Peer Review Results for the H295R Cell-Based Assay for Steroidogenesis

Prepared for:

U.S. Environmental Protection Agency
Exposure Assessment Coordination and Policy Division
Office of Science Coordination and Policy
1200 Pennsylvania Avenue, N.W.
Washington, DC 20460

Prepared by:

Eastern Research Group, Inc.
14555 Avion Parkway
Suite 200
Chantilly, VA 20151-1102

12 June 2008

EPA Contract No. EP-W-05-014
Work Assignment 4-06

TABLE OF CONTENTS

Page

1.0 INTRODUCTION 1-1
1.1 Peer Review Logistics 1-3
1.2 Peer Review Experts 1-3

2.0 PEER REVIEW COMMENTS ORGANIZED BY CHARGE QUESTION 2-1
2.1 General Comments 2-1
2.2 Comments on the Clarity of the Stated Purpose of the Assay 2-1
2.3 Comments on the Biological and Toxicological Relevance of the Assay as Related to its Stated Purpose 2-3
2.4 Comments on the Clarity and Conciseness of the Protocol in Describing the Methodology of the Assay such that the Laboratory can a) Comprehend the Objective, b) Conduct the Assay, c) Observe and Measure Prescribed Endpoints, d) Compile and Prepare Data for Statistical Analyses, and e) Report Results 2-5
2.4.1 Comprehend the Objective 2-5
2.4.2 Conduct the Assay 2-6
2.4.3 Observe and Measure Prescribed Endpoints 2-9
2.4.4 Compile and Prepare Data for Statistical Analyses 2-10
2.4.5 Report Results 2-10
2.4.6 Provide Additional Advice Regarding the Protocol 2-11
2.5 Comments on Whether the Strengths and/or Limitations of the Assay Have Been Adequately Addressed 2-13
2.6 Comments on the Impacts of the Choice of a) Test Substances, b) Analytical Methods, and c) Statistical Methods in Terms of Demonstrating the Performance of the Assay 2-15
2.7 Comments on Repeatability and Reproducibility of the Results Obtained with the Assay, Considering the Variability Inherent in the Biological and Chemical Test Methods 2-18
2.8 Comments on Whether the Appropriate Parameters were Selected and Reasonable Values were Chosen to Ensure Proper Performance of the Assay, with Respect to the Performance Criteria 2-19
2.9 Comments on Whether the Data Interpretation Criteria are Clear, Comprehensive, and Consistent with the Stated Purpose 2-21
2.10 Please Comment on the Overall Utility of the Assay as a Screening Tool in the EDSP Tier 1 Battery 2-22
2.11 Additional Comments and Materials Submitted 2-26

3.0 PEER REVIEW COMMENTS ORGANIZED BY REVIEWER 3-1
3.1 Safa Moslemi Review Comments 3-1
3.2 Damian Romero Review Comments 3-10
3.3 Thomas Sanderson Review Comments 3-17
3.4 Matt Vijayan Review Comments 3-27

TABLE OF CONTENTS (Continued)

Page

Appendix A: CHARGE TO PEER REVIEWERS A-1

Appendix B: INTERIM FINAL VALIDATION REPORT B-1

Appendix C: SUPPORTING MATERIAL C-1

1.0 INTRODUCTION

In 1996, Congress passed the Food Quality Protection Act (FQPA) and amendments to the Safe Drinking Water Act (SDWA), which requires EPA to:

"…develop a screening program, using appropriate validated test systems and other scientifically relevant information, to determine whether certain substances may have an effect in humans that is similar to an effect produced by naturally occurring estrogen, or other such endocrine effect as the Administrator may designate."

To assist the Agency in developing a pragmatic, scientifically defensible endocrine disruptor screening and testing strategy, the Agency convened the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC). Using EDSTAC (1998) recommendations as a starting point, EPA proposed an Endocrine Disruptor Screening Program (EDSP) consisting of a two-tier screening/testing program with in vitro and in vivo assays. Tier 1 screening assays will identify substances that have the potential to interact with the estrogen, androgen, or thyroid hormone systems using a battery of relatively short-term screening assays. The purpose of Tier 2 tests is to identify and establish a dose-response relationship for any adverse effects that might result from the interactions identified through the Tier 1 assays. The Tier 2 tests are multi-generational assays that will provide the Agency with more definitive testing data.

One of the test systems recommended by the EDSTAC was the sliced testes assay. Its purpose in the Tier-1 battery was to provide a sensitive in vitro assay to detect chemicals that may affect the endocrine system by inhibiting the enzymes responsible for the inhibition of enzymes in the steroid hormone synthesis pathway. After encountering two substantial issues with the standardization of the sliced testes assay—high variability and the inability to distinguish general cytotoxicity from Leydig cell toxicity—EPA abandoned the sliced testes assay in favor of the H295R. The H295R assay offered a number of substantial advantages over the sliced testes assay and other cell-based assays. Like other cell-based assays it does not use animal tissue and is capable of detecting inducers as well as enzyme inhibitors. Unlike the other cell-based assays, it contains all of the enzymes of the steroidogenic pathway.

Although peer review of the steroidogenesis assay was performed on an individual basis (i.e., its strengths and limitations evaluated as a stand alone assay), it is noted that the steroidogenesis assay, along with a number of other in vitro and in vivo assays, will potentially constitute a battery of complementary screening assays. A weight-of–evidence approach is also expected to be used among assays within the Tier-1 battery to determine whether a chemical substance has a positive or negative effect on the estrogen, androgen or thyroid hormonal systems. The Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) Scientific Advisory Panel (SAP) has already conducted a peer review of the EPA's recommendations for the Tier-1 battery. The steroidogenesis assay was one of the assays recommended by EPA contingent upon satisfactory validation and peer review of the assay.

The purpose of this peer review was to review and comment on the steroidogenesis assay for use within the EDSP to detect chemicals that may affect the endocrine system by inhibiting the enzymes responsible for the inhibition of enzymes in the steroid hormone synthesis pathway. Unlike other peer reviews, EPA did not have time to produce an Integrated Summary Report (ISR); therefore, peer reviewers were asked to focus on the Interim Final Validation Report and to a lesser extent on the prevalidation reports for conducting this review.

The remainder of this report is comprised of the **unedited** written comments submitted to ERG by the peer reviewers in response to the peer review charge (see Appendix A). Section 2.0 presents peer review comments organized by charge question, and Section 3.0 presents peer review comments organized by peer review expert. The Interim Final Validation Report is presented in Appendix B and additional supporting materials are included in Appendix C.

The final peer review record for the steroidogenesis assay will include this peer review report consisting of the peer review comments, as well as documentation indicating how peer review comments were addressed by EPA, and the final EPA work product.

1.1 Peer Review Logistics

ERG initiated the peer review for the steroidogenesis assay on April 30, 2008. ERG held a pre-briefing conference call on May 15, 2008 to provide the peer reviewers with an opportunity to ask questions or receive clarification on the review materials or charge and to review the deliverable deadlines. Peer review comments were due to ERG on or before June 4, 2008.

1.2 Peer Review Experts

ERG researched potential reviewers through its proprietary consultant database; via Internet searches as needed; and by reviewing past files for related peer reviews or other tasks to identify potential candidates. ERG also considered several experts suggested by EPA. ERG contacted candidates to ascertain their qualifications, availability and interest in performing the work, and their conflict-of-interest (COI) status. ERG reviewed selected resumes, conflict-of-interest forms, and availability information to select a panel of experts that were qualified to conduct the review. ERG submitted a list of candidate reviewers to EPA to either (1) confirm that the candidates identified met the selection criteria (i.e., specific expertise required to conduct the assay) and that there were no COI concerns, or (2) provide comments back to ERG on any concerns regarding COI or reviewer expertise. If the latter, ERG considered EPA's concerns and as appropriate proposed substitute candidate(s). ERG then selected the five individuals who ERG determined to be the most qualified and available reviewers to conduct the peer review.

Following initiation of the review, one of the five selected reviewers indicated they would no longer be able to conduct the review due to availability constraints. ERG initially began efforts to identify a replacement reviewer; however, after discussions with EPA, it was determined that the four remaining reviewers provided the breadth of expertise required to provide EPA with adequate feedback. Therefore, a replacement reviewer was not identified.

A list of the peer reviewers and a brief description of their qualifications is provided below.

- ***Safa Moslemi, Ph.D.,*** is currently an Assistant Professor in Biochemistry at the University of Caen, a researcher in the Biochemistry and Molecular Biology Laboratory studying estrogens and reproduction, and a researcher in the Extra- cellular Matrix and Pathology Laboratory at the Institut de Biologie Fondamentale et Appliquée (I.B.F.A.). He received his Ph.D. in 1993 from Ecole Nationale Supérieure des Industries Agricoles et Alimentaires (E.N.S.I.A), and in 1998 a diploma of the capacity for research management from University of Caen, France. Dr. Moslemi's research domains include purification, characterization and cloning of equine aromatase; analysis and quantification of sexual steroids in the physiological medium; development of non-steroidal inhibitors against aromatase; evaluation of xenobiotics as endocrine disruptors using different in vitro models; and the role of estrogen and their receptors on the biomarkers expression in differentiated and dedifferentiated chondrocytes.

- ***Damian Romero, Ph.D.,*** is currently an Assistant Professor in the department of Biochemistry at the University of Mississippi Medical Center. He received his PhD degree in Molecular Biology at the University of Buenos Aires, Argentina in 2000. Dr. Romero received his post-doctoral training in the Departments of Medicine, and Physiology and Biophysics at the University of Mississippi Medical Center. For the past 15 years, his research has focused in the molecular mechanisms that regulate mineralocorticoid biosynthesis by adrenal and extra-adrenal tissues. Dr. Romero is a member of professional societies such as the Endocrine Society, American Physiological Society and American Heart Association. Dr. Romero reviews manuscripts for multiple peer-reviewed journals and is a member of the IACUC committee for the G.V. Montgomery VA Medical Center. Dr. Romero has published multiple peer-reviewed articles in journals such as *Endocrinology, Journal of Endocrinology, Physiological Genomics, American Journal of Physiology, Journal of Steroid Biochemistry and Molecular Biology, Molecular Reproduction and Development, Molecular and Cellular Endocrinology, Steroids,* etc.

- ***Thomas Sanderson, Ph.D.,*** obtained his bachelors degree (BSc 1989) from the Faculty of Chemistry and Pharmacochemistry, Free University of Amsterdam, the Netherlands. He went on to complete a PhD degree (PhD 1994) in Pharmacology and Toxicology at

the Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, BC, Canada. His doctoral research focused on the toxic effects of dioxins and PCBs on various wild and domestic avian species. This avian toxicology work was continued during his postdoctoral research training (1994-1996) at the National Food Safety and Toxicology Centre at Michigan State University, MI, USA. It is here that his research interests turned towards endocrine disruption and steroid hormone synthesis and metabolism. During his assistant professorship (1997-2005) at the Institute for Risk Assessment Sciences, University of Utrecht, the Netherlands, he established a research program to study the effects of xenobiotics on the steroid biosynthesis pathway in humans and wildlife. Key research accomplishments are the identification of aromatase, the enzyme responsible for the conversion of androgens to estrogens, as an important target for endocrine disrupters. His work demonstrated that induction or inhibition of aromatase activity or expression by various pesticides, medicinal drugs and naturally occurring phytochemicals posed an alternate non-receptor mediated mechanism by which chemicals can cause pro- or antiestrogenic/androgenic effects in humans and wildlife. As associate professor (2005) at the Institut Armand-Frappier in Montréal, QC, Canada, Thomas Sanderson is focusing his research on the effects of chemicals on the regulation of expression and catalytic activity of several key enzymes involved in the biosynthesis of potent steroid hormones, such as aromatase, steroid 5-alpha reductase and steroid 17-alpha-hydroxylase/17-20-lyase.

- ***Matt Vijayan, Ph.D.,*** is a Professor and University Research Chair in the Department of Biology at the University of Waterloo. The focus of Dr. Vijayan's laboratory is elucidating the molecular and biochemical strategies involved in allowing animals to cope with stress, including the mechanism(s) of action of endocrine disruptors. Using fish as a model, studies are aimed at understanding the mechanisms of action of endogenous (stress hormones) and exogenous (xenobiotics) signals on the cellular stress response process. Dr. Vijayan's research includes characterization of the stress axis and its modulation by contaminants, including glucocorticoid biosynthesis, receptor (GR) dynamics and GR signaling pathways as well as feedback regulation of cortisol in circulation. His research also utilizes functional genomics and proteomics approaches to identify stressor-specific and stressor-non-specific regulatory networks that are involved

in stress adaptation. Dr. Vijayan has developed a rainbow trout-specific low density targeted cDNA microarray (~210 genes) for characterizing the functional basis of gene expression patterns seen with stressors, including endocrine disruptors.

2.0 PEER REVIEW COMMENTS ORGANIZED BY CHARGE QUESTION

Peer review comments received for the steroidogenesis assay are presented in the sub-sections below and are organized by charge question (see Appendix A). Peer review comments are presented in full, unedited text as received from each reviewer.

2.1 General Comments

Safa Moslemi: This report « Multi-Laboratory validation of the H295R Steroidogenesis Assay to identify modulators of testosterone and estradiol production » aimed to develop and standardize of the assay as a screen of steroidogenesis using H295R cell line to identify chemicals that act as endocrine disruptors. You find here after the answers and comments to the charge questions.

2.2 Comments on the Clarity of the Stated Purpose of the Assay

Safa Moslemi: Yes. The steroidogenic screen assay consist to detect any natural and human-made substance that would disrupt endogenous estrogens and/or androgens production. In this way, the assay will complement the other Tier 1 assays which aim to identify xenobiotics that could be classified as endocrine disruptors of both human and wildlife. The use of H295R cell line present several advantages making this model unique when compared with other models. Besides its availability, this model allows the detection of both increases and decreases in the production of testosterone (T) and estradiol (E2) in the presence of chemicals, and to follow the direct potential impact of a chemical on cell viability/cytotoxicity. Furthermore, H295R cells express a wide range of steroidogenic enzymes found in the adult adrenal cortex and the gonads, including those required to produce, cholesterol, mineralocorticoids, glucocorticoids, androgens and estrogens. Thus, this cell line enable the research of any target site within the steroidogenic pathway downstream of cholesterol in addition to T and E2 investigated in this work.

Damian Romero: The stated purpose of the assay is "The steroidogenic assay is intended to identify xenobiotics that have as their target site(s) the endogenous components that comprise the intracellular biochemical pathway beginning with the sequence of reactions occurring after the

receptor, up through and including the production of the terminal steroid hormones, i.e. testosterone (males) and estradiol/estrone (females)" (page 16). Although it is clear in general terms, it could be improved in the following structures:

1. "…the endogenous components…" it could be replaced with "intracellular components", since now the assay is performed using the H295R cell line, in contrast to the original design using rodent sliced testes which perhaps required a broader description.
2. "… occurring after the receptor…" although it is probably referring to LH/FSH receptors it could be confused with androgen/estrogen receptors, it could be clarified to avoid confusion with the later ones.
3. "…terminal steroid hormones…" may be replaced with "…terminal sex steroid hormones" to avoid confusion with other steroid hormones, especially since the cell line used in the assay also secretes gluco- and mineralo-corticoids.

Thomas Sanderson: The 'interim H295R validation report' states that section 408 of the FFDCA requires EPA to:

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

This directive has been redefined by the EPA resulting in the development a two-tier testing system for endocrine disruptors, which would cover disruption of the androgen, estrogen and thyroid hormone systems. The interim report defines the steroidogenesis assay as a screening tool for the detection of any substance that would disrupt estrogen and/or androgen gonadal steroid hormone production. The definition goes on to say that the steroidogenesis assay is intended to detect any disruption of the intracellular biochemistry involved in the formation of the gonadal estrogens and androgens, but excluding any disruptions that may occur before the receptor (question: which receptor(s)?), effects on storage of sex hormones or effects on the hypothalamic-pituitary-gonadal axis. However, it remains unclear which steps involved in steroidogenesis are considered to be part of this tier 1 screening tool and which are not. For example, are interactions of chemicals with cell surface receptors that may modulate

steroidogenesis included or not (the definition states *...after the receptor...*)? …and what about effects on cholesterol storage/release/*de novo* synthesis?

Matt Vijayan: Yes, it is clear that the H295R assay is being used as a screening tool to detect substances that will impact estrogen and/or androgen production.
Pg.17. Not sure what you mean by "storage or release of gonadal steroid hormones"?
Pg.17, last but one line. "…identify chemicals that will act to alter steroidogenic process". This assay will not identify chemicals that will act upstream of cAMP, including trophic hormone stimulation.

2.3 Comments on the Biological and Toxicological Relevance of the Assay as Related to its Stated Purpose

Safa Moslemi: Actually, many of chemicals (especially xenoestrogens) have the potential to disrupt endocrine processes at tow major levels; first at sex hormone receptors, particularly the estrogen receptor, and second at steroidogenic enzymes involved in both steroid synthesis and metabolism. The second level being more appropriate target since most of the environmental chemicals, when introduced in the organism, are present at low concentrations and show relatively, when compared with estradiol, low affinity for estrogen receptors. Therefore, estrogen receptor pathway could not be considered as an ideal endpoint to study endocrine disruption of xenobiotics. Utility of H295R cells has been well established as an unique model for study of steroidogenic pathways but also to test and evaluation of xenobiotics since these cells express genes that encode for all enzymes of steroidogeneis especially those involved in sexual mal and female hormones synthesis; androgens and estrogens. Furthermore, this model permits to evaluate, in the same cells, the potential cytotoxicity of chemicals allowing the discrimination between effects that are due to cytotoxicity or due to the direct interaction of chemicals with steroidogenic enzymes.

Damian Romero: The H295R steroidogenesis assay is biologically and toxicologically relevant to the stated purpose. The assay would fit perfectly in the Tier 1battery of assays to screen for endocrine disruptors. The assay has a series of strengths that would make it an excellent screening tool for endocrine disruptors of sex steroid hormone synthesis. However,

results obtained with this test should always be interpreted along with the results obtained with all the other assays of the Tier 1 battery.

It is important to stress that chemicals that generate a negative result in the H295R steroidogenic assay could be false negatives and they should not be considered safe without a complete evaluation of them with the other Tier 1 battery assays. This in vitro system lacks that ability to study complex interactions that could occur in vivo such as metabolism of tested compounds, biodistribution, interaction with other endocrine systems that may modulate sex hormones steroidogenesis, etc.

Thomas Sanderson: Estrogens and androgens are crucial hormones for human development and are involved in numerous processes in almost all tissues of the body. Dramatic changes in sex hormone levels during critical periods of development are known to adversely affect human development and health. Thus interferences with sex hormone production induced by chemicals or otherwise may have deleterious repercussion for the organism. However, sex hormone levels vary considerable among individuals; they also vary considerably during the day and month and are relatively easily affected by stress, diet and other life-style factors. They are also produced by several tissues in the body, not just the gonads.

One major question is whether any small change in hormone production in an isolated *in vitro* system has any relevance for the health outcome of an exposed organism. This remains unaddressed in the documents available for review.

Also, how relevant is the use of an adrenocortical cell line for what is intended to be the screening of chemical effects on the gonadal sex hormones? Although steroidogenic enzymes such as CYP19 and CYP17, for example, are the same in these tissues, they are not regulated in the same way in the adrenal cortex as in the gonads (Bulun *et al.*, 2003; Simpson, 2003).

Furthermore, various other factors that may influence basal secretion of steroid hormones, such as extent of conjugation to sulfates/glucuronides, activity of transport proteins and mechanisms that control storage capacity will be different in different tissues. As the steroidogenesis assay only looks at one final outcome, namely the amount of estradiol and testosterone secreted, it is not possible to make biologically meaningful statements on the relevance of any observed disruption for the organisms as a whole. There are so many factors not directly related to steroidogenesis that could influence the assay system as it is currently described and intended to

be used, that the issue of 'false positives' is likely to be an important concern, particularly once dealing with unknown complex environmental samples.

Matt Vijayan: The H295R cell line is derived from a human adrenocortical carcinoma cell line and used a model system to examine sex steroid production. The rationale being that the full complement of the steroid biosynthetic pathway is present in this cell system. Also, this cell line produces sex steroid constitutively for its use to detect inhibitors of steroid production. Consequently, use of this model system as a screening tool to identify substance that can potentially impact steroid biosynthetic pathway leading to hormone production is valid. However, the extrapolation of this information, from a carcinoma cell line, to the impact (or lack off) of these test substances *in vivo* requires further validation as the endocrine physiology of the gonads and the adrenal gland is different from this undifferentiated cell system. Specifically, this cell system lacks the membrane receptors required for trophic hormone (gonadotropin) stimulation, which is an essential component of the sex steroid biosynthetic cascade.

2.4 Comments on the Clarity and Conciseness of the Protocol in Describing the Methodology of the Assay such that the Laboratory can a) Comprehend the Objective, b) Conduct the Assay, c) Observe and Measure Prescribed Endpoints, d) Compile and Prepare Data for Statistical Analyses, and e) Report Results

Safa Moslemi: Protocol is well described and the methodology generally presented in a comprehensible manner allowing the reader to fallow all steps cited above.

2.4.1 Comprehend the Objective

Damian Romero: The objective of the assay is clearly stated.

Thomas Sanderson: The objectives of the H295R steroidogenesis assay are not very clearly described in the appendices I and II. The sections *Purpose* and *Scope* are not very informative. Under *Purpose*, for example, the purpose of the document is described, not the purpose of the actual assay the document is meant to describe. Under *Scope and Application*, the reader does not find an easy guide to what the assay is about. It is also not helpful that appendix I has an appendix I and II and that appendix II has an appendix I.

It would be more logical to have a single protocol that covers the four main aspects of the H295R steroidogenesis assay (1) Cell culture (2) Exposure to test compounds (3) Analysis of estradiol and testosterone (4) Data analysis and presentation/interpretation of the results.

Matt Vijayan: YES, the protocol is easy to follow and the objectives are clear.

2.4.2 Conduct the Assay

Damian Romero: Both Standard Operating Procedures (SOP), "Culturing of the H295R human adrenocortical carcinoma cell line" (SOP#1) and "Exposure of the H295R human adrenocortical carcinoma cells…" (SOP#2), are clear and allow the operator to conduct the assay. However there are some points that need clarification and/or need to be improved as indicated below:

1. SOP#1, page 5, item 2.2: it says "Do not freeze cells upon arrival…", unless cells arrive to the lab growing (which is fairly uncommon for ATCC cultures) there should be not much difference in keeping them for a short period of time in liquid nitrogen. In other case, an appropriate reason should be stated, since the requirement to immediately begin to culture the cells may create a burden to the lab that may be not necessary. Furthermore, in SOP#1, page 13, item 5.1.2 it is indicated to remove the vial of cells from the liquid nitrogen storage. In any case, it would be perhaps also useful to stress that cells should always be stored in liquid nitrogen to avoid any confusion.
2. SOP#2, page 6, item 2, the examples about the nomenclature of the cultures should be checked. It seems that numbers 5.4, 5.7 and 5.2 should be 4.5, 7.5 and 2.5 following the example in Appendix II of SOP#1. Idem in item 5.1 and 5.2 on pages 8 and 9 of SOP#2.
3. SOP#2, page 10, item 5.2.1, in the Reagents section it is stated "… for 6 and 11 generations…", since the term "generation" is not clearly defined in the text, and not used anywhere else, it would be better to maintain consistency to use and refer to "passages" throughout the text.
4. SOP#2, page 10, item 5.2.5, it says "General rule: use 1 petri dish…", since the statements is giving a quantitative recommendation is very important to indicate the size (probably a 100 mm diameter) of the cell culture dish. Also, for consistency, refer to "petri dish" as "cell culture dish" since "petri dish" is not used anywhere else in the protocol.

5. SOP#2, page 12, item 5.3.2, when the protocol refers to add the media and chemical compounds for testing it is not very clear. From the text and Table 3, it seems that it is recommended to add 1 ml of media per well and then add 1 µl of stock chemical solution to the well. This procedure could be a great source of error if it is performed in that way. In these cases, it is greatly preferred to make a "master mix" (i.e. 5 ml media plus 5 µl of the tested chemical stock solution) and then dispense 1 ml of media plus tested chemical or solvent per well. It would be very useful to clearly indicate that this is the preferred method to add the tested chemicals to the wells of the cell culture plate.
6. SOP#2, page 14, item 6.1.1, in the "Equipment" section it is indicated the use of a "Fluoroskan Ascent Fluorometric Microtiter Plate Reader", it would be better to indicate that the protocol have been extensively tested and validated with that particular piece of equipment but that any fluorometer microplate reader that have the adequate filters may be used.
7. SOP#2, page 14, item 6.1.1, in the "Materials" sections it says 200,000 cells per well and it should be 300,000 cells per well to be consistent throughout the protocol.
8. SOP#2, page 15, item 6.1.2, it should be useful to indicate in 6.1.1 "Materials" section the brand and catalog number of the plate sealers to be used.
9. SOP#2, page 16 and subsequent, item 7 and subitems, it is recommended to spike the sample with ^{3}H-testosterone for recovery calculation. From the text it is not clear if the same protocol should be used for estradiol extraction and recovery calculation. Should the same solvent be used for estradiol extraction? Should ^{3}H-estradiol be used for estradiol recovery calculation?
10. SOP#2, page 16 and subsequent, item 7 and subitems, anhydrous ether is the solvent recommended for steroid extraction. The use of ether is a serious hazard concern since it is highly flammable having a flammability rating of 4, the highest possible. Although in SOP#2, section 3, it is stated that "Special safety requirements need to be considered when working with ether …" a more serious advice should be given since this solvent is an extremely serious hazard known to have caused multiple laboratory accidents. It is highly recommended to explore the use of other solvents that are not as hazardous as ether that would make the protocol safer and easier to perform.

11. SOP#2, page 17, item 7.1.2, in point "5" it says "add 10 µl of the ^{3}H-labeled hormone" , since this point is in the middle of the extraction procedure it would be useful to indicate that this tube would be used to calculate the CPMs of the "CPM spike tube".
12. SOP#2, page 19, item 7.2, the formula for the "final hormone concentration" should be updated to include the volume of reconstituted sample used in the assay and the necessary corrections to refer to the final volume of media of 1 ml. The formula should be:

$$Final\ hormone\ concentration = 150\ pg * 0.69 * \frac{1000\ \mu l}{450\ \mu l} * \frac{250\ \mu l}{vol\ used\ in\ assay\ (\mu l)} * 10$$

Using the same example as in SOP#2 with the following additions:
1000 µl refers to the final volume of media in the well
Vol used in assay (µl) is the volume in microliters used in the hormone concentration assay

The formula needs to be checked since the concentration of secreted steroid hormones is one of the quality criteria of the assay.

General considerations for cell culture that could be stressed through the SOP:

1. Perform all operations with cells in a GENTLY manner.
2. ALWAYS remove media/reagents from the well border.
3. ALWAYS add media/reagents resting the pipette against the well wall.
4. NEVER vortex or vigorously shake cells.

Thomas Sanderson: The protocol is described at length in appendices I and II. However, no methodological information whatsoever is provided in the interim report, which is a major limitation to the comprehension of the results discussed in this document.

Questions concerning the protocol for H295R cell culture: is Nu-Serum available world-wide? Also, what are the batch to batch variations in the sex hormone content of the Nu-Serum? Finally, why was not the use of a steroid-free medium recommended?

Matt Vijayan: The protocol is provided in sufficient detail and the methodology is well laid out for any laboratory to conduct the assay.

2.4.3 Observe and Measure Prescribed Endpoints

Damian Romero: The prescribed endpoints can be easily measured following the protocol. One strength of the protocol that will allow its widespread use is the possibility to use any testosterone/estradiol detection method already in use in the laboratory if it reaches the quality controls specified in the protocol.
As explained below the routine inclusion of controls to test for each chemical positively or negatively affecting the steroid quantification should be advisable to generate more confidence in the assay performance.

Thomas Sanderson: One of the most important aspects of the H295R steroidogenesis assay, the analysis of testosterone and estradiol, is poorly defined in the provided documents. The choice of analysis method is left entirely to the implementing laboratory. It is known that ELISAs and RIAs can have very different outcomes dependent on the sample dilution, kit and antibodies used, not to mention the numerous confounding factors (solvent, cross-reactive components). The issues of cross-reactivity, how to deal with conjugated metabolites, and how to reliably compare between hormone levels determined by RIA or LC-MS are left undiscussed. It is highly inconsistent that there is an elaborate protocol for the 'consistent' use of a standard method such as the LIVE/DEAD cytotoxicity kit while no detailed attention is given to the crucial hormone analysis methodology.

Matt Vijayan: The endpoints involve collection of medium for measuring steroids and the cells for cytotoxicity assay. The steroid measurement may involve hormone extraction from the medium and a methodology is provided for consistency in extraction efficiency among laboratories. A protocol that can eliminate this extraction step may be better suited for wider application. This can be ascertained by testing the interference of the test substances with the antibody cross-reactivity. The cytotoxicity assay is also well explained and easy to carry out. However, I am not clear what greater than 100% cell viability means (Figs. 9.5 and 9.6 and pg. 56) as this could potentially confound the results.

2.4.4 Compile and Prepare Data for Statistical Analyses

Damian Romero: The worksheet design is adequate for data compilation and statistical analysis submission.

Thomas Sanderson: There does not appear to be a section dedicated to this important aspect of the steroidogenesis assay in the protocols (Appendices I, II, III). Some information can be retrieved from the interim report (Chapter 8), raising questions: How is normality tested, is it by a standard chi-square test? The steroidogenesis assays essentially requires only that a deviation from the basal secretion of estradiol or testosterone is tested statistically. This can be done by using a two-tailed t-test or its non-parametric version, the Mann Whitney U test, with or without correction for multiple comparisons, if required (Zar, 1999). The text, however, mentions the non-parametric version of ANOVA, the Kruskal-Wallis test. An ANOVA-style test is not really appropriate for concentration-response data. Can this be clarified?

Matt Vijayan: The data compilation is explained clearly and a data sheet template is also provided. However, there appears to be some confusion around data normalization. The magnitude of change from control either shown as actual concentration change or percent change would be appropriate for inter-laboratory validation.

2.4.5 Report Results

Damian Romero: Although the preliminary report deals extensively with data analysis and report the protocol "Exposure of H295R…." does not address this point satisfactorily.
From the extensive preliminary report addressing several analysis techniques based on the data generated with the core chemicals it seems that the use of "Fold change" in combination with "Percent of control" to be the most adequate way to report the results. This procedure was applied with the supplementary chemicals and in the report is shown to have worked very satisfactorily.
After reaching a consensus, the protocol "Exposure of H295R…." should include a section indicating how the results are going to be analyzed, how chemicals are going to be classified, etc.

Similarly, the section 6.1.3 of SOP#2 "Exposure of H295R…." should include clear cutoffs in order for a chemical to be further analyzed regarding steroid synthesis. In the preliminary report a cutoff of more than 80% viability was used and it seems to be an excellent choice since chemicals which further decrease viability would probably have non-specific effects on steroidogenesis.

Thomas Sanderson: There does not appear to be a section dedicated to this aspect of the steroidogenesis assay and this should be included. For example, how should concentration-response data be expressed and presented? There are several figures in the interim report that are not interpretable: Figures 3.4 and 3.5 express testosterone concentrations as a % of the maximal response to prochloraz/fadrozole but the percentages are negative. What is considered the *maximum* response in these figures, and what does -20% of the maximum response of prochloraz/fadrozole mean?
The same problem returns in figures 10.1-10.4. Zero % of any response is zero, -20% of a response is impossible. A consistent approach would be to express all data as a % of basal hormone secretion, as this conforms to the aim of the assay as currently defined.

Matt Vijayan: The results and the statistical analyses are clearly explained and easy to follow.

2.4.6 Provide Additional Advice Regarding the Protocol

Safa Moslemi: In order to improve protocol, the following advices are proposed:
1) The choice of solvent for steroid extraction should be précised (ethyl ether or dichloromethane).
2) During extraction procedure of steroids with ethyl ether a rapid freezing of aqueous phase (after step 9 and before step 10) facilitate the separation of aqueous (inferior) and organic (supernatant) phases (see Appendix ii page 17).
3) Collected solvent phases could be washed by distilled water to eliminate hydrophilic contaminants and to reduce background of the detection by RIA or ELISA.
4) Hormone purification, at least for protocol validation and before using a detection system based on antibody (ELISA, RIA), being necessary to avoid cross reaction observed with chemicals especially with trilostane and to reduce background of assay.

5) What is the maximum passage of cell culture to be respected; 10 passages (Appendix I page 7) or 7 passages (Appendix ii page 7) ?

6) When possible, replace methanol (highly toxic) by ethanol for cytotoxicity analysis (see page 36, point 5 & table)

7) There is a confusion between E2 decrease in the presence of 3 □M prochloraz when determining performance criteria for each laboratory. In page 35, table 7.1, the average change in hormone production relative to the solvent control (SC=1) was not reported (n/a, please spell this acronym) whereas in page 36, point 4, a change of 50% in E2 reduction was reported!

8) Also, there is a difference in the reported induction of E2 by 10 □M forskolin ; ≥10-times induction of E2 in page 36 point 3, different from ≥15-times induction in table 7.1, page 35!

9) In page 44, table 9.1, the Minimum Detectable Level (MDL) and the measuring system used (RIA, ELISA or others) of each laboratory should be reported.

Damian Romero: 1. Cell suspension: H295R cells have a strong tendency to clump after trypsinization and this could be one of the reasons for the relatively high degree of variation observed in some of the protocols and specially with inexperienced laboratories. Although it is indicated in several parts of the protocol, I think it should be further highlighted the necessity to gently but consistently resuspend the cells after homogenization and/or centrifugation. The use of a pipette that can hold all the volume of media containing the cells that needs to be resuspended is very important. Also, the use of pipetting device that can aspirate/deliver liquid at an adequate speed to ensure good cell resuspension.

2. Use of "master mixes" to add test compounds: As indicated above if the use of "master mixes" was not the routine procedure it is greatly advice to use them to reduce the error due to pipetting small volumes.

3. Addition of media to cell culture plate wells: The volume added to each well of media plus test compound is very important since it later on it is assumed to be exactly 1 ml for all calculations. To reduce the error, it is greatly advice to prepare a "master mix" as indicated above and then dispense 1 ml per well using a 1 ml pipette, giving even preference to the use of a micropipette. The use of pipettes that can hold larger volumes, i.e. 5 ml pipettes, could add a significant error to the assay due to volume variations between wells of the media dispensed.

4. Since basal estradiol synthesis is very low and ,as indicated through the text, it is difficult to evaluate inhibitors of estradiol synthesis. One possible alternative would be to test the chemicals

in cells treated simultaneously with forskolin where it would be easier to observe a decrease in the forskolin-mediated estradiol induction than a reduction of already low basal levels.
5. The crossreactivity of the core chemicals was evaluated in section 9.2.3 "Confounding factors". However, since the assay is planned to be used with a series of chemicals, it would be recommended to routinely test each of the chemicals or samples to be tested using the H295R steroidogenic assay for positive (as it was tested for the core chemicals) as well as negative interference effects. Each chemical should be tested at least at the higher concentration used for both interfering effects: a) positive: media which have had no contact with cells supplemented with the chemical at the highest concentration tested; b) negative: media which have had no contact with cells spiked with either testosterone or estradiol and supplemented with the chemical at the highest concentration tested. This test should be run routinely for each tested chemical/sample and will help to identify chemicals/samples that either increase or decrease the apparent concentration of each steroid in the determination assay.

Thomas Sanderson: *Analysis of sex hormones.* The greatest weakness in the protocols is the lack of detail on sex hormone analysis methodology. This reviewer is of the opinion that LC-MS would be, by far, the preferred analysis tool for the detection of testosterone and estradiol. LC-MS would avoid the problems that will be (and already have been) encountered with inappropriate cross-reactivity of test samples/chemicals with the antibodies used in sex steroid ELISAs and RIAs. Please see also comments on trenbolone under point 7. The validation of a sensitive LC-MS method should be a logical part of the H295R steroidogenesis assay as currently defined. Furthermore, a single LC-MS analysis could detect a number of steroids in addition to estradiol and testosterone at little additional effort/expense, thus improving the 'expandability' of the H295R tool for other hormone endpoints.

2.5 Comments on Whether the Strengths and/or Limitations of the Assay Have Been Adequately Addressed

Safa Moslemi: The advantages and disadvantages of H295R regarding to other cell lines should be detailed especially to JEG-3 and JAR placental choriocarcinoma cell lines. For instance, JEG-3 and JAR placental choriocarcinoma cell lines appear relatively more sensitive to cytotoxic effects of chemicals than H295R cell line (Letcher et al, 1999). This rises the question about the suitable model (more sensitive or less sensitive to cytotoxicity) to screen chemicals for their

endocrine disruption effect since the endocrine disruption of chemicals is tested at non cytotoxic concentrations and this might affect interpretation of results, chemical classification and determination of their tolerability concentrations (threshold) in organism.

Damian Romero: The strengths and limitations have been adequately addressed in the protocol. The major strengths of the assay are that: 1) H295R cells are commercially available, 2) it is an in vitro system that does not require the use of live animals, 3) H295R cells are of human origin which would make results more relevant to human endocrinology and cell physiology, 4) the protocol is relatively easy to perform allowing its wide use as an screening tool, 5) the possibility to use any steroid determination methods that successfully passed the quality control criteria using validated standards.
The limitations of the assay are that: 1) the system does not allow to study complex interactions that occur in vivo, 2) the system does not allow to study the regulation of the hypothalamic-pituitary axis, 3) the system does not allow to detect very weak inducers or inhibitors, 4) the system does not allow to study the effect of metabolites of the tested chemicals generated *in vivo*.

Thomas Sanderson: There is a brief discussion of strengths and weaknesses, but lacks detail and supporting scientific references. The main strength mentioned in the interim report is that the H295R cell line is a pluripotent cell lines that expresses all the enzymes necessary for the production of testosterone and estradiol. However, the fact that numerous other steroid hormone synthesis pathways are also present, although acknowledged, is not discussed. The implications of the presence of these other pathways (aldosterone, cortisol synthesis) may be far reaching for the reliable application of the proposed H295R steroidogenesis assay, as all these pathways are interconnected (at least in adrenocortical cells, not necessarily in gonadal cells). There is no critical discussion of the potential drawbacks of choosing an *adrenocortical* cell line to study effects of chemicals on *gonadal* testosterone and estradiol production. There is no scientifically supported discussion of the possible differences in regulation of steroidogenesis in adrenocortical cells and gonadal cells, yet it is known these are qualitatively and quantitatively very different. Several of the above points have been discussed in detail in several publications from my own lab in recent years (Sanderson and van den Berg, 2003; Sanderson, 2006).

Matt Vijayan: The strengths have been addressed adequately but the limitation of the assay requires mention (see pg. 20 line 3 onwards). For instance this assay will only detect changes that happens post-receptor activation. This is a drawback to this cell system because *in vivo* the steroidogenic cells secrete steroids in response to trophic hormone stimulation. This assay completely bypasses the receptor signaling which is an essential step in steroid biosynthesis. So substances that can affect steroid production by altering trophic hormone signaling will not be evaluated by this cell system. Also, the high constitutive production of the hormone is abnormal *in vivo* as this usually happens only in response to trophic hormone stimulation. So it is unknown whether the changes seen with the test substances can be mimicked *in vivo* to the same extent (or may be even greater) and will require confirmation with animal models or other relevant cell or tissue systems. Also, the high constitutive levels of steroids, for instance testosterone, may deplete the precursor available for steroid synthesis and may be limiting the steroid biosynthetic capacity in response to test (inducer) substances. The changes in the magnitude of steroid synthesis with forskolin, smaller change for testosterone because basal secretion is high and higher for E2 because of lower basal secretion, clearly support this contention. This requires testing perhaps by supplementing the medium with cholesterol.

2.6 Comments on the Impacts of the Choice of a) Test Substances, b) Analytical Methods, and c) Statistical Methods in Terms of Demonstrating the Performance of the Assay

Safa Moslemi: Yes, there is in general a good choice of different chemicals, analytical and statistical methods. However, information concerning the effect type on T and E2 production should be updated for some chemicals (danazol, finasteride, flutamide, Glyphosate, RU-486/mifepristone, spironolactone, taxol etc in table 6.3, pages 33 & 34). For instance, danazol is known to : inhibit aromatase transcription in ectopic human endometrial tissue (Fechner et al, 2007), inhibit aromatase activity of endometriosis-derived stromal cells (Murakami et al, 2006), induce a marked up-regulation of free T and down stream 17□-E2 in hereditary angioedema (Thon et al, 2007). Glyphosate (Roundup) showed also to inhibit aromatase in vitro (Richard et al, 2005; Benachour et al, 2007). Moreover, protocol and analytical method should be revised for E2 evaluation in the presence of inhibitor chemicals since production of E2 was not evident during validation assay in H295R cells after 5 passage (see additional comments in point 9).

Damian Romero: The test substances, and analytical and statistical methods chosen were appropriate to validate the assay. However, a decision should be done regarding the most adequate methods for data analysis and data report and this should be clearly stated in the SOP "Exposure of H295R….".

Thomas Sanderson:

Choice of chemicals:

The compounds selected appear to be largely appropriate for validation of the assay, although the information given in Table 6.2 and 6.3 to support the choice was not very helpful. Specifically, the information under heading 'mode of action' and 'effect type' is not clear. Under *mode of action* a target may be mentioned but no information is given concerning the effect on that target. For example, is an *ER binder* an agonist or antagonist? Is trilostane really a strong inducer of T and E2 production (which seems unlikely given its 3bHSD inhibition potential), or is this erroneously based on the results of the present interim report, which indicates that the apparent induction of E2 and T is the artefact of cross-reactivity with the immunoassay kit? Danazol is said to have unknown effects, however, it is a well known (no longer used) medication against endometriosis withdrawn for its anabolic/androgenic effects. How is vinclozolin an inducer and inhibitor of T production at the same time? Also, flutamide, genistein, glyphosate, RU486 and spironolactone are missing relevant information on their mode of action. Table 6.3 also needs references, and abbreviations need to be defined.

Analytical methods:

The use of immunoassays for the determination of testosterone and estradiol raises major concerns. There are numerous commercial antibody-based kits on the market, which all have different specificities for the target molecules. The testosterone detection kits usually show considerable 5-30% cross-reactivity with DHT and/or androstenedione. Estradiol kits generally show 10-15% cross-reactivity with estrone and for both hormones cross-reactivity with their sulfate conjugates can be as high as 100% (although extraction of free hormone circumvents this problem). Earlier studies with H295 cells indicate that these cells produce relatively large quantities of androstenedione and 11beta-OH androstenedione (the latter not usually found in gonadal tissues or healthy adrenal cortex)(Gazdar *et al.*, 1990). Without knowing how much

androstenedione and other potentially cross-reactive steroids and metabolites are present in the cellular system it is difficult to assign any reliable value to the concentrations determined by immunoassay. A concentration of 5 pg/ml testosterone may in fact be more than 50% androstenedione, or something else, or not. These types of uncertainty need to be eliminated. Trilostane which is a steroid with an androgen base structure, not surprisingly, interferes with the testosterone immunoassay kits. There will be more steroid-like molecules (in environmental extracts and other unknowns) that will interfere with immunoassay based hormone analysis especially if one considers that these test compounds/extracts are added to the assay system in micromolar quantities whereas the endogenous hormones are present in picomolar quantities. Thus even a cross-reactivity of less than 1% would cause major interference. See also comment on trenbolone under point 7.
If immunoassay based analytical methods will be continued to be used all these pitfalls will need to be addressed.

Statistical methods:

Statistical testing hypotheses have not been explicitly defined. See also comments under point 3e.

Matt Vijayan: Little is known about the impact of most of the test substances on steroid production. The lack of response to a known inducer of sex steroid production in gonadal tissue, for instance human chorionic gonadotropin (hcG), suggests that this system has limitations because of the type of tissue involved (adrenal carcinoma). Also, I am surprised that neither arylhydrocarbon receptor ligands (for instance PCBs) nor metals were used as a test substance to validate steroid output using this model system, especially since several studies have shown that metals and PCBs inhibit steroidogenesis. It may also be worthwhile using DMSO as a test substance especially since it is being used as a solvent control.
The analytical methods and the statistical methods are appropriate to demonstrate the performance.

2.7 Comments on Repeatability and Reproducibility of the Results Obtained with the Assay, Considering the Variability Inherent in the Biological and Chemical Test Methods

Safa Moslemi: Excepting some within- and among-laboratory CVs which being highly elevated (Tables 9.3 & 9.3, pages 46 & 47), assay is generally sufficiently reproducible as demonstrated by statistical analysis and by fixing up to 30% of CV for inter- and intra- assay variations and by demonstrating conformance with the QC plates and data performance criteria outlined in chapter 7.1.2 such as basal production of T and E2 fixed at least at 2.5-times MDL, minimum induction- and inhibition- folds in the presence of forskolin and prochloraz, and cytotoxicity up to 20%. However, the low basal level of E2 production which is sometimes near to the MDL of the detection system used remains problematic. See additional comments below in Point 9.

Damian Romero: Results obtained following the protocols seem to be repeatable and reproducible. However, several suggestions are indicated under the "What additional advice, if any, can be given regarding the protocol?" heading that would probably help to further increase assay performance and reproducibility.

Thomas Sanderson: The reproducibility of the test system appears to be relatively poor. This may be partly due to the variability inherent in the use of cell lines in culture, but is also likely to be due to the various immunoassay-based hormone analysis methods used. The latter influence may be reduced by selecting a single method of detection, preferably not immunoassay based. Furthermore, the steroidogenesis assay depends on basal hormone secretion; results may be more consistent if estradiol and testosterone production are monitored after exposure of cells to an early precursor hormone in the biosynthesis pathway, such as pregnenolone. This would change the nature of the assay, but in a way would make the assay a more steroidogenesis-focused assay, as it would eliminate early variables such availability of cholesterol as precursor for the steroidogenesis of the sex hormones.

Matt Vijayan: Foe the most part the assay is sufficiently repeatable and reproducible. However, I am concerned with the high CV among laboratories and also within laboratories. The within lab CV is particularly high for prochloraz and this could be because it is inhibiting the basal steroid production. As the constitutive levels are being inhibited this may lead to error as the levels may

differ due to autoregulation that is inherent in this system. I would recommend using a test group where the inhibition is tested using acute-stimulated (forskolin or 8bromocAMP) steroid production as a model. This might reduce the variability and make the data set more comparable between the laboratories. For instance there is a large variability in EC50 for foskolin between the different labs (Table 10.3). The advantage of using a cell line is the consistency in response no matter where it is used but the high CV (ranging from 57 – 89%; Table 9.2) shown here suggests that the basal production of hormone is subjected to autoregulation leading to differences in the magnitude of response. In this regard, the basal hormone levels may be a key variable that need to be within a narrow range among laboratories prior to screening for substances modulating sex steroid production.

2.8 Comments on Whether the Appropriate Parameters were Selected and Reasonable Values were Chosen to Ensure Proper Performance of the Assay, with Respect to the Performance Criteria

Safa Moslemi: Yes, except for the allowable location of the acceptable range for SC-inducers and inhibitors which should be, I think, between the Center (mean range of hormone concentration of solvent control SC) and respectively the upper and the lower part of the linear standard curve (figure 7.2). Actually, allowable location for inducers and inhibitors should not cross and should be within 50% range (and not 75% range) of the linear range of standard curve.

Damian Romero: The performance criteria are adequate and would allow the assay to be performed at multiple laboratories without major problems.
Table 7.2 and Table I.2 from the SOP "Exposure of H295R…." should be checked for consistency since many of the parameters differ between both of them.

Thomas Sanderson: The performance criteria are outlined in Table 7.1 of the interim report. However, a performance criterion for inhibition of estradiol production by prochloraz is lacking; this needs to be addressed. Forskolin is used as a positive control for induction of testosterone and estradiol. This is a reasonable choice. It must, however, be kept in mind that forskolin increases the production of these two hormones via a very specific mechanism, by stimulating intracellular cAMP levels causing induction of various steroidogenic enzymes and ultimately increased synthesis of the sex hormones, but also of cortisol. There are however, many other

mechanisms by which testosterone and or estradiol concentrations can be affected in H295R cells. (preferential inhibition of aldosterone/cortisol synthesis, increased bioavailability of cholesterol or decreased conjugation pathways, increased membrane permeability etc.). The steroidogenesis as currently set up will not be able to distinguish between any of these mechanisms, which in itself is not the intention. But it does means that comparing an induction response by a sample/unknown to that caused by forskolin as a performance criterion may in numerous instances be comparing apples to oranges. This makes the use of the Percent Control concept (Chapter 10.3) fundamentally flawed.

The interim report mentions that forskolin may not be the best choice of inducer because its effect on testosterone production is relatively weak. This likely due to the fact that forskolin strongly induces aromatase activity, which consumes testosterone to form estradiol. A better response may be obtained if the assay is adapted to use a (pregnenolone) precursor to avoid the limitation of substrate availability to the various steroidogenic enzymes of interest.

The interim report also mentions that trenbolone is being considered as replacement for forskolin as a positive control for induction of testosterone production. The immediate question is whether trenbolone, which is a steroid with a structure very similar to testosterone, is not in fact causing cross-reactivity with the immunoassays for testosterone. Has this been ruled-out? The next question would be how trenbolone, a potent AR agonist, is able to induce testosterone levels in H295R cells? AR agonists do not normally have any effect on testosterone formation in these cells.

Matt Vijayan: The test substances chosen were appropriate to demonstrate the performance of the assay (forskolin and prochloraz as inducer and inhibitor, respectively). However, the magnitude of change is very different for testosterone and estradiol. This difference may be related to the difference in their basal secretion rate (high for T and low for E2). Consequently, changes in E2 levels may not be a good performance indicator for testing inhibitors of steroidogenesis. It may be useful to use other inducers such as cAMP analogue and 25 hydoxycholesterol to obtain stimulated steroid production levels to validate the performance assay. Also, supplementing medium with cholesterol may be required to confirm that this precursor is not a limiting factor for steroid production in this cell system given the high basal secretion for testosterone.

For testing the performance for inhibitors it may be necessary to use inhibition of stimulated-steroid production as the end point at least in the case of E2 secretion.
This cell line is derived from adrenal carcinoma and consequently would be a suitable system for detecting corticosteroid production. Hence, a stimulated (ACTH or 8bromocAMP) cortisol production may be useful as a positive control for cell system validation among laboratories to meet the QA/QC criteria.
The CV for SCs that is acceptable for QC is relatively high. I would suggest a CV≤20% as acceptable for replicate measures within a laboratory.
On Table7.1 the performance criteria for estradiol with forskolin is given as ≥15 times SC, whereas on pg. 36 it is shown as ≥10-times induction of E2 production.
Pg. 44. 9.1.1.1. line 5, change to Lab 5
Table 9.1 – change "second" Lab 4 to Lab 5.

2.9 Comments on Whether the Data Interpretation Criteria are Clear, Comprehensive, and Consistent with the Stated Purpose

Safa Moslemi: Yes, However care must be taken when extrapolating results from in vitro to in vivo effects, see additional comments in point 9.

Damian Romero: As indicated above, although the preliminary report deals extensively with data analysis and report the protocol "Exposure of H295R…." does not address this point satisfactorily since it does not have guidelines on how to interpret the data.
From the extensive preliminary report addressing several analysis techniques based on the data generated with the core chemicals, it seems that the use of "Fold change" in combination with "Percent of control" to be the most adequate way to report the results. This procedure was applied with the supplementary chemicals and in the report is shown to have worked very satisfactorily.
After reaching a consensus, the protocol "Exposure of H295R…." should include a section indicating how the results are going to be analyzed, how chemicals are going to be classified, etc.

Thomas Sanderson: For data interpretation criteria I am dependent on the information dispersed over Chapters 7.3 and 8 and 10. Using the H295R steroidogenesis assay as a semi-quantitative screening tool is a reasonable approach. The classification of inducers into weak, medium,

strong and very strong seems too elaborate. Given the large variability and uncertainties in hormone determinations and mechanisms of induction, as well as the limited meaningfulness of fluctuation in hormone levels that are less than 2-fold it would be preferable to reduce this classification to weak (2-5 fold) and strong (>5-fold) inducers, and consider anything less than 2-fold as 'possible' inducers. Expressing results using the PCmax/PC50 concept is, as mentioned under point 7, not likely to be very useful.

Matt Vijayan: The data interpretation is clear and consistent with the objective of the report. However, I am not convinced with the categorization of test substances as weak, medium, strong or very strong, because of some of the limitations of the cell system. For instance the lack of response (or weak response) may be due to the high basal hormone production in the case of testosterone or the low secretion for E2. This needs to be further tested, refined and validated for both testosterone and estradiol. Also, the dose-response curves will have to be tightened (narrower range) based on the initial screening.

2.10 Please Comment on the Overall Utility of the Assay as a Screening Tool in the EDSP Tier 1 Battery

Safa Moslemi: 1) Although H295R cell line express all steroidogenic enzymes founded in gonads and other tissues of both sexes, gene/protein expression of these enzymes depend on species, sexes, tissue, age and physiologic conditions. Therefore, extrapolation of in vitro to in vivo effects requires further investigations.

2) The sexual distinctions are not qualitative differences but rather result from quantitative divergence in hormones concentrations and differential expression of steroid hormones receptors. This results in differential sensitivity of female and male tissues in regard to steroidal hormone. Thus, when evaluating xenobiotcs on androgen and estrogens synthesis (induction and/or inhibition) using H295R cell line in vitro, the sexual sensitivity dimorphism which occur naturally in vivo should be considered in the classification of the chemicals as moderate, middle or highly endocrine disruptor. For instance, a chemical which is considered as highly endocrine disruptor for mal by inhibiting estrogen production might be classified as middle or moderate for female since mal and female have not the same sensitivity toward endogenous and therefore altered estrogen.

3) Another point which merit to be discussed is the differentiation of H295R cell line in relation to passage. Indeed, H295R cell line have the physiological characteristics of zonally undifferentiated human fetal cells, with the ability to produce the steroid hormones of each of the three phenotypically distinct zones found in the adult adrenal cortex (Gazdar AF, et al, 1990). Validated protocol should be able to answer to these questions. a) Does the number of passage affect the differentiation of these zones in different manner ? b) Has the morphology of these zones been studied at structural level after different passages ? c) Is the different in absolute production of hormones that occur as a function of cell passage due to the zones differentiation ? d) What is the relative basal amounts of each class of steroid (cholesterol, mineralocorticoids, glucocorticoids, androgens and estrogens) produced in these cells at different passages ? For example, it is important to know whether the glucocorticoid or androgen/estrogen pathway is predominant in the passage cells used in this assay. So, the suitable passage to study and evaluate each class of steroid hormone should be known.

4) A xenobiotic might present differential effect (inducers or inhibitor) on steroidogenic enzymes and therefore androgen/estrogen ratio appears more precise in this evaluation than the individual variation of each steroid. There is another reason which justifies the evaluation of androgen/estrogen ratio. Actually, H295R assay showed
its limit to detect decreases in E2 production after exposure to an inhibitor. Indeed, E2 production is already faint in this model as reported in table 9.1, page 44 by all laboratories participated and this renders difficult the classification of chemicals regarding their effect on E2. Thus, variation of androgen/estrogen ratio should better reflect chemical effects on steroidogenesis of sexual hormone production in H295R cell line and which might be further extrapolated to the variation of androgen/estrogen ratio in healthy and exposed men and women in order to evaluate xenobiotics as endocrine disruptors.

5) In this assay, T and E2 variation was evaluated at basal level and did not include the addition of a specific upstream precursor such as progesterone and/or dehydroepiandrosterone which could induce enzymatic activities involved in the T and E2 production. So, xenobiotic effect may be different from the case in which an inducer or a steroid precursor being added. For instance, 2378-tetrachlorodibenzo-p-dioxin (TCDD) had no effect on basal aromatase or cholesterol side-chain cleavage activity, but did reduce the inducibility of both activities by 8-bromo-cyclic AMP in H295R (Sanderson & Van den Berg, 1998; Sanderson et al, 2001). In H295R assay, addition of a precursor seems necessary for E2 but not for T production since the

basal production of the latter was often high in H295R cell line. This is confirmed in final report "Development of an assay using the H295R cell Line to…" where the exposure to 100 □M progesterone cause a significant elevation of E2 production in culture medium when compared to SC (page 32, table 6.2). Supplementation of progesterone could resolve the background problem encountered with E2 evaluation in H295R assay. Another solution will be to evaluate total E2 (free and conjugated) since a conjugation of E2 via an estrogen sulfotransferase is not excluded in H295R cell line as evoked in Draft report "Standardization and refinement of the H295R cell …" page 47 section 10.3.

Damian Romero: The H295R steroidogenic assay would be an invaluable tool that would complement other assays of the Tier 1 battery. The assay has multiple advantages including a relative easy to perform, inexpensive an reproducible in vitro screening tool, that do not rely on live animals or animal tissues that may allow the screening of multiple compounds in a relative short period. The assay would identify chemicals with endocrine disruptor characteristics that could be further evaluated with other assays of the Tier 1 battery.

Thomas Sanderson: The EPA has been charged with the development and implementation of a screening tool for interferences with steroidogenesis. Given this task, the EPA has decided to develop a tool that will evaluate effects on gonadal sex hormone production, specifically testosterone and estradiol. To do this the H295R human adrenocortical carcinoma cell line was chosen as *in vitro* biological system. The H295R steroidogenesis assay is intended to detect chemicals and/or complex environmental samples extracts that may interfere with basal secretion of testosterone and estradiol.

As a system to study effects of chemicals on steroidogenesis the H295R cell line has great potential as it is capable of producing mineralocorticoids, glucocorticoids, androgens and estrogens. The steroid biosynthesis pathway is highly complex and also highly interconnected. This limits the usefulness of only evaluating effects on one or two specific hormones as there are a large number of influences unrelated to the steroidogenesis of those hormones that may cause small fluctuations in their secretion by the cell system. The system, as designed, is not really a steroidogenesis assay although is may pick up inhibitors and inducers of testosterone and estradiol synthesis. However, because the system, as designed, ignores all the other steroid hormones, including other active androgens and estrogens known to be produced in these cells, a

bigger picture is not obtained, limiting the interpretability of any observed alterations in solely levels of testosterone and estradiol. The H295R system also in no way reflects the conditions of a gonadal system in which mineralo- and glucocorticoid pathways do not play a role. The regulation of the various steroidogenic enzymes will also be different in different tissues, again limiting the interpretability of any observed effects on induction of testosterone or estradiol secretion (if steroidogenesis related).

For the EPA to have a true steroidogenesis assay (a system that detect the ability of chemicals to interfere with the biosynthesis of steroid hormones) the H295R cell line could provide a very useful model with some alterations to the design: (1) H295R cells would be analyzed for 4 key steroid hormones (aldosterone, cortisol, estradiol and testosterone) using a single analytical technique such as LC-MS, (2) pregnenolone would be used as precursor for all steroids (3) effects on the relative production of the 4 hormones would be relatively easy to interpret as they would provide clues on which of these 4 essential steroid hormones and which steps of the steroidogenic pathway are affected.

The way the H295R cell system is being proposed to be used is like a black box. It will be difficult to interpret the meaning of any outcomes that may be observed on testosterone and estradiol levels, and this is further compounded by the drawbacks of using immunoassay-based detection methods. A more focused definition of the purpose of a tier 1 assay for steroidogenesis would be recommendable; allowing for the development of a H295R cell-based steroidogenesis assay that would provide less ambiguous information about the steroidogenesis disruption potential of chemicals or unknown environmental extracts.

Matt Vijayan: The H295R steroidogenic assay has been validated for its steroid production capacity and as a tool for screening substances that modulate sex steroid production. The multi-laboratory validation suggests that the assay has potential as a screening tool for sex steroid disruptors. However, the assay has limitations and some of them are related to the cell system itself. For instance the high basal unstimulated sex steroid production is not physiologically relevant but provides a model for testing the capacity for substances to induce or inhibit steroidogenesis. This assay focuses only on the signaling pathway downstream of trophic hormone stimulation. While the mechanism for the high basal testosterone output is unclear, it remains to be seen if that would modify the steroid production capacity in response to stimulators or inhibitors. Also, the low basal E2 production seen with these cells does not

provide an ideal model to test inhibitors of steroidogenesis. This can be easily tested by examining the capacity of the modulators to inhibit forskolin- or 8bromo-cAMP-stimulated E2 production. The huge CV reported for between laboratory comparisons may have to do with the difference in basal hormone production and associated differences in the magnitude of response to know inducers and inhibitors as well as test substances. Overall, the assay has the potential to be a screening tool for steroidogenesis but requires further testing and refinement.

2.11 Additional Comments and Materials Submitted

Safa Moslemi:

References

Sanderson JT, Slobbe L, Lansbergen GD, Safe S, Van den Berg M. Toxicol. Sci. 2001, 61, 41-48.

Sanderson JT and Van den Berg. Organohalogen Comp. 1998, 37, 77-80.

Letcher RJ, van Holsteijn I, Drenth HJ, Norstrom RJ, Bergman A, Safe S, Raymond Pieters R and van den Berg M.. Toxicol. Appl. Pharmacol. 1999, 160: 10-20.

Fechner S, Husen B, Thole H, Schmidt M, Gashaw I, Kimming R, Winterhager E, GrümmerR, Fertil Steril, 2007, 88, 1029-1038.

Murakami K, Nomura K, Shinohora K, Kasai T, Shozu M, Inoue M. Fertil Steril, 2006, 86, 291-297.

Thon V, Härle P, Schölmerich J, Kuklinek P, Lokaj J, Straub RH. Allergy, 2007, 62, 1320-1325.

Richard S, Moslemi S, Sipahutar H, Benachour N, Séralini GE. Environ Health Perspect, 2005, 113, 716-720.

Benachour N, Sipahutar H, Moslemi S, Gasnier C, Travert C, Séralini GE. Arch. Environ. Contam. Toxicol. 2007, 53, 126-133.

Editorial comments :

General comments:

The final report should be re-written and experimental protocols (cell culture, treatment with chemicals, steroid and results analysis …..) should be included in detail.

Specific comments:

1) Page 4, Table of tables, Table 3.1: "Time lines" instead of "time line". Page 5, add legends of Table 9.3 and Table 9.4 in the list, Table 9.3 became table 9.5.

2) Add in "Abbreviations" page 13 : EDSP, EDSTAC, EPA, FFDCA, QA, EDS, HCG, QC, SOP, SC, CV, OD, MDL…

3) Page 19, line 2, reference Hecker et al, 2006 a or b?,

4) Page 20, in section " 3.3 the H295R steroidogenesis assay", line 3, reference Hecker et al, 2006 a or b?)

5) Page 21, in section " 3.4 pre-validation studies overview", line 5, reference Hecker et al, 2007a or b? Line 9, (Figures 3.3, 3.4, 3.5) instead of (Figure 3).

6) Page 25, add point at the end of last paragraph.

7) Page 28, delete (KFDA) from National Institute of Toxicological Research.

8) Why the number referring to the code number for 18 supplementary chemicals tested by 6 laboratories did not appear in table 6.1? I think that it is preferable, to avoid confusion, to add directly the name of each chemical tested or to be tested by each laboratory in table 6.1.

9) In table 6.2, there is a confusion between the "Mode of action" and the "Effect type" of each chemical; should be precised "Effect type on T and/or E2 production"? Correct hCG instead of hcG for Human chorionic gonadotropin, Pharmaceutical instead of pharmacytical for "product class" of Letrozole and put "Spermatide" instead of "unknown" for the "mode of action" of Nonoxynol-9.

10) Table 6.3, add horizontal line after 2,4-Dinotrophanol, What is the "product class" of Cyanoketone? In "product class" of finasteride, delete r from prostrate (prostate is correct). In "effect type" of vinclozoline, add E2 after "weak inducer of …..".

11) Page 35, table 7.1, if "n/a" is meaning "non analysed" this should be corrected because this is down in the next page (page 36), see "4. Inhibitor: Prochloraz (3 μM): ……

12) Page 44, in section " 9.1.1.1 Basal hormone Production (Blanks)", line 5, Lab 6 instead of Lab 5. In the table 9.1, change the second Lab 4 into Lab 5.

13) Page 46, Figure 9.1, use other patterns for different histograms to distinguish the laboratories.

14) Page 53, in the legend of figure 9.5, “Lab 5 (instead of Lab 6): No data available for aminogluthetimide, atrazine, (add Benomyl here), forskolin and hCG”.

15) Page 54, in the legend of figure 9.6, “Lab 5 (instead of lab 6) : No data available…..”

16) Page 55, in section “9.2.2.1 Testosterone”, line 10, “(10 μM) instead of (100 μM) at all laboratories with the exception of Lab 6 (instead of Lab 4)”. In figure 9.7 add the results of Paraben and delete it from figure 9.9 since this chemical induces testosterone. In the legend of the figure 9.7, “Lab 5 (instead of Lab 6) : No data available …”

17) Page 56, in the legend of figure 9.8, “Lab 5 (instead of Lab 6): No data available ….”

18) Page 57, Figure 9.9, change patterns for Lab 5 and Lab 6. In the legend, molinate instead of molinat. “Lab 5 instead of Lab 6: No data available …”

19) In the legends of figures 9.10, 9.11, 9.12, “Lab 5 instead of Lab 6 : No data available.”

20) Page 61, Table 9.5 instead of 9.3.

21) Page 63, in section “10.1 Fold-Change Evaluation”, ….”the effect evoked” instead of “effect the evoked”. (Table 10.1).

22) Page 65, Table 10.2, add unity (μM) after “Lowest observed effect concentrations”, change “Nonythenol” to Nonoxynol-9.

23) Page 68, line 7, …..”while chemical (delete “the”and add “exposure that resulted in an”) elevated hormone production”…….Actually, one should reed …..”while chemical exposure that resulted in an elevated “……

24) Page 70, Table 10.4, add the concentrations used for forskolin (10 μM) and prochloraz (3 μM).

25) Pages 71-74, Figure 10.1-10.4, there are confusion between chemicals, controls and production of T and E2 related to inducer (Forskolin) and to inhibitor (prochloraz). I suggest to eliminate testosterone and estradiol (in the parentheses) and replace the Y axis by % of testosterone and /or estradiol.

26) Page 77, Figure 11.1, change Danazole to Danazol.

27) Page 78, Figure 11.2, in X axis, add unity for the concentrations used.

28) Page 84, second paragraph, line 9, in vitro instead of vitro. At the end of paragraph ….”would provide a more detail in the dose-response relationships”.

29) Add in References page 87 : EPA 1997, EPA 1998, EDSTAC 1998,EDMVAC 2005. Add title for references when needed.

Thomas Sanderson:

References

Bulun, S. E., Sebastian, S., Takayama, K., Suzuki, T., Sasano, H., and Shozu, M. (2003). The human CYP19 (aromatase P450) gene: update on physiologic roles and genomic organization of promoters. *J Steroid Biochem Mol Biol* **86**, 219-224.

Gazdar, A. F., Oie, H. K., Shackleton, C. H., Chen, T. R., Triche, T. J., Myers, C. E., and Chrousos, G. P. (1990). Establishment and characterization of a human adrenocortical carcinoma cell line that expresses multiple pathways of steroid biosynthesis. *Cancer Res.* **50**, 5488-5496.

Sanderson, J., and van den Berg, M. (2003). Interactions of xenobiotics with the steroid hormone biosynthesis pathway. *Pure Appl Chem* **75**, 1957-1971.

Sanderson, J. T. (2006). The steroid hormone biosynthesis pathway as a target for endocrine-disrupting chemicals. *Toxicol Sci* **94**, 3-21.

Simpson, E. R. (2003). Sources of estrogen and their importance. *J Steroid Biochem Mol Biol* **86**, 225-230.

Zar, J. H. (1999). *Biostatistical Analysis*. Prentice-Hall, N.J.

3.0 PEER REVIEW COMMENTS ORGANIZED BY REVIEWER

Peer review comments received for the steroidogenesis assay are presented in the sub-sections below and are organized by reviewer. Peer review comments are presented in full, **unedited** text as received from each reviewer.

3.1 Safa Moslemi Review Comments

Comments :

This report « Multi-Laboratory validation of the H295R Steroidogenesis Assay to identify modulators of testosterone and estradiol production » aimed to develop and standardize of the assay as a screen of steroidogenesis using H295R cell line to identify chemicals that act as endocrine disruptors. You find here after the answers and comments to the charge questions.

Charge Questions:

1. Is the stated purpose of the assay clear?

Yes. The steroidogenic screen assay consist to detect any natural and human-made substance that would disrupt endogenous estrogens and/or androgens production. In this way, the assay will complement the other Tier 1 assays which aim to identify xenobiotics that could be classified as endocrine disruptors of both human and wildlife. The use of H295R cell line present several advantages making this model unique when compared with other models. Besides its availability, this model allows the detection of both increases and decreases in the production of testosterone (T) and estradiol (E2) in the presence of chemicals, and to follow the direct potential impact of a chemical on cell viability/cytotoxicity. Furthermore, H295R cells express a wide range of steroidogenic enzymes found in the adult adrenal cortex and the gonads, including those required to produce, cholesterol, mineralocorticoids, glucocorticoids, androgens and estrogens. Thus, this cell line enable the research of any target site within the steroidogenic pathway downstream of cholesterol in addition to T and E2 investigated in this work.

2. Is the assay biologically and toxicologically relevant to the stated purpose?

Actually, many of chemicals (especially xenoestrogens) have the potential to disrupt endocrine processes at tow major levels; first at sex hormone receptors, particularly the estrogen receptor, and second at steroidogenic enzymes involved in both steroid synthesis and metabolism. The second level being more appropriate target since most of the environmental chemicals, when introduced in the organism, are present at low concentrations and show relatively, when compared with estradiol, low affinity for estrogen receptors. Therefore, estrogen receptor pathway could not be considered as an ideal endpoint to study endocrine disruption of xenobiotics. Utility of H295R cells has been well established as an unique model for study of steroidogenic pathways but also to test and evaluation of xenobiotics since these cells express genes that encode for all enzymes of steroidogeneis especially those involved in sexual mal and female hormones synthesis; androgens and estrogens. Furthermore, this model permits to evaluate, in the same cells, the potential cytotoxicity of chemicals allowing the discrimination between effects that are due to cytotoxicity or due to the direct interaction of chemicals with steroidogenic enzymes.

3. Does the protocol describe the methodology of the assay in a clear, and concise manner so that the laboratory can:

a) comprehend the objective;

b) conduct the assay;

c) observe and measure prescribed endpoints;

d) compile and prepare data for statistical analyses; and

e) report the results?

Protocol is well described and the methodology generally presented in a comprehensible manner allowing the reader to fallow all steps cited above.

What additional advice, if any, can be given regarding the protocol?

In order to improve protocol, the following advices are proposed :

1) The choice of solvent for steroid extraction should be précised (ethyl ether or dichloromethane).

2) During extraction procedure of steroids with ethyl ether a rapid freezing of aqueous phase (after step 9 and before step 10) facilitate the separation of aqueous (inferior) and organic (supernatant) phases (see Appendix ii page 17).
3) Collected solvent phases could be washed by distilled water to eliminate hydrophilic contaminants and to reduce background of the detection by RIA or ELISA.
4) Hormone purification, at least for protocol validation and before using a detection system based on antibody (ELISA, RIA), being necessary to avoid cross reaction observed with chemicals especially with trilostane and to reduce background of assay.
5) What is the maximum passage of cell culture to be respected; 10 passages (Appendix I page 7) or 7 passages (Appendix ii page 7) ?
6) When possible, replace methanol (highly toxic) by ethanol for cytotoxicity analysis (see page 36, point 5 & table)
7) There is a confusion between E2 decrease in the presence of 3 μM prochloraz when determining performance criteria for each laboratory. In page 35, table 7.1, the average change in hormone production relative to the solvent control (SC=1) was not reported (n/a, please spell this acronym) whereas in page 36, point 4, a change of 50% in E2 reduction was reported!
8) Also, there is a difference in the reported induction of E2 by 10 μM forskolin ; ≥10-times induction of E2 in page 36 point 3, different from ≥15-times induction in table 7.1, page 35!
9) In page 44, table 9.1, the Minimum Detectable Level (MDL) and the measuring system used (RIA, ELISA or others) of each laboratory should be reported.

4. Have the strengths and/or limitations of the assay been adequately addressed?

The advantages and disadvantages of H295R regarding to other cell lines should be detailed especially to JEG-3 and JAR placental choriocarcinoma cell lines. For instance, JEG-3 and JAR placental choriocarcinoma cell lines appear relatively more sensitive to cytotoxic effects of chemicals than H295R cell line (Letcher et al, 1999). This rises the question about the suitable model (more sensitive or less sensitive to cytotoxicity) to screen chemicals for their endocrine disruption effect since the endocrine disruption of chemicals is tested at non cytotoxic concentrations and this might affect interpretation of

results, chemical classification and determination of their tolerability concentrations (threshold) in organism.

5. Were the (a) test substances, (b) analytical methods, and (c) statistical methods chosen appropriate to demonstrate the performance of the assay?

Yes, there is in general a good choice of different chemicals, analytical and statistical methods. However, information concerning the effect type on T and E2 production should be updated for some chemicals (danazol, finasteride, flutamide, Glyphosate, RU-486/mifepristone, spironolactone, taxol etc in table 6.3, pages 33 & 34). For instance, danazol is known to : inhibit aromatase transcription in ectopic human endometrial tissue (Fechner et al, 2007), inhibit aromatase activity of endometriosis-derived stromal cells (Murakami et al, 2006), induce a marked up-regulation of free T and down stream 17β-E2 in hereditary angioedema (Thon et al, 2007). Glyphosate (Roundup) showed also to inhibit aromatase in vitro (Richard et al, 2005; Benachour et al, 2007). Moreover, protocol and analytical method should be revised for E2 evaluation in the presence of inhibitor chemicals since production of E2 was not evident during validation assay in H295R cells after 5 passage (see additional comments in point 9).

6. Considering the variability inherent in biological and chemical test methods, were the results obtained with this assay sufficiently repeatable and reproducible?

Excepting some within- and among-laboratory CVs which being highly elevated (Tables 9.3 & 9.3, pages 46 & 47), assay is generally sufficiently reproducible as demonstrated by statistical analysis and by fixing up to 30% of CV for inter- and intra- assay variations and by demonstrating conformance with the QC plates and data performance criteria outlined in chapter 7.1.2 such as basal production of T and E2 fixed at least at 2.5-times MDL, minimum induction- and inhibition- folds in the presence of forskolin and prochloraz, and cytotoxicity up to 20%. However, the low basal level of E2 production which is sometimes near to the MDL of the detection system used remains problematic. See additional comments below in Point 9.

7. With respect to performance criteria, were appropriate parameters selected and reasonable values chosen to ensure proper performance of the assay?

Yes, except for the allowable location of the acceptable range for SC-inducers and inhibitors which should be, I think, between the Center (mean range of hormone concentration of solvent control SC) and respectively the upper and the lower part of the linear standard curve (figure 7.2). Actually, allowable location for inducers and inhibitors should not cross and should be within 50% range (and not 75% range) of the linear range of standard curve.

8. Are the data interpretation criteria clear, comprehensive, and consistent with the stated purpose?

Yes, However care must be taken when extrapolating results from in vitro to in vivo effects, see additional comments in point 9.

9. Please comment on the overall utility of the assay as a screening tool described in the introduction of the ISR to be used by the EPA to identify chemicals that have the potential to interact with the endocrine system.

1) Although H295R cell line express all steroidogenic enzymes founded in gonads and other tissues of both sexes, gene/protein expression of these enzymes depend on species, sexes, tissue, age and physiologic conditions. Therefore, extrapolation of in vitro to in vivo effects requires further investigations.

2) The sexual distinctions are not qualitative differences but rather result from quantitative divergence in hormones concentrations and differential expression of steroid hormones receptors. This results in differential sensitivity of female and male tissues in regard to steroidal hormone. Thus, when evaluating xenobiotcs on androgen and estrogens synthesis (induction and/or inhibition) using H295R cell line in vitro, the sexual sensitivity dimorphism which occur naturally in vivo should be considered in the classification of the chemicals as moderate, middle or highly endocrine disruptor. For instance, a chemical which is considered as highly endocrine disruptor for mal by inhibiting estrogen production might be classified as middle or moderate for female since mal and female have not the same sensitivity toward endogenous and therefore altered estrogen.

3) Another point which merit to be discussed is the differentiation of H295R cell line in relation to passage. Indeed, H295R cell line have the physiological characteristics of zonally undifferentiated human fetal cells, with the ability to produce the steroid hormones of each of the three phenotypically distinct zones found in the adult adrenal cortex (Gazdar AF, et al, 1990). Validated protocol should be able to answer to these questions. a) Does the number of passage affect the differentiation of these zones in different manner ? b) Has the morphology of these zones been studied at structural level after different passages ? c) Is the different in absolute production of hormones that occur as a function of cell passage due to the zones differentiation ? d) What is the relative basal amounts of each class of steroid (cholesterol, mineralocorticoids, glucocorticoids, androgens and estrogens) produced in these cells at different passages ? For example, it is important to know whether the glucocorticoid or androgen/estrogen pathway is predominant in the passage cells used in this assay. So, the suitable passage to study and evaluate each class of steroid hormone should be known.

4) A xenobiotic might present differential effect (inducers or inhibitor) on steroidogenic enzymes and therefore androgen/estrogen ratio appears more precise in this evaluation than the individual variation of each steroid. There is another reason which justifies the evaluation of androgen/estrogen ratio. Actually, H295R assay showed
its limit to detect decreases in E2 production after exposure to an inhibitor. Indeed, E2 production is already faint in this model as reported in table 9.1, page 44 by all laboratories participated and this renders difficult the classification of chemicals regarding their effect on E2. Thus, variation of androgen/estrogen ratio should better reflect chemical effects on steroidogenesis of sexual hormone production in H295R cell line and which might be further extrapolated to the variation of androgen/estrogen ratio in healthy and exposed men and women in order to evaluate xenobiotics as endocrine disruptors.

5) In this assay, T and E2 variation was evaluated at basal level and did not include the addition of a specific upstream precursor such as progesterone and/or

dehydroepiandrosterone which could induce enzymatic activities involved in the T and E2 production. So, xenobiotic effect may be different from the case in which an inducer or a steroid precursor being added. For instance, 2378-tetrachlorodibenzo-*p*-dioxin (TCDD) had no effect on basal aromatase or cholesterol side-chain cleavage activity, but did reduce the inducibility of both activities by 8-bromo-cyclic AMP in H295R (Sanderson & Van den Berg, 1998; Sanderson et al, 2001). In H295R assay, addition of a precursor seems necessary for E2 but not for T production since the basal production of the latter was often high in H295R cell line. This is confirmed in final report "Development of an assay using the H295R cell Line to…" where the exposure to 100 µM progesterone cause a significant elevation of E2 production in culture medium when compared to SC (page 32, table 6.2). Supplementation of progesterone could resolve the background problem encountered with E2 evaluation in H295R assay. Another solution will be to evaluate total E2 (free and conjugated) since a conjugation of E2 via an estrogen sulfotransferase is not excluded in H295R cell line as evoked in Draft report "Standardization and refinement of the H295R cell …" page 47 section 10.3.

References

Sanderson JT, Slobbe L, Lansbergen GD, Safe S, Van den Berg M. Toxicol. Sci. 2001, 61, 41-48.

Sanderson JT and Van den Berg. Organohalogen Comp. 1998, 37, 77-80.

Letcher RJ, van Holsteijn I, Drenth HJ, Norstrom RJ, Bergman A, Safe S, Raymond Pieters R and van den Berg M.. Toxicol. Appl. Pharmacol. 1999, 160: 10-20.

Fechner S, Husen B, Thole H, Schmidt M, Gashaw I, Kimming R, Winterhager E, GrümmerR, Fertil Steril, 2007, 88, 1029-1038.

Murakami K, Nomura K, Shinohora K, Kasai T, Shozu M, Inoue M. Fertil Steril, 2006, 86, 291-297.

Thon V, Härle P, Schölmerich J, Kuklinek P, Lokaj J, Straub RH. Allergy, 2007, 62, 1320-1325.

Richard S, Moslemi S, Sipahutar H, Benachour N, Séralini GE. Environ Health Perspect, 2005, 113, 716-720.

Benachour N, Sipahutar H, Moslemi S, Gasnier C, Travert C, Séralini GE. Arch. Environ. Contam. Toxicol. 2007, 53, 126-133.

Editorial comments :

General comments:

The final report should be re-written and experimental protocols (cell culture, treatment with chemicals, steroid and results analysis …..) should be included in detail.

Specific comments:

1) Page 4, Table of tables, Table 3.1: "Time lines" instead of "time line". Page 5, add legends of Table 9.3 and Table 9.4 in the list, Table 9.3 became table 9.5.

2) Add in "Abbreviations" page 13 : EDSP, EDSTAC, EPA, FFDCA, QA, EDS, HCG, QC, SOP, SC, CV, OD, MDL…

3) Page 19, line 2, reference Hecker et al, 2006 a or b?,

4) Page 20, in section " 3.3 the H295R steroidogenesis assay", line 3, reference Hecker et al, 2006 a or b?)

5) Page 21, in section " 3.4 pre-validation studies overview", line 5, reference Hecker et al, 2007a or b? Line 9, (Figures 3.3, 3.4, 3.5) instead of (Figure 3).

6) Page 25, add point at the end of last paragraph.

7) Page 28, delete (KFDA) from National Institute of Toxicological Research.

8) Why the number referring to the code number for 18 supplementary chemicals tested by 6 laboratories did not appear in table 6.1? I think that it is preferable, to avoid confusion, to add directly the name of each chemical tested or to be tested by each laboratory in table 6.1.

9) In table 6.2, there is a confusion between the "Mode of action" and the "Effect type" of each chemical; should be precised "Effect type on T and/or E2 production"? Correct hCG instead of hcG for Human chorionic gonadotropin, Pharmaceutical instead of pharmacytical for "product class" of Letrozole and put "Spermatide" instead of "unknown" for the "mode of action" of Nonoxynol-9.

10) Table 6.3, add horizontal line after 2,4-Dinotrophanol, What is the "product class" of Cyanoketone? In "product class" of finasteride, delete r from prostrate (prostate is correct). In "effect type" of vinclozoline, add E2 after "weak inducer of …..".

11) Page 35, table 7.1, if "n/a" is meaning "non analysed" this should be corrected because this is down in the next page (page 36), see "4. Inhibitor: Prochloraz (3 μM): ……

12) Page 44, in section " 9.1.1.1 Basal hormone Production (Blanks)", line 5, Lab 6 instead of Lab 5. In the table 9.1, change the second Lab 4 into Lab 5.

13) Page 46, Figure 9.1, use other patterns for different histograms to distinguish the laboratories.

14) Page 53, in the legend of figure 9.5, "Lab 5 (instead of Lab 6): No data available for aminogluthetimide, atrazine, (add Benomyl here), forskolin and hCG".

15) Page 54, in the legend of figure 9.6, "Lab 5 (instead of lab 6) : No data available….."

16) Page 55, in section "9.2.2.1 Testosterone", line 10, "(10 μM) instead of (100 μM) at all laboratories with the exception of Lab 6 (instead of Lab 4)". In figure 9.7 add the results of Paraben and delete it from figure 9.9 since this chemical induces testosterone. In the legend of the figure 9.7, "Lab 5 (instead of Lab 6) : No data available …"

17) Page 56, in the legend of figure 9.8, "Lab 5 (instead of Lab 6): No data available …."

18) Page 57, Figure 9.9, change patterns for Lab 5 and Lab 6. In the legend, molinate instead of molinat. "Lab 5 instead of Lab 6: No data available …"

19) In the legends of figures 9.10, 9.11, 9.12, "Lab 5 instead of Lab 6 : No data available."

20) Page 61, Table 9.5 instead of 9.3.

21) Page 63, in section "10.1 Fold-Change Evaluation", …."the effect evoked" instead of "effect the evoked". (Table 10.1).

22) Page 65, Table 10.2, add unity (μM) after "Lowest observed effect concentrations", change "Nonythenol" to Nonoxynol-9.

23) Page 68, line 7, ….."while chemical (delete "the"and add "exposure that resulted in an") elevated hormone production"…….Actually, one should reed …."while chemical exposure that resulted in an elevated "……

24) Page 70, Table 10.4, add the concentrations used for forskolin (10 μM) and prochloraz (3 μM).

25) Pages 71-74, Figure 10.1-10.4, there are confusion between chemicals, controls and production of T and E2 related to inducer (Forskolin) and to inhibitor (prochloraz). I suggest to eliminate testosterone and estradiol (in the parentheses) and replace the Y axis by % of testosterone and /or estradiol.

26) Page 77, Figure 11.1, change Danazole to Danazol.

27) Page 78, Figure 11.2, in X axis, add unity for the concentrations used.

28) Page 84, second paragraph, line 9, in vitro instead of vitro. At the end of paragraph …."would provide a more detail in the dose-response relationships".

29) Add in References page 87 : EPA 1997, EPA 1998, EDSTAC 1998,EDMVAC 2005. Add title for references when needed.

3.2 Damian Romero Review Comments

H295R CELL-BASED ASSAY FOR STEROIDOGENESIS

CHARGE QUESTIONS

1. Is the stated purpose of the assay clear?

The stated purpose of the assay is "The steroidogenic assay is intended to identify xenobiotics that have as their target site(s) the endogenous components that comprise the intracellular biochemical pathway beginning with the sequence of reactions occurring after the receptor, up through and including the production of the terminal steroid hormones, i.e. testosterone (males) and estradiol/estrone (females)" (page 16). Although it is clear in general terms, it could be improved in the following structures:

1. "…the endogenous components…" it could be replaced with "intracellular components", since now the assay is performed using the H295R cell line, in contrast to the original design using rodent sliced testes which perhaps required a broader description.
2. "… occurring after the receptor…" although it is probably referring to LH/FSH receptors it could be confused with androgen/estrogen receptors, it could be clarified to avoid confusion with the later ones.
3. "…terminal steroid hormones…" may be replaced with "…terminal sex steroid hormones" to avoid confusion with other steroid hormones, especially since the cell line used in the assay also secretes gluco- and mineralo-corticoids.

2. Is the assay biologically and toxicologically relevant to the stated purpose?

The H295R steroidogenesis assay is biologically and toxicologically relevant to the stated purpose. The assay would fit perfectly in the Tier 1battery of assays to screen for endocrine disruptors. The assay has a series of strengths that would make it an excellent screening tool for endocrine disruptors of sex steroid hormone synthesis. However, results obtained with this test should always be interpreted along with the results obtained with all the other assays of the Tier 1 battery.

It is important to stress that chemicals that generate a negative result in the H295R steroidogenic assay could be false negatives and they should not be considered safe without a complete evaluation of them with the other Tier 1 battery assays. This *in vitro* system lacks that ability to study complex interactions that could occur *in vivo* such as metabolism of tested compounds, biodistribution, interaction with other endocrine systems that may modulate sex hormones steroidogenesis, etc.

3. Does the protocol describe the methodology of the assay in a clear, and concise manner so that the laboratory can:

a) comprehend the objective;

The objective of the assay is clearly stated.

b) conduct the assay;

Both Standard Operating Procedures (SOP), "Culturing of the H295R human adrenocortical carcinoma cell line" (SOP#1) and "Exposure of the H295R human adrenocortical carcinoma cells…" (SOP#2), are clear and allow the operator to conduct the assay. However there are some points that need clarification and/or need to be improved as indicated below:

13. SOP#1, page 5, item 2.2: it says "Do not freeze cells upon arrival…", unless cells arrive to the lab growing (which is fairly uncommon for ATCC cultures) there should be not much difference in keeping them for a short period of time in liquid nitrogen. In other case, an appropriate reason should be stated, since the requirement to immediately begin to culture the cells may create a burden to the lab that may be not necessary. Furthermore, in SOP#1, page 13, item 5.1.2 it is indicated to remove the vial of cells from the liquid nitrogen storage. In any case, it would be perhaps also useful to stress that cells should always be stored in liquid nitrogen to avoid any confusion.
14. SOP#2, page 6, item 2, the examples about the nomenclature of the cultures should be checked. It seems that numbers 5.4, 5.7 and 5.2 should be 4.5, 7.5 and 2.5 following the example in Appendix II of SOP#1. Idem in item 5.1 and 5.2 on pages 8 and 9 of SOP#2.
15. SOP#2, page 10, item 5.2.1, in the Reagents section it is stated "… for 6 and 11 generations…", since the term "generation" is not clearly defined in the text, and not used anywhere else, it would be better to maintain consistency to use and refer to "passages" throughout the text.

16. SOP#2, page 10, item 5.2.5, it says "General rule: use 1 petri dish…", since the statements is giving a quantitative recommendation is very important to indicate the size (probably a 100 mm diameter) of the cell culture dish. Also, for consistency, refer to "petri dish" as "cell culture dish" since "petri dish" is not used anywhere else in the protocol.
17. SOP#2, page 12, item 5.3.2, when the protocol refers to add the media and chemical compounds for testing it is not very clear. From the text and Table 3, it seems that it is recommended to add 1 ml of media per well and then add 1 µl of stock chemical solution to the well. This procedure could be a great source of error if it is performed in that way. In these cases, it is greatly preferred to make a "master mix" (i.e. 5 ml media plus 5 µl of the tested chemical stock solution) and then dispense 1 ml of media plus tested chemical or solvent per well. It would be very useful to clearly indicate that this is the preferred method to add the tested chemicals to the wells of the cell culture plate.
18. SOP#2, page 14, item 6.1.1, in the "Equipment" section it is indicated the use of a "Fluoroskan Ascent Fluorometric Microtiter Plate Reader", it would be better to indicate that the protocol have been extensively tested and validated with that particular piece of equipment but that any fluorometer microplate reader that have the adequate filters may be used.
19. SOP#2, page 14, item 6.1.1, in the "Materials" sections it says 200,000 cells per well and it should be 300,000 cells per well to be consistent throughout the protocol.
20. SOP#2, page 15, item 6.1.2, it should be useful to indicate in 6.1.1 "Materials" section the brand and catalog number of the plate sealers to be used.
21. SOP#2, page 16 and subsequent, item 7 and subitems, it is recommended to spike the sample with ^{3}H-testosterone for recovery calculation. From the text it is not clear if the same protocol should be used for estradiol extraction and recovery calculation. Should the same solvent be used for estradiol extraction? Should ^{3}H-estradiol be used for estradiol recovery calculation?
22. SOP#2, page 16 and subsequent, item 7 and subitems, anhydrous ether is the solvent recommended for steroid extraction. The use of ether is a serious hazard concern since it is highly flammable having a flammability rating of 4, the highest possible. Although in SOP#2, section 3, it is stated that "Special safety requirements need to be considered when working with ether …" a more serious advice should be given since this solvent is

an extremely serious hazard known to have caused multiple laboratory accidents. It is highly recommended to explore the use of other solvents that are not as hazardous as ether that would make the protocol safer and easier to perform.

23. SOP#2, page 17, item 7.1.2, in point "5" it says "add 10 µl of the ^{3}H-labeled hormone" , since this point is in the middle of the extraction procedure it would be useful to indicate that this tube would be used to calculate the CPMs of the "CPM spike tube".
24. SOP#2, page 19, item 7.2, the formula for the "final hormone concentration" should be updated to include the volume of reconstituted sample used in the assay and the necessary corrections to refer to the final volume of media of 1 ml. The formula should be:

$$Final\ hormone\ concentration = 150\ pg * 0.89 * \frac{1000\ \mu l}{450\ \mu l} * \frac{250\ \mu l}{vol\ used\ in\ assay\ (\mu l)} * 10$$

Using the same example as in SOP#2 with the following additions:
1000 µl refers to the final volume of media in the well
Vol used in assay (µl) is the volume in microliters used in the hormone concentration assay

The formula needs to be checked since the concentration of secreted steroid hormones is one of the quality criteria of the assay.

General considerations for cell culture that could be stressed through the SOP:

5. Perform all operations with cells in a GENTLY manner.
6. ALWAYS remove media/reagents from the well border.
7. ALWAYS add media/reagents resting the pipette against the well wall.
8. NEVER vortex or vigorously shake cells.

c) observe and measure prescribed endpoints;

The prescribed endpoints can be easily measured following the protocol. One strength of the protocol that will allow its widespread use is the possibility to use any testosterone/estradiol detection method already in use in the laboratory if it reaches the quality controls specified in the protocol.

As explained below the routine inclusion of controls to test for each chemical positively or negatively affecting the steroid quantification should be advisable to generate more confidence in the assay performance.

d) compile and prepare data for statistical analyses; and

The worksheet design is adequate for data compilation and statistical analysis submission.

e) report the results?

Although the preliminary report deals extensively with data analysis and report the protocol "Exposure of H295R…." does not address this point satisfactorily.

From the extensive preliminary report addressing several analysis techniques based on the data generated with the core chemicals it seems that the use of "Fold change" in combination with "Percent of control" to be the most adequate way to report the results. This procedure was applied with the supplementary chemicals and in the report is shown to have worked very satisfactorily.

After reaching a consensus, the protocol "Exposure of H295R…." should include a section indicating how the results are going to be analyzed, how chemicals are going to be classified, etc. Similarly, the section 6.1.3 of SOP#2 "Exposure of H295R…." should include clear cutoffs in order for a chemical to be further analyzed regarding steroid synthesis. In the preliminary report a cutoff of more than 80% viability was used and it seems to be an excellent choice since chemicals which further decrease viability would probably have non-specific effects on steroidogenesis.

What additional advice, if any, can be given regarding the protocol?

1. Cell suspension: H295R cells have a strong tendency to clump after trypsinization and this could be one of the reasons for the relatively high degree of variation observed in some of the protocols and specially with inexperienced laboratories. Although it is indicated in several parts of the protocol, I think it should be further highlighted the necessity to gently but consistently resuspend the cells after homogenization and/or centrifugation. The use of a pipette that can hold all the volume of media containing the cells that needs to be resuspended is very important. Also, the use of pipetting device that can aspirate/deliver liquid at an adequate speed to ensure good cell resuspension.

2. Use of "master mixes" to add test compounds: As indicated above if the use of "master mixes" was not the routine procedure it is greatly advice to use them to reduce the error due to pipetting small volumes.
3. Addition of media to cell culture plate wells: The volume added to each well of media plus test compound is very important since it later on it is assumed to be exactly 1 ml for all calculations. To reduce the error, it is greatly advice to prepare a "master mix" as indicated above and then dispense 1 ml per well using a 1 ml pipette, giving even preference to the use of a micropipette. The use of pipettes that can hold larger volumes, i.e. 5 ml pipettes, could add a significant error to the assay due to volume variations between wells of the media dispensed.
4. Since basal estradiol synthesis is very low and ,as indicated through the text, it is difficult to evaluate inhibitors of estradiol synthesis. One possible alternative would be to test the chemicals in cells treated simultaneously with forskolin where it would be easier to observe a decrease in the forskolin-mediated estradiol induction than a reduction of already low basal levels.
5. The crossreactivity of the core chemicals was evaluated in section 9.2.3 "Confounding factors". However, since the assay is planned to be used with a series of chemicals, it would be recommended to routinely test each of the chemicals or samples to be tested using the H295R steroidogenic assay for positive (as it was tested for the core chemicals) as well as negative interference effects. Each chemical should be tested at least at the higher concentration used for both interfering effects: a) positive: media which have had no contact with cells supplemented with the chemical at the highest concentration tested; b) negative: media which have had no contact with cells spiked with either testosterone or estradiol and supplemented with the chemical at the highest concentration tested. This test should be run routinely for each tested chemical/sample and will help to identify chemicals/samples that either increase or decrease the apparent concentration of each steroid in the determination assay.

4. Have the strengths and/or limitations of the assay been adequately addressed?

The strengths and limitations have been adequately addressed in the protocol.

The major strengths of the assay are that: 1) H295R cells are commercially available, 2) it is an *in vitro* system that does not require the use of live animals, 3) H295R cells are of human origin

which would make results more relevant to human endocrinology and cell physiology, 4) the protocol is relatively easy to perform allowing its wide use as an screening tool, 5) the possibility to use any steroid determination methods that successfully passed the quality control criteria using validated standards.
The limitations of the assay are that: 1) the system does not allow to study complex interactions that occur *in vivo*, 2) the system does not allow to study the regulation of the hypothalamic-pituitary axis, 3) the system does not allow to detect very weak inducers or inhibitors, 4) the system does not allow to study the effect of metabolites of the tested chemicals generated *in vivo*.

5. Were the (a) test substances, (b) analytical methods, and (c) statistical methods chosen appropriate to demonstrate the performance of the assay?
The test substances, and analytical and statistical methods chosen were appropriate to validate the assay. However, a decision should be done regarding the most adequate methods for data analysis and data report and this should be clearly stated in the SOP "Exposure of H295R….".

6. Considering the variability inherent in biological and chemical test methods, were the results obtained with this assay sufficiently repeatable and reproducible?
Results obtained following the protocols seem to be repeatable and reproducible. However, several suggestions are indicated under the "What additional advice, if any, can be given regarding the protocol?" heading that would probably help to further increase assay performance and reproducibility.

7. With respect to performance criteria, were appropriate parameters selected and reasonable values chosen to ensure proper performance of the assay?
The performance criteria are adequate and would allow the assay to be performed at multiple laboratories without major problems.
Table 7.2 and Table I.2 from the SOP "Exposure of H295R…." should be checked for consistency since many of the parameters differ between both of them.

8. Are the data interpretation criteria clear, comprehensive, and consistent with the stated purpose?

As indicated above, although the preliminary report deals extensively with data analysis and report the protocol "Exposure of H295R…." does not address this point satisfactorily since it does not have guidelines on how to interpret the data.
From the extensive preliminary report addressing several analysis techniques based on the data generated with the core chemicals, it seems that the use of "Fold change" in combination with "Percent of control" to be the most adequate way to report the results. This procedure was applied with the supplementary chemicals and in the report is shown to have worked very satisfactorily.
After reaching a consensus, the protocol "Exposure of H295R…." should include a section indicating how the results are going to be analyzed, how chemicals are going to be classified, etc.

9. Please comment on the overall utility of the assay as a screening tool in the EDSP Tier 1 battery.
The H295R steroidogenic assay would be an invaluable tool that would complement other assays of the Tier 1 battery. The assay has multiple advantages including a relative easy to perform, inexpensive an reproducible *in vitro* screening tool, that do not rely on live animals or animal tissues that may allow the screening of multiple compounds in a relative short period. The assay would identify chemicals with endocrine disruptor characteristics that could be further evaluated with other assays of the Tier 1 battery.

3.3 Thomas Sanderson Review Comments

A Peer Review of the H295R Steroidogenesis Assay as an appropriate Tier 1 screening tool for endocrine disruptors

J Thomas Sanderson

1. Is the stated purpose of the assay clear?

The 'interim H295R validation report' states that section 408 of the FFDCA requires EPA to:

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

This directive has been redefined by the EPA resulting in the development a two-tier testing system for endocrine disruptors, which would cover disruption of the androgen, estrogen and thyroid hormone systems. The interim report defines the steroidogenesis assay as a screening tool for the detection of any substance that would disrupt estrogen and/or androgen gonadal steroid hormone production. The definition goes on to say that the steroidogenesis assay is intended to detect any disruption of the intracellular biochemistry involved in the formation of the gonadal estrogens and androgens, but excluding any disruptions that may occur before the receptor (question: which receptor(s)?), effects on storage of sex hormones or effects on the hypothalamic-pituitary-gonadal axis. However, it remains unclear which steps involved in steroidogenesis are considered to be part of this tier 1 screening tool and which are not. For example, are interactions of chemicals with cell surface receptors that may modulate steroidogenesis included or not (the definition states *...after the receptor...*)? …and what about effects on cholesterol storage/release/*de novo* synthesis?

2. Is the assay biologically and toxicologically relevant to the stated purpose?

Estrogens and androgens are crucial hormones for human development and are involved in numerous processes in almost all tissues of the body. Dramatic changes in sex hormone levels during critical periods of development are known to adversely affect human development and health. Thus interferences with sex hormone production induced by chemicals or otherwise may have deleterious repercussion for the organism. However, sex hormone levels vary considerable among individuals; they also vary considerably during the day and month and are relatively easily affected by stress, diet and other life-style factors. They are also produced by several tissues in the body, not just the gonads.

One major question is whether any small change in hormone production in an isolated *in vitro* system has any relevance for the health outcome of an exposed organism. This remains

unaddressed in the documents available for review.

Also, how relevant is the use of an adrenocortical cell line for what is intended to be the screening of chemical effects on the gonadal sex hormones? Although steroidogenic enzymes such as CYP19 and CYP17, for example, are the same in these tissues, they are not regulated in the same way in the adrenal cortex as in the gonads (Bulun *et al.*, 2003; Simpson, 2003).

Furthermore, various other factors that may influence basal secretion of steroid hormones, such as extent of conjugation to sulfates/glucuronides, activity of transport proteins and mechanisms that control storage capacity will be different in different tissues. As the steroidogenesis assay only looks at one final outcome, namely the amount of estradiol and testosterone secreted, it is not possible to make biologically meaningful statements on the relevance of any observed disruption for the organisms as a whole. There are so many factors not directly related to steroidogenesis that could influence the assay system as it is currently described and intended to be used, that the issue of 'false positives' is likely to be an important concern, particularly once dealing with unknown complex environmental samples.

3. Does the protocol describe the methodology of the assay in a clear and concise manner so that the laboratory can:

a) comprehend the objective?

The objectives of the H295R steroidogenesis assay are not very clearly described in the appendices I and II. The sections *Purpose* and *Scope* are not very informative. Under *Purpose*, for example, the purpose of the document is described, not the purpose of the actual assay the document is meant to describe. Under *Scope and Application*, the reader does not find an easy guide to what the assay is about. It is also not helpful that appendix I has an appendix I and II and that appendix II has an appendix I.

It would be more logical to have a single protocol that covers the four main aspects of the H295R steroidogenesis assay (1) Cell culture (2) Exposure to test compounds (3) Analysis of estradiol and testosterone (4) Data analysis and presentation/interpretation of the results.

b) conduct the assay?

The protocol is described at length in appendices I and II. However, no methodological information whatsoever is provided in the interim report, which is a major limitation to the comprehension of the results discussed in this document.

Questions concerning the protocol for H295R cell culture: is Nu-Serum available world-wide? Also, what are the batch to batch variations in the sex hormone content of the Nu-Serum? Finally, why was not the use of a steroid-free medium recommended?

c) observe and measure prescribed endpoints?

One of the most important aspects of the H295R steroidogenesis assay, the analysis of testosterone and estradiol, is poorly defined in the provided documents. The choice of analysis method is left entirely to the implementing laboratory. It is known that ELISAs and RIAs can have very different outcomes dependent on the sample dilution, kit and antibodies used, not to mention the numerous confounding factors (solvent, cross-reactive components). The issues of cross-reactivity, how to deal with conjugated metabolites, and how to reliably compare between hormone levels determined by RIA or LC-MS are left undiscussed. It is highly inconsistent that there is an elaborate protocol for the 'consistent' use of a standard method such as the LIVE/DEAD cytotoxicity kit while no detailed attention is given to the crucial hormone analysis methodology.

d) compile and prepare data for statistical analyses?

There does not appear to be a section dedicated to this important aspect of the steroidogenesis assay in the protocols (Appendices I, II, III). Some information can be retrieved from the interim report (Chapter 8), raising questions: How is normality tested, is it by a standard chi-square test? The steroidogenesis assays essentially requires only that a deviation from the basal secretion of estradiol or testosterone is tested statistically. This can be done by using a two-tailed t-test or its non-parametric version, the Mann Whitney U test, with or without correction for multiple comparisons, if required (Zar, 1999). The text, however, mentions the non-parametric version of

ANOVA, the Kruskal-Wallis test. An ANOVA-style test is not really appropriate for concentration-response data. Can this be clarified?

e) report the results?

There does not appear to be a section dedicated to this aspect of the steroidogenesis assay and this should be included. For example, how should concentration-response data be expressed and presented? There are several figures in the interim report that are not interpretable: Figures 3.4 and 3.5 express testosterone concentrations as a % of the maximal response to prochloraz/fadrozole but the percentages are negative. What is considered the *maximum* response in these figures, and what does -20% of the maximum response of prochloraz/fadrozole mean?

The same problem returns in figures 10.1-10.4. Zero % of any response is zero, -20% of a response is impossible. A consistent approach would be to express all data as a % of basal hormone secretion, as this conforms to the aim of the assay as currently defined.

What additional advice, if, any, can be given regarding the protocol?

Analysis of sex hormones

The greatest weakness in the protocols is the lack of detail on sex hormone analysis methodology. This reviewer is of the opinion that LC-MS would be, by far, the preferred analysis tool for the detection of testosterone and estradiol. LC-MS would avoid the problems that will be (and already have been) encountered with inappropriate cross-reactivity of test samples/chemicals with the antibodies used in sex steroid ELISAs and RIAs. Please see also comments on trenbolone under point 7. The validation of a sensitive LC-MS method should be a logical part of the H295R steroidogenesis assay as currently defined. Furthermore, a single LC-MS analysis could detect a number of steroids in addition to estradiol and testosterone at little additional effort/expense, thus improving the 'expandability' of the H295R tool for other hormone endpoints.

4. Have the strengths and/or limitations of the assay been adequately addressed?

There is a brief discussion of strengths and weaknesses, but lacks detail and supporting scientific references. The main strength mentioned in the interim report is that the H295R cell line is a pluripotent cell lines that expresses all the enzymes necessary for the production of testosterone and estradiol. However, the fact that numerous other steroid hormone synthesis pathways are also present, although acknowledged, is not discussed. The implications of the presence of these other pathways (aldosterone, cortisol synthesis) may be far reaching for the reliable application of the proposed H295R steroidogenesis assay, as all these pathways are interconnected (at least in adrenocortical cells, not necessarily in gonadal cells). There is no critical discussion of the potential drawbacks of choosing an *adrenocortical* cell line to study effects of chemicals on *gonadal* testosterone and estradiol production. There is no scientifically supported discussion of the possible differences in regulation of steroidogenesis in adrenocortical cells and gonadal cells, yet it is known these are qualitatively and quantitatively very different. Several of the above points have been discussed in detail in several publications from my own lab in recent years (Sanderson and van den Berg, 2003; Sanderson, 2006).

5. Were the (a) test substances, (b) analytical methods and (c) statistical methods chosen appropriate to demonstrate the performance of the assay?

Choice of chemicals:

The compounds selected appear to be largely appropriate for validation of the assay, although the information given in Table 6.2 and 6.3 to support the choice was not very helpful. Specifically, the information under heading 'mode of action' and 'effect type' is not clear. Under *mode of action* a target may be mentioned but no information is given concerning the effect on that target. For example, is an *ER binder* an agonist or antagonist? Is trilostane really a strong inducer of T and E2 production (which seems unlikely given its 3bHSD inhibition potential), or is this erroneously based on the results of the present interim report, which indicates that the apparent induction of E2 and T is the artefact of cross-reactivity with the immunoassay kit? Danazol is said to have unknown effects, however, it is a well known (no longer used) medication against endometriosis withdrawn for its anabolic/androgenic effects. How is vinclozolin an inducer and inhibitor of T production at the same time? Also, flutamide,

genistein, glyphosate, RU486 and spironolactone are missing relevant information on their mode of action. Table 6.3 also needs references, and abbreviations need to be defined.

Analytical methods:

The use of immunoassays for the determination of testosterone and estradiol raises major concerns. There are numerous commercial antibody-based kits on the market, which all have different specificities for the target molecules. The testosterone detection kits usually show considerable 5-30% cross-reactivity with DHT and/or androstenedione. Estradiol kits generally show 10-15% cross-reactivity with estrone and for both hormones cross-reactivity with their sulfate conjugates can be as high as 100% (although extraction of free hormone circumvents this problem). Earlier studies with H295 cells indicate that these cells produce relatively large quantities of androstenedione and 11beta-OH androstenedione (the latter not usually found in gonadal tissues or healthy adrenal cortex)(Gazdar *et al.*, 1990). Without knowing how much androstenedione and other potentially cross-reactive steroids and metabolites are present in the cellular system it is difficult to assign any reliable value to the concentrations determined by immunoassay. A concentration of 5 pg/ml testosterone may in fact be more than 50% androstenedione, or something else, or not. These types of uncertainty need to be eliminated.

Trilostane which is a steroid with an androgen base structure, not surprisingly, interferes with the testosterone immunoassay kits. There will be more steroid-like molecules (in environmental extracts and other unknowns) that will interfere with immunoassay based hormone analysis especially if one considers that these test compounds/extracts are added to the assay system in micromolar quantities whereas the endogenous hormones are present in picomolar quantities. Thus even a cross-reactivity of less than 1% would cause major interference. See also comment on trenbolone under point 7.

If immunoassay based analytical methods will be continued to be used all these pitfalls will need to be addressed.

Statistical methods:

Statistical testing hypotheses have not been explicitly defined. See also comments under point 3e.

6. Considering the variability inherent in biological and chemical test methods, were the results obtained with this assay sufficiently repeatable and reproducible?

The reproducibility of the test system appears to be relatively poor. This may be partly due to the variability inherent in the use of cell lines in culture, but is also likely to be due to the various immunoassay-based hormone analysis methods used. The latter influence may be reduced by selecting a single method of detection, preferably not immunoassay based. Furthermore, the steroidogenesis assay depends on basal hormone secretion; results may be more consistent if estradiol and testosterone production are monitored after exposure of cells to an early precursor hormone in the biosynthesis pathway, such as pregnenolone. This would change the nature of the assay, but in a way would make the assay a more steroidogenesis-focused assay, as it would eliminate early variables such availability of cholesterol as precursor for the steroidogenesis of the sex hormones.

7. With respect to the performance criteria, were appropriate parameters selected and reasonable values chosen to ensure proper performance of the assay?

The performance criteria are outlined in Table 7.1 of the interim report. However, a performance criterion for inhibition of estradiol production by prochloraz is lacking; this needs to be addressed. Forskolin is used as a positive control for induction of testosterone and estradiol. This is a reasonable choice. It must, however, be kept in mind that forskolin increases the production of these two hormones via a very specific mechanism, by stimulating intracellular cAMP levels causing induction of various steroidogenic enzymes and ultimately increased synthesis of the sex hormones, but also of cortisol. There are however, many other mechanisms by which testosterone and or estradiol concentrations can be affected in H295R cells. (preferential inhibition of aldosterone/cortisol synthesis, increased bioavailability of cholesterol or decreased conjugation pathways, increased membrane permeability etc.). The steroidogenesis as currently set up will not be able to distinguish between any of these mechanisms, which in itself is not the intention. But it does means that comparing an induction response by a sample/unknown to that caused by forskolin as a performance criterion may in numerous instances be comparing apples to oranges. This makes the use of the Percent Control concept (Chapter 10.3) fundamentally flawed.

The interim report mentions that forskolin may not be the best choice of inducer because its effect on testosterone production is relatively weak. This likely due to the fact that forskolin strongly induces aromatase activity, which consumes testosterone to form estradiol. A better response may be obtained if the assay is adapted to use a (pregnenolone) precursor to avoid the limitation of substrate availability to the various steroidogenic enzymes of interest.

The interim report also mentions that trenbolone is being considered as replacement for forskolin as a positive control for induction of testosterone production. The immediate question is whether trenbolone, which is a steroid with a structure very similar to testosterone, is not in fact causing cross-reactivity with the immunoassays for testosterone. Has this been ruled-out? The next question would be how trenbolone, a potent AR agonist, is able to induce testosterone levels in H295R cells? AR agonists do not normally have any effect on testosterone formation in these cells.

8. Are the data interpretation criteria clear, comprehensive, and consistent with the stated purpose?

For data interpretation criteria I am dependent on the information dispersed over Chapters 7.3 and 8 and 10. Using the H295R steroidogenesis assay as a semi-quantitative screening tool is a reasonable approach. The classification of inducers into weak, medium, strong and very strong seems too elaborate. Given the large variability and uncertainties in hormone determinations and mechanisms of induction, as well as the limited meaningfulness of fluctuation in hormone levels that are less than 2-fold it would be preferable to reduce this classification to weak (2-5 fold) and strong (>5-fold) inducers, and consider anything less than 2-fold as 'possible' inducers. Expressing results using the PCmax/PC50 concept is, as mentioned under point 7, not likely to be very useful.

9. Please comment on the overall utility of the assay as a screening tool in the EDSP Tier 1 battery.

The EPA has been charged with the development and implementation of a screening tool for interferences with steroidogenesis. Given this task, the EPA has decided to develop a tool that will evaluate effects on gonadal sex hormone production, specifically testosterone and estradiol.

To do this the H295R human adrenocortical carcinoma cell line was chosen as *in vitro* biological system. The H295R steroidogenesis assay is intended to detect chemicals and/or complex environmental samples extracts that may interfere with basal secretion of testosterone and estradiol.

As a system to study effects of chemicals on steroidogenesis the H295R cell line has great potential as it is capable of producing mineralocorticoids, glucocorticoids, androgens and estrogens. The steroid biosynthesis pathway is highly complex and also highly interconnected. This limits the usefulness of only evaluating effects on one or two specific hormones as there are a large number of influences unrelated to the steroidogenesis of those hormones that may cause small fluctuations in their secretion by the cell system. The system, as designed, is not really a steroidogenesis assay although is may pick up inhibitors and inducers of testosterone and estradiol synthesis. However, because the system, as designed, ignores all the other steroid hormones, including other active androgens and estrogens known to be produced in these cells, a bigger picture is not obtained, limiting the interpretability of any observed alterations in solely levels of testosterone and estradiol. The H295R system also in no way reflects the conditions of a gonadal system in which mineralo- and glucocorticoid pathways do not play a role. The regulation of the various steroidogenic enzymes will also be different in different tissues, again limiting the interpretability of any observed effects on induction of testosterone or estradiol secretion (if steroidogenesis related).

For the EPA to have a true steroidogenesis assay (a system that detect the ability of chemicals to interfere with the biosynthesis of steroid hormones) the H295R cell line could provide a very useful model with some alterations to the design: (1) H295R cells would be analyzed for 4 key steroid hormones (aldosterone, cortisol, estradiol and testosterone) using a single analytical technique such as LC-MS, (2) pregnenolone would be used as precursor for all steroids (3) effects on the relative production of the 4 hormones would be relatively easy to interpret as they would provide clues on which of these 4 essential steroid hormones and which steps of the steroidogenic pathway are affected.

The way the H295R cell system is being proposed to be used is like a black box. It will be difficult to interpret the meaning of any outcomes that may be observed on testosterone and estradiol levels, and this is further compounded by the drawbacks of using immunoassay-based detection methods. A more focused definition of the purpose of a tier 1 assay for steroidogenesis would be recommendable; allowing for the development of a H295R cell-based steroidogenesis

assay that would provide less ambiguous information about the steroidogenesis disruption potential of chemicals or unknown environmental extracts.

References

Bulun, S. E., Sebastian, S., Takayama, K., Suzuki, T., Sasano, H., and Shozu, M. (2003). The human CYP19 (aromatase P450) gene: update on physiologic roles and genomic organization of promoters. *J Steroid Biochem Mol Biol* **86**, 219-224.

Gazdar, A. F., Oie, H. K., Shackleton, C. H., Chen, T. R., Triche, T. J., Myers, C. E., and Chrousos, G. P. (1990). Establishment and characterization of a human adrenocortical carcinoma cell line that expresses multiple pathways of steroid biosynthesis. *Cancer Res.* **50**, 5488-5496.

Sanderson, J., and van den Berg, M. (2003). Interactions of xenobiotics with the steroid hormone biosynthesis pathway. *Pure Appl Chem* **75**, 1957-1971.

Sanderson, J. T. (2006). The steroid hormone biosynthesis pathway as a target for endocrine-disrupting chemicals. *Toxicol Sci* **94**, 3-21.

Simpson, E. R. (2003). Sources of estrogen and their importance. *J Steroid Biochem Mol Biol* **86**, 225-230.

Zar, J. H. (1999). *Biostatistical Analysis*. Prentice-Hall, N.J.

3.4 Matt Vijayan Review Comments

CHARGE QUESTIONS

1. Is the stated purpose of the assay clear?

Yes, it is clear that the H295R assay is being used as a screening tool to detect substances that will impact estrogen and/or androgen production.

Pg.17. Not sure what you mean by "storage or release of gonadal steroid hormones"?

Pg.17, last but one line. "…identify chemicals that will act to alter steroidogenic process". This assay will not identify chemicals that will act upstream of cAMP, including trophic hormone stimulation.

2. Is the assay biologically and toxicologically relevant to the stated purpose?

The H295R cell line is derived from a human adrenocortical carcinoma cell line and used a model system to examine sex steroid production. The rationale being that the full complement of the steroid biosynthetic pathway is present in this cell system. Also, this cell line produces sex steroid constitutively for its use to detect inhibitors of steroid production. Consequently, use of this model system as a screening tool to identify substance that can potentially impact steroid biosynthetic pathway leading to hormone production is valid. However, the extrapolation of this information, from a carcinoma cell line, to the impact (or lack off) of these test substances *in vivo* requires further validation as the endocrine physiology of the gonads and the adrenal gland is different from this undifferentiated cell system. Specifically, this cell system lacks the membrane receptors required for trophic hormone (gonadotropin) stimulation, which is an essential component of the sex steroid biosynthetic cascade.

3. Does the protocol describe the methodology of the assay in a clear, and concise manner so that the laboratory can:

A) comprehend the objectives;

YES, the protocol is easy to follow and the objectives are clear.

B) conduct the assay;

The protocol is provided in sufficient detail and the methodology is well laid out for any laboratory to conduct the assay.

C) observe and measure prescribed endpoints;

The endpoints involve collection of medium for measuring steroids and the cells for cytotoxicity assay. The steroid measurement may involve hormone extraction from the medium and a methodology is provided for consistency in extraction efficiency among laboratories. A protocol that can eliminate this extraction step may be better suited for wider application. This can be ascertained by testing the interference of the test substances with the antibody cross-reactivity. The cytotoxicity assay is also well explained and easy to carry out. However, I am not clear what greater than 100% cell viability means (Figs. 9.5 and 9.6 and pg. 56) as this could potentially confound the results.

D) compile and prepare data for statistical analyses;

The data compilation is explained clearly and a data sheet template is also provided. However, there appears to be some confusion around data normalization. The magnitude of change from

control either shown as actual concentration change or percent change would be appropriate for inter-laboratory validation.

E) report the results?

The results and the statistical analyses are clearly explained and easy to follow.

4. Have the strengths and/or limitations of the assay been adequately addressed?

The strengths have been addressed adequately but the limitation of the assay requires mention (see pg. 20 line 3 onwards). For instance this assay will only detect changes that happens post-receptor activation. This is a drawback to this cell system because *in vivo* the steroidogenic cells secrete steroids in response to trophic hormone stimulation. This assay completely bypasses the receptor signaling which is an essential step in steroid biosynthesis. So substances that can affect steroid production by altering trophic hormone signaling will not be evaluated by this cell system. Also, the high constitutive production of the hormone is abnormal *in vivo* as this usually happens only in response to trophic hormone stimulation. So it is unknown whether the changes seen with the test substances can be mimicked *in vivo* to the same extent (or may be even greater) and will require confirmation with animal models or other relevant cell or tissue systems. Also, the high constitutive levels of steroids, for instance testosterone, may deplete the precursor available for steroid synthesis and may be limiting the steroid biosynthetic capacity in response to test (inducer) substances. The changes in the magnitude of steroid synthesis with forskolin, smaller change for testosterone because basal secretion is high and higher for E2 because of lower basal secretion, clearly support this contention. This requires testing perhaps by supplementing the medium with cholesterol.

5) Were the (a) test substances, (b) analytical methods, and (c) statistical methods chosen appropriate to demonstrate the performance of the assay?

Little is known about the impact of most of the test substances on steroid production. The lack of response to a known inducer of sex steroid production in gonadal tissue, for instance human chorionic gonadotropin (hcG), suggests that this system has limitations because of the type of tissue involved (adrenal carcinoma). Also, I am surprised that neither arylhydrocarbon receptor ligands (for instance PCBs) nor metals were used as a test substance to validate steroid output using this model system, especially since several studies have shown that metals and PCBs

inhibit steroidogenesis. It may also be worthwhile using DMSO as a test substance especially since it is being used as a solvent control.
The analytical methods and the statistical methods are appropriate to demonstrate the performance.

6. Considering the variability inherent in biological and chemical test methods, were the results obtained with this assay, sufficiently repeatable and reproducible?

Foe the most part the assay is sufficiently repeatable and reproducible. However, I am concerned with the high CV among laboratories and also within laboratories. The within lab CV is particularly high for prochloraz and this could be because it is inhibiting the basal steroid production. As the constitutive levels are being inhibited this may lead to error as the levels may differ due to autoregulation that is inherent in this system. I would recommend using a test group where the inhibition is tested using acute-stimulated (forskolin or 8bromocAMP) steroid production as a model. This might reduce the variability and make the data set more comparable between the laboratories. For instance there is a large variability in EC50 for foskolin between the different labs (Table 10.3). The advantage of using a cell line is the consistency in response no matter where it is used but the high CV (ranging from 57 – 89%; Table 9.2) shown here suggests that the basal production of hormone is subjected to autoregulation leading to differences in the magnitude of response. In this regard, the basal hormone levels may be a key variable that need to be within a narrow range among laboratories prior to screening for substances modulating sex steroid production.

7. With respect to performance criteria, were appropriate parameters selected and reasonable values chosen to ensure performance of the assay?

The test substances chosen were appropriate to demonstrate the performance of the assay (forskolin and prochloraz as inducer and inhibitor, respectively). However, the magnitude of change is very different for testosterone and estradiol. This difference may be related to the difference in their basal secretion rate (high for T and low for E2). Consequently, changes in E2 levels may not be a good performance indicator for testing inhibitors of steroidogenesis. It may be useful to use other inducers such as cAMP analogue and 25 hydoxycholesterol to obtain stimulated steroid production levels to validate the performance assay. Also, supplementing

medium with cholesterol may be required to confirm that this precursor is not a limiting factor for steroid production in this cell system given the high basal secretion for testosterone.
For testing the performance for inhibitors it may be necessary to use inhibition of stimulated-steroid production as the end point at least in the case of E2 secretion.
This cell line is derived from adrenal carcinoma and consequently would be a suitable system for detecting corticosteroid production. Hence, a stimulated (ACTH or 8bromocAMP) cortisol production may be useful as a positive control for cell system validation among laboratories to meet the QA/QC criteria.
The CV for SCs that is acceptable for QC is relatively high. I would suggest a CV≤20% as acceptable for replicate measures within a laboratory.
On Table7.1 the performance criteria for estradiol with forskolin is given as ≥15 times SC, whereas on pg. 36 it is shown as ≥10-times induction of E2 production.
Pg. 44. 9.1.1.1. line 5, change to Lab 5
Table 9.1 – change "second" Lab 4 to Lab 5.

8. Are the data interpretation criteria clear, comprehensive, and consistent with the stated purpose?

The data interpretation is clear and consistent with the objective of the report. However, I am not convinced with the categorization of test substances as weak, medium, strong or very strong, because of some of the limitations of the cell system. For instance the lack of response (or weak response) may be due to the high basal hormone production in the case of testosterone or the low secretion for E2. This needs to be further tested, refined and validated for both testosterone and estradiol. Also, the dose-response curves will have to be tightened (narrower range) based on the initial screening.

9. Please comment on the overall utility of the assay as a screening tool in the EDSP Tier 1 battery.

The H295R steroidogenic assay has been validated for its steroid production capacity and as a tool for screening substances that modulate sex steroid production. The multi-laboratory validation suggests that the assay has potential as a screening tool for sex steroid disruptors. However, the assay has limitations and some of them are related to the cell system itself. For

instance the high basal unstimulated sex steroid production is not physiologically relevant but provides a model for testing the capacity for substances to induce or inhibit steroidogenesis. This assay focuses only on the signaling pathway downstream of trophic hormone stimulation. While the mechanism for the high basal testosterone output is unclear, it remains to be seen if that would modify the steroid production capacity in response to stimulators or inhibitors. Also, the low basal E2 production seen with these cells does not provide an ideal model to test inhibitors of steroidogenesis. This can be easily tested by examining the capacity of the modulators to inhibit forskolin- or 8bromo-cAMP-stimulated E2 production. The huge CV reported for between laboratory comparisons may have to do with the difference in basal hormone production and associated differences in the magnitude of response to know inducers and inhibitors as well as test substances. Overall, the assay has the potential to be a screening tool for steroidogenesis but requires further testing and refinement.

Appendix A

CHARGE TO PEER REVIEWERS

CHARGE TO PEER REVIEWERS
for
INDEPENDENT PEER REVIEW OF THE H295R CELL-BASED ASSAY FOR STEROIDOGENESIS
AS A POTENTIAL SCREEN IN THE ENDOCRINE DISRUPTOR SCREENING PROGRAM (EDSP) TIER-1 BATTERY

April 29, 2008

Background:

According to Section 408(p) of the EPA's Federal Food Drug and Cosmetic Act, the purpose of the EDSP is to:

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

Subsequent to passage of the Act, the EPA formed the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC), a panel of scientists and stakeholders that was charged by the EPA to provide recommendations on how to implement the EDSP. Upon recommendations from the EDSTAC, the EPA expanded the EDSP using the Administrator's discretionary authority to include the androgen and thyroid hormone systems as well as wildlife.

One of the test systems recommended by the EDSTAC was the sliced testes assay. Its purpose in the Tier-1 battery was to provide a sensitive *in vitro* assay to detect chemicals that may affect the endocrine system by inhibiting the enzymes responsible for the inhibition of enzymes in the steroid hormone synthesis pathway. After encountering two substantial issues with the standardization of the sliced testes assay—high variability and the inability to distinguish general cytotoxicity from Leydig cell toxicity—EPA abandoned the sliced testes assay in favor of the H295R. The H295R assay offered a number of substantial advantages over the sliced testes assay and other cell-based assays. Like other cell based assays it does not use animal tissue and is capable of detecting inducers as well as enzyme inhibitors. Unlike the other cell-based assays, it contains all of the enzymes of the steroidogenic pathway.

Although peer review of the H295R assay will be done on an individual basis (i.e., its strengths and limitations evaluated as a stand alone assay), it is noted that the H295R assay along with a number of other *in vitro* and *in vivo* assays will potentially constitute a battery of complementary screening assays. A weight-of–evidence approach is also expected to be used among assays within the Tier-1 battery to determine whether a chemical substance has a positive or negative effect on the estrogen, androgen or thyroid hormonal systems.

The FIFRA Scientific Advisory Panel (SAP) has already conducted a peer review of the EPA's recommendations for the Tier-1 battery. The H295R assay was one of the assays recommended by EPA contingent upon satisfactory validation and peer review of the assay.

This peer review will focus on the scientific work EPA performed to validate the H295R assay. Each peer reviewer is asked to focus his/her review on this issue. Unlike other peer reviews EPA did not have time to produce an Integrated Summary Report (ISR), so peer reviewers will be

asked to focus on the interim final validation report and to a lesser extent on the prevalidation reports for conducting this review. It should be noted that in order to meet the August 2008 deadline for implementation, EPA is requesting review of the interlaboratory study on the 12 core chemicals and the 18 supplementary chemicals that were tested in the lead laboratory. When the other participating labs have completed testing of the 18 supplementary chemicals a final report will be prepared which will also undergo peer review.

Charge Questions:

Your review and comments should be directed to each of the following questions:

1. Is the stated purpose of the assay clear?

2. Is the assay biologically and toxicologically relevant to the stated purpose?

3. Does the protocol describe the methodology of the assay in a clear, and concise manner so that the laboratory can:
   a) comprehend the objective;
   b) conduct the assay;
   c) observe and measure prescribed endpoints;
   d) compile and prepare data for statistical analyses; and
   e) report the results?

   What additional advice, if any, can be given regarding the protocol?

4. Have the strengths and/or limitations of the assay been adequately addressed?

5. Were the (a) test substances, (b) analytical methods, and (c) statistical methods chosen appropriate to demonstrate the performance of the assay?

6. Considering the variability inherent in biological and chemical test methods, were the results obtained with this assay sufficiently repeatable and reproducible?

7. With respect to performance criteria, were appropriate parameters selected and reasonable values chosen to ensure proper performance of the assay?

8. Are the data interpretation criteria clear, comprehensive, and consistent with the stated purpose?

9. Please comment on the overall utility of the assay as a screening tool in the EDSP Tier 1 battery.

Appendix B

INTERIM FINAL VALIDATION REPORT

Multi-Laboratory Validation of the H295R Steroidogenesis Assay to Identify Modulators of Testosterone and Estradiol Production (PDF) (91 pp, 1.1MB)

Appendix C

SUPPORTING MATERIALS

Appendix 1. H295R Cell Culture Protocol

H295R Cell Culture Protocol (PDF) (22 pp, 162K)

Appendix 2. H295R Cell Exposure Protocol

H295R Cell Exposure Protocol (PDF) (25 pp, 221K)

Appendix 3. Data Entry Sheet Template

Data Entry Sheet Template (PDF) (10 pp, 51K)

Appendix 4. Raw data sheets Laboratory 1

Raw data sheets Laboratory 1 (XLS) (352K)

Appendix 5. Raw data sheets Laboratory 2

Raw data sheets Laboratory 2 (XLS) (476K)

Appendix 6. Raw data sheets Laboratory 3

Raw data sheets Laboratory 3 (XLS) (326K)

Appendix 7. Raw data sheets Laboratory 4

Raw data sheets Laboratory 4 (XLS) (246K)

Appendix 8. Raw data sheets Laboratory 5

Raw data sheets Laboratory 5 (XLS) (43K)

Appendix 9. Raw data sheets Laboratory 6

Raw data sheets Laboratory 6 (XLS) (175K)

Background Documents

Final Report - Development of an Assay Using the H295R Cell Line to Identify Chemical Modulators of Steroidogenesis and Aromatase Activity (PDF) (45 pp, 433K)

Draft Report – Standardization and Refinement of the H295R Cell- Based Assay to Identify Chemical Modulators of Steroidogenesis and Aromatase Activity (Repeat of Inter-Laboratory Study) (PDF) (52 pp, 683K)

Preliminary Report - Response of H295R Cells to the Exposure to Four Model Chemicals (Prochloraz, Aminoglutethimide, Forskolin & Vinclozolin) (PDF) (8 pp, 152K)