# 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

# Expert Panel Final Report

# September 2002

**Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM)**

**National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)**

**National Institute of Environmental Health Sciences**
**National Institutes of Health**
**U.S. Public Health Service**
**Department of Health and Human Services**

*[This Page Intentionally Left Blank]*

# Table of Contents

**Table of Contents** ..... **A-3**
**ED Expert Panel Roster** ..... **A-4**
**Executive Summary** ..... **A-5**

**I. *In Vitro* Estrogen Receptor (ER) Binding Assays** ..... **A-17**
1.0 Recommendations and Priority for Validation Studies ..... A-17
2.0 Minimum Procedural Standards for *In Vitro* ER Binding Assays ..... A-18
3.0 Recommendations for *In Vitro* ER Binding Test Method Protocols for Validation Studies ..... A-23
4.0 Recommended List of Substances to be Used for Validation of *In Vitro* ER Binding Assays ..... A-24
References ..... A-25
Appendix 1: Ligand Titration Assay ..... A-26
Appendix 2: Ligand Competition Array ..... A-29

**II. *In Vitro* Estrogen Receptor (ER) Transcriptional Activation (TA) Assays** ..... **A-34**
1.0 Recommendations and Priority for Validation Studies ..... A-34
2.0 Minimum Procedural Standards for *In Vitro* ER TA Assays ..... A-36
3.0 Recommendations for *In Vitro* ER TA Test Method Protocols for Validation Studies ..... A-41
4.0 Recommended List of Substances to be Used for Validation of *In Vitro* ER TA Assays ..... A-43
Table 1: Substances in Common Between *In Vitro* ER Binding Assays and *In Vitro* ER TA Agonism and Antagonism Assays ..... A-47
Additional Panel Recommendations ..... A-48

**III. *In Vitro* Androgen Receptor (AR) Binding Assays** ..... **A-50**
1.0 Recommendations and Priority for Validation Studies ..... A-50
2.0 Minimum Procedural Standards for *In Vitro* AR Binding Assays ..... A-54
3.0 Recommendations for *In Vitro* AR Binding Test Method Protocols for Validation Studies ..... A-58
4.0 Recommended List of Substances to be Used for Validation of *In Vitro* AR Binding Assays ..... A-60

**IV. *In Vitro* Androgen Receptor (AR) Transcriptional Activation (TA) Assays** ..... **A-62**
1.0 Recommendations and Priority for Validation Studies ..... A-62
2.0 Minimum Procedural Standards for *In Vitro* AR TA Assays ..... A-75
3.0 Recommendations for *In Vitro* AR TA Test Method Protocols for Validation Studies ..... A-80
4.0 Recommended List of Substances to be Used for Validation of *In Vitro* AR TA Assays ..... A-82
Summary ..... A-84
Additional Comments on the BRD ..... A-85
References ..... A-88

# ED Expert Panel Roster

- **George Daston, Ph.D.**, (Panel Chair), Research Fellow, Miami Valley Laboratories, Procter & Gamble, Cincinnati, OH
- **Nira Ben-Jonathan, Ph.D.**, Professor of Cell Biology, Neurobiology and Anatomy, University of Cincinnati Medical School, Cincinnati, OH
- **Terry Brown, Ph.D.**, Professor, School of Public Health, Johns Hopkins University, 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, Dept. of Environmental Medicine, University of Rochester School of 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.**, Biostatistics Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC
- **Walter Piegorsch, Ph.D.**, Professor, Dept. of Statistics, University of South Carolina, Columbia, SC
- **Bernard Robaire, Ph.D.**, Professor, Dept. of Pharmacology & Therapeutics, McGill University, Montreal, Quebec, Canada
- **Stephen Safe, Ph.D.**, Professor, Dept. of Veterinary Physiology & Pharmacology, Texas A & M University, College Station, TX
- **John Stegeman, Ph.D.**, Senior Scientist, Chairman, 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, Environmental Health Science, Tulane University, New Orleans, LA (not present at the meeting)
- **Elizabeth M. 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.**, Senior Associate Dean of Academic Affairs, Professor, Environmental Health Sciences, School of Public Health, Johns Hopkins University, Baltimore, MD
- **Tim Zacharewski, Ph.D.**, Associate Professor, Dept. of Biochemistry & Molecular Biology, Michigan State University, East Lansing, MI

# EXECUTIVE SUMMARY

## Introduction

In response to public concern that pesticides may interfere with endocrine processes in humans and wildlife, Congress in 1996 directed the U.S. EPA, through the Food Quality Protection Act (FQPA) (Public Law 104-170), to develop a screening program for evaluating the potential of pesticides and other substances to induce hormone-related health effects. In 1998, the U.S. EPA Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) issued a report recommending that the agency evaluate the effects of these substances on both human and ecological (wildlife) health, and that a screening and testing program be implemented for identifying endocrine disruptors. In response, the U.S. EPA proposed the Endocrine Disruptor Screening Program (EDSP) (*Federal Register*, Vol. 63, No. 248, pp. 71541-71568, December 28, 1998, available at *http://www.epa.gov/fedrgstr/EPA-TOX/1998/December/Day-28/t34298.htm)*. The proposed EDSP consists of a Tier 1 screening battery of tests that is designed to identify substances capable of interacting with the endocrine system, and different Tier 2 testing assays that are designed to confirm and extend the Tier 1 results. If, based on a weight of evidence evaluation of the results from the Tier l screening battery, the test substance is identified as a potential endocrine disruptor, Tier 2 *in vivo* tests are conducted to provide detailed information on concentration response relationships and specific abnormal effects that may result. The proposed Tier 1 *in vitro* assays include estrogen receptor (ER) and androgen receptor (AR) assays. Currently, the U.S. EPA proposes that either a binding assay or a transcriptional activation (TA) assay be used. These *in vitro* assays are relevant for screening purposes because they might identify substances that alter natural endocrine processes by binding with estrogen and/or androgen receptors, resulting in agonist and/or antagonist activity.

To assess the current validation status of these *in vitro* methods, the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), which provides operational support to the Interagency Coordinating Committee for the Validation of Alternative Methods (ICCVAM), prepared Background Review Documents (BRD) on:

- *In vitro* ER binding assays
- *In vitro* ER TA assays
- *In vitro* AR binding assays
- *In vitro* AR TA assays

As part of the ICCVAM evaluation, the U.S. EPA asked for development of minimum performance criteria that could be used to define acceptable *in vitro* ER/AR binding and TA assays. It was envisioned that these criteria would be based on the performance of existing standardized *in vitro* ER binding assays, and would be used to assess the acceptability of future new or revised assays.

An extensive literature search for relevant publications was conducted and a formal request through the U.S. *Federal Register* (Vol. 66, No. 57, pp. 16278-16279, March 23, 2000) was made for data and information from completed, ongoing, or planned studies using or evaluating ER/AR binding and TA assays. BRDs, prepared for each of the four types of assays, are included:

- A review of the different protocols used for each type of assay

- A review of the critical procedural components for each type of assay and proposed minimum procedural standards
- A prioritized list of assays recommended for validation
- A list of substances proposed for future validation studies

During development of the BRDs, ICCVAM and NICEATM determined that no validation studies have been completed on the assays being considered. With agreement from the U.S. EPA, NICEATM and ICCVAM decided to proceed with an expert panel evaluation of the current status of ER/AR binding and TA assays and the development of recommendations for their future validation.

**ICCVAM Expert Panel Review Meeting, May 21-22, 2002**

An Expert Panel meeting, sponsored by the National Institute of Environmental Health Sciences (NIEHS) and the NTP and organized by NICEATM in collaboration with ICCVAM, was held on May 21-22, 2002 at the Sheraton Imperial Hotel in Research Triangle Park, NC. The Panel was charged with assessing the current validation status of these four types of *in vitro* endocrine disruptor screening methods and to develop 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.

***In Vitro* ER Binding Assays**

The Panel reviewed 14 different *in vitro* ER binding assays in which 635 different substances had been tested in one or more of the assays. The sources of ER used in these assays included cytosol prepared from MCF-7 cells and from the uteri of mice, rats, and rabbits; intact MCF-7 cells; purified recombinant human ERα and ERβ, and fusion proteins between glutathione-S-transferase and the binding domains of the human ERα, mouse ER, chicken ER, anole ER, and rainbow trout ER.

The Panel agreed that the *in vitro* ER binding assays considered in the BRD still required standardization and that the available data were not adequate to assess the validation status of the test methods. The Panel recommended that test methods using recombinant ER receptors (both α and β subtypes) should be given the highest priority for further assay standardization and validation. Either human or rat receptors were considered acceptable. For screening for possible ecological effects, recombinant receptors from wildlife were considered to be potentially more relevant and should be evaluated.

There was consensus among the Panel on the adequacy of the following minimum procedural standards described in the *In Vitro* ER Binding BRD:

- All studies utilizing animals should be approved by the Institutional Animal Care and Use Committee (IACUC) or its equivalent.
- The dissociation constant ($K_d$) of the reference estrogen must be determined with each set of assays.

- The reference estrogen for the assays should be 17β-estradiol.
- Test substances should be prepared in water or 95-100% ethanol but dimethyl sulfoxide (DMSO) could be used, if necessary. A set of solvent/vehicle-only controls (with final solvent concentrations identical to those used in the reaction mixtures containing the test substance) must be included in each set of assays.
- The limit concentration should be 1 mM but the solubility characteristics of each test substance must be taken into consideration. In addition, possible denaturation of the receptor at high test substance concentrations, and for some substances (e.g., surfactants), at concentrations below 1 mM, need to be considered when the results of positive assays are interpreted.
- The concentration range of test substances should consist of at least seven different concentrations spaced at one order of magnitude apart from each other (e.g., 1, 10, 100 nM, 1, 10, 100 μM, 1 mM). However, if a lower maximum concentration is tested due to solubility constraints or excessive toxicity, the number of concentrations tested can be adjusted to account for the reduced concentration range.
- The control solvent/vehicle volume must be the same as that used in the reaction mixtures containing the test substances and should remain constant throughout the concentration range tested.
- A substance (e.g., tamoxifen, coumestrol, estriol) with a binding affinity of two or three orders of magnitude below that of 17β-estradiol should be used as a concurrent positive control.
- Triplicate measurements should be performed at each concentration (negative and positive controls, reference estrogen, test substance) tested.
- For data analysis, it is essential that both the $B_{max}$ (number of binding sites or specific binding capacity) and the $K_d$ values be computed.
- For an assay to be acceptable, the reference estrogen/positive control responses must be acceptable, based on historical data.
- If an $IC_{50}$ cannot be obtained after testing to the limit dose or the highest dose possible, the test substance is usually classified as being “negative” for *in vitro* ER binding. However, when test substances induce a significant reduction in binding, but do not achieve at least a 50% reduction in the binding of the reference estrogen to the ER, such responses should be noted and the substances classified as "equivocal" until additional information becomes available about the significance of this category of dose response curves.
- The test report should include information on the test substance, the solvent used, the type and source of the ER, the test conditions, the results, a discussion of the results, and a determination as to whether the substance is positive or negative.
- Replicate studies are not mandated but questionable data needs to be confirmed by re-testing the substance.
- The assays should be performed following Good Laboratory Practice guidelines.

Additional or modified minimum procedural standards that were recommended by the Panel included:

- Classification of a test substance as “positive” for binding should be based on the use of statistical inferences pertinent to the characteristics of the assay. The state of the art for making statistical inferences with endocrine disruptor data requires more detailed research and study.

- Hexa-tritium labeled 17β-estradiol (i.e., [2,3,6,7,16,17-$^3$H] 17β-estradiol) should be used as the ligand for all assays.
- In situations where more than one solvent could be used, preference should be given to water, then ethanol, followed by DMSO.
- The type of protein assay and the concentration of the protein used in the assays must be specified in the study report.
- Sodium molybdate and a cocktail of protease inhibitors must be added to the assay to protect the ER from degradation.
- The use of dextran-coated charcoal was recommended as the most suitable method for separating bound from free labeled 17β-estradiol.
- For data analysis, the use of a ligand titration array (Raffelsberger and Wittliff, 1997) would provide simultaneous evaluation of a laboratory's performance and determination of the estrogen binding properties (e.g., $IC_{50}$, $K_d$, and $K_i$ values) of both the reference estrogen and the test substances.

The Panel concluded that each of the binding parameters (i.e., $B_{max}$, $K_d$, $K_i$ values) of the reference estrogen and test substances should be measured in each ligand titration/binding assay. Since the binding of a test substance to the receptor is often not directly proportional to the concentration of the substance being tested, a non-linear response is often obtained. Thus, it seemed appropriate that a non-linear statistical model would be the best approach for the analysis of this data. However, the Panel concluded that a decision on this statistical approach needed further evaluation before the most appropriate statistical method could be identified. This evaluation would be facilitated by the collection of data generated by different laboratories using the same defined protocol and testing the same set of substances.

The Panel did not recommend the inclusion of a metabolic activation system in *in vitro* ER binding assays.

The Panel was comfortable with the "Example Protocol for the *In Vitro* Estrogen Receptor (ER) Competitive Binding Assay using Rat Uterine Cytosol (RUC)" included in the ER Binding BRD (Section 12, Annex - based on the U.S. EPA RUC Protocol, 2002), provided that it is amended to include the additional details presented in the discussion on minimum procedural standards in the BRD and the points discussed in the ER Binding Group report responses to Questions 1 and 2. This amended protocol can serve as a prototype for developing protocols for other ER binding assays, such as those using a purified ER protein.

The Panel endorsed the 33 substances recommended in the *In Vitro* ER Binding BRD for the validation of *in vitro* ER binding assays. While this list provided substances across the range of binding activities, the Panel recommended that the list be reviewed to ensure that it represents the diversity of chemical classes and the range of potencies that the U.S. EPA is interested in screening. The Panel recommended that the proportion of negative ER binding substances in the list should be increased to at least 25% to enable the specificity of the assay to be accurately determined.

In regard to future validation studies, the Panel concluded that both inter- and intra-laboratory reproducibility must be assessed using the same reference ER preparation and identical set of test substances.

***In Vitro* ER TA Assays**

The Panel reviewed 95 different ER TA assays (63 mammalian reporter gene assays, 10 mammalian proliferation assays, 22 yeast strain reporter gene assays) in which 698 different substances had been tested in one or more of the assays. The source of the ER used in these assays included unspecified ER from human, mouse, and rat; or ERα and ERβ subtypes found endogenously or transiently/stably transfected into various cell lines. The luciferase and chloramphenicol acetyltransferase reporter genes were used in the mammalian cell line assays and the β-galactosidase reporter gene was used in the yeast strain assays.

The Panel, while indicating that an assay using a stably transfected cell line appeared to be more amenable to high throughput screening, was not convinced that such an assay was the most appropriate. Stable cell lines lose their stability over time, are limited in availability, and are difficult to isolate. To resolve this issue, the Panel recommended that a comparative study be conducted in which the response of cell lines transiently or stably transfected with the same ER receptor/reporter gene constructs be compared using a selected set of test substances. A third cell line expressing an endogenous ER and transfected with the same reporter construct should be included in this study.

There was consensus among the Panel on the adequacy of the following minimum procedural standards provided in the *In Vitro* ER TA BRD:

- The reference estrogen should be 17β-estradiol; the transcriptional activation response with this substance must be demonstrated by a full concentration response curve.
- Test substances should be prepared preferably in absolute ethanol or culture medium but DMSO could be used, if necessary. A set of solvent/vehicle-only controls (with final solvent concentrations identical to those used in the reaction mixtures containing the test substance) must be included in each set of assays.
- The solvent/vehicle volume must be the same as that used in the reaction mixtures containing the test substance, and should remain constant throughout the concentration range tested.
- A relatively active antagonist (e.g., ICI 182,780) should be used as the positive control for antagonist studies.
- The limit concentration should be 1 mM but the solubility characteristics of each test substance must be taken into consideration. (One Panel member felt that this concentration was excessive and that 0.1 mM should be adequate).
- The concentration range of test substances should consist of at least seven different concentrations spaced at one order of magnitude apart from each other (e.g., 1, 10, 100 nM, 1, 10, 100 μM, 1 mM). However, if a lower maximum concentration is tested due to solubility constraints or excessive cytotoxicity, the number of concentrations tested can be reduced to account for the altered concentration range. At least one Panel member suggested that five concentration levels would be adequate.
- Triplicate measurements should be performed at each test substance concentration level.

- Classification of a test substance as "positive" for agonist or antagonist activity in transcriptional activation assays should be based on the generation of a concentration response curve.
- Historical data should be used as part of the assay acceptance criteria (i.e., reference substances for agonism and antagonism must give appropriate responses).
- The test report should include information on the test substance, the solvent used, the ER, the reporter plasmid, the cell line, the test conditions, the results, a discussion of the results, and a determination as to whether the substance is positive or negative.
- Replicate studies are not mandated but questionable data needs to be confirmed by re-testing of the substance. However, one Panel member recommended testing each substance at least three times in different experiments.
- The assays should be performed following Good Laboratory Practice guidelines.

Additional or modified minimum procedural standards that were recommended by the Panel included:

- Concentrations to be tested for the reference positive control, 17β-estradiol, should range from 1 pM to 1 μM, and clear guidance is needed about the expected response.
- With regard to the preparation of test substances, the level of solvent that does not adversely affect assay response should be determined before testing by performing appropriate pre-validation studies using the reference estrogen.
- A relatively weak estrogenic agonist (e.g., estriol) should be included as an additional control for agonist studies.
- A measure of cellular cytotoxicity should be incorporated into the assay to help define the upper limit for test material concentrations, similar to the Maximum Tolerated Dose (MTD) approach used in *in vivo* studies.
- To ensure that a positive agonist response reflects a receptor-mediated activity, the test substance could be re-tested with ICI 182,780 (the candidate ER antagonist) present in the culture medium.
- For transient transfection assays, there is a need to include a constitutive reporter gene assay to control for transfection efficiency.
- A standard definition for "relative activity" must be decided upon so that a positive/ negative call for agonism and antagonism can be made.
- Suitable diagnostics must be performed on any statistical procedure to ensure that the model fits the data before it is finally chosen for analysis of the data.
- The test report should also include the complete DNA sequences of constructs and vectors, the transfection methods used, the cell passage number(s) during the study, and the $CO_2$% level in the incubator.

The Panel did not recommend the inclusion of an exogenous metabolic activation system in *in vitro* ER TA studies at this time. However, the Panel recommended that available information on the Phase I/Phase II metabolic capabilities of the cell lines employed in validation studies, as well as available information on the metabolism of the validation chemicals be compiled. The Panel concluded that studies should be conducted to obtain information on the importance of metabolism for systems ultimately employed in screening.

The Panel concluded that the protocols provided in the *In Vitro* ER TA BRD were adequate, provided that the minimum procedural standards are included, and that a laboratory with cell or yeast culture expertise should be able to perform the assays. However, there are a number of issues relating to standardization that will have to be added to the protocols.

In terms of the list of reference substances recommended in the *In Vitro* ER TA BRD for future *in vitro* ER TA assay validation studies, the Panel recommended that selection be based on solid scientific rationale such as a combination of existing ER binding, antagonist, or agonist data. To the extent possible, all the same reference substances should be used for validation of both the *in vitro* ER binding and the *in vitro* ER TA assays. The Panel also recommended that the proportion of negative substances be increased to more effectively evaluate the specificity of these assays. During development of the final list, consideration should be given to substances selected by the U.S. EPA and the Organisation for Economic Co-operation and Development (OECD) for validating *in vivo* endocrine disruptor assays. Also, the Panel encouraged the development of a centralized repository of chemicals with verified purity for future validation studies.

The Panel recommended that a sequential testing strategy be evaluated for utility during the pre-validation of *in vitro* ER/AR binding and TA agonism/antagonism assays. In this approach, if a substance induces a positive response in any assay, then testing in any of the other binding/TA assays would not need to be conducted. In support of this strategy, the Panel concluded that further classification of the activity of a positive test substance using additional binding/TA endpoints would provide little additional information that would assist with prioritization and the design of subsequent *in vivo* studies.

***In Vitro* AR Binding Assays**

The Panel reviewed 11 different *in vitro* AR binding assays in which 109 different substances had been tested in one or more of the assays. The sources of AR used in these assays included cytosol from calf uterus, rat epididymis and prostate, and MCF-7 cells; rat epididymal nuclear fraction; COS-1 cells transiently transfected with a human AR; LnCaP cells and human genital fibroblasts with an endogenous AR; and semi-purified recombinant human AR.

The Panel acknowledged the lack of an existing acceptable standardized *in vitro* AR binding assay protocol, and that the published data were inadequate for assessing the reliability and comparative performance of these assays. Based on the available data, the Panel recommended that the highest priority for future efforts is the development of a high-throughput assay using a purified, recombinant full-length AR protein. A truncated AR protein (consisting of the AR ligand binding domain) has been purified but this protein is less desirable than the full-length protein because it appears that domains other than the actual binding domain modulate the binding of substances to the AR. However, in contrast to the human ER protein, the commercial availability of the intact AR cDNA is apparently limited due to restrictions by the exclusive license of the patent. The Panel recommended that the appropriate government agencies investigate the status of patents and licenses attendant to the use of the human and rat AR and provide guidance as to how the scientific community should proceed with the development of such AR assays. An alternative approach might be to proceed with a truncated AR protein that is commercially available but that has not yet been evaluated for sensitivity and reliability.

There was consensus among the Panel on the adequacy of the following minimum procedural standards provided in the *In Vitro* AR Binding BRD:

- All studies utilizing animals should be approved by the IACUC or its equivalent.
- The $K_d$ of the reference androgen must be determined with each set of assays.
- Test substances should be prepared preferably in water, and then absolute ethanol but DMSO can be used, if necessary. A set of solvent/vehicle-only controls (with final solvent concentrations identical to those used in the reaction mixtures containing the test substance) must be included in each set of assays.
- The control solvent/vehicle volume must be the same as that used in the reaction mixtures containing the test substance, and should remain constant throughout the concentration range tested.
- The limit concentration should be 1 mM but the solubility characteristics of each test substance must be taken into consideration. In addition, possible denaturation of the receptor at high test substance concentrations, and for some substances (e.g., surfactants), at concentrations below 1 mM, need to be considered when the results of positive assays are interpreted.
- The concentration range of test substances must consist of at least seven different concentrations spaced at one order of magnitude apart from each other (e.g., 1, 10, 100 nM, 1, 10, 100 μM, 1 mM). However, if a lower maximum concentration is tested due to solubility constraints or excessive cytotoxicity, the number of concentrations tested can be reduced to account for the altered concentration range.
- A substance (e.g., cyproterone acetate) with a binding affinity of two or three orders of magnitude below that of the reference androgen should be used as a positive control.
- Triplicate measurements should be performed at each concentration level.
- For data analysis, it is essential to compute both the $B_{max}$ and the $K_d$ values.
- If an $IC_{50}$ cannot be obtained after testing to the limit dose or the highest dose possible, the test substance is usually classified as being “negative” for *in vitro* AR binding. However, when test substances induce a significant reduction in binding, but do not achieve at least a 50% reduction in the binding of the reference androgen to the AR, such responses should be noted and the substances classified as "equivocal" until additional information becomes available about the significance of this category of dose response curves.
- The test report should include information on the test substance, the solvent used, the type and source of the AR, the test conditions, the results, a discussion of the results, and a determination as to whether the substance is positive or negative.
- Replicate studies are not mandated but questionable data needs to be confirmed by re-testing of the substance.
- The assays should be performed following Good Laboratory Practice guidelines.

Additional or modified minimum procedural standards that were recommended by the Panel included:

- Classification of a test substance as “positive” for binding should be based on the use of statistical inferences pertinent to the characteristics of the assay. The state of the art for making statistical inferences with endocrine disruptor data requires more detailed research and study.

- The endogenous ligand, 5α-dihydrotestosterone (DHT), was recommended as the reference androgen for recombinant protein-based assays where metabolism of DHT would not occur. Due to its high affinity, lack of metabolism, and low non-specific protein binding, R1881 was recommended as the reference androgen for most other assays. However, as R1881 binds to the progesterone receptor (PR), binding assays based on cells or tissues that contain this receptor should include triamcinolone acetonide to block its binding to the PR. Alternatively, mibolerone, which has a low affinity for PR, was considered appropriate as the reference androgen for such assays.
- An additional positive control substance with a binding affinity within two orders of magnitude of the limit of sensitivity of the assay should be included also.
- The dextran-coated charcoal procedure should be used for the separation of free and bound ligand.
- The type of protein assay and the concentration of the protein used in the assays must be specified.
- Sodium molybdate and a cocktail of protease inhibitors must be added to the assay to protect the AR from degradation.
- For data analysis, the use of a ligand titration array (Raffelsberger and Wittliff, 1997) would provide simultaneous evaluation of a laboratory's performance and determination of the androgen binding properties (e.g., $IC_{50}$, $K_d$, and $K_i$ values) of both the reference androgen and the test substances.
- The study report should include information on the chemical and radiochemical purity of the radiolabeled androgen, as well as information on the assay used for protein determination.

The addition of an exogenous metabolic activation system was not recommended for current use by the Panel in *in vitro* AR binding assays.

For the same reasons described for the *in vitro* ER binding assays, the Panel concluded that the statistical approaches for the analysis of data generated using *in vitro* AR binding assays required further investigation. The Panel recommended that prevalidation studies be conducted to evaluate an *in vitro* AR binding assay using purified AR. Data generated from these prevalidation studies could be used by the biostatisticians to develop the most reliable and robust statistical models for data analysis.

The Panel concluded also that, although the rat prostate cytosol (RPC) protocol was sufficiently detailed, this assay should not be a priority for further validation. Rather, the Panel was of the opinion that the simplest and most preferred assay would be one in which purified AR is fixed to multiwell plates. One commercial source of the AR (PanVera Corporation, Madison, WI, USA) is available, but this AR is a truncated protein and has not been evaluated for sensitivity and reliability.

The Panel recommended that the same reference substances should be used for validation of both *in vitro* AR binding and *in vitro* AR TA assays. Furthermore, the Panel recommended the inclusion of additional weakly positive reference substances representing the range of possible environmental exposures and an increase in the proportion of negative substances. The Panel also recommended that bicalutamide, a substance that binds to the AR but does not activate its

transcription, and finasteride, a commercially available 5α-reductase inhibitor which does not bind to the AR, be included as additional assay controls.

In regard to future validation studies, the Panel concluded that both inter- and intra-laboratory reproducibility must be assessed for the same reference AR preparation using an identical set of test substances.

***In Vitro* AR TA Assays**

The Panel reviewed 17 different AR TA assays (15 mammalian reporter gene assays, 1 mammalian proliferation assay, 1 yeast strain reporter gene assay) in which 147 different substances had been tested in one or more of the assays. The source of the AR used in these assays included AR from the human, mouse, and rat. The luciferase and chloramphenicol acetyltransferase reporter genes were used in the mammalian cell line assays and the β-galactosidase reporter gene was used in the yeast strain assay.

It was the consensus of the Panel that no current *in vitro* AR TA protocol was optimal for assessing AR agonist and antagonist activities. However, the Panel concluded that one cell line, described in the *In Vitro* AR TA BRD, containing an endogenous AR and transduced with an adenovirus containing the reporter gene was the most promising approach for development of an assay to assess AR agonist and antagonist activity. The adenovirus method is straightforward and avoids time consuming procedures associated with transient transfection methodology. Important additional developments that are needed include the identification and use of a cell line that lacks high response levels of the glucocorticoid and progesterone receptors, and the use of a reporter vector that shows greater specificity for the AR.

There was consensus on the adequacy of the following minimum procedural standards provided in the *In Vitro* AR TA BRD:

- The transcriptional activation of the reference androgen must be demonstrated by a full concentration response curve.
- The reference androgen should be R1881.
- The active antagonist hydroxyflutamide was recommended as a positive control for antagonist studies.
- Test substances should be prepared preferably in water or absolute ethanol but DMSO could be used, if necessary. A set of solvent/vehicle-only controls (with solvent concentrations identical to those used with reaction mixtures containing test substances) must be performed in each set of assays.
- The solvent/vehicle volume must be the same as that used in the reaction mixtures containing the test substances.
- An androgen that is two orders of magnitude less potent than R1881 should be used as a positive control.
- The limit concentration should be 1 mM but the solubility characteristics of each test substance must be taken into consideration. Concern was expressed by some Panel members that this concentration might be excessive and that 30 μM would be adequate.
- The concentration range of test substances must consist of at least seven different concentrations spaced at one order of magnitude apart from each other (e.g., 1, 10, 100 nM,

1, 10, 100 μM, 1 mM). However, if a lower maximum concentration is tested due to solubility constraints or excessive cytotoxicity, the number of concentrations tested can be reduced to account for the altered concentration range.

- Triplicate measurements should be performed at each test substance concentration level.
- Classification of a test substance as "positive" for agonist or antagonist activity in transcriptional activation assays should be based on the generation of a concentration response curve.
- Reference substances for agonism and antagonism assays must give appropriate responses based on historical data.
- The test report should include information on the test substance, the solvent used, the AR, the reporter plasmid, the cell line, the test conditions, the results, a discussion of the results, and a determination as to whether the substance is positive or negative.
- Replicate studies are not mandated but questionable data needs to be confirmed by re-testing of the substance.
- The assays should be performed following Good Laboratory Practice guidelines.

Additional or modified minimum procedural standards recommended by the Panel included:

- Serum free and phenol red free media should be used rather than charcoal stripped serum.
- The stability of the cell lines must be monitored using selection media.
- DHT should be used as an additional reference androgen.
- Levels of cytotoxicity must be evaluated in each assay.
- A nonselective reporter (e.g., MMTV) should be used.
- A suitable nonlinear regression model such as the Hill equation must be used to estimate the potency ($EC_{50}$ or $IC_{50}$ values) and slope of the concentration-response curve with a 95% confidence interval.
- Statistical analysis of the data should indicate agonist and inhibitory test chemical effects on transcription that meet or exceed the 2-3 fold induction or 50% inhibition level compared to the respective controls.
- Diagnostics need to be performed on the model by checking for suitability and normality of the curve. If necessary, suitable data transformations need to be performed.
- The test report should also include the type of method used for isolating the DNA for making constructs; the cell passage number; the volume of the test substance applied to the test mixture; information on controls for the activity of other steroid receptors and controls for cytotoxicity; source of plasticware and other materials used in the assay; entry and exit analytical assay results for all test article and control compounds; and the response in absolute units such as light units for luciferase activity with the error indicated and as fold induction, if this is deemed appropriate.

In terms of statistical analysis, the Panel concluded that nonlinear statistical models (e.g., the Hill equation) appear to be the most useful models for estimating the potency and the slope of the concentration-response curve for agonists and antagonists.

The Panel did not recommend the inclusion of a metabolic activation system for *in vitro* AR TA assays.

In terms of the list of substances recommended in the *In Vitro* AR TA BRD for future *in vitro* AR TA assay validation studies, the Panel identified a more limited list of potential candidate substances for use in prevalidation studies. This list included substances that could affect luciferase reporter gene transcription activity independent of the AR (e.g., by inhibiting RNA or protein synthesis).

Although recognizing that these *in vitro* endocrine disruptor assays are proposed as components of a screening test battery of test methods where the results will be used in making weight of evidence decisions, the Panel recommended determination of the predictive value of these assays for estimating *in vivo* responses. Therefore, the Panel recommended that substances proposed for validation of the *in vivo* test methods should also be evaluated in the *in vitro* assays included in the screening battery and, to the extent possible, and vice-versa.

The Panel encouraged the development of a centralized repository of substances with verified purity that could be distributed to laboratories developing or conducting validation studies. The purpose of this repository is to ensure the comparability of data generated during the validation of the different *in vitro* assays and to provide a source of coded samples for validation.

# I. *In Vitro* Estrogen Receptor (ER) Binding Assays

## 1.0 Recommendations and Priority for Validation Studies

The *In Vitro* ER Binding BRD reviews the comparative performance, reliability, advantages, and disadvantages for different *in vitro* ER binding assays, and recommends a relative priority for further development and/or validation based on this information.

**1.1 Considering the intended use of the assays as a toxicological screen, is the Panel aware of other advantages and disadvantages for the assays discussed in the BRD?**

The Panel did not discuss any additional advantages and disadvantages not covered in the BRD.

**1.2 Considering the intended use of the assays as a toxicological screen, does the Panel agree with the relative priority recommended for these sets of assays? Does the Panel recommend any changes in priority, or have specific recommendations for prioritization?**

The Panel agrees with the BRD's conclusions that assays using recombinant receptors (both subtypes) should be given the highest priority for assay standardization and validation. Human or rat receptors would be acceptable. For screening for possible effects in wildlife, recombinant receptors from other species might be more relevant. The Panel believes that a consistent, standardized preparation of the receptor is essential for quality control and in making valid comparisons across laboratories and experiments. The fluorescent polarization assay has not been in wide use and there are only limited data for comparison. Availability of specialized equipment and reagents is also of concern.

**1.2.1 Is rat uterine cytosol the best source of estrogen receptors for the binding assays?**

The rat uterine cytosol (RUC) is not considered the best source of ER for the ligand binding assay. A standardized preparation of the ER is of the utmost importance for quality control and comparison of results across laboratories.

**1.2.2 Should the binding of substances to different receptor subtypes be addressed in the binding assays?**

The use of either human or rat recombinant proteins, both $\alpha$ and $\beta$, is a high priority. Recombinant receptors from other species are recommended for screening substances that pose particular hazard to wildlife.

**1.2.3 Should a metabolic activation system be included in the binding assays?**

The inclusion of a metabolic activation system in *in vitro* ER binding assays is not recommended at this time. The type of metabolic activation system developed will depend on which *in vitro* assay(s) are considered validated for detecting endocrine disrupting substances. The Panel recommends, while validation is being conducted, that available information on the metabolism of the validation chemicals be compiled, including the degree to which metabolism is known to alter estrogenic activity. Once the importance of metabolic activation in the ability of substances to disrupt endocrine function has been demonstrated and validated in *in vitro* ER binding assays, appropriate methods for including metabolic activation in the assays can be developed and validated.

## 2.0 Minimum Procedural Standards for *In Vitro* ER Binding Assays

**2.1 To facilitate assay standardization, the BRD proposes minimum procedural standards that should be incorporated into *In Vitro* ER Binding Assay protocols (Section 12.2). Does the Panel agree with the adequacy of the proposed procedural standards? If not, what changes should be made to each standard and why?**

The Panel agrees with the critical methodological issues proposed in the BRD, and endorsed the fact that any assays using animals must be undertaken under the guidance of the relevant Institutional Animal Care and Use Committee (IACUC).

### 2.1.1 Binding Constant ($K_d$) of the Reference Estrogen

The Panel agrees that the dissociation constant must be determined with each set of assays and that 17β-estradiol should be used as the reference estrogen. Furthermore, the Panel recommends that the hexa-tritium labeled 17β-estradiol (i.e., [2,3,6,7,16,17-$^3$H] 17β-estradiol) be used as the ligand for all assays because it is the most potent naturally occurring estrogen in the human body, and because of the high specific radioactivity available commercially, which increases the sensitivity of both the ligand titration assay and the ligand competition assay considerably. Furthermore, there was consensus that recombinant ER preparations, particularly human ER, be employed in the validation and screening assays.

The Panel agrees that 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, to demonstrate assay validation and transferability across laboratories. The ability of a laboratory operator to achieve a specific binding capacity and $K_d$ value for 17β-estradiol of a reference receptor protein within accepted limits for that type of preparation is a critical measure of the robustness of the procedure at that institution. These data will be essential to the establishment of a Quality Assurance Program (Assay Proficiency) for the evaluation of endocrine disruptor substances at numerous laboratories. Also, the Panel recommends replicate determinations and replicate assays, and the establishment of decision criteria for evaluating inter- and intra-laboratory reproducibility across the range of chemicals selected for testing. Attention should be given to the level of nonspecific binding in the ER preparation selected.

The Panel recommends that straightforward procedures for determination of both the $K_d$ value of the radio-labeled reference ligand ([2,3,6,7,16,17-$^3$H] 17β-estradiol), and the $K_d$ value of an unlabeled test substance, such as the ligand titration array (Raffelsberger and Wittliff, 1997), be considered. Software programs such as Compete® (Lundon Software, Inc., Cleveland Heights, OH), GraphPad Prism® (GraphPad Software, Inc., San Diego, CA), LIGAND (Munson and Rodbard, 1980) and OneSite® (Lundon Software, Inc., Cleveland Heights, OH), will determine the specific binding capacity and $K_d$ values of the reference ER preparation, and analyze the competition results and compute the $K_d$ value of the unlabeled putative estrogen mimic examined in the assay. **Appendices 1** and **2** provide an example of the curves

that might be generated using this approach. As shown in **Appendix 1**, the $K_d$ value for the reference estrogen ([2,3,6,7,16,17-$^3$H]17β-estradiol), determined by Ligand Titration Assay, is computed according to various models. This $K_d$ value is used in the calculations of the Ligand Competition Array (**Appendix 2**) to arrive at an apparent dissociation constant ($K_d$ value) of the unlabeled test substance (in this example, estrone) for the reference receptor preparation (in this example, recombinant ERα). Note the good agreement in the $K_d$ values computed for unlabeled estrone (4.0 versus 2.6 x $10^{-9}$ M) with the two models and with the $K_i$ value (1.96 x $10^{-9}$ M). The $IC_{50}$ value is also provided. As noted below, however, more study is necessary for specifying the precise statistical characteristics of ER binding data when fitting nonlinear regression curves and estimating pertinent parameters such as $K_d$ and the $IC_{50}$.

### 2.1.2 Reference Estrogen

The Panel agrees that only the native estrogen ligand, 17β-estradiol, should be employed as the reference estrogen, for the reasons indicated.

### 2.1.3 Preparation of Test Substances

The Panel agrees with the BRD that test substances should be prepared in water, 95-100% ethanol, or dimethyl sulfoxide (DMSO), depending upon their solubility (**Section 12.2.3** of the BRD). The Panel recommends that preference should be given to 95-100% ethanol and that a set of solvent (vehicle)-only controls (with solvent concentrations identical to those used with reactions containing test substances) must be performed in each set of assays. It is known that sex-hormone receptor preparations from various tissue origins, including recombinant expression systems, exhibit different performance characteristics in the presence of the same solvent, again emphasizing the need for solvent controls.

In situations where more than one solvent could be used, preference should be given to water, then 95-100% ethanol, and then DMSO, in that order.

### 2.1.4 Concentration Range of Test Substances

The Panel agrees that the limit concentration should be 1 mM but the solubility characteristics and potential toxicity of each test substance (e.g., denaturation of the receptor) must be taken into consideration. If the limit concentration is used, seven test substance concentrations at log intervals should be tested. If a lower maximum concentration is tested due to solubility constraints or excessive toxicity, the number of concentrations tested can be adjusted to account for the altered concentration range. Concern was expressed among the Panel that denaturation of the receptor could occur at high test substance concentrations, and that this needed to be considered when the results of positive assays are interpreted. One Panel member stated during the meeting that some substances (e.g., surfactants) at concentrations below 1 mM might produce results that will be erroneously interpreted as positive in receptor binding assays, because of the loss of tertiary structure of the receptor. Whatever limit dose is chosen, care must be taken to ensure that only soluble concentrations of the test substance are used.

Also, the Panel agrees that at least seven different concentrations of the test substance within the range proposed in the BRD should be examined to increase the likelihood of developing a competition curve satisfactory for $IC_{50}$ analysis. For substances exhibiting solubility problems, both 95-100% ethanol and DMSO should be evaluated as solvents, perhaps with gentle warming at 50-55°C, to achieve the higher concentrations.

**2.1.5 Solvent and Positive Controls**

The Panel agrees that it is essential that the solvent (vehicle) volume in the solvent control assays be the same as that used in the reactions containing the test substances, and further that the solvent (vehicle) volumes remain constant throughout the competition curve development.

Regarding the positive control substance, the Panel strongly agrees that a substance with a binding affinity of two or three orders of magnitude below that of 17β-estradiol should be used as a positive control.

In order to assess the sensitivity of the assay system in each laboratory, it is suggested that a positive control should be used that is known to be a weak binder. Such a positive control could be either a naturally occurring estrogen, such as estriol or coumestrol, or a synthetic estrogen mimic, like tamoxifen. This positive control should be tested at three different concentrations.

There was discussion about employing a reference substance to determine the lower limit of detection of the assay but no recommendation was made.

**2.1.6 Within-Test Replicates**

The Panel agrees that triplicate measurements should be performed at each dose level to increase the likelihood of developing a competition curve satisfactory for $IC_{50}$ analysis, particularly during prevalidation and validation studies.

**2.1.7 Dose Spacing**

The Panel agrees with the recommended dose spacing of one order of magnitude in the concentrations of the candidate estrogen mimics, and in the use of half-log doses in certain cases.

**2.1.8 Data Analysis**

The Panel agrees with the requirement to determine and compute both the $B_{max}$ (number of binding sites or specific binding capacity) and the $K_d$ value. There was general agreement that the approach presented in the BRD is acceptable for screening substances, which inhibit estrogen binding. However, the use of alternative approaches such as a ligand titration array (Raffelsberger and Wittliff, 1997), which provides simultaneous evaluation of a laboratory's performance and determination of the estrogen binding properties (e.g., $IC_{50}$, $K_d$ value, $K_I$ value) of both reference and test substances, is recommended. This approach is equally valid for androgen

receptor analyses of putative androgenic mimics. The study by Villeneuve et al. (2000), on the derivation and application of relative potency estimates based on *in vitro* bioassay results could serve as a possible template for calculating $EC_{50}$ values for partial agonists or for substances for which the slope of the binding curve is atypical.

In any case, more detail is needed on statistical models for nonlinear regression analysis to compute $K_d$, $K_i$ and $IC_{50}$ values. This includes the nature of the statistical characteristics of the data (distribution, variance patterns, specific nonlinear models, etc.), and how to fit the models. When doing so, confidence limits must be calculated for $K_d$, $K_i$, and $IC_{50}$ values. From these data, details on how to make pertinent and valid statistical inferences should be specified.

### 2.1.9 Assay Acceptance Criteria

In general, there is agreement among the Panel with each of the acceptance criteria presented in the BRD. However, it is highly recommended that a reference ER preparation with established binding parameters be employed for the determination of the $K_d$ value and specific binding capacity by the laboratories chosen for validation of ER-based procedures. The Panel recommends that evaluation of the same reference receptor preparation with an identical set of test substances by various laboratories (as well as by individuals within the same laboratory) involved in this key process will provide a means of assessing both inter- and intra-laboratory reproducibility. Distribution of sets of results from either androgen or estrogen-based procedures to laboratories involved in validation is recommended to assess and compare the uniformity in the mode of calculation of desired parameters.

The positive control response must fall within defined limits, and assays should be performed in compliance with standard Good Laboratory Practice guidelines.

### 2.1.10 Evaluation and Interpretation of Results

The Panel agrees with the approach presented in the BRD. Because the choice of a standard ER preparation will have a significant influence on the $IC_{50}$, $K_d$, and $K_i$ values determined, the Panel recommends that recombinant human ER be employed for the validation and screening assays.

There is general agreement among the Panel with the designation of "equivocal" for substances that do not bring about a 50 % reduction in specific estrogen binding, particularly because other protein molecules such as heat shock proteins, co-activators, and co-repressors are known to influence estrogen binding properties (e.g., Horwitz et al., 1996; McKenna et al., 1999; McKenna and O'Malley, 2002). Because these proteins might be unavailable or present in suboptimal concentrations in certain ER preparations, negating accurate assessment of a substance's estrogen mimicry, discordance in results from ERα and ERβ assays should be considered in substance classification.

Furthermore, the classification of a test substance as "positive for binding" should be based on the use of statistical inferences pertinent to the positive characterization. This may require new research and development into valid statistical methods for making such a characterization. The state of the art for making statistical inferences with endocrine disruptor data is nascent and requires more detailed research and study.

**2.1.11 Test Report**

The Panel agrees with the detailed delineation of the test substance, the solvent/vehicle and the ER, but it recommends that the type of protein assay be specified and that the concentration of protein used in the reactions be reported. The Panel agrees also that the remaining information requested for test conditions, results, discussion of results and conclusion was adequate to achieve the stated goals of the survey and its validation.

The Panel recommends establishing a new range of reference $IC_{50}$, $K_d$, and $K_i$ values with a standardized ER preparation using a test set of substances. These data will be far more useful in the evaluation of putative estrogen mimics by various laboratories involved in the validation process, rather than using historical values of these parameters collected with various receptor preparations.

Historical data can also be used to assess the biological significance of results for a current test that has shown to be statistically significant.

**2.1.12 Replicate Studies**

The Panel agrees with the recommendation stated in the BRD.

**2.2 Considering the intended use of the assays as a toxicological screen, are there other minimum procedural standards that should be included? If so, what are they and why?**

The Panel recommends standardizing the type of protein assay and the conditions, and highly recommends that both the androgen and estrogen receptor-based assay reactions contain 10 mM sodium molybdate as well as a cocktail of protease inhibitors such as those described on page B3-5 of the BRD, to minimize degradation of the receptor protein.

Although the hydroxylapatite (HAP) procedure has been used by numerous investigators, it is fraught with problems, not the least of which is that the receptor-ligand complexes are bound to the matrix, require retention during washing, and must retain complex association during elution. These are important variables to control. In contrast, the U.S. Food and Drug Administration (FDA) approved the testing of human tissues with the radioligand (estrogen) binding assay using the Dextran-Coated Charcoal procedure. The latter procedure with dextran-coated charcoal allows the receptor-ligand complexes to remain in the original reaction medium while removing the unbound ligand. Published clinical cancer studies utilizing these FDA-approved procedures in Assay Proficiency Surveys (e.g., Fisher et al. 1980, 1983, 2001; Wittliff et al. 1981, 1998) indicate the reproducibility and transferability of this assay.

The Panel recommends the use of the ligand titration assay using dextran-coated charcoal as the preferred procedure.

**3.0 Recommendations for *In Vitro* ER Binding Test Method Protocols for Validation Studies**

**3.1 A standardized *In Vitro* ER Binding Assay protocol using rat uterine cytosol (RUC) is provided in Appendix B of the BRD. This assay is proposed for validation studies by the U.S. EPA and other sponsors. Section 12.3 discusses additional detail that should be added based on the minimum procedural standards in Section 12.2. In addition, an example of the *In Vitro* ER Binding RUC Assay (based on the U.S. EPA protocol), which incorporates the recommended minimum procedural standards, is provided in Section 12 Annex of the *In Vitro* ER Binding BRD. Considering the intended use of the assays as a toxicological screen, would the current protocols, with the additions detailed in Section 12.2 and 12.3, provide a level of detail to appropriately minimize interlaboratory variability? If not, what revisions or additions should be made to the protocols?**

The Panel is comfortable with the BRD protocol for the ligand binding assay, provided that it is amended to include the additional details presented in the discussion on minimum procedural standards (e.g., protease inhibitors, protein concentration assays [noting interference]) of the BRD and the points discussed in response to **Questions 1** and **2**.

**3.2 In addition to the minimum procedural details listed in Section 12.2, are there other protocol elements that should be considered for other *In Vitro* ER Binding assays recommended for validation as a toxicological screen, including those protocols provided in Appendix B?**

Assays are routinely performed using isotopes for comparative displacement binding assays. The Panel appreciates that some laboratories may have difficulties using isotopes because of licensing restrictions and efforts to limit production of isotope waste. Anisotropy for displacement of fluorescent estrogen may be used in non-isotope assays. However, experience is presently limited on the strengths and limitations of this end point.

There is also a concern about using a potent estrogen as the only standard reference, given that many of the agents that we wish to evaluate are at the weak end of the potency spectrum. The standard reference serves a role distinct from that of the positive control. The response of the positive control measures the assay's performance and stability over time. The use of a *concurrent* positive control is a hallmark of *in vitro* tests used in the regulatory arena worldwide. The use of a concurrent positive control is a stated "requirement" in protocols submitted to ICCVAM. The measurements of positive control performance need to address the assay endpoint(s) of interest (e.g., RBA values). They provide the basis for assessing the acceptability of the assay trial and thus the use of data from "unknowns" tested concurrently. Positive control data provide one basis for comparison of assay performance across laboratories. The positive control selected must be able to demonstrate both an under and over response relative to its historical values. The positive control suggested in Question 2 would provide the ability to measure both over and under prediction.

**3.3 Considering the intended use of the assays as a toxicological screen, is the Panel aware of other available standardized protocols for assays recommended for validation?**

The Panel's consensus is that purified recombinant ERα be the primary receptor used to develop the first assay for validation. To date, there are no estrogen mimics that are selective for either ERα or ERβ. In this assay, both ligand titration and competition end points should be developed for test chemicals. The ERα was suggested in response to Question 2 and the use of a ligand binding assay with dextran-coated charcoal is recommended to separate bound from free labeled 17β-estradiol. Among species comparisons could be facilitated by employing glutathione-S-transferase fusion proteins consisting of the d-e-f domains from the respective ERs (GST).

**4.0 Recommended List of Substances to be Used for Validation of *In Vitro* ER Binding Assays**

**Section 12.4** of the BRD provides a list of substances recommended for use in validation studies of *in vitro* ER binding assays.

**4.1 Considering the intended use of the assays as a toxicological screen, does the Panel agree with the selection criteria, adequacy and appropriateness of substances recommended for validation studies in terms of the following issues? If not, what substances should be added or deleted?**

**4.1.1 Number and Distribution of Substance Across the Range of Measurable ER Binding Activity, Including Negatives**

**4.1.2 The Number and Range of Substances by Chemical Class**

**4.1.3 The Number and Range of Substances by Product Class**

The Panel endorses the list of chemicals provided in the BRD, but with the following caveats.

There are concerns about how well the list represented the kinds of substances that the U.S. EPA is interested in screening, including the diversity of chemical classes, the range of potencies. The Panel also indicates that it will not be possible to determine the specificity of the assay with a list of substances 90% of which are positive. The recommendation is that at least 25% of the test substances be negative for ER binding.

**References**

Fisher B, Redmond C, Brown A, Wolmark N, Wittliff J, Fisher ER, Plotkin D, Bowman D, Sachs S, Wolter J, Frelick R, Desser R, Licalzi N, Geggie P, Campbell T, Elias EG, Prager D, Koontz P, Volk H, Dimitrov N, Gardner B, Lerner H, Shibata TA. 1981. Treatment of primary breast cancer with chemotherapy and tamoxifen. New England Journal of Medicine 305:1-6.

Fisher B, Redmond C, Brown A, Wickerham DL, Wolmark N, Allegra J, Escher G, Lippman M, Savlov E, Wittliff J, Fisher ER, Plotkin D, Bowman D, Wolter J, Bornstein R, Desser R, Frelick R. 1983. Influence of tumor estrogen and progesterone receptor levels on the response to tamoxifen and chemotherapy in primary breast cancer. J. Clinical Oncology 1:227-241.

Horwitz KB, Jackson TA, Rain DL, Richer JK, Takimoto GS, Tung L. 1996. Nuclear receptor coactivators and corepressors. Molecular Endocrinology 10:1167-1177.

McKenna NJ, Lanz RB, O'Malley BW. 1999. Nuclear receptor coregulators: Cellular and molecular biology. Endocrine Reviews 20:321-344.

McKenna NJ and O'Malley BW. 2002. Minireview: Nuclear receptor coactivators – An update. Endocrinology 143:2461-2465.

Munson PJ and Rodbard D. 1980. LIGAND - A versatile computerized approach for characterization of ligand-binding systems. Anal Biochem 107:220-239.

Raffelsberger W and Wittliff JL. 1997. A novel approach for comparing ligand binding results from titration and competition analyses to study hormone mimics. J. Clinical Ligand Assay 20: 275-280.

Villeneuve DL, Blankenship AL, Giesy JP. 2000. Derivation and application of relative potency estimates based on in vitro bioassay results. Environ Toxicol and Chem 19:2835-2843.

Wittliff J, et al. 1981. Methods of steroid receptor analyses and their quality control in the clinical laboratory. In: Physiopathology of Endocrine Diseases and Mechanisms of Hormone Action. New York:Alna R. Liss, Inc., 397-411.

## Appendix 1
## Ligand Titration Assay

***Data File: C303S.DAT***
***Protocol File: 3HE2A.PRO***

File ID: hERalpha with 3HE2/tris
1.1.1 12:33:03

Protein (mg/ml): 0.2
Weighting: Unity
Nonspecific Binding: # Points = 6
Correlation Coefficient, R = 0.9974
Linear and Non-Linear Regressions: # Points = 6

| | **Model 1** | **Model 2** | **Model 3** |
|---|---|---|---|
| | Scatchard | Non-Linear One Site | Non-Linear One Site with NSB |
| BMAX [FM/MG PROTEIN] | 4.039E+03 | 3.898E+03 | 2.956E+03 |
| KD [MOLAR] | 3.864E-10 | 3.567E-10 | 2.239E-10 |
| NSB [LITERS/MG PROTEIN] | -- | -- | 5.806E-04 |
| BMAX STANDARD ERROR +/- | -- | 179.2 | 296.5 |
| KD STANDARD ERROR +/- | -- | 4.797E-11 | 4.010E-11 |
| NSB STANDARD ERROR +/- | -- | -- | 1.273E-04 |
| SUM OF SQUARES OF RESIDUALS | 6.543E-26 | 5.716E-26 | 2.014E-26 |
| CORRELATION COEFFICIENT | -.9582 | -- | -- |

***Data File: C303S.DAT***
***Protocol File: 3HE2A.PRO Assay Protocol***

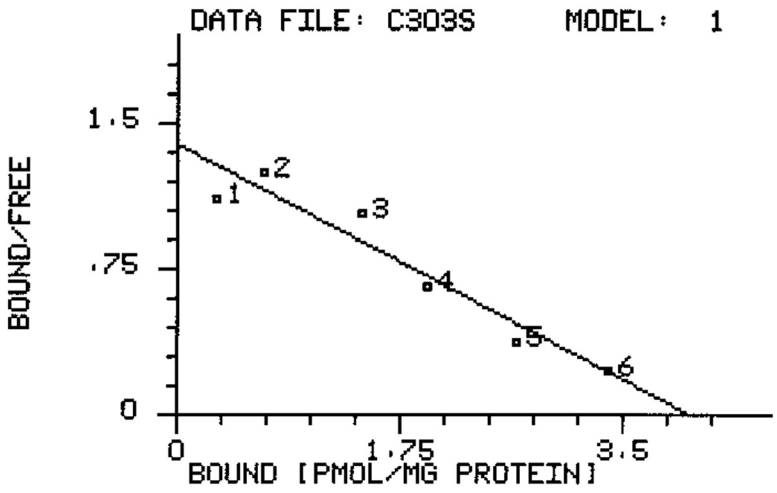
Cytosol or Membrane Protein Volume (ml) .....0.1000
Total Incubation Volume (ml) .....0.1500
Separation Solution Added (e.g., DCC) (ml).....0.3000
Aliquot Counted (ml).....0.3000
Efficiency of Dose Tubes (~).....100.0000
Efficiency of Total Bound and NSB Tubes (~).....100.0000
Specific Activity Value .....158.0000
Specific Activity Units ..... DPM/fmol

***Data File: C303S.DAT***
***Protocol File: 3HE2A.PRO***

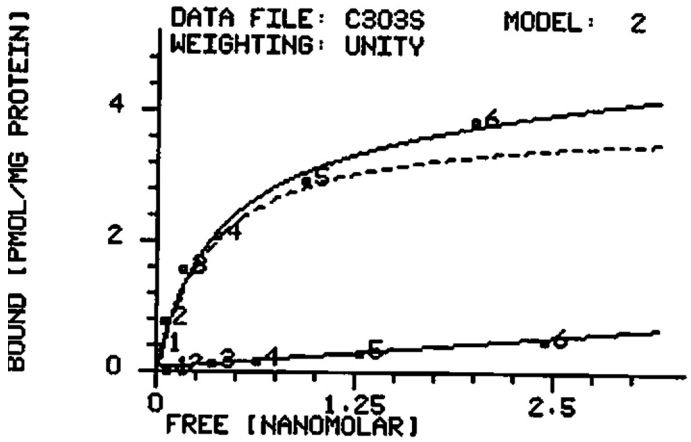
Measured Data Table

| DOSE # | TOTAL COUNTS | TOTAL COUNTS | TOTAL COUNTS | %CV TOTAL | TOTAL BOUND | TOTAL BOUND | TOTAL BOUND | %CV BOUND | NSB | NSB | %CV NSB |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2134 | 2192 | 2216 | 1.93 D | 637 D | 781 | 811 | 0.00 | 88 | 73 | 13.18 |
| 2 | 4039 | 4246 | 4305 | 3.33 | 1613 | 1770 | 1426 | 10.74 | 118 | 129 | 6.30 |
| 3 | 9579 | 10171 | 8522 | 8.86 | 3619 | 2551 | 3692 | 19.43 | 287 | 248 | 10.31 |
| 4 | 15462 | 16613 | 16445 | 3.84 | 4453 | 4478 | 4428 | 0.56 | 379 | 350 | 5.63 |
| 5 | 32123 | 32675 | 30750 | 3.11 | 5629 | 6646 D | 5127 | 11.72 | 659 | 587 | 8.17 |
| 6 | 58390 | 58960 | 62415 | 3.63 | 7886 | 9501 | 6856 | 16.50 | 917 | 1212 | 15.60 |
| 7 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0.00 |
| 8 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0.00 |
| 9 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0.00 |
| 10 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0.00 |
| 11 | 0 | 0 | o | 0 00 | 0 | 0 | 0 | 0 00 | 0 | 0 | 0 00 |
| 12 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0.00 |
| 13 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0.00 |
| 14 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0.00 |
| 15 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0.00 |
| 16 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0.00 |
| 17 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0 | 0. 00 | 0 | 0 | 0 00 |
| 18 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0 | 0.00 | 0 | 0 | 0.00 |

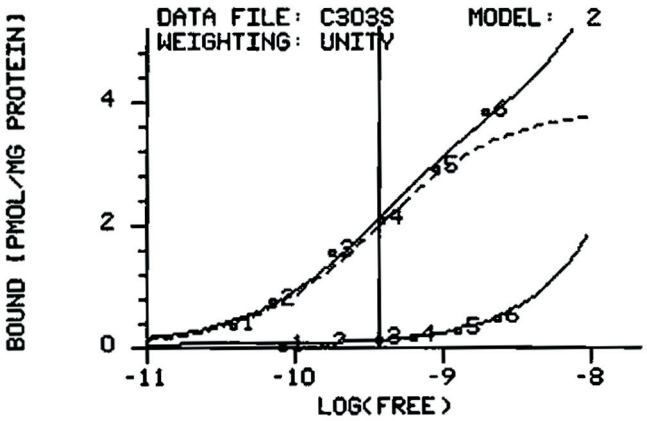
***Data File: C303S.DAT***
***Protocol File: 3HE2A.PRO***



***Data File: C303S.DAT***
***Protocol File: 3HE2A.PRO***




***Data File: C303S.DAT***
***Protocol File: 3HE2A.PRO***



# Appendix 2
# Ligand Competition Array

## *Summary Table*
## *Page 1*

```
CM1 FILE: YER303L6.CM1        DT1 FILE: YER303L6.DT1        PR1 FILE: YER303L6.PR1
                              WEIGHT:     UNITY
Protein Conc [mg/ml]: 0.2   Assay ID: hERalpha with estrone
Model:            1           2           3           4           5
                1-measNSB   1-compNSB   2-measNSB   2-compNSB       1-SITE
                >----------------NON-LINEAR-----------------< >----LINEAR----<
Kd1  (M)<c> :  3.864E-10   3.864E-10      --          --      IC50:  1.220E-08
Kd2  (M)<c> :     --          --          --          --        Ki:  1.957E-09
Kd3  (M)    :  4.031E-09   2.613E-09      --          --      Hill:  -.663E+00
Kd4  (M)    :     --          --          --          --       Kic:  1.697E-09
Bmax1(m/mg) :  2.968E-12   2.936E-12      --          --      #Pts:      6
Bmax2(m/mg) :     --          --          --          --
NSB         :     --       3.964E-02      --          --
Kd3      SE :  9.561E-10   6.595E-10      --          --
Kd4      SE :     --          --          --          --
Bmax1    SE :  2.267E-13   2.062E-13      --          --
Bmax2    SE :     --          --          --          --
NSB      SE :     --       4.950E-03      --          --
SSE         :  2.291E-21   9.816E-22      --          --
```

## *Summary Table*
## *Page 2*

```
CM1 FILE: YER303L6.CM1        DT1 FILE: YER303L6.DT1        PR1 FILE: YER303L6.PR1

# Competitor Concentrations:  7
Model:            1           2           3           4           5
                1-measNSB   1-compNSB   2-measNSB   2-compNSB       1-SITE
                >----------------NON-LINEAR-----------------< >----LINEAR----<
SSE         :  2.291E-21   9.816E-22      --          --          --
Runs #      :     5           6          --          --          --
 +Res       :     4           3          --          --          --
 -Res       :     3           4          --          --          --
DF          :     5           4          --          --          --
Runs Test   :  p > 0.05    p > 0.05       --          --
F-comparison:                          3 vs 1      4 vs 2
F-Test      :
Error Code  :     0           0          13          13           0
```

***Data File: YER303L6***
***Page 1***

| REPLICATE NUMBER | TOTAL COUNTS ADDED PER TUBE | TOTAL COUNTS BOUND NO COMPETITOR | NSB COUNTS |
|---|---|---|---|
| 1 | 58390 | 7886 | 917 |
| 2 | 58960 | 9501 | 1212 |
| 3 | 62415 | 6856 | 0 |
| 4 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 |
| Reps | 3 | 3 | 2 |
| Mean | 59921.7 | 8081.0 | 1064.5 |
| Mean [M] | 2.535E-09 | 5.128E-10 | 6.754E-11 |
| %CV | 3.63 | 16.50 | 19.60 |

***Data File: YER303L6***
***Page 2-3***

| DOSE # | TOTAL COUNTS | TOTAL COUNTS | TOTAL COUNTS | TOTAL COUNTS | TOTAL COUNTS | TOTAL COUNTS | REPS | MEAN VALUE | MEAN MOLAR | %CV |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 6227 | 6184 | 0 | 0 | 0 | 0 | 2 | 6205.5 | 3.938E-10 | 0.49 |
| 2 | 4742 | 4542 | 0 | 0 | 0 | 0 | 2 | 4642.0 | 2.945E-10 | 3.05 |
| 3 | 3922 | 3435 | 0 | 0 | 0 | 0 | 2 | 3678.5 | 2.334E-10 | 9.36 |
| 4 | 2841 | 3007 | 0 | 0 | 0 | 0 | 2 | 2924.0 | 1.855E-10 | 4.01 |
| 5 | 1893 | 1793 | 0 | 0 | 0 | 0 | 2 | 1843.0 | 1.169E-10 | 3.84 |
| 6 | 1942 | 2178 | 0 | 0 | 0 | 0 | 2 | 2060.0 | 1.307E-10 | 8.10 |
| 7 | 1442 | 1554 | 0 | 0 | 0 | 0 | 2 | 1498.0 | 9.505E-11 | 5.29 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 23 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 26 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 27 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 28 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 29 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 30 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 31 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 32 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 33 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 34 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 35 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |
| 36 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.000E+00 | 0.00 |

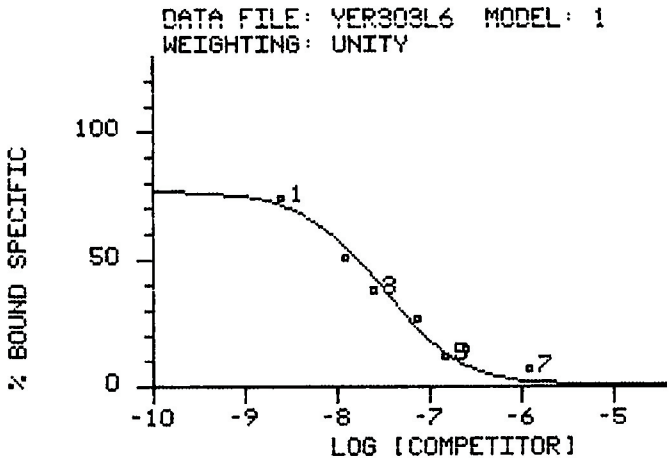
***Protocol File: YER303L6***

Cytosol or Membrane Protein Volume (ml). .....0.1000
Total Incubation Volume (ml) . .....0.1500
Separation Solution Added (e.g., DCC) (ml) .....0.3000
Aliquot Counted (ml) .....0.3000
Specific Activity Value . .....157.6
Specific Activity Units ..... .DPM/fmol
Efficiency of Dose Tubes (%).....100
Efficiency of Round and NSB tubes (%) .....100
Loading Sequence .....Sequential
Method of Weighting ..... UNITY
Printer Type ..... 24 PIN
Kd linear model.....3.864E-10
Kd1 <c>.....3.864E-10
Kd2 <c>.....0
Bmax units . ..... [moles/mg]
Report Sequence: GHBCDEFIKM

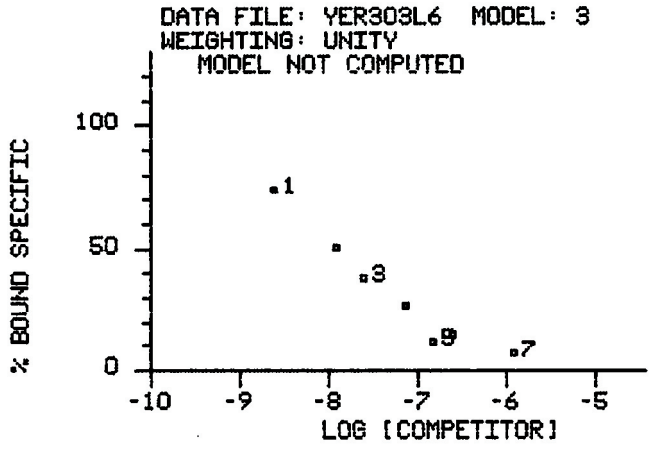
***Concentration File: YER303L6***

| DOSE | CONCENTRATION [M] | DOSE | CONCENTRATION [M] |
|---|---|---|---|
| 1 | 2. 530E-09 | 19 | -- |
| 2 | 1.270E-08 | 20 | -- |
| 3 | 2.530E-08 | 21 | -- |
| 4 | 7.600E-08 | 22 | -- |
| 5 | 1.520E-07 | 23 | -- |
| 6 | 2.530E-07 | 24 | -- |
| 7 | 1.270E-06 | 25 | -- |
| 8 | -- | 26 | -- |
| 9 | -- | 27 | -- |
| 10 | -- | 28 | -- |
| 11 | -- | 29 | -- |
| 12 | -- | 30 | -- |
| 13 | -- | 31 | -- |
| 14 | -- | 32 | -- |
| 15 | -- | 33 | -- |
| 16 | -- | 34 | -- |
| 17 | -- | 35 | -- |
| 18 | -- | 36 | -- |

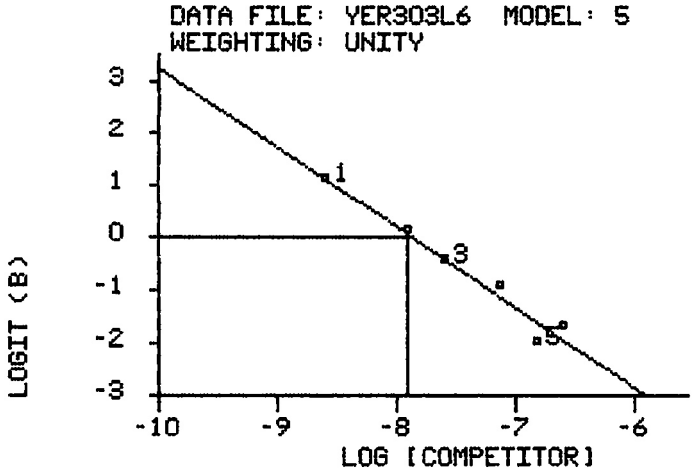
***Data File: YER303L6***
***Protocol File: YER303L6***




***Data File: YER303L6***
***Protocol File: YER303L6***



***Data File: YER303L6***
***Protocol File: YER303L6***



## II. *In Vitro* Estrogen Receptor (ER) Transcriptional Activation (TA) Assays

### 1.0 Recommendations and Priority for Validation Studies

The *in vitro* ER TA BRD reviews the comparative performance, reliability, advantages, and disadvantages for different *in vitro* ER TA assays, and recommends a relative priority for further development and/or validation based on this information.

**1.1 Considering the intended use of the assays as a toxicological screen, is the Panel aware of other advantages and disadvantages for the assays discussed in the BRD?**

The Panel has not identified additional advantages and disadvantages over and above those discussed in the BRD.

**1.2 Considering the intended use of the assays as a toxicological screen, does the Panel agree with the relative priority recommended for these sets of assays? Does the Panel recommend any changes in priority, or have specific recommendations for prioritization?**

The Panel, while agreeing that a stably transfected cell line is an attractive test method for prioritization (because such assays would be more amenable to high throughput screening), is not convinced that the BRD made a strong enough argument for this approach as having the highest priority. Reasons for this conclusion are:

1. Stable cell lines are notorious for losing their stability over time and therefore require continuous selection.
2. Their availability is limited and a highly responsive stable cell line is difficult to isolate.
3. While stably transfected cell lines might be potentially less challenging to use, they could have inherent confounding issues such as the effect of multiple receptor subtype activation.

Therefore, before making such assays a priority, appropriate comparative data relative to assays using transiently transfected cell lines (which generally have a higher level of responsiveness) should be generated using a select set of test substances.

**1.2.1 Which receptor types (species, isoform) are the best for *In Vitro* ER TA Assays?**

The Panel expresses a general preference for the use of the human ERα and ERβ subtypes in *in vitro* ER TA screening assays. However, if patent issues preclude the utilization of human gene sequences in commercial applications, consideration should be given to rat and possibly porcine receptors. Also, the potential for evaluation of receptors (ERα, β, or γ) from species of environmental relevance (e.g., fish) should be considered.

**1.2.2 Should preference be given to cells with endogenous ER, transiently transfected ER expression vectors, or stably transfected ER expression vectors?**

The Panel does not think that there was enough information in the BRD to make a judgment as to the superiority of one cell system over another; too few positive and negative chemicals have been tested for agonism or antagonism activity in multiple *in vitro* ER TA assays using different mammalian cell lines. Furthermore, the *in vivo* endocrine disrupting activity of a chemical would most likely be tissue-, cell-,

and promoter-specific. Therefore, the intrinsic responsiveness of a cell line cannot be generalized based on the result of a single assay system, due to the potential differences in coactivator populations, cross-talk with other receptors, and other signal transduction pathways between cell types.

The availability of stable cell lines that are already established should be investigated.

**1.2.3 Which response elements (species, sequence) are the best for the reporter vectors?**

The Panel believes that the effect of the reporter construct itself should not be underestimated in the validation of these assays. Issues such as single versus multiple estrogen response elements (EREs), other enhancer elements, and different minimal promoters are of importance in the evaluation and optimization of an assay system. However, the vitellogenin ERE (vit ERE) consensus sequence is recommended based on its broad-based responsiveness, although concerns about the activity of the vit ERE in some mammalian cells have been expressed by the Panel. The use of a reporter construct containing multiple EREs is recommended to maximize the sensitivity of the resulting assay and to minimize missing weak responders.

The Panel is also of the opinion that the development of a series of transient transfection assays for individual receptor subtypes (e.g., ERα, β) is worthwhile, especially in the context of the use of chimeric receptor-reporter constructs which would allow for greater mechanistic specificity.

Some Panel members commented on the fact that the use of such chimeric ligand binding domain (LBD) constructs could potentially minimize effects due to differing promoters/EREs.

Estrogenic compounds can stimulate transcriptional activation of the AP-1 (fos-jun) complex (that bind to regulatory sequences in the promoter of various target genes to modulate transcription) through a cooperative interaction of the ER with this complex. The fact that different ER ligand complexes have different affinities for fos-jun and other co-activators should be a consideration in the selection and evaluation of cell lines and assays for determination of ER TA screening assays.

Suggestions and Recommendations:

1. As an important part of the validation process, a study needs to be conducted to determine if stably transfected cell lines really do perform better than transiently transfected cells. This study would involve a comparison of a stably transfected mammalian line with one transiently transfected with the ER and other reporter plasmids, along with one expressing an endogenous ER.
2. If stably transfected cell lines are used, there should be a standard procedure for ensuring the maintenance of minimum response criteria to selected standards, including 17β-estradiol. Furthermore, treatment with the required selection

antibiotic should be performed on a regular basis to ensure maintenance of the inserted receptor and/or reporter construct.

3. These assays do not measure toxicity. The incorporation of an appropriate measure of cytotoxicity into the assay system is recommended.

## 2.0 Minimum Procedural Standards for *In Vitro* ER TA Assays

### 2.1 Considering the intended use of the assays as toxicological screens, does the Panel agree with the adequacy of the proposed procedural standards recommended for *In Vitro* ER TA Assays?

#### 2.1.1 Transcriptional Activation of the Reference Estrogen

The Panel agrees that the transcriptional activation-inducing ability of 17β-estradiol must be demonstrated and that consistency in the level of response is appropriate as a criterion for assay acceptance. There was a question of whether the level of response should be from a single dose or if it would be necessary to obtain a full dose response curve. The Panel generally felt a full dose-response curve would be more informative.

#### 2.2.2 Reference Estrogen

17β-Estradiol is appropriate as the reference estrogen. The Panel recommends that preliminary studies be performed with multiple transactivation assays to statistically define assay performance expectations for 17β-estradiol dose response curves (i.e., maximum fold induction, $EC_{50}$ values, confidence limits). Concentrations to be tested for the reference positive control, 17β-estradiol, should range from 1 pM to 1 μM to establish a full dose-response curve. Clear guidance is needed with regard to expectations for the extent of response that should be observed, which can be determined based on preliminary studies.

#### 2.1.3 Preparation of Test Substances

The test substances should be prepared in water, 95-100% ethanol, or dimethyl sulfoxide (DMSO), depending upon their solubility. Preference should be given to the solvent that allows testing of the maximal concentration of the test substance, without exceeding the limit dose. However, in situations where more than one solvent could be used, preference should be given to water, then 95-100% ethanol, and then DMSO, in that order. The Panel suggests that guidelines be provided with regard to the concentration of solvent in the stock solution. Even when using 95-100% ethanol or DMSO, substances to be tested could be prepared in stock solutions where their concentration approaches solubility limits. This could introduce variation from laboratory to laboratory. In addition, it should be stated that standards or positive controls need to be dissolved in the same solvent and to the same maximal concentration. One Panel member indicated a preference for 95-100% ethanol at concentrations of 0.01 to 0.1% (v/v) because some substances dissolved in DMSO have been observed to exhibit lower activity than when dissolved in ethanol.

The Panel concludes that it would be prudent to perform a prevalidation of *in vitro* ER TA assays with the reference estrogen for assessment of the level of solvent that does not adversely affect assay response.

### 2.1.4 Concentration Range of Test Substances

The Panel generally agrees that, for both agonism and antagonism assays, the limit concentration should be 1 mM as long as the solubility characteristics and potential cytotoxicity of each test substance is taken into consideration. However, concentrations greater than 10 µM are often problematic due to solubility issues in aqueous media, compounding the level of toxicity. Thus, one Panel member recommended 0.1 mM as the limit concentration. At a minimum, the solubility of the substance should be reported and the concentration used in the test should not exceed the limit of solubility.

If the limit concentration is used, the Panel generally agrees that seven test substance concentrations at log intervals should be tested. However, one Panel member suggested that it is likely that covering a concentration range of five orders of magnitude will be appropriate and that it is unclear what would be gained by using anything other than one order of magnitude between the doses. This is especially true since the assay results will only be assessed in a semi-quantitative manner (i.e., $IC_{50}$ or $EC_{50}$ values should not be used to rank compouds regarding possible potency).

If a lower maximum concentration is tested due to solubility constraints or excessive cytotoxicity, the number of concentrations tested can be adjusted to account for the altered concentration range.

In addition, a measure of cellular cytotoxicity incorporated 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 type of assessment might mitigate the need to go to concentrations higher than 10 µM.

The Panel also recommends that a sequential testing strategy be adopted. For example, if the substance is positive for agonist activity, there would seem to be little value in testing it for ER antagonist activity or AR-mediated activities. A positive result in any of these assays will likely warrant further examination in *in vivo* screens and tests other than transactivation assays. However, the validity of this approach should be evaluated in the prevalidation phase and a decision made on its applicability after sufficient data has been reviewed.

Assuming that each test substance is tested as an agonist and an antagonist, substances demonstrating a positive agonist response could also be tested with ICI 182,780 (the candidate ER antagonist) to make sure that the effect reflects a receptor-mediated activity. Assuming that ICI 182,780 does block the action of the

test substance under the test conditions used, it will also help to indicate the presence of toxicity if the signal level is significantly below that of ICI 182,780 alone.

During deliberations of the Panel, it was suggested that the concentrations (doses) of chemicals tested be similar for ER binding and ER transcriptional activation assays. However, the concentrations presented to the receptor in cell based systems could be markedly different from that presented to the receptor in media. Differences in chemical properties in serum or in solvent can markedly affect the uptake of chemicals by cells. Thus, the magnitude of response may not be directly comparable between cell-free binding assays and cell based TA assays. When considered important, radiolabeled test chemicals should be used to determine the amount of chemical taken up by the cells.

### 2.1.5 Solvent and Positive Controls

The Panel concludes that the recommendations in the BRD are appropriate. ICI 182,780 appears to completely block 17β-estradiol at 0.1 μM. However, the collective experience of the Panel is that obtaining ICI 182,780 might require permission from its producer in England and the maximum amount available might possibly be only 10 mg/year. One Panel member expressed reservations concerning the use of tamoxifen as a positive control. In addition, as with the reference standard for agonist activity, clear guidelines for the positive antagonist and the expected extent of antagonism when testing the substance should be provided. Perhaps running a parallel toxicity study on the same plates with a compound like Alamar Blue or a metabolism assay would add additional information.

During the deliberations of the Panel, it was suggested that a weak estrogen (e.g., estriol) should be included routinely in assays along with 17β-estradiol as a reference compound. Whether this will add necessary and sufficient information should be evaluated in prevalidation assays. If the inclusion of a weak estrogen improves the ability to interpret the results obtained with validation chemicals, or with unknowns that are weak estrogens, then the inclusion of a weak estrogen reference positive control is strongly recommended.

### 2.1.6 Within-Test Replicates

The Panel recommends that, initially, each test chemical concentration be tested in triplicate. The validity of this approach, however, should be evaluated with statistical consultation.

### 2.1.7 Dose Spacing

The Panel generally agrees with the recommended dose spacing of one order of magnitude in the concentrations of the candidate test substances but did not think the use of half-log doses in certain cases would be useful.

### 2.1.8 Data Analysis

#### 2.1.8.1 Response Variable for Evaluating the Potential Agonism/ Antagonism of a Substance

In **Section 5-3** of the BRD, it is stated that there are several different definitions available for the "relative activity" of a test substance. The Panel recommends that a consensus be reached on one definition that can be considered as the standard definition for all future assays. This is vital because important decisions, including the final call (positive/negative agonism, and positive/negative antagonism) regarding a chemical are based on the chosen definition.

#### 2.1.8.2 Assumptions Made in a Statistical Analysis

Most statistical procedures make certain assumptions regarding the underlying data. For instance, if ANOVA type methods are used, then it is necessary to demonstrate that there is no heteroscedasticity and that the data are approximately normal. The Panel recommends that before any statistical procedure is used, suitable diagnostics are performed to make sure all underlying assumptions regarding the procedure are true. If the assumptions are not true then suitable data transformations might be performed before analyzing the data.

#### 2.1.8.3 Estimation of $EC_{50}$/$IC_{50}$ Values and a "Steepness" Parameter

If it is appropriate to perform nonlinear modeling, then a suitable nonlinear model should be selected. In some cases, the Hill equation might be suitable. Once a model is selected, appropriate model diagnostics need to be performed to ensure that the model fits the data and the various underlying assumptions such as normality and homoscedatsticity are true. Diagnostics should also be performed to detect curvature effects, typically by using standard asymptotic confidence intervals. However, in some situations, especially in the presence of curvature effects, these confidence intervals might not be appropriate. In such cases, a resampling procedure such as jackknife or bootstrap[1] might be used.

#### 2.1.8.4 Combining $EC_{50}$ and $IC_{50}$ Estimates from Different Laboratories

To obtain estimates of mean $EC_{50}$ and mean $IC_{50}$ values from different laboratories, the average across laboratories should not be computed but rather estimates should be obtained using mixed effects nonlinear models, treating laboratories as the random effects. This approach takes into

[1] cf. Simonoff and Tsai, "Jackknife-based estimators and confidence regions in nonlinear regression," Technometrics, 28, 103-112, 1986, and Zhang, J., Peddada, S. D., Rogol, A. D, "Estimation of parameters in nonlinear regression models," Statistics for the 21st century: Methodologies for applications of the future, edited by C. R. Rao and G. Szekely, Marcel-Dekker, New York, NY, 2000.

account within and between laboratory variability. This methodology also allows for a determination of corresponding standard errors.[2]

**2.1.8.5 Uniformity Trials**

To understand the underlying variability in the data, which is important for proper data analysis, the Panel recommends undertaking a set of carefully planned comprehensive interlaboratory negative control studies.

**2.1.9 Assay Acceptance Criteria**

Reference substances for agonism and antagonism should give responses within appropriate confidence limits. These confidence limits should be determined in preliminary studies (see **Section 2.1.2**). The Panel recommends that guidelines be provided for a certain level of agonism or antagonism expected for the reference standards and that responses in these ranges should be required if the assay is to be accepted. There should also be a minimal fold ($B_{max}$) and/or minimal fold/between experiment variance ratio for the assay. This will be needed for each assay type recommended.

Incorporation of an evaluation of cell cytotoxicity/cell proliferation should be included in the protocol to ensure that responses at non-toxic doses only are analyzed, and acceptance criteria need to be established for when cytotoxicity affects the performance of the assay.

The Panel recommends that the assays be performed following Good Laboratory Practice guidelines.

**2.1.10 Evaluation and Interpretation of Results**

The interpretation of positive results for a substance as an agonist or antagonist should incorporate some elements of a dose-response relationship in comparison to the reference standards. Simply classifying a substance as an ER agonist based on significance above the concurrent control without consideration of a dose-response is not sufficient. In addition, assay performance criteria must be within an acceptable range.

**2.1.11 Test Report**

In addition to the information required for the test report, as listed in the BRD, 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, detection of luciferase activity, etc. must also be reported. If it is a transfection assay, there is also a need to include a constitutive reporter gene assay to control for transfection efficiency between wells. The passage number of the cell line should be monitored, as well as the $CO_2$ level during culturing

[2] cf. Davidian and Giltinan, Nonlinear models for repeated measurement data, Chapman and Hall, London, UK, 1995

and treatment. The $EC_{50}$ value for agonism or the $IC_{50}$ value for antagonism, together with corresponding confidence limits, must be reported.

**2.1.12 Replicate Studies**

The Panel concludes that the appropriate extent of replication should be determined after prevalidation studies have been conducted and the resulting data statistically evaluated. However, it was generally thought that replicate studies in a screening assay are probably not required as long as the expected response from the reference standards has been observed and a statistically meaningful dose response is observed for the test substance(s). If either of these provides data that do not conform to expectations, the assay should be repeated, as per standard practice. Defining the minimum standards for replication will need to wait until the extent of variation within a test has been carefully evaluated. In contrast, one Panel member believes that each substance should be tested at least three times in different experiments because it could be toxic in one assay (due to low cell density, fitness, etc.) and then the activity is detected in another assay. If repeated, the incorporation of more closely spaced dosing/treatment concentrations in the replicate assay based on the initial test results should also facilitate better analysis of the overall dose-response of the test substance.

**2.2 Considering the intended use of the assays as a toxicological screen, are there other minimum procedural standards that should be included? If so, what are they and why?**

This issue has been discussed in preceding sections, as appropriate.

**General Statements or Comments in Regard to the BRD -** The Panel concludes that:

1. There was inconsistency in the statements on pages 12-1 and 12-11 of the BRD concerning stable versus transiently transfected cell lines.
2. There was no discussion of individual assays for ERα and β.
3. *In vitro* ER TA assays are not a toxicological screen as stated in the questions to the Panel. They are simply a measure of transactivation. Further research is needed to understand the link between the results from this *in vitro* assay and a toxicological effect in an organism.

**3.0 Recommendations for *In Vitro* ER TA Test Method Protocols for Validation Studies**

Protocols provided by scientists with expertise in *in vitro* ER TA test methods were provided in **Appendix B** of the BRD. **Section 12.3** discusses additional details that should be added, based on the minimum procedural standards in **Section 12.2**.

**3.1 Considering the intended use of the assays as a toxicological screen, would the current protocols, with the additions detailed in Sections 12.2 and 12.3, provide a level of detail to appropriately minimize interlaboratory variability? If not, what revisions or additions should be made to the protocls?**

The consensus among the Panel is that the protocols adequately described the needed procedures and that any laboratory with cell/yeast culture and basic molecular experience should be able to produce dependable results.

Some interlab variability may be expected due to laboratory specific culture techniques (for example, cell counting, determination of the percentage of confluence, ability to seed plates evenly, etc.). Perhaps standards or performance criteria should be specified for such activities.

**3.2 In addition to the minimum procedural details listed in Section 12.2, are there other protocol elements that should be considered for *In Vitro* ER TA Assays recommended for validation as a toxicological screen, including those protocols provided in Appendix B?**

The Panel believes that the level of detail was generally sufficient, but recognizes that performance would depend on the experience of the staff in each laboratory.

During Panel deliberations, one member suggested that additional procedural details should be added if volatile chemicals are tested. Another member suggested that problems with volatiles might be reduced if specifications are made for incubators without circulation fans in the chamber.

The Panel concludes that the following issues might need to be added or expanded on in the protocols:

1. Standards should be provided for uniform counting and plating of cells among wells and between experiments.
2. Discussion and review of methods should be included for making dextran coated charcoal (DCC)-stripped sera and perhaps even a recommendation for commercial sources.
3. Discussion and review of known sources of estrogen contamination in the laboratory should be included. This might include plastic ware as well as media additives and commercial prepared media. Some tubes, filter units, antibiotic mixtures, and pre-made media in polycarbonate bottles are examples of suspect and/or problematic items. Performance criteria should be established for determining background levels of hormone active contaminants (e.g., by comparing the reporter activity for ICI 182,780 versus the blank). This could also include methods to maintain a hormone free lab environment.
4. The issue of culturing some cell stocks in estrogen rich media and then withdrawing them to an estrogen free media may need to be expanded or emphasized. Suggested washing techniques, including the number of days for withdrawal, etc., need to be detailed. Again, performance criteria should be established to show that each experiment was conducted using estrogen free conditions (e.g., by comparing the reporter activity for ICI 182,780 versus the blank).
5. Another issue related to hormone active contamination is that while an estrogen free environment is required for reliable estrogen activity assays, it does not seem to be an issue for androgen receptor (AR) activity determinations. However, it should be noted that there is potential for cross talk between "estrogenic" media contaminants and other signaling pathways. Thus, it is not clear what effects this may have on androgen activity assays. Perhaps the AR methods should use hormone free procedures as a precaution (no phenol red, DCC sera, etc.).
6. The metabolism of chemicals selected for validation, or unknowns, is an important concern in cell-based assays. Oxidative (Phase I) or conjugation (Phase II) reactions

can convert pro-estrogens to active estrogens or inactivate parent chemicals that are active estrogens. Cells in primary cultures have inconsistent capacities for xenobiotic metabolism. Cell lines often have limited capacity for xenobiotic biotransformation. Given the number of chemicals involved and the number of enzymes potentially involved, the Panel suggests that the scope of effort to determine pathways, products, and activities of products is beyond what would be feasible in a validation study. However, the Panel recommends that available information on the P450 complement and Phase II enzyme complement be compiled for the cell lines employed in this validation process. The Panel also recommends that available information on the metabolism of the validation chemicals is compiled, including the degree to which metabolism alters estrogenic activity. It is further recommended that studies to obtain such information for systems ultimately employed in screening be planned and performed, when applicable. While metabolism could affect the magnitude of the signal of parent estrogens, it is unlikely to negate the possibility of detecting such activity. If metabolic conversion of proestrogens to an active derivative occurs at very low rates, then the estrogenic activity could be missed.

**3.3 Considering the intended use of the assays as a toxicological screen, is the Panel aware of other available standardized protocols for assays recommended for validation?**

One Panel member indicated the availability of a standardized protocol for the MVLN assay, which uses an MCF-7 cell line derivative known as MVLN. This cell line, which harbors an endogenous ER, has been stably transfected with the luciferase reporter gene under control of the vit ERE. Estrogen-specific transcription activity is directly related to luciferase activity. Cells, seeded into a microtiter plate, are treated with the test substance and incubated overnight. The following day the cells are lysed and luciferase activity in the supernatant is measured in a luminometer. A copy of the protocol has been added to Appendix B of the ER TA BRD.

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

**4.1 Does the Panel agree with the selection criteria, adequacy, and appropriateness of substances recommended for validation studies in terms of the following issues? If not, what substances should be added or deleted?**

In general, the Panel agrees with the selection criteria, adequacy, and appropriateness of the chemicals chosen for the validation studies. However, several specific concerns were raised and the Panel made recommendation to address these concerns.

### 4.1.1 Number and Distribution of Substance Across the Range of ER TA Responses, Including Negatives

The basis in the BRD for the selection of chemicals to use in the validation of *in vitro* ER TA assays was not based on ER binding only. Rather, selection was based on the median $EC_{50}$ values of the chemicals to induce the expression of a reporter gene, as reported in published papers and publicly available reports. **Table 12-1** in **Section 12** (pages 12-9 through 12-10) of the BRD lists thirty-one substances that are recommended for validation of *in vitro* ER TA agonist assays. The median

$EC_{50s}$ for these chemicals range from 8.85 (methoxychlor) to 0.000011 μM (17α-ethinyl estradiol), although no indication of the variation around these values was provided (many of the values were derived from a single study). This represents a six log range of $EC_{50}$ values from $10^{-5}$ μM to 10 μM. In addition, the list included one equivocal chemical and five chemicals that gave negative results when tested.

The Panel concludes that the distribution of the recommended agonist chemicals across the range of "potency" responses, based on the $EC_{50}$ values, for agonist activity in reporter gene assays was appropriate.

In **Table 12-2** in **Section 12** (page 12-10) of the BRD, 21 substances are listed as being recommended for validation of *in vitro* ER TA antagonism assays. Based on the published/publicly available data, the substances were categorized qualitatively as positive or negative for ER antagonist activity; 17 of the substances are positive and four are negative.

The Panel concludes that this group of recommended antagonist chemicals seemed appropriate, although selection is based only on their qualitative classification.

However, the Panel has some concerns and qualifications to these responses. Of the 31 chemicals recommended to validate *in vitro* ER TA agonist assays and of the 21 chemicals recommended to validate *in vitro* ER TA antagonist assays, only six chemicals are in common. These are α-zearalanol, zearalenone, phloretin, bisphenol A, coumestrol, and atrazine. The "scientific" basis for these chemicals being the ones that are common between these two assays is not apparent. Are they in common for some reason or is it just by chance? Is the basis for selection that they cover a range of relative binding affinity (RBA) values for binding to the ER?
The Panel recommends that consideration be given to choosing a set of chemicals that will be used and that the selection process be based on a solid scientific rationale such as a combination of existing ER binding, antagonist, or agonist data. (A particular chemical need not have published/public data available for its performance in all three assays, but should have data available for at least one or two such properties).

Of the 31 chemicals recommended to validate *in vitro* ER TA agonist assays and the 21 chemicals recommended to validate *in vitro* ER TA antagonist assays, only five are expected to be negative in agonist assays and only four are expected to be negative in antagonist assays.

The Panel recommends the inclusion of additional negative chemicals in the list to more effectively evaluate the specificity of the assays.

Of the 31 chemicals recommended for validation of *in vitro* ER TA agonist assays and of the 21 chemicals recommended for validation of *in vitro* ER TA antagonist assays, only 16 and eight chemicals, respectively, are on the list of 33 chemicals

recommended for validation of *in vitro* ER binding assays. (see **Table 1** and **Section 12.4**, pages 12-11 through 12-16 of the *In Vitro* ER Binding BRD).

The Panel recommends that complete overlap exist for chemicals to be tested in both the *in vitro* ER binding and transcriptional assays, or at the very least, that a core of chemicals that is common to all *in vitro* assays be developed. The scientific basis for this selection should be distribution across a range of RBAs, $EC_{50}$ values for TA agonism activity, and positive/negative responses for TA antagonism activity.

The Panel does not expect that data for all three effects will be available for each chemical. However, the list should include chemicals covering RBA values ranging from >10 to <0.0001 plus negatives (see **Table 12-1** of the *In Vitro* ER Binding BRD); chemicals with *in vitro* ER TA agonism potencies, as shown in **Table 12-1** of the *In Vitro* ER TA BRD, ranging from $10^{-5}$ μM to >1.0 μM, including negatives; and substances classified as positive and negative for *in vitro* ER TA antagonism activity. During the Expert Panel meeting, there was discussion concerning the possible need for a prevalidation study with regard to identifying the most appropriate *in vitro* ER TA assays to use. If a prevalidation study is undertaken, it would be appropriate to conduct it using a smaller number of chemicals. Nevertheless, the basis of selection of such a smaller group of chemicals should also be based on a solid scientific rationale. **Table 1** can serve as the starting point for chemical selection.

Another concern pertains to coordination of the selected substances with those being proposed for use by the Endocrine Disruptor Methods Validation Subcommittee (EDMVS) of the U.S. Environmental Protection Agency (EPA) for the *In Vivo* Screening Assays. The overlap of chemicals should also be reviewed for *in vivo* Tier I and Tier II studies under consideration by the EDMVS and the Organisation for Economic Co-operation and Development (OECD) Global Harmonization Program so that responses can ultimately be compared throughout the entire screening and testing battery.

The Panel recommends close collaboration and cooperation regarding the chemical selection process by ICCVAM with the *in vivo* test validation studies being reviewed by the EDMVS and OECD.

### 4.1.2 The Number and Range of Substances by Chemical Class

The 31 chemicals selected for the validation of *in vitro* ER TA agonism assays and the 21 chemicals selected for the validation of *in vitro* ER TA antagonism assays represent a wide range of chemical classes. However, a couple of notable deficiencies were identified. In particular, no phthalates or polychlorinated biphenyls were included. In addition, only two polycyclic aromatic hydrocarbons were selected, two for antagonism assays and only one for agonism assays. However, evaluation of these substances could follow the prevalidation or validation steps.

The Panel recommends that these deficiencies in the chemical list be considered as a revised list of substances is developed as recommended above. However, it may be more appropriate to address these issues following the initial prevalidation studies referred to above.

Specific compounds in **Tables 12.1** and **12.2** of the BRD identified as chiral are (this may not be all): zearalenone, β-zearalenol, *o,p*'-DDT (dichlorodiphenyltrichl oroethane), naringenin, and heptachlor. Specific substances in **Tables 12.1** and **12.2** that may have multiple isomers present (positional isomers that may or may not be chiral): zearanol, chlordane (cis and trans), methoxychlor (*p,p*' and *o,p*'), dicofol (*p,p*' and *o,p*'). Chiral components that might be present are: both zearanol isomers, cis and trans chlordane, *o,p*'-methoxychlor, *o,p*'-dicofol. It is possible for one enantiomer to have agonist and another antagonist activities and for the racemate to be neutral.

The Panel recommends that chiral compounds be evaluated as the racemate and as individual enantiomers, as available. Efforts should be made to test only pure isomers when possible (e.g., cis and/or trans chlordane, *p,p*'-methoxychlor, *p,p*'-dicofol, etc.) and to provide analytical data from suppliers indicating what the isomer and/or enantiomer ratio is so data can be related to others in the validation study. The Panel felt that investigation of chiral issues may, in many cases, be deferred until prevalidation studies identify an optimal assay(s)/protocol(s) to validate further.

It is not clear that the Chemical Abstracts Service Registry Numbers (CAS RN) given for zearanol, chlordane, methoxychlor, and dicofol are the designation for the isomer mix or for one pure isomer.

The Panel recommends that CAS numbers should be checked since they are sometimes different for commercial grade mixtures compared to pure compounds.

### 4.1.3 The Number and Range in Substances by Product Class

The chemicals selected cover a range of products from the pharmaceuticals, natural products, chemical intermediates, and pesticides. Natural product chemicals appear at a somewhat higher frequency and pesticides seem appropriately represented.

The Panel feels that the range of products is appropriate for a validation study.

**Table 1**
**Substances in Common Between *In Vitro* ER Binding Assays and *In Vitro* ER TA Agonism and Antagonism Assays**

| *Substance* | *Median EC50 (mM)** | *Median RBA** | *Agonist* | *Antagonist* |
|---|---|---|---|---|
| Diethylstilbestrol | 0.0000189 | 214 | Agonist | ------- |
| 17β-Estradiol | 0.000098 | 100 | Agonist | ------- |
| Estrone | 0.00063 | 48 | Agonist | ------- |
| Zearalenone | 0.002 | 44 | Agonist | positive |
| Coumestrol | 0.0168 | 1.9 | Agonist | negative |
| Estriol | 0.0348 | 16 | Agonist | ------- |
| 4-*tert*-Octylphenol | 0.05 | 0.20 | Agonist | ------- |
| Genistin | 0.075 | 0.56 | Agonist | ------- |
| Phloretin | 0.03 | 0.069 | Agonist | positive |
| Bisphenol A | 0.45 | 2.6 | Agonist | positive |
| *o,p*'-DDT | 0.66 | 0.013 | Agonist | --------- |
| Naringenin | 1.0 | 0.008 | Agonist | -------- |
| *o,p*'-DDT | 2.0 | 0.003 | Agonist | -------- |
| Methoxychlor | 8.85 | 0.001 | Agonist | -------- |
| Progestrone | equivocal | 0.0003 | Agonist | -------- |
| Atrazine | negative | 0.0003 | Agonist | negative |
| 4 Hydroxytamoxifen | positive | 175 | -------- | positive |
| Tamoxifen | positive | 3.1 | -------- | positive |
| 4-Octylphenyl | positive | 0.005 | -------- | positive |

*Values from Tables in *In Vitro* ER Binding and *In Vitro* ER TA BRDs.

**Additional Panel Recommendations**

The Panel's recommendations are not only for *in vitro* ER TA assays but also some general comments that are applicable to the performance and use of these types of *in vitro* assays.

1. The Panel recommends the development of datasets for statistical analysis (i.e., confidence limits), to assess reliability and to determine the impact of variability. Pre-validation studies should be carried out to provide data for evaluation by the statisticians.

2. There is a need to assess the ability of these *in vitro* screens to predict *in vivo* responses. One way to accomplish this is to make sure that substances to be tested in the *in vitro* screens are also tested in the *in vivo* screens and tests so that information and the "weight of the evidence" can be assessed for particular chemicals.

3. More comprehensive *in vitro* methods are needed to detect endocrine disrupters. The regulatory community does not need to depend forever on the current assays. As methods develop and are refined, their utility in the screening process need to be evaluated. Different protocols need to be further optimized and the most reliable assays identified.

4. The Panel encourages the development of a centralized repository of chemicals with verified purity that can be used across assays.

5. It is well established that ER mediated gene expression is species, tissue, cell and promoter context specific. Consequently, the agonist and/or antagonist activities of a ligand cannot be generalized or extrapolated to all genes based on a single assay. Moreover, *in vitro* ER TA assays use artificial reporter genes (i.e., engineered with multiple ERE and minimal promoters), further limiting their utility for predicting in vivo ligand activity. Therefore, a sequential testing strategy is recommended for *in vitro* AR TA agonism/antagonism assays. If a substance induces a positive effect in any of these assays, testing in additional *in vitro* ER and AR binding or TA agonism/antagonism assays should not be conducted before proceeding to short term Tier 1 *in vivo* studies. The rational for this recommendation is that a positive response in these assays only demonstrates a functional consequence and in TA assays, an agonist/antagonist classification for this artificial response system. It is highly likely that the substance will elicit a broad spectrum of agonist and antagonist gene expression responses *in vivo* and therefore, further classification of the activity of the substance using TA assays will provide little additional information that will assist with prioritization and subsequent in vivo study design.

   If the substance fails to induce an agonist response in an *in vitro* AR TA assay, antagonist activity should be investigated since some substances may only exhibit antagonist activity in the TA assay being used. Similarly, if a substance exhibits agonist or antagonist activity in an *in vitro* ER TA assay, it is questionable if testing for AR activity will provide significant additional information since the positive ER-mediated activity will trigger further short-term Tier 1 *in vivo* testing.

   The Panel recommends that a sequential testing strategy be evaluated as part of a pre-validation assessment of *in vitro* AR TA assays in order to determine the value of performing

agonism and antagonism studies for estrogen and androgen receptors and how information from these assays are used to decide subsequent short-term Tier 1 *in vivo* testing.

6. Standardization and validation across laboratories performing these studies for regulatory decision-making is critical and must occur before these assays are used for regulatory purposes. The work that is outlined in the BRD suggests that within laboratory variability for some of the assays is acceptable and the studies are reproducible. However, little between laboratory standardization has been performed. A formal validation process is needed in order to establish a "gold standard" study for use and not have many personal variants of similar assays.

# III. *In Vitro* Androgen Receptor (AR) Binding Assays

## 1.0 Recommendations and Priority for Validation Studies

The AR Binding BRD reviews the comparative performance, reliability, advantages, and disadvantages for different *in vitro* AR binding assays, and recommends a relative priority for further development and/or validation based on this information.

### 1.1 Considering the intended use of the assays as a toxicological screen, is the Panel aware of other advantages and disadvantages for the assays discussed in the BRD?

Many of the advantages and disadvantages are presented in the BRD, and these, for the most part, are reasonable. However, additional comments by the Panel are found below.

The Panel rejects the two principal recommendations of assay protocols put forward in the BRD. Neither the rat prostate cytosol (RPC) nor the transient transfection of COS cells with the human AR expression vector were deemed to be optimal for *in vitro* AR binding assays. Rather, based upon scientific rationale, the Panel proposes and recommends that a high-throughput assay for AR binding be developed using purified, recombinant full-length AR protein. This recommendation is concordant with that of the *in vitro* ER binding Panel members for validation of binding assays based upon purified, recombinant receptor proteins.

For the purposes of organization, the responses to this question have been organized by the type of assay (i.e., cytosol-based, cell-based, and use of purified AR). Other comments/issues relevant to many of the assays are also summarized below. Recommendations, where applicable, are also included. For the purposes of the BRD, it might be useful to construct a summary table listing the type of assay along with the relative advantages and disadvantages of each assay.

**A. Cytosol-based Assays (RPC Assay):**

The RPC assay has historically been the assay most frequently used in published studies of AR binding. The experimental protocol for this assay was described in greatest detail within the BRD. This assay is currently being used by the U.S. Environmental Protection Agency (EPA) for testing the ability of test chemicals to bind to the AR. Although the Panel recognized several advantages of this assay, it was reluctant to endorse this assay for future studies based upon a significant number of disadvantages.

Among the advantages of the RPC assay was its description as the "gold standard" for *in vitro* assays of AR binding. As such, this is particularly useful as a reference. Another advantage that should not be minimized is the fact that the rat prostate expresses endogenous AR and the AR functions to regulate specific genes in the prostate. Although this latter advantage may seem obvious, it is significant with respect to other, heterologous cell-based assay systems in which the AR is expressed from a foreign expression vector.

A significant number of disadvantages were recognized within the RPC assay. This assay was viewed as an inappropriate use of animals in a single type of assay for which the small (500-700 mg) ventral prostate tissue was harvested. The assay is time consuming, for, among other reasons, the animals must be castrated 18-24 h prior to the harvest of tissue to reduce the binding of endogenous androgen to the receptor. The RPC is an assay of rat AR and not human AR. The AR is notoriously labile in *in vitro* systems, including the RPC assay, and significant methodological precautions must be taken to avoid its degradation. The RPC assay measures ligand binding, and not any functional aspect of the AR. A cell-free binding assay by nature cannot distinguish between agonists and antagonists. The *in vitro* binding of chemicals and ligands to the AR in the RPC assay occurs at 4°C (i.e., it does not occur under physiologic conditions of temperature or intracellular environment). RPC is a crude tissue cytosol preparation that contains many proteins in addition to AR, including other endogenous steroid receptors that may interfere with the assay. In addition, some metabolism of the reference ligand and/or test chemical may occur even in cytosol preparations incubated at 4°C.

Although the RPC assay has been widely used for many years to assess the binding properties of the rat AR, the Panel recommended that other AR binding assays be considered and developed in place of the RPC assay.

**B. Cell-based Assays (COS + hAR Assay):**

A second assay described in the BRD was based upon AR binding studies conducted in heterologous cells (e.g., COS monkey kidney cells) that were transfected with a human AR expression plasmid to express the receptor. The advantage of such a system is that it models whole cell, physiologic conditions for ligand binding. The fact that only the AR and not other receptors are expressed in transfected COS cells is an advantage. Moreover, the AR in this assay system is most often human, but the AR of other species can also be expressed in the COS cells to assess the binding of chemicals that may be relevant to a particular animal, fish or amphibian species.

Although numerous research laboratories have utilized this assay to characterize basic functional properties of the human AR, a number of disadvantages can be cited relevant to its use in evaluating and validating the binding of chemicals to AR. As mentioned above, the hAR expression vector must be transfected into COS cells. In addition to being labor intensive, this procedure has the potential of being highly variable between laboratories, especially in the absence of detailed methodological protocols. Only a fraction (which also can be quite variable) of the cells will express AR, and the expression is most often artificially high within individual cells. When gene expression occurs following transient transfection, the gene is not subject to normal restrictions of chromatin structure. COS cells do not normally express AR and therefore, the intracellular environment is artificial compared to an androgen-responsive cell. If polymorphisms for the hAR have functional significance, the effects of this genetic variability will not be appreciated in cells that express a single form of AR from an expression vector. In an intact cell system, access of different

ligands to the AR may be determined by differential kinetics of transport across the cell membrane, of metabolism, and of binding to cellular components. Similarly, it may not be clear if metabolism of the test chemical has taken place to a metabolite that has no binding, some binding or even higher binding to AR. The derivation of COS cells from monkey kidney also suggests that species differences in the degree and type of metabolism that takes place in these cells may not accurately reflect the human situation in target cells.

In summary, the Panel suggests that the use of a stable-transfected cell line would preclude some of the difficulties inherent to the use of transient transfection assays. However, the Panel recognizes that stable cell lines are also prone to instability over time in culture. The Panel recommends consideration of other human cells lines that might be amenable to establishing transient or stable transfection/expression of the human AR.

**C. Assay Using Purified hAR:**

A third assay system that was not included in the BRD, but which garnered the most support among members of the Panel, was the development of an *in vitro*, high-throughput AR binding assay based upon the use of a purified, recombinant human AR. By definition, this assay would be the most efficacious and time-saving of the potential assay systems. It eliminates any problems associated with the use of animals or cells. Large amounts of recombinant protein could be produced in bulk and supplied as a homogeneous, uniform preparation to all test laboratories. Assurances regarding the steady production and availability of recombinant AR protein would, however, need to be ascertained from potential suppliers. The use of purified, recombinant protein can be readily adapted to high-throughput methods of analysis. Disadvantages regarding the potential absence of biologically significant post-translational modifications of the recombinant AR protein and the absence of a putative biologically relevant environment during the conduct of the binding assays were noted.

The Panel enthusiastically recommends, with highest priority, the development of an assay using purified, recombinant AR from human, rat or another mammalian species.

**1.2 Considering the intended use of the assays as a toxicological screen, does the Panel agree with the relative priority recommended for these sets of assays? Does the Panel recommend any changes in priority, or have specific recommendations for prioritization?**

As stated previously, the Panel strongly encourages the development of an AR binding assay based upon the use of a purified, recombinant form of AR.

**1.2.1 In considering prioritization, is RPC the best source of AR for binding assays?**

No. The Panel considers the disadvantages of the RPC assay to be such that this assay should not be assigned a high priority. The disadvantages that were cited above include the fact that RPC contains other steroid receptors that may interfere

with the assay for AR binding and that ligand metabolism may take place in cytosol preparations. The RPC cannot substitute for human AR, or for AR in those wildlife where significant exposure to androgenic chemicals may occur. Considering the number of disadvantages attributed to the RPC assay, alternatives should be sought.

**1.2.2 Should the binding of substances to different receptor subtypes be addressed in the binding assays?**

Yes. This is prudent if it is known that subtypes with known functional significance exist. Although only a single type of AR is known to exist in the human, this may be particularly relevant to other non-mammalian species, such as the rainbow trout.

**1.2.3 Should a metabolic activation system be included in the binding assays?**

No, at least not currently in the context of routine AR binding assays. The Panel considers the evaluation of the binding of specific test chemicals to AR to be the first priority such that the binding of chemical derivatives of the parent substance resulting from metabolism were irrelevant to the present mandate of the proposed studies. Moreover, there is currently no definition of an activation system appropriate for each tissue or species of concern. The Panel recognizes the usefulness of having a system in which the binding assay was coupled to metabolic activation, if there are indications of an AR-binding chemical that was generated as a metabolite of the parent substance.

NOTE: The Panel bases its recommendations of AR binding assays on the basis of scientific considerations and relevance. However, it is critical to point out that there are other issues that influence the implementation of our recommendations. Most significantly, the human AR cDNA sequence is protected by at least two different U.S. Government patents. Furthermore, the commercial/non-academic use of hAR in cis-trans functional assays of AR is further governed by a license granted to Ligand Pharmaceuticals (San Diego, CA). The rat AR cDNA sequence may/is also protected by a patent. Although a commercial source of recombinant AR-ligand binding domain protein is currently available, the reliability of this preparation in binding assays has not been established. A full-length recombinant form of AR from any species is presently not available. An AR sequence from a species closely related to human may be necessary to allow the development of such an assay.

The appropriate government agencies should investigate the status of patents and licenses attendant to the use of the human and rat AR and should provide guidance for the use and development of AR assays in the public/private domains.

## 2.0 Minimum Procedural Standards for *In Vitro* AR Binding Assays

**2.1 To facilitate assay standardization, the BRD proposes minimum procedural standards that should be incorporated into *In Vitro* AR Binding Assay protocols (Section 12.2). Considering the intended use of the assays as a toxicological screen, does the Panel agree with the adequacy of the proposed procedural standards? If not, what changes should be made to each standard and why?**

The Panel agreed with the critical methodological issues proposed in the BRD for *in vitro* AR binding assays, and endorsed the fact that any assays using animals must be undertaken under the guidance of the relevant IACUC.

### 2.1.1 Binding Constant ($K_d$) of the Reference Androgen

There was consensus that a specific binding capacity, $B_{max,}$ and the dissociation constant, $K_d$, values for the reference receptor protein is a critical measure of the robustness of the procedure. The $K_d$ should be clearly established for the reference androgen in each assay and all test laboratories should be able to generate comparable data within acceptable limits. At present, data from different laboratories do not establish a well-defined, highly replicated $K_d$ for R1881 in any of the test systems for AR binding. A minimum number of concentrations of ligand should be identified for generating a $K_d$. The Panel recommends adoption of seven concentrations of ligand for analysis, as is implied later in the BRD.

The Panel agrees that a detailed assay protocol must be provided for performing each type of assay, accompanied by criteria for evaluation and acceptance of results, to demonstrate assay validation and transferability across laboratories. The ability of a laboratory operator to obtain a $B_{max}$ and $K_d$ value for the reference androgen of a reference receptor protein within accepted limits for that type of preparation is a critical quality control parameter in that laboratory. These data will be essential to the establishment of a Quality Assurance Program for endocrine disruptor substance evaluation at numerous laboratories. The Panel recommends that straightforward procedures for determination of both the $K_d$ value of the radiolabeled reference androgen, and the $K_d$ value of an unlabeled test substance, such as the ligand titration array (Raffelsberger and Wittliff, 1997), be considered. Software programs such as Compete® (Lundon Software, Inc., Cleveland Heights, OH), GraphPad Prism® (GraphPad Software, Inc., San Diego, CA), LIGAND (Munson and Rodbard, 1980) and OneSite® (Lundon Software, Inc., Cleveland Heights, OH), will determine the specific binding capacity and $K_d$ values of the reference AR preparation, and analyze the competition results and compute the $K_d$ value of the unlabeled test substance examined in the assay. The Panel concludes that additional studies are necessary to specify the precise statistical characteristics of AR binding when fitting nonlinear regression curves and estimating pertinent parameters such as $K_d$ and the $IC_{50}$.

### 2.1.2 Reference Androgen

The choice of a reference androgen is, in part, dependent on the assay being used. If the assay is based on a purified AR, then using the natural ligand [i.e., 5α-

dihydrotestosterone (DHT)] would make the most sense and should be the preferred ligand for standardization. If the test is based on a crude cytosolic preparation or on a cell line, then selecting a molecule that does not undergo significant metabolism is important.

The use of R1881 is based upon its properties as a high affinity AR ligand, its lack of metabolism and its low nonspecific protein binding in whole cell and crude cellular extracts. Thus, in this case, R1881 seems appropriate as a ligand. However, R1881 is a synthetic substance and may not recapitulate all of the properties of the endogenous ligands, testosterone or DHT. R1881 will also bind to the progesterone receptor (PR) in binding assays based on cells or tissues that contain this receptor, as for instance RPC. Since the specific binding of R1881 to AR is confounded by the presence of PR in the sample (e.g., RPC), triamcinolone acetonide should be used to block binding of R1881 to PR.

During the Panel deliberations, none of the members expressed any knowledge of known interference between triamcinolone acetonide and other chemicals. However, the Panel does not believe that it has sufficient understanding to predict if the potent synthetic glucocorticoid, triamcinolone acetonide, will interfere in the subsequent evaluations of androgen mimics. An alternative might be to use the synthetic progestin, promegestone (R5020) that has a high affinity for PR, as a means of diminishing the contribution of PR binding. Another alternative would be to use mibolerone. This ligand interacts less avidly with PR. Therefore, if an assay is chosen based on radioactivity measurement and uses PR-containing cells or tissue, consideration should be given to the use of mibolerone, rather than R1881. This would avoid the concomitant use of triamcinolone acetonide to block binding to PR in the AR binding assay.

### 2.1.3 Preparation of Test Substances

All test substances should be standardized and prepared according to rigorous quality controls for purity and concentration. The test substances should be prepared in water, 95-100% ethanol, or dimethyl sulfoxide (DMSO), depending upon their solubility. Preference should be given to the solvent that allows testing of the maximal concentration of the test substance, without exceeding the limit dose. However, in situations where more than one solvent could be used, preference should be given to water, then 95-100% ethanol, and then DMSO, in that order. The Panel strongly agrees that 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. In the experience of the Panel, many test substances will require DMSO as a solvent, and again solvent-only controls must be performed. In situations where more than one solvent could be used, preference should be given to water or 95-100% ethanol, in that order. Preparation of stock solutions should be performed under rigorous quality control standards. The stability of stock solutions should be determined on an on-going basis and it may be necessary to prepare some chemical solutions fresh at each time of use.

The Panel recommends that the U.S. EPA establish an inventory of high purity chemicals that can be provided to laboratories as required for validation and test studies.

### 2.1.4 Concentration Range of Test Substances

In general practice, substances should be tested over a wide range of concentrations. It is desirable to have a concentration high enough to produce some effect in an assay even though very high concentrations (mM range) of a test chemical may be unrealistic when compared to levels found in the environment or to those obtained after normal exposure. The Panel agrees with the recommendations that at least seven different concentrations of the test substances within the range 1 nM to 1 mM should be examined to increase the likelihood of developing a competition curve satisfactory for $IC_{50}$ analysis. Furthermore, it is important to have at least one concentration below the $IC_{50}$ of the test substance. The limit concentration should be 1 mM, taking into consideration the solubility characteristics and possible toxicity (e.g., denaturation of the receptor) of the test substance. If a lower maximum concentration is tested due to solubility constraints or excessive toxicity, the number of concentrations tested can be adjusted to account for the altered concentration range. The concentration range should be governed by practical considerations of the chemistry of each substance, which determine its solubility in a specific solvent.

### 2.1.5 Solvent and Positive Controls

A set of solvent controls, with solvent concentrations identical to those used with the reaction mixtures containing the test substance must be included in each set of assays. Solvents should be the standard ones used (i.e., water, 95-100% ethanol, DMSO). The solvents should not have any effect on AR binding at the concentrations used in the assays and should be of utmost purity. The solvent volumes must remain constant throughout the concentration range tested. The positive control substance should have a binding affinity within two orders of magnitude of the limit of sensitivity of the assay. This control is critical to the assessment of the reproducible sensitivity of the assay within and between laboratories and is of particular relevance in determining the ability to assign substances with low AR binding affinity as different from no binding. A second positive control (e.g. cyproterone acetate) is recommended since this substance has an RBA within the range of 1-10% of the RBA of the reference androgen.

### 2.1.6 Within-Test Replicates

The $IC_{50}$ value should be based on triplicate measurements at each dose level.

### 2.1.7 Dose Spacing

The Panel agrees with the recommended dose spacing of one order of magnitude in the concentrations of the candidate test substances, and in the use of half-log doses in certain cases.

**2.1.8 Data Analysis**

The Panel recommends the essential requirement to determine and compute both the $B_{max}$ and the $K_d$ value for AR binding in each assay. Alternative approaches such as ligand titration array provide a simultaneous evaluation of a laboratory's performance and determination of the AR binding properties (e.g., $IC_{50}$, $K_d$, and $K_i$ values) of each test substance. It may be useful to determine the noncompetitive, competitive and uncompetitive nature of AR binding with specific chemicals that demonstrate unexpected binding curves. More detail is needed on statistical models for nonlinear regression to assess $K_d$, $K_i$, and $IC_{50}$ values. The mode of calculation and assumptions for the statistical methods must be justified. This includes the nature of the statistical characteristics of the data (distribution, variance patterns, specific nonlinear models, etc,) and how to fit the models. When doing so, confidence limits must be calculated for $K_d$, $K_i$ and $IC_{50}$ values. From these, details on how to make pertinent and valid statistical inferences should be specified.

A possible approach for developing these statistical characteristics is to conduct a set of carefully-designed, comprehensive interlaboratory negative control studies. These would enable better understanding of the underlying statistical variability in AR binding data.

**2.1.9 Assay Acceptance Criteria**

There is agreement among the Panel with each of the BRD acceptance criteria. In addition, the Panel recommends that the assays be performed in compliance with standard Good Laboratory Practice guidelines.

**2.1.10 Evaluation and Interpretation of Results**

The approach presented in the BRD is accepted by the Panel. The designation of "equivocal" for substances that do not bring about a 50% reduction in specific AR binding is accepted. The classification of a test chemical as "positive for binding" requires the use of statistical methods.

**2.1.11 Test Report**

All of the BRD recommendations are accepted by the Panel. In addition, 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 $IC_{50}$, $K_d$, and $K_i$ values with a standardized AR preparation using a selected set of test substances should be established.

**2.1.12 Replicate Studies**

The Panel agrees with the recommendations contained in the BRD.

**2.2 Considering the intended use of the assays as a toxicological screen, are there other minimum procedural standards that should be included? If so, what are they and why?**

Specific recommendation regarding the type of protein assay and the conditions would

be useful. The Panel highly recommends that both the androgen and estrogen receptor-based assay reactions contain 10 mM sodium molybdate as well as a cocktail of protease inhibitors such as those described on page B3-5 of the BRD. Although the hydroxylapatite (HAP) procedure has been used by numerous investigators for separation of free and bound ligand, and is recommended for the RPC assay, problems with this separation procedure may arise. One such problem is that the receptor-ligand complexes are bound to the matrix, their retention is required during washing and the association of receptor-ligand complexes must be retained during elution. These represent significant variables that must be controlled. By contrast, the U.S. Food and Drug Administration (FDA) approved the testing of human tissues with the dextran-coated charcoal procedure in combination with a ligand titration assay. The latter procedure with dextran-coated charcoal allows the receptor-ligand complexes to remain in the original reaction medium while removing the unbound ligand. The results may be far more reproducible with this method than those obtained with the HAP assay.

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

**3.1 A standardized *In Vitro* AR Binding Assay protocol using rat prostate cytosol (RPC) is provided in Appendix B of the BRD. This assay is proposed for validation studies by the U.S. EPA and other sponsors. Section 12.3 discusses additional details that should be added, based on the minimum procedural standards in Section 12.2. In addition, an example *In Vitro* ER Binding RPC Assay (based on the U.S. EPA Protocol) is provided in Section 12 Annex of the BRD. Considering the intended use of the assays as a toxicological screen, would the current protocols, with the additions detailed in Sections 12.2 and 12.3, provide a level of detail to appropriately minimize interlaboratory variability? If not, what revisions or additions should be made to the protocols?**

The Panel is in agreement regarding the lack of an existing, standardized, acceptable protocol for an AR binding assay. Little effort and no synthesis were put forward by the BRD with respect to Question 3. The text on pages 12-8/12-9 of the BRD merely provided a list of the four documents in Appendix B and indicated that there was a need to review these protocols "for completeness and adequacy for their intended purpose." Appendix B1 is a detailed description of the protocol presently being used by the U.S. EPA to validate the RPC assay. Appendix B2 is a brief description of the COS cell binding assay. The information, as provided, does not allow another researcher to reproduce the work nor does it provide the rationale for inclusion of most of the steps. Appendix B3 is similar in scope to B1, but is much less well written and has numerous severe omissions/errors. This protocol should not receive further consideration. Appendix B4 does not provide a protocol, but rather gives a valuable list of general concerns, cautions and guidelines on how to put such an assay together.

The standardized protocol for the RPC assay is provided in great detail. Although the RPC assay has been designated as the "gold standard", this is the more difficult of the assays to perform in a standardized format. The relative simplicity of the transfected cell assay (COS + hAR/rAR) is amenable to high throughput and requires simple methods and minimal volumes and variations of buffers and solutions. If this assay is to be pursued further, a

standard transfection protocol based upon commercially available transfection reagents and a standardized cell line would be necessary for these assays. The production of a stable cell line expressing AR would avoid the problems inherent in transient transfection assays.

Perhaps, the simplest and least amenable assay to inconsistency would be one in which the AR protein would be fixed in multi-well plates and tracer and test ligands added in appropriate amounts to develop data appropriate for the accurate calculation of the $K_d$, $K_i$, and $IC_{50}$ values, or other pertinent data analyses. The desirability of moving away from radioactive tracer ligands and toward more environmentally and safer fluorescent ligands also needs to be considered.

**3.2 In addition to the minimum procedural details listed in Section 12.2, are there other protocol elements that should be considered for other *In Vitro* AR/ER Binding Assays recommended for validation as a toxicological screen, including those protocols provided in Appendix B?**

- If a radionuclide is to be used as the tracer ligand, its chemical and radiochemical purity must be stated.
- The maximal time of storage at -80°C/-20°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.
- Details are missing from the COS-cell binding assay. Some of these include preparation, purity and stability of the AR vector, more detailed timing on cell transfections, confluency of cells, transfection efficiencies (what these should be or if it makes a difference). Some rationale for the choice of timing, incubation conditions, etc., should also be given, especially since equilibrium conditions are sought.
- If a cytosolic protein preparation is to be used, a cocktail of protease inhibitors, rather than a single inhibitor, is to be used to increase stability of the AR.
- The simplest and least amenable assay to inconsistency 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 binding analyses.
- The desirability of moving away from radioactive tracer ligands and toward more environmentally and safer fluorescent ligands also needs to be considered in assay development.
- In developing such an assay, it is important to avoid situations that may render the assay less readily acceptable at the international level (e.g., having to comply with patent regulations and regulations regarding the use of radionuclides).

**3.3 Considering the intended use of the assays as a toxicological screen, is the Panel aware of other available standardized protocols for assays recommended for validation?**

Pan Vera Corporation (Madison, WI, USA) is selling an 'AR Ligand Binding Domain: Activity Assay' based on radioactivity measurement. The sensitivity and reliability of this assay are not apparent. Furthermore, the present Pan Vera AR assay uses only the ligand binding domain recombinant protein; this is much less desirable than the use of the full-length AR in either an *in vitro* or *in vivo* assay. There is no indication that a full-

length recombinant AR will be available in the near future. The present ER, progesterone receptor (PR) and glucocorticoid receptor (GR) competitive binding assays from Pan Vera Corporation are based upon full-length recombinant proteins and do not use radioactivity, rendering them much closer to the idealized assay described above. Given the development of non-AR assays, it would be surprising that Pan Vera Corporation does not have under development the type of assay discussed above as the optimal one. The Panel is not aware of any other assay under development that would meet the desired criteria.

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

**4.1 Section 12.4 provides a list of substances recommended for use in validation studies of *In Vitro* AR Binding Assays. Considering that the intended use of the assays are as a toxicological screen, does the panel agree with the selection criteria, adequacy and appropriateness of substances recommended for validation studies, in terms of the following issues? If not, what substances should be added or deleted?**

Generally:

- The Panel members essentially accept and/or recommend the list of test chemicals for validation of the assays.
- The same range and types of substances are recommended for validation of both AR binding and TA (transcriptional assays), if both assay types are selected for further validation.

Specifically:

- Weak-positive reference chemicals, which comparably represent the breadth of environmental chemicals, should be available.
- There are several "obvious" substances missing from the list. Anti-androgenic chemicals such as flutamide (or hydroxyflutamide if used *in vitro*) and bicalutamide that bind to AR but do not activate its transcriptional activity, should be used as model chemicals. Finasteride (the commercially available 5α-reductase inhibitor) does not bind to AR and should be added as a negative control.
- An assumption has been made regarding the mode of action with AR, such as competitive ligand binding (i.e., substances bind to the same site as endogenous androgens). As mentioned previously, consideration should also be given to non-competitive and uncompetitive mechanisms.
- One or more of the estrogens (ethinyl estradiol, estrone, or diethylstilbestrol) could be omitted from the list, as 17β-estradiol is already included.

### 4.1.1 The Number and Distribution of Substances Across the Range of Measurable AR Binding Activity, Including Negatives

A total of 31 chemicals are recommended in the AR binding BRD. They are almost equally divided among those with higher binding affinities in the range of 10 – 0.1 RBA relative to the R1881 and those with considerably lower binding affinities in the range of 0.01 – 0.0001 RBA. However, only three of the chemicals listed are negative and it is necessary to expand this number.

### 4.1.2 The Number and Range of Substances by Chemical Class

Of the 31 chemicals listed in the BRD, 21 are steroids of endogenous biological origin or are pharmaceuticals. Many of the other chemicals are organochlorines. Chemicals in the class of polychlorinated biphenyls, phthalates, and heavy/organo metals are not represented and should be.

### 4.1.3 The Number and Range of Substances by Product Class

The steroids are represented by 21 chemicals and the remainder represents pesticides, herbicides, and fungicides. Additional polychlorinated biphenyls, phthalates, and heavy/organo metals should be included.

# IV. *In Vitro* Androgen Receptor (AR) Transcriptional Activation (TA) Assays

## 1.0 Recommendations and Priority for Validation Studies

The *In Vitro* AR TA BRD reviews the comparative performance, reliability, advantages, and disadvantages for different *in vitro* AR TA assays, and recommends a relative priority for further development and/or validation based on this information.

**1.1 Considering the intended use of the assays as a toxicological screen, is the Panel aware of other advantages and disadvantages for the assays discussed in the BRD?**

Advantages and disadvantages of these assays are discussed below.

**1.2 Considering the intended use of the assays as a toxicological screen, does the Panel agree with the relative assessments of the protocols as described in the BRD? Should any of these be considered for further evaluation? What specific aspects have not been addressed that should be considered in the formulation of the ideal protocol for screening potential AR agonists and antagonists?**

Only one type of assay was recommended in the BRD for further study. This resulted in part from the expected limitations imposed by patent restrictions on the AR clone. The recommended protocol involves the use of the endogenous AR and a stably integrated reporter. Only one assay referenced in the BRD meets these criteria and made use of an MDA-kb2 cell line. The MDA-MB-453 cell assay discussed is not a stable cell line. An analysis of these assays required a review of the original publications as fold induction of luciferase activity with 5α-dihydrotestosterone (DHT) was provided for the MDA-MB-453 cells, but the response to cortisol was not provided. Both of these cell lines are complicated by the lack of steroid receptor specificity in transactivation of the mouse mammary tumor virus luciferase (MMTV-Luc) reporter and by a lack of specificity in terms of androgen induction by the AR.

MDA-kb2 is a breast cancer cell line with endogenous AR and glucocorticoid receptor (GR) and has a stably integrated MMTV-luciferase reporter (Wilson et al. 2002) and derives from MDA-MB-453 cells. The response was 3-9 fold for DHT and 1-19.5 fold for dexamethasone. Hydroxyflutamide was used to inhibit AR agonists to differentiate activities mediated by AR and GR. A disadvantage of the assay is that weak AR agonist activity could be difficult to detect due to the weak response to DHT (only up to 9 fold). After 40 passes of the cells, the luciferase response to 1 nM DHT dropped to 5-6 fold. Thus, relatively low cell passage numbers would be required for transactivation assays and a need to continually monitor the cell line for genetic drift and loss of activity. Another potential problem is that hydroxyflutamide has only a moderate binding affinity for AR so its inhibitory effect is lost in the presence of high agonist activity; 1 μM hydroxyflutamide did not inhibit the activity of 10 nM DHT. In addition, it might be difficult to differentiate activity mediated by GR from the ineffectiveness of hydroxyflutamide to inhibit the agonist response. The reporter vector was activated by 17β-estradiol by the AR. Overall, the method may fail to detect weak AR agonists present at relatively low concentrations and is complicated by the presence of endogenous GR.

The issue of estradiol activation of the AR in the *in vitro* assays is complicated by the fact that there is no physiological evidence to support that estradiol is an AR agonist. Therefore, the question arises whether the positive *in vitro* response to estradiol as an AR agonist is meaningful. The AR is known to be promiscuous in terms of ligand binding and can accommodate a wide range of steroids. For various mechanistic reasons that are beyond the scope of this report, estradiol often results in a positive response *in vitro*. An optimal assay would show that estradiol is not an AR agonist at concentrations below 100 nM; however, thus far the majority of assays show estradiol as an AR agonist at concentrations of 10 nM or higher. It is recognized that agonists working through this *in vitro* mechanism may be false positives compared to *in vivo* results. Ideally, the *in vitro* assays should predict *in vivo* activity.

MDA-MB-453 cells are human breast cancer cells that contain endogenous AR and GR. These are not stably transfected cells in contrast to what is indicated in the BRD. For each assay, cells are transduced with a recombinant adenovirus containing the MMTV-Luc reporter. The response was 24 fold with 0.1 nM DHT and 248 fold with 1 µM dexamethasone (Hartig et al. 2002). The level of induction by DHT meets the suggested minimum performance criteria of fold induction by the control androgen. No method was proposed to account for the high transactivation of the reporter by GR. Presumably, this would be done by selectively inhibiting AR mediated activity with hydroxyflutamide. The presence of endogenous GR and its high activity complicates the assessment of AR agonist activity. The presence of GR would probably not interfere with assessing antagonist activity unless the AR antagonist interacts with GR. AR and GR agonists would be identified simultaneously using these cells. The use of adenovirus infection is advantageous in that virus infections are relatively straightforward. Development of the recombinant adenovirus is complex and the recombinant virus must be made available by the U.S. Environmental Protection Agency (EPA) to laboratories, which might limit applicability of the assay. Despite the limitations of this adenovirus infection protocol, it was deemed the best method of those presented in the BRD to proceed with further evaluation. This decision was tempered by inherent limitations resulting from the presence of GR in the cells. The Panel suggests that additional studies be performed to develop this assay using cells that lack the GR and to develop an adenovirus vector for a different androgen responsive reporter vector that shows greater specificity for the AR.

The other stable cell line (protocol B6) presented also had low fold induction with DHT. All the remaining protocols involve transient transfection and are subject to patent restrictions. Below are outlined some of the major advantages or disadvantages of these assays.

**ADVANTAGES AND DISADVANTAGES**

1. **Different cell lines - presence of other steroid receptors and cell metabolism**
   One of the primary differences among the transient transfection protocols provided in the BRD is the use of different cell lines. For transient and stable transfection experiments, the optimal cell line would be one that expresses only the AR either endogenously or after transfection. HepG2, HeLa, and Chinese hamster ovary (CHO) cells express relatively high levels of GR, making it more difficult to specifically assess

AR activity by different ligands. It should be noted that while the presence of the GR is likely required for a viable cell, some cell lines such as monkey kidney CV1 cells have a sufficiently low level of GR that does not interfere with assays for AR transcriptional activity. Three of the protocols (B2, B5 and B6) utilize CHO cells which respond to cortisol suggesting the presence of GR. Except for the use of hydroxyflutamide to selectively inhibit AR mediated transactivation as described above, the protocols do not provide a method to differentiate reporter gene transactivation between AR and GR. The presence of the GR in the MBA cell line complicates its use as a screening assay for detecting AR agonist activity. This complication can be addressed by co-incubating with an AR antagonist such as hydroxyflutamide but this would require a parallel set of assays for all test substances, essentially doubling the effort. The LNCaP cell proliferation assay should not be considered since this cell line contains a mutant AR that does not discriminate agonists from antagonists. The assays proposed to measure increases or inhibition of cell proliferation are not transcriptional activation assays and could be impacted through multiple cellular pathways that do not necessarily involve the AR, and thus could not be recommended. Additional complications with cellular proliferation assays for the evaluation of endocrine activity is that cellular proliferation may be a consequence of non-AR receptor mediated mechanisms through the activation of cellular signaling cascades (e.g., phorbol esters like tetradecanoylphorbol-13 acetate (TPA) through protein kinase C).

The cell lines also differ in metabolic activity. For example, HepG2 cells derived from liver cells may retain some metabolic activity that could bioactivate or bioinactivate test substances. On the other hand, most metabolic activity, specifically P450 activities, might be lost when liver cells are cultured. Residual activity could confound interpretation of *in vitro* results. Differences in cell metabolism make R1881 the control agonist of choice as suggested in the BRD although DHT should be included as an additional positive control. A potential complication of R1881 is that it is less stable than DHT in solution. Control stocks must be prepared frequently and maintained in the dark in ethanol at -20°C. For yeast, there are potential differences in metabolism from mammalian cells and different chemical transport activities by transporters such as P-glycoprotein homologues. Thus, some parent substances may not gain access to the yeast cell to interact with the AR.

***Recommendation:***

CHO, HepG2, and HeLa cells are less advisable due to the presence of the GR. HepG2 cells have some metabolic activity which could inactivate test substances and/or agonist ligands. CV1 cells have relatively low metabolic activity and no detectable endogenous GR, but also no endogenous AR. There is evidence that CV1 cells metabolize the parent forms of several environmental antagonists to their active forms that interact with AR. The control androgen should be R1881, as recommended in the BRD, due to possible metabolism of natural androgens, but DHT should be included as an additional control. However, the inclusion of a metabolic activation system in *in vitro* AR TA assays is not recommended at this time. The type of metabolic activation

system developed will depend on which *in vitro* assay(s) are considered validated for detecting endocrine disrupting substances.

2. **Stable versus transient expression**
Cell lines with stably integrated reporter vectors and endogenous AR are advantageous because there are fewer patent issues. It should be noted, however, that cell lines with stably integrated AR plasmids are subject to AR patents restrictions since the AR plasmid DNA was integrated. A potential problem with stable cell lines is that they are not completely stable and tend to lose the integrated plasmids. They usually require continuous drug selection that may be expensive. The advantage of stable cell lines is that time consuming transfection procedures are not required. Variance in signal response tends to be less in stable versus transient transfection assays. On the other hand, transient transfection methods have been streamlined in recent years through the use of multi-well plates and commercially available, highly efficient transfection reagents that have simplified the process and improved reproducibility. Transient transfections require the continuous expansion of DNA plasmids although less amounts of these plasmids are needed as protocols are scaled to multi-well plates.

***Recommendation:***
Transiently transfected cell lines would be advisable due to their greater sensitivity if patent issues can be resolved. If patent issues persist, cells with endogenous AR with a stably expressing reporter vector are optimum. It may be necessary to demonstrate 10-20 fold induction by a control agonist in order to detect weak agonists. Other considerations are stability of the stably transfected cell line, and absence of the complicating factor of endogenous GR. The Panel suggests an approximate minimally acceptable level of 10-20 fold to allow for sufficient sensitivity to detect weak agonists. The stable cell line should contain a reporter with a response element that is relatively specific to the AR. If stable cell lines can be developed with sufficient sensitivity, these would be advisable but probably must be maintained in drugs to select for cells containing the integrated plasmids. A central laboratory for the generation and disposition of stable lines should be pursued, since during the course of generating these lines, integration of transfected receptor constructs may occur at different locations within the genome leading to potentially unique response profiles across laboratories. A central source of adenovirus vector would be needed for the adenovirus protocol because propagation of the virus can be technically challenging. Thus far, a particular stable cell line could not be recommended because of weak induction of luciferase activity by the control agonist.

3. **Reporter plasmid specifically for androgen-bound AR**
The optimal androgen responsive reporter vector would not allow for cross specificity with other receptor subtypes such as GR and the progesterone receptor (PR), as exhibited by MMTV. An androgen specific consensus human response element (HRE) is preferred and the use of multiple HREs in the construct could increase sensitivity. The MMTV reporter vector was used in the majority of the transient (protocol B1, B2, and others) or stable (MDA-MB) assays presented in the BRD. Protocols B5 and

B6 use four copies of the rat ventral prostate prostatic binding protein gene subunit C3 linked to the luciferase gene. None has been shown to be specifically activated by the androgen-bound AR. In almost all the protocols listed, 17β-estradiol activated the reporter indicating that AR can activate the MMTV and other reporters in the presence of a variety of ligands, depending on its concentration. This is a potential problem in screens that assay relatively high doses of substances and could result in false positives. Other reporters not included in the report could have greater specificity for activation by the androgen-bound AR. Establishing the optimum reporter that demonstrates specificity for AR and for the androgen-bound AR requires further investigation. Some reporters that have been investigated include the sex-limited protein gene and probasin. However, even these can show activation by other receptors.

***Recommendation:***
Other reporter vectors different from MMTV that show specificity for activation by the androgen-bound AR should be investigated. Possibilities include the rat prostate C3 gene promoter and enhancer promoter regions for the prostate specific antigen gene, sex limited protein gene, and probasin, although these reporters may be less sensitive than MMTV. Ideally, the reporter should not be activated by the AR in the presence of 17β-estradiol or cortisol. The alternative is to indicate a cutoff concentration, otherwise all estrogenic substances may be identified as AR agonists. This may be difficult since the cutoff concentration for steroid hormones will differ from the cutoff values for unknown environmental substances.

**4. Sensitivity to detect low concentration substances in screening**
The transient transfection assays were more sensitive than stably transfected lines. Stable cell lines often have less than 10 fold induction with the control agonist whereas transient transfection can have 50-100 fold induction. Low induction by the control agonist may make reliable detection of antagonists more difficult and the detection of weak agonists impossible. During the deliberations of the Panel, it was suggested that stable cell lines with generally lower fold induction could reliably detect AR antagonists. The CHO stable cell line (protocol B6) had only 5 fold induction with the control agonist which would be too weak a response for testing unknown substances. The HeLa cell assay of Wang and Fondell (2001) had 4.5-7 fold induction with 100 nM DHT, which is unacceptably too low and would not be useful in detecting weak or strong agonists or antagonists. The CHO assay with CAT activity of 100 fold with 0.1 nM DHT was 35 fold with cortisol, reflecting the complication of endogenously expressed GR. Fold induction for the MDA stable cell line ranged from 3-10 fold.

***Recommendation:***
Based on the experience of the Panel, an induction of 10-20 fold should be achieved by the control agonist in a concentration range of 0.1-1 nM R1881 or DHT to allow for the detection of weak agonists. An effective means of separating out confounding effects of other steroid hormone receptors is essential.

5. **Suitability for large scale screening**
All of the protocols could be scaled to multiwell plate assays. This is a requirement for the optimal assay as it allows for large-scale screening and the use of transfection reagents that would otherwise be too costly.

***Recommendation:***
The assay should be established in a multiwell format.

6. **Ability to detect a weak active agonist or antagonist in a complex mixture**
This could be difficult using a stable cell line that typically has less than 10 fold induction with the control agonist. Detection of weak acting substances may require an assay with at least a 10-20 fold induction with the control androgen at a concentration of 0.1 nM DHT or R1881.

***Recommendation:***
The most sensitive assay would be achieved using transient cotransfection assays because the sensitivity of these assays exceeds that of the stable lines. However, considering the potential patent restrictions on the AR and the cis/trans cotransfection methodology, the optimum assay would be a cell line with endogenous AR that is sufficiently sensitive to detect weak acting agonists and antagonists. Further research should be directed toward making more sensitive stable cell lines.

7. **Ability to discriminate agonist and antagonist**
Each of the assays, except for the yeast assay, discriminate AR agonists and antagonists. All cell assays documented in the BRD demonstrated the antagonist activity of casodex (bicalutamide) and hydroxyflutamide in the presence of an AR agonist such as DHT. Each also indicated agonist activities of DHT, testosterone and other known agonists. However, AR agonist activities were also reported for cortisol, dexamethasone, and 17β-estradiol, none of which are AR agonists *in vivo*. These latter responses reflect difficulties due to lack of specificity of the reporter and AR ligand binding.

***Recommendation:***
All assays showed 17β-estradiol was an AR agonist, which is not observed *in vivo*. This results primarily from lack of specificity of the MMTV reporter to ligand activation of the AR and a lack of absolute AR specificity for binding steroids. Many of the assays showed AR agonist activity of cortisol which is neither an agonist nor an antagonist and results in part from the presence of GR. All of the assays showed DHT as an agonist and hydroxyflutamide and casodex to be antagonists. The ideal AR screening assay will not show AR activation by 17β-estradiol or cortisol.

8. **Sufficient fold induction by androgen to detect antagonist activity**
A reporter assay should show at least 10-20 fold induction with 0.1-1 nM of the control agonist. It was not possible to evaluate fold induction for many of the assays in the BRD. Transient transfection experiments would be superior to stable cell lines in terms of sensitivity (i.e., extent of fold induction).

***Recommendation:***
Thus far, most stable cell lines show less than 10 fold induction. Transient transfected cells or adenovirus infected cells are more sensitive and would be advantageous.

**9. Transferability, patent/proprietary issues**

The AR protein and coding sequence have been patented. Ligand Pharmaceuticals (San Diego, CA) holds an exclusive license based on the U.S. patent and Karo Bio AB (Huddinge, Sweden) has licensing rights on the European patent. The AR patent covers any AR vectors that have the recombinant derived human AR sequence. The patents therefore apply to cells stably transfected with a human AR plasmid. Presently, Ligand Pharmaceuticals will not allow other companies to use this technology. They also have patent rights over the cis/trans duel transient transfection assays. It needs to be clarified whether these patents nullify all of the transient transfection protocols proposed. Under the present conditions, while the protocols could be repeated in independent academic labs, they could not be repeated by commercial companies. If the gold standard protocol does not require transferability to commercial laboratories, transient transfection methods could be further considered. It is stated on page 6-9 in the BRD that patent issues preclude the use of transient transfection assays. It should be noted that this also applies to cell lines with an integrated AR plasmid. The patent restrictions on many of the assays are a significant disadvantage. It needs to be clarified whether a gold standard assay should be available to commercial companies involved in developing AR screening assays.

***Recommendation:***
A stable cell line with endogenous AR and stably expressed reporter would avoid patent issues unless there is a patent associated with the reporter. The MMTV lacks specificity for activation by AR and also for low androgen specificity for AR activation. However, a more suitable reporter may lead to additional patent problems. The U.S. government supported the original research that determined the AR coding sequence. The U.S. EPA could investigate their ability to use its rights under that funding to get a license from Ligand Pharmaceuticals to make the transient transfection methodology using the AR expression plasmid a viable alternative.

**10. The use of yeast in endocrine disruptor screening**

The yeast assay B4 is complicated by the possibility of unusual metabolism of ligands, by problems associated with cellular uptake and transport of steroids and test substances, and by an inability to distinguish agonists from antagonists. The BRD is correct that yeast would not be the optimal approach for these assays.

***Recommendation:***
Yeast assays should be avoided.

**11. Reproducibility and expense of transfection methods**

There was insufficient information to compare the assays described in the BRD in terms of reproducibility within and between labs. Most methods use expensive but highly

effective commercially available transfection reagents. These can be cost effective using the multiwell plate format. All of the methods proposed appear to be applicable to multiwell technology. The calcium phosphate transfection method is the least expensive but requires precise handling and close attention to details and may not be easily transferable between laboratories, but should be amenable to the multiwell format.

***Recommendation:***
The use of multiwell plates is recommended to keep down transfection costs and to provide for large scale screening. Reproducibility would need to be evaluated.

**12. Endogenous mutant AR with loss of androgen specificity**
A cell proliferation assay using the human prostate cancer cell line LNCaP was mentioned, although the protocol was not provided. This cell line is unacceptable due to a mutation in the endogenously expressed AR that alters its ligand specificity.

***Recommendation:***
A cell line must express an AR with wild-type sequence.

**13. Are results directly applicable to humans and wildlife species?**
The results with all of the mammalian-based cell assays are similarly applicable to *in vivo* conditions, although some cells metabolize ligands more efficiently than others. It would be optimal to have minimal metabolism of control steroids. On the other hand, it would be advantageous to have metabolism of substances that reflects what occurs *in vivo*. In almost all of the assays, the human AR was used, making the data relevant to human. Data obtained using the human AR should also be relevant to wildlife species as the ligand binding domain of the AR across species is highly conserved. Mammalian cells should be used for the assay of human AR activity as human AR has not been shown to be active in fish cells, for example. This most likely reflects the low sequence homology of the AR amino terminal region that contains the major activation domain of the AR. Thus far, every active substance examined in both wildlife and mammalian assays has been detected in mammalian assays. However, this is not true for wildlife assays. For example, flutamide is not detected as an AR antagonist in some fish assays despite finding hydroxyflutamide as the primary metabolite. No substances have been reported that are only positive in wildlife assays. The data support what is known about the function of these hormones in both sex differentiation and development and AR action in adults in mammals and other species. All currently available AR transient and stable transfection assays suffer from the same weakness, that they may not accurately predict response in humans or whole animals because *in vitro* assays cannot adequately measure absorption, metabolism, distribution, and elimination, as well as target-tissue specific factors that influence AR function.

***Recommendation:***
Depending on the extent of cell metabolism of the test substances, results from transient or stable transfection assays should be applicable to humans and wildlife. Assays should

use the human AR that requires transcriptional analysis be performed in mammalian cells.

**14. Are controls provided to test for cytotoxicity when assessing antagonists?**

About half of the protocols provide β-galactosidase control vectors as a control for cytotoxicity. A more appropriate control would be to include transfection of a constitutively active luciferase vector such as CMV-Luc or pSG5-Luc. This is advantageous over β-galactosidase in controlling for cytotoxicity because the use of luciferase vectors also controls for direct effects of the test substances on luciferase enzymatic activity. If only β-galactosidase is used, the possibility exists that the test substance directly inhibits β-galactosidase activity and has nothing to do with gene expression. In contrast, a direct inhibitor of luciferase may be mistaken for an AR antagonist if β-galactosidase is used as the enzyme assay. By using luciferase as the cytotoxicity control, direct effects on enzyme activity and on cytotoxicity are included in the control assay. Tests for cytotoxicity are especially important at high concentrations of test ligands so it may be appropriate that cytotoxicity and luciferase activity tests are limited to samples at concentrations over 1 μM. Controls should also include those for the vehicle used for hormone and chemical additions.

***Recommendation:***

Control plasmids such as CMV-Luc or pSG5-Luc should be included to assess cytotoxicity and inhibition of luciferase activity. Alternatively, cytotoxicity tests might be limited to substances that show antagonist activity since it might apply to substances at high concentrations.

**15. Other endogenous steroid receptors that complicate the assays**

A complication of using the stable cell line MDA-kb2 is that it contains sufficiently high levels of GR to interfere with the assay. This could also be a complication of HepG2 and HeLa cells depending on which reporter vector is used. Because the MMTV Luc or CAT reporters respond to GR better than they do to AR, it becomes more difficult to conclude that a response is significant.

***Recommendation:***

Cell lines should be used that lack relatively high levels of other steroid receptors, in particular the GR and PR. These receptors share with the AR a similar DNA binding domain sequence allowing them to activate in many cases the same enhancer-promoter sequence.

**16. Are the results in general agreement with other assays and known activities?**

Almost all the assays show that 17β-estradiol activates AR. However, there is no evidence to support that this substances activates AR *in vivo*. The optimum assay would be expected to not show AR activation by 17β-estradiol. Better protocols are needed that allow for AR activation only by known androgen agonists.

***Recommendation:***
Most of the assays show cortisol and 17β-estradiol as AR agonists. Cortisol activity results in part from the presence of GR in some cell lines, and in part because of the lack of specificity of response elements in the reporter vector. AR agonist activity of 17β-estradiol does not agree with its known *in vivo* activity. The results raise the question at what concentration should a substance be considered a real AR agonist.

**17. What is the minimum acceptable fold induction for a control agonist?**
For most of the assays presented in the BRD, it was not possible to determine the fold induction achieved by the control androgen. At least a 10-20 fold induction would be considered acceptable. Otherwise, it may be difficult to assess weak agonists or antagonist activity.

***Recommendation:***
Induction should be at least 10-20 for a control AR agonist at 0.1-1 nM R1881 or DHT.

**18. Can the assay be accommodated in a multi-well format necessary for large scale screening?**
All of the assays could be accommodated by this format.

***Recommendation:***
A multi-plate assay is required.

**19. Error range, variance**
RELIABILITY = REPRODUCIBILITY WITHIN AND BETWEEN LABS
It was not possible to determine this from the data provided in the BRD.

***Recommendation:***
The data should be within 20% error.

**20. Statistical considerations**
**a. Recommendations for agonist and antagonist classification**
Substances can display both AR agonist and antagonist activities depending on the concentration, the assay system, or the presence of endogenous androgen. In some instances, the duel activities are real and reflect endogenous activities. One example is hydroxyflutamide, a classical AR antagonist, which at high concentrations has agonist activity. Whether this occurs *in vivo* is not clear but could be reflected in the hydroxyflutamide withdrawal syndrome, where prostate cancer patients improve after removing treatment with the antagonist. Another example is the drug medroxyprogesterone acetate (MPA), which when administered to pregnant women can induce clitoromegaly in a female fetus (an agonist response) and hypospadias in a male fetus (an antagonist response). A substance like MPA with weak AR agonist activity can be androgenic in the developing female fetus where low androgen levels are present. However, a weak agonist may be antiandrogenic in the developing male

fetus because, by being a weaker agonist than DHT, it binds AR and elicits a weaker agonist response, appearing as an antagonist.

For substances that display agonist activity, the effective concentration (EC) should be provided that increases transcriptional activity by 2-3 fold. For substances that display antagonist activity, the $IC_{25}$-$IC_{50}$ could be used (i.e., the concentration required to inhibit ligand-induced transcriptional activity by 25-50%). In some cases, antagonist activity may be detected but not achieve 25-50% inhibition. It is not clear whether decreases in transcriptional activity of 10% or less should be considered physiologically relevant. For antagonist assays, the concentration of agonist (i.e., R1881 or DHT) used to assay antagonist activity should be concentrations of R1881 or DHT that induce approximately 75% transcriptional response in the assay system but not more than 1 nM. In some cases, dose response relationships may not be observed due to sample impurities or metabolism of the substance. In such situations, further analysis is suggested using more purified preparations or cell-free *in vitro* competitive binding assays. The results from Tier 1 screening assays should be considered together in the decision of whether Tier 2 testing is required or not.

**b. Recommendations for statistical analysis**

A critical aspect of the analysis of *in vitro* and *in vivo* Tier 1 screening assays is to assess the results from all screening assays *in toto* and subjectively establish whether evaluation by Tier 2 testing is necessary. There appear to be two major considerations in the evaluation of *in vitro* AR transcriptional activation data. First is agonist or antagonist potency. The EC at which a 2-3 fold increase in transcriptional activity is observed or a 25-50% decline in gene expression ($IC_{25}$-$IC_{50}$ concentration) could be used to rank order potency. The second consideration of equal importance is whether the response varies with dose and if so, how steep is the dose-response curve, as discussed further below. The reported data for these assays should be the concentration where gene expression is increased by 2-3 fold or decreased by 25-50%, and the slope of the tangent line to the dose-response curve. As such, the reported data from these assays will minimize the tendency to label substances as significant *in vitro* endocrine disruptors and maximize their input toward assessing the results from all Tier 1 screens.

The goal in the *in vitro* screening studies is to determine the concentration of a substance that induces an alteration in gene expression that is biologically meaningful. Here, it is assumed that a 2-3 fold increase or 25-50% decrease in response is indicative of a potential *in vivo* response. This approach, combined with determining how steep the dose-response curve is from this point downward (or upward for agonists), should provide sufficient information from these assays together with results from other *in vitro* and *in vivo* assays to assess whether Tier 2 testing is warranted. The ultimate decision as to whether to proceed to Tier 2 testing should not be a quantitative assessment where results from each of the assays in the Tier 1 screen are given a score and a weight. Data from individual assays in the Tier 1 screen should be assessed *in toto*.

**i. Response variable for evaluating the potential agonism/antagonism of a test substance**

On page 5-3 of the BRD, several definitions are indicated for relative activity of a test substance. Classification of a test substance as an agonist or antagonist depends on how relative activity is defined. Estimates of quantities such as $EC_{50}$ and $IC_{50}$ values for a given test substance depend on the definition used for relative activity. Meaningful comparisons of $EC_{50}$ /$IC_{50}$ across studies and chemicals require that all studies use the same definition for relative activity. For instance in **Table 12-1** on page 12-11 of the BRD, median $EC_{50}$ values are provided. The median $EC_{50}$ for a given chemical is meaningful only if all the participating labs used the same definition for relative activity. Thus, definition of the response variable of interest should be standardized for future assays. This is vital because all important decisions, including the final determination of positive/negative agonism, and positive/negative antagonism, regarding a chemical are based on the chosen definition.

**ii. Assumptions made in a statistical analysis**

Most statistical procedures make certain assumptions regarding the underlying data. For instance, if ANOVA type methods are used then it must be determined that there is no heteroscedasticity and that the data are approximately normal. The Panel recommends that before any statistical procedure is used, suitable diagnostics be performed to verify that all underlying assumptions regarding the procedure are true. If the assumptions are not true, then suitable data transformations might be performed before analyzing the data.

**iii. Estimation of $EC_{50}$/$IC_{50}$ and steepness parameter**

**Section 12.2.11** (pages 12-5 and 12-7) of the BRD state that $EC_{50}$ and $IC_{50}$ values should be reported along with their confidence intervals. The standard approach is to fit the data with a suitable non-linear model, such as a Hill equation, which gives an objective estimate of the $EC_{50}$ /$IC_{50}$ values, as well as a confidence interval. In most cases, an objective estimate of the $EC_{50}$ /$IC_{50}$ values based on a Hill equation or other suitable nonlinear model is required. If it is appropriate to perform nonlinear modeling based on a suitable dose response curve, a suitable nonlinear model should be used. Suitable model diagnostics should be performed to ensure that the model fits the data and the various underlying assumptions such as normality and homoscedasticity are true. Also, diagnostics should be performed to detect curvature effects and a suitable estimation procedure should be used for estimating confidence intervals for the parameters of interest. In some cases, especially in the presence of curvature effects, the standard asymptotic confidence intervals might not be appropriate.

In such cases, a resampling procedure such as jackknife or bootstrap might be used.[3]

**iv. Combining $EC_{50}$ and $IC_{50}$ estimates from different labs**

To obtain estimates of mean $EC_{50}$ and mean $IC_{50}$ values from different laboratories, the average across the laboratories should not be simply computed. Instead, estimates using mixed effects nonlinear models, treating laboratories as the random effects, should be used. This approach takes into account within and between laboratory variability. This methodology also allows for a determination of the corresponding standard errors.[4]

**v. Positive agonists and antagonists**

In situations where it is not possible to obtain a nonlinear model, the Panel recommends performing a statistical trend test. If the trend is significant, confidence intervals should be computed at each dose. If a confidence interval contains 10% of the maximum response, that would suggest that this substance should be further evaluated.

**vi. Uniformity trials**

To understand the underlying variability in the data, which is important for proper data analysis, the Panel recommends conducting a set of carefully planned comprehensive, interlaboratory negative control studies.

**Conclusions to Question 1**

Theoretically, stable transfection assays are preferable to transient transfection assays for high throughput screening. A stably transfected reporter system allows for non-radioisotopic detection using a reporter gene. Stably transfected lines would need to be established that allow for sufficient sensitivity and reproducibility across laboratories. Stable assays will require an agonist response to 0.1-1 nM control androgen such as R1881 or DHT that is sufficient to detect weak androgens. Induction of at least 10-20 fold would facilitate this detection; however, lower fold induction may be adequate. The Panel recommends further analysis of the adenovirus infection method using MDA-MB-453 cells, noting that while these cells contain endogenous AR, they also contain significant levels of GR, complicating analysis of potential AR agonist or antagonist activity.

This recommendation does not preclude the use of transient transfection methodology; however, potential patent restrictions require further clarification. Transient transfection assays can be developed for high-throughput screening and are advantageous because of higher fold induction

[3] cf. Simonoff and Tsai, "Jackknife-based estimators and confidence regions in nonlinear regression," Technometrics, 28, 103-112, 1986, and Zhang, J., Peddada, S. D., Rogol, A. D, "Estimation of parameters in nonlinear regression models," Statistics for the 21st century: Methodologies for applications of the future, edited by C. R. Rao and G. Szekely, Marcel-Dekker, New York, NY, 2000.

[4] Davidian M, and Giltinan D (1995) Nonlinear models for repeated measurement data, Chapman and Hall, London, UK.

and can include appropriate controls for transfection efficiency and toxicity, and can be designed to be more receptor specific. However, transient transfection assays may require greater cost and effort, are technically more difficult for laboratories that do not routinely do this type of assay, and are limited by patent restrictions.

Yeast-based assays are not acceptable because of reduced ability to detect certain substances either because of an inability of the substances to cross the yeast cell wall or because of active transport mechanisms. The yeast-based assay also does not accurately discriminate between agonists and antagonists. Stable assays with greater sensitivity should be pursued in cell lines that are not complicated by other endogenous steroid receptors.

## 2.0 Minimum Procedural Standards for *In Vitro* AR TA Assays

### 2.1 Considering the intended use of the assays as toxicological screens, does the Panel agree with the adequacy of the proposed procedural standards recommended for *in vitro* AR TA assays?

#### 2.1.1 Transcriptional Activation of the Reference Androgen

The Panel agrees that the transcriptional activation-inducing ability of R1881 must be demonstrated and that consistency in the level of response is appropriate as a criterion for assay acceptance. A full dose-response curve should be generated.

#### 2.1.2 Reference Agonist and Antagonist

The gold standard reference agonist for validation should be R1881 due to possible metabolism of natural androgens in different cell lines. DHT should also be included in all assays. The concentration of the agonist selected should be within the linear region of the dose-response curve of 50 - 70% induction. The concentration of the agonist selected, as well as the ability to identify significant effects, will depend on the assay, but should be within 0.1-10 nM DHT or R1881.

The reference antagonist should be hydroxyflutamide (not flutamide). Casodex (bicalutamide) should be included in the list, although casodex can be difficult to obtain. There should be 70-90% inhibition in the presence of 0.1 nM or 1 nM R1881 or DHT. Depending on the reporter, the reference androgen concentration should be 0.1 or 1 nM R1881 or DHT for maximal induction. The inhibitory concentration ($IC_{25}$- $IC_{50}$) of hydroxyflutamide is approximately 500 nM. The $IC_{25}$- $IC_{50}$ should be defined in terms of the androgen concentration against which it is inhibiting. Antagonist activity should be expressed in terms of the $IC_{25}$- $IC_{50}$, or the response rate at the $p < 0.05$ level of significance.

#### 2.1.3 Preparation of Test Substances

The test substances should be prepared in water, 95-100% ethanol, or dimethyl sulfoxide (DMSO), depending upon their solubility. Preference should be given to the solvent that allows testing of the maximal concentration of the test substance, without exceeding the limit dose. However, in situations where more than one

solvent could be used, preference should be given to water, then 95-100% ethanol, and then DMSO, in that order. It would be prudent to perform a pre-validation of the transcriptional activation assay with the reference androgen for assessment of the level of solvent that does not adversely affect assay response. Substances should be dissolved at 1-10 mg/mL in water, 95-100% ethanol, or DMSO and solubility verified. Appropriate solvent controls should be included in all screening assays.

### 2.1.4 Concentration Range of Test Substances

For both agonism and antagonism assays, the limit concentration should be 1 mM but the solubility characteristics and potential cytotoxicity of each test substance must be taken into consideration. If the limit concentration is used, seven test substance concentrations at log intervals should be tested. If a lower maximum concentration is tested due to solubility constraints or excessive cytotoxicity, the number of concentrations tested can be adjusted to account for the altered concentration range. A measure of cellular cytotoxicity will help define the upper limit for test material concentration similar to a Maximum Tolerated Dose (MTD) approach used in *in vivo* studies. This may mitigate the need for the higher concentrations which appear excessive in these *in vitro* systems. However, it is important to consider that agonist activities detected only at doses >1 μM for endogenous steroids or test substances should not be considered significant unless it is a pharmaceutical. At high concentrations, nonspecific interactions occur that could lack physiological relevance, depending on the exposure concentration. It is important to keep in mind the sensitive dose response relationships of endocrine activity.

### 2.1.5 Solvent and Positive Controls

In each experiment, there should be at least two positive controls, DHT and R1881. Controls for cytotoxicity should be performed for all samples that show apparent antagonist activity. This will involve the higher concentrations of ligands. Controls for cytotoxicity can be accomplished by including an internal constitutively active control reporter plasmid such as CMV-Luc. Levels of cytotoxicity exceeding 10% are unacceptable. Results are not useful at concentrations where substances are cytotoxic. If cell lines are used that contain other endogenous steroid receptors, inhibitors must be added that selectively inhibit ligand binding to that receptor. AR agonist activity could be selectively inhibited using hydroxyflutamide; however, this approach requires substantially more work. Hydroxyflutamide or bicalutamide (casodex) controls in the absence of added agonist could be included in antagonist experiments to control for possible endogenous androgen activity in the media used for the assay.

### 2.1.6 Within-Test Replicates

Each test chemical concentration should be tested in triplicate.

### 2.1.7 Dose Spacing

The Panel agrees with the recommended dose spacing of one order of magnitude in the concentrations of the test substances.

### 2.1.8 Data Analysis

A uniform method for expressing response should be established. The data should be expressed in absolute units such as light units for a luciferase assay. The data can be multiplied by a scaler for convenience (i.e., x 0.001). The data can in addition be expressed as fold induction but fold induction alone is not sufficient. The use of modern plate readers with high sensitivity can lead to fluctuations in background levels that can have profound effects on the apparent fold induction. For agonist activity, the $EC_{50}$ value, the lowest effective concentration where there is a 50% increase in response to the stimulatory ligand, may not be appropriate. $EC_{50}$ values can be misleading if the substance is not a full agonist or if the substance cannot be tested at high enough concentrations due to solubility limits or toxicity. $EC_{50}$ values depend on the conditions of the assay and can vary between laboratories even under standardized conditions. Instead, the lowest observed effective concentration (LOEC) at which a significant (2-3 fold) response is observed over background could be reported. Percent of control would not be acceptable unless the absolute relative light units (RLU) are given for the control (i.e., such that the RLU for all responses can be easily calculated). However, percent of control may be acceptable for comparing multiple experiments in which maximal induction levels vary. For transient transfections that include a proper constitutively expressed control for transfection efficiency and toxicity, the data could be expressed as corrected units. Control values should be monitored to ensure that assay responsiveness remains within historically accepted limits. Cells may loose their effectiveness over time and may need to be reestablished, and DNA used in transient transfections may degrade over time. An internal standard reporter vector is not required for stable cell lines but should be required for transient transfections. The benefit of an internal standard is that it can also be used to monitor for chemical toxicity.

### 2.1.9 Assay Acceptance Criteria

Test chemical entry and exit assays for all dose formulations must be within 10% of the intended concentration. An entry assay assesses the identification and concentration of the test article in the dose formulation prior to the start of the experiment, whereas the exit assay assesses these same parameters at the end of the experiment. The need for these analyses may depend on existing analytical methods for their assessment; complex mixtures may not be appropriate for these analytical analyses. The 10% level is based on standard analytical chemistry assessments of dosing solutions to insure the concentration is actually the concentration that was used. Compliance with standard Good Laboratory Practice (GLP) guidelines is advisable. An unacceptable experiment would have replicate variability exceeding 25-30%, cytotoxicity measurements exceeding 10% of the response level or positive and/or negative control levels that do not meet predetermined criteria, including fold induction of 20x, internal replicate variability of 20% or less, inappropriately negative or positive response. A valid experiment would have appropriate responses from positive (DHT, R1881) and negative (solvent) controls. The response should be within the acceptable limits as defined by historical data. If the response is outside the historically established range, it is not acceptable. This will also help the laboratory

monitor assay drift. Toxicity should be monitored. Concentrations of chemicals that cause 10% toxicity should not be considered. The transient transfection efficiency would be useful to know; however, it would not need to be established in every assay.

***Specificity of response:***
The reporter gene assay should show activation by the classical androgens DHT and testosterone, and by the synthetic androgens R1881 and mibolerone. Ideally, the assay should not show agonist activity with 17β-estrradiol or cortisol up to concentrations of 10 nM. The assay should show the classical antagonist response of hydroxyflutamide and casodex and should not allow response to other receptors. This requires that the GR and PR are expressed at low levels in the test cell line and that an AR selective reporter vector is used. The MMTV reporter used in the majority of the assays presented in the BRD is nonselective and can be activated by other hormones if the appropriate receptors are present.

***Sensitivity to detect weakly active substances:***
Thus far, stable cell lines may lack the sensitivity to detect weakly active substances. This deficiency likely requires a transient transfection assay that demonstrates at least 10 fold induction with the control androgens and levels of variance that allow detection of alterations in gene expression of at least 2 fold stimulation or 25-50% inhibition of activity. It is difficult to determine what fold change in reporter gene activity is indicative of a change in gene expression *in vivo*. However, a 2-3 fold increase or 25-50% inhibition would imply a significant change in AR functional activity.

***What is the minimum fold induction acceptable for active androgens?***
The minimum acceptable induction is 2-3 fold over the no hormone control. For many of the assays presented in the BRD, this was difficult to evaluate and some were clearly not acceptable. A positive response would be at least 2 fold over background levels. Alternatively, rather than set cutoff limits based on fold induction or antagonism, a more statistical approach might be considered.

***Acceptable variance:***
The percent coefficient of variation (%CV) should not be greater than 20%. This is calculated as standard deviation/mean x 100.

### 2.1.10 Evaluation and Interpretation of Results

An acceptable limit for acceptance as a positive or negative response would be a change, relative to the control, of 2-fold induction for an agonist and 25-50% inhibition for an antagonist.

### 2.1.11 Test Report

The test report should include the recommendations in the BRD plus the following changes and additions:

a. Information should be included on controls for the activity of other steroid receptors and controls for cytotoxicity. The source of plasticware and other materials used in the assay should be listed. The cell passage number should be recorded.
b. Chemical names of known test substances and structures are appropriate.
c. The solvent does not require justification unless it is other than water, 95-100% ethanol or DMSO. The solvent used should be indicated. For the AR source, the supplier should be indicated if it is a noncommercial source.
d. Procedures for making constructs should indicate only the type of method used for isolating the DNA, not the detailed procedure.
e. The structure of the response elements in the reporter vector should not be needed, simply the name and reference.
f. The methodology for making the reporter plasmid should not be required.
g. The reference androgen should only need a rationale if it is other than DHT or R1881. The assays should require the use of the standard recommended androgens.
h. The concentration and volume should be indicated for the test substance.
i. At least two replicates of the experiments are needed and the assays performed in triplicate. More experiments are required if the experiments are not in agreement.
j. The response should be indicated in absolute units such as light units for luciferase activity with the error indicated, and in addition, as fold induction if this is deemed appropriate.
k. Statistical analysis of the data should indicate agonist and inhibitory test chemical effects on transcription that meet or exceed the 2-3 fold induction or 50% inhibition level compared to the respective controls.

#### 2.1.12 Replicate Studies

Replicate studies are not mandated but questionable data needs to be confirmed by re-testing the substance. In such situations, the incorporation of stricter treatment concentrations in follow-up assays based on the initial experiment should facilitate better analysis of the overall dose-response of the test material.

### 2.5 Considering the intended use of the assays as a toxicological screen, are there other minimal procedural standards that should be included? If so, what are they and why?

1. An internal standard reporter vector such as CMV-Luc should be used to control for cytotoxicity and direct effects of the test substance on luciferase enzyme activity. However, other methods could also be acceptable. Cytotoxicity controls should only be needed for high concentrations of substances that show apparent antagonistic effects.
2. GLP guidelines should be required. The use of GLP will improve overall results and minimize potential sources of error. These include making sure the balance and pipettes measure accurately, reagents are not past expiration, and minimizing mistakes in data transfer or transcribing from one location to another. Entry and exit assays for test article and control substances should be included as out-of-normal results cannot be accurately interpreted without them. However, due to the associated costs and lack of appropriate

analytical standards for many substances that will be tested, this approach should not be a requirement for screening assays.

3. Methods for establishing and propagation of a stable cell line are available in the scientific literature.
4. The use of charcoal-stripped serum should not be required if the cells can be maintained during the assay in serum free media. Propagation of the cells should not require charcoal-stripped serum. In cell assays that require the presence of serum during hormone incubations, serum that has been stripped of endogenous hormones would allow for a more sensitive assay.
5. Some information should be provided concerning the stability of stably expressing cell lines. Stable cell lines should be sufficiently stable to retain the integrated plasmids and response to control agonists and antagonists. Details about the drug requirements for maintaining the stable cell lines should be indicated. The MDA cell line described in the BRD appeared to be stable for at least 80 passages. Stability of any cell line should be closely monitored and ultimately a cutoff passage number should be determined. The cost of drugs necessary to maintain a stable cell line may be less than the costs of reagents necessary for a transient transfection screening assay.
6. Steroid or chemical metabolism should be established for positive responses. This could be done by including additional entry and exit assays for control ligands and test chemicals using satellite cultures of the cells plus media before and after culture. This approach would not add much to the cost as most of the cost is spent in the initial set-up and validation of each assay.

**3.0 Recommendations for *In Vitro* AR TA Test Method Protocols for Validation Studies**

Protocols provided by scientists with expertise in *in vitro* AR TA test methods are provided in Appendix B of the BRD. Section 12.3 discusses additional details that should be added, based on the minimum procedural standards in Section 12.2.

**3.1 Considering the intended use of the assays as a toxicological screen, would the current protocols, with the additions detailed in Section 12.2 and 12.3, provide a level of detail to appropriately minimize interlaboratory variability? If not, what revisions or additions should be made to the protocols?**

In general, the details on the effectiveness of the different assays were scarce in the BRD. Important details needed to compare different assays include fold induction by the control androgen, intra- and inter-assay coefficients of variability, stability of cell responsiveness over time and passage number, and a standardized method for comparing potencies of agonists and antagonists in the different assays.

**3.2 In addition to the minimum procedural details listed in Section 12.2, are there other protocol elements that should be considered for *In Vitro* AR TA Assays recommended for validation as a toxicological screen, including those protocols provided in Appendix B?**

The Panel recommends that the adenovirus infection method be further explored with the goal to eliminate activation by endogenous GR and use a more selective reporter. The other mammalian-based assays may also be appropriate for use in validation studies. The yeast-based assay has inherent limitations and complications related to the presence of a yeast

cell wall and active transport mechanisms that differ from those found in mammalian cells. The yeast-based assays do not discriminate between agonists and antagonists and should not be considered. Many of the mammalian-based assays are limited by several major considerations. There are patent issues associated with the transient co-transfection assays and cell lines that have the AR plasmid stably integrated as a result of transfection. If the patent issues cannot be resolved, almost all of the assays proposed will not be useful. On the other hand, it may be that a transient cotransfection assay could be used as a gold standard by which other assays would be judged in terms of response. One of the stable cell lines (protocol B6) is also subject to patent issues because the AR plasmid was transfected. The usefulness of the MDA stable line assay by Wilson et al. (2002) was complicated by the presence of endogenous GR, use of the MMTV-Luc reporter, and the low fold induction to the control androgen agonist (less than 10 fold) and the positive responses to 17β-estradiol and cortisol.

**3.3 Considering the intended use of the assays as a toxicological screen, is the Panel aware of other available standardized protocols for assays recommended for validation?**

***Assay 1:***
The N/C two-hybrid interaction assay in mammalian cells makes use of GAL4 and VP16 fusion proteins with the AR ligand binding domain and the AR N-terminal region. Assays are also being developed using the GAL4-AR ligand binding domain expressed with full-length AR. These assays have been modified for use in HeLa cells in a multiwell format (He et al. 2000). The advantage of the assay is that it distinguishes agonists and antagonists and can be performed in a multiwell format. The assay does not have false positives resulting from 17β-estradiol or cortisol because it depends on the androgen-specific interaction between the $NH_2$- and carboxyl-terminal regions of the AR. Limitations of the assay are that it is subject to the same patent restrictions that apply to other transient cotransfection assays that use the AR expression vector and that apply to stable cell lines with an integrated AR plasmid. The N/C assay has greater than 20 fold induction with 0.1 nM DHT, a sensitivity significantly greater than that achieved by stable assays presented in the BRD. The reporter vector is a GAL-Luc reporter with which no other steroid receptors are active.

***Assay 2:***
Other naturally occurring androgen response elements may have greater specificity to activation by the androgen-bound AR as opposed to activation by AR binding of 17β-estradiol. Without this specificity, assays using MMTV-Luc, while highly sensitive, have the disadvantage of false positives. Other response elements may have lower response than that achieved by the MMTV reporter (50-100 fold for MMTV-Luc). An assay such as this would also be subject to AR patent restrictions and may be further complicated by restrictions on specific reporter vectors.

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

**4.1 Does the Panel agree with the selection criteria, adequacy and appropriateness of substances recommended for validation studies for agonists and antagonists? If not, what substances should be added or deleted?**

The listed substances are primarily steroids and pesticides of known AR agonist or antagonist activity. The objective of this transcriptional screening assay is the correct identification of substances that act as AR agonists or antagonists. As such, the chemicals with known AR agonist or antagonist activity could be abbreviated to those listed below to determine the ability of the test system to correctly identify their activity and their correct rank order. In addition, it is equally important to determine the ability of the test system to correctly identify known and predicted confounders, such as chemicals whose activity would be expected to alter luciferase production or activity independent of AR binding. In this way, confidence in the test system to correctly identify androgen active substances and correctly identify indirect or cytotoxic activity is maximized. The assay of liquid and gaseous volatiles is apparently not subject to screening in these test systems, otherwise modifications to the basic protocols might be needed.

***Recommendations for Substances to be used in Pre-Validation Studies of In Vitro AR TA Agonists and Antagonists Assays***

R1881
DHT
testosterone
androstenedione
dexamethasone
cortisol
17β-estradiol
progesterone
medroxyprogesterone acetate
hydroxyflutamide
casodex (bicalutamide)
cyproterone acetate
fluoxymesterone
Linuron
*p,p*'-DDE (1,1 Dichloro-bis[4-chlorophenyl]ethylene)
finasteride
possibly other weak agonists yet to be determined

***Substances to be Included that have Known or Predicted Activity that Could Affect Luciferase Transcription***

cycloheximide *(protein synthesis inhibitor)*
actinomycin D *(RNA synthesis inhibitor)*
sodium azide *(cytotoxicant)*
specific inhibitors of luciferase activity *(none known to the panel)*
TPA *(ligand independent activation)*

It should not be necessary to include chemicals from other classes such as heavy metals, acids, bases, insoluble solids or reactive agents. Instead, it is critical to demonstrate accurate detection of known agonists and antagonists and to interpret cytotoxicity and indirect effects on luciferase synthesis and activity. Based on the possibility of ligand independent activation of AR cell systems, the phorbol ester TPA could be included as a negative control for agonism. The addition of a classic metabolic inducer like phenobarbitol or a protein synthesis inhibitor like cycloheximide as controls for antagonism might also be worthwhile. For full validation efforts, a more diverse set of chemicals such as that presented in the BRD could be considered.

Some of the suggested substances may not be readily available commercially. Substances on the list should be available from commercial sources, although hydroxyflutamide and casodex can be difficult to obtain. The U.S. EPA could provide this standard set of chemicals for validation purposes.

Some of the substances listed in the BRD are not the active forms of the chemical, including flutamide, methoxychlor, procymidone, vinclozolin and DDT. The U.S. EPA would need to provide the active forms such as HPTE (from methoxychlor) and M2 (from vinclozolin).

The list should include substances such as cortisol, 17β-estradiol and progesterone that rule out activity of other endogenous steroid receptors and also substances with known or predicted confounding mechanisms.

#### 4.1.1 The Number and Distribution of Substances Across the Range of Measurable AR/ER Binding Activity, Including Negatives

There is a need for weak substances but most that are listed in the BRD are precursors of the active forms and most of the active forms are not readily available.

#### 4.1.2 The Number and Range of Substances by Chemical Class

There is a need to reduce the number of substances within various chemical classes and increase the number of chemical classes and/or predicted confounding mechanisms that are controlled for with the recommended cytotoxicity control procedure.

#### 4.1.3 The Number and Range of Substances by Product Class

This was not considered by the Panel.

**Summary**

It is the overall conclusion of the Panel that no specific protocol was optimum for assessing AR agonist and antagonist activities. Major problems with the protocols presented in the BRD include:

1. Confounding effects of other endogenous steroid receptors in stable cell lines
2. Questions concerning the robustness of stable cell lines to detect weak androgens
3. AR patent issues relating to the transient and stable cell lines that utilize AR expression vectors and the cis-trans cotransfection methodology, and
4. Specificity of reporter vector response elements to reveal AR mediated transcriptional activation or inhibition.

With these considerations in mind, of the protocols provided by the BRD, the Panel concludes that the adenovirus infection method provides the most promising avenue for assessing AR agonist and antagonist activities and should be further considered. This assay was considered advantageous because it avoids time consuming transfection procedures as adenovirus infection protocols are straightforward. The assay showed a robust response of up to ~80 fold induction (as indicated during the meeting). Important aspects in the improvement of this protocol would be the identification and use of a cell line that lacks high response levels to the glucocorticoid and progesterone receptors. It would also make use of a reporter vector that shows greater specificity for the AR. The lack of absolute specificity for androgen binding by the AR is reflected in a general lack of specificity in hormone response in these *in vitro* assays; however, this does not parallel the *in vivo* situation. An ideal *in vitro* protocol would accurately reflect what is known about the *in vivo* physiological properties of steroid hormones. It was determined that nonlinear statistical models (e.g., the Hill equation) be used to estimate potency and steepness of the dose-response curve for full agonists and antagonists and that trend analysis be used to establish the significance of data that does not follow classical dose response relationships. The Panel recommendation for the list of chemicals for use in validation contained fewer substances than that suggested in the BRD. During the deliberations at the meeting, it was indicated that the U.S. EPA will consider supplying chemicals for validation studies, so additional weak acting agonists and antagonists could be included.

## Additional Comments on the BRD

**viii-ix:**

ER  and ER  are not isoforms. They represent the products of different genes and in some regions have little homology.

**ES-3:**

It is not necessarily difficult to reproducibly transfect the same amount of DNA. It requires close adherence to the protocol and high quality cells. Also, cells that are stably transfected with an AR plasmid are also subject to patent restrictions, in contrast to what is stated on this page. The only way around this is to use the endogenously expressed AR with a transfected reporter plasmid. The reporter plasmid may or may not have additional patent restrictions.

**1-6:**

The AR cDNA does not contain an androgen responsive promoter, but rather, a promoter (CMV) that is responsive to numerous ubiquitous transcription factors insuring a high rate of transcription of the AR in the transfected cell.

**1-7:**

In the paragraph beginning, "In a series of deletion …", should be changed to "…, while the ligand binding domain served an inhibitory function in the absence of androgen binding".

**1-9; line 4:**

The AR mutation in the LNCaP cell line "would definitely impact" on its use in screening assays, not "might impact". The LNCaP cells could not be used because this mutant AR has lost its specificity for binding androgen.

**1-10:**

It has not been shown that the AR dissociates from corepressor proteins on the binding of agonist. Also, androgen response elements are not always located "upstream" but are often within intron regions. Also "… including those necessary for cell proliferation, normal 'male' fetal development, or adult homeostasis." On this page, the AR gene is on the long arm of the X chromosome at q11-12, not the short arm as stated (see Quigley et al. 1995).

**1-11; line 5 from the bottom:**

...antagonist.

**1-12; last paragraph:**

The "AR system" is not highly conserved in vertebrates. The fish AR has not been shown to be active in mammalian cells. This is most likely due to low sequence homology in the $NH_2$-terminal region of the AR in vertebrates. This contrasts the ER where the rainbow trout ER is active in mammalian cells. The ligand binding domain is relatively highly conserved so this could be reworded to say "the ligand binding domain" rather than the AR "system".

**1-13; last line:**
Relevance is defined…

**2-7; 6th line from bottom:**
Luminescence is measured in a luciferase assay, not fluorescence.

**3-2:**
It is hydroxyflutamide not flutamide that is the AR antagonist although flutamide is given to prostate cancer patients. In most places in the text, flutamide should be replaced with hydroxyflutamide. Flutamide does not bind the AR and must be metabolized *in vivo* to the active form of the AR antagonist.

**6-4:**
The Poulin et al. 1991 reference is missing from the reference list.

**6-4:**
What is 17-alpha oxidase? The two endogenous steroid hormone biosynthetic enzymes that alter steroids at the 17 position are 17-beta-hydroxysteroid dehydrogenase (oxidizes testosterone to androstenedione) and 17-ketosteroid reductase (reduces androstenedione to testosterone). If liver metabolizing enzyme activity is meant here, are the authors referring to CYP2C11 activity? What is the product of the reaction and subsequent androgen agonist activity of the product (i.e. potential to interfere with the assay)?

**6-7:**
MDA-MB453-kb2 only has the reporter vector integrated and expresses endogenous AR. Otherwise it would not escape patent restrictions.

**6-9:**
AR patent issues also apply to cell lines with an integrated AR plasmid, no matter which plasmid was used as long as it contains the recombinant human AR sequence.

**7-5:**
For the data in Table 7-2 to be meaningful, the concentration of androgen should be indicated. The more androgen used in the studies, the higher amount of antagonist is required for inhibition. Unless all of the assays used the same concentration of androgen, the data from the different assays are not directly comparable.

**12-2:**
Binding of testosterone or DHT to TeBG could potentially be a problem in the assays if serum is included in the cell culture medium during the hormone incubations.

**14-1:**
Androgen is not technically a male hormone. It is a class of male hormones. The male hormones are testosterone and DHT. Just like estrogen is not a hormone. The hormone is 17β-estradiol (see **14-3**).

**14-4:**

It should be hypospadias and this anomaly only applies to males.

**A2-7:**

MDA-MB-453 is not a stable cell line. The cells must be transduced with adenovirus carrying the MMTV-luciferase reporter for each assay.

**A3:**

In several manuscripts, a truncated constitutively active human AR (AR1-660) was used to estimate cytotoxicity. This constitutively active human AR induces transactivation of the MMTV-Luc reporter and serves as an ideal cytotoxicity control as both nonspecific effects on transcription and luciferase enzyme activity are assessed. Alternatively, a CMV-Luc construct could be used to accomplish the same objective.

**References**

Hartig, PC, Bobseine, KL, Britt, BH, Cardon, MC, Lambright, CR, Wilson, VS, and Gray, LE, Jr. 2002. Development of two androgen receptor assays using adenoviral transduction of MMTV-Luc reporter and/or hAR for endocrine screening. Toxicol Sci 66:82-90.

He B, Kemppainen JA, Wilson EM. 2000. FXXLF and WXXLF sequences mediate the NH2-terminal interaction with the ligand binding domain of the androgen receptor. J Biol Chem 275: 22986-22994.

Quigley CA, Debellis A, Marschke KB, el-Awady MK, Wilson EM, French FS. 1995. Androgen receptor defects – historical, clinical, and molecular perspectives. Endocrine Reviews 16: 271-321.

Wang, Q, and Fondell, JD. 2001. Generation of a mammalian cell line stably expressing a tetracycline-regulated epitope-tagged human androgen receptor: Implications for steroid hormone receptor research. Anal Biochem 289:217-230.

Wilson, VS, Bobseine, K, Lambright, CR, and Gray, LE, Jr. 2002. A novel cell line, MDA-kb2, that stably expresses an androgen- and glucocorticoid-responsive reporter for the detection of hormone receptor agonists and antagonists. Toxicol Sci 66:69-81.