the indicators of hormonal activity are thyroid hormones (T3, T4 and TSH). A positive result for indicators of hormonal activity could be statistically significant changes in thyroid hormone profiles. A positive result for indicators of hormonal activity alone should be considered with caution as it is possible that these endpoints may have detected weak effects that were not detected by the apical endpoints in OECD TG 407 but may then be detected in longer-term studies.

Other *in vivo* indicators for hormonal activity, such as changed levels of oestrogen and androgens, are not assessed in any OECD guideline studies, but such data may be available from non-guideline studies. Also, results from the Uterotrophic and the Hershberger assay (TGs 440 and 441) can provide evidence for *in vivo* (anti-)oestrogenic or (anti-)androgenic MoA, respectively.

The adverse effects related to data showing *in vivo* ED mode of action can be studied using assays from mammalian level 5 of the OECD CF e.g. the extended one-generation reproductive toxicity Study (TG 443) or the developmental neurotoxicity study (TG 426) or from wildlife level 4 or 5 e.g. FSDT (TG 234) or a fish full life cycle toxicity test (e.g. USEPA 850.1500).

4.4 Alert from epidemiological studies

Suspicion of endocrine disrupting properties raised from epidemiological studies should lead to an initial testing of the chemical in relevant *in vitro* test systems and - if indicated by the results -

subsequent testing at the higher OECD conceptual framework levels as described under ‘Suspicion from *in vitro* test systems’. However, negative results in the *in vitro* assays should be considered carefully as a metabolite may be responsible for the effect suspected in humans or the effects may be caused by other mechanisms.

If the suspicion is based on adverse effects that may be due to endocrine disruption *in vivo* testing of the substance for effects of relevance for the effect in humans may also be warranted. Negative *in vivo* results should be viewed with caution as adverse effect observed in humans may for example be due to mixed exposure including other substances.

4.5 Alert from field investigations

If a chemical is suspected of having endocrine disrupting properties based on observations made in field investigations, the obvious choice is to test the properties of the chemical in *in vitro* test systems with subsequent testing at the higher OECD conceptual framework levels if the results from the *in vitro* testing indicate the need for further testing.

It has to be borne in mind, however, that the suggested testing strategies might not have been able to catch all of the endocrine disrupting effects that have been observed in field studies.

Looking at the chemicals that have caused endocrine disruption in organisms in nature, it could be a problem that this suggested testing strategy would not have been able to catch all of these chemicals because only OECD validated test systems were applied and these tests mostly covers EATS MoA. Two examples of this are TBT and DDT/DDE: in nature, TBT has caused effects and subsequent population declines mainly in mollusc species and molluscs are not yet included in the battery of OECD validated test systems. The validation of reprotoxicity tests and full life-cycle tests with molluscs has been initiated and included in the OECD working program. DDT/DDE caused eggshell thinning and subsequent population declines mainly in birds of prey and fish eating birds and because of the marked differences in sensitivities to these effects among different families of birds, these effects would have had a very low chance of being revealed in studies utilizing the traditional avian test species (quails, hens and ducks are almost insensitive to these effects). DDT/DDE would, however, have been identified as a chemical of very high concern due to its persistence and bioaccumulative properties.

5. "Test package" for already registered substances in REACH

Under REACH chemical substances produced/imported in tonnages of 1000 or more per manufacturer/importer per year (High Production Volume Chemicals, HPVC) have been registered by 1 December 2010. Substances covered by the next tonnage level (100-1000 tonnes per year per manufacturer/importer) should be registered by 31 May 2013.

For substances registered at these two higher tonnages-bands there may be extensive human and environmental exposure. It seems therefore relevant to set up a "test package" for these substances. As these substances are registered at the higher-tonnage levels this means that in theory there are adverse effect data from long-term repeated dose toxicity data in mammals (e.g. OECD TG 408) and fish (e.g. OECD TG 210) as well as data from the two-generation reproduction toxicity study in rats (OECD TG 416). In addition, for some substance there may be some available data from published studies. Identification of a substance as an ED requires both mode of action data and adverse effect data, but as mentioned previously for the substances already registered under REACH, data on ED mode of action are most likely not available. Furthermore, adverse effect data of relevance for ED may be quite limited, especially if the testing was performed before the last revision of OECD TG 416 which only addressed a limited number of ED related parameters. In addition, it is noted that TG 408 and TG 210 does not really target adverse effect data of relevance for ED. Finally, the Commission REACH review and the Annual ECHA reports concerning dossier evaluation have revealed that the expected standard information is only available for a minor part of the registered substances (DK EPA 2013, pers. com).

If the registration file or other REACH related processes show alerts for ED, further studies should be selected and conducted based on the OECD GD and the general ED testing strategy described in section 4. Generally, this will have to be a case-by-case evaluation to be performed by experts within the field.

However, taking the limitations in the current testing methods and the available standard information for substances registered at high tonnage levels into account as well as the potential for high exposure from these substances to humans and/or the environment, it seems warranted to request more comprehensive investigations of the endocrine disrupting properties of these substances. It is therefore recommended that these substances are prioritized based on alerts of endocrine effects raised by QSARs, read across, and *in vitro* data and substances with ED alerts are tested by using validated and adopted standard test methods that specifically include ED relevant endpoints, i.e. the fish sexual development test (TG 234) and the extended one-generation reproductive toxicity study (TG 443). A QSAR-screening may inform the prioritisation of substances for testing in the extended one-generation reproductive toxicity study (TG 443) e.g. whether to request this test at the 100-1000 tpa or at the > 1000 tpa level where we find it should be required.

6. Conclusions and recommendations

The existing information/data requirements in REACH, PPPR and BPR are not sufficient to adequately detect substances with endocrine disrupting properties. Proposals for changes in the existing REACH, PPPR and BPR information/data requirements are therefore provided. The proposals include enhancement of standard test methods as well inclusion of new methods and the following is recommended:

- For all substances QSAR studies and testing using *in vitro* assays for interaction with different ED modalities e.g. ER, AR and steroidogenesis interference, should be conducted to elucidate whether there are alert(s) for further testing for ED effects.

Human toxicity

- *Repeated dose 28-days study (OECD TG 407):* The optional endpoints, i.e. weight of uterus and ovaries, oestrous cyclicity, histopathologic changes in mammary glands and pituitary and circulating levels of T3, T4 and TSH are all of relevance for identification of the effects of EDs and were included in the OECD validation. It is therefore recommended that they are mandatory.
- *Repeated dose 90-days study (OECD TG 408):* It is recommended to include as mandatory similar endpoints as in OECD TG 407, incl. the current optional ED related endpoints.
- *Screening for reproductive/developmental toxicity (TG 421) or combined repeated dose/reproductive toxicity screening test (TG 422):* Measurements of AGD and assessment of nipple retention is recommended on postnatal day 13/14. Thus, an extension of the testing period from postnatal day 3 to 13/14, i.e. 10-11 day is recommended. Also, it is recommended to include assessment of thyroid hormone levels in the dams during pregnancy and in the offspring at the termination of the study.
- *Prenatal developmental toxicity study (OECD TG 414):* Inclusion of measurements of AGD is recommended.
- The use of the *extended one-generation reproductive toxicity study* (OECD TG 443) instead of TG 416 study would significantly enhance the ability for detection of endocrine disrupters. It is therefore strongly recommended to replace the existing two-generation reproduction toxicity study (TG 416 from 2001) with the extended one-generation reproductive toxicity study (OECD TG 443 from 2011).
- *The extended one-generation reproductive toxicity study* (OECD TG 443) shall be proposed by the registrant if there are indications of potential endocrine toxicity from a repeated dose toxicity study, TG 421/422, or the substance has a close structural relationship with a known endocrine toxicant or there are clear SAR alerts or positive valid predictions from (Q)SAR models or *in vitro* testing.

Environmental toxicity

- Alert of oestrogenic agonism, androgenic agonism or steroidogenesis inhibition should trigger a FSDT (TG 234). Both TG 229 and TG 230 could inform about *in vivo* MoA but not about adversity. Androgenic antagonism could be confirmed as *in vivo* MoA in the (AFSS (OECD GD 148)). A FLCTT or MMGT would be needed to check for adversity. TG 229 has been validated for oestrogenic antagonistic effects but a full lifecycle test (FLCTT or MMGT) should investigate the connection between MoA and adversity. Thyroidal effects

could be tested *in vivo* using the AMA (TG 231) and adversity confirmed by a FLCTT with endocrine relevant endpoints included or a MMGT or a LAGDA. It is recommended that this proposed strategy is followed for REACH, PPPR and BPR.

A mode of action based outline for a testing strategy when there is a concern for endocrine disrupting properties of a chemical substance has also been developed. It is recommended that the strategy should depend on the basis for the concern – whether it is based on indications of endocrine mode of action or *in vivo* alerts for potential endocrine disrupting effects. This strategy may be used by the industry for providing more comprehensive information with regard to the endocrine disrupting properties of a substance or by the authorities e.g. for the requirement of additional information/testing from industry.

Higher tonnage substances have already been or will soon be registered under REACH and as these substances may lead to extensive human and/or environmental exposure, it is recommended to require information on these substances in order to elucidate whether they have endocrine disrupting properties. A “test package” that includes OECD TG 234 and/or TG 443 is suggested. The prioritization should be based on alerts of endocrine effects raised by QSARs, read across, and *in vitro* data as well as all any other available information.

7. References

Ankley GT, Gray LE. Cross-species conservation of endocrine pathways: A critical analysis of tier 1 fish and rat screening assays with 12 model chemicals. Environ Toxicol Chem. 2013 Feb 7. doi: 10.1002/etc.2151.

Borch J, Axelstad M, Vinggaard AM, Dalgaard M. Diisobutyl phthalate has comparable antiandrogenic effects to di-n-butyl phthalate in fetal rat testis. Toxicol Lett. 2006 Jun 1;163(3):183-90.

Cao, J., Y. Lin, L.H. Guo, A.Q. Zhang, Y. Wei, and Y. Yang. 2010. Structure-based investigation on the binding interaction of hydroxylated polybrominated diphenyl ethers with thyroxine transport proteins. Toxicology 277(1-3):20-8.

Danish EPA, 2011: Establishment of Criteria for Endocrine Disruptors and Options for Regulation, 17 May 2011. (Criteria proposal and Annex A, B and C).

Dean A, Smith LB, Macpherson S, Sharpe RM. The effect of dihydrotestosterone exposure during or prior to the masculinization programming window on reproductive development in male and female rats. Int J Androl. 2012 Jun;35(3):330-9.

Hass U, Scholze M, Christiansen S, Dalgaard M, Vinggaard AM, Axelstad M, Metzdorff SB, Kortenkamp A 2007 Combined exposure to anti-androgens exacerbates disruption of sexual differentiation in the rat. Environmental Health Perspectives 115 Suppl 1:122-128

Hass U, Christiansen S, Axelstad M, Sørensen KD and Boberg J. Input for the REACH-review in 2013 on endocrine disrupters, Final report 21 March 2013, p. 1-50

Kortenkamp A, Martin O, Faust M, Evans R, McKinlay R, Orton F, Rosivatz E. (2012). State of the Art Assessment of Endocrine Disrupters. Final Report. [Online] Available at: http://ec.europa.eu/environment/endocrine/documents/4_SOTA%20EDC%20Final%20Report%20V3%206%20Feb%2012.pdf

OECD test guidelines (human health): http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788

OECD 2004 OECD HPV (High Production Volume) programme (OECD 2004)

OECD. (2006). Report of the Validation of the Updated Test Guideline 407: Repeat Dose 28-day Oral Toxicity Study in Laboratory Rats. Series on Testing and Assessment No 59, ENV/JM/MONO(2006)26.

OECD (2007a). Guidance Document on the Uterotrophic Bioassay - Procedure to Test for Antioestrogenicity. OECD Series on Testing and Assessment. No. 71. ENV/JM/MONO(2007)15.

OECD (2009a). Guidance document for histologic evaluation of endocrine and reproductive tests in rodents. OECD Series on Testing and Assessment no. 106. Organisation for Economic Cooperation and Development, Paris. 26 pp.

OECD (2010a). Guidance document on the diagnosis of endocrine-related histopathology in fish gonads. OECD Series on Testing and Assessment no. 123. 114 pp.

OECD (2010d). Draft detailed review paper on environmental endocrine disruptor screening: the use of estrogen and androgen receptor binding and transactivation assays in fish. Series on Testing and Assessment no. 135. ENV/JM/MONO(2010)34.

OECD (2012). Guidance Document on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption. Series on Testing and Assessment No. 150, ENV/JM/MONO(2012)22

OECD (2012b). Detailed Review Paper on the State of the Science on Novel In Vitro and In Vivo Screening and Testing Methods and Endpoints for Evaluating Endocrine Disruptors. Series on Testing and Assessment No 178. ENV/JM/MONO(2012)23.

OECD (2012c) Fish Toxicity Testing Framework. Series on Testing and Assessment No. 171. ENV/JM/MONO(2012)16

Saillenfait AM, Sabate JP, Gallissot F. Effects of in utero exposure to di-n-hexyl phthalate on the reproductive development of the male rat. Reproductive Toxicology 2009;28:468–76.

Saillenfait AM, Roudot AC, Gallissot F, Sabaté JP, Chagnon MC. Developmental toxic potential of di-n-propyl phthalate administered orally to rats. J Appl Toxicol. 2011a Jan;31(1):36-44. doi: 10.1002/jat.1565.

Saillenfait AM, Roudot AC, Gallisot F, Sabaté JP. Prenatal developmental toxicity studies on di-n-heptyl and di-n-octyl phthalates in Sprague-Dawley rats. Reproductive Toxicology 2011b;32:268–76.

Saillenfait AM, Gallissot F, Sabaté JP, Remy A. Prenatal developmental toxicity studies on diundecyl and ditridecyl phthalates in Sprague-Dawley rats. Reprod Toxicol. 2013 Jun;37:49-55. doi: 10.1016/j.reprotox.2013.01.004. Epub 2013 Jan 31.

Schmutzler, C., A. Bacinski, I. Gotthardt, K. Huhne, P. Ambrugger, H. Klammer, C. Schlecht, C. Hoang-Vu, A. Gruters, W. Wuttke, H. Jarry, and J. Kohrle. 2007. The ultraviolet filter benzophenone 2 interferes with the thyroid hormone axis in rats and is a potent in vitro inhibitor of human recombinant thyroid peroxidase. Endocrinology 148(6): 2835-44.

WHO 2002. Global assessment of the state-of-the-science of endocrine disruptors. WHO/PCS/EDC/02.2. Geneva: World Health Organisation,

Zoeller, R.T. 2007. Environmental chemicals impacting the thyroid: targets and consequences. Thyroid 17, 811-7.

Zoeller TR, Crofton KM (2000). Thyroid hormone action in fetal brain development and potential for disruption by environmental chemicals. Neurotoxicology 21(6):935-946.

Annex 1: ED relevant environmental tests

OECD TG 229: Fish Short-Term Reproduction Assay.
This test is performed on zebrafish (*Danio rerio*), fathead minnow (*Pimephales promelas*) or Japanese medaka (*Oryzias latipes*). The test is designed to detect chemicals that affect fish reproduction via the hypothalamic pituitary gonadal axis. The test exposes groups of fish that have shown to be successful spawner´s during the 1-2 week acclimation period. The exposure duration is 21 days. Besides apical endpoints, the test include histological changes in the gonads and endpoints indicative of ED effects such as change in secondary sex characteristics (SSC), vitellogenin concentration and gonadal histopathology. The number of eggs produced, are registered daily during the pre-exposure and exposure to measure reproductive effects.
The test has, in the validation processes, been able to identify oestrogen receptor agonists (vitellogenin induction in males, depression of male SSC in fathead minnow or Japanese medaka, specific gonad histopathologic findings (OECD 2010a; OECD 2012)), oestrogen receptor antagonists (vitellogenin depression in females (assuming no systemic toxicity) and specified gonad histopathologic findings (OECD 2010a)), androgen agonists (induction of male SSC in female fathead minnow or Japanese medaka and specified gonad histopathologic findings (OECD 2010a), androgen antagonists (depression of male SSC in fathead minnow or Japanese medaka, specified gonad histopathologic findings (OECD 2010a)) inhibitors of steroid synthesis like for example aromatase inhibitors (vitellogenin depression in females (assuming no systemic toxicity) and specified gonad histopathologic findings (OECD 2010a)). Beside the above mentioned endpoints, fecundity is potentially sensitive to ED effects but not diagnostic. It should, however be noticed that due to high variability in the fecundity, the power of this endpoint can be low compared to standard power advices (OECD 2012c). The results of the TG 229 can be divided into apical endpoints as survival and fecundity and more specific endocrine endpoints as changes in vitellogenin concentrations and secondary sex characteristics. The test is placed at level 3 in OECD CF as a screening test because it primarily gives information about ED mechanisms. It is not generally possible to link the ED MoA observed with adverse effects (changes in fecundity can be adverse but it is not an endocrine specific endpoint).
A positive result on the ED specific endpoints vitellogenin and/or secondary sex characteristics will classify a substance as suspected ED in accordance with the Danish criteria proposal.
A negative result on the ED specific endpoints should be compared with existing data, both *in vitro* and *in vivo*. If these data are also negative, the substance is probably not affecting the oestrogenic and androgenic pathways in adult fish. To test for effects during sexual development a Fish Sexual Development Test (OECD TG 234) should be considered (test description below) and to test for thyroidal effects, the amphibian metamorphosis assay (TG 231) should be considered (test description below).

OECD TG 230: 21-Day Fish Assay.
TG 230 is comparable to TG 229 using same exposure scenario and same fish species. The exceptions are that fecundity and histopathology are not included. The ED specific endpoints are vitellogenin concentrations and secondary sex characteristics (SSC). During validation studies, the test has been able to identify oestrogen receptor agonists (vitellogenin induction in males), androgen

agonists (induction of male SSC in female fathead minnow or Japanese medaka) and aromatase inhibitors and oestrogen antagonists (depression in female vitellogenin concentration). The core endocrine endpoints are indicators of hormonal activity and there are no apical measures of adverse effects that can be attributed to a single EATS modality (OECD GD 171 (OECD 2012c)). TG 230 is placed on OECD CF level 3 as a screening assay.
A positive result on the ED specific endpoints vitellogenin and/or secondary sex characteristics will classify a substance as suspected ED in accordance with the Danish criteria proposal.
A negative result on the ED specific endpoints should be compared with existing data, both in vitro and in vivo. If these data are also negative, the substance is probably not affecting the oestrogenic and androgenic pathways in adult fish. The test has difficulties in responding to androgen receptor antagonists. In case of suspicion of this MoA, the Androgenised Female Stickleback Screen (AFSS) could be an alternative (AFSS Guidance Document No. 148 in the Series on Testing and Assessment, ENV/JM/MONO(2011)29). To test for effects during sexual development a Fish Sexual Development Test (OECD TG 234) should be considered (test description below) and to test for thyroidal effects, the amphibian metamorphosis assay (TG 231) should be considered (test description below).

OECD TG 231: Amphibian Metamorphosis Assay (AMA)
TG 231 is a 21 day screening test for thyroid activity in amphibians using *Xenopus laevis*. It should be noted that there are several types of thyroid disruption, where not all interact with the thyroid receptor. As described in OECD GD 150, TG 231 does not, however, allow unequivocal diagnosis of which type of thyroid disruption is occurring. It includes a specific endpoint (thyroid gland histopathology) for some types of thyroid activity, but also includes apical measurements (hind limb length, snout-vent length, developmental stage and wet weight), which are used to determine other thyroid-responsive endpoints: advanced development, asynchronous development or delayed development. The first two of these are considered by some authorities to be diagnostic of thyroid activity, while the latter is only diagnostic if non-specific systemic toxicity is absent (OECD 2012). The evaluation of the positive and negative responses of the endpoints included can be complicated (for details see OECD GD). In relation to read across between amphibians and mammals, by extrapolation of thyroid effects between mammalian and amphibian screening models very good correlations have been observed (26 of 27 substances active in TG 231 were also active in mammalian screens and all substances active in mammalian screens turned out active in TG 231 (OECD 2012)).

OECD TG 234: Fish Sexual Development Test (FSDT)
TG 234 includes Japanese medaka, the three-spined stickleback and zebrafish. The test is in principle an enhancement of TG 210: Fish, Early Life Stage Toxicity Test, where the exposure is continued until the fish are sexually differentiated, i.e. about 60 days post-hatch (dph) for Japanese medaka, the three-spined stickleback and zebrafish (the exposure period can be shorter or longer for other species that are validated in the future), and endocrine-sensitive endpoints are added. The FSDT assesses early life-stage effects and potential adverse consequences of putative endocrine disrupting chemicals (e.g. oestrogens, androgens and steroidogenesis inhibitors) on sexual development. Two core endpoints are measured as indicators of endocrine-associated

developmental aberrations, the vitellogenin concentrations and sex ratios (proportions of sex) determined via gonad histology. Gonadal histopathology (evaluation and staging of oocytes and spermatogenetic cells) is optional. Additionally, the genetic sex is determined whenever possible (e.g. in Japanese medaka and the three spined stickleback). The presence of a genetic sex marker is a considerable advantage as it increases the power of the sex ratio statistics and enables the detection of individual phenotypic sex reversal. The ED MoA covered to date are oestrogens, androgens and steroidogenesis inhibitors. Other apical endpoints that should be measured include hatching rate, survival, length and body weight. Positive response in the combination of the two core endocrine endpoints, vitellogenin concentration and phenotypic sex ratio enable the test to indicate the mode of action of the test chemical. Due to the population-relevant change in phenotypic sex ratio, the FSDT can be also used for hazard and risk assessment. However, if the test is used for hazard or risk assessment, the stickleback should not be used because the validation data available so far showed that in this species the alterations of phenotypic sex ratio by the test substances were uncommon (OECD TG 234).
A negative response in the core endocrine endpoints indicates that the tested substance does not interact with the sexual development of fish. It should be noted that TG 234 has not been validated for androgen antagonists and thyroid effects and that reproduction is not covered. Therefore a FLCTT with endocrine relevant endpoints included or a MMGT should be considered.

OECD TG 206: Avian Reproduction Test

This TG does not contain endpoints which solely respond to endocrine disrupters, and it has not been specifically validated with EDs. However, some of the endpoints in this apical test are nevertheless potentially affected by EATS EDs. Of particular interest in the context of oestrogens, androgens and steroidogenesis disrupters are egg production, embryo viability, and hatchability, but other endpoints may also be responsive to some EDs (e.g. growth may respond to some thyroid disrupters; % cracked eggs and egg shell thickness may respond to chemicals interfering with the control of shell deposition) (OECD 2012). The test exposes birds for 20 weeks via food, where after eggs are collected for 10 weeks, hatched and chickens are maintained for 14 days. Species include Bobwhite quail, Japanese quail and mallard duck (OECD TG 206). A positive result in OECD TG 206 indicate that a chemical may be an ED if they can be plausibly linked to an endocrine MoA established on the basis of prior screening. However, more conclusive data in this regard would be obtainable from an ATGT. If screening data are unavailable or negative, it should not be concluded that a positive OECD TG 206 is the result of endocrine disruption. On the other hand, a negative OECD TG 206 combined with negative screening data should lead to a conclusion that a chemical is probably not an ED in birds. A negative OECD TG 206 set against a background of a positive screen might, however, raise concerns if the chemical is strongly bioaccumulative, known to be involved in epigenesis, or suspected of having effects on sexual development, when an ATGT should be considered (OECD 2012).

OECD GD 148: Androgenised Female Stickleback Screen (AFSS)

The androgenised female stickleback screen (AFSS) is designed to identify chemicals that interact with the androgen receptor, especially as antagonists. Identification of test chemicals as androgen receptor antagonists is based on their ability to block the biological activity (induction of spiggin)

of a model androgen receptor agonist dihydrotestosterone (DHT) in female adult three spined sticklebacks. The test includes seven treatment groups: a water-only control, solvent control, negative control (test chemical at a "high" concentration), positive control (5 μg DHT/L), high test chemical concentration plus DHT, medium test chemical concentration plus DHT, and low test chemical concentration plus DHT. Chemical exposures are initiated after a 1 week acclimation period, and the test is terminated after a 21-d chemical exposure (OECD 2012c). The AFSS will detect chemicals that inhibit DHT-induced spiggin production in female sticklebacks. Several published studies have shown that one important class of endocrine-active chemicals that will do this is androgen receptor antagonists. In this regard the AFSS fills a niche not covered by TGs 229 and 230, neither of which have mechanistic endpoints that serve to specifically identify androgen receptor antagonists (OECD 2012c)

Fish Lifecycle Toxicity Test (FLCTT) (USEPA OPPTS 850.1500))

The existing FLCTT as described by Benoit (1981), USEPA (1996) and others does not contain endpoints which solely respond to endocrine disrupters. However, many of the endpoints in this apical test are nevertheless affected by EATS EDs. Of particular interest in the context of oestrogens, androgens and steroidogenesis disrupters are time to sexual maturity, sex ratio of adults, fecundity and fertility, but other endpoints may also be responsive to some EDs (e.g. growth may respond to some thyroid disrupters). The sex ratio of adults could though be regarded as an ED specific adverse endpoint if no systemic toxicity is observed at the affected test concentration (s). This assay is designed primarily as an apical test for chemicals with suspected reproductive or long-term toxicity. It has not been adopted for publication as a OECD TG, but has been widely used for several decades by regulatory agencies for assessing possible chronic effects in fish. The endpoints are all apical measures of development, growth or reproduction. Exposure of the test organisms (fathead minnow Pimephales promelas, in the case of Benoit 1981, but other species can be successfully used with minor changes in the protocol, including sheepshead minnow (Cyprinodon variegatus), zebrafish (Danio rerio), and medaka (Oryzias latipes) usually continues from the freshly fertilised eggs of the F0 generation to the fry or young fish of the F1 generation (4-8 weeks post-hatch in the case of fathead minnow – Benoit, 1981). It should be noted that it would be relatively straightforward to include ED-specific endpoints in this test. Depending on the species and test objectives, these could include inter alia sex hormones, thyroid hormones, vitellogenin, spiggin, secondary sex characteristics, gonadal histopathology, and genetic sex. It would be desirable to include such ED-specific endpoints before using the FLCTT to investigate a possible ED. (the above text is from OECD GD 150 (OECD 2012)).

OECD draft TG: Larval Amphibian Growth and Development Assay (LAGDA)

This assay is a partial lifecycle test with the clawed frog *Xenopus laevis*. It covers the stages of larval/juvenile growth and sexual development, but not those of reproduction and embryonic development. It could therefore be thought of as the amphibian near-equivalent of the Fish Sexual Development Test (FSDT) (TG 234), although it also includes endpoints that are specifically responsive to thyroid disrupters. It does not include all processes which may respond to EATS EDs (especially reproduction), and it is currently unknown whether the LAGDA is therefore less responsive to some of these chemicals than a lifecycle test (OECD 2012)

OECD draft test guideline: Medaka Multi-Generation Test (MMGT)

The MMGT is a 24 weeks test with Japanese medaka running from F0 reproducing adults to F2 pre-reproducing adults, and hence encompasses two complete generations. The test includes several ED specific endpoints as well as ED specific apical endpoints: (i.e. vitellogenin, SSC, gonad histopathology, fecundity, fertility and gonad sex reversal). It is expected to be responsive to most chemicals with EATS modalities (OECD 2012). As the test includes all developmental and reproductive phases as well as maternal transfer it is expected that a positive result on the ED endpoints could be used for both hazard- and risk assessment and a fully negative result can give information about a tested substance not to affect adversely the endocrine system in fish.

OECD draft test guideline: Avian Two-Generation Test (ATGT)

An avian two-generation test (ATGT) is under development including ED relevant endpoints. ATGT is a life cycle assay with the Japanese quail *Coturnix japonica* that runs for 21 weeks, from 4 week old F0 reproducing adults to 2 week old F2 chicks, and hence encompasses more than one complete generation. It is therefore expected to be responsive to most chemicals with ED modalities, although the full extent of its applicability awaits further validation. It should be noted that if the assay gives a positive result, this may be due to a positive indicator of hormonal activity (i.e. endocrine organ weights; endocrine organ gross pathology and histopathology; feather dimorphism; time to first egg lay; sexual behaviour; sex hormones; tibiotarsus length), a positive for apical endpoints (egg production; general health and toxic signs; sex ratio; body weight; shell thickness/cracking/strength; fertility; embryo viability; hatchability), or a positive for both types of endpoint (OECD 2012). A positive apical response in the ATGT indicates that a substance is a developmental or reproductive toxicant which may or may not be an ED. A combination of a positive apical response and a positive endocrine-specific endpoint (e.g. feather dimorphism) is strong evidence that the chemical is an actual ED, especially if the two types of endpoint are causally related and if positive mechanistic data are also available (OECD 2012).