

US EPA ARCHIVE DOCUMENT

# **TECHNICAL REVIEW DOCUMENT**

**for**

## **ENDOCRINE DISRUPTOR SCREENING PROGRAM (EDSP):**

### **PROPOSED TIER 1 SCREENING BATTERY**



**U.S. Environmental Protection Agency  
Office of Prevention, Pesticides, and Toxic Substances  
Office of Science Coordination and Policy**

**March 7, 2008**

## Table of Contents

1.0	Introduction .....	4
2.0	EDSP Background .....	4
3.0	EDSP Assay Validation Processes and Principles.....	7
3.1	Assay Validation Process .....	8
3.2	Assay Validation Principles .....	10
4.0	Selection of the Proposed Tier 1 Screening Battery .....	11
4.1	Basis for Including Assays in the EDSP Proposed Tier 1 Screening Battery ..	12
4.1.1	Assays for detection of compounds that affect the estrogen signaling pathway.....	13
4.1.2	Assays for detection of compounds that affect the androgen signaling pathway.....	16
4.1.3	Assays for detection of compounds that affect steroid synthesis .....	18
4.1.4	Assays for detection of chemicals that affect the HPG axis .....	20
4.1.5	Assays for detection of chemicals that affect the HPT axis .....	21
4.2	Basis for Not Including Potential Screening Assays from the Proposed EDSP Tier 1 battery.....	23
5.0	Performance Review of the EDSP Tier 1 Screening Battery.....	26
6.0	Summary.....	26
7.0	Charge Questions for the FIFRA SAP .....	31
8.0	References.....	32
Appendix A – Fact Sheets for Assays Included in the Proposed EDSP Tier 1 Screening Battery.....		35
Appendix A1 – Amphibian Metamorphosis .....		36
Appendix A2 – Androgen Receptor Binding .....		39
Appendix A3 – Aromatase (Recombinant).....		41
Appendix A4 – Estrogen Receptor Binding.....		43
Appendix A5 – Estrogen Receptor Stably Transfected Transcriptional Activation.....		45
Appendix A6 – Fish Short-term Reproduction (Fish Assay).....		46
Appendix A7 – Hershberger .....		49
Appendix A8 – Pubertal Female .....		51
Appendix A9 – Pubertal Male .....		53
Appendix A10 – Steroidogenesis (H295R) .....		55
Appendix A11 – Uterotrophic.....		57
Appendix B - General Principles for Evaluating the Results from a Tier 1 Screening Battery According to EDSTAC .....		59
Appendix C: Fact Sheets for Assays Not Included in the Proposed EDSP Tier 1 Screening Battery.....		61
Appendix C1 – Adult Male .....		62
Appendix C2 – Aromatase (Placental).....		65
Appendix C2 – <i>In Utero</i> Through Lactational .....		67
Appendix C4 – Steroidogenesis (Sliced testes).....		69
Appendix D - A Comparison of the Screening Assays Recommended by EDSTAC and those Proposed by EPA for the EDSP Tier 1 Battery.....		71

## Table of Lists

Table 1: Screening Assays Recommended by EDSTAC for the EDSP Tier 1 Battery....	7
Table 2: Screening Assays Proposed by EPA for the EDSP Tier 1 Battery .....	12
Table 3: Modes of Action Covered by Screening Assays Proposed in the Tier 1 Battery .....	13
Table 4: Profiles Diagnostic for Various Modes of Action (MOA) in the Proposed Tier 1 Screening Battery.....	21
Table 5: Summary of the Strengths in Detecting EAT effects in the Proposed Tier 1 Screening Battery According to Mode of Action (MOA).....	27

# TECHNICAL REVIEW DOCUMENT for ENDOCRINE DISRUPTOR SCREENING PROGRAM (EDSP): PROPOSED TIER 1 SCREENING BATTERY

## 1.0 Introduction

The purpose of this Technical Review Document is to serve as a basic guide and source of information for members of the Federal Insecticide, Fungicide, Rodenticide Act (FIFRA) Scientific Advisory Panel (SAP) who will be charged to review the proposed Endocrine Disruptor Screening Program (EDSP) Tier 1 Screening (TIS) battery. Additional information in the form of appendices within this document as well as other pertinent documents (e.g., Integrated Summary Reports of the validation process and assay Peer Review Records) will be provided to further support and facilitate SAP review and responses to the charges posed by EPA.

In general, this document includes the following information:

- 1) overview of the approach employed by the EDSP for screening and testing substances for effects on the estrogen, androgen and thyroid (EAT) hormonal systems, including the initial recommendations from the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) for the Tier 1 screening battery,
- 2) overview of assay validation procedures, and
- 3) overview of the proposed Tier 1 screening battery, including scientific criteria for assay selection.

## 2.0 EDSP Background

Passage of the Food Quality Protection Act (FQPA) in 1996 and subsequent amendments to the Safe Drinking Water Act (SDWA) and Federal Food, Drug, and Cosmetic Act (FFDCA) required EPA to:

*develop a screening program, using appropriate validated test systems and other scientifically relevant information, to determine whether certain substances may have an effect in humans that is similar to an effect produced by a naturally occurring estrogen, or other such endocrine effect as the Administrator may designate [21 U.S.C. 346a(p)].*

In response to this mandate, the Agency established a multi-stakeholder federal advisory committee, EDSTAC. This committee was asked to provide advice to the Agency on how to design a screening and testing program for endocrine disrupting chemicals. In 1998, the EDSTAC published their final report (EDSTAC, 1998), which

included three overarching recommendations which shaped the Agency's current EDSP:

- 1) Expand the evaluation of additional modes of action beyond estrogenic activity to include test systems that detect androgen, thyroid, and hypothalamic-pituitary-gonadal (HPG) disruption.
- 2) Expand the number of organisms to include wildlife (*i.e.*, amphibian, fish, reptiles, birds, invertebrates), in addition to humans.
- 3) Incorporate a two-tiered approach, whereby Tier 1 would consist of a suite of complementary and less complex assays designed to effectively and efficiently screen substances for interactions along the EAT hormonal pathways. If results from Tier 1 indicate that a substance does exhibit the potential to interact with the E, A or T pathways, then more complex and definitive dose-response testing would likely be done in Tier 2.

Tier 1 screening, according to EDSTAC, should:

- Maximize sensitivity to minimize false negatives while permitting an acceptable level of false positives.
- Include a range of organisms representing known or anticipated differences in metabolic activity. The Tier 1 battery should include assays from representative vertebrate classes to reduce the likelihood that important pathways for metabolic activation or detoxification are not overlooked.
- Be designed to detect all known modes of action for the endocrine endpoints of concern. All chemicals known to affect the action of EAT hormonal axes should be detected.
- Include a sufficient range of taxonomic groups to account for known differences in endocrine systems.
- Incorporate sufficient diversity among the endpoints and assays to reach conclusions based on "weight-of-evidence" considerations.

Tier 2 testing, according to EDSTAC, is to provide a more definitive approach that will include a broad range of taxa exposed through various routes and during sensitive life-stages, so that the adverse consequences related to EAT hormonal function can be characterized with greater specificity. This will be done in the larger context of testing for developmental and reproductive toxicity potential by any mechanism (including EAT) using study designs that provide a comprehensive assessment of relevant functions.

Tier 2 tests should:

- identify hazard and
- establish quantitative relationships between dose and adverse effects.

EPA considered the recommendations from EDSTAC and, in accordance with the Administrator's discretionary authority, the Agency adopted the two-tiered testing strategy and expanded the EDSP to include the androgen and thyroid hormonal systems, as well as wildlife. EPA's proposed EDSP is described in detail in a 1998 Federal Register Notice (EPA, 1998). The proposed EDSP was reviewed in 1999 by a joint committee of the Agency's Science Advisory Board (SAB) and the FIFRA SAP. A final report was submitted to EPA in 1999 (SAB/SAP, 1999).

As recommended by EDSTAC, EPA has implemented its EDSP in three major parts. This document deals only with selection of the assays to be included in the Tier 1 battery, which is considered part of the assay validation process. The three parts of the EPA's EDSP are briefly summarized as follows:

- 1) **Priority setting.** EPA is prioritizing chemicals to undergo screening in the battery of Tier 1 assays. EPA described its priority setting approach for the first 50-100 chemicals to be tested in the Federal Register of September 27, 2005 (70 FR 56449), and published a draft initial list of 73 chemicals to undergo Tier 1 screening in the Federal Register for public review on June 18, 2007 (72 FR 33486). The Agency expects to finalize this initial list of chemicals before screening is initiated in 2008. The battery of Tier 1 assays will be used to screen these chemicals.
- 2) **Procedures.** EPA intends to commence Tier 1 screening of the first group of pesticide chemicals by issuing test orders under FFDCA section 408(p) to chemical companies identified as the manufacturer or processor of the identified chemicals, including the pesticide registrant. EPA has drafted implementation policies that describe the procedures that EPA will use to issue orders, the procedures that the recipients will use to respond, and the procedures for data protection and compensation. These and other related procedures or policies were published in the Federal Register of December 13, 2007 (72 FR 70842). In addition, EPA developed a draft template for the test order and a draft information collection request (ICR) to obtain the necessary clearances under the Paperwork Reduction Act (PRA). The Agency expects to finalize the policies and ICR before screening is initiated in 2008.
- 3) **Assay validation.** Considering that none of the Tier 1 screening assays proposed by EDSTAC were "validated test systems" in 1998, the EPA embarked on an ambitious effort to develop, standardize and validate many of the initially proposed assays, as well as more novel assays that have emerged since the EDSTAC report was published. The assays recommended by EDSTAC are

listed in Table 1. Although this process was largely coordinated by and funded through the Agency's EDSP, other offices within the Agency, especially the Office of Research and Development (ORD), as well as domestic industry and academic institutions, provided necessary research and technical support. In addition, methods development, standardization, and validation work was also conducted internationally. For example, the uterotrophic and Hershberger assays were vetted through various forums involving member countries in the Organization for Economic Cooperation and Development (OECD).

To provide independent technical expertise and recommendations to EPA throughout the assay validation processes, the EDSP sought guidance from other federal advisory committees such as the Endocrine Disrupter Methods Validation Sub-committee and the Endocrine Disrupter Methods Validation Advisory Committee from 2001 through 2006. Afterwards, the FIFRA SAP filled the advisory role.

Table 1 below provides the battery of Tier 1 assays recommended by EDSTAC.

**Table 1: Screening Assays Recommended by EDSTAC for the EDSP Tier 1 Battery.**

<b>Battery Recommended by EDSTAC</b>
<i><b>In vitro</b></i>
Estrogen receptor (ER) binding – rat uterus
Estrogen receptor transcriptional activation
Androgen receptor (AR) binding – rat prostate
Androgen receptor (AR) transcriptional activation
Steroidogenesis – minced rat testes
<i><b>In vivo</b></i>
Uterotrophic (rat)
Hershberger (rat)
Pubertal female (rat)
Amphibian metamorphosis (frog)
Fish gonadal recrudescence
<b>Alternative Assays Recommended by EDSTAC</b>
Aromatase – Human placental
Pubertal male (rat)
Adult male (rat)
<i>In utero</i> through lactational (rat)

### 3.0 EDSP Assay Validation Processes and Principles

Most of the assays considered for the proposed Tier 1 battery have been through a validation process in accordance with Section 408(p) of the FFDCA which requires EPA to use validated test systems. The two exceptions are the ER binding assay and the H295R assay for steroidogenesis, which replaced the minced testes. These assays will complete the final step in the validation process (Peer Review) prior to the initiation of testing. Moreover, the Interagency Coordinating Committee for the Validation of



Alternative Methods (ICCVAM) Authorization Act of 2000 requires all federal agencies to ensure that new and revised test methods are valid prior to their use (ICCVAM, 2000).

### 3.1 Assay Validation Process

In general, EPA has followed a five-stage validation process:

**First Stage - *Test Development*** was an applied research function which culminated in an initial protocol. As part of this stage, EPA prepared a Detailed Review Paper (DRP) or an analogous document (e.g., Background Review Document) to explain the purpose of the assay, the context in which it will be used, and the scientific basis upon which the protocol endpoints and relevance are based. The DRP reviewed the scientific literature for candidate protocols and evaluated them with respect to a number of considerations, such as whether the candidate protocols meet the intended purpose of the assay, the costs and other practical considerations. The DRP also addressed any deficiencies in the assay and, if possible, described an initial protocol for the initiation of the second stage of validation.

**Second Stage - *Prevalidation*** was the stage during which the protocol was refined, optimized, standardized and initially assessed for transferability and performance. Depending upon the completeness of the assay, additional studies were conducted during this second stage. The initial assessment of transferability was generally a trial in a second laboratory to determine whether or not another laboratory could follow the protocol and execute the study.

**Third Stage - *Inter-laboratory Validation*** studies were conducted in independent laboratories using the optimized, standardized protocol. The results of these multi-laboratory studies were used to determine inter-laboratory variability and to develop performance criteria.

**Fourth Stage - *Peer Review*** included, an independent scientific review by qualified experts. Peer review of individual screening assays was conducted in compliance with EPA's Peer Review Handbook (EPA, 2006). EPA uses peer review as an important component of the scientific process, as it involves the critical evaluation of scientific and technical work products by independent experts for the purpose of improving the quality, credibility and acceptability of regulatory decisions. For most assays, an external letter review was organized under an EPA peer review contract; three assays (i.e., uterotrophic, Hershberger and ER stably transfected transcriptional activation assays), however, went through OECD's validation process and peer review. The EPA contractor compiled a list of qualified peer review candidates who were independent of those who performed the work or who have been involved in the development or refinement of the screening assays, including those who have provided EPA with expert advice throughout the validation process.

From the pool of candidate reviewers, the contractor established a balanced peer review panel consisting of five peer reviewers for each assay. Although EPA was notified of the identity of the peer reviewers, the Agency did not have contact with them before or during the peer review process, aside from participating in a contractor-mediated teleconference to kick off the review and answer logistical questions. The contractor provided each reviewer with a peer review package from the EPA that included the assay Integrated Summary Report (ISR), other supporting review material specific to the assay under review, and a list of charges to the reviewers common to all assay reviews.

Each ISR served as the main document during peer review, providing an overview of the development, prevalidation and inter-laboratory testing of individual assays proposed in the EDSP Tier 1 screening battery.

In general, the ISR for each assay included:

- 1) a historical overview of the assay,
- 2) key prevalidation steps and results used to establish the *relevance* of the assay,
- 3) a standardized assay protocol and
- 4) results of inter-laboratory validation studies that were conducted to demonstrate the *reliability* of the assay.

ISRs for assays considered in the proposed battery and peer reviewed by an EPA contractor are included in the SAP review package. A peer review record for each of these assays was produced by the contractor and submitted to EPA. Each peer review record included the names, affiliations and qualifications of the peer review panel members, responses to the charges and any additional comments, information and materials received from each reviewer. Additionally, EPA has provided written responses to comments made by the reviewers which have been summarized and included in the SAP review package along with the peer review record for each assay.

Peer review summary reports are also included in the SAP review package for the Uterotrophic, Hershberger, and ER stably transfected transcriptional activation assays, which were peer reviewed according to the OECD process.

The EPA peer review record and OECD peer review summary report for each assay were also opportunities for the EDSP to clarify the strengths and limitations of each assay, which are summarized in respective Assay Fact Sheets in Appendix A. The Assay Fact Sheets are intended to serve as a quick reference for the SAP and, therefore, also include the purpose, design, endpoints, data interpretation and summary of key peer review comments.

**Fifth Stage - Regulatory Acceptance** is an adoption of an assay for regulatory use by an agency. The inclusion of an assay in the proposed Tier 1 battery and final adoption of the battery by EPA constitutes the regulatory acceptance stage of the process.

### 3.2 Assay Validation Principles

In general, “methods validation” has been defined as “the process by which the relevance and reliability of a test method is evaluated for a particular use” (OECD, 1996; NIEHS, 1997).

*Relevance* describes whether a test is meaningful and useful for a particular purpose (OECD, 1996). It refers to the ability of an assay to measure the biological effect of interest. For Tier 1 EDSP assays, relevance can be defined as the ability of an assay to detect chemicals with the potential to interact with one or more of the EAT hormonal systems.

*Reliability* is defined as the reproducibility of results from an assay within and between or among laboratories.

The validation principles described by ICCVAM and OECD were originally designed for assays that would replace other assays. The OECD has recognized the need for flexibility in order to apply these methods more broadly and, therefore, provided recommendations documented in Guidance Document 34 (OECD, 2005). The EPA has adapted these principles for the EDSP as described in an EPA validation paper (EPA, 2007). The principles are as follows:

- 1) Provide scientific and regulatory rationale for the assay, including a clear statement of its purpose.
- 2) Address the relationship of the endpoints determined by the bioassays to the *in vivo* biological effect and toxicity of interest.
- 3) Provide a formal detailed protocol and make it available in the public domain. Include sufficient detail to enable the user to adhere to it and perform data analysis and apply decision criteria.
- 4) Evaluate within-test, intra-laboratory and inter-laboratory variability and how these parameters vary with time.
- 5) Demonstrate the assay’s performance using a series of reference chemicals coded to exclude bias.
- 6) Describe the limitations of the assay.
- 7) Obtain data in accordance with Good Laboratory Practices (GLPs).

- 8) Make publicly available all data supporting the assessment of the validity of the assays including the full data set collected during the validation studies and publish results in independent, peer-reviewed scientific journals.

#### 4.0 Selection of the Proposed Tier 1 Screening Battery

As recommended by the EDSTAC, the EPA is proposing a Tier 1 screening battery that is designed to detect whether a chemical substance interacts with the EAT hormonal systems. In selecting this battery, EPA has identified a combination of assays that provide complementary measurements that detect the endocrine disrupting potential of a chemical. Both *in vitro* and *in vivo* assays are employed to provide corroborating information supporting a weight-of-evidence approach among assays within the battery.

The proposed screening battery presented here has been designed to ensure that interaction with the EAT hormonal systems will be detected. The proposed battery is intended to fulfill the EDSTAC recommendations that a range of taxonomic groups be included to account for differences in endocrine systems and metabolic activity, and that sufficient diversity of endpoints is included to maximize sensitivity and minimize false negatives. In recommending assays for the battery, EDSTAC had to choose among assays that were highly specific for a hormonal activity and assays that may be less specific but more sensitive and apical (e.g., a more comprehensive assessment of functions that are relevant to development, reproduction, or chronic health). Although the proposed battery is composed of both types of assays, the EPA opted, as EDSTAC also recommended, to emphasize the latter (i.e., more sensitive) since this aspect corresponds with the overall mission of detecting potential endocrine-mediated effects regardless of mechanism of action.

The EPA is also sensitive to the issue of whole animal testing and continues to strive to reduce animal testing. Currently, the state of the science is such that *in vitro* and *in silico* methods cannot fully replace the role of *in vivo* test methods. Therefore, the EDSP Tier 1 battery as recommended includes both *in vitro* and *in vivo* tests.

Table 2 below provides EPA's proposed Tier 1 battery.

**Table 2: Screening Assays Proposed by EPA for the EDSP Tier 1 Battery**

<b><i>In vitro</i></b>
<sup>1</sup> Estrogen receptor (ER) binding – rat uterus
Estrogen receptor $\alpha$ (hER $\alpha$ ) transcriptional activation - Human cell line (HeLa-9903)
Androgen receptor (AR) binding – rat prostate
<sup>1,2</sup> Steroidogenesis – Human cell line (H295R)
<sup>2</sup> Aromatase – Human recombinant
<b><i>In vivo</i></b>
Uterotrophic (rat)
Hershberger (rat)
Pubertal female (rat)
Pubertal male (rat)
<sup>2</sup> Amphibian metamorphosis (frog)
<sup>2</sup> Fish short-term reproduction

<sup>1</sup>ER and H295R have not completed peer review yet and inclusion in the battery is contingent on successful review of these assays.

<sup>2</sup>Assays modified from the original assays recommended by EDSTAC.

A science-based approach to interpretation of the results of the battery will generally follow the principles recommended by EDSTAC, as shown in Appendix B. The primary principles to be considered include:

- ☐ Interpretation of the battery will be considered in light of the results of all assays in the battery, using a weight-of-evidence approach, taking into consideration *in vitro/in vivo* discrepancies (if any), metabolism, and route of exposure.
- ☐ When all screening assays are performed and all assays are negative, it may be concluded that the chemical will not likely interact with EAT hormonal processes included in the battery.

#### 4.1 Basis for Including Assays in the EDSP Proposed Tier 1 Screening Battery

The screening battery as proposed is intended to work as a whole. The basis for selecting a candidate assay to include in the battery involved: (1) the capacity of that assay to detect estrogenic- and androgenic-mediated effects by various modes of action including receptor binding (agonist and antagonist) and activation/transcription, steroidogenesis, and hypothalamic-pituitary-gonadal (HPG) feedback, and (2) the degree that *in vitro* and *in vivo* assays complemented one another in the battery as summarized in Table 3 below. In addition, rodent and amphibian *in vivo* assays were selected for the proposed battery based on their capacity to detect direct and indirect effects on thyroid function (hypothalamic-pituitary-thyroidal, HPT, feedback). Thus, the robustness of the proposed Tier 1 Screening Battery is based on the strengths of each individual assay and their complementary nature within the battery to detect effects on EAT hormonal function. The details of these strengths are discussed in Sections 4.1.1

to 4.1.5 and summarized in Section 6.0. Appendix A provides Fact Sheets for Assays included in the proposed battery.

**Table 3: Modes of Action Covered by Screening Assays Proposed in the Tier 1 Battery**

Assays	Modes of Action							
	E	Anti-E	A <sup>1</sup>	Anti-A	Steroidogenesis		HPG	HPT
					T <sup>1</sup>	E <sup>1</sup>		
<i>In vitro</i>								
ER Binding	■	■						
ER α Transcriptional Activation	■							
AR Binding			■	■				
Steroidogenesis H295R					■	■		
Aromatase Recombinant						■		
<i>In vivo</i>								
Uterotrophic	■							
Hershberger			■	■				
Pubertal Male			■	■	■		■	■
Pubertal Female	■	■				■	■	■
Amphibian Metamorphosis								■
Fish Short-term Reproduction (male & female)	■	■	■	■	■	■	■	

<sup>1</sup>A = Androgen; T = Testosterone; E = Estrogen

#### 4.1.1 Assays for detection of compounds that affect the estrogen signaling pathway.

The earliest concern for endocrine disruptors was related to environmental chemicals that could bind to the nuclear estrogen receptor and thereby interfere with the natural estrogen signaling pathway. As noted in the introduction, it was this concern that led to the statutory requirement in the FQPA to screen pesticide chemicals for estrogenic effects and is, therefore, the first mode of action that EDSTAC and EPA considered in designing a battery for the EDSP. Estrogen is important for reproductive function in both males and females, including sexual differentiation of the brain; development of secondary female sex characteristics; and structural and functional development as well as overall homeostasis of a large number of other systems.



Five assays within the battery are capable of detecting whether or not a chemical affects estrogen receptor function. Together these assays will detect chemicals with estrogenic and anti-estrogenic activity and include: 1) estrogen receptor (ER) binding, 2) ER transcriptional activation, 3) uterotrophic, 4) pubertal female, and 5) fish short-term reproduction. Of the five assays, the two *in vitro* assays (ER binding and ER transcriptional activation) identify the ability of the test chemical to interact with the estrogen receptor providing mechanistic information about how the chemical interacts at the cellular level. The three *in vivo* assays provide confirmatory evidence for the effects of the chemical following *in vivo* exposure via subcutaneous injection, oral gavage, and aquatic medium, respectively. The uterotrophic assay has been shown to detect weak estrogens with subcutaneous treatment, while the female pubertal assay can detect both estrogen agonists and antagonists by examining the age at vaginal opening among other estrogen-dependent endpoints. While many chemicals are active via the subcutaneous route, in some instance, oral exposure is more effective because of differences in absorption and metabolism. The different routes of exposure associated with the uterotrophic, female pubertal and fish short-term reproduction assays may provide guidance when designing any subsequent Tier 2 tests. Interpreting the results of the subset of estrogenic/anti-estrogenic assays within the battery is accomplished by examining the results of all the tests as the sum of all of the datasets is far greater than the information provided by any single assay alone. A brief description as well as the value of each of the five assays for ER follows.

#### ER Binding Assay

The ER receptor binding assay utilizing rat uterine cytosol (RUC) is a rapid *in vitro* assay that measures the affinity of a test chemical to bind to the estrogen receptor (or ligand binding domain of the estrogen receptor in the case of the human recombinant ER binding assays). It is a mechanistic assay that measures ligand-receptor interactions. It cannot distinguish between agonists, antagonists, or chemicals that have mixed agonist/antagonist activity or functional consequences of the interaction. Yet, the technical simplicity of the assay and its rapid turn-around time are conducive for screening large numbers of chemicals. Thus, the assay is a valuable tool for identifying chemicals that can compete with endogenous estrogen for ER binding. The practical use of this assay and its relevance to *in vivo* effects is well documented in the scientific literature. The assay has been standardized and examined with a number of positive and negative test compounds and the inter-laboratory studies are underway. EPA anticipates the completion of validation, including peer review, prior to the time that screening is expected to start in August 2008.

#### ER Transcriptional Activation Assay

The ER transcriptional activation (TA) assay is a method to detect the interaction and response of a chemical on the estrogen receptor. TA assays are based upon the expression of a reporter gene induced by a chemical following the ligand-receptor binding and subsequent transcriptional activation. As part of the Endocrine Disruption Testing and Assessment Task Force activity under the OECD Test Guidelines Program,

Chemicals Evaluation and Research Institute (CERI) of Japan developed and validated a stably transfected transactivation assay with ER $\alpha$  using the hER-HeLa-9903 (HeLa) cell line. This assay complements the ER binding assay as it can identify ER agonists. However, it has yet to be validated for ER antagonists. An OECD Test Guideline for the HeLa-9903 ER agonist assay has been developed and has been submitted for member country approval. Approval by the National Coordinators of the Test Guideline program is expected in April 2008. Thus, this assay is included in the battery as a more contemporary and functional approach to detect ER $\alpha$  agonists than the receptor binding assay.

#### Uterotrophic Assay

The uterotrophic assay is an *in vivo* assay that evaluates the ability of a chemical to elicit uterine growth consistent with the effects of estrogen agonists. The assay is generally conducted using adult ovariectomized females but may be conducted with sexually immature intact female rats in which endogenous estrogens are minimal. The sole endpoint measured is an increase in uterine weight in response to estrogen-induced water imbibition and hypertrophy. By using a subcutaneous (sc) route of exposure, the uterotrophic assay contributes information on a specific estrogen-related biological response that precludes any first-pass liver metabolism. Thus, data from this assay can complement the *in vitro* ER assays where metabolic activity is either non-detectable (ER binding) or minimal (TA assay) or has first passed through the liver as in the *in vivo* female pubertal assay. In regard to the latter, chemicals that are estrogenic prior to metabolism (e.g., bisphenol A) would be positive in the ER binding, TA assay, and uterotrophic assays, but not necessarily positive in the female pubertal assay. In contrast, chemicals that need to be metabolized in order to be estrogenic (e.g., methoxychlor) may be weak or likely missed in ER binding and TA assays, but positive in the female pubertal assay, and possibly positive in the uterotrophic assay at higher dose levels.

#### Pubertal Female Assay

The pubertal female assay is an *in vivo* assay that is sensitive to estrogens and antiestrogens. It is the only assay currently validated that can detect estrogen receptor agonists. For example, chemicals such as methoxychlor, nonylphenol, and octylphenol advance the age of vaginal opening (Laws *et al.*, 2000). However, there is a paucity of data describing the effects of known anti-estrogenic compounds in this assay. This is largely attributable to the fact that very few, if any, environmental anti-estrogens have been identified. The presumptive response to an anti-estrogen would be a delay in vaginal opening (VO). It should be noted that tamoxifen, a selective estrogen receptor modulator (SERM) with mixed agonist/antagonist activity was examined. This compound was found to advance VO because of the expected tamoxifen-induced estrogenic effects on the uterus and vagina. Since the age of VO can also be delayed by an effect on the HPG-axis, change in VO is not necessarily diagnostic for specific ER effects. However, estrogens also accelerate the age at first estrus and can induce vaginal cornification. In addition, when used in combination with the *in vitro* ER assays



and the uterotrophic assay, the distinction between an ER mechanism and other HPG mechanisms is readily apparent, which may provide guidance when designing any subsequent Tier 2 tests.

In the female pubertal assay, the animals are treated by oral gavage. This may be in contrast to the uterotrophic assay where sc dosing is an option. Considering the oral route of exposure, the pubertal female assay contributes information on a specific estrogen-related biological response for which absorption, distribution, metabolism and excretion (ADME) are fully taken into account and is crucial to the identification of antiestrogens and selective estrogen receptor modulators (SERMs).

#### Fish Short-Term Reproduction

The fish short-term reproduction assay with fathead minnows is designed to detect changes in spawning, morphology, and specific biochemical endpoints that reflect disturbances in the HPG axis, including estrogen agonists and antagonists. Although some endpoints may be highly diagnostic for estrogen signaling interaction (e.g., vitellogenin induction in males), not all endpoints in the assay are intended to unequivocally identify specific endocrine modes of action. However, collectively, the endpoints observed do allow inferences to be made with regard to possible endocrine disturbances involving the estrogen hormonal pathway and, thus, provide guidance for further testing.

Vitellogenin (egg yolk protein) production is primarily controlled through estrogen interaction with the estrogen receptor; hence, it is directly related to a mechanism of concern. It is a well-established endpoint, and commercial availability of ELISA kits specific to the fathead minnow have made vitellogenin production readily measurable. Induction of vitellogenin in male fish is an extremely sensitive and specific indication of ER agonists because males normally have very low circulating concentrations of endogenous estrogen and therefore vitellogenin. Reproductively active females normally have moderate circulating concentrations of vitellogenin which can be decreased by ER antagonists. Estrogens and anti-estrogens can also affect egg production in the fish assay. Changes in fecundity combined with alterations in gonadal histopathology provide a good indication of reproductive health and have been demonstrated to be sensitive to estrogenic and anti-estrogenic exposures.

#### **4.1.2 Assays for detection of compounds that affect the androgen signaling pathway.**

Androgens are critical for sexual differentiation and development of secondary sex characteristics in the male, as well as for a wide variety of functions in both males and females. To date, a number of environmental chemicals have been shown to act as androgens or antiandrogens. Four assays within the battery are capable of detecting whether or not a chemical affects androgen receptor function. Together these assays will detect chemicals with androgenic and anti-androgenic activity and include: 1) AR binding, 2) Hershberger, 3) pubertal male and 4) fish short-term reproduction.

The *in vitro* AR binding assay provides mechanistic information on the cellular (nuclear) mode of action. The three *in vivo* assays provide confirmatory evidence for the effects of a chemical on the reproductive system. Specifically, the Hershberger assay is diagnostic for both androgenic and anti-androgenic activity. The male pubertal and fish reproduction assays reflect changes in AR regulation, but due to their apical nature are also sensitive to chemicals that may affect other modes of action involved in HPG function. Again, interpreting results of these assays within the battery is supported by examining the results of all the assays, as the sum of the four datasets is far greater than the information provided by any one assay.

#### AR Binding Assay

The androgen receptor binding assay (AR binding), utilizing rat prostate cytosol, is a rapid *in vitro* assay that measures the affinity of a test chemical to bind to the androgen receptor. It is a mechanistic assay that measures only ligand-receptor interactions. As with the ER binding assay, its technical simplicity along with its rapid turn-around time are conducive for screening large numbers of chemicals. Thus, the assay is a valuable tool for identifying chemicals that can compete with the endogenous ligand.

While the AR binding assay detects both agonists and antagonists, it cannot distinguish between the two. Thus, it can be used in conjunction with the Hershberger assay which can distinguish agonists from antagonists.

#### Hershberger Assay

The Hershberger assay is a short-term *in vivo* screen that evaluates the ability of a chemical to elicit biological activities consistent with either androgen agonists or antagonists by utilizing changes in the weights of five androgen-dependent tissues: 1) ventral prostate, 2) seminal vesicle, 3) levator ani-bulbocavernosus (LABC) muscle, 4) Cowper's glands, and 5) glans penis. Specifically, an increase in tissue weights is diagnostic of androgenic activity. In contrast, an anti-androgenic chemical will block any increase in tissue weights when co-administered with an androgen such as testosterone propionate. The Hershberger contributes to the battery by providing information on a specific androgen-related biological response and, being an *in vivo* assay, integrates ADME into the responses. The assay has been used to identify the anti-androgenic effects of several chemicals including vinclozolin and flutamide (Gray *et al.*, 1999; McIntyre, *et al.*, 2001).

#### Pubertal Male Assay

The male pubertal assay is an *in vivo* test sensitive to disruptions by chemicals that act as androgens or antiandrogens or interfere with androgen synthesis. Importantly, as an *in vivo* assay, it can detect chemicals which require metabolism in order to interact with the AR. For example, chemicals such as vinclozolin delay the age

of preputial separation and decrease the growth of androgen dependent tissues (Gray *et al.* 1994, 1999). The male pubertal assay is reproducible and sensitive for chemicals which alter androgenic hormone action and provides useful confirmatory information for AR agonists and antagonists which are detected in the *in vitro* AR receptor assay.

#### Fish Short-Term Reproduction

Secondary sex characteristics of fathead minnows are endpoints that are affected by androgenic/anti-androgenic substances. Specifically, females will develop external male secondary sex characteristics (nuptial tubercles) when exposed to an AR agonist. This endpoint not only is quite specific for this mode of action, but is very sensitive in that females typically do not express these characteristics. In contrast, AR antagonists decrease the expression of male secondary sex characteristics in male fathead minnows. Changes in secondary sex characteristics in fathead minnows are biologically relevant, unique and robust. Inter-laboratory comparisons of these endpoints have been reproducible. Androgens and anti-androgens also effectively inhibit egg production in the fish assay with concurrent alterations in gonad histopathology.

#### **4.1.3 Assays for detection of compounds that affect steroid synthesis**

A number of environmental compounds have been shown to interfere with the synthesis of estrogens (e.g., estradiol) and androgens (e.g., testosterone). In this regard, a number of *in vitro* assays for steroidogenesis were considered for the battery, with the decision to include the H295R cell line as it offers the potential to identify chemicals that induce or inhibit testosterone and estradiol synthesis. In addition, since many environmental compounds are known to inhibit aromatase, the decision was made to validate a human recombinant aromatase assay. These two *in vitro* assays, in addition to three *in vivo* assays (pubertal female, pubertal male and fish short-term reproduction), are expected to provide sufficient information for making informed decisions as to whether or not a compound interferes with the production of these two important biologically active steroids.

Again, data from both the *in vitro* and *in vivo* assays provide the necessary information to determine whether or not the compound affects steroidogenesis. If hormone production is affected only in the *in vivo* protocols, with no *in vitro* verification, it is likely that the compound impairs hypothalamic-pituitary function and subsequently alters gonadotropin synthesis/secretion. The indication of hypothalamic-pituitary effects are discussed below.

#### H295R for Steroidogenesis

H295R is a human adrenocortical carcinoma cell line that possesses all of the key enzymes throughout the steroidogenic pathways. Several studies have shown that these enzymes and their mRNA and products can all be readily measured in a high-throughput format. For the purposes of the screening battery, the measurement of

testosterone and estradiol produced with or without the test compound are the key endpoints. The H295R cell's ability to metabolize xenobiotics is presumed to be low, although this has not been characterized. The assay provides a straightforward and inexpensive way to detect chemicals that affect steroid hormone synthesis either by inhibiting the enzymes in the pathway, leading to decreased production of one or more of the hormones measured, or inducing the production of enzymes, leading to increased production of one or more of the hormones measured. Although other assays in the battery may detect the adverse effects of chemicals that interfere with steroid hormone synthesis, they can not identify the specific component of the pathway that was altered. The assay has been standardized and examined with a number of positive and negative test compounds. The protocol has been published (Hecker *et al.*, 2006) and the inter-laboratory comparisons have been completed. It is anticipated that validation, including peer review, will be successfully completed prior to the time that screening is expected to start in August 2008.

Because the H295R assay detects increases and decreases in aromatase activity, it is a potential candidate to eventually replace the recombinant aromatase assay as an *in vitro* screen in the proposed Tier 1 battery. It is expected that as increasing amounts of data become available during screening of the first 73 chemicals, the performance (*e.g.*, sensitivity and specificity) of these assays within the context of the battery can be determined.

#### Human recombinant aromatase

The human recombinant aromatase assay is an inexpensive, rapid method to detect chemicals that inhibit aromatase activity and thus block the conversion of androgens to estrogens. The ability of the H295R assay to also detect inhibitors of aromatase may make this assay redundant as discussed above. However, it is included in the proposed battery at this time for two reasons: 1) the sensitivity and specificity of the H295R has not been fully characterized relative to that of the recombinant aromatase assay, and 2) the H295R assay has not completed the assay validation process and peer review. The most relevant limitation of the recombinant aromatase assay is its inability to detect inducers of aromatase gene transcription. Either the H295R assay or the aromatase assay is required in the Tier 1 battery since these are the only assays that have been shown to be sensitive enough to detect the activity of xenobiotics that weakly inhibit aromatase and estrogen synthesis.

#### The Pubertal Female and Pubertal Male Assays

Changes in the numerous hormone-dependent endpoints in the male and female pubertal assays will detect the effects of a chemical that interferes with endogenous steroid hormone production by the testes and ovaries, respectively. Although the testosterone dependent endpoints measured in the pubertal male assay do not provide the diagnostic information necessary to discern impaired steroidogenesis, the *in vitro* data from the H295R cell assay does. Similarly, the male pubertal assay was shown to detect chemicals that affect steroidogenesis prior to the formation of estrogen (*e.g.*,

ketoconazole). The female pubertal assay will detect effects of altered aromatase activity. For example, changes in the endpoints in this assay were identified following fadrazole exposure (Marty *et al.*, 1999). However, a low dose of another purported aromatase inhibitor, fenarimol, was without effect, which may have been due to the inherent estrogenicity of the compound (Vinggaard *et al.*, 2005). Nonetheless, while both the male and female pubertal protocols do possess the capacity to readily detect compounds that affect steroidogenesis, the proper diagnosis for a steroidogenic mode of action can only be made with supporting *in vitro* data. It has also been shown that measurement of androgen levels and testis synthesis of androgens enhances the sensitivity of the male assay to detect chemicals that block androgen synthesis (Blystone *et al.*, 2007).

#### Fish Short-Term Reproduction

Interference in the steroid synthesis pathways is detected by several endpoints in the fish assay. Proliferation of interstitial cells (Leydig cells) in the male testes, reduction of circulating concentrations of steroids, decreased plasma vitellogenin in females, and impaired reproduction would all signal potential steroid synthesis modulation. Although the specific mode of action would not be discernible, the assay is still sensitive to effects involving the steroidogenic pathways.

#### **4.1.4 Assays for detection of chemicals that affect the HPG axis**

The EDSTAC determined that evaluating the effect of environmental chemicals on the hypothalamic-pituitary-gonadal axis (HPG) was also important. To address this issue, the battery includes the male and female pubertal assays and the fish short-term reproduction assay, which includes both males and females. The hypothalamic-pituitary regulation of reproductive development and function is sensitive to a number of environmental compounds, as there are a variety of target mechanisms that can be affected. It is well known that many pharmaceuticals can interfere with hypothalamic regulation of gonadal function and ultimately gonadal hormone and gamete production. Similarly, environmental compounds such as dithiocarbamates (Stoker *et al.*, 2001), formamidines (Goldman *et al.*, 1993), chlorotriazines (Cooper *et al.*, 2000), among others, have been found to interfere with endocrine function by altering the hypothalamic regulation of pituitary hormone secretion. By this mode of action, it has been shown that many of these same chemicals can interfere with reproductive development and aging.

The battery, as designed, does not have an *in vitro* test for hypothalamic effects simply because of the myriad of pathways and mechanisms associated with normal function of this highly complex process. However, it is possible to use the combined results of the *in vivo* tests included in the proposed battery to determine deductively that the HPG axis was altered. Thus, if a chemical is found to delay puberty in both male and female rats, but not alter either ER or AR binding or steroidogenesis, the delay in puberty is likely due to impaired hypothalamic-pituitary function. This is exactly the profile produced by compounds that act on the central nervous system such as



dithiocarbamate thiram (impairs norepinephrine synthesis and GnRH release) and atrazine (impairs LH secretion) when assessed in the male and female pubertal assays. Pharmaceuticals such as bromocriptine, pimozide and haloperidol, which alter dopaminergic receptor function (Marty *et al.*, 1999; Female Pubertal ISR), have also been evaluated in the pubertal assays. In every case, these compounds that act on the central nervous system were found to alter normal pubertal development. Thus, appropriate identification of a compound that affects hypothalamic-pituitary function may require the type of comparisons shown in Table 4.

**Table 4: Profiles Diagnostic for Various Modes of Action (MOA) in the Proposed Tier 1 Screening Battery.**

Assay	Was a Change Detected?			
	Yes	No	No	No
ER Binding	Yes	No	No	No
AR Binding	No	Yes	No	No
Steroidogenesis	No	No	Yes	No
Male Pubertal	No	Yes	Yes	Yes
Female Pubertal	Yes	No	Yes	Yes
Fish	Yes	Yes	Yes	Yes
<b>Likely MOA:</b>	<b>ER</b>	<b>AR</b>	<b>Steroidogenesis</b>	<b>HPG</b>

The fish short-term reproduction assay with fathead minnows is designed to detect changes in spawning, morphology and specific biochemical endpoints that reflect alterations in the HPG axis, including (anti-) estrogen and (anti-) androgen pathways. Again, as with the male and female pubertal assays, it is important to recognize that the assay is not intended to differentiate, quantify or confirm the mode of action, but to provide indirect evidence that certain endocrine regulated processes involving the estrogen and androgen hormonal pathways may be sufficiently perturbed to warrant more definitive examination. Although some endpoints may be highly diagnostic (*e.g.*, vitellogenin induction in males and tubercle formation in females), not all endpoints in the assay are intended to unequivocally identify specific cellular mechanisms of action but, collectively, the endpoints observed in the assays included in the proposed battery allow inferences to be made with regard to possible endocrine disturbances and, thus, provide guidance for further testing.

#### 4.1.5 Assays for detection of chemicals that affect the HPT axis

In addition to identifying environmental compounds that have the potential to alter the hormonal regulation of reproductive function involving the estrogen and androgen hormonal pathways, certain assays included in the proposed Tier 1 screening battery will also provide relevant information about the potential of a chemical to interfere with thyroid function. Thyroid hormone is essential for normal development and for maintenance of normal physiological functions in vertebrates. Delivery of thyroid hormones to tissues and cells is highly regulated during early development and in the adult, and it is governed by complex physiological processes involving the hypothalamic-pituitary-thyroid (HPT) axis, including peripheral organs/tissues. Environmental factors, such as the presence of specific toxicants, can perturb this

system at various points of regulation, inducing a variety of responses that can be detected with thyroid-related endpoints in the assays. Three assays have been identified as useful for this purpose: 1) pubertal female, 2) pubertal male and 3) amphibian metamorphosis.

### Pubertal Male and Female Assays

The pubertal male and female assays include multiple endpoints that can detect an interaction of a test chemical with the thyroid hormone system, including serum T4 and TSH concentrations, thyroid organ weight, thyroid histology and liver weight. Both the male and the female assays have been shown to detect thyrotoxicants that act by various mechanisms that interfere with the synthesis and elimination of thyroid hormones. While the male and female pubertal assays include the same thyroid endpoints, examining the thyroid axis in both sexes provides the opportunity to detect gender differences in response to treatment. It is not clear whether the male or the female is more sensitive to toxicants that interfere with the thyroid axis at this early age. However, the male pubertal assay may be more robust than the female because the male is dosed longer. It has been shown in prevalidation studies that the male may be more sensitive to chemicals that induce hepatic clearance of thyroid hormone based on the response to lower dose levels. In the male assay, a food restriction study showed that a reduction in terminal weight of 9% or greater relative to controls could result in a decrease in circulating T4 concentrations (Laws *et al.*, 2007). This effect will need to be considered when interpreting the battery for the thyroid mode of action. However, the redundancy of the *in vivo* assays should assist with interpretation of such data, as one change in a single endpoint will not be interpreted as a positive result if the other assays find no effect involving the HPT axis.

### Amphibian Metamorphosis

The amphibian metamorphosis assay (AMA) is a screening assay intended to identify substances which may interfere with the normal function of the HPT axis. The AMA represents a generalized vertebrate model to the extent that it is based on the conserved structure and function of thyroid systems among species. The amphibian metamorphosis assay provides a well-studied, thyroid-dependent process which responds to substances active along the HPT axis, and it is the only proposed assay for the Tier 1 battery that assesses thyroid activity in a species undergoing morphological development.

The AMA is based on the principle that the dramatic morphological changes that occur during post-embryonic development are dependent upon the normal functioning of the HPT axis, and that interference with these processes leads to measurable effects. During tadpole metamorphosis, thyroid hormone (TH) influences virtually every tissue in the animal's body initiating diverse morphological, physiological and biochemical changes that include cell proliferation, differentiation and death. The result is *de novo* organ formation, organ loss, and extensive tissue remodeling. Given the dependence of metamorphosis on TH, and the strict biochemical control under which these

processes occur, the timing and character of these processes can serve as experimental endpoints representative of thyroid axis function and, as such, are exploited in the AMA. Additionally, although post-embryonic development appears quite different in mammals and most amphibians (direct development versus metamorphosis), there is a high level of evolutionary conservation of the thyroid system among vertebrates and the underlying cellular and molecular pathways that control these processes are similar, if not identical. The evolutionarily conserved nature of the vertebrate thyroid system enhances the ability to use an amphibian, particularly anurans, as a general model for evaluating HPT axis interference such that the results can be extrapolated to other vertebrate species.

The primary endpoints in the AMA are developmental stage, hindlimb length, and thyroid histology. Each endpoint can be affected by chemicals that interact with the HPT axis. For example, antagonists of thyroid production, iodination and action, such as perchlorate and methimazole, will delay development and induce diagnostic lesions in the thyroid gland. Thyroid agonists (e.g., native thyroid hormone) will accelerate development. Additionally, unlike the mammalian assays that have been developed to detect interactions along the HPT axis, the AMA has the ability to detect chemicals that act on peripheral tissues. For example, inhibition of monodeiodinases that transform T4 to T3 can cause asynchronous development, detected by an inability to assign a developmental stage to a tadpole. The detection of this mechanism is important because, in this case, development can be affected without concomitant effects on thyroid histology or circulating thyroid hormone.

#### **4.2 Basis for Not Including Potential Screening Assays from the Proposed EDSP Tier 1 battery**

Throughout the validation process there were some screening assays (sliced testes steroidogenesis and *in utero* through lactational) in which technical difficulties associated with assay design or protocol development could not be immediately resolved and, therefore, they were not considered feasible and, subsequently, assay validation was suspended. For some other screening assays (placental aromatase and adult male), the validation process was completed and they were initially considered as candidate assays for the Tier 1 battery. However, they were either deemed impractical and replaced with a more contemporary assay, or weak or limited in providing strength to the battery in a complimentary manner with other candidate assays. The basis for not including some EDSTAC-recommended screening assays in the proposed Tier 1 battery is discussed in more detail in the following sections. Appendix C provides fact sheets for assays not included in the proposed battery.

##### Sliced Testes Steroidogenesis Assay

A steroidogenesis assay using fragments of rat testicular tissue was evaluated by the EPA. The variability of this assay and lack of specificity of cytotoxicity tests to differentiate between Leydig and other parenchymal cells in the tissue cultures were problematic. Thus, after presentation of prevalidation data to the Endocrine Disruptor



Methods Validation Advisory Committee (EDMVAC, 2005), it was recommended to suspend assay development and direct attention to developing and standardizing the steroidogenesis H295R cell-based assay.

#### Placental Aromatase Assay

The aromatase assay was validated using human placental aromatase (derived from fresh tissue) and a recombinant human placental aromatase (available commercially). Both sources of the enzyme were equivalent when tested during the validation process. Since the enzyme is readily available, the recombinant assay is the more practical assay of choice as proposed in the Tier 1 screening battery. However, the option exists for the test sponsor's to submit a request to the EPA to use the placental aromatase assay *in lieu* of the recombinant assay.

#### In Utero through Lactational Assay

The EPA presented the alternative *in utero* through lactational (IUL) rat screening assay to the FIFRA SAP in February 2007 to consider whether the IUL assay, as represented by Protocol C and tested with methoxychlor, was suitable as an alternative Tier 1 screening assay and whether the assay validation process should continue using Protocol C or some other protocol (e.g., A or B).

In general, the SAP considered Protocol C too complex for a Tier 1 screen and not in accord with the EDSTAC criteria of a Tier 1 assay. Although it was felt that Protocol B or a modification of it could be validated as a simpler screen, there was concern that none of the proposed protocols (A, B, C) could be standardized and validated to find utility as a routine Tier 1 screen.

In consideration of the SAP response to charges in the final report (SAP, 2007), the EPA concluded that the IUL assay was too complex, lengthy, and costly to serve as a Tier 1 screening assay at this time. The EPA recognizes the importance of prenatal and neonatal developmental periods and notes that *in utero* through lactational exposure is currently covered in the EDSP by multi-generational tests in Tier 2. Thus, further development and validation of the alternative IUL screening rat assay has been suspended and is not being considered in the proposed Tier 1 battery.

#### Adult Male Assay

The 15-day intact adult male rat assay was considered as an alternative assay in the proposed EDSP Tier 1 screening battery as recommended by EDSTAC; however, after individual assay peer review and considering other alternative assays in the battery, EPA chose not to include the assay in the proposed battery at this time based, in part, on several biological and technical weaknesses and limitations.

As proposed by industry, an anticipated strength of the adult male assay was its unique suite of reproductive hormonal assays (O'Connor *et al.*, 2002). By using adult

male rats with mature HPG and HPT axes, the hormonal assay results were expected to be less variable and, therefore, more interpretable than results collected from immature rats. Moreover, the hormonal results were to serve as additional primary endpoints and, in combination with reproductive organ weight and associated histomorphological changes, provide a unique profile of information (*i.e.*, “fingerprint”) to differentiate the mechanistic nature of the responses following test chemical exposure. Two test chemicals, an anti-androgen (linuron) and a compound that has an indirect effect on the thyroid (phenobarbital), were evaluated simultaneously in three independent contract laboratories (ISR for the Adult Male Rat). Results of the inter-laboratory validation exercise indicated that the concept of a chemical responsive “fingerprint”, especially for relatively weak anti-androgens, was not supported. Despite enhancing and controlling for some technical aspects within the adult male assay protocol (e.g., 15 animals/dose group, source of animals and feed, hormonal assay kits, timing of necropsy, common statistical analyses), there was extensive variation in the serum hormone results among the contract laboratories and in reference to historical control values that were established primarily by one industrial laboratory. Moreover, within some contract laboratories, test-chemical organ weight and/or associated histopathological changes were not supported by respective hormonal changes, especially for those hormones involved in the androgen pathway (e.g., LH, testosterone, DHT). Thus, it was apparent that the mature life-stage did not provide an added strength to the adult male assay for the measurement of a suite of reproductive hormones; consequently, the hormonal endpoints were relegated to a secondary or supporting role within the assay as was discussed in the Adult Male ISR and supported by peer review.

Despite the limitation, the EPA considered the scenario of replacing the female pubertal assay with the alternative adult male assay in the Tier 1 battery as recommended by the EDSTAC. As a result, however, the number of female *in vivo* endpoints within the battery would have been markedly diminished since the uterotrophic assay provides only a single *in vivo* endpoint (*i.e.*, uterine weight). Moreover, the uterotrophic assay does not involve an intact HPG axis or screen potential differences between genders regarding the HPT axis. Alternatively, if the pubertal female assay was kept in the battery, there would be a better representation of female endpoints but numerous male litter mates would not be utilized according to the current pubertal female protocol. Thus, to balance the number of mammalian female with male *in vivo* endpoints in the battery and provide efficient and effective use of both female and male pups, the EPA chose to include the alternative pubertal male assay in the proposed Tier 1 battery as recommend by the EDSTAC along with the pubertal female assay to be run in parallel. The complementary nature of running both the pubertal female and male assays in the proposed battery, as discussed in Section 4.1, adds more to the robustness of the battery than replacing the pubertal female with the adult male assay.

## 5.0 Performance Review of the EDSP Tier 1 Screening Battery

When a joint committee of the SAB and SAP submitted their review of the EDSP in a final report in 1999, it was recommended that EPA convene a panel of independent scientists to review all the screening data for the first series of test chemical substances (SAB/SAP, 1999). It is expected that the panel would consider whether or not the battery adequately identified those chemicals with EAT activity and whether or not there were clear indications for triggering Tier 2 tests. Subsequent to performance review of the battery, selected protocols may be modified, or assays eliminated and replaced. Prior to performance review of the battery, it is also reasonable to expect that new *in vitro* and, perhaps, *in vivo* assays will emerge that will serve to replace some of the current screening assays. Depending, in part, on availability, EPA may pursue protocol development, standardization and validation of more novel assays for eventual use as replacement assays in a second generation EDSP Tier 1 screening battery.

## 6.0 Summary

As recommended by EDSTAC, the EPA is proposing a battery that is designed to identify the effect of environmental chemicals on specific endocrine modes of action, including estrogen, androgen, thyroid and HPG. In selecting this battery, the EPA has identified a combination of tests that provide complementary measurements that detect the endocrine disrupting potential of a chemical. Both *in vitro* and *in vivo* tests are employed to provide corroborating information supporting a weight-of-evidence of approach. These protocols were developed and standardized through an extensive and ambitious validation effort that included individual assay peer review. The validation and peer review processes helped to clarify the strengths and limitations for these assays and their intended mode of action which include the detection of:

- 1) Estrogenicity and anti-estrogenicity through the estrogen receptor binding, transcriptional activation, uterotrophic, fish and pubertal female assays;
- 2) Androgenicity and anti-androgenicity through the androgen receptor binding, Hershberger, fish, and pubertal male assays;
- 3) Inhibition and induction of the steroidogenic pathway from cholesterol through testosterone with the H295R cell-based assay, and pubertal male and fish assays;
- 4) Inhibition of aromatase with the aromatase, H295R, pubertal female and fish assays;
- 5) The HPG axis through neuroendocrine mediated effects on the control of gonadal function with the pubertal male and female rat and fish assays; and the

- 6) HPT axis through agonistic and antagonistic effects and chemical induced changes in thyroid homeostasis with the pubertal male and female rat and amphibian metamorphosis assays.

Table 5 lists the assays that provide information on these modes of action and a qualitative estimate of the strength of the battery for each assay and mode of action (MOA). It is intended to summarize and show how these individual assays may be used in developing a weight-of-evidence evaluation of a chemical's effect on the selected MOAs. The weight-of-evidence evaluation would be used to determine whether or not further testing is required in Tier 2. It should be noted that although it is necessary to take into account the weight-of-evidence in discussing the composition of the Tier 1 battery, the EPA is still developing recommendations for how the Agency will apply a weight-of-evidence approach in determining whether or not Tier 2 testing is necessary. Thus, the SAP is being asked at this time to comment only on the adequacy of the battery to cover all known modes of action recommended by the EDSTAC so that the EPA will be able to further develop and make final recommendations on the process of determining the outcome of Tier 1 screening.

**Table 5: Summary of the Strengths in Detecting EAT effects in the Proposed Tier 1 Screening Battery According to Mode of Action (MOA)**

MOA	Assays	Individual Assay – Strength (per validation) <sup>1</sup>	Overall Strength of Battery for MOA Prediction
<b>Estrogenic Activity</b> (ER agonist)	ER receptor binding: <i>in vitro</i>	+++	<b>Strong</b> (Good predictive ability based on several assays - both <i>in vivo</i> and <i>in vitro</i> validated assays)
	ER gene expression assay: <i>in vitro</i> , (validation ongoing by OECD)	+++	
	Uterotrophic assay: <i>in vivo</i> ; uterine weight and histology (subcutaneous or oral; subcutaneous usually more sensitive)	+++	
	Pubertal female assay: <i>in vivo</i> ; induction of pseudoprecocious puberty in female rat (oral, sometimes more sensitive than uterotrophic)	+++	
	Fathead minnow assay: <i>in vivo</i> ; increased vitellogenin synthesis in male fish	+++	
<b>Anti-estrogenic Activity</b> (ER	ER receptor binding: <i>in vitro</i>	+++	<b>Weak</b> (Primarily

MOA	Assays	Individual Assay – Strength (per validation) <sup>1</sup>	Overall Strength of Battery for MOA Prediction
antagonist)	<p>ER gene expression assay: <i>in vitro</i>; (but OECD did not validate the assay for this mode)</p> <p>Pubertal female assay: <i>in vivo</i>; delayed VO and increased diestrus in female rat</p> <p>Fathead minnow assay: <i>in vivo</i>; decreased vitellogenin synthesis in female fish</p>	<p>-</p> <p>+</p> <p>+</p>	reliant on a single <i>in vitro</i> assay – some support from other validated assays)
<b>Androgenic Activity</b> (AR agonist)	<p>AR receptor binding: <i>in vitro</i></p> <p>AR gene expression assay: <i>in vitro</i>; (not validated by OECD; EPA anticipates conducting validation efforts in the near future)</p> <p>Hershberger assay: <i>in vivo</i>; organ weights (oral; most sensitive <i>in vivo</i> assay)</p> <p>Pubertal male assay: <i>in vivo</i>; induction of pseudoprecocious puberty in male rat and increased organ weights (oral; little data on a single very potent androgen)</p> <p>Fathead minnow assay: <i>in vivo</i>; induced male traits in female fish and infertility</p>	<p>+++</p> <p>-</p> <p>+++</p> <p>+</p> <p>+++</p>	<b>Strong</b> (Good predictive ability based on several assays - both <i>in vivo</i> and <i>in vitro</i> validated assays)
<b>Anti-androgenic Activity</b> (AR Antagonist)	<p>AR receptor binding: <i>in vitro</i></p> <p>AR gene expression assay: <i>in vitro</i> (not validated, by OECD; EPA anticipates conducting validation efforts in the near future)</p> <p>Hershberger assay: <i>in vivo</i>; organ weights (oral; most sensitive <i>in vivo</i> assay).</p>	<p>+++</p> <p>-</p> <p>+++</p>	<b>Strong</b> (Good predictive ability based on several assays - both <i>in vivo</i> and <i>in vitro</i> validated assays)

MOA	Assays	Individual Assay – Strength (per validation) <sup>1</sup>	Overall Strength of Battery for MOA Prediction
	<p>Pubertal male assay: <i>in vivo</i>; delay in puberty in male rat and reduced organ weights (oral; reasonable data base with fairly weak anti-androgens)</p> <p>Fathead minnow assay: <i>in vivo</i>; attenuated male secondary sex characteristics (effects are seen, but not clearly diagnostic of this mode at low concentrations).</p>	<p>+++</p> <p>++</p>	
<b>Inhibition of steroidogenesis</b> – testosterone	<p>H295R cell assay: <i>in vitro</i></p> <p>Pubertal male assay: <i>in vivo</i>; delayed PPS and reduced serum testosterone in male (little data from validation exercise but data are available in the literature from ORD)</p> <p>Fathead minnow assay: <i>in vivo</i>; decreased hormone levels, fertility, increased interstitial cells in testes, and sexual traits</p>	<p>+++</p> <p>++</p> <p>+++</p>	<b>Moderate to Strong</b> (Good predictive based on more than one assay – both <i>in vivo</i> and <i>in vitro</i> )
<b>Inhibition of aromatase - estrogen synthesis</b>	<p>Recombinant aromatase assay H295R cell assay; <i>in vitro</i></p> <p>Pubertal female assay: <i>in vivo</i>; inhibition of aromatase in ovary delays puberty, but only works for potent inhibitors (false negatives were found)</p> <p>Fathead minnow assay.</p>	<p>+++</p> <p>+++</p> <p>+</p> <p>++</p>	<b>Moderate</b> (Primarily reliant on two <i>in vitro</i> assays – some support from other validated <i>in vivo</i> assays)
<b>Altered Hypothalamic-pituitary function</b>	<p>Pubertal male assay: <i>in vivo</i>; delay in puberty in male rat and reduced organ weights (oral dosing, data from ORD studies)</p> <p>Pubertal female assay: <i>in vivo</i>: delay in puberty in female rat and</p>	<p>++</p> <p>++</p>	<b>Moderate</b> (Good predictive based on three assays – only <i>in vivo</i> )



MOA	Assays	Individual Assay – Strength (per validation) <sup>1</sup>	Overall Strength of Battery for MOA Prediction
	abnormal estrous cycles (oral dosing, data from ORD studies)  Fathead minnow assay: <i>in vivo</i> ; should affect fertility and hormones.	+	
<b>Anti-thyroid activity</b>	Pubertal male assay: <i>in vivo</i> ; reduced serum T4 and T3, increased TSH and thyroid histology in male rat	++	<b>Strong</b> (For some modes of action)
	Pubertal female assay: <i>in vivo</i> ; reduced serum T4 and T3, increased TSH and thyroid histology in female rat	++	<b>Weak</b> (For some modes of action)
	Amphibian metamorphosis: <i>in vivo</i> ; delayed metamorphosis and altered thyroid histology (uncertain sensitivity to weak chemicals and some modes of action or specificity).	+++	
<b>Thyromimetic activity</b>	Amphibian metamorphosis: <i>in vivo</i> ; accelerated metamorphosis and altered thyroid histology (uncertain sensitivity to weak chemicals).	++	<b>Weak</b> (Primarily reliant on a single assay)

<sup>1</sup>+ and – signs indicate detection or no detection respectively. Number of + signs indicates degree of detection (+=some ability to +++= robust).

Appendix D provides a comparison of the battery recommended by EDSTAC, as well as the battery that EPA is proposing for Tier 1 screening. EPA's proposed battery is based on the recommendations of EDSTAC as well as knowledge gained from the validation process.

## 7.0 Charge Questions for the FIFRA SAP

The SAP is asked to review and provide comment on each of the charges listed below:

1. Please comment on the ability of the proposed Tier 1 screening battery to provide sufficient information to determine whether or not a substance potentially interacts with the estrogen, androgen, and thyroid hormonal systems based on the modes of action covered within the battery:
  - a. Estrogenicity: acting agonistically by potentiating the estrogen signal.
  - b. Anti-estrogenicity: acting antagonistically by attenuating the estrogen signal.
  - c. Androgenicity: acting agonistically by potentiating the androgen signal.
  - d. Anti-androgenicity: acting antagonistically by attenuating the androgen signal.
  - e. Steroidogenesis effects: acting agonistically or antagonistically by modulating normal steroidogenic processes including aromatase.
  - f. Hypothalamic/pituitary/gonadal effects: acting agonistically or antagonistically by modulating processes not captured in the above categories.
  - g. Hypothalamic/pituitary/thyroid effects: acting agonistically or antagonistically by modulating processes associated with direct thyroid hormone receptor interaction as well as those processes involved indirectly (e.g., synthesis, secretion, elimination of thyroid hormones) in thyroid function.
2. EPA proposed a Tier 1 screening battery that includes many assays that are complementary in nature and in their coverage of the EAT hormonal systems (the strengths of one assay offset the limitations of another), albeit by different taxa, life-stages, endpoints, exposure and use of *in vitro* and *in vivo* methods executed at different levels of biological organization (e.g., cytosolic receptor binding, cell-based assays, whole organism).
  - a. Please comment on how well the proposed battery minimizes the potential for “false negatives” and “false positives.”
  - b. Please comment on whether there are unnecessary redundancies for MOAs across the battery.
  - c. Please comment on whether a different combination of validated assays would be more effective in achieving the purpose of the battery than that proposed by EPA.



## 8.0 References

- Blystone, C.R.; Lambright, C.S.; Furr, J.; Wilson, V.S.; Gray, L.E., (2007). Iprodione delays male rat pubertal development, reduces serum testosterone levels, and decreases *ex vivo* testicular testosterone production. *Toxicology Letters*. 174:74-81.
- Cooper, R.L.; Stoker, T.E.; Tyrey, L.; Goldman, J. M.; McElroy, W.K. (2000). Atrazine disrupts hypothalamic control of pituitary-ovarian function. *Toxicological Sciences*. 53: 297-307.
- EDMVAC (2005). Consensus Recommendation on the Continued Development of the Sliced Testis Steroidogenesis Assay.  
[http://www.epa.gov/scipoly/oscpendo/pubs/edmvac/leblanc\\_letter\\_edmvac\\_06\\_05.pdf](http://www.epa.gov/scipoly/oscpendo/pubs/edmvac/leblanc_letter_edmvac_06_05.pdf).
- EDSTAC (1998). Endocrine Disruptor Screening and Testing Advisory Committee, Final Report, Volume I-II.  
<http://www.epa.gov/scipoly/oscpendo/pubs/edspoverview/finalrpt.htm>.
- EPA (1998). Federal Register Notice (FRN), volume 63, page 71542; Dec 28, 1998.
- EPA (2006). U.S. Environmental Protection Agency Peer Review Handbook, 3rd Edition, EPA/100/B-06-002, Prepared by Members of the Peer Review Advisory Group, for EPA's Science Policy Council, available at  
<http://www.epa.gov/peerreview/pdfs/Peer%20Review%20HandbookMay06.pdf>.
- EPA (2007) Validation of Screening and Testing Assays Proposed for the EDSP. Version 5.4. July, 2007.
- Goldman, J. M. and Cooper, R. L. (1993). Assessment of toxicant-induced alterations in the luteinizing hormone control of ovulation. In: *Methods in Toxicology*. Vol III, Part B. Female Reproductive Toxicology ed. by R.E. Chapin and J. Heindel, Academic Press, Orlando, pp79-91.
- Gray, L.E. Jr.; Ostby, J.S.; Kelce, W.R. (1994). Developmental effects of an environmental antiandrogen: The fungicide vinclozolin alters sex differentiation of the male rat. *Toxicol Appl Pharmacol* 129:46-52.
- Gray, L.E. Jr.; Ostby, J., Monosson, E.; Kelce, W.R. (1999). Environmental antiandrogens: Low doses of the fungicide vinclozolin alter sexual differentiation of the male rat. *Toxicol Ind Health* 15:48-64.

Hecker, M.; Newsted, J.L.; Murphy, M.B.; Higley, E.B.; Jones, P.D.; Wu, R.; Giesy, J.P. (2006). Human adrenocarcinoma (H295R) cells for rapid in vitro determination of effects on steroidogenesis: hormone production. *Toxicol. Appl. Pharmacol.* 15;217(1):114-124.

ICCVAM (2000). Interagency Coordinating Committee for the Validation of Alternative Methods (ICCVAM) Authorization Act of 2000, 42 U.S.C. 285, Public Law 106-545, signed December 19, 2000.

Laws, S.C.; Carey, S.A.; Ferrell, J.M.; Bodman, G.J.; Cooper R.L. (2000) Estrogenic activity of octylphenol, nonylphenol, bisphenol A, and methoxychlor in rats. *Toxicological Sciences.* 54(1): 154-1667.

Laws, S.; Stoker, T.; Ferrel, J.; Hotchkiss, M.; Cooper, R. (2007). Effects of altered food intake during pubertal development in male and female Wistar rats. *Toxicological Sciences.* 100(1) 194-202.

Marty, M.S.; Crissman, J.W.; Carney, E.W.; (1999). Evaluation of the EDSTAC female pubertal assay in CD rats using 17beta-estradiol, steroid biosynthesis inhibitors, and a thyroid inhibitor. *Toxicol Sci.* Dec;52(2):269-77.

McIntyre, B.S.; Barlow, N.J.; Foster, P.M. (2001). Androgen-mediated development in male rat offspring exposed to flutamide in utero: Permanence and correlation of early postnatal changes in anogenital distance and nipple retention with malformations in androgen-dependent tissues. *Toxicol Sci* 62:236-249.

NIEHS (1997) National Institute of Environmental Health Sciences. "Validation and Regulatory Acceptance of Toxicological Test Methods, A Report of the ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods." Research Triangle Park, NC. NIH Report 97-3981. March, 1997.

O'Connor, J.C.; Cook, J.C.; Marty, M.S.; Davis, L.G.; Kaplan, A.M.; Carney, E.W. (2002). Evaluation of Tier I screening approaches for detecting endocrine-active compounds (EACs). *Crit. Rev. Toxicol.* 32, 521-549.

OECD (1996). Organisation for Economic Cooperation and Development. Final Report of the OECD Workshop on Harmonization of Validation and Acceptance Criteria for Alternative Toxicological Test Methods. August, 1996.

OECD (2005). Organisation for Economic Cooperation and Development. Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. Guidance Document No. 34. June 2005.

SAB/SAP (1999). Review of the Endocrine Disruptor Screening Program by a Joint Subcommittee of the Science Advisory Board and Scientific Advisory Panel. EPA-SAB-EC-99-013.

SAP (2007). Status of the in utero through lactational assay in the endocrine disruptor screening program (EDSP). Scientific Advisory Panel. SAP Minutes No. 2007-04. <http://www.epa.gov/scipoly/sap/meetings/2007/2007/february/sap-2007-04-report-f.pdf>.

Stoker, T.E.; Goldman, J.M.; Cooper, R.L. (2001). Delayed ovulation and pregnancy outcome: effect of environmental toxicants on the neuroendocrine control of the ovary. *Environmental Toxicology and Pharmacology*, 9:117-129.

Vinggaard, A.M.; Jacobsen, H.; Metzdorff, S.B.; Andersen, H.R.; Nelleman, C. (2005) Antiandrogenic effects in short-term *in vivo* studies of the fungicide fenarimol. *Toxicology* 207: 21-34.

## **Appendix A – Fact Sheets for Assays Included in the Proposed EDSP Tier 1 Screening Battery**

The fact sheets for individual assays are arranged in alphabetical order as indicated:

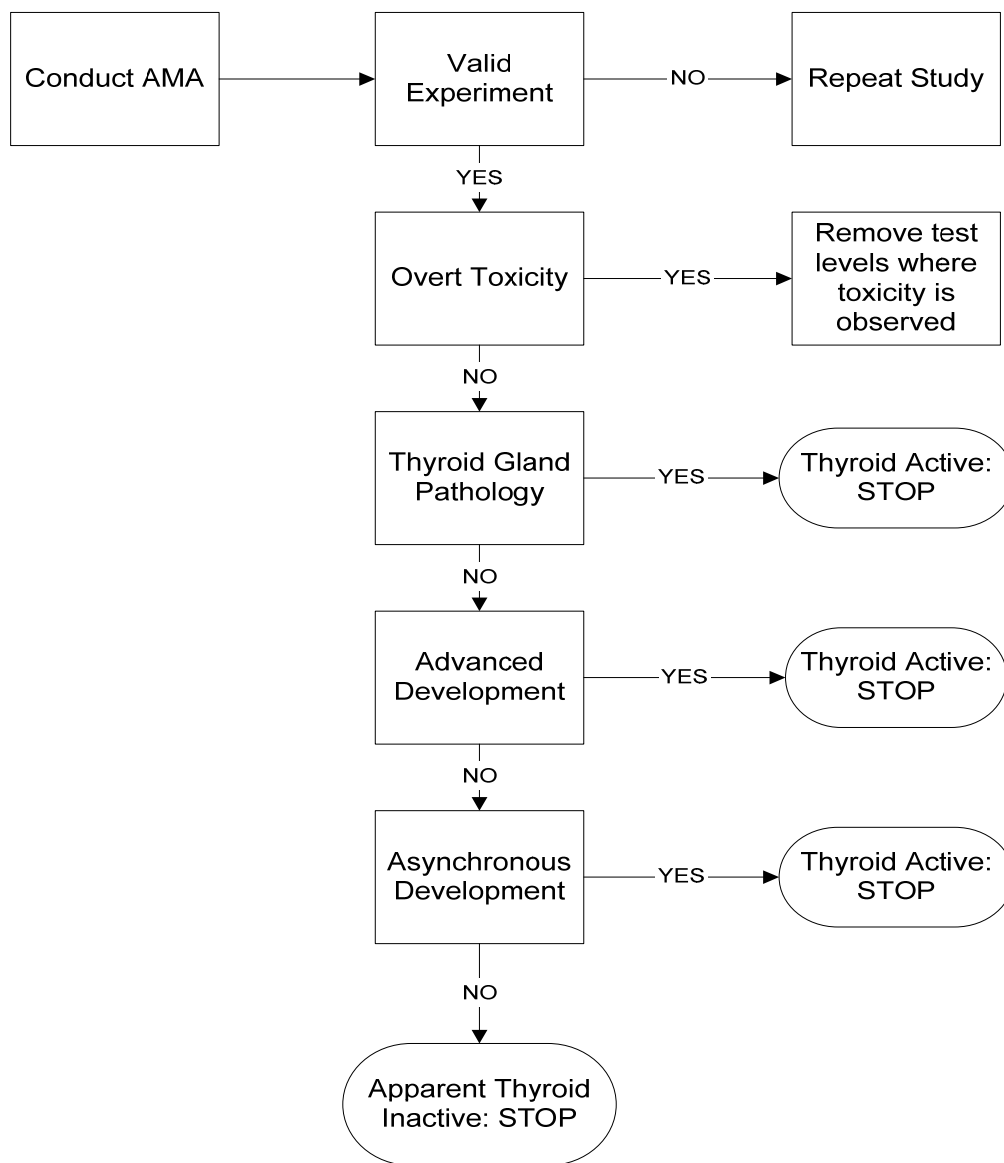
- Appendix A1 – Amphibian Metamorphosis
- Appendix A2 – Androgen Receptor Binding
- Appendix A3 – Aromatase (Recombinant)
- Appendix A4 – Estrogen Receptor Binding
- Appendix A5 – Estrogen Receptor Stably Transfected Transcriptional Activation
- Appendix A6 – Fish Short-term Reproduction (Fish Assay)
- Appendix A7 – Hershberger
- Appendix A8 – Pubertal Female
- Appendix A9 – Pubertal Male
- Appendix A10 – Steroidogenesis (H295R)
- Appendix A11 – Uterotrophic

## Appendix A1 – Amphibian Metamorphosis

Amphibian Metamorphosis	
Purpose	The AMA is a screening assay intended to empirically identify substances which may interfere with the normal function of the hypothalamic-pituitary-thyroid (HPT) axis. The AMA represents a generalized vertebrate model to the extent that it is based on the conserved structure and functions of thyroid systems. It is not intended to quantify or confirm endocrine disruption, or to provide a quantitative assessment of risk, but only provide evidence that thyroid regulated processes may be sufficiently perturbed to warrant more definitive testing.
Design	The general experimental design entails exposing <i>Xenopus laevis</i> tadpoles at NF stage 51 to a minimum of three different aqueous concentrations of a test chemical and a dilution water control for 21 days. There are four replicate tanks at each test substance concentration or treatment. Larval density at test initiation is 20 tadpoles per test tank for all treatment groups.
Endpoints	<p>Daily mortality</p> <p>Morphological endpoints</p> <p>Whole body length/snout-vent length (d 7 and 21)</p> <p>Hind limb length (d 7 and 21)</p> <p>Wet weight (d 7 and 21)</p> <p>Developmental stage (d 7 and 21)</p> <p>Histology</p> <p>Thyroid gland (d 21)</p>
Interpretation	Results are evaluated for evidence of interaction of the test chemical with the HPT axis as follows. Data values and the study report are evaluated for deviations from the test method or performance criteria to evaluate the validity of the study. If necessary, test concentrations with overt toxicities are removed from the data set. Significant histological findings in thyroid tissue deem the assay positive. If no thyroid gland histopathology is observed, then developmental landmarks are evaluated. If development is accelerated or asynchronous, the test is deemed positive. The assay is considered negative if no effects are detected in thyroid gland histology or morphological landmarks of development. The following decision flow chart diagrams the interpretation logic.

## Amphibian Metamorphosis

### Decision Flow Chart



Amphibian Metamorphosis	
Main peer review comments	<ul style="list-style-type: none"> <li>• Assay is relevant to its purpose.</li> <li>• Data interpretation needs to be better described.</li> <li>• Protocol is generally clear and appropriate, but notable changes and clarifications are needed to improve the assay.</li> </ul> <p>[EPA accepts the recommendations of the peer review panel and will revise the protocol guidance accordingly.]</p>
Strengths (within the context of the proposed battery)	<ul style="list-style-type: none"> <li>• intact <i>in vivo</i> system on an animal undergoing morphological development allowing for evaluation of parent compounds and degradates</li> <li>• amphibian metamorphosis is a well-studied developmental process that is dependent on thyroid hormone, thus effects on metamorphic development are relatively specific indicators of HPT axis perturbation</li> <li>• conserved nature of the components and functions of the amphibian HPT axis are relevant for other vertebrate classes</li> <li>• apical assay covering several modes of HPT axis interaction, including central homeostatic mechanisms and peripheral mechanisms</li> <li>• redundant endpoints, maximizing chance for detection while minimizing false negatives</li> <li>• provides toxicological data in a taxon (amphibians) underrepresented in available Agency protocols</li> <li>• well-established relationship between endpoints and endocrine system</li> <li>• endpoints easy to measure</li> </ul>
Limitations (within the context of the proposed battery)	<ul style="list-style-type: none"> <li>• inherent difficulties in testing some substances not amenable to aquatic systems</li> <li>• sensitivity of the assay has not been fully characterized</li> <li>• non-thyroidal toxicities have the potential to affect some of the morphological endpoints of the assay</li> </ul>

## Appendix A2 – Androgen Receptor Binding

Androgen Receptor Binding	
Purpose	The AR binding assay is a sensitive <i>in vitro</i> test to detect chemicals that may affect the endocrine system by binding to the androgen receptor. It will give added confidence that positive results seen in the Hershberger assay are truly due to an AR binding mechanism.
Design	Cytosol isolated from the rat prostate provides the source of the androgen receptor. Test chemical and R1881, a strong ligand, compete for binding with the AR when incubated together overnight. The assay measures the binding of [ <sup>3</sup> H]-R1881 in the presence of eight test chemical concentrations. Unlabeled R1881 serves as a strong positive control producing the standard curve. Dexamethasone is run with each block of test chemicals as a weak positive control. Solvent is the negative control.
Endpoints	The DPM of [ <sup>3</sup> H]-R1881 is measured by liquid scintillation counter. Data for the standard curve and each test chemical will be plotted as the percent [ <sup>3</sup> H] R1881 bound versus the molar concentration through use of a four parameter non-linear regression program.
Interpretation	<p>Performance criteria have been set for the top, bottom, and slope for R1881 and the weak positive, dexamethasone.</p> <ul style="list-style-type: none"> <li>• If the binding curve crosses 50% (competes to displace the standard ligand by 50%) the test chemical is considered to be a binder.</li> <li>• Chemicals for which the binding curve crosses 75%, but not 50%, are considered to be equivocal.</li> <li>• Chemicals which do not fit the model or for which the binding curve does not cross 75% are considered to be non-binders.</li> </ul>
Main peer review comments	<ul style="list-style-type: none"> <li>• The AR assay is highly relevant to detect substances that bind to the AR receptor in humans.</li> <li>• Reproducibility of data for weaker chemicals by less proficient laboratories was a significant problem [EPA response: Performance standards have been developed that will screen out weaker laboratories.]</li> <li>• An adequate range and number of test chemicals were selected for validation</li> <li>• There are better assays for this purpose than this assay including recombinant assays and transcriptional activation assays. [EPA response: These other assays have not been validated and, therefore, are not ready for use in a regulatory program at this time.]</li> </ul>



Androgen Receptor Binding	
Strengths (within the context of the proposed battery)	<ul style="list-style-type: none"><li>• Specific for identifying an interaction with the AR receptor providing mechanistic information</li><li>• More sensitive than typical <i>in vivo</i> assays</li><li>• Rapid</li><li>• Inexpensive</li><li>• It will give added confidence that positive results seen in the Hersberger and other male <i>in vivo</i> assays are truly due to an AR</li></ul>
Limitations (within the context of the proposed battery)	<p>Although it detects both agonists and antagonists although it cannot predict the consequences of binding</p> <ul style="list-style-type: none"><li>• Other steroids that are not natural ligands for the AR may bind at high concentrations</li><li>• Chemicals that denature the receptor may be identified as false positives</li><li>• No metabolic capability</li></ul>

## Appendix A3 – Aromatase (Recombinant)

Aromatase (Recombinant)	
Purpose	The aromatase assay detects chemicals that inhibit aromatase activity. Aromatase is the enzyme that metabolizes androgens such as testosterone to estrogens
Design	Androstenedione and [ $1\beta$ - $^3\text{H}$ ]-androstenedione (ASDN) serve as substrate for human recombinant microsomal aromatase. Full activity control (ASDN in medium, no inhibitor), background activity control (no NADPH), positive control (4-hydroxyandrostenedione at eight concentrations) and test chemical (8 concentrations) are run in the reaction for 15 minutes, and the reaction products produced are measured and plotted as percent enzyme activity (inhibition curve) through use of a non-linear regression program.
Endpoints	The formation of $^3\text{H}_2\text{O}$ , one of the co-reaction products along with estrone, is measured by liquid scintillation counter.
Interpretation	Chemicals that reduce enzyme activity levels by 50% or more (as determined by the inhibition curve calculated by a four parameter non-linear regression program) are considered to be inhibitors of aromatase. Chemicals that fit the inhibition curve but allow 50-75% activity, <i>i.e.</i> , reduce activity by 25-50%, are considered equivocal. Chemicals that do not fit the model or that fit the model but reduce inhibition by 25% are considered to be non-inhibitors of aromatase.
Main peer review comments	<ul style="list-style-type: none"> <li>• Comments supported the use of the assay for the intended purpose, the clarity of the protocol, the data interpretation procedure, and performance criteria; however, one reviewer noted that <math>K_i</math> determination would be superior to <math>\text{IC}_{50}</math>.</li> <li>• The chemicals and analytical methods used in the validation of the assay were appropriately chosen.</li> <li>• The protocol could be further optimized for small volumes resulting in less cost and waste and greater convenience. [EPA response: This option will be permitted in the revised protocol.]</li> <li>• There are better assays that could have been selected such as cell-based assays which would be advantageous in that</li> </ul>

Aromatase (Recombinant)	
	they would detect both induction and inhibition. [EPA response: EPA is currently validating the H295R assay. See H295R fact sheet.]
Strengths (within the context of the proposed battery)	<ul style="list-style-type: none"><li>• Highly specific to inhibition of aromatase activity providing mechanistic information.</li><li>• More sensitive than typical in vivo assays</li><li>• Rapid</li><li>• Inexpensive</li><li>• Capable of high throughput</li><li>• Provides useful information for the interpretation of in vivo assays.</li></ul>
Limitations (within the context of the proposed battery)	<ul style="list-style-type: none"><li>• Cannot detect chemicals that induce aromatase activity.</li><li>• False positives could result from chemicals that denature the enzyme.</li><li>• Limited/no ability to metabolize xenobiotics</li></ul>

## Appendix A4 – Estrogen Receptor Binding

Estrogen Receptor Binding	
Purpose	The purpose of the estrogen receptor (ER) binding assay is to identify test chemicals that can bind to the estrogen receptor.
Design	<p>A saturation radioligand binding experiment is conducted to demonstrate that the ER binding assay is working under optimal conditions within a given laboratory. This assay is conducted by measuring the equilibrium binding of increasing concentrations of <math>^3\text{H}</math>-estradiol to rat cytosolic or human recombinant ER<math>\alpha</math>. Nonlinear regression analysis of the data provides estimates of the affinity of the receptor for 17<math>\beta</math>-estradiol (<math>K_d</math>) and the concentration of receptors (<math>B_{\text{max}}</math>).</p> <p>A competitive ER binding assay is conducted by measuring the equilibrium binding of a single concentration of <math>^3\text{H}</math>-17 <math>\beta</math>-estradiol at various concentrations (over a range of at least six orders of magnitude) of a test chemical in rat cytosolic or human recombinant ER. After equilibration, the amount of radioactivity bound to the ER is measured as an indicator of how much was displaced by the test compound at each concentration. Data analysis provides an estimate of the potency of the test chemical for binding to the ER relative to 17 <math>\beta</math>-estradiol. 17 <math>\beta</math>-estradiol is run as a reference standard with each run, as are a weak positive and a non-binder.</p> <p>In each portion of the study, three replicate data points are collected at each concentration in one run, and three independent runs are performed to constitute one assay.</p>
Endpoint	Binding curve fit to a four-parameter Hill equation, where the parameters are top, bottom, slope, and $\log(\text{IC}_{50})$ ( <i>i.e.</i> , base-10 log of the molar concentration of test chemical which inhibits 50% of binding by the radioligand).
Interpretation	<p>Performance criteria have been set for the top, bottom, and slope for 17<math>\beta</math>-estradiol and the weak positive, norethynodrel, for the competitive ER binding portion of the assay. Within-run variability is also subject to a performance criterion.</p> <p>Classification of test chemicals:            Positive: A <math>\log(\text{IC}_{50})</math> value can be obtained from an unconstrained curve fitted to the Hill equation that has a slope of approximately -1.0.            Equivocal: Acceptable binding curve reaches 25% displacement</p>

Estrogen Receptor Binding	
	<p>of radioligand but not 50% at the highest concentration. Also applied if slope is unusually steep or shallow.</p> <p>Negative: Acceptable binding curve does not reach 25% displacement of the radioligand; or curve cannot be fit and no data point shows displacement of more than 25%.</p>
Main peer review comments	<ul style="list-style-type: none"><li>• [Peer review of this assay expected by the end of August 2008.]</li></ul>
Strengths (within the context of the proposed battery)	<ul style="list-style-type: none"><li>• Quick (two days)</li><li>• Uses relatively few, or no, animals (depending on source of receptor – rat uterine cytosol or human recombinant)</li><li>• Specific for identifying an interaction with estrogen receptor (<i>i.e.</i>, provides mechanism-related information)</li></ul>
Limitations (within the context of the proposed battery)	<ul style="list-style-type: none"><li>• Although it detects both agonists and antagonists, it cannot distinguish the consequences of binding.</li><li>• Does not account for potential metabolic activation or deactivation of test chemical.</li></ul>

## Appendix A5 – Estrogen Receptor Stably Transfected Transcriptional Activation

Estrogen Receptor Stably Transfected Transcriptional Activation	
Purpose	Provides an in vitro cell-based assay to detect chemicals that bind to the estrogen receptor and alters gene transcription.
Design	The HeLa-9903 (HeLa) cell line (derived from a human cervical tumor) has been stably transfected with the hER $\alpha$ expression construct (full length receptor) and a reporter bearing five tandem repeats of a vitellogenin estrogen response element driven by a mouse metallothionein promoter TATA element.
Endpoints	Measurement of bioluminescence reflecting changes in gene transcription.
Interpretation	Data are reported as EC <sub>50</sub> , PC (percent of positive control response) 50, and PC10. The final data interpretation criteria are being finalized at OECD.
Main peer review comments	<ul style="list-style-type: none"> <li>• Robust assay providing similar results to other TA assays.</li> <li>• Criteria for a positive response need further definition. [This is being addressed in the OECD test guideline.]</li> <li>• Need better guidance on acceptable performance criteria. [This is also being addressed in the OECD test guideline.]</li> <li>• Performance only demonstrated for estrogen agonists. [Japan is planning to validate the antiestrogen MOA but this effort is complicated by lack of clear reference chemicals.]</li> </ul>
Strengths (within the context of the proposed battery)	<ul style="list-style-type: none"> <li>• Can detect binders to ER<math>\alpha</math></li> <li>• Has the potential to distinguish agonists from antagonists.</li> <li>• Assay incorporates a quantitative biological response as well as concentration at which the response occurs.</li> </ul>
Limitations (within the context of the proposed battery)	<ul style="list-style-type: none"> <li>• Only detects binders to ER<math>\alpha</math></li> <li>• Not yet validated for ER antagonists</li> <li>• Limited metabolism</li> </ul>



## Appendix A6 – Fish Short-term Reproduction (Fish Assay)

Fish Short-term Reproduction (Fish Assay)	
Purpose	The fish short-term reproduction assay is a screening assay intended to identify changes in morphology, histopathology, spawning, and specific biochemical endpoints which may reflect interference with the normal function of the hypothalamic-pituitary-gonadal (HPG) axis. It is not intended to quantify or confirm endocrine disruption, or to provide a quantitative assessment of risk, but rather to provide suggestive evidence that endocrine-regulated processes may be sufficiently perturbed to warrant more definitive testing.
Design	The fish short-term reproduction assay entails exposing reproductively mature fathead minnows ( <i>Pimephales promelas</i> ) to a minimum of three concentrations of a test chemical and appropriate control(s) for 21 days. Successful spawning is established during a pre-exposure period of at least 14 days. Each of the four replicate tanks in each treatment level contains four females and two males.
Endpoints	<p>Survival</p> <p>Behavior</p> <p>Body length</p> <p>Body weight</p> <p>Fecundity*</p> <ul style="list-style-type: none"> <li>• # of spawns</li> <li>• # of eggs/female reproductive day</li> </ul> <p>Fertilization Success*</p> <ul style="list-style-type: none"> <li>• # fertile eggs/female reproductive day</li> <li>• % fertile eggs</li> </ul> <p>Gonadal Histopathology*</p> <p>Gonadosomatic Index (GSI)*</p> <p>Appearance and Secondary Sex Characteristics*</p> <ul style="list-style-type: none"> <li>• Overall body coloration, vertical banding</li> <li>• Fatpad (weight, score, index)</li> <li>• Tubercles (count, score)</li> <li>• Ovipositor size</li> </ul> <p>Biochemical measures*</p> <ul style="list-style-type: none"> <li>• Vitellogenin</li> <li>• Estradiol</li> <li>• Testosterone</li> </ul> <p>* key endpoints</p>

## Fish Short-term Reproduction (Fish Assay)

Interpretation	<p>The fish short-term reproduction assay as presented is intended to serve in a screening capacity to provide an indication of potential endocrine activity, not to confirm any specific mechanism, mode of action, or adverse effect. Therefore, a significant effect in one or more of the key endpoints of this assay (fecundity, fertilization success, histopathology, GSI, biochemical measures, and secondary sex characteristics) should be considered indicative of possible endocrine system disturbance. The suite of endpoints included is necessary to provide a fully comprehensive assessment of the disrupting potential to the HPG-axis in a representative fish.</p> <p>It is important to note however that if a given exposure level results in substantial mortality or other overt signs of toxicity, responses in other endpoints may be due to general toxicity, not necessarily mediated primarily via interaction with the endocrine system. The lower treatment level(s) should be examined for effects outside of the range of general toxicity. If all test concentrations exhibit mortality, then the assay would need repeating before inference on possible endocrine activity can be made.</p> <p>It is recognized that some endpoints may be responsive to non-endocrine stresses in addition to endocrine-mediated pathways, particularly fecundity. Although reductions in fecundity indicate adverse organismal and, potentially, population level effects (<i>i.e.</i>, reproductive toxicity), these cannot be definitively distinguished from direct endocrine-mediated effects by this assay when changes in other core endpoints are not present. Nevertheless, reductions in fecundity are considered a positive effect in this assay because they may be endocrine-mediated and should be considered in concert with results of the other assays in the Tier 1 battery. Results that would be considered equivocal for this single assay should be considered indications of potential endocrine activity and evaluated in light of the weight-of-evidence from the other assays in the Tier I battery of assays for the EDSP.</p>
Main peer review comments	<ul style="list-style-type: none"> <li>• Agreed that the assay is biologically and toxicologically relevant to the stated purpose.</li> <li>• Agreed that the overall design and endpoints selected are generally highly appropriate for screening for HPG perturbing chemicals, particularly (anti-) estrogenic and (anti-) androgenic compounds.</li> <li>• Recommend that fish are as similar as possible in egg production at the beginning of exposure.</li> <li>• Recommend that fish are sexually mature and of similar and optimal age for reproduction and to avoid mistaking immature males for females.</li> <li>• Recommend clarifying guidance for equal distribution of spawning groups among treatments to avoid bias.</li> </ul>

Fish Short-term Reproduction (Fish Assay)	
	<ul style="list-style-type: none"> <li>Suggest clarifying use of behavior observations.</li> </ul> <p>[EPA accepts the recommendations given and will revise the protocol guidance accordingly.]</p>
Strengths (within the context of the proposed battery)	<ul style="list-style-type: none"> <li>Incorporates a standard, easily acquired laboratory model species, <i>Pimephales promelas</i>, and utilizes common aquatic toxicology methods;</li> <li>Straightforward, cost effective, reasonably short-term assay;</li> <li>Detects (anti-)estrogen and (anti-)androgen perturbations in addition to disruptors of the entire HPG axis using reproductively active male and female fish;</li> <li>Employs an intact HPG axis and hence is relevant to other taxa when conserved elements of the HPG axis are considered;</li> <li>Reproducible results demonstrated in multiple laboratories;</li> <li>Informs the appropriate concentration range to be used in Tier 2 testing, which avoids the need for an additional range-finding study and reduces the number of animals needed.</li> </ul>
Limitations (within the context of the proposed battery)	<ul style="list-style-type: none"> <li>Inherent technical difficulties testing substances that are poorly soluble in water in aquatic systems, and methods for delivering such substances to the test system. (Generally addressed on OECD Guidance Document 23);</li> <li>Some measurements (e.g., plasma steroids) will require specialized technical expertise</li> </ul>

## Appendix A7 – Hershberger

Hershberger	
Purpose	A short-term in vivo assay to detect androgenic or antiandrogenic chemicals or chemicals that inhibit 5 $\alpha$ -reductase.
Design	There are two versions of the Hershberger Assay: an immature version and a peripubertal version employing castrated rats. In addition, each version can be run to detect AR agonists or antagonists. When screening for potential androgenic activity, the test substance is administered daily by oral gavage or subcutaneous injection for a period of ten consecutive days. Test substances are administered to a minimum of two treatment groups of experimental animals using one dose level per group. When screening for potential antiandrogenic activity, the test substance is administered daily by oral gavage or subcutaneous injection for a period of ten consecutive days in concert with daily TP doses (0.2 or 0.4 mg/kg/d) by sc injection. Graduated test substance doses are administered to a minimum of two treatment groups of experimental animals using one dose level per group. In both the agonist and antagonist procedures, the animals are necropsied approximately 24 hours after the last dose.
Endpoints	The assay is based on statistically significant changes in weight in androgen dependent tissues. The five tissues weighed in the castrated male are ventral prostate (VP), seminal vesicle (SV) (plus fluids and coagulating glands), levator ani-bulbocavernosus (LABC) muscle, paired Cowper's glands (COW) and the glans penis (GP). In the immature version the GP cannot be detached and measured but the testes and epidimides are weighed in the intact weanling.
Interpretation	A positive result is a statistically significant change in the weight of two of the tissues.

Hershberger	
Main peer review comments	<ul style="list-style-type: none"><li>• Assay could be used effectively to detect androgen agonists, antagonists and inhibitors of 5<math>\alpha</math>-reductase.</li><li>• Additional work should be undertaken to characterize the rate of false positives. [Addressed by OECD.]</li><li>• Clear guidance should be provided on the MTD [Addressed by OECD in the test guideline.]</li><li>• The data interpretation procedure needs further definition [Addressed by OECD in the test guideline.]</li></ul>
Strengths (within the context of the proposed battery)	<ul style="list-style-type: none"><li>• Relatively rapid screen that is quite specific to androgenic effects (see limitations)</li><li>• In vivo procedure incorporates metabolism; thus, it can detect chemicals that need activation.</li><li>• Oral administration will model a primary exposure route and incorporates ADME</li></ul>
Limitations (within the context of the proposed battery)	<ul style="list-style-type: none"><li>• The growth response of the individual androgen-dependent tissues is not entirely of androgenic origin, <i>i.e.</i>, compounds other than androgen agonists can alter the weight of certain tissues. However, the growth response of several tissues concomitantly substantiates a more androgen-specific mechanism.</li></ul>

## Appendix A8 – Pubertal Female

Pubertal Female	
Purpose	This assay is capable of detecting chemicals with estrogenic/anti-estrogenic activity, or agents which alter pubertal development via changes in steroidogenesis, or hypothalamic-pituitary regulation of the ovary and thyroid homeostasis.
Design	Test chemical is administered daily by gavage from post-natal day (PND) 22 to PND 42 (21 days) to 15 females per dose. Two doses plus vehicle control are employed. The animals are weighed daily, and examined for vaginal opening from PND 22 until opening is complete. After vaginal opening, vaginal smears are taken daily. Additional measures are taken at necropsy.
Endpoints	<p>Growth (daily body weight)</p> <p>Age and weight at vaginal opening</p> <p>Organ weights:</p> <ul style="list-style-type: none"> <li>Uterus (blotted)</li> <li>Ovaries (paired)</li> <li>Thyroid</li> <li>Liver</li> <li>Kidneys (paired)</li> <li>Pituitary</li> <li>Adrenals (paired)</li> </ul> <p>Histology</p> <ul style="list-style-type: none"> <li>Uterus</li> <li>Ovary</li> <li>Thyroid (colloid area and follicular cell height)</li> <li>Kidney</li> </ul> <p>Blood Chemistry, standard panel</p> <p>Hormones</p> <ul style="list-style-type: none"> <li>Serum or plasma thyroxine (<math>T_4</math>), total</li> <li>Serum or plasma thyroid stimulating hormone (TSH)</li> </ul> <p>Estrus cyclicity</p> <ul style="list-style-type: none"> <li>Age at first estrus after vaginal opening</li> <li>Length of cycle</li> <li>Percent of animals cycling</li> <li>Percent of animals cycling regularly</li> </ul>
Interpretation	Results are evaluated for evidence of interaction of the test chemical with the endocrine system, primarily estrogen- and thyroid-related. Body weight, organ weight, and hormone values for the control animals are subject to performance criteria for mean and coefficient of variation. Thyroid endpoints are generally interpreted separately from the sex-hormone-related endpoints.



Pubertal Female	
Main peer review comments	<ul style="list-style-type: none"> <li>• On the whole, the purpose and protocol are clear</li> <li>• Vaginal opening is a sensitive endpoint for assessing estrogen function, alteration of steroidogenesis or HPG axis.</li> <li>• Uterine and ovarian weights in cycling animals are variable due to the estrous cyclicity and may not be useful endpoints.</li> <li>• The overall detection of the effects of the test chemicals was comparable across laboratories although not always on an endpoint by endpoint basis.</li> </ul>
Strengths (within the context of the proposed battery)	<ul style="list-style-type: none"> <li>• Intact mammalian <i>in vivo</i> system, and thus addresses ADME concerns.</li> <li>• Apical assay covering several modes of interaction, including ones not covered by other assays in battery</li> <li>• Redundant endpoints, maximizing chance for detection while minimizing false negatives</li> <li>• Covers pubertal period of development</li> <li>• Well-established relationship between endpoints and endocrine system</li> </ul>
Limitations (within the context of the proposed battery)	<ul style="list-style-type: none"> <li>• Protocol is not as diagnostic for specific MOAs as other assays in the battery such as uterotrophic for ER agonist.</li> <li>• Although a toxic negative chemical has not been identified, several chemicals positive for one of the MOAs have been found to be negative for the other MOAs evaluated in this assay.</li> </ul>

## Appendix A9 – Pubertal Male

Pubertal Male	
Purpose	Provide information obtained from an <i>in vivo</i> mammalian system that is useful in determining the potential of chemicals or mixtures to interact with the endocrine system. Detect chemicals with antithyroid, androgenic, or antiandrogenic [androgen receptor (AR) or steroid-enzyme-mediated] activity or agents which alter pubertal development via changes in gonadotropins, prolactin, or hypothalamic function.
Design	Test chemical is administered daily by gavage from post-natal day (PND) 23 to PND 53 (31 days) to 15 males per dose level at two dose levels plus vehicle control. The animals are weighed daily, and examined for preputial separation from PND 30 until separation is complete. The other measurements are taken at necropsy.
Endpoints	<p>Growth (daily body weight)</p> <p>Age and weight at preputial separation</p> <p>Organ weights</p> <ul style="list-style-type: none"> <li>seminal vesicle plus coagulating gland</li> <li>ventral prostate</li> <li>dorsolateral prostate</li> <li>levator ani plus bulbocavernosus muscle complex</li> <li>epididymis</li> <li>testis</li> <li>thyroid</li> <li>liver</li> <li>kidney</li> <li>adrenal</li> <li>pituitary</li> </ul> <p>Blood Chemistry, standard panel</p> <p>Hormone levels</p> <ul style="list-style-type: none"> <li>serum testosterone, total</li> <li>serum thyroxine, total</li> <li>serum thyroid stimulating hormone</li> </ul> <p>Histology</p> <ul style="list-style-type: none"> <li>epididymis</li> <li>testis</li> <li>thyroid</li> <li>kidney</li> </ul>

Pubertal Male	
Interpretation	Results are evaluated for evidence of interaction of the test chemical with the endocrine system, primarily androgen- and thyroid-related. Body weight, organ weight, and hormone values for the control animals are subject to performance criteria for mean and coefficient of variation. Thyroid endpoints are generally interpreted separately from the androgen-related endpoints.
Main peer review comments	<ul style="list-style-type: none"> <li>• Assay is relevant to its purpose.</li> <li>• Hormone assays should be standardized and centralized QC standards maintained by EPA. [EPA will provide better guidance for standardization but will not maintain a centralized standard.]</li> <li>• Several endpoints are variable so the redundancy of endpoints is good.</li> </ul>
Strengths (within the context of the proposed battery)	<ul style="list-style-type: none"> <li>• Intact mammalian <i>in vivo</i> system and thus addresses ADME concerns.</li> <li>• Apical assay covering several modes of interaction, including ones not covered elsewhere</li> <li>• Redundant endpoints, maximizing chance for detection while minimizing false negatives</li> <li>• Covers pubertal period of development</li> <li>• Well-established relationship between endpoints and endocrine system</li> <li>• Endpoints easy to measure</li> </ul>
Limitations (within the context of the proposed battery)	<ul style="list-style-type: none"> <li>• Variability of hormone measurements, particularly testosterone</li> <li>• Relatively long duration</li> <li>• Although a toxic negative chemical has not been identified, several chemicals positive for one of the MOAs have been found to be negative for the other MOAs evaluated in this assay.</li> </ul>

## Appendix A10 – Steroidogenesis (H295R)

Steroidogenesis (H295R)	
Purpose	Provides an <i>in vitro</i> cell-based assay to detect chemicals that affect the synthesis of the sex steroid hormones.
Design	H295R cells are incubated with 7 concentrations of test chemical in triplicate overnight at 37°C along with 2 concentrations of prochloraz and forskolin as positive controls.
Endpoints	17 $\beta$ - estradiol and testosterone content of the supernatant are analyzed using appropriate steroid hormone assays. Cell viability is measured by live/dead assay.
Interpretation	Final guidance for data interpretation will be provided in the integrated summary report. Currently fold induction or inhibition is the basis for expressing the outcome of the assay. The criteria for differentiating positive and negative outcomes will likely be minimum fold change, but may be statistically significant difference between controls levels
Main peer review comments	Peer review of this assay is expected in the spring of 2008.
Strengths (within the context of the proposed battery)	<ul style="list-style-type: none"> <li>• Only <i>in vitro</i> assay that can evaluate effects on the entire steroidogenesis pathway—cells have all of the enzymes necessary for steroidogenesis.</li> <li>• Rapid and inexpensive</li> <li>• Detects chemicals that inhibit and induce steroidogenesis</li> <li>• Response is two dimensional (effective concentration and magnitude of response) and can distinguish among strong, moderate and weak inducers and inhibitors</li> <li>• Cells readily available from the ATCC</li> </ul>

Steroidogenesis (H295R)	
Limitations (within the context of the proposed battery)	<ul style="list-style-type: none"><li>• Limited metabolism</li></ul>

## Appendix A11 – Uterotrophic

Uterotrophic	
Purpose	To detect estrogenic chemicals through a simple in vivo assay .
Design	There are two versions of the uterotrophic assay an immature version and an ovariectomized adult version. In both versions, two concentrations of test substance are administered orally or sc to ovariectomized or immature female rats for a minimum of three consecutive days. Estrogenic substances cause a uterotrophic response that is due to the imbibition of water and the growth of cells.
Endpoints	Uterine weight is measured and compared with controls.
Interpretation	A statistical increase in uterine weight compared with controls is a positive result.
Main peer review comments	<ul style="list-style-type: none"> <li>• The rat uterus is biologically relevant for detecting estrogenic effects</li> <li>• Validation program was inadequate in several aspects <ul style="list-style-type: none"> <li>◦ Inadequate number of negative substances tested</li> <li>◦ Phytoestrogen levels need to be addressed</li> <li>◦ Program seriously flawed by not conforming to ICCVAM/ECVAM</li> </ul> </li> </ul> <p>[These comments were responded to by OECD.]</p> <ul style="list-style-type: none"> <li>• The protocol needs additional refinement <ul style="list-style-type: none"> <li>◦ Dose setting procedure needs to be clarified</li> <li>◦ Definition of what constitutes a positive result</li> </ul> </li> </ul> <p>[These were clarified in the OECD Test Guideline]</p> <ul style="list-style-type: none"> <li>• The antiestrogen procedure cannot be claimed to be validated because only one strong chemical was tested. [OECD agreed with this point. Finding pure antiestrogens for validation of this MOA is a real problem.]</li> </ul>
Strengths (within the context of the proposed battery)	<ul style="list-style-type: none"> <li>• Relatively rapid screen that is quite specific to estrogenic effects (see limitations)</li> <li>• In vivo procedure incorporates metabolism; thus, it can detect chemicals that need activation.</li> <li>• Oral administration will model a primary exposure route and incorporates ADME</li> </ul>



Uterotrophic	
Limitations (within the context of the proposed battery)	<ul style="list-style-type: none"><li>• Uterotrophic response is not due exclusively to estrogenic chemicals, so a uterotrophic response should be confirmed by corroborating information such as ER binding or transcriptional activation.</li></ul>

## Appendix B - General Principles for Evaluating the Results from a Tier 1 Screening Battery According to EDSTAC

The EDSTAC (EDSTAC, 1998) has provided guidance on the use of the “weight-of-evidence” approach for interpretation of the results for a Tier- battery. The “weight-of-evidence” approach makes explicit the assumption that results of some assays, in some taxa, at some level of severity, are intrinsically “worth” more than others and should, therefore, carry more weight in decisions following Tier 1 screening. There are several specific criteria to be met by the decision process assuming appropriate dose and route of exposure as recommended by EDSTAC. Although EPA is in the process of developing guidelines for interpretation of the results of the Tier 1 screening battery, specific principles recommended by EDSTAC are listed below, which are provided for information only since some of the principles may no longer be relevant.

1. If functionally equivalent information is available (e.g., from the sorting and prioritization phase), it may be appropriate that only those Tier 1 screening assays which evaluate the endocrine activity of concern (based on prior information) of a chemical substance or mixture would be performed (i.e., only a subset of assays would be run. Similarly, the results of the Tier 1 screening assays may require that only a subset of the Tier 2 test be conducted.
2. If all assays are performed and all assays are negative, then the chemical substance or mixture does not have endocrine activity that can affect the EAT hormonal pathways at this time.
3. *In vitro* assays cannot and shall not be “gatekeepers”, they cannot constitute a “decision node”; they are useful as information for possible mechanisms or sites of action but not as “yes/no” determinants to do or not to do Tier 1 *in vivo* assays or proceed to Tier 2 testing because:
  - a) *in vitro* assays mediated by receptor binding evaluate only one of many possible sites and modes of action.
  - b) negative results may mean relatively little due to limitations of the assay. For example, lack of metabolic capability, solubility, etc. (i.e., false negatives)
  - c) positive results may be false positives.
4. Results from *in vivo* assays have more weight than results from *in vitro* assays; *in vitro* assays since:
  - a) *in vitro* assays will generate false negatives as well as false positives based on differences in access to the target tissue, metabolism etc., relative to *in vivo* assays;
  - b) *in vivo* results are considered to be more relevant.
5. Results from *in vitro* assays that assess endocrine activity with and without metabolic activation are worth more than results from *in vitro* assays without metabolic activation since the former can assess the activity of metabolites

generated within the culture if the correct metabolic activation is used (e.g., rat liver S9) and the latter can only assay the parent compound.

6. Results from apical *in vivo* assays are worth more than the results from specific *in vivo* assays since they indirectly assay many more sites of action to get to the same endpoint (e.g., uterotrophic assay in ovariectomized adult females (specific assay; chemical substances act at level of the uterus) versus in intact immature females (apical assay; chemical substances can act at level of the HPG or HPT axes and/or uterus). A positive specific assay provides mechanistic information but other mechanisms of action may also be present and go undetected; a negative specific assay is less informative.
7. Biologically plausible results are worth more than biological implausible results (obviously dependent on the state of current scientific knowledge).
8. Statistical significance is a useful tool but must be interpreted within the context of biological significance. For example, an observed association which does not achieve statistical significance, but which is consistent with results from related assays suggesting a common mechanism of action, might be interpreted as biologically significant. This means the use of any particular criterion such as  $P = 0.05$  should be carefully considered and there may be no hard and fast rule for weighing by statistical significance.
9. A consistent pattern of positive or negative results in various related assays is worth more than a single isolated positive or negative result. For example, positive results in the binding to ER transcriptional activation *in vitro* and positive results in the apical or specific uterotrophic assay *in vivo* are worth more than a positive result for receptor binding or transcriptional activation but no uterotrophic response.
10. The decision which will emerge from Tier 1 is:
  - a) The chemical substance does not require further testing for EAT activity at this time.
  - b) The chemical substance should be tested further for EAT activity at this time and proceed to Tier 2.

## **Appendix C: Fact Sheets for Assays Not Included in the Proposed EDSP Tier 1 Screening Battery**

The fact sheets for individual assays are arranged in alphabetical order as indicated:

- C1 – Adult male
- C2 – Aromatase (Placental)
- C3 – *In utero* through lactational
- C4 – Steroidogenesis (Sliced testes)

## Appendix C1 – Adult Male

Adult Male	
Purpose	To detect interactions with the endocrine system, especially chemicals that may be AR agonists/antagonists, steroid biosynthesis inhibitors, and gonadotropin and thyroid modulators either directly or indirectly through intact HPG or HPT axes. Versatility of the assay may also permit detection of potential ER agonists/antagonists, progesterone agonists/antagonists and prolactin modulators through neuroendocrine pathways.
Design	Adult male rats (~10 wks) are treated daily for 15 days by oral gavage at three dose levels (low, intermediate, high) plus a vehicle-control (0.25% methylcellulose) at a dose volume of 5 ml/kg. Dose concentrations (mg/kg/d) are adjusted daily based on body weight for all animals in all groups (n=15/group). On Day 15, final body weight is determined and animals are anesthetized and decapitated. Target organs and blood are collected within a 3 hour window during mid-morning.
Endpoints	<p>Clinical observations, food consumption and body weight, daily</p> <p>Organ weights</p> <ul style="list-style-type: none"> <li>Liver</li> <li>Testes</li> <li>Epididymides</li> <li>Prostate (total)</li> <li>Seminal vesicles with coagulating gland containing fluid</li> <li>Accessory sex glands (prostate plus seminal vesicles with coagulating gland)</li> <li>Thyroid</li> </ul> <p>Hormone concentrations (assays may be run based on nature of test chemical and organ weight and histology results)</p> <ul style="list-style-type: none"> <li>Testosterone</li> <li>Dihydrotestosterone</li> <li>Estradiol</li> <li>Follicle-Stimulating Hormone</li> <li>Luteinizing Hormone</li> <li>Prolactin</li> <li>Thyroid-Stimulating Hormone</li> <li>Thyroxine</li> <li>Triiodothyronine</li> </ul> <p>Histology</p> <ul style="list-style-type: none"> <li>Testes</li> <li>Epididymides</li> <li>Thyroid</li> </ul>

Adult Male	
Interpretation	<p>Final body weight, organ weights (absolute and relative to final body weight) and hormone concentrations in the treated groups are statistically compared to those in the control group. A trend analysis is also done to determine the dose-response relationship for organ weights and hormone concentrations. Relevant historical control data may be used to further confidence in the performance of the bioassay results for organ weights and hormone concentrations in the vehicle-control group. Determination of whether the results in the treatment groups are endocrine-related first involves whether the final body weight decrement relative to the control group is within the limits of interpretation of an endocrine-related effect rather than an acute toxic effect secondary to an extreme decrease in final body weight during treatment. Primary effects associated with organ weights and histomorphology are assessed statistically (organ weights) and biologically (organ weights and histomorphology) to determine if there are endocrine-related responses due to treatment. Statistical and biological evaluations of hormone concentrations are used secondarily to support primary effects; they are not used alone within the bioassay. A weight-of-evidence approach with biological plausibility is considered among the multiple endpoints within the bioassay to conclude whether or not the bioassay has detected an interaction between the test chemical and the E, A or T hormonal pathways.</p>
Main peer review comments	<ul style="list-style-type: none"> <li>• Historical control data outdated and limited to one industrial laboratory</li> <li>• Within- and between-laboratory CVs for all endpoints were properly analyzed and relatively consistent for organ weight endpoints but highly variable for the hormonal endpoints</li> <li>• Running the full suite of hormonal assays is not justified</li> <li>• Hormonal assays were not standardized</li> <li>• Pre-validation studies involved extensive testing of a wide range of chemicals but inter-laboratory study involved too few chemicals covering a limited number of MOA (<i>i.e.</i>, anti-androgen and thyroid toxicant)</li> <li>• Negative or ambiguous results with relatively weak estrogenic and androgenic test compounds during pre-validation and inter-laboratory study</li> </ul>



Adult Male	
Strengths	<ul style="list-style-type: none"><li>• Intact mammalian <i>in vivo</i> system taking into account ADME</li><li>• Flexible to cover multiple MOA, receptor and non-receptor mediated</li><li>• Dose setting is readily achieved without confounding factors (e.g., growth and maturation of the HPG and HPT axes)</li><li>• Short dosing duration of 2 weeks at 3 dose levels</li><li>• Simple design comparable to sub-acute toxicology study</li><li>• Multiple and complimentary male reproductive organs as primary endpoints with secondary hormonal endpoints; therefore, minimizing false negatives</li></ul>
Limitations or weaknesses	<ul style="list-style-type: none"><li>• Insensitive to relatively weak estrogenic and androgenic compounds</li><li>• Extensive variability in hormone assay measurements, especially those relevant to androgen-dependent organs (LH, testosterone, DHT).</li><li>• Hormonal results considered secondary that may or may not support primary organ weight and histological results.</li></ul>

## Appendix C2 – Aromatase (Placental)

Aromatase (Placental)	
Purpose	The aromatase assay detects chemicals that inhibit aromatase activity. Aromatase is the enzyme that metabolizes androgens such as testosterone to estrogens
Design	Androstenedione and [ $1\beta$ - $^3\text{H}$ ]-androstenedione (ASDN) serve as substrate for placental microsomal aromatase. Full activity control (ASDN in medium, no inhibitor), background activity control (no NADPH), positive control (4-hydroxyandrostenedione at eight concentrations) and test chemical (8 concentrations) are run in the reaction for 15 minutes, and the reaction products produced are measured and plotted as percent enzyme activity (inhibition curve) through use of a non-linear regression program.
Endpoints	The formation of $^3\text{H}_2\text{O}$ , one of the co-reaction products along with estrone, is measured by liquid scintillation counter.
Interpretation	Chemicals that reduce enzyme activity levels by 50% or more (as determined by the inhibition curve derived from a four parameter non-linear regression model) are considered to be inhibitors of aromatase. Chemicals that fit the model and allow 50-75% activity, <i>i.e.</i> , reduce activity by 25-50%, are considered equivocal. Chemicals that do not fit the model or fit the model and reduce activity by less than 25% (allow activity levels greater than 75%) are considered to be non-inhibitors of aromatase.
Main peer review comments	<ul style="list-style-type: none"> <li>• Comments supported the use of the assay for the intended purpose, the clarity of the protocol, the data interpretation procedure, and performance criteria; however, one reviewer noted that <math>K_i</math> determination would be superior to <math>\text{IC}_{50}</math>.</li> <li>• The chemicals and analytical methods used in the validation of the assay were appropriately chosen.</li> <li>• The protocol could be further optimized for small volumes resulting in less cost and waste and greater convenience.</li> <li>• There are better assays that could have been selected such as cell-based assays which would be advantageous in that they would detect both induction and inhibition.</li> </ul>

Aromatase (Placental)	
Strengths	<ul style="list-style-type: none"><li>• Highly specific to inhibition of aromatase activity providing mechanistic information.</li><li>• More sensitive than typical in vivo assays</li><li>• Rapid</li><li>• Inexpensive</li><li>• Provides useful information for the interpretation of in vivo assays.</li></ul>
Limitations	<ul style="list-style-type: none"><li>• Cannot detect chemicals that induce aromatase activity.</li><li>• False positives could result from chemicals that denature the enzyme.</li><li>• Limited/no ability to metabolize xenobiotics</li><li>• Procuring fresh placenta and preparing microsomes is an added burden compared with the recombinant assay.</li><li>• Using human tissue raises concerns about pathogens</li></ul>

Appendix C2 – *In Utero* Through Lactational

<i>In Utero</i> Through Lactational	
Purpose	Determine developmental and reproductive consequences of exposure to chemicals that affect the estrogen, androgen and thyroid hormonal systems during development <i>in utero</i> , during lactation and after weaning until puberty.
Design	<p>Protocol C:</p> <p>F0 females gavaged orally from gd 6 - pnd 21 (weaning).  F1 offspring divided into two female and two male cohorts after weaning: 1) immediate necropsy of males, 2) Uterotrophic cohort: subcutaneous injection (1/litter) from pnd 22-24, 3) Pubertal female cohort: oral gavage (4/litter, 2 dosed/2 not dosed) from pnd 21 – 42, and 4) Pubertal male cohort: oral gavage (4/litter, 2 dosed and 2 not dosed) from pnd 21 – 70.</p> <p>F0 females and F1 offspring dosed at 3 dose levels plus a vehicle control.</p>
Endpoints	<p>Maternal:</p> <p>During in-life, body weights, feed consumption and clinical observations are taken.  At necropsy (pnd 21), final body, liver and thyroid weights, count of uterine implantation sites, serum T4/TSH and thyroid histology are collected and analyzed.</p> <p>Offspring:</p> <p>Body weights on pnd 0, 4, 7, 14, 18, 21, 22, 24, (necropsy uterotrophic), 42 (necropsy female pubertal) or 70 (necropsy male pubertal).  Anogenital distance on pnd 21 and necropsy.  Uterotrophic (dosed from pnd 22-24): Ovarian and uterine weight, uterine histology and serum T4 and TSH concentrations.  Female pubertal (dosed from pnd 22 – 42): acquisition of vaginal patency, weights of reproductive organs and thyroid, histology on ovaries, uterus and thyroid and serum T4 and TSH concentrations.  Male pubertal (dosed from pnd 22 – 70): acquisition of preputial separation, retained nipples and areolae, weights of reproductive organs and thyroid, histology on testis, epididymis, and thyroid, serum T4 and TSH concentrations.</p>

<i>In Utero</i> Through Lactational	
Interpretation	No data interpretation criteria were adopted by the EPA.
Main peer review comments	<p>The EPA presented the alternative <i>in utero</i> through lactational (IUL) rat screening assay to the FIFRA SAP in February 2007 to consider whether the IUL assay, as represented by Protocol C and tested with methoxychlor, was suitable as an alternative Tier 1 screening assay and whether the assay validation process should continue using Protocol C or some other protocol (e.g., A or B).</p> <p>In general, the SAP considered Protocol C too complex for a Tier 1 screen and not in accord with the EDSTAC criteria of a Tier 1 assay. Although it was felt that Protocol B or a modification of it could be validated as a simpler screen, there was concern that none of the proposed protocols (A, B, C) could be standardized and validated to find utility as a routine Tier 1 screen.</p>
Strengths	<p>Evaluates the effects of chemical exposure <i>in utero</i>, during lactation and after weaning until puberty.</p> <p>Pre- and postnatal endpoints are appropriate and sensitive to endocrine disrupting activates.</p>
Limitations	<ul style="list-style-type: none"><li>• Complex</li><li>• Long</li><li>• Costly</li></ul>

## Appendix C4 – Steroidogenesis (Sliced testes)

Steroidogenesis (Sliced Testes)	
Purpose	Provides an in vitro assay to detect chemicals that affect the synthesis of the sex steroid hormones.
Design	Fresh testes are sliced into approximately 50-100 mg fragments. Each fragment is incubated in 9 mm test tubes for 4 hours. After each hour, the supernatant is collected and fresh media with or without test chemical (as appropriate) was added. A composite sample was prepared and analyzed at the end of the four hour incubation period. Aminoglutethimide serves as the positive control and 2,4-dinitrophenol served as the cytotoxicity control.
Endpoints	Testosterone is measured by radioimmunoassay (RIA). Cell viability was determined by the LDH assay.
Interpretation	Fold or percent inhibition is the basis for expressing the outcome of the assay. No data interpretation criteria were adopted by the EPA.
Main peer review comments	At the recommendation of the EDMVAC, EPA terminated efforts on this assay prior to the interlaboratory validation phase.
Strengths	<ul style="list-style-type: none"> <li>• Only assay that is specific for the entire steroidogenesis pathway.</li> <li>• Rapid and inexpensive</li> <li>• Detects chemicals that inhibit steroidogenesis</li> <li>• Results appear to correlate well with other in vitro and in vivo data</li> </ul>

Steroidogenesis (Sliced Testes)	
Limitations	<ul style="list-style-type: none"><li>• Cannot identify chemicals that <b>induce</b> steroidogenesis.</li><li>• Cannot detect chemicals that are toxic specifically to the Leydig cell.</li><li>• Limited metabolism</li></ul>



**Appendix D - A Comparison of the Screening Assays Recommended by EDSTAC and those Proposed by EPA for the EDSP Tier 1 Battery.**

<b>Battery Recommended by EDSTAC</b>	<b>Proposed by EPA</b>
<b><i>In vitro</i></b>	<b><i>In vitro</i></b>
Estrogen receptor (ER) binding – rat uterus	<sup>3</sup> Estrogen receptor (ER) binding – rat uterus
Estrogen receptor transcriptional activation	Estrogen receptor $\alpha$ (hER $\alpha$ ) transcriptional activation - Human cell line (HeLa-9903)
Androgen receptor (AR) binding – rat prostate	Androgen receptor (AR) binding – rat prostate
Androgen receptor (AR) transcriptional activation	<sup>4</sup> ---
<sup>1</sup> Steroidogenesis – minced/sliced rat testes	<sup>2,3</sup> Steroidogenesis – Human cell line (H295R)
<b><i>In vivo</i></b>	<b><i>In vivo</i></b>
Uterotrophic (rat)	Uterotrophic (rat)
Hershberger (rat)	Hershberger (rat)
Pubertal female (rat)	Pubertal female (rat)
Amphibian metamorphosis (frog)	Amphibian metamorphosis (frog)
Fish gonadal recrudescence	<sup>2</sup> Fish short-term reproduction
<b>Alternative Assays Recommended by EDSTAC</b>	
<sup>1</sup> Aromatase – Human placental	<sup>2</sup> Aromatase – Human recombinant
Pubertal male (rat)	Pubertal male (rat)
<sup>1</sup> Adult male (rat)	---
<sup>1</sup> <i>In utero</i> through lactational (rat)	---

<sup>1</sup>Basis for not including these assays in the proposed battery is discussed in Section 4.2.

<sup>2</sup>Assays modified from original assays suggested recommended by EDSTAC.

<sup>3</sup>ER and H295R have not completed peer review yet and inclusion in the battery is contingent on successful review of these assays.

<sup>4</sup>No AR transcriptional activation assay is validated at this time.