

**Department of Health and Human Services  
National Institutes of Health  
National Institute of Environmental Health Sciences  
Interagency Coordinating Committee on the Validation of Alternative Methods  
(ICCVAM) Special Emphasis Panel Meeting**

Summary minutes of the Expert Panel Meeting on the Evaluation of the Validation Status of *In Vitro* Test Methods for Detecting Endocrine Disruptors.

### **Introduction**

A public meeting of an independent Expert Panel was convened on May 21-22, 2002, at the Sheraton Imperial Hotel, Research Triangle Park, NC to review the current status of *in vitro* methods used to measure estrogen and androgen receptor binding and estrogen and androgen transcriptional activation assays. The meeting was organized by ICCVAM and the National Toxicology Program (NTP) Interagency Center for the Evaluation of Toxicological Methods (NICEATM) and sponsored by the National Institute of Environmental Health Sciences (NIEHS) and the NTP. A comprehensive report of the peer review panel is provided as an attachment to these minutes.

The following scientists served on the expert panel:

- George Daston, Ph.D., (Panel Chair), Research Fellow, Procter & Gamble Miami Valley Laboratories, Cincinnati, OH
- Nira Ben-Jonathan, Ph.D., Professor of Cell Biology, Neurobiology & Anatomy, University of Cincinnati Medical School, Cincinnati, OH
- Terry Brown, Ph.D., Professor, Johns Hopkins University, School of Public Health, Baltimore, MD
- Grantley Charles, Ph.D., Toxicology and Environmental Research & Consulting, Dow Chemical Co., Midland, MI
- Robert Combes, Ph.D., Professor, FRAME, Nottingham, United Kingdom
- Kevin Gaido, Ph.D., Scientist II, CIIT, Research Triangle Park, NC
- Thomas Gasiewicz, Ph.D., Professor, University of Rochester School of Medicine, Dept. of Environmental Medicine, Rochester, NY
- John P. Giesy, Ph.D., Professor, Michigan State University, East Lansing, MI
- John W. Harbell, Ph.D., Vice President and Chief Scientific Officer, Institute for In Vitro Sciences, Inc., Gaithersburg, MD
- Tohru Inoue, M.D., Ph.D., Director, Center for Biological Safety Research, National Institute of Health Sciences, Tokyo, Japan
- William R. Kelce, Ph.D., F.A.T.S., Senior Scientist, Pharmacia, Corp., Kalamazoo, MI
- Ellen M. Mihaich, Ph.D., D.A.B.T., Senior Environmental Toxicologist, Rhodia, Inc., Raleigh, NC
- Shyamal Peddada, Ph.D., National Institute of Environmental Health Sciences, Biostatistics Branch, Research Triangle Park, NC

- Walter Piegorsch, Ph.D., Professor, Dept. of Statistics, University of South Carolina, Columbia, SC
- Bernard Robaire, Ph.D., Professor, McGill University, Dept. of Pharmacology & Therapeutics, Montreal, Quebec, Canada
- Stephen Safe, Ph.D., Professor of Veterinary Physiology & Pharmacology, Texas A & M University, College Station, TX
- John Stegeman, Ph.D., Senior Scientist, Chair, Biology Dept., Woods Hole Oceanographic Institution, Woods Hole, MA
- Anne Marie Vinggaard, Ph.D., Senior Scientist, Institute of Food Safety & Toxicology, Danish Veterinary & Food Administration, Soborg, Denmark
- Tom Weise, Ph.D. Assistant Professor, Tulane University, New Orleans, LA (not present at the meeting)
- Elizabeth Wilson, Ph.D., Professor of Pediatrics, Biochemistry, & Biophysics, University of North Carolina, Chapel Hill, NC
- James L. Wittliff, Ph.D., F.A.C.B., Professor of Biochemistry & Molecular Biology, University of Louisville, Louisville, KY
- James D. Yager, Ph.D., Associate Dean, Professor of Toxicology/EHS, Johns Hopkins University, School of Public Health, Baltimore, MD
- Tim Zacharewski, Ph.D., Associate Professor, Michigan State University, Dept. of Biochemistry and Molecular Biology, East Lansing, MI

The following ICCVAM agency representatives were present:

- Dr. Karen Hamernik, (Endocrine Disruptor Working Group - EDWG) U.S. Environmental Protection Agency
- Dr. David Hattan, (EDWG) U.S. Food and Drug Administration
- Dr. Jerold Heindel, (EDWG) National Institute of Environmental Health Sciences
- Dr. Abigail Jacobs, (EDWG) U.S. Food and Drug Administration
- Dr. Leonard Schechtman, (EDWG) (ICCVAM Chair) U.S. Food and Drug Administration
- Dr. William Stokes, (EDWG) (Director, NICEATM) National Institute of Environmental Health Sciences, Executive Secretary

The following additional members of the ICCVAM Endocrine Disruptor Working Group (EDWG) were present:

- Dr. Paul Brown, U.S. Environmental Protection Agency
- Dr. Sally Perreault-Darney, U.S. Environmental Protection Agency
- Dr. Julius Thigpen, National Institute of Environmental Health Sciences

The following NICEATM Staff were present:

- Mr. Brad Blackard, ILS, Inc.
- Ms. Sue Brenzel, ILS, Inc.
- Ms. Loretta Frye, National Institute of Environmental Health Sciences
- Ms. Christina Inhof, ILS, Inc.

- Ms. Linda Litchfield, ILS, Inc.
- Ms. Debbie McCarley, National Institute of Environmental Health Sciences
- Mr. Steve Myers, ILS, Inc.
- Mr. Michael Paris, ILS, Inc.
- Dr. Barbara Shane, ILS, Inc.
- Dr. Judy Strickland, ILS, Inc.
- Dr. Ray Tice, ILS, Inc.
- Dr. Errol Zeiger, Zeiger Consulting/ILS, Inc.

The following members of the public were present:

- Ms. Gina Alvino, Humane Society of U.S.
- Dr. Naohiro Araki, Otsuka Pharmaceuticals
- Dr. Rick Becker, American Chemistry Council
- Dr. George Clark, Xenobiotic Detection Systems, Inc.
- Dr. Ralph Cooper, U.S. Environmental Protection Agency
- Dr. Paul Foster, U.S. Environmental Protection Agency
- Dr. L. Earl Gray, U.S. Environmental Protection Agency
- Dr. Susie Humphreys, U.S. Food and Drug Administration
- Mr. Jim Kariya, U.S. Environmental Protection Agency
- Mr. Robert Kavlock, U.S. Environmental Protection Agency
- Dr. Elena Klaymenova, CIIT
- Ms. Christy Lambright, U.S. Environmental Protection Agency
- Dr. Susan Laws, U.S. Environmental Protection Agency
- Dr. Mitsuru Lida, Otsuka Pharmaceuticals
- Dr. Po Yung Lu, Oak Ridge National Library/U.S. Dept. of Energy
- Mr. John McArdle, Alternatives Research & Development Foundation
- Dr. Kazuhiko Nishioka, Jetro
- Dr. Zafar Randawa, Otsuka Pharmaceuticals
- Dr. Madhampyanda Sar, CIIT
- Mr. Jim Stevens, Syngenta
- Ms. Kris Thayer, World Wildlife Foundation
- Mr. Gary Timm, U.S. Environmental Protection Agency
- Dr. Gail Tudor, University of North Carolina, Chapel Hill
- Ms. Catherine Willett, Phyionix Pharmaceuticals, Inc.
- Dr. Mary Wolfe, National Institute of Environmental Health Sciences
- Dr. Yoji Ikawa, Otsuka Pharmaceutical

The purpose of this meeting was to evaluate the validation status of *in vitro* test methods for detecting endocrine disruptors. The Expert Panel was asked to evaluate four background review documents (BRDs) prepared by National Toxicology Program Interagency Center for the Evaluation of Toxicological Methods (NICEATM).

The four BRDs reviewed and discussed were:

- Current Status of Test Methods for Detecting Endocrine Disruptors: *In Vitro* Estrogen Receptor (ER) Binding Assays

- Current Status of Test Methods for Detecting Endocrine Disruptors: *In Vitro* Estrogen Receptor Transcriptional Activation (ER TA) Assays
- Current Status of Test Methods for Detecting Endocrine Disruptors: *In Vitro* Androgen Receptor (AR) Binding Assays
- Current Status of Test Methods for Detecting Endocrine Disruptors: *In Vitro* Androgen Receptor Transcriptional Activation (AR TA) Assays.

## Introductions

Dr. George Daston, Panel Chair, called the meeting of the Expert Panel (Panel) to order at 9:00 a.m. and asked each attendee to state their name and affiliation. Dr. Daston stated that the public would be given the opportunity to speak at various times during the meeting. Each speaker from the public would be limited to seven (7) minutes, and anyone addressing the group should state their name for the benefit of the transcriptionist.

Dr. William Stokes, Executive Secretary for the Special Emphasis Panel, read the Statement of Conflict of Interest and explained policies and procedures regarding confidentiality and avoidance of conflict of interest, as follows: *“The members of this special emphasis panel serve as individual scientists and not as representatives of any organization. Each member is to exercise judgment as to whether a potential conflict of interest might exist relative to one or more of the topics being discussed due to his or her occupational affiliation, professional activity or financial interest. Should there be a potential conflict of interest, the member is to recuse his or herself from participating in the discussion of panel recommendations and/or decisions on the topic. You will be signing a conflict of interest certification which declares that during this panel meeting you did not participate in discussion of panel recommendations and/or decisions that involve a particular matter that could have a direct and predictable effect on: 1) Any organization, institution or university system in which a financial interest exists for yourself, spouse, parent, minor child or partner. 2) Any organization in which you, your spouse, parent, minor child or partner serves as an officer, director, trustee or employee or is otherwise similarly associated. 3) Any organization with which you, your spouse, parent, minor child or parent [sic] is negotiating or have any arrangements concerning prospective employment or other such associations. Panel members are asked to identify at the beginning of this meeting the nature of any such conflicts.”*

Dr. Elizabeth Wilson (Univ. of N. Carolina) responded that she was one of the scientists who was involved in cloning the androgen receptor (AR) in 1998. “This resulted in the awarding of a patent for the androgen receptor sequence to the University of North Carolina in Chapel Hill. This patent has also been awarded to the University of Chicago.” Dr. Wilson went on to say “And at the moment Ligand Pharmaceuticals holds an exclusive license on this patent. And because of that potential complication, I plan to offer my comments in terms of scientific expertise, but I will recuse myself from any decisions relating to protocols.”

Dr. Robert Combes (FRAME, UK) responded that although he did not have a financial conflict of interest, he did have a bias against the use of animal tests that would affect his

recommendations. He works for an organization that promotes non-animal methods “and therefore, if there is a recommendation, a choice between two assays that are scientifically equivalent, but one uses less animals or no animals at all, then I would promote the one that doesn't use animals or is more welfare conscious. So I don't think I need to recuse myself.”

### **Overview of the ICCVAM Test Method and Evaluation Process**

Dr. Stokes, (Director, NICEATM, NIEHS) provided a brief background on ICCVAM and NICEATM, and described the purpose of the meeting. He explained that this is an Expert Panel rather than a Peer Review Panel because no specific methods have been standardized and evaluated in validation studies.

ICCVAM was established as an *ad hoc* committee in 1994 in response to revisions in the 1993 NIH Revitalization Act (P.L. 103-43) that mandates that the NIEHS develop criteria for validation and regulatory acceptance of test methods, and develop a process to achieve regulatory acceptance of scientifically valid methods. The *ad hoc* committee issued its report in 1997, and the ICCVAM committee was formally established that year to implement P.L. 103-43 directives. In 2000, the ICCVAM Authorization Act (P.L. 106-545) established ICCVAM as a permanent committee.

Member agencies of ICCVAM include those involved in regulatory and research activities (CPSC; DOA; DOI; DOT; EPA; FDA; OSHA) and those involved in non-regulatory research (ATSDR; DOD; DOE; NCI; NIEHS; NIOSH; NLM; NIH, OD). NICEATM is located at NIEHS and is responsible for providing operational and technical support to ICCVAM.

The purposes of ICCVAM, as set forth in P.L. 106-545, are to:

- Increase efficiency and effectiveness of Federal agency test method review;
- Eliminate unnecessary duplicative efforts and share experiences between Federal regulatory agencies;
- Optimize utilization of scientific expertise outside the Federal Government;
- Ensure that new and revised test methods are validated to meet the needs of Federal agencies; and
- Reduce, refine, or replace the use of animals in testing, where feasible.

The duties and responsibilities of ICCVAM are to:

- Review and evaluate new, revised or alternative test methods;
- Facilitate interagency and international harmonization of test methods;
- Facilitate and provide guidance on test method development, validation criteria, and validation processes;
- Facilitate acceptance of scientifically valid test methods;
- Submit test method recommendations to Federal agencies;
- Consider petitions from the public for review and evaluation of validated test methods.

An ICCVAM Endocrine Disruptor Working Group (EDWG) comprised of government scientists that is co-chaired by Drs. David Hattan and Marilyn Wind, worked with NICEATM to develop the questions that were addressed to the panel. This group also recommended experts to serve on the panel and the members reviewed the BRDs for completeness. The EDWG will review the recommendations proposed by the Expert Panel and develop draft ICCVAM recommendations. ICCVAM recommendations and the Panel's report will be forwarded to the U.S. EPA and other Federal Agencies for consideration.

The background and history of the ICCVAM evaluation of *in vitro* estrogen receptor (ER) and androgen receptor (AR) methods were described. In 2000, the U.S. Environmental Protection Agency (U.S. EPA) requested that ICCVAM conduct an independent scientific peer review of *in vitro* ER and AR binding and transcriptional activation methods. In March 2001, ICCVAM published a *Federal Register* notice requesting data and information on these methods, and the nomination of experts that might serve on the peer-review Panel. At the same time, the four BRDs -- *in vitro* ER binding; *in vitro* ER transcriptional activation; *in vitro* AR binding; *in vitro* AR transcriptional activation were being prepared by NICEATM. During this review it was noted that there were no standardized test methods that had undergone formal validation studies. In April 2002, a *Federal Register* notice announced the dates of this meeting, the availability of the BRDs, and a request for public comments.

### **Charge to the Expert Panel and Organization of the Review**

Dr. Stokes explained the charge to the Expert Panel. The Panel was requested to review the BRDs and provide conclusions and recommendations on the following:

1. Assays that should be considered for further evaluation in validation studies and their relative priority.
2. Adequacy of the proposed minimum procedural standards for each of the four types of assays.
3. Adequacy of available test method protocols for assays recommended for validation studies.
4. Adequacy and appropriateness of the substances recommended for use in validation studies.

### **Overview of the biology of estrogen and androgen receptor binding and transcriptional activation assays**

Dr. Vickie Wilson, (Research Biologist, Reproductive Toxicology Division NHEERL, U.S. EPA) provided an overview of the biology of estrogen and androgen receptor binding and transcriptional activation assays. She discussed the concept of the binding of a hormone ligand with a receptor, resulting in a conformational change of the receptor. The ligand-receptor complex dimerizes and is then able to bind to a DNA response element resulting in the transcription or inhibition of the transcription of a gene. Ultimately a protein is produced that has some biological function in the organism. The receptor binding assays measure whether a test substance binds in place of the natural

hormone to the receptor. The transcriptional activation (TA) assays measure the next step in the pathway, namely, the transcription of a gene. A compound that initiates transcription, following receptor binding, is known as an agonist while one that blocks transcription after binding to the receptor is known as an antagonist.

### **Receptor Binding Assays**

Dr. Vickie Wilson described two general types of receptor binding assays. In the first type of assay, a saturation binding experiment is performed in which increasing amounts of radiolabeled hormone are added to the receptor until binding is saturated. This experiment allows for the determination of the equilibrium dissociation constant for a radioligand ( $K_d$ ), an indicator of the binding affinity of the radiolabeled hormone to the receptor, and for the maximum number of binding sites in the receptor preparation ( $B_{max}$ ). The second type of experiment is a competitive binding experiment in which increasing amounts of the test substance are added to the receptor in the presence of a single concentration of the radiolabeled reference hormone that is usually at or just below the  $K_d$  value. The components are allowed to come to equilibrium, the bound radioligand is separated from the free radioligand, and the quantity of radioligand bound receptor is determined at each concentration of test compound. An  $IC_{50}$ , which is the molar concentration of test substance that reduces the binding of the radiolabeled hormone to the receptor by 50%, can be calculated. The relative binding affinity (RBA) of the test substance, which is the ratio between the  $IC_{50}$  of the substance and the  $IC_{50}$  of the reference hormone, can then be calculated. To determine if the observed binding inhibition is truly due to competitive inhibition; assays can be performed to experimentally determine the affinity of the unlabelled substance ( $K_i$ ) to the receptor. Similar experiments can be performed for estrogen and androgen binding substances.

Dr. Vickie Wilson then described examples of the different kinds of binding curves that can be obtained when substances displace the hormone from the receptor. She stressed the problems associated with the testing of relatively high concentrations of the test substance and the situation that can occur when one obtains a precipitous decline in the binding over a very narrow range in concentration of the test substance. Under the latter conditions, it might be necessary to determine the  $K_i$  value experimentally. In this case, increasing concentrations of the test substance are added to several different concentrations of radiolabeled hormone to generate a number of lines in a double reciprocal plot. The pattern of the lines indicates the type of inhibition, for example, competitive inhibition versus non-competitive inhibition. The slopes of the lines are then plotted and the intercept of the line on the X-axis is the  $-K_i$ . Dr. Vickie Wilson then briefly discussed the various sources of the ER receptor and some of the general strengths and limitations of binding assays.

### **Transcriptional Activation Assays**

Dr. Vickie Wilson described four major categories of TA assays and the methods of transfecting the receptor (androgen or estrogen) and reporter gene (luciferase) into the cell lines. Transcriptional activation is quantified by the measurement of an androgen

responsive promoter attached to a reporter gene such as luciferase. For antagonism assays, Dr. Vickie Wilson emphasized the importance of first measuring the TA of the reference ligand to determine the linear part of the dose response curve and to establish the appropriate concentration of the reference ligand to use in subsequent assays. Then increasing concentrations of the test substance are added to the cells that are simultaneously being exposed to a specific concentration of the reference ligand. She emphasized the need for the use of media controls and performance of a cytotoxicity assay to determine that decreases in reporter gene activity are not due to cell death. Dr. Vickie Wilson discussed some of the strengths and limitations of the different types of transcriptional activation assays. In closing, she stated that data evaluation is critical and criteria need to be established to determine whether a compound is positive or negative.

### **Organization of the Panel review**

During the course of the meeting the Panel addressed the questions concerning the completeness and utility of the BRD and the performance of each particular assay. Four sub-groups of the Expert Panel were responsible for addressing the questions for each BRD, and drafting responses for consideration by the entire Panel.

Prior to the presentations and discussions by each of the four groups, Dr. Barbara Shane (NICEATM, ILS) provided a brief summary of the assays described in the BRD, the minimum procedural standards for an assay, and the substances suggested for validation of the assay(s).

Each group presented their draft responses for each of the questions assigned for their BRD. After each presentation, the Panel discussed the draft positions and offered additional comments and suggestions. The Chairman summarized the discussion for each question and sought consensus from the Panel on the topic. Public comments were accepted following the Panel's discussion of each BRD.

#### **A. Estrogen Receptor (ER) Binding Assays**

*Primary reviewers: G. Daston, Group Chair (Procter & Gamble); N. Ben-Jonathan (Univ. of Cincinnati); R. Combes (FRAME, UK); J. Harbell (Institute for In Vitro Sciences, Inc.); S. Safe (Texas A&M Univ.); J.L. Wittliff (Univ. of Louisville); W. Piegorsch (Univ. of S. Carolina).*

#### **Summary of the ER Binding Background Review Document**

Dr. Shane described the approach used to compile the BRDs. She stated that the on-line databases searched for publications on ER binding were Medline, Cancerlit, Toxline, Agricola, NIOSHTIC, Embase, CABA, Biosis, and Life Sciences. The key words *screen, tests, batteries, bind, ligand, estrogen, and receptor* were used in the search. This yielded 260 records of which 74 contained relevant information for inclusion in the BRD. The data abstracted from all records included the assay description, substance name, CASRN, and citation. Where available, the substance purity,  $K_i$  ( $\mu\text{M}$ ),  $\text{IC}_{50}$  ( $\mu\text{M}$ ),



standard deviation of IC<sub>50</sub> (μM), relative binding affinity (RBA), and highest dose tested for negative data (μM), were included in the BRD. If the RBA was not provided in the record, it was calculated from the available information in the report.

The database contained information on 14 assays with data for 635 unique chemicals. Of these chemicals, 235 (37%) were tested in 2 or more assays, and 51 (8%) were tested in 7 or more assays. The chemicals were assigned to chemical and product classes; 17 chemical classes and 7 product classes each had at least 10 entries. The most frequent chemical class was the polychlorinated biphenyls; the most frequent product class was pharmaceuticals.

The 14 assays included uterine cytosol from the mouse, rabbit, and rat; MCF-7 cell cytosol; intact MCF-7 cells; purified human (h) and rat (r) receptors, hER<sub>α</sub>, hER<sub>β</sub>, or purified hER<sub>α</sub> using a fluorescent polarization assay (hER<sub>α</sub>+FP), rER<sub>α</sub>, and rER<sub>β</sub>; and glutathione (GST) constructs containing the "def" (binding domain) domains of the receptor from anole chicken, human, mouse, and rainbow trout. All assays used radioactivity to measure binding except the fluorescent polarization (hER<sub>α</sub>+FP) assay, which used a fluorescently labeled estrogen.

Comparative performance analyses were performed following log transformation of positive RBA values of substances tested in two or more assays. The data was analyzed quantitatively using two- and three-way ANOVA and qualitatively for relative sensitivity by comparing the different RBA values of each substance in each assay to that of the substance in the rat uterine cytosol (RUC) assay. It was concluded that the numbers of substances tested in multiple assays was too limited for an adequate comparison to be conducted. Comparative inter-laboratory reproducibility analyses concluded that there was little variation in RBA values of the same substance tested in different labs and in different assays. However, this conclusion was based on data obtained with potent substances only.

The three assays with the most promise use purified human ER's. Either the human ER (hER<sub>α</sub>) or human ER (hER<sub>β</sub>) proteins with radiolabelled 17β-estradiol or the ER protein with a fluorescently-labeled estrogen (hER<sub>α</sub>-FP) are the most appropriate assays. The RUC assay could be used for comparison purposes. These assays were recommended in the BRD because of their greater sensitivity, direct relevance to humans, and their elimination of the use of animals, and in the case of the fluorescent polarization assay, the elimination of radioisotopes. A revised U.S. EPA RUC protocol incorporating minimum procedural standards was also proposed in the BRD. For future validation studies, 35 substances were suggested for testing.

### ***1. Recommendations and Prioritization of Assays for Validation Studies***

The Panel agreed that assays using recombinant human or rat estrogen receptor alpha or beta should have the highest priority for validation and standardization. Recombinant receptors from other species would be more relevant for screening for possible effects in

wildlife. A standardized preparation of the receptor is essential for quality control and to enable comparison across laboratories.

In general, the ER $\alpha$  and ER $\beta$  proteins produce similar results, and the differences between them are primarily quantitative. There are a few examples of substances that bind to only one of the ERs. The Panel recommended that once a basic assay using a recombinant ER is developed, either ER $\alpha$  or ER $\beta$ , other types of ERs could be substituted in the protocol. However, there is a preference for the use of recombinant hER.

Despite the suggestion in the BRD that the rat uterine cytosol (RUC) assay be used as a “benchmark” assay, the Panel identified a number of disadvantages. These include its bias towards ER $\alpha$ , animal welfare considerations, the difficulties of standardizing a cytosolic preparation from an animal due to the effect of age, weight, strain, etc., of the animals and the use of many animals even though this is an *in vitro* assay. Despite these drawbacks, there is much information using RUC for ER testing, and therefore the Panel recommended that this test be considered for comparative purposes only. In addition the Panel recommended that the minimum performance criteria that have been developed for this assay (see later) be applied when validating the other assay types using recombinant ER $\alpha$  or ER $\beta$  proteins.

The Panel recommended that receptors for species other than human and rat should be considered for ecotoxicological concerns. There have been no species identified for use in general ecotoxicology screening, but the identification of such an environmentally relevant ER is important and should be considered in future plans. There are three ERs in fish, but it is not known how results with human and rat ERs reflect the binding of substances to any of these fish proteins. Amphibian liver ER has been proposed for the testing of substances that could alter endocrine disruption in amphibians and reptiles. Some European laboratories have made reference preparations of these non-mammalian ERs.

The Panel acknowledged that while an assay using fluorescent polarization (FP) would be advantageous, this assay currently is not in wide use, and there are limited data available for comparison with other methods. The FP assay also requires specialized equipment and a fluorescently tagged estrogen. A fluorescent estrogen would obviate the use of radioactivity, the use of which is being phased out in many European countries. Although the FP method has many attributes, the Panel recommended that methods that use radioactivity should be used for the present.

The Panel agreed that incorporation of metabolic activation capabilities into the test system should be considered. However, it would be difficult to obtain the ideal *in vitro* system, since most *in vitro* metabolic activation systems only contain enzymes and co-factors for phase I metabolism, which generates molecules that have binding activity. As there is no significant phase II (detoxifying) activity in many of these preparations, inclusion of metabolic activation would be expected to generate false positives. Another difficulty would be how to incorporate such a system into the assays. Thus, although the

inclusion of a metabolic activation system would be desirable, the Panel did not recommend it until extensive development of this aspect of the assay was undertaken in the future.

While there are no known patent issues pertaining to hERs, there are some commercial assays that use these proteins. The question of patent issues should be investigated because they could affect any test system that would be selected.

## **2. *Minimum procedural standards for in vitro ER binding assays***

The Panel was in agreement regarding most of the procedural standards in the BRD but also proposed revisions to the following standards:

### Dissociation Constant of the Reference Estrogen:

- The dissociation constant must be determined with each set of assays. Hexa-tritium labeled 17 $\beta$ -estradiol- (i.e., 2,3,6,7,16,17-<sup>3</sup>H) 17 $\beta$ -estradiol should be used as the ligand for all assays because it is the most potent naturally occurring estrogen in the human body, and because it is available commercially with a high specific activity. Such a potent preparation will considerably increase the sensitivity of both the ligand titration array and the ligand competition assays.

### Preparation of Test Substances:

- Test substances should be prepared in water, absolute ethanol, or dimethyl sulfoxide (DMSO), depending upon their solubility. Preference should be given to the solvent that allows testing of the highest concentration of the test substance, without exceeding the limit dose.

### Concentration Range of Tests Substances

- It was recommended that the highest dose tested should depend on the solubility constant ( $K_{ow}$ ) of the substance. This concentration may or may not be 1mM as originally proposed in the BRD. The substance at the highest dose will then be diluted by seven orders of magnitude in log decrements to obtain the relevant dilutions of the test substance for the assay. This will permit the generation of a dose response curve.

### Solvent and positive controls:

- A set of solvent (vehicle)-only controls (with solvent concentrations identical to those used with reactions containing test substances) must be included in each set of assays.
- The solvent (vehicle) volumes must remain constant throughout the concentration range tested.
- A naturally occurring estrogen, such as estriol or coumestrol, or an estrogen mimic, tamoxifen is recommended as a positive control.
- The positive control should be tested at 3 dose levels whenever each assay is run.
- Because it is anticipated that many of the substances that will be tested in the future will be weak, the inclusion of a weakly positive control substance should be considered if only one positive control substance will be used. The routine use of a

weakly positive control would establish the lower level of sensitivity of the assay, and confidence in low-level responses. However, no recommendation of a specific substance was made.

#### Within Test Replicates:

- Triplicate measurements should be performed at each dose level.

#### Data Analysis:

- More details are needed on statistical models for non-linear regression to assess  $K_d$ ,  $K_i$ , and  $IC_{50}$  values.
- The statisticians noted that it is also important to calculate standard errors or other confidence levels associated with the  $K_d$ ,  $K_i$ , and  $IC_{50}$  values and that these calculations may not be trivial.
- The use of alternative approaches such as the ligand titration array which provides simultaneous evaluation of a laboratory's performance and determination of the estrogen binding properties (e.g.,  $K_d$ ,  $K_i$ , and  $IC_{50}$  values) of both reference and test substance was recommended.

#### Assay Acceptance Criteria:

- A detailed assay protocol must be provided for performing each type of assay (i.e., ligand titration and competition), with criteria for evaluation and acceptance of results, with demonstrated assay validation and lab transferability.
- Achieving a specific binding capacity and  $K_d$  value for the reference receptor protein is a critical measure of the robustness of the procedure. These data are essential to the establishment of a Quality Assurance Program (assay proficiency).
- A reference ER preparation, with established binding parameters must be employed for the determination of the  $K_d$  value and specific binding capacity by the laboratories chosen for the validation of the ER binding assay.

#### Evaluation and interpretation of results:

- The method of calculation of the statistical parameters and assumptions must be justified. The classification of a test substance as "positive for binding" will require the use of statistical models. Historical data can also be used to assess the biological significance of results for a current test that has shown to be statistically significant.
- The Panel did not come to a clear consensus on the definitions of a weak response or a negative response. However they were agreeable on an equivocal response.

#### Test Report:

- Solubility information should be included in the test report
- A description of the justification for the chemical concentrations used must be included in the report.
- A clear identification of the test chemical by name and Chemical Abstracts Service Registry Number (CASRN) is required. The chemical structure may also be desirable in some cases, especially where the substance is chiral or if the CASRN points to a substance that is not pure. A proposal was originally made to use the IUPAC name

for the chemical but it was noted that it is often difficult to determine this nomenclature and the common name would be sufficient.

- The Panel recommended establishing a new range of reference  $K_d$ ,  $K_i$ , and  $IC_{50}$  values with a standardized ER preparation using a test set of substances.

The Panel recommended the following additional minimum procedural standards:

- The assay used for protein determination should be specified and the concentration of protein used in the reactions reported.
- 10 mM sodium molybdate as well as a cocktail of protease inhibitors should be employed to minimize degradation of the receptor protein during the assay.
- Dextran-coated charcoal is preferred over the hydroxyapatite procedure for separating the free from the bound radiolabeled  $17\beta$ -estradiol.
- In performing the binding assays, a range of 50-100 fM of hER, which corresponds to 5-10 pg/mg of extracted protein, was recommended. There is less experience with hER, so no protein range could be recommended.

There was extensive discussion regarding the need for, and use of, concurrent positive controls during the performance of the assays. With the exception of one member of the Panel, the Panel agreed that concurrent positive controls are essential. The purpose of the positive control is to measure the performance of the test and of the laboratory. The reference ligand,  $17\beta$ -estradiol should not be used as the positive control in the ER assays because it would then be compared against itself. Although there was agreement with the need for positive controls and the need for consistency among ER and AR assays, there was no consensus regarding the minimum numbers and types of controls to be used, specifically as they related to substances with low activity. The advantage of including control substances that would be expected to elicit low and mid-range responses would be the ability of determining the limit of detection of the test in the laboratory on a specified day. This would aid in concluding whether a test result is called positive or negative. The ideal situation would be the inclusion of three or four positive controls spanning a range of different binding affinities to measure test and laboratory performance. This is especially important because of the increasing variability in the response as one moves towards the lower end of the dose-response curve. Recommended positive controls were estrone and estriol, which are one and two logs less potent *in vitro* than  $17\beta$ -estradiol. Reasons presented for limiting the assays to one positive control substance is cost and level of effort.

The Panel recommended that each BRD contain a separate paragraph or section describing pertinent statistical analysis, and especially the evaluation of low-activity chemicals. However, the biostatisticians on the Panel stated that currently insufficient data are available to address all the statistical or data evaluation issues that would enable them to recommend specific statistical analyses. Before specific statistical procedures and action levels can be identified, more details are needed about the methods and their performance criteria. It will be necessary to evaluate confidence limits, standard errors etc., to better understand the data. Different data and statistical analyses will be required depending on whether the test will be used simply as a yes/no indicator, than if the results will be used in a quantitative manner.

Classification of a substance as positive will require a formal statistical procedure if the test substance does not produce a clear-cut sigmoid curves. For these reasons, a large database of substances that are negative or elicit weak responses needs to be established. This database could then be used to build the appropriate statistical models for the various measurements and endpoints. Pre-validation studies, or studies before entering pre-validation, can be used to generate this needed data.

When undertaking a receptor binding assay it is important that the  $K_d$  and  $B_{max}$  be determined. There was an unresolved question whether  $B_{max}$  should be determined every time a binding assay is performed, or just for every lot of receptor. This determination ensures that the reference preparation of receptor is performing properly and that values can be compared across laboratories. Titration assays are justified because chemicals may interact with, and damage, the receptor in a non-ligand-binding manner. The  $K_i$  should also be calculated, and this can easily be done using commercial software packages. Such calculations show good agreement with the values obtained using a Scatchard plot. The statisticians noted that it is also important to calculate standard errors or other confidence levels associated with the  $K_i$ , and that these calculations may not be trivial. It was proposed that the  $K_d$  and  $B_{max}$  values for a number of model chemicals be established as part of prevalidation studies.

### **3. *Recommendations for In Vitro ER Binding Test Method Protocols for Validation Studies***

The Panel reiterated its recommendation that an assay using a recombinant ER protein, preferably the human ER , should be developed. The assay protocol could be modeled on the BRD recommended RUC assay protocol, which is similar to the U.S. EPA protocol currently being used to measure ER binding of 21 substances by three laboratories. The cytosol-based assay can be refined to accommodate a purified protein instead of a cytosolic preparation. The Panel proposed that a concurrent positive control be included in the protocol as:

- It is a hallmark of *in vitro* tests used in the regulatory arena worldwide.
- It is a stated “requirement” in protocols submitted to ICCVAM.
- It measures the assay’s performance and stability over time.
- It provides the basis for assessing the acceptability of the assay trial and thus the use of data from “unknowns” tested concurrently.
- It provides a basis for comparison of assay performance across laboratories.

### **4. *Recommended List of Substances to be Used for Validation of In Vitro ER Binding Assays***

The Panel expressed concerns regarding the composition of the list of substances for validation as to whether it included the kinds of substances that the U.S. EPA is interested in screening. They stated that an adequate representation of substances across chemical classes and across the range of potencies must be considered. Also, there

should be an appropriate ratio of estrogens to non-estrogens in the list. The Panel made specific recommendations as follows:

- The number of negative substances must be increased from the current 9% to at least 25% in order to determine the specificity of an assay. Ideally, at least half the test substances should be non-estrogens, with about 60 compounds included in a validation set. A wider range of negative substances that belong to a wider range of chemical classes is needed so that appropriate criteria for negative results can be developed. This is especially important since many of the positive chemicals that will be encountered in the testing of industrial or environmental chemicals are likely to be weak, and the test needs to be sufficiently sensitive to detect these substances.
- Presently, there is insufficient information available to evaluate the utility of the binding assays at low potency ranges.
- An underrepresented class, the phthalates, was recommended by a number of Panel members as a group of substances that should be added to the list of negative substances, although no specific phthalates were identified.
- It was recommended that the EPA should maintain a repository of the chemicals to be used in the validation studies. A suggestion was made that there be two lists of substances, one set of substances would be used to test the protocol (which includes the pre-validation studies), and a second, more extensive set of substances for use in the validation studies.
- There was limited discussion regarding quality assurance issues. It was recommended by one Panel member that entry and exit assays be incorporated into the testing. This refers to the analysis of test chemical stock solutions before and after the assay to assure the identity and purity of the chemical, and its stability in solution. This recommendation was not generally acceptable to the Panel. It was noted that where many diverse chemicals are being screened, the analytical chemistry could be more complex and more expensive than the biological tests.

### **Public Comments Session (Morning, May 21)**

Mr. Richard Becker (American Chemistry Council) commented that patent and proprietary issues, and restrictions on the use of certain methods, were often stumbling blocks to international acceptance of methods. The Panel was requested to give consideration to these concerns and to the availability of methods and materials.

### **B. Estrogen Receptor (ER) Transcriptional Activation (TA) Assays**

*Primary reviewers: J. Stegeman, Group Chair (Woods Hole Oceanographic Institute.); G. Charles (Dow Chemical Co.); E. Mihaich (Rhodia, Inc.); T. Weise (Tulane and Xavier Univs.; not present at meeting); J. Yager (Johns Hopkins School of Public Health); T. Zacharewski (Michigan State Univ.); S. Peddada (National Institute of Environmental Health Sciences).*

## Summary of the ER Transcriptional Activation Background Review Document

Dr. Shane briefly summarized the assays described in the BRD. The same on-line databases, as searched for the ER binding reports, were searched for relevant publications for the ER Transcriptional Activation BRD. Key words included *screen, tests, batteries, bind, ligand, agoni, antagoni, transcription, estrogen, and receptor*. The search yielded 258 records; data were available from 86 for inclusion in the BRD. For the agonism assays, the qualitative positive or negative response, a measure of relative activity, EC<sub>50</sub>, (μM), and cell growth information were extracted. For the antagonism assays, the qualitative response, relative activity, and the IC<sub>50</sub> (μM) were extracted.

The BRD database contains 95 assays, and data on 703 unique chemicals. Of these chemicals, 634 were tested for agonism; 228 (36%) were tested in 2 assays and 51 (8%) were tested in 5 assays. Of the 255 chemicals tested for antagonism, 94 (37%) were tested in 2 assays and 8 (3%) were tested in 5 assays. The database of chemicals was comprised of 15 chemical classes and 11 product classes, for which there were 10 or more entries. The most frequently tested substances in the chemical and product classes were polychlorinated biphenyls and pesticides (including metabolites), respectively.

The 95 assays included 63 permutations of 9 human cell lines: BG-1, HEC-1, HEK293, HeLa, HepG2, MCF-7, MDA-MB-231, T47D, and Ishikawa cell lines and three other mammalian cell lines: CHO-K1, COS-1, and ELT-3. The ER<sub>α</sub> and ER<sub>β</sub> proteins were purified from human, mouse, and rat ER (unspecified). The activity of luciferase or chloramphenicol acetyltransferase reporter genes were used as a measure of TA. There were 10 mammalian cell proliferation assays that used Ishikawa, MCF-7, T47D, and ZR-75 cell lines. In addition, there were 22 yeast assays involving 13 *S. cerevisiae* strains with the hER, hER<sub>α</sub>, hER<sub>β</sub>, mER, and rtER receptors, and a -galactosidase reporter gene.

Comparative performance and reliability analyses of EC<sub>50</sub> or IC<sub>50</sub> values were not conducted because the numbers of substances tested in multiple assays, or multiple times using the same assay in the same or different labs, were too limited for an adequate comparison.

Based on these considerations of the available data, recommendations for minimum procedural standards were prepared for the BRD. In addition 31 chemicals were recommended for use in future validation studies in agonism assays, and 21 were recommended for use in future validation studies in antagonism assays.

### ***1. Recommendations and Prioritization of ER TA Assays for Validation Studies***

The Panel agreed that there was too little information to recommend one assay over another. No specific cell line could be preferentially recommended for ER TA because there was not enough data presented in the BRD from the different mammalian cell lines. One concern was that the activity of a chemical will probably be species-, tissue-, cell-, and promoter-specific, and therefore its response can not be generalized based on results



from any single assay. Also, potential differences in co-activator populations, cross talk with other receptors, and other signal transduction pathways between different cell types, etc., could alter the response in a cell. As a result of this complexity, there are a number of aspects of the various cell lines that will have to be investigated further before any decision can be made on the most appropriate cell line for an assay.

Discussion then ensued on whether a stably transfected or transiently transfected ER cell line or a cell line with an endogenous ER should be recommended. A stably transfected cell line would seem preferable but no conclusion can be drawn until appropriate comparative data are collected on cell lines with each of these different types of receptors.

The difficulties with stably transfected cell lines are the frequent instability of the constructs, problems encountered in maintaining highly responsive lines, and the limited availability of these lines. Since transiently transfected cell lines have more flexibility, they may be more appropriate for screening. The Panel suggested that an important part of the validation process would be a study to determine if stably transfected lines perform better or are more sensitive than transiently transfected cell lines. It was recommended, therefore, that before any test validation is begun, a research and development effort is implemented to compare the responses of stably and transiently transfected cell lines to the same small group of chemicals. This would involve a comparison of the response of a mammalian line stably transfected with a receptor and reporter with one transiently transfected with the same ER and reporter plasmids. In addition, the response of a cell line with an endogenous receptor needs to be evaluated alongside these transfected cell lines. If stable cell lines are selected for validation, there should be a standard procedure for evaluating their performance and the stability of the constructs. Stability can be monitored by antibiotic selection.

The Panel agreed that in the development of an assay, a number of different constructs with different components transfected into different cell lines need to be evaluated, optimized, and at a minimum, clearly defined for each assay. These include the components of the reporter construct, the number of EREs, the presence of other enhancers in the construct, the types of promoters, as well as the co-activators and co-repressors in the cell line. As a beginning to pre-validation, it was recommended that a series of transient transfection assays for individual receptor subtypes be developed and evaluated.

The Panel agreed that the ER $\alpha$  and ER $\beta$  are the most appropriate receptor types, but if patent issues arise with the use of human ER, the rat ERs would be an acceptable alternative. Supporting data needs to be obtained to determine whether the use of ER alone, would be sufficient.

The Panel was of the opinion that although the vitellogenin response element (vitERE) responds to substances that bind to the progesterone or corticosterone receptors found in some cell lines, this estrogen response element should be used due to its sensitivity. For optimized sensitivity, multiple vitERE constructs were recommended. Chimeric ligand binding domain ER's should also be considered for these preliminary studies due to their mechanistic specificity.

Although the metabolic activities of the various cell lines need to be considered, most cell lines used in these assays have not been characterized with regard to their metabolism of xenobiotics. To characterize the metabolism for a range of chemical structures is an enormous undertaking although it can be done with a few model chemicals. It is also possible that certain test substances can induce metabolism. Therefore, the metabolic characterization of untreated cells may not be relevant. A caution was presented regarding the exogenous metabolic activation systems, and those inherently present in the cell lines being used, that they may not mimic those found in the relevant *in vivo* target tissues.

## 2. *Minimum Procedural Standards for In Vitro ER TA Assays*

The Panel agreed with the recommended minimum procedural standards in the BRD, with recommended the following additions:

### Concentration Range of Test Substances:

- The Panel agreed that the limit concentrations could be 1mM as long as the solubility characteristics and cytotoxicity of the test substance is taken into consideration. There was a consensus, however, that, in general, concentrations of the test substances above 10  $\mu$ M should not be used because this concentration is excessive and often problematic due to solubility issues in aqueous media. A concentration range from 1 nM to 10  $\mu$ M should be sufficient for a screening study. The Panel recommended that since certain chemicals (such as tamoxifen) can be estrogenic at low doses and anti-estrogenic at high doses, tests should be performed over a wide dose range, and single-dose experiments be avoided. Incorporation of a measure of cellular cytotoxicity into the assay could help define the upper limit for test material concentrations, similar to the Maximum Tolerated Dose (MTD) approach used in *in vivo* studies. This measure of potential cell cytotoxicity/cell proliferation should be a part of the data collected to ensure non-toxic doses are being used.
- Since solubility could affect absorption of the test substance by the cell, it might be necessary, to evaluate the uptake of the substance using isotope-labeled substances.

### Solvent and Positive Controls:

- The Panel suggested that guidelines be provided with regard to the concentration of solvent in the stock solution. Whether ethanol or DMSO is used, compounds to be tested could be prepared in stock solutions so that the test substance concentration approaches the solubility limits. However, this approach could introduce variation from laboratory to laboratory and thus should be standardized. In addition, controls need to be in the same carrier solvent as the test substances. A pre-validation of the

TA assay should be performed with the reference estrogen, to assess the level of solvent that does not adversely affect assay response.

- The Panel agreed that ICI 182,780 appears to completely block 17 $\beta$ -estradiol at 0.1  $\mu$ M and thus, it should be used as the positive antagonist. However, availability of ICI 182,780 may be limited. Clear guidelines should be given for the positive antagonist and the expected extent of antagonism when testing the compound.
- Each test substance that is positive in the agonist assay could also be tested with ICI 182,780 to confirm a receptor-mediated activity.

#### Within-test Replicates:

- The test must be run in triplicate at each concentration.

#### Data Analysis: (for more details see Expert Panel Report)

- The Panel recommended that preliminary studies be performed with multiple transactivation assays to statistically define assay performance expectations for 17 $\beta$ -estradiol dose response curves (i.e., maximum fold induction, EC<sub>50</sub> values, confidence limits).
- The biostatisticians on the Panel stated that currently, insufficient data are available to address all the statistical or data evaluation issues that would enable them to recommend specific statistical analyses. Before specific statistical procedures and action levels can be identified, more details are needed about the methods and their performance criteria. It will be necessary to evaluate confidence limits, standard errors etc., to better understand the data. Different data and statistical analyses will be required depending on whether the test will be used simply as a yes/no indicator, than if the results will be used in a quantitative manner.

#### Assay Acceptance Criteria:

- The transcriptional activation-inducing ability of 17 $\beta$ -estradiol must be demonstrated. A consistent minimum response would be an appropriate criterion for assay acceptability.
- Reference compounds for agonism and antagonism should give responses within appropriate confidence limits. These confidence limits should be determined in preliminary studies. Guidelines should be provided for a certain expected range in response for the reference standards in agonism and antagonism assays, and responses in these ranges should be required if the assay is to be accepted.

#### Evaluation and Interpretation of Results:

- The interpretation of positive results for a compound as an agonist or antagonist should incorporate some elements of dose-response in comparison to the reference standards. Simply classifying a substance as an ER agonist based on a significant response above the concurrent control without consideration of a dose-response is not sufficient.

#### Test Report:

- The complete DNA sequence of constructs and vectors used for receptor and reporter genes should be identified.

- All assay parameters regarding cells, plasmids, culture methods, transfection methods, and a method for measuring luciferase activity must also be reported. For a transfection assay, a constitutive reporter gene assay must be included to control for transfection efficiency between wells. The passage number of cells should be tracked. The % CO<sub>2</sub> in the incubator must be monitored. EC<sub>50</sub>/IC<sub>50</sub> values, fold change, and confidence limits must be reported.
- Solubility information should be included in the test report.
- A description of the justification for the chemical concentrations used must be included in the report.

#### Additional Minimal Procedural Standards

Cell Toxicity: The Panel discussed what level of toxicity would be acceptable for inclusion of the data if cell toxicity was observed. No agreement was reached on the definition of toxicity, nor how it should be measured. Two suggested endpoints were overt cell death or decreased expression of a specific marker product. The measurement used may be dependent on the test system. Although some Panel members proposed a 10% killing as a cut off value, no consensus was reached regarding this value or any other specific value. It was agreed, however, that some value(s) would have to be defined. Methods for quantifying cytotoxicity in the TA assays included measurement of the activity of the gene product of a co-transfected  $\beta$ -galactosidase or luciferase gene that fluoresces at a different wavelength than the *luc* reporter gene used in the same cell. CMV-driven *luc* plasmids were suggested as the carrier of the co-transfected gene, although these plasmids might be affected by some test substances and therefore respond to non-endocrine transcription signals.

Corrections to the BRD: There were two observations in the BRD that require clarification. Firstly, there is an inconsistency in the statements on page 12-1 and 12-11 [in the BRD] concerning stable vs. transiently transfected cells. Secondly, there was no discussion of individual assays for ER<sup>+</sup> and ER<sup>-</sup>.

Discussion ensued as to whether a tiered strategy should be adopted for the TA assays. For example, if the compound is positive for agonist activity in the TA assay, is there any value in testing it for ER antagonist activity or AR-mediated activities? It was pointed out that a positive result in any of these assays will likely warrant further examination in tests other than transactivation assays. However, other Panel members disagreed with such a tiered strategy because the assays will be used as part of an integrated test battery and thus the elimination of one of the endpoints (agonism or antagonism) would be equivalent to losing part of the data. No consensus on a tiered approach was reached.

A discussion then followed as to whether data in the *in vitro* assays would trigger the testing of a substance in an *in vivo* assay. The Chair then asked Gary Timm (U.S. EPA) to clarify the roles of these tests.

According to Mr. Timm, the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) report, and the U.S. EPA Endocrine Disruptor Testing Program

(EDSP) proposals view all the Tier 1 tests as an integrated battery, and no single test result will trigger Tier 2 testing, or the designation of a chemical as a potential endocrine disrupter. The EDSTAC report had a preference for TA assays for mechanistic reasons, but binding assays were regarded as equally acceptable. The *in vitro* Tier 1 assays are not considered a sub-tier for the *in vivo* assays because the EPA proposes to evaluate the results from the entire Tier 1 battery in a weight-of-evidence approach. The composition of the specific Tier 1 battery to be used has not yet been determined. This determination will be based on the outcomes of the validation studies for each test method proposed for the battery. EDSTAC proposed that the Tier 1 *in vitro* and *in vivo* tests be run simultaneously, but recognized the role of *in vitro* tests in identifying chemicals for *in vivo* testing. Mr. Timm stated that the EPA does not contemplate running thousands of chemicals through the entire Tier 1 battery. Mr. Timm stated that the *in vitro* assays would not be used for priority setting.

Mr. Timm went on to say that the data that are presently being generated in Tier 1 testing by contractual arrangements will be publicly available; but when, and in what form the data will be released has not yet been determined. The EPA is sensitive to the potential problems associated with piece-meal release of the data and will probably release the data on a chemical-by-chemical basis, not by test. The EPA is also concerned with potential confidential business information issues that may attach to some of the data and is working to resolve this issue.

Comments were made that, in the future, gene expression profile patterns may be used to identify endocrine-active substances, and to distinguish estrogens from androgens, and agonists from antagonists. Gene panels can be developed for different tissues. Binding assays, as they are currently performed, may be considered relatively "old science."

### **3. *Recommendations for In Vitro ER TA Test Method Protocols for Validation Studies***

The Panel agreed that the protocols are described adequately but the details of the protocols are contingent on the incorporation of the minimum procedural standards. The Panel was of the opinion that any laboratory with cell/yeast culture and basic molecular experience should be able to produce dependable results. Some inter-laboratory variability may be expected due to laboratory specific techniques (e.g., cell counting). Acceptance standards should be specified for culture techniques such as cell counting, determination of % confluency, ability to seed plates evenly, etc. to limit inter-lab variability. Additional procedure details should be added if volatile chemicals are tested. The following topics need to be added or expanded in the protocols:

- Standards for uniform counting and plating of cells in wells between experiments.
- Review of methods for making DCC stripped sera or a recommendation for commercial sources of this serum.
- Review of known sources of estrogen contamination in the lab.
- Discussion regarding the culturing of cells in estrogen rich media and withdrawal to an estrogen-free medium.
- Discussion of washing techniques and number of days for withdrawal.

- Discussion of procedures to demonstrate that the lab and each particular experiment is performed under estrogen-free conditions (e.g., ICI vs. blank reporter activity).

### **Other Available Standardized Protocols**

Dr. Thomas Weise in a written contribution (Dr. Weise was not able to attend in person) suggested that the MVLN Assay, that uses MCF-7 cells stably transfected with the vitellogenin-luciferase reporter plasmid, is among those that should be considered further. [Copies of the procedure were made available to the Panel members]. The Panel agreed that it should be one of the assays validated with the other proposed assays.

The Panel emphasized that standardization and validation of assays across laboratories is critical and must occur before these assays are used for regulatory purposes. A formal validation process is needed in order to establish a “gold standard” study for use and not just to have personal variants of similar assays.

#### **4. *Recommended List of Substances to be Used for Validation of In Vitro ER TA Assays***

The Panel was of the opinion that the distribution of the recommended substances seemed appropriate, but more thought should go into the final compilation of the list that is used for validation. The following criteria should be considered:

- Inclusion of more chemicals expected to be negative.
- More overlap of chemicals used for validation of the ER binding assays and the ER TA agonist and antagonist assays.
- Close collaboration and cooperation regarding chemical selection with the *in vivo* test validation studies being reviewed by the EPA’s Endocrine Disrupter Methods Validation Subcommittee is encouraged.
- Possible inclusion of phthalates, polychlorinated biphenyls, and additional polycyclic aromatic hydrocarbons. These classes were originally omitted from the list of substances due to their limited availability from a commercial source and difficulties with their disposal.
- All substances for validation should come from one EPA repository.
- Chiral compounds (i.e., compounds that cannot be superimposed upon their mirror images and are thus asymmetrical) need to be included in the validation list as different components of a racemic mixture may elicit different responses. There is a possibility that one enantiomer could be an agonist while the other is an antagonist with the racemate is neutral. These substances should be included in the validation list, but be omitted from the pre-validation list.

### **Public Comments Session (Afternoon, May 19)**

Dr. George Clark (Xenobiotic Detection Systems, Inc.) presented information describing the construction and performance of his company’s chemical-activated luciferase expression (CALUX) screening system for ER transcriptional activation. This assay uses a stably transfected cell line, BG1, which contains a luciferase reporter gene. Information on this assay was submitted to NICEATM for inclusion in the ER TA BRD. Based on

the information presented, the test system is amenable to high-throughput screening and is highly reproducible. The cells express predominantly ER (95%) with a low amount of ER (5%). This test system is available commercially and the company can supply the cells, or multi-well plates that are coated with the cells, for use by the customer. Alternatively, the company also provides testing services.

A Panel member asked what approach will be used by regulatory agencies in identifying companies such as Xenobiotic Detection Systems, Inc., that might be developing or have developed an *in vitro* method useful for screening. Dr. Stokes responded from an NIEHS and ICCVAM perspective and stated that Dr. Clark's report would be made available to the public and forwarded to Federal Agencies so that it can be considered by individuals or organizations that wish to support validation. Once methods have gone through validation, the data can then be submitted to EPA and/or ICCVAM for further evaluation.

### **C. Androgen Receptor (AR) Binding Assays**

*Primary reviewers: T. Brown, Group Chair (Johns Hopkins Univ.); T. Gasiewicz (Univ. of Rochester Medical Center); T. Inoue (National Institute of Health Sciences, Japan); B. Robaire (McGill Univ., Canada); A.M. Vinggaard (Danish Veterinary and Food Administration, Denmark); W. Piegorsch (Univ. of S. Carolina).*

#### **Summary of the AR Binding BRD**

Dr. Shane provided an overview of the AR binding BRD. The same on-line databases were searched for relevant publications for inclusion in the AR binding BRD, but using the following key words, *screen, tests, batteries, bind, ligand, androgen, and receptor*. The search yielded 108 records; data from 23 were included in the BRD. The same data as abstracted in the ER BRD were collected from the records.

The BRD database contains 11 assays, and data on 108 unique chemicals from AR binding. Of these chemicals, 33 (31 %) were tested in 2 assays, and 11 (10%) were tested in 6 assays. The chemicals were assigned to chemical and product classes with nonphenolic steroids and pharmaceuticals being the most frequent chemical and product classes, respectively.

The 11 assays included: calf uterine cytosol, rat prostate cytosol, and rat epididymal cytosol and the nuclear fraction from rat epididymal cells, MCF-7 cell cytosol; COS-1 cells transfected with the hAR (COS-1+hAR), LNCaP cells, and; intact human genital fibroblast (HGF) cells; purified recombinant human AR (rhAR).

Comparative performance and reliability analyses of IC<sub>50</sub> or RBA values were not conducted because the numbers of substances tested in multiple assays, or multiple times using the same assay in the same or different laboratories, was too limited for an adequate comparison. Thirty-one chemicals were suggested for validation; three (10%) of which were negative.

## ***1. Recommendations and Prioritization of Assays for Validation Studies***

The Panel discussed the advantages and disadvantages of the three different assay systems used to measure AR binding namely, the rat prostate cytosol assay, the cell-based assay using COS-1 cells transfected with a human AR, and the assay using purified human AR (hAR). The Panel recommended that an assay using purified recombinant hAR or rat AR (rAR) (or other species) be developed. The Panel did not recommend that a metabolic activation system be incorporated into the assay system at this time.

The Panel was concerned that a potential difficulty in using purified AR is that the human AR cDNA sequence is protected by patent, and commercial use of the hAR in functional assays is restricted by a license. It is not known whether the rat AR cDNA sequence is also protected by patent restrictions. As a result, the full-length recombinant AR is not presently available for use in an AR binding assay. A recombinant human AR protein is available, but it only contains the ligand binding domain of the protein. The reliability of this protein in the binding assay has not been established.

The Panel noted that an assay with whole cells that contains an endogenous AR is unlikely to be restricted by patents, and that some of these cell lines express significant amounts of AR. However, they noted that cells containing endogenous receptors do not always express levels of AR that are as high as transduced or transfected cells, and they may have other inherent disadvantages such as stability. The relative simplicity of the transfected cell assay (e.g. COS + hAR/rAR) is amenable to high throughput screening and requires simple methods, minimal volumes of reagents, and few variations in buffers and solutions. One possible source of recombinant AR might be derived from nonhuman primates.

Dr. Hattan wondered what the implications would be if a substance was positive for binding or TA using the human receptor *in vitro*, but was negative in the *in vivo* rodent tests. Could such a response be based solely on the different sources of the receptor? The Panel thought that the similarities in receptors between humans and rodents are such that it would be the rare exception where differences between the *in vitro* and *in vivo* responses were based solely on the composition and responsiveness of the receptor. Because of their homologies, the receptors are expected to have similar binding characteristics, although the binding kinetics could be affected by the contribution of other parts of the receptor molecule besides the binding domain. In the situation that was described, the activity of the substance *in vitro* can be further examined using the rodent receptor. This may be an important consideration because of possible post-translational changes to the receptor that does not occur *in vitro*. It was noted that it is not unusual to get positive *in vitro* and negative *in vivo* test results because of the differences in their sensitivities. The problem is not so much one of biology as it is of public perception of the relevance of the *in vitro* test.



Following this discussion, the Panel recommended Government agencies should, in light of the status of the patents and licenses, provide guidance for the development and use of AR assays in the public and private domains.

The Panel unanimously agreed that rat prostate cytosol (RPC) was not the best source of the AR for these assays because:

- The RPC contains other steroid receptors that may interfere with the assay for AR binding.
- Some metabolism of the test substance may occur even in cytosol preparations.
- RPC cannot substitute for hAR, or AR in those wildlife species where significant exposure to androgenic chemicals may occur.
- The AR is extremely unstable in cytosolic preparations and in fact, the protein is usually degraded so that only the AR binding domain remains intact.
- Although the RPC has been the most utilized assay to measure AR binding, this is the more difficult of the assays to perform in a standardized format.

The Panel recommended that the simplest and most consistent assay would be one in which the AR protein would be fixed in multiwell plates and tracer and test ligands added in appropriate amounts to develop data for a Scatchard (or equivalent) analysis. They also recommended that there should be a move away from radioactive tracer ligands toward more environmentally friendly and safer fluorescent ligands.

The Panel recommended that irrespective of which assay was developed and validated, that it should be acceptable at the international level (*e.g.*, It should not have to comply with patent regulations and regulations regarding the use of radionuclides).

## 2. ***Minimum Procedural Standards for In Vitro AR Binding Assays***

The Panel agreed with the AR binding BRD regarding minimum procedural standards, with the following additions and revisions:

### Dissociation Constant:

- The  $B_{\max}$  and  $K_d$  of the reference androgen should be determined in each assay and all laboratories should be able to generate comparable values within accepted limits. These values are a critical measure of the robustness of the procedure and the abilities of the laboratory.
- The minimum number of concentrations used to obtain the  $K_d$  should be stated.
- Straightforward procedures, such as ligand titration arrays for determining the  $K_d$  value of the radiolabeled reference ligand and the unlabeled test substance should be considered.

### Reference Androgen:

- 5 $\alpha$ -Dihydrotestosterone (DHT) is recommended as the reference androgen for an assay based on a purified receptor while methyltrienolone (R1881) or mibolerone is recommended for an assay based on cytosol or cells.

- Triamcinolone acetonide or a synthetic progesterone receptor (PR) agonist to block binding to the PR should be used in assays where PR is present and R1881 is used in the assay. Alternatively mibolerone could be used.

#### Preparation of Test Substances

- Preparation of stock solutions should be performed under rigorous quality control. The stability of stock solutions must be established.

#### Concentration range of test substances:

- At least 5 concentrations of the test substance should be examined to increase the likelihood of obtaining a satisfactory competition curve for estimation of the IC<sub>50</sub>.
- The limit dose should be 1 mM, taking into consideration the solubility characteristics of the compound.

#### Solvent and positive controls:

- As discussed for the ER binding assays, preference should be given to the solvent that allows testing of the maximal concentration of the test substance without exceeding the limit dose.
- A set of solvent-only controls (with solvent concentrations identical to those used with reactions containing test substances) must be included in each set of assays.
- The solvent volumes must remain constant throughout the concentration range tested.
- The positive control compound should have a binding affinity within two orders of magnitude of the limit of sensitivity of the assay. A second positive control within 1-10% of the RBA of the reference androgen should be included.
- One minimum procedural standard that was discussed at some length was the use of a positive control that is close to the level of detection of the assay. There was no clear consensus as to whether this is necessary. It would depend on whether one wants to categorize a substance as binding to the AR or whether one wants to determine an IC<sub>50</sub> value. The routine use of a weakly positive control would establish the lower level of sensitivity of the assay, and confidence in low-level responses.

#### Within test replicates:

- Triplicate measurements should be performed at each dose level.

#### Data analysis:

- More details are needed on statistical models for non-linear regression to assess  $K_d$ ,  $K_i$ , and IC<sub>50</sub> values.
- Mode of calculation and assumptions for the statistical methods must be justified.
- The designation of "equivocal" for compounds that do not bring about a 50% reduction in specific androgen binding is acceptable.
- The classification of a test substance as "positive for binding" requires the use of statistical methods.
- The biostatisticians on the Panel stated that currently insufficient data are available to address all the statistical or data evaluation issues that would enable them to recommend specific statistical analyses. Before specific statistical procedures and action levels can be identified, more details are needed about the methods and their

performance criteria. It will be necessary to evaluate confidence limits, standard errors etc., to better understand the data. Different data and statistical analyses will be required depending on whether the test will be used simply as a yes/no indicator, than if the results will be used in a quantitative manner.

- It may be useful to determine whether binding is through a non-competitive, competitive, or uncompetitive mechanism for substances that demonstrate an unusual binding curve. This determination is most easily accomplished by adding different concentrations of the test substance to different concentrations of radiolabeled hormone to generate a number of curves as proposed in the ligand binding array. The slopes of the lines are then plotted and the intercept of the line with the X axis is the  $K_i$ .

#### Assay Acceptance

- The Panel recommends that the assays be performed in compliance with Good Laboratory Practice (GLPs).

#### Additional Minimum Procedural Standards

- The assay used for protein determination should be specified and the concentration of protein used in the reactions reported.
- The chemical and radiochemical purity and the supplier of the radiolabeled androgen should be stated.
- A new range of reference  $IC_{50}$ ,  $K_d$ , and  $K_i$  values with a standardized AR preparation using a set of test compounds should be established.

### **3. *Recommendations for In Vitro AR Binding Test Method Protocols for Validation Studies***

The Panel concluded that there is no existing, standardized, acceptable protocol for an AR binding assay. However, the RPC protocol, which was well written, could be used as a model for the development of a protocol using a purified AR (either the entire protein or the binding domain if the entire protein cannot be used). The protocol described for the COS cell binding assay did not have the necessary details that are required for future testing of AR binding substances. In addition to the minimum procedural standards recommended by the Panel, the following considerations should be taken into account before a final protocol is developed.

- If a transfected cell line is adopted, a standard transfection protocol based on commercially available transfection agents and a standardized cell line would be necessary.
- The production of a stable cell line expressing the AR would avoid the problems inherent in transient transfection assays.

#### **Additional Protocol Elements**

The Panel agreed that the following details should be included in the RPC protocol:

- The maximal time of storage at  $-80^{\circ}\text{C}/-20^{\circ}\text{C}$  of cytosol, cells, or other material used as the source of AR should be indicated.

- The type of tubes/culture dish for homogenization and storage of cytosol or of cells should be indicated.
- Information on preparation and purity of the AR vector should be provided. Protocol elements for the COS cell binding assay (*e.g.*, preparation and stability of the vector, detailed timing on cell transfections, confluency of cells, transfection efficiencies, rationale for the choice of timing, incubation conditions, etc.) should be provided.
- If a cytosolic protein preparation is to be used, a cocktail of protease inhibitors, must be included to increase stability of the AR.

#### **Other Available Standardized Protocols**

The Panel suggested that if PanVera<sup>®</sup> is developing an AR binding assay using the AR ligand binding domain (LBD), this assay should be considered for validation.

However, since only the LBD is being used, it is not apparent what the sensitivity and reliability of this assay will be. Use of only the ligand binding domain recombinant protein is much less desirable than use of full-length AR in either an *in vitro* or *in vivo* assay because there is scientific evidence that the LBD interacts with other domains of the AR protein during the binding process.

There is no indication that a full-length recombinant AR will be available in the near future. Competitive binding assays for ER, PR, and GR that are available from PanVera<sup>®</sup> are based upon full-length recombinant proteins and do not use radioactivity.

The Panel is not aware of any other assay under development that would meet the desired criteria described in C.1 in the BRDs.

#### **4. *Recommended List of Substances to be Used for Validation of In Vitro AR Binding Assays***

The Panel was in agreement with the list of chemicals proposed in the BRD with the following additions and considerations:

- The same range and types of substances should be used for validation of both AR binding and AR TA assays.
- Anti-androgenic chemicals flutamide (or hydroxyflutamide, if used *in vitro*) and bicalutamide that bind to AR but do not initiate transcriptional activity, should be included in the list.
- Finasteride (the commercially available 5'-reductase inhibitor which does not bind to AR) should be added as a negative control.
- One or more of the estrogens (ethinyl estradiol, estrone or DES) can be omitted from the list, as 17 $\beta$ -estradiol is included.
- A number of negative substances should be added to the list. For example, phthalates, which can be activated *in vivo*, but do not bind to the AR should be added.
- A few substances that have been tested *in vivo* for which the *in vitro* database is extremely small or non-existent should be considered for testing.

- Additional non-binding chemicals need to be included in the recommended list of chemicals for validation studies. Androgen antagonists that do not have high binding activities should be included.

### **Public Comments Session (Morning, May 20)**

Dr. Yoji Ikawa (Otsuka Pharmaceutical Co., Japan) presented information describing the construction of cell lines and performance of the company's EcoScreen Transfection Assay (transiently transfected) and ER/AR-EcoScreen (stably transfected) assay systems for AR transcriptional activation. This information had previously been submitted to NICEATM for inclusion in the relevant BRD. The EcoScreen Transfection Assay is designed for high throughput screening, but ER/AR-EcoScreen cannot be used for high throughput screening. Testing was successfully performed using a liver cytosolic fraction from homogenized cells that had been centrifuged at 9,000x gravity (S9 preparations) for metabolic activation. This test system is available commercially from Otsuka Pharmaceutical Co.

### **D. Androgen Receptor (AR) Transcriptional Activation (TA) Assays**

*Primary reviewers: E. Wilson, Group Chair (Univ. of N. Carolina); K. Gaido (CIIT Centers for Health Research); W. Kelce (Pharmacia Corp.); S. Peddada (National Institute of Environmental Health Sciences).*

#### ***Summary of the AR TA BRD***

Dr. Barbara Shane summarized the information that had been included in the AR TA BRD. The same on-line databases were searched to retrieve publications with data on AR TA. The following key words, *screen, tests, batteries, bind, ligand, agonist, antagonist, transcription, androgen, and receptor* were included in the search, which yielded 108 records; data from 27 records were available for inclusion in the BRD. Similar data as described for the ER TA BRD were abstracted from all the records.

The BRD database contains data on 146 unique chemicals from 17 assays. Of these, 109 were tested for agonism; 49 (45%) were tested in 2 assays and 17 (16%) were tested in 4 assays. Of the 87 chemicals tested for antagonism, 22 (26%) were tested in 2 assays and 6 (7%) were tested in 4 assays. The most frequent chemical and product classes were nonphenolic steroids (35 substances) and pharmaceuticals (55 substances), respectively.

The 17 assays were comprised of 15 mammalian cell-based assays using six human cell lines: (HeLa, HepG2, MDA-MB-453, MDA-MB-453-kb2, PC-3, and PALM), two mammalian cell lines (CHO and CV-1), and one carp cell line (EPC). The ARs were derived from human, mouse, and rainbow trout. The luciferase and chloramphenicol acetyltransferase reporter genes were used. There was one mammalian cell proliferation

assay that used the LNCaP-FGC cell line. In addition, there was one yeast assay using *S. cerevisiae* YPH500 with the hAR receptor and a  $\beta$ -galactosidase reporter gene.

Comparative performance and reliability analyses of EC<sub>50</sub> or IC<sub>50</sub> values were not conducted because the numbers of substances tested in multiple assays, or multiple times using the same assay in the same or different laboratories, were too limited for an adequate comparison.

Based on these considerations of the available data, recommendations for minimum procedural standards were prepared for the BRD. In addition, 28 chemicals were recommended for use in future validation studies of agonism assays, and 25 were recommended for use in future validation studies of antagonism assays.

### **AR TA Group Presentation and Discussion**

The discussion of the AR TA BRD was led by Dr. George Daston because Dr. Elizabeth Wilson withdrew her participation in decisions regarding the AR methods due to her potential conflict of interest. Dr. William Kelce presented the draft conclusions and recommendations for the AR TA assays.

#### ***1. Recommendations and Prioritization of Assays for Validation Studies***

The Panel decided that they could not recommend a specific assay at this time because the available assays are not yet ready for standardization nor validation. There is a need for further methods development and standardization before a specific assay can be recommended for validation.

The Panel agreed with the BRD recommendation that a stable cell line be used for testing. The Panel proposed that the MDA-MB-453 cell line, which harbors an endogenous AR and which has been transduced with an adenovirus carrying the reporter gene, be developed further. This cell line has a high sensitivity with a 24-fold induction of luciferase in the presence of DHT. However, this cell line is deficient in that:

- It lacks specificity for the AR (activated by glucocorticoid (GR) and progesterone receptors (PR)).
- A 248-fold induction with dexamethasone has been reported due to the presence of GR. The presence of AR can be overcome by adding hydroxyflutamide that blocks its activity. This would entail the use of an additional set of reagents for each substance being tested to distinguish AR activity from GR activity.
- The AR in this MDA-MB-453 cell line has not been sequenced to confirm that it is intact and has no mutation.
- A central source of adenovirus, for transduction purposes, will be required by the testing laboratories because propagation of adenovirus is technically challenging.

A discussion ensued on the difficulties of recommending any of the cell lines discussed in the BRD because of their lack of sensitivity (less than 10-fold induction), lack of specificity due to the activation of the endogenous GR by the MMTV ERE (HepG2,

HeLa, CHO cells), and the activation of the AR by 17 $\beta$ -estradiol. The LNCaP cells contain a mutant AR that does not discriminate agonists from antagonists and yeasts have different metabolic and cell wall transport proteins potentially limiting "exposure." The stably transfected cell lines are unstable and require continuous selection with an antibiotic, which is costly, and by the 40<sup>th</sup> passage their sensitivity has dropped to a 5 to 6-fold induction.

Since all of the cell lines discussed in the BRD have drawbacks, the Panel recommended that ideally the chosen cell line should have the following characteristics:

- Little metabolic activity.
- An endogenous wild-type hAR (little or no PR protein; cells apparently require some low level of GR for survival).
- Adenovirus infected or stable expression of a specific ARE-Luc reporter (Use of the promoter from the C3 prostate binding protein, sex-limited protein, and probasin genes have an advantage over the MMTV promoter because of their specificity, but they are not ideal because they are less sensitive than the MMTV). At least a 20-fold induction with 0.1-1 nM R1881/DHT is needed for maximum sensitivity. Minimal agonist activity with estrogens and glucocorticoids.
- Large scale screening capability (multi-well format).
- No patent restrictions. Use of a constitutively active luciferase reporter (CMV-Luc, pSG5-Luc) to monitor cytotoxicity.
- Control to measure any direct inhibition of luciferase activity.
- A 20% inter- and intra-assay coefficient of variation.
- A cell line in which weak agonists increase induction of luciferase activity by at least two to three fold and antagonists decrease induction of the enzyme by at least 25%.

Discussion ensued regarding the potential problem associated with the presence of GR and the MMTV promoter. It was pointed out that it is unlikely that many cells would survive without glucocorticoids, so the solution to this problem would be the use of a cell with a different promoter. Discussion also ensued about the difficulties of using a yeast cell line due to the different metabolic pathways in these cells compared to mammalian cells and the transport of substances into the cells. The latter could be overcome by manipulating the permeability of the cell wall through mutagenesis of the genes coding for cell wall proteins. It was the consensus of the Panel that yeast should not be used for the assay.

## **2. *Minimum Procedural Standards for In Vitro AR TA Assays***

The Panel was in agreement with the minimal procedural standards outlined in the BRD but added that the following standards must be included for future assays.

### Reference Androgens:

- R1881 should be the reference agonist because it is not metabolized.
- 5 $\alpha$ -DHT should be included as one of the positive controls in all tests. Maximal transcriptional activity of R1881 should be obtained with a concentration of ~0.1-1 nM.

- Hydroxyflutamide should be used as the reference antagonist.
- The IC<sub>50</sub> should be ~ 500 nM with a ~70-90% inhibition occurring with 1-5 µM hydroxyflutamide.

Concentration Range of Test Substances:

- For both agonism and antagonism the limit dose should be 1nM but the solubility characteristics and potential cytotoxicity must be taken into consideration.
- Seven concentrations at log intervals should be tested.
- A measure of cell toxicity will help define the upper limit for test material concentration similar to the Maximum Tolerated Dose approach.
- Data should be expressed in relative light units (RLU) or fold induction relative to the background control (RLU for background control must be stated). A suitable nonlinear regression model such as the Hill equation must be used to estimate the potency (EC<sub>50</sub> or IC<sub>50</sub>) and slope of the dose-response curve with the calculation of a 95% confidence interval.
- Diagnostics need to be performed on the model by checking for suitability and normality of the curve. If necessary, suitable transformations need to be performed.
- For agonist or antagonist activity that does not exhibit a full dose-response, (*e.g.*, partial agonist) a trend analysis to detect a dose-response must be used. This can be followed up with confidence interval estimation at each dose level if the trend is significant. If the trend is not significant, then no further action is necessary. Significant trends imply potential activity and may be examined further.
- The biostatisticians on the Panel stated that currently insufficient data are available to address all the statistical or data evaluation issues that would enable them to recommend specific statistical analyses. Before specific statistical procedures and action levels can be identified, more details are needed about the methods and their performance criteria. It will be necessary to evaluate confidence limits, standard errors etc., to better understand the data. Different data and statistical analyses will be required depending on whether the test will be used simply as a yes/no indicator, than if the results will be used in a quantitative manner.

Assay Acceptance criteria:

- At least a 10-fold induction with the control androgen is required to ensure sensitivity to detect weakly active substances.
- The concentration of R1881 used in the antagonist assays should induce transcriptional activity ~75% of the maximal response using a concentration of ~0.1-1.0 nM R1881.
- For a substance to be classified as a positive agonist it must induce at least a 2-3 fold increase in transcriptional activity over background levels.
- For a positive antagonist response, a substance must inhibit at least 25-50% agonist-induced transcriptional activity (using concentrations of R1881 that are ~75% maximal activity).
- The inter- and intra-assay % coefficients of variation should not exceed ~20%.



Evaluation and interpretation of results:

- There should be no activation with other steroid hormones [(17 $\beta$ -estradiol, glucocorticoids (cortisol, corticosterone), progesterone] due to the presence of other receptors (GR or PR) in the cell line.
- The assays should be performed under GLPs.

Test Report:

- Information on controls for the activity of other steroid receptors and controls for cytotoxicity.
- Source of supplies (e.g., plasticware used in the assays).
- Cell passage number.
- IUPAC chemical names sufficient (structures not required).
- Solvent (justification if other than ethanol or DMSO).
- DNA isolation method (not detailed procedure).
- Name and reference for reporter vector (structure not needed).
- Justification for reference androgen only if is not R1881 or DHT.
- Statistical analysis (e.g., Hill Equation) for potency and steepness of the dose-response curve.
- Solubility information should be included in the test report.
- A description of the justification for the chemical concentrations used must be included in the report.

Additional Minimal Procedural Standards

- Serum free and phenol red free media should be used rather than charcoal stripped serum when possible based on cell viability.
- The stability of the stable cell lines must be monitored using selection media.
- The cell doubling time must be monitored.
- Cytotoxicity controls using one of the following plasmids (CMV-Luc, pSG5-Luc) must be included up to the highest dose. Cytotoxicity above 10% is not acceptable.
- Controls for direct inhibition of luciferase activity must be included.

The Panel discussed the possible methods for quantifying cytotoxicity in the assay. The approach that seemed to have the greatest promise included the measurement of the activity of the gene product of a co-transfected luciferase gene that fluoresces at a different wavelength than the *luc* reporter gene used in the same cell. CMV-driven *luc* plasmids were suggested as the carrier of the co-transfected gene, although these plasmids may be affected by some test substances, and therefore respond to non-endocrine transcription signals.

The issue of entry and exit assays was discussed. This is a measure of the concentration of the chemical in the stock solution before and after the binding or TA assay is performed. There was also the question of whether this analysis should be performed after the test chemical is added to the cells and media. The analysis would then be performed before the binding or TA assay was run and then again after the assay was run. This analysis would indicate whether the test substance was degraded during the assay's incubation, whether it was absorbed to the glassware or plasticware, and also whether the

substance was metabolized during the course of the incubation. This latter point was very important for many of the substances that are AR antagonists, since the parent compound is inactive but the metabolite is the active form. These entry and exit assays would increase the cost of performing the ER and AR binding and TA assays particularly if they were performed on the substance after it was dissolved in the media used in the assays. The Panel thought that this additional analysis and expense would place too large a burden on the laboratories running the assays.

### **3. *Recommendations for In Vitro AR TA Test Method Protocols for Validation Studies***

The Panel was of the opinion that the three test method protocols lacked sufficient detail.

The yeast-based assays are not appropriate because they:

- Cannot distinguish an agonist from an antagonist.
- Have a cell wall that affects active transport.

Transfection-based assays may not be appropriate due to patent restrictions. The adenovirus assay may be appropriate but it needs to be improved.

#### **Other Protocol Elements**

Additional information that needs to be included in the protocols are:

- Fold induction by the control androgen.
- Intra- and inter-assay coefficients of variability.
- Stability of cell responsiveness over time and passage number.
- A standardized method for comparing potencies of agonists and antagonists in the different assays.

#### **Other Available Standardized Protocols**

The Panel pointed out that the N/C interaction assay had not been mentioned in the BRD. In this assay expression vectors are made of the GAL4 and VP16 genes with the N terminal end of the AR and AR ligand binding domain. These vectors are transfected into HeLa cells which can then be used to measure TA. The advantages of this assay are that the HeLa cell line is conducive to a multi-well format, both 17 $\beta$ -estradiol and cortisol are negative in the assay, its sensitivity is significantly greater than that achieved with stable assays with a 20 fold induction in response with 0.1 nM DHT, and it has a GAL-Luc reporter with which no other steroid receptors are active. The disadvantage of the assay is that it is subject to the same patent restrictions that apply to other transient co-transfection assays that use the AR expression vector and that apply to stable cell lines with an integrated AR plasmid.

#### 4. **Recommended List of Substances to be Used for Validation in In Vitro AR TA Assays**

The Panel recommended that the following substances be included in the list for validation of the AR TA assay:

- Agonists: R1881, DHT, testosterone, androstenedione, fluoxymesterone.  
Antagonists: hydroxyflutamide, casodex (bicalutamide), cyproterone acetate, *p,p'*-DDE, linuron.  
Mixed activity: progesterone (PR agonist), medroxyprogesterone acetate (GR and PR agonist).  
No activity: dexamethasone (GR agonist), cortisol (GR agonist), 17 $\beta$ -estradiol (ER agonist)<sup>1</sup>.  
Negative controls: cycloheximide (protein synthesis inhibitor), actinomycin D (RNA synthesis inhibitor), sodium azide (cytotoxicant), specific inhibitors of luciferase activity, TPA (ligand independent activation).

The Panel recommended that heavy metals, acids and bases, insoluble solids or reactive agents, liquid and gaseous volatiles were not required for validation. However, there was a question as to whether organotin compounds are positive in the assay. A concern with testing metals is the concentration of EDTA in the assay system. There was a consensus that as long as this concentration of EDTA is kept at 1.5mM or lower there would be no problem in testing metals.

More weak compounds could be included but inactive parent compounds such as flutamide, methoxychlor, vinclozolin, and DDT should be deleted. Although the respective active intermediates of the above mentioned compounds, namely hydroxyflutamide, HPTE, the major metabolite of methoxychlor, M2, a metabolite of vinclozolin, and *p,p'*-DDE are active in the assay, only hydroxyflutamide and *p,p'*-DDE were recommended for testing because HPTE and M2 are difficult to obtain.

As mentioned previously, the working group suggested that the U.S. EPA should provide a standard set of chemicals for validation purposes.

#### **Public Comments Session (Afternoon, May 20)**

Dr. Daston asked if there were any public comments before adjournment of the meeting.

Dr. Gray (U.S. EPA). When developing or recommending an “ideal” protocol, it is important to distinguish between required and desirable features. There is a need to challenge the assays with weak agonists and antagonists. However, there are no known, weakly acting non-steroidal androgen agonists. With respect to measuring fold-induction during the TA assays, it is important to examine the variability of the response. He also

---

<sup>1</sup> 17 $\beta$ -Estradiol is listed here as having no activity in AR-TA assays despite the many reports of positive responses in the literature reviewed for the BRD. The reason for this listing, according to the Work Group members, is that it does not induce transcriptional activation *in vivo*, and the positive responses seen in the *in vitro* systems are artifacts of the recombinant systems used.

requested that the Panel expand on the list of negative compounds that they would recommend for testing in the AR TA assays so that he could challenge the two assays that he was presently evaluating in his laboratory.

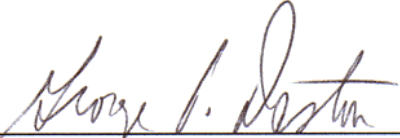
Mr. Rick Becker (American Chemical Council) echoed the request of Dr. Gray that the Panel attempt to determine which of the procedural standards were desirable and which ones were necessary in the development of these assays.

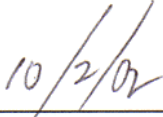
Mr. Becker also stated that the recommendations and report by the Panel is critical. It is clear that there are no validated assays and research will be needed to develop such assays. The Panel is asked that their report contain practical recommendations to help identify and provide valid assays for screening.

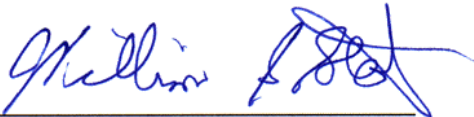
Dr. Stokes thanked the Panel on behalf of the NTP and ICCVAM for their thoughtful deliberations and careful evaluation of the test methods and background review documents. The Chair adjourned the Panel Meeting at 2:15pm.

**May 21-22, 2002**  
**Expert Panel Evaluation of the Validation Status of *In Vitro* Test Methods for  
Detecting Endocrine Disruptors: Estrogen Receptor and Androgen Receptor  
Binding and Transcriptional Activation Assays**

“These Summary Minutes have been read and approved by the Chair of the Expert Panel Evaluation of the Validation Status of *In Vitro* Test Methods for Detecting Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays, as certified below.”

  
\_\_\_\_\_  
Dr. George Daston  
Panel Chair

  
\_\_\_\_\_  
Date

  
\_\_\_\_\_  
Dr. William Stokes  
Executive Secretary

  
\_\_\_\_\_  
Date

*[This page intentionally left blank.]*