Executive Summary

Endocrine disrupting substances are defined as chemicals that interfere with the normal function of hormones, either during development or during the life of an animal, resulting in abnormal development, growth, or reproduction (Ankley et al. 1998; Combes 2000; EPA 1998; Gray et al. 1998). Concern regarding these substances arises from observations of reproductive and developmental abnormalities in animal populations exposed to high levels of certain persistent pollutants in the environment. In addition, human health consequences including increases in the incidence of birth defects, cancers in hormonally-receptive tissues, and decreased fertility have been attributed to exposure of humans to endocrine disruptors. In response to these concerns, Congress directed the U.S. Environmental Protection Agency (EPA) in 1996 to validate and implement a screening and testing program to evaluate the potential of these substances to cause hormone-related health effects (Public Law [P.L.] 104-170). Based on advice from the EPA Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC), EPA proposed the Endocrine Disruptor Screening Program (EDSP) (EPA 1998). The EDSP consists of a Tier 1 screening battery of *in vitro* and *in vivo* assays that is designed to identify substances capable of interacting with the endocrine system. Tier 2 of the EDSP is a battery of *in vivo* assays that provides detailed information on concentration response relationships and specific abnormal effects. Based on a weight-of-evidence evaluation of the results from the Tier 1 screening battery, Tier 2 *in vivo* tests are conducted. Included among the proposed Tier 1 *in vitro* assays are estrogen receptor (ER) and androgen receptor (AR) binding and transcriptional activation (TA) assays.

In April 2000, EPA asked the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) to evaluate the validation status of *in vitro* ER and AR binding and TA assays. ICCVAM, which is charged by law (P.L. 106-545) to evaluate the scientific validity of new, revised, and alternative test methods proposed for specific regulatory uses, agreed to evaluate the assays based on their potential interagency applicability and public health significance. Because a large number of *in vitro* methods were known to exist, it was expected that at least some of these would have been adequately validated and could be rapidly included in the EDSP following a review of existing data and verification of their validity. The EPA also asked for the development of minimum performance standards that could be used to define acceptable *in vitro* ER and AR binding and TA assays. It was envisioned that these standards would be based on the performance of validated *in vitro* ER- and AR-based assays.

The National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) subsequently compiled all available relevant data and information on the *in vitro* methods of interest. A comprehensive review of these data determined that there were no adequately validated *in vitro* ER- or AR-based assays, and therefore, no assays could serve as the basis for establishing minimum performance standards. It was also discovered that there was little consistency among available protocols, and that no assay protocol was adequately detailed and standardized. Minimum procedural standards were therefore proposed that should be incorporated in the standardized protocols for each of the four types of assays. These minimum procedural standards include critical elements such as

dose selection criteria, number of replicates per test, appropriate positive and negative controls, and criteria for an acceptable test.

Four draft Background Review Documents (BRDs) were developed and organized according to published guidelines for submission of test methods to ICCVAM (ICCVAM 1999). Each BRD (NIEHS 2002a, 2002b, 2002c, 2002d) contained:

- a description of the types of test methods used to measure the endpoints of interest and the available data substantiating their scientific validity;
- published and submitted data on substances tested in the test methods being considered;
- an evaluation of the comparative reliability and performance of the test methods being considered;
- specific protocols for test methods provided by interested scientists;
- a prioritized list of test methods recommended for validation;
- proposed minimum procedural standards for the types of test methods being considered;
- a list of substances proposed for future validation studies.

The final *in vitro* ER binding BRD summarized and evaluated data on 638 different substances tested at least once in one or more of 14 different test methods. The *in vitro* ER TA BRD summarized and evaluated data on 698 different substances tested at least once in one or more of 95 different test methods. The *in vitro* AR binding BRD summarized and evaluated data on 108 different substances tested at least once in one or more of 11 different test methods. The *in vitro* AR TA BRD summarized and evaluated data on 145 different substances tested at least once in one or more of 18 different test methods.

ICCVAM asked its Endocrine Disruptor Working Group (EDWG) to assist NICEATM with the technical evaluation of the four types of *in vitro* endocrine disruptor assays. The EDWG, which is comprised of knowledgeable scientists from participating ICCVAM agencies, was charged with:

- identifying and recommending scientists for the Expert Panel;
- reviewing the four draft BRDs for completeness and accuracy;
- developing questions for the Expert Panel to consider during their deliberations;
- developing draft ICCVAM recommendations based on the conclusions and recommendations of the Expert Panel.

An Expert Panel consisting of 24 scientists was selected based on advice from the EDWG. The expertise of the members included relevant areas such as reproductive toxicology, androgen and/or estrogen receptor binding and TA assays, validation of alternative *in vitro* methods, ecotoxicology, and biostatistics. The Expert Panel members were from the United States, the United Kingdom, Canada, Japan, and Denmark, and included scientists from industry, academia, and government.

The Expert Panel was charged with reviewing the information and recommendations provided in the four draft BRDs, and developing conclusions and recommendations on the following:

- specific test methods that should undergo further evaluation in validation studies, and their relative priority for evaluation;
- the adequacy of the proposed minimum procedural standards;
- the adequacy of protocols for specific test methods recommended for validation;
- the adequacy and appropriateness of substances proposed for validation studies.

The Expert Panel met in public session on May 21-22, 2002, in Research Triangle Park, North Carolina. The Expert Panel presented the evaluations, conclusions, and recommendations for each of the four types of assays. Opportunities for public comment were provided during the meeting. After consideration of the public comments, the Expert Panel reached consensus on each of its recommendations. The Expert Panel's written evaluations and recommendations were consolidated into an independent report, which is included in this document as **Appendix A**.

Following the Expert Panel meeting, the four draft BRDs were revised to address corrections and omissions noted by the Expert Panel and published as final versions, which are available on the ICCVAM/NICEATM website *http://iccvam.niehs.nih.gov/methods/endocrine.htm*. Based on the recommendations of the Expert Panel, the EDWG, with the assistance of NICEATM, developed draft minimum procedural standards and lists of proposed substances for validation of ER and AR binding and TA assays.

In October 2002, the final report of the Expert Panel and the EDWG's draft list of proposed substances were made available to the public for comment (67 *FR* 204: 64902-64903, October 22, 2002). Following their review of the public comments, the EDWG and ICCVAM finalized their recommendations on minimum procedural standards, test methods for future validation, and substances that should be used to standardize and validate the test methods. This information is provided in this report. The final Expert Panel report, public comments, and other relevant documents are appended to this report, and are available also on the ICCVAM/NICEATM website *http://iccvam.niehs.nih.gov/methods/endocrine.htm*.

Recommendations

ICCVAM concurs with the recommendations of the Expert Panel with regard to the four different types of assays. The major recommendations, organized by assay type, are:

***In Vitro* ER Binding Assays**

- Recombinant rat or human ERs (α and β subtypes) should be given the highest priority for further test method standardization, prevalidation, and validation. Recombinant receptors are superior to crude cytosolic preparations because they can be prepared and distributed as standardized products with significantly less contamination. This will result in greater reproducibility and facilitate comparison of results across laboratories. To screen for possible ecological effects, recombinant receptors from wildlife are considered to be potentially more relevant and their use should be evaluated.
- Although it would be advantageous to use nonradioactive methods such as fluorescent polarization to assess ER binding, this method has not been widely used and specialized equipment is required. However, once a test method using recombinant ER proteins has been validated, there should be an effort to optimize a fluorescence-based method to replace the use of radioactivity.
- *In vitro* ER binding assay protocols should be standardized to incorporate the recommended minimum procedural standards (see **Section 3.1**). Exceptions should be justified with scientific rationale. Following protocol standardization, prevalidation studies should be conducted to optimize a reproducible protocol. Once this has been achieved, validation studies to assess the reliability and comparative performance of the test method should be conducted.

- Proposed *in vitro* ER binding test methods should be evaluated in validation studies using, at a minimum, the 53 substances listed in **Section 3.2**. This list includes substances that cover a range of activities, from negative to weakly positive to strongly positive, with 40 (75%) positive and presumed positive and 13 (25%) negative and presumed negative substances. The list also represents a wide range of relevant chemical and product classes (see **Section 2.0**). Following validation studies using the 53 substances, ICCVAM recommends that data should be generated on the remainder of the substances in the list of 78. The additional data will aid in the assessment of the usefulness of an *in vitro* test battery for prioritizing substances for subsequent *in vivo* studies.

***In Vitro* ER TA Assays**

- A comparative study should be conducted to determine whether transiently or stably transfected cell lines are more appropriate for a routine test system. Transiently transfected systems generally have a higher level of responsiveness, while stably transfected cell lines have a lower level of responsiveness but are generally more amenable to high-throughput screening. Such a study should use cell lines with the same ER reporter gene constructs. A third cell line expressing an endogenous ER and transfected with the same reporter construct should be included in this study.
- *In vitro* ER TA assay protocols should be standardized to incorporate the recommended minimum procedural standards (see **Section 4.1)**. Exceptions should be justified with scientific rationale. Following protocol standardization, pre-validation studies should be conducted to optimize a reproducible protocol. Once this has been achieved, validation studies to assess the reliability and comparative performance of the protocol should be conducted.
- To facilitate the comparison of *in vitro* ER-based assays, the same minimum list of 53 substances (provided in **Section 4.2**) recommended for ER binding assays should be used in the validation of *in vitro* ER TA agonist and antagonist assays. For ER TA agonism and antagonism assays, 34 (64%) and 11 (21%) of the substances, respectively, are reported to be positive or presumed positive, and 19 (36%) and 42 (79%) of the substances, respectively, are presumed negative. Following validation studies using the 53 substances, ICCVAM recommends that data should be generated on the remainder of the substances included in the list of 78. The additional data will aid in the assessment of the usefulness of an *in vitro* test battery for prioritizing substances for subsequent *in vivo* studies.

***In Vitro* AR Binding Assays**

- A recombinant protein should be used as the source of the AR. Recombinant receptors are superior to crude cytosolic preparations because the recombinant protein can be standardized, which contributes to improved quality control and comparison of results across laboratories. Thus, the highest priority for future research and development efforts should be given to the development of a test method using a recombinant full-length AR protein. Patents on the AR protein have hindered development of this assay.
- *In vitro* AR binding assay protocols should be standardized to incorporate the recommended minimum proce-dural standards (see **Section 5.1**). Exceptions should be justified with scientific rationale. Following protocol standardization, prevalidation studies should be conducted to optimize a reproducible protocol. Once

this has been achieved, validation studies to assess the reliability and comparative performance of the protocol should be conducted.

- Proposed *in vitro* AR binding assays should be evaluated in validation studies using, at a minimum, the 44 substances listed in **Section 5.2**. This list consists of 33 (75%) positive and presumed positive substances and 11 (25%) presumed negative substances for AR binding. Following validation studies using the 44 substances, ICCVAM recommends that data should be generated on the remainder of the substances included in the list of 78. The additional data will aid in the assessment of the usefulness of an *in vitro* test battery for prioritizing substances for subsequent *in vivo* studies.

***In Vitro* AR TA Assays**

- None of the *in vitro* AR TA assays reviewed by the Expert Panel were considered optimal for assessing AR agonist and antagonist activities. The highest priority for future efforts should be a cell line containing an endogenous AR that is transduced with an adenovirus containing a reporter vector that shows high specificity for the AR. The chosen cell line should not respond to, or have minimal response levels for, the glucocorticoid and progesterone receptors. Because of patent restrictions, it may be necessary that a cell line with an endogenous AR be used for validation. Transduction of a reporter construct in a virus particle is more efficient and reproducible than transfection of a construct.
- *In vitro* AR TA assay protocols should be standardized to incorporate the recommended minimum procedural standards (see **Section 6.1)**. Exceptions should be justified with scientific rationale. Following protocol standardization, prevalidation studies should be conducted to optimize a reproducible protocol. Once this has been achieved, validation studies to assess the reliability and comparative performance of the protocol should be conducted.
- To facilitate *in vitro* AR-based test method comparisons, the same minimum list of 44 substances (provided in **Section 6.2**) recommended for *in vitro* AR binding assays should be used in the validation of *in vitro* AR TA agonist and antagonist assays. For AR TA agonism and antagonism assays, 20 (45%) and 20 (45%) of the substances, respectively, are reported to be positive and presumed positive, and 24 (55%) and 24 (55%) of the substances, respectively, are presumed negative. Following validation studies using the 44 substances, ICCVAM recommends that data should be generated on the remainder of the substances included in the list of 78. The additional data will aid in the assessment of the usefulness of an *in vitro* test battery for prioritizing substances for subsequent *in vivo* studies.

Other Recommendations

ICCVAM agrees with the Expert Panel that the development and validation of *in vitro* ER and AR binding and TA assays should emphasize the use of recombinant-derived proteins. Based on current knowledge and experience, it appears that continuing to use animal-derived ER or AR in *in vitro* endocrine disruptor test methods requires scientific justification. The advantages of using recombinant-derived receptors for binding test methods include:

- Standardized recombinant protein can be prepared and used by multiple laboratories, which will contribute to improved inter- and intra-laboratory reproducibility and an enhanced ability to compare results across laboratories.

- Recombinant-derived receptors avoid the disadvantages of animal-derived receptors, which include:
  - The receptors, particularly the ARs, are unstable in tissue extracts.
  - The cytosolic extracts contain many proteins, including other endogenous steroid receptors that can interfere with the performance of the assay.
  - Animals have to undergo surgery before isolation of the tissue of interest. For AR binding assays, males are castrated, and, for ER binding assays, females undergo an ovariectomy before removal of the requisite tissues and isolation of the respective receptors.
  - Animals need to be killed to obtain either the uterus (ER binding) or prostate (AR binding) glands.
- The inclusion of a metabolic activation system in *in vitro* ER and AR binding and TA assays is not recommended at this time, as the type of metabolic activation system developed will depend on which *in vitro* assays are selected. Available information on the metabolism of the validation substances should be compiled, including the degree to which metabolism is known to alter estrogenic and androgenic activity *in vivo*. Once the importance of metabolic activation in the ability of substances to disrupt endocrine function has been demonstrated, and valid *in vitro* ER and AR binding and TA assays have been identified, appropriate methods for including metabolic activation in the assays can be developed and validated.
- The current analyses for making statistical inferences with *in vitro* endocrine disruptor data require more detailed research and study. Appropriate prevalidation studies should be conducted to generate data necessary for biostatisticians to develop appropriate statistical methods for analyzing binding and TA agonist and antagonist assay data.
- Although these *in vitro* endocrine disruptor assays are proposed as components of a screening test battery where the results will be used in making weight-of-evidence decisions, the predictive value of these *in vitro* assays for estimating *in vivo* responses should be determined. To facilitate this determination, ICCVAM recommends that all 78 substances (see **Section 2.0**) should be evaluated in each *in vitro* assay. It is only through this effort that the performance of the *in vitro* test methods for predicting responses in animals can be evaluated and decisions made as to whether and how *in vitro* assays can reduce or replace animal use. Such data will also be needed to determine the usefulness of the *in vitro* battery for prioritizing substances for further testing.
- A centralized repository of the 78 substances with verified purity should be organized to facilitate future validation studies. The purpose of this repository is to provide a source of coded samples, of known purity, for validation studies. This approach would greatly enhance evaluation of the comparative reliability and performance of different versions of *in vitro* ER and AR binding and TA assays.
- Federal agencies are encouraged to support research and development of new technologies (e.g., genomics) that may provide more accurate assessments and/or advantages in terms of time and cost.