

Disposition to Peer Review Panel Comments on the H295R Steroidogenesis Assay

Comment	Commenter	Comment	EPA Response
Topic: 2.2 Comments on the Clarity or the Stated Purpose of the Assay.			
1	SM	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.	No response needed.
2	DR		All suggestions made by this reviewer have been accepted and the text was revised accordingly (p16, last paragraph).
3	TS	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	While H295R cells express gonadotropin receptors, it is unlikely that these receptors significantly influence the production of sex

		<p>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?</p>	<p>steroids by these cells. In this study we could demonstrate that addition of HCG did not affect sex steroid production, and similarly a different study (Rao et al 2004; Biol Reprod. 71: 579-587) found no effects of LH receptor agonists on steroid production in these cells (the only effect observed in this study was a weak increase [~50% relative to controls] in DHEA sulfate, indicating some minor increase in DHEA metabolism but no effect on androstenedione, cortisol or DHEA). While expression of the sex steroid receptors has been reported for the H295R cells, it is unclear at the current time what their potential influence on hormone production by the cell could be. However, given that the H295R Steroidogenesis Assay is utilized as part of a test battery including estrogen and androgen receptor binding assays, this does not seem to be of relevance with regard to the aim of this assay.</p> <p>However, GtH and steroid receptor mediated effects are beyond the scope of this assay. The purpose of the assay is to determine whether or not a chemical has the potential to alter the synthesis of steroid hormones, especially T and E by disrupting the function of the enzymes involved in this</p>
--	--	--	---

			process. Finally, effects on cholesterol storage/release/de novo synthesis are beyond the scope of the EDSP.
4	MV	<p>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.</p> <p>Pg.17. Not sure what you mean by “storage or release of gonadal steroid hormones”?</p> <p>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</p>	<p>It has been stated in the text that this assay only aims to address steroidogenic processes downstream of the GtH receptors: that comprises the intracellular biochemical pathway beginning with the sequence of reactions occurring after the gonatotropic hormone receptors (FSHR and LHR), up through and including the production of the terminal sex steroid hormones” (p16, last paragraph). However, changed sentence to more clearly reflect this: “that act to alter steroidogenic process downstream of the gonadotropin hormone receptors in humans and wildlife”</p> <p>Statement “storage and release of gonadal steroid hormones” has been removed from the text.</p>
Topic: 2.3 Comments on the Biological and Toxicological Relevance of the Assay as Related to its Stated Purpose			
1	SM	<p>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</p>	No response needed.

		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 steroidogenesis especially those involved in sexual male 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.	
2	DR	<p>The H295R steroidogenesis assay is biologically and toxicologically relevant to the stated purpose. The assay would fit perfectly in the Tier 1 battery 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.</p> <p>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.</p>	No response necessary.
3	TS	One major question is whether any small change in hormone production in an isolated in vitro system has any relevance for	This is an important consideration for the H295R assay as it is in any in vitro test

		the health outcome of an exposed organism. This remains unaddressed in the documents available for review.	system. That is the purpose of the validation studies using chemicals that are known to disrupt these processes <i>in vivo</i> and other <i>in vitro</i> assays. In addition, that is why this assay has to be interpreted in light of the entire battery results.
		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). 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.	As noted above, this assay is used in combination with a series of other in vitro and in vivo assay in the Tier I battery, and its purpose is not to elucidate complex interactions but to identify effects at a very specific level.
4	MV	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.	No response needed.

		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.	We agree. The following statement has been included into Section 3.2 to reflect this limitation: “It should be noted, however, that the H295R cells lack or only marginally express the membrane receptors required for trophic hormone (gonadotropin) stimulation, which are involved in the regulation of sex steroid biosynthesis in <i>in vivo</i> systems.”
	Topic: 2.4 Protocol (General Comments)		
1	SM	Protocol is well described and the methodology generally presented in a comprehensible manner allowing the reader to follow all steps cited above.	No response needed.
	Topic: 2.4.1 Protocol (Comprehend the Objective)		
1	DR	The objective of the assay is clearly stated.	No response needed.
2	TS	<p>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.</p> <p>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.</p>	<p>Appendices I and II are standard operating procedures describing step by step the processes to be undertaken to successfully conduct the assay. Purpose and scope were provided in the validation study report. There will be a complete and separate H295R protocol after this assay undergoes peer review both by the US-EPA and OECD which will combine all of these aspects. .</p>

3	MV	YES, the protocol is easy to follow and the objectives are clear.	No response needed.
Topic: 2.4.2 Protocol (Conduct the Assay)			
1	DR	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:	No response needed.
		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.	Sentence removed.
		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.	Numbering system has been revised and corrected.
		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.	Revised text as follows: “Passage 4.5 to 10.5. NCI-H295R cells (ATCC Cat # CRL-2128) cultured under standard conditions as described in the H295R culture protocol”

		<p>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.</p>	<p>Changed as recommended by this reviewer.</p>
		<p>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 □l 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 □l 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.</p>	<p>The H295R Exposure has been changed such that an option to use a mastermix was included (P11; Section 5.4.2): “Note: Alternatively, a mastermix containing 3.94 mL of medium plus 4 µL of the respective chemical stock solution in DMSO can be prepared prior to dosing the cells. Then, 1 mL of the appropriate mastermix is to be dispensed per replicate well of each dose. If this approach is chosen, omit step 4.”</p>
		<p>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.</p>	<p>Changed text as follows: “Plate reading fluorometer (note: this protocol has been tested with the Fluoroskan Ascent Fluorometric Microtiter Plate Reader (Thermo Electron Corporation))”</p>
		<p>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.</p>	<p>“200,000” corrected to “300,000”</p>
		<p>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</p>	<p>No plate sealer is used during this process. The plate is simply covered using the plate</p>

		number of the plate sealers to be used.	lid. The text has been changed as follows to reflect this (P15, Section 6.1.2): "... cover the plate with lid to prevent evaporation"
		9. SOP#2, page 16 and subsequent, item 7 and subitems, it is recommended to spike the sample with 3H-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 3H-estradiol be used for estradiol recovery calculation?	Only 3H Testosterone is added, and it assumed that the recovery rate observed is also valid for E2. While we agree that it would be more precise to use a 3H E2 spike this is not feasible because this would require a separate extraction for E2. Overall, the primary purpose of this spike is to correct for errors (e.g. pipetting) that occurred during the conduct of the extraction procedure, and for this purpose adding only one 3H hormone should be able to sufficiently address this issue.
		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.	Ether has been shown to be one of the most reliable solvents for the extraction of sex steroid hormones. Alternatives such as dichloromethane are even more of a health hazard. However, we agree with this reviewer's comment on the potential risks associated with the use of ether, and a line was included stating that "appropriate care must be taken when working with hazardous chemicals such as ether".
		11. SOP#2, page 17, item 7.1.2, in point "5" it says "add 10 <input type="checkbox"/> l of the 3H-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	Text has been included as recommended by this reviewer.

		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:	The formula has been completed by including adjustment to 1mL (1000uL). However, as the concentration as derived from the hormone ELISA already is adjusted to final concentration per mL this was not included in the formula. A statement has been included into the Example section to reflect this.
		The formula needs to be checked since the concentration of secreted steroid hormones is one of the quality criteria of the assay.	No changes in hormone concentrations occurred because the concentration derived from the hormone ELISA was already adjust to pg/mL.
		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.	These considerations have been added to the scope/important considerations section of both protocols.

2	TS	<p>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.</p> <p>Questions concerning the protocol for H295R cell culture:</p>	<p>The purpose of the APPENDICES I & II was the detailed description of the specific procedures required to successfully conduct the assay. We felt it easier to read the report if the details of the cell culture methods (appendix I) and the exposure protocol (appendix II) were kept separate. There will be a complete and separate H295R protocol after this assay underwent peer review both by the US-EPA and OECD which will combine all of these aspects. .</p>
		<p>is Nu-Serum available world-wide?</p>	<p>To our knowledge: Yes! All laboratories that participated in this validation study were able to obtain Nu serum with no problem. The countries that conducted these studies included Canada, China (Hong Kong), Denmark, Germany, Korea, Japan and the U.S.</p>
		<p>Also, what are the batch to batch variations in the sex hormone content of the Nu-Serum?</p>	<p>There is some variation, and average levels of e.g. E2 seem to be somewhere between 5 and 80 pg/ml. To control for any variation of hormone production resulting from changes in steroid levels in Nu serum, we have included the appropriate controls in the QC plates and we recommend measuring E2 values in the NuSerum so that the results can be appropriately evaluated before running the assay (i.e.,</p>

			before freeze down).
		Finally, why was not the use of a steroid-free medium recommended?	The cells need certain amounts of steroid precursors in the serum to be able to grow properly and to prime the appropriate production of sex steroids. We chose to use NuSerum because of the limited availability or unavailability of Ultrulser SF and NuSerum has been used commonly.
3	MV	The protocol is provided in sufficient detail and the methodology is well laid out for any laboratory to conduct the assay.	No response needed.
Topic: 2.4.3 Observe and Measure Prescribed Endpoints			
1	DR	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.	No response needed.
		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.	This is already done by including the QC plates with every assay conducted.
2	TS	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	For this reason QC criteria were defined. Most of the labs that participated in this study used different hormone detection assays (RIA, ELISA, Time Resolved Fluorescence, LC-MS), because different methods are preferred or accessibility is limited by various labs around the world. EPA cannot recommend a specific assay. Also, there are restrictions in the use of tests

		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.	utilizing radioisotopes, and mass spectrometry based technologies are not trivial and feasible for high throughput tests. For this reason QC criteria were defined that are applicable to all methods. While this may have been the source for some of the variation among labs, it should be acknowledged here that regardless of the assay used always the same type of response was observed.
3	MV	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.	This will be difficult if people intend to use antibody based assays. This is especially true because Most antibodies cross-react with the metabolization products (e.g. sulfate- and glucose-conjugates) of the original hormones.
		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.	It has been shown that some compounds such as forskolin enhance cell growth resulting in changes in fluorescence that may yield numbers greater than 100% However, as explained in the report, for this reason-hormone concentrations were normalized to % viable cell in each well.
Topic: 2.4.4 Compile and Prepare Data for Statistical Analyses			
1	DR	The worksheet design is adequate for data compilation and statistical analysis submission.	No response needed.
2	TS	There does not appear to be a section dedicated to this important aspect of the steroidogenesis assay in the protocols (Appendices	We are puzzled the comment that an ANOVA followed by the Dunnett’s is not

		<p>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?</p>	<p>appropriate as the design. Data do not violate any of the assumptions of the ANOVA. We have checked with two biostatisticians and they were in agreement that the statistical approach featured in this study is appropriate for the evaluation of the data. For a statistical analysis of the concentration-response data when values are expressed as fold- or percent-change, the issue here relates to effects of prochloraz or any other compound that will suppress T or E production. As the values fall with higher doses, there will be a systematic relationship between the mean and variance, affecting the assumption of homogeneity of variance. These relative values can be addressed by a form of data transformation- either a log transformation or arc sine transformation (used for proportions or percentages, where values are a percentage of controls). Before being transformed, individual control values and values for treated cells are calculated as differences from the overall control mean. The transformed data can then be used in an ANOVA, accompanied by Dunnett's test when the ANOVA shows an overall significant response.</p>
3	MV	<p>The data compilation is explained clearly and a data sheet template is also provided. However, there appears to be some</p>	<p>As stated in Section 8, the data has been expressed as changes relative to the SCs by</p>

		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.	dividing the concentration in each well by the average SC value. This is equivalent to the expression of the data as % response relative to the SC.
Topic: 2.4.5 Report Results			
1	DR	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.	A section describing data analysis and report has been included into the “H295R Exposure Protocol” (P22, Section 8 “Data Processing and Reporting, and Statistics”)
		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.	The newly added Section 8 (Data Processing and Reporting, and Statistics) lists these criteria.
2	TS	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?	A section describing data analysis and report has been included into the “H295R Exposure Protocol” (P22, Section 8 “Data Processing and Reporting, and Statistics”)
		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	This was explained in detail in the figure headings

		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.	The use of – and + was simply to demonstrate the direction of the effect. The final report will set the control level at 1.0 so that inhibition will show as fractions and induction levels as fold change above 1.
3	MV	The results and the statistical analyses are clearly explained and easy to follow.	No response needed.
Topic: 2.4.6 Provide Additional Advice Regarding the Protocol			
1	SM	In order to improve protocol, the following advices are proposed:	See below.
		1) The choice of solvent for steroid extraction should be specified (ethyl ether or dichloromethane).	This information had been specified in the H295R Exposure Protocol “Ether, Anhydrous (J.T. Baker Cat# 9244-22); See P20, Section 7.2.1
		2) During extraction procedure of steroids with ethyl ether a rapid freezing of aqueous phase (after step 9 and before step 10) facilitates the separation of aqueous (inferior) and organic (supernatant) phases (see Appendix ii page 17).	A line has been added to the protocol to describe the alternative method using rapid freezing of the water phase.
		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.	Added the following step to the protocol: “12. Wash ether with 1 mL nanopure water to remove possibly present hydrophilic contaminants. Cap, vortex, and centrifuge as above, and transfer ether to new glass vial. (Note: May not be necessary. Can be omitted if laboratory can demonstrate that there is no difference in analytical results between washed and non water treated samples).”

	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.	A check for cross-reactivity will be conducted in future work.												
	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) ?	Has been corrected to 10 passages throughout the protocols.												
	6) When possible, replace methanol (highly toxic) by ethanol for cytotoxicity analysis (see page 36, point 5 & table)	The use of methanol is recommended by the manufacturer.												
	7) There is a confusion between E2 decrease in the presence of 3 μ 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!	<p>The text has been changed both in the report and the protocols. The performance criteria now read as follows throughout the text:</p> <table border="1"> <thead> <tr> <th></th><th>Testosterone</th><th>Estradiol</th></tr> </thead> <tbody> <tr> <td><i>Basal Production</i></td><td>≥ 2.5-times MDL</td><td>≥ 2.5-times MDL</td></tr> <tr> <td><i>Induction (10uM forskolin)</i></td><td>≥ 2-times SC</td><td>≥ 10-times SC</td></tr> <tr> <td><i>Inhibition (3uM prochloraz)</i></td><td>≤ 0.5-times SC</td><td>≤ 0.5-times SC</td></tr> </tbody> </table>		Testosterone	Estradiol	<i>Basal Production</i>	≥ 2.5 -times MDL	≥ 2.5 -times MDL	<i>Induction (10uM forskolin)</i>	≥ 2 -times SC	≥ 10 -times SC	<i>Inhibition (3uM prochloraz)</i>	≤ 0.5 -times SC	≤ 0.5 -times SC
	Testosterone	Estradiol												
<i>Basal Production</i>	≥ 2.5 -times MDL	≥ 2.5 -times MDL												
<i>Induction (10uM forskolin)</i>	≥ 2 -times SC	≥ 10 -times SC												
<i>Inhibition (3uM prochloraz)</i>	≤ 0.5 -times SC	≤ 0.5 -times SC												
	8) Also, there is a difference in the reported induction of E2 by 10 μ M forskolin ; ≥ 10 -times induction of E2 in page 36 point 3, different from ≥ 15 -times induction in table 7.1, page 35!	The text has been changed to “10-fold” throughout all documents.												
	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.	Minimum detectable hormone concentrations and hormone detection systems used by each group have been												

			added to Table 9.1.
2	DR	<p>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.</p>	This has been explicitly mentioned in the “Important Consideration” section of the H295R Culture Protocol (P7; Section 2.2).
		<p>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.</p>	See previous response to this comment (Section 2.4.2, DR, Comment 5).
		<p>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.</p>	A section has been added to the “Important Considerations” section of the H295R Exposure Protocol emphasizing the importance of using pipettes that allow precise pipetting of required volumes during dosing (P7, Section 2.1).
		<p>4. Since basal estradiol synthesis is very low and ,as indicated through the text, it is difficult to evaluate inhibitors of estradiol synthesis..</p>	This was considered – and explored - previously but it has been decided to avoid competitive inhibition experiments with forskolin stimulated cells because of the

			risk of not being able to identify weak inhibitors and/or inducers
		<p>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.</p>	<p>Point a) of this reviewer (positive: media which have had no contact with cells supplemented with the chemical at the highest concentration tested) is already an integral part of the testing protocol. However, to emphasize the need for this “interference” testing a section has been added to the “Important Considerations” section of the H295R Exposure Protocol. Furthermore, a separate chapter on “Conduct of Chemical-Hormone Assay Interference Test” has been included following the suggestions of this referee (P19, Section 7.1).</p>
3	TS	<p>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</p>	<p>OECD member countries wanted to be able to choose the hormone analysis system. Not all laboratories have access to LC-MS and not all can use radioisotopes.</p>

		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.	
Topic: 2.5 Comments on Whether the Strengths and/or Limitations of the Assay Have Been Adequately Addressed			
1	SM	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.	Letcher et al, 1999 actually stated that these cell lines are "too sensitive to organochlorine" toxicity to make the useful to detect their effect on aromatase activity.
2	DR	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.	No response needed.
		The limitations of the assay are that: 1) the system does not allow to study complex interactions that occur in vivo, 2) the	We agree; however, the limitations of the assay will be offset by other assays in 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.	battery.
3	TS	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.	An extended discussion if this has been included into the Summary and Conclusions section of the report (Section 12). We agree that there is the possibility that the assay may capture or is influenced by pathways other than strictly gonadal steroidogenesis. However, we feel that the response to the chemicals used in validation study that was highly comparable to those observed with other gonad based in vitro and in vivo tests demonstrate the validity of the H295R Steroidogenesis Assay as a test to identify inducers and inhibitors of sex steroid synthesis. Finally, there are other assays in the test battery that are part of the weight of evidence approach of EDSP's Tier I testing.
		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	See response to this reviewer's comments above with regard to the relevance of the findings obtained with the H295R Steroidogenesis Assay.

		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).	
4	MV	The strengths have been addressed adequately but the limitation of the assay requires mention (see pg. 20 line 3 onwards).	A discussion of these limitations has been included into the “Conclusions and Summary“ section of the report. See also responses below.
		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.	The data obtained from the H295R assay will always be interpreted as a part of the larger battery.
		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 is being investigated and the protocol modified, if appropriate.

		This requires testing perhaps by supplementing the medium with cholesterol.	
	Topic: 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		
1	SM	Yes, there is in general a good choice of different chemicals, analytical and statistical methods.	No response needed.
		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 β -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).	The information in the tables has been updated as suggested by this reviewer, and the appropriate references have been included in the report.
2	DR	The test substances, and analytical and statistical methods chosen were appropriate to validate the assay.	No response needed.
		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...."	A section on "Data Processing and Reporting , and Statistics" has been included in the H295R Exposure Protocol (P.22, Section 8).
3	TS	<i>Choice of chemicals:</i> The compounds selected appear to be largely appropriate for	Information has been updated and errors corrected. As for the inducing potential of

		<p>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.</p>	<p>trilostane for T please refer to the discussion provided in the report (Section 9.2.3).</p>
		<p><i>Analytical methods:</i> 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</p>	<p>This is not entirely true. While some kits may show this extent of cross-reactivity, all kits used in this study (four different antibody based kits were used by the groups) did not reveal such cross-reactivities. Also, the comparability of the results demonstrated the validity of the assay, regardless of the possible cross-reactivities that may have occurred for individual assays. However, a section on the Requirements for</p>

		<p>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.</p> <p>Trilostane which is a steroid with an androgen base structure, not surprisingly, interferes with the testosterone immunoassay kits.</p>	<p>the “Performance of Hormone Detection Systems” has been included to Appendix I of the H295R Exposure Protocol (P28) providing QC guidance for the specific criteria each hormone assay has to fulfill before it can be considered for use in the H295R Assay.</p>
		<p>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.</p> <p>If immunoassay based analytical methods will be continued to be used all these pitfalls will need to be addressed.</p>	<p>This is exactly the reason why the protocol requires testing for the interference of these samples with the hormone assay. We are currently conducting some analyses of environmental samples, and after these are completed we will be able to answer the question whether it is likely that such samples interfere with hormone assays or not. As for individual chemicals, we could demonstrate that there are only a very few compounds that do so (even EE2, which is the synthetic analog to E2 , does not cross-react with our E2 AB). It is our opinion that the assay as it has been revised adequately addresses these issues.</p>
		<p><i>Statistical methods:</i></p> <p>Statistical testing hypotheses have not been explicitly defined. See also comments under point 3e.</p>	<p>See response 2 (same reviewer) to section 2.4.4 above.</p>

4	MV	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).	The data obtained from the H295R assay will always be interpreted as a part of the larger battery.
		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.	Experiments with AhR receptor agonists such as TCDD had no specific effects 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). Thus, these compounds were not considered of priority with regard to the validation of the assay. Exposure experiments with increasing doses of DMSO have been conducted during the early phase of the development of this assay. These studies found no direct effects on the production of T and E2 at increasing concentrations of this solvent but reported decreases in cell viability at concentrations greater or equal to 1% DMSO.
		The analytical methods and the statistical methods are appropriate to demonstrate the performance.	No response needed.
Topic: 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			
1	SM	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	No response needed.

		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.	This is being investigated and the protocol modified, if appropriate.
2	DR	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.	See responses to this reviewer’s comments in the subsequent section.
3	TS	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.	While we agree that some of the intra laboratory CVs were rather high the results obtained in response to all chemicals tested were – with very few exceptions – very comparable in terms of type and sensitivity of the response. EPA is not mandating the use of a particular hormone detection systems because OECD member countries have different constraints with respect to the types of systems they can use or which are preferred. For example, there are constraints in some countries with regard to the use of radioisotopes.
		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	This is being investigated as a result of the peer review comments. The protocol will be modified, if appropriate.

		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.	
4	MV	Foe the most part the assay is sufficiently repeatable and reproducible.	No response needed.
		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 forskolin between the different labs (Table 10.3).	This is being investigated and the protocol modified, if appropriate. The stimulation with forskolin was considered – and explored - previously but it has been decided to avoid competitive inhibition experiments with forskolin or cAMP stimulated cells because of the risk of not being able to identify weak inhibitors and/or inducers.
		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.	While we agree that some of the intra laboratory CVs were rather high the results obtained in response to all chemicals tested were – with very few exceptions – very comparable in terms of type and sensitivity of the response. Furthermore, CVs were calculated based on data obtained during two to three repeat experiments.
	Topic: 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															
1	SM	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.	The selection of the acceptable range for the SC within the linear portion was selected such that it still would allow for the precise determination of a 3- to 5-fold increase or decrease in hormone production after exposure to inducers or inhibitors would be possible. While we agree that a more stringent range would be preferable this would be difficult to implement for routine testing in high-throughput settings.												
2	DR	The performance criteria are adequate and would allow the assay to be performed at multiple laboratories without major problems.	No response needed.												
		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.	Tables have been revised and are now identical.												
3	TS	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.	<div>As stated above all tables have been checked for consistency and were corrected accordingly. The performance criteria now read as follows throughout the text:</div> <table><tr><th></th><th>Testosterone</th><th>Estradiol</th></tr><tr><td>Basal Production</td><td>≥ 2.5-times MDL</td><td>≥ 2.5-times MDL</td></tr><tr><td>Induction (10uM forskolin)</td><td>≥ 2-times SC</td><td>≥ 10-times SC</td></tr><tr><td>Inhibition (3uM</td><td>≤ 0.5-times SC</td><td>≤ 0.5-times SC</td></tr></table>		Testosterone	Estradiol	Basal Production	≥ 2.5-times MDL	≥ 2.5-times MDL	Induction (10uM forskolin)	≥ 2-times SC	≥ 10-times SC	Inhibition (3uM	≤ 0.5-times SC	≤ 0.5-times SC
	Testosterone	Estradiol													
Basal Production	≥ 2.5-times MDL	≥ 2.5-times MDL													
Induction (10uM forskolin)	≥ 2-times SC	≥ 10-times SC													
Inhibition (3uM	≤ 0.5-times SC	≤ 0.5-times SC													

			<i>prochloraz</i>)	
		<p>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 mean 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.</p>	<p>The purpose of the assay is to flag a chemical as a potential inducer or inhibitor of E2 or T, and not to identify the comparability of its MOA with that of a model compound. As the QC plate with its model inducer and inhibitor reflects the conditions of the cells at the time during which an experiment was conducted, in our opinion it represents a valid reference for each experiment to which the results can be compared. Although the assay may identify chemicals as false positives, these may be acceptable because the data will be interpreted and balanced in light with the other assays in the EDSP battery. With regard to the PC approach, this evaluation tool does not aim to predict a specific mode of action but rather provides a basis for the consistent estimate of the strength of an effect under the specific condition under which the assay was performed.</p>	
		<p>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</p>	<p>The reviewer has a valid point here.</p>	

		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.	Trenbolone was not substituted for forskolin in the assay. This will be clarified in the final report.
4	MV	The test substances chosen were appropriate to demonstrate the performance of the assay (forskolin and prochloraz as inducer and inhibitor, respectively).	No response needed.
		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 hydroxycholesterol 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.	This is being investigated and the protocol modified, if appropriate.
		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.	Testing inhibition by adding stimulatory compounds such as forskolin or cAMP was considered – and explored – previously. However, it has been decided to avoid inhibition experiments with stimulated cells

			because of the risk of not being able to identify weak inhibitors and/or inducers.
		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 aim of the assay is to identify inducers/inhibitors of sex steroid production, not corticosteroids. Also, the cells have been reported previously to not respond to ACTH stimulation.
		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.	Although CV's above 20% seem high, this was the performance data obtained during the validation program and the laboratories were able to distinguish among inhibitors, inducers and inactive chemicals.
		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.	This was corrected to through ">10-times" throughout all documents.
		Pg. 44. 9.1.1.1. line 5, change to Lab 5	Corrected.
		Table 9.1 – change "second" Lab 4 to Lab 5.	Corrected.
Topic: 2.9 Comments on Whether the Data Interpretation Criteria are Clear, Comprehensive, and Consistent with the Stated Purpose			
1	SM	Yes, However care must be taken when extrapolating results from in vitro to in vivo effects, see additional comments in point 9.	No response needed.
2	DR	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.	As stated above, a section on "Data Processing and Reporting, and Statistics" has been included in the H295R Exposure Protocol (P22, Section 8).
		From the extensive preliminary report addressing several analysis techniques based on the data generated with the core	No response needed.

		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.	This has been done in the final protocol. The data interpretation procedure will use statistically significant difference in fold change at the criterion for a positive result.
3	TS	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.	We agree. As noted above, the data interpretation procedure has been simplified to inducer (statistically significant increase in fold change over control), inhibitor (statistically significant decrease compared with control) or inactive (no significant difference from control).
		Expressing results using the PCmax/PC50 concept is, as mentioned under point 7, not likely to be very useful.	See previous response to this comment (Section 2.8, 2 nd reviewer (TS), 2 nd Comment).
4	MV	The data interpretation is clear and consistent with the objective of the report.	No response needed.
		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.	Agree, see response to comment 3 above.
		For instance the lack of response (or weak response) may be due to the high basal hormone production in the case of testosterone	This is being investigated and the protocol modified, if appropriate.

		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.	For primary purpose of this assay, screening of chemicals with unknown endocrine toxicity and potency, this will not be feasible because of the risk of missing relevant concentrations. We agree that refinement experiments should then be conducted with tighter dose-ranges around the active concentrations. It is to be decided by the EDSP whether this is desired or not.
	Topic: 2.10 Please Comment on the Overall Utility of the Assay as a Screening Tool in the EDSP Tier 1 Battery		
1	SM	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.	We agree.
		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 xenobiotics 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	Again, this is true but beyond the scope of this assay. That is the reason why this assay is only to be used in context with a screening battery that than in its entity will hopefully be able to address the question whether in vivo effects are likely.

		middle or moderate for female since male and female have not the same sensitivity toward endogenous and therefore altered estrogen.	
		<p>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.</p> <p>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.</p>	These are basic research questions that go beyond the Agency's needs for a screening assay.
		<p>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</p>	These are basic research questions that go beyond the Agency's needs for a screening assay.

		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.	
		<p>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</p>	This is being investigated and the protocol modified, if appropriate.

		Draft report “Standardization and refinement of the H295R cell ...” page 47 section 10.3.	
2	DR	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 and 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.	No response needed.
3	TS	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 it 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	See previous comments to these comments raised by this reviewer (e.g. Section 2.2, 3 rd Reviewer (TS)).

		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.	<p>Response to 1): Using labor and cost intensive techniques such as LC/MS is not feasible for large scale screening purposes. For the validity of antibody based assay for the analysis of T and E2 as described in the protocols please refer to our responses above (e.g. Section 2.4.3, 2nd Comment (TS)).</p> <p>Response to 2): This is being investigated and the protocol modified, if appropriate.</p> <p>Response to 3): The purpose of the assay is to flag a chemical as a potential inducer or inhibitor of E2 or T, regardless of the specific steroidogenic pathway. However, we agree that the inclusion of the corticosteroid hormones would provide useful information for the interpretation of the results in addition to the screening purpose of the assay.</p>
		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	<p>As could be demonstrated in the validation studies, the H295R Steroidogenesis Assay successfully detected inducers and inhibitors of T and E2 production that were in accordance among laboratories and with effects previously reported in the literature for the tested chemicals.</p> <p>With regard to the use of antibody based</p>

		<p>ambiguous information about the steroidogenesis disruption potential of chemicals or unknown environmental extracts.</p>	<p>hormone detection systems: For this reason QC criteria were defined. Most of the labs that participated in this study used different hormone detection assays (RIA, ELISA, Time Resolved Fluorescence, LC-MS), because different methods are preferred or accessibility is limited by various labs around the world. EPA cannot recommend a specific assay. Also, there are restrictions in the use of tests utilizing radioisotopes, and mass spectrometry based technologies are not trivial and feasible for high throughput tests. For this reason QC criteria were defined that are applicable to all methods. While this may have been the source for some of the variation among labs, it should be acknowledged here that regardless of the assay used always the same type of response was observed. The definition of the purpose of the assay has been changed as follows: “The H295R Steroidogenesis Assay is intended to identify xenobiotics that target intracellular components that comprise the steroidogenic pathway beginning with the sequence of reactions occurring after the gonatotropin hormone receptors (FSHR and LHR) through the production of testosterone and estradiol/estrone. The steroidogenic assay is not intended to identify substances that</p>
--	--	---	---

			affect steroidogenesis due to effects on the hypothalamus or pituitary gland.”
4	MV	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.	We agree, but this would be out of scope of the assay. As stated above, the H295R Steroidogenesis Assay will be utilized in junction with a battery of other assays, and it is not aimed to identify complex interactions of chemicals with all aspects of the hypothalamus-pituitary-gonad axis.
		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.	This is being investigated and the protocol modified, if appropriate. The stimulation with forskolin was considered – and explored - previously but it has been decided to avoid competitive inhibition experiments with forskolin or cAMP stimulated cells because of the risk of not being able to identify weak inhibitors and/or inducers.
		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.	While we agree that some of the intra laboratory CVs were rather high the results obtained in response to all chemicals tested were – with very few exceptions – very comparable in terms of type and sensitivity

			<p>of the response. The primary objective of this assay is to flag chemicals as potential inducers and/or inhibitors of T and E2 production in a Tier I battery, and not to provide highly quantitative and mechanistic insights. As a semi-quantitative or non-quantitative assay, therefore, it is our opinion that the H295R Steroidogenesis Assay is highly reproducible as could be demonstrated by the findings across laboratories presented in this study. See also response to Section 2.7, 2nd reviewer (TS), 1st Comment.</p> <p>This is being investigated and the protocol modified, if appropriate.</p>
		Overall, the assay has the potential to be a screening tool for steroidogenesis but requires further testing and refinement.	No response needed.

Peer Review Panel Members:

SM = Safa Moslemi
DR = Damian Romero
TS = Thomas Sanderson
MV = Matt Vijayan