

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 34 industrialised countries in North and South America, Europe and the Asia and Pacific region, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in ten different series: **Testing and Assessment; Good Laboratory Practice and Compliance Monitoring; Pesticides and Biocides; Risk Management; Harmonisation of Regulatory Oversight in Biotechnology; Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; and Safety of Manufactured Nanomaterials.** More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (www.oecd.org/ehs/).

This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organisations.

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The Participating Organisations are FAO, ILO, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD. UNDP is an observer. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

This publication is available electronically, at no charge.

**For this and many other Environment,
Health and Safety publications, consult the OECD's
World Wide Web site (www.oecd.org/ehs/)**

or contact:

**OECD Environment Directorate,
Environment, Health and Safety Division
2 rue André-Pascal
75775 Paris Cedex 16
France**

Fax: (33-1) 44 30 61 80 E-mail: ehscont@oecd.org

E-mail: ehscont@oecd.org

FOREWORD

This document presents the Peer Review Report (PRR) of the validation of the *Androgen Receptor Mediated Stably Transfected Transcriptional Activation (AR-STTA) assay to detect androgenic and anti-androgenic activities of chemicals*. It also includes a statement of the Working Group of National Coordinators of the Test Guidelines Programme on the follow-up to the PRR.

The project for developing a Test Guideline for an AR STTA assay was proposed by Japan and included in the work plan of the Test Guidelines Programme in 2008. A draft validation report was submitted to the Validation Management Group for non animal testing in December 2010. The PRR was endorsed by the Working Group of National Coordinators of the Test Guidelines Programme at its meeting held on 12-14 April 2011. The Joint Meeting of the Chemicals Committee and Working Party on Chemicals, Pesticides and Biotechnology (Joint Meeting) agreed to its declassification on 5 October 2011.

This document is published under the responsibility of the Joint Meeting.

Agreement of the Working Group of the National Coordinators of the Test Guidelines Program on the Follow-up to the Validation Peer Review Report

The Validation Peer Review Report of the *Androgen Receptor Mediated Stably Transfected Transcriptional Activation (AR-STTA) Assay to Detect Androgenic and Anti-androgenic Activities of Chemicals* was submitted for endorsement to the Working Group of National Coordinators of the Test Guidelines Program (WNT) at its April 2011 meeting.

Considering the major recommendations of the Peer Review Panel (summarized below), i.e.:

- a) A dedicated inter-laboratory study should be carried out, using the final test protocol to test substances covering a broad range of activity, especially including non-active substances and weak agonists and antagonists. The number of substances already tested (five test chemicals) in the inter-laboratory validation, and the affinity range that they cover, is not sufficient;
- b) The following discussion points should be added to the validation report:
 - a. advantages of the AR-STTA assay over similar AR activation assays (i.e., lack of Glucocorticoid receptors in this cell line eliminates cross-talk with AR, and more discussion of positive results in AR-STTA that are negative in AR binding assays),
 - b. potential interference of partial agonists with antagonist effects, and proposed solutions to elucidate such interference,
 - c. potential impacts of differences between protocols used for the pre-validation and the inter-laboratory validation studies,
 - d. the lack of a cytotoxicity measurement in the agonist assay, which masks identification of true negatives from false negatives;
- The protocol should be revised to:
 - e. add acceptance and assessment criteria for the positive control (5 α -Dihydrotestosterone (DHT)),
 - f. precisely define the decision criteria for classification, especially considering cytotoxic effects (e.g. introduce the option of equivocal/not conclusive results, since cytotoxicity can interfere with the detection of androgenic and especially anti-androgenic responses),
 - g. Explore the biological and statistical appropriateness of the PC10 in more detail,
 - h. Include a list of proficiency chemicals for both the androgenic and the anti-androgenic assay,

the WNT agreed that, before finalizing the development of the draft Test Guideline for an AR-STTA assay,

- The Validation Management Group for non-animal testing should address the above recommendations as appropriate, in particular the recommendation to test more substances in a new inter-laboratory validation, while ensuring a good balance of substances with androgenic and anti-androgenic activity, negative and positive control substances;
- The cell line should be made freely available.

**SUMMARY REPORT OF THE PEER REVIEW PANEL
ASSESSING THE
ANDROGEN RECEPTOR MEDIATED STABLY TRANSFECTED TRANSCRIPTIONAL
ACTIVATION (AR-STTA) ASSAY
TO DETECT ANDROGENIC AND ANTI-ANDROGENIC ACTIVITIES**

Prepared for

**Environment Directorate
Organisation for Economic Co-operation and Development (OECD) Paris**

March 03, 2011

PREAMBLE

This document presents the summary report of the assessment of the validation of the Androgen Receptor Mediated Stably Transfected Transcriptional Activation (AR-STTA) *in vitro* assay to detect androgenic and anti-androgenic activities of chemicals by an independent Peer Review Panel.

The AR-STTA assay utilises the AR-EcoScreen™ cell line, which is derived from a Chinese hamster ovary derived cell line (CHO-K1). Three stable constructs have been inserted in this cell line. The constructs are a human AR expression construct (encoding the full-length human receptor), a firefly luciferase reporter construct bearing four tandem repeats of a prostate C3 gene-responsive element driven by a minimal heat shock protein promoter and a constitutive Renilla luciferase expression vector.

The assay is used to detect changes in the activity of the androgen receptor. The activated receptor translocates into the nucleus where it binds to specific DNA response elements and transactivates a luciferase reporter gene, resulting in increased cellular expression of luciferase enzyme. Luciferin is a substrate that is transformed by the luciferase enzyme to a bioluminescence product that can be quantitatively measured with a luminometer.

In order to better understand the purpose of the assay, the AR-STTA has to be considered in view of the OECD conceptual framework for testing and assessment of potential endocrine disrupting chemicals. This framework comprises the five levels

- sorting & prioritization based upon existing information (Level 1)
- *in vitro* assays providing mechanistic data (Level 2)
- *in vivo* assays providing data about single endocrine mechanisms and effects (Level 3)
- *in vivo* assays providing data about multiple endocrine mechanisms and effects (Level 4)
- *in vivo* assays providing data on effects from endocrine & other mechanisms (Level 5).

As a transcriptional activation (TA) assay providing mechanistic data, the AR-STTA is a Level 2 assay.

While several TA and receptor binding assay are currently about to enter or are undergoing validation exercises, only the “Stably Transfected Transcriptional Activation (TA) using HeLa-9903 cell line for detecting estrogenic activity of chemicals”, the AR-STTA assay counterpart for estrogenic activities, has been adopted as an OECD test guideline (TG 455) in 2009. Indeed, the AR-STTA assay is the first androgenic activity detecting assay undergoing OECD peer review.

Summary Report of the Peer Review Panel assessing the Androgen Receptor Mediated Stably Transfected Transcriptional Activation (AR-STTA) assay to detect androgenic and anti-androgenic activities of chemicals

The peer review process

1. The Peer Review Panel (Panel) was formed in December 2010 in order to provide an independent review of the validation of the Androgen Receptor Mediated Stably Transfected Transcriptional Activation (AR-STTA) assay to detect androgenic and anti-androgenic activities of chemicals.

The work of the Panel was coordinated by a panel chair contracted by the OECD. In addition to experts invited by the Secretariat, potential Panel members were nominated by the Working Group of the National Coordinators of the Test Guidelines Programme (WNT) and then approached by the chair. The members of the Panel are listed in Annex 1.

Furthermore, Dr. Masahiro Takeyoshi, from the Chemicals Evaluation Research Institute (CERI), Japan, who is familiar with the AR-STTA, was nominated to support the panel in case of open issues needing clarification.

2. The Panel was asked to evaluate the data collected on the test method, and to answer specific charge questions. These questions addressed the eight OECD validation criteria set out in the OECD Guidance Document. Panel members were asked to base their review on the document 'Draft Report of Pre-validation and Inter-laboratory Validation For Androgen Receptor (AR) Mediated Stably Transfected Transcriptional Activation (AR-STTA) Assay to Detect Androgenic and Anti-androgenic Activities', provided as the file '100903_AR_Validation_Report_FINAL.doc'.

As background information, they were also provided with the OECD Guidance Document on the Validation and International Acceptance of New or Updated Methods for Hazard Assessment, Series on Testing and Assessment, Number 34, 2005 (last access on February 14, 2011, under [http://www.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono\(2005\)14&doclanguage=en](http://www.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono(2005)14&doclanguage=en)).

The charge to the Panel was to assess to what extent the eight OECD validation criteria set out in the OECD Guidance Document had been met. The charge questions as provided to the Panel are listed in Annex 2. A summary of the Panel's responses to the individual questions is presented in paragraphs 6 to 44. For transparency, the individual comments from the Panel members are provided anonymously and in an edited form in Annex 3.

3. During the review process, the Panel held two teleconferences (December 16, 2010, and January 28, 2011) which were organised and coordinated by the chair. Furthermore, Dr. Masahiro Takeyoshi was contacted twice (January 17 and 28, 2011) in order to provide clarifications on issues raised by members of the Panel. The reply to the first inquiry, which comprised three issues, is provided in Annex 4a. This reply was discussed during the second teleconference. In this conference, further clarification on two aspects was requested. Again, Dr. Masahiro Takeyoshi was asked to provide help. The questions and the respective reply are provided in Annex 4b.

Each Panel member provided written responses on the charge questions to the Panel chair by February 11, 2011. Based on these responses, a draft report taking into account all individual comments was compiled by the chair and provided to the Panel for review and comments (February 15, 2011). Furthermore, it contained some clarifications on specific issues, which were added by the chair. The Panel commented on the draft report until February 22. Accounting for this feedback and resolving remaining open issues, the

final report was drafted by the chair and send to the PRP for approval on February 24. This report presents the resulting approved responses of the Panel to each of the charge questions.

General Panel responses

4. The reviewers acknowledged the work which has been carried out in order to validate the AR-STTA. The following general editorial issues were raised:
- The report should be updated to the current general situation, e.g. REACH came into force in 2007 (see paragraph 18).
 - Tests under Reprotect are pre-validated only (see paragraph 18).
 - Unnecessarily repetitions should be removed (see e.g. paragraph 30).
 - The report should be edited (e.g. paragraph 78 is incomplete; see also the replies to charge question 3 in Annex 3).
 - The report title indicated that the report is a draft, while the name of the provided file suggested a final version. It was assumed, that the report title was not up-to-date.

Panel responses to the charge questions: The eight OECD principles and criteria for test method validation

5. The PRP reached consensus on all eight charge questions.

Charge question 1: A rationale for the test method should be available, including a clear statement of scientific basis and the regulatory purpose and need for the test method

6. The Panel agreed that this criterion had been met. The rationale for the test method is clearly stated with regard to the scientific basis and regulatory purpose. In addition, it was commented that the assay has been well-established and that the approach (using the Renilla luciferase reporter) is especially suitable.

Charge question 2: The relationship between the test method endpoint(s) and the biological effect and to the toxicity of interest should be described, addressing also limitations of the test methods

7. The Panel agreed that this criterion has been partly met.
8. Potential advantages of the AR-STTA assay as compared to similar assays should be described. In particular, more information on the lack of Glucocorticoid receptor (GR) in the cell line should be provided, especially in view of GR cross talk interference with AR, which has been described for other cell lines developed for the similar purposes. It is proposed that the AR-STTA may therefore be more robust favouring it over similar assays. Potential consequences when comparing AR-STTA assay results with results of other AR assays are addressed in paragraph 40.
9. Potential interference of partial agonists with antagonist effects has not been addressed and should be discussed, eventually proposing solutions to elucidate such interference.
10. As the transcriptional activation of endogenous genes *in vivo* might involve additional factors and mechanisms, an important limitation of the AR-STTA is that not necessarily all transcriptional responses *in vivo* are reflected. This aspect should be described.

11. The fact that the assay will not detect effects mediated by AR mechanisms that are independent of direct interactions with DNA / AREs, e.g. activation of intracellular signalling pathways by membrane-associated AR, should be addressed. In this context, the Renilla system should be discussed in the light of recent findings published in peer-reviewed literature. Firefly luciferase specific inhibitors that affect the catalytic activity of the enzyme have been described. In addition, chemicals might stimulate post-transcriptional effects, especially on protein stability, as has been shown for the “superinduction” of luciferase luminescence by genistein. In particular, AR-independent effects might have played a role in the effects observed for fluoranthene and medroxyprogesterone.

12. The robustness of viral promoter, such as SV40, should be discussed. As a potential consequence, mild cytotoxic effects might not be detected in the Renilla or firefly Luciferase cytotoxicity control.

13. A major limitation of the agonist assay is the lack of cytotoxicity measurements. As a consequence, true negative chemicals can hardly be discriminated from false negative due to masking by cytotoxicity, i.e. cytotoxicity is observed at concentrations below the concentration needed to elicit an agonistic response. A discussion of this aspect should be added, especially focusing on hydroxyflutamide and the use of the cytotoxicity data from the antagonist assay for the interpretation of the agonist assay.

14. Finally, limitations related to the substance Methyltrienolone (R1881) were identified. The availability of such an anabolic steroid might be very restricted due to potential import restrictions. Therefore, it does not qualify as a control, reference or proficiency chemical. As a potential replacement the substance mestanolone (MDHT) could be considered. Furthermore, the activity of R1881 might be discussed in view of the work of Hanson *et al.* (2007) in ‘Frontiers in Bioscience 12, (1387-1394)’.

Charge question 3: A detailed protocol for the test method should be available

15. The Panel agreed that this criterion had been partly met.

16. Assessment/Acceptance criteria for the positive control DHT are lacking and should be added. In this context, it is suggested to include a graph of the expected DHT concentration-response curve in the protocol.

17. Regarding the data analysis, the applied approaches are likely to be unfamiliar to test laboratories. More guidance on the data analysis, e.g. by making a validated data analysis template publicly available, should be provided. In addition, it was recommended that in case of incomplete concentration-response curves, approximations of inhibiting and effective concentration should be clarified and improved. Furthermore, it should be clarified if other software besides GraphPad can be used.

18. The protocol used in the between-laboratory study differed from the protocol used in the pre-validation study. Nevertheless, the data have been combined for analysis. The potential impacts of protocol differences should be discussed, especially demonstrating, possibly supported by data, why these differences allow data combination. In particular, an explanation why the Renilla reporter was not used in the pre-validation should be provided.

19. The AR-EcoScreen cell line and especially the construct that was stably transfected to allow constitutive expression of the Renilla luciferase should, if possible, be described in more detail.

20. As substances were dissolved in ethanol, ethanol should be included in the protocol as a potential solvent.

21. It is recommended to propose proficiency chemicals for both the androgenic and the anti-androgenic assay.

22. The decision criteria to conclude on the classification should be more clearly defined. In this respect, cytotoxic effects should be considered in more detail. Since cytotoxicity can interfere with the detection of androgenic and especially anti-androgenic responses (see e.g. results of ketoconazole, DDT and methoxychlor) it is proposed to account for this, e.g. by introducing the option of equivocal/not conclusive results or by considering such substances as being outside of the applicability domain of these assays.

23. Testing concentration up to 10µM may not allow detection of weakly active compounds. Therefore testing up to higher concentrations, up to the solubility limit and up to the onset of cytotoxicity should be considered.

Charge question 4: Within- and between-laboratory reproducibility of the test method should be demonstrated

24. The Panel agreed that this criterion had been partly met.

25. Regarding the between-laboratory reproducibility, the amount of substances and the affinity range covered by these substances was not appropriate. Assuming that potential candidate substances are available, more substances should be tested. Additional substances should cover a broad affinity range, especially including weak agonists and antagonists as well as negative substances, and substances with incomplete concentration-response curves. Some reviewers signalled willingness to support the selection of such substances.

26. An explanation should be provided why data on cytotoxicity for the substance R1881 are lacking.

27. Different concentrations of test substances have been tested in the pre-validation and the inter-laboratory study. An explanation for this difference should be provided.

Charge question 5: Demonstration of the test method's performance should be based on testing of representative, preferably coded reference chemicals

28. The Panel agreed that this criterion had been partly met.

29. Too few substances that did not appropriately cover the entire range of possible responses of the AR-STTA assay were tested for between-laboratory reproducibility. For example, in the agonist assay only the structurally similar substances DHT, i.e. the positive control, and R1881 have been tested in several laboratories. Re-enforcing the recommendations of paragraph 24, the panel is of the opinion that it will be inevitable to test additional carefully selected substances in a dedicated inter-laboratory study. In this regard it should be noted that this potential shortcoming has already been identified in paragraph 15 of the reviewed validation report '... though the number of chemicals used under multi-laboratory validation study may not be sufficient.' Furthermore, additional testing is considered necessary in order to allow proposing a list of proficiency chemicals, as proficiency chemicals should be supported by a sound and broad database.

In regard of the number of chemicals to be tested, the general issue of the in practice less stringent requirements for *in vivo* assays than *in vitro* methods was raised. For example, while for the inter-laboratory reproducibility assessment of an *in vivo* assay a limited number of chemicals might be

considered sufficient, usually a higher number of chemicals is expected for the inter-laboratory reproducibility assessment of an *in vitro* assay. This difference is considered to be driven essentially by animal welfare consideration and, if at all, only marginally by scientific reasoning. As the scientific requirements for *in vitro* and *in vivo* test should be similar, but in practice are not, this constitutes a fundamental obstacle for an objective review of validation studies in a general framework that equally addresses *in vivo* as well as *in vitro* tests.

Charge question 6: The performance of test methods should have been evaluated in relation to existing relevant toxicity data as well as information from the species of concern

30. The Panel agreed that this criterion had been partly met.
31. AR-binding data should be discussed in more detail, especially focusing on substances with positive responses in the AR-STTA assay that are either not detectable in the *in vitro* binding assay for technical reasons (comparison with published binding data), e.g. for vinclozolin or nonylphenol, or mediated by other mechanisms than direct AR-binding. The presented discussion is not convincing due to differences in maximum concentrations used in these assays.
32. Inconsistencies in the presented data have been found, e.g. in Paragraph 116 and Table 14 it is stated that 4-Hydroxytamoxifen was, despite significant binding to the AR receptor, not detected as an AR agonist or AR antagonist. However in Table 8, it is listed as an antagonist in the AR-STTA assay.
33. The ICCVAM list is several years old and consequently not necessarily up-to-date. Furthermore, some classifications are questionable, especially for antagonists. It is recommended, if possible, to update the results presented here, e.g. when the ICCVAM list is revised.
34. The assay is not discussed in relation to the Hershberger assay. Comparison with the Hershberger assay and/or a discussion of problems in the comparison to the Hershberger is proposed.
35. The biological relevance of PC10, especially in combination with no concentration-response, as well as its statistical significance is questioned. The use of a more pronounced PC that potentially better balances biological relevance and statistical significance should be explored in more detail, taking into account that comparison with the ICCVAM list and with AR binding data needs further discussion (see paragraphs 8, 31, 33 and 41).
36. The approach taken to calculate EC50 and IC50, i.e. irrespective of the completeness of the concentration-response curve, is questioned. Calculation based on comparison to the concentration-response of the positive control (i.e. DHT) expressed as percentage of control should be explored.
37. For the antagonist assay $lin.IC_{30}$ was used in the validation study. The reason for using $lin.IC_{30}$ is not clear, as, e.g. in Table 11, no differences between $lin.IC_{50}$ or $lin.IC_{30}$ are apparent.
38. The presentation of the data, data analyses and interpretation should be improved and be consistent throughout the report. In particular, the information if data are based on single or several experiments should be provided. Furthermore, it should be clarified if the agonist and antagonist experiments were performed in parallel or independent of each other.
39. Negative classification in the agonist assay because of insufficient induction of the luciferase activity, although EC50 values can be calculated, should be discussed in more detail considering that low inductive activities at low concentrations could still be biologically significant.

40. It is proposed that the lack of GR (see paragraph 9) might explain some classification discordances in the comparison with the ICCVAM list, e.g. for corticosterone, potentially favouring the results of the AR-STTA assay.

Charge question 7: All data supporting the assessment of the validity of the test method should be available for expert review

41. The Panel agreed that this criterion had been met. However, for clarity and ease of evaluation the presentation of PC-, IC- and EC-values should be improved. Furthermore, some additions are suggested (i.e. a list of the substances and the respective primary data used to generate Table 9, Figure 9, Table 11 and Figure 11).

Charge question 8: Ideally, all data supporting the validity of a test method should have been obtained in accordance with the principles of Good Laboratory Practice (GLP)

42. The Panel agreed that this criterion had been partly met.

43. The pre-validation was not conducted according to GLP. However, this is in general considered acceptable. In this context, it was also mentioned that the fact that the pre-validation data have been published in a peer-reviewed journal should be considered.

44. It was mentioned that elsewhere, i.e. not in the review document, the information that the pre-validation was conducted according to GLP principles or in the spirit of GLP has been provided. Therefore, it is recommended to include this information in the validation report.

Recommendations

45. The Panel agrees that this report provides a summary of their views on the status of the validation of the Androgen Receptor Mediated Stably Transfected Transcriptional Activation (AR-STTA) assay to detect androgenic and anti-androgenic activities of chemicals, as detailed in the responses to the questions posed to the Panel and based on the information related to the test method validation provided to the Panel.

46. The report of the Panel, along with the provided review document on Androgen Receptor Mediated Stably Transfected Transcriptional Activation (AR-STTA) assay should form the basis for decisions on whether the validation meets the OECD principles for validation. The Panel recommends that the OECD considers the Panel report as guidance for recommending additional work required to fully meet all OECD principles.

47. Future work should focus especially on the following aspects. Most importantly, it is strongly recommended

- c) to carry out a dedicated inter-laboratory study testing substances covering a broad range of activity, especially including non-active substances as well as weak agonist and antagonists with the final test protocol (see paragraphs 25 and 29).

Furthermore, discussion of several issues, which so far have not or only partially been addressed, should be added to the validation report. These issues are

- d) potential advantages of the AR-STTA assay over similar AR activation assays (see paragraphs 8 and 31),
- e) the possibility of interference of partial agonists with antagonist effects, eventually proposing solutions to elucidate such interference (see paragraph 9),

- f) the potential impact of differences in the protocols of the pre-validation study and the inter-laboratory study (see paragraph 18),
- g) the lack of a cytotoxicity measurement in the agonist assay (see paragraph 13).

Regarding the protocol, the major recommendations were:

- h) to add acceptance and assessment criteria for the positive control DHT (see paragraph 16) and
- i) to more precisely define the decision criteria for classification, especially considering cytotoxic effects, e.g. by introducing the option of equivocal/not conclusive results, since cytotoxicity can interfere with the detection of androgenic and especially anti-androgenic responses (see paragraph 22).

In addition:

- j) the biological and statistical appropriateness of the PC10 should be explored in more detail (see paragraph 35)
- k) and a list of proficiency chemicals for both the androgenic and the anti-androgenic assay should be proposed.

Finally, the recommendations in the paragraphs 10, 11, 12, 14, 17, 19, 20, 23, 26, 27, 32, 34, 36-41 and 44 should be addressed.

ANNEX 1

Members of the peer review panel

Panel member	Affiliation
Miriam Jacobs	European Food Safety Agency (EFSA), Parma, Italy
Tsutomu Nishihara	Hyogo University of Health Sciences, Kobe, Japan
Michael Oelgeschläger	Federal Institute for Risk Assessment (BfR), Berlin, Germany
Vickie Wilson	United States Environmental Protection Agency (US EPA), Research Triangle Park, North Carolina, US

ANNEX 2

Charge questions

PRP Charge questions:

The eight OECD principles and criteria for test method validation

1. **A rationale for the test method should be available, including a clear statement of scientific basis and the regulatory purpose and need for the test method.**
2. **The relationship between the test method endpoint(s) and the biological effect and to the toxicity of interest should be described, addressing also limitations of the test methods.**
3. **A detailed protocol for the test method should be available.**
The protocol should be sufficiently detailed and should include, *e.g.*, a description of the materials needed, such as specific cell types or construct or animal species that could be used for the test (if applicable), a description of what is measured and how it is measured, a description of how data will be analysed, decision criteria for evaluation of data and what are the criteria for acceptable test performance.
4. **Within- and between-laboratory reproducibility of the test method should be demonstrated.**
Data should be available revealing the level of reproducibility and variability within and among laboratories over time. The degree to which biological variability affects the test method reproducibility should be addressed.
5. **Demonstration of the test method's performance should be based on testing of representative, preferably coded reference chemicals.**
6. **The performance of test methods should have been evaluated in relation to existing relevant toxicity data as well as information from the species of concern.**
7. **All data supporting the assessment of the validity of the test method should be available for expert review.**
The detailed test method protocol should be readily available and in the public domain. The data supporting the validity of the test method should be organised and easily accessible to allow for independent review(s), as appropriate. The test method description should be sufficiently detailed to permit an independent laboratory to follow the procedures and generate equivalent data. Benchmarks should be available by which an independent laboratory can itself assess its proper adherence to the protocol.
8. **Ideally, all data supporting the validity of a test method should have been obtained in accordance with the principles of Good Laboratory Practice (GLP).**
Aspects of data collection not performed according to GLP should be clearly identified and their potential impact on the validation status of the test method should be indicated.

ANNEX 3

**Collated initial comments from the peer review panel
assessing the
Androgen Receptor Mediated Stably Transfected Transcriptional Activation (AR-STTA) Assay
to Detect Androgenic and Anti-androgenic Activities**

ISSUES	COMMENTS AND RECOMMENDATIONS
<p>General comments</p>	<p>Reviewer #1: The AR-STTA appears to be a suitable assay for the identification of androgenic as well as anti-androgenic activities. The Chemicals Evaluation and Research Institute (CERI) has done a tremendous amount of work testing a large number of different chemicals to provide evidence that this assay generates relevant information. Unfortunately the method was modified for the inter-laboratory validation study. During the validation study cytotoxicity was determined using the Renilla activity of the AR-EcoScreen™ cell line instead of the cLuc-EcoScreen™ cell used during pre-validation. In addition, different concentration ranges were tested in the pre-validation and validation studies. Although the rationale behind these modifications is explained and understandable, the sole purpose of a validation study is to test for the reliability and relevance of the method established during a pre-validation phase. In particular, the identification of acceptable variations of the test protocol should not be part of a validation study testing coded chemicals in multiple labs. In addition, the inter-laboratory study has a limited informative value since only 5 chemicals were tested. Multi-lab studies with coded chemicals are of central importance for the validation of a test method. This part of the validation needs to be improved and might require additional testing of coded chemicals. There are also additional points that need to be clarified, in particular: The description of the Renilla construct and its use as an endpoint for cytotoxicity (including the lack of cytotoxicity measurements in the agonist assay), the criteria for positive responses and some inconsistencies in the presented data.</p> <p>Reviewer #2, #3, #4: -</p>

VALIDATION CRITERIA	
<p>1. Rationale for the test Method</p>	<p>Reviewer #1: The rationale for the test method is clearly stated. The scientific background and its regulatory purpose are well described. From a regulatory perspective this method is intended to be used as a screening procedure to detect androgenic as well as anti-androgenic activities of chemicals and to provide mechanistic information as a level 2 in vitro assay. As minor point: The text does not reflect the current situation and should be updated, e.g. in paragraph 18 it is stated that “Reach is expected to enter into force in 2007”.</p> <p>Reviewer #2: The cell is appropriate, the AR construct is full length, and the renilla luciferase vector is an optimum vector to utilise for the AR. An AR ag/antagonist assay is very much needed by regulatory authorities globally; no validated equivalent is so far available. However editorially some of the paragraphs are out of date or redundant (e.g. paragraph 18: REACH entered into force some years ago now; the tests under Reprotect are ‘prevalidated’ only; or paragraph 30), and the report generally needs better editing for clarity in English. Paragraph 78 is incomplete.</p> <p>Reviewer #3: The rationale and a clear statement of the scientific basis, and the purpose and need are clearly identified within the text of the report.</p> <p>Reviewer #4: The technique used for the proposed assay system has been well-known and well-established in the variety fields. The regulatory purpose and needs were well documented in the validation report.</p> <p>The Panel agreed that this criterion has been met.</p>
<p>2. Relationship between the test method endpoint(s) and the biological effect and to the toxicity of</p>	<p>Reviewer #1: Limitations of the test method are not sufficiently described. The activation of a stably transfected human androgen receptor and its subsequent nuclear localisation results in binding of the ARE and activation of luciferase expression. This process represents the best understood mechanism of androgen activity, although the transcriptional activation of endogenous genes in vivo might involve additional factors and mechanisms. Thus, an important limitation of this test (like any test that is based on transcriptional activation of an artificial promoter in an “artificial” cellular</p>

<p>interest addressing also limitations</p>	<p>environment) is that it does not necessarily reflect transcriptional responses in vivo. In addition, the assay will not detect effects mediated by AR mechanisms that are independent of direct interactions with DNA / AREs, e.g. activation of intracellular signalling pathways by membrane-associated AR. For the antagonist assay, cytotoxicity is determined using a stably transfected Cho-K1 cell line that constitutively expresses firefly luciferase (cLuc-EcoScreen) or the determination of constitutive Renilla luciferase expression of the AR-EcoScreen cell line. However, for the purpose of the test system the Renilla system might generate problems. In particular, firefly luciferase specific inhibitors have been described in the literature (see Thorne et al. 2010 and references therein) that affect the catalytic activity of the enzyme. In addition, chemicals might stimulate post-transcriptional effects, in particular, on protein stability as has been shown for the “superinduction” of luciferase luminescence by genistein (Sotoca et al., 2010). In this respect, the comparison of regulated and constitutive firefly luciferase activity in distinct cellular systems, as done in the extensive pre-validations studies, might be more appropriate to identify and exclude AR unrelated effects. AR-independent effects might have played a role in the effects observed for fluoranthene. In the antagonistic assay, fluoranthene significantly stimulated cell viability (luciferase activity in the cLuc-EcoScreen cell line) at concentrations of 10⁻⁵ M (Fig. 10.2) that apparently also stimulated a positive response in the agonist assay (Fig.8). In addition the “superinduction” of luciferase luminescence by medroxyprogesterone (not tested in the antagonist assay) might reflect an ARE independent activity comparable to the direct effects of genistein on the firefly luciferase protein. A major limitation of the agonist assay is the complete lack of cytotoxicity measurements. Thus, it might be not always clear, if the chemical is a true negative, or if one cannot detect agonistic activities, because the substance is cytotoxic at concentrations below the concentration needed to elicit an agonistic response. This problem is described for hydroxyflutamide (paragraph 98) but should be discussed in more general terms, since it is not clear, if the cytotoxicity data from the antagonist assay could also be used for the interpretation of the agonist assay, if the tests are done in parallel.</p> <p>Reviewer #2: Some areas require improvement regarding potential limitations or the comparative lack of limitations compared to similar assays with the same endpoints, i.e. provide more clarifying information on the lack of GR in the cell line. This is a potential issue regarding metabolic capacity and with GR cross talk interference with AR in other cell lines developed for AR agonism/antagonism possibly being implicated in differences in low dose/higher/biphasic dose responses of ligands that are perhaps due to different receptor mechanisms occurring in the same cell with both agonistic or antagonistic activity in the AR at one dose and agonistic or antagonistic activity in the GR at another. Should this be the case it might explain lack of literature concordance with for e.g. corticosterone, and might perhaps suggest the AR STTA assay is more robust in this respect than other similar AR assays. In any case, it would be helpful to discuss the fact that partial agonists can interfere with antagonist effects and how this can be elucidated more clearly. Of particular merit is the Renilla luciferase vector. Not only does this act as a highly effective internal cytotoxicity control, but in reviewing the literature, is probably one of the best vectors to insert into an androgenic cell line, especially PC3, with AR. Certainly,</p>
--	---

	<p>according to the literature, for the AR performing far better and more accurately than the luciferase or beta galactosidase construct used for many transient and stably transfected cell lines.</p> <p>Reviewer #3: This assay was designed to evaluate the ability of test compounds to impact androgen-mediated gene transcription. This is a well characterized pathway of androgen action through the androgen receptor. The assay uses a cell line which has been engineered to stably express both human AR and an androgen responsive promoter/luciferase reporter construct. In theory, binding of the compound to the AR induces dimerisation of 2 ligand-bound AR which can then bind and activate the ARE promoter and stimulates production of the reporter gene, luciferase (a protein not normally found in the parent cell line). Luciferase activity should be proportional to the potency of the test compound. Conversely antagonist activity measures the ability of a compound to reduce the normal luciferase activity of a well-characterized agonist such as DHT. Some strengths of the assay include that it is an in vitro cell based assay so no animals are used for the assay, is a cost effective alternative to in vivo assays for prioritization and/or grouping compounds for further chemical testing and, unlike in vitro binding assays, has the ability to distinguish agonist from antagonist responses. Some limitations include that results are limited to the compound tested unless known metabolites are also tested and that test concentrations are limited by the solubility and cytotoxicity of the test compound.</p> <p>Reviewer #4: There is a comparison with results of AR binding assay and that of the proposed assay. This comparison demonstrates good relationship as mechanism basis. The comparison with in vivo assay systems, such as Hershberger assay (in vivo) may not be appropriate since Hershberger assay still has a complex biological system other than AR mediated effects that can alter weights of male accessory reproductive organs.</p> <p>The Panel agreed that this criterion has been partly met.</p>
<p>3. Availability of a detailed test method protocol</p>	<p>Reviewer #1: For most parts, the test protocol is clearly described, although the description of the two different protocols that have been used during pre-validation and the multi-lab study makes it somewhat difficult to follow. There are some points that need to be addressed in more detail, i.e. the description of the Renilla expression construct and the criteria for positive responses.</p> <p>The AR-EcoScreen cell line needs to be described in more detail, in particular the construct that was stably transfected to allow constitutive expression of the Renilla luciferase. It is somewhat confusing that the cLuc-EcoScreen was used for the cytotoxicity measurements during pre-validation, since constitutive Renilla expressing vectors are commonly used to standardize for transfection efficiency or unspecific effects on transcriptional or post-transcriptional events. Since</p>

	<p>obviously the same cell line was used in all studies the reasons why Renilla was not used from the very beginning should be explained / discussed.</p> <p>Different doses are tested in the pre-validation and inter-lab studies. The rationale behind these changes should be explained in detail.</p> <p>As mentioned before the different possibilities of unspecific modulation of luciferase activities need to be addressed. In addition, one should consider the possibility that mild cytotoxic effects might not be detected in the Renilla or the Luciferase cytotoxicity control. In general, viral promoters like the SV40 promoter that is also used to express the human AR protein in the AR-EcoScreen cell line, are very robust and might not be as sensitive as the ARE-reporter. Sodium azide and 17α-ethinylestradiol were dissolved in water or ethanol. It is stated in paragraph 51 that 0.1% ethanol did not affect the cells and, thus, ethanol should be included in the SOP.</p> <p>A common list of proficiency chemicals for the androgenic and anti-androgenic assay might be useful, since both assays would have to be performed in parallel anyway.</p> <p>The presentation of the data and the data analyses is somewhat confusing. PC10, PC50, EC50, lin.IC30 or lin.IC50 values are provided and discussed. It is also not always clear if data is based on single or several studies and if the agonist and antagonist studies were performed in parallel or independent of each other.</p> <p>In general, the data interpretation procedure should be stated more clearly and should be consistent throughout the report. For the agonist assay, the sensitivity of method is only acceptable, if PC10 values are used (Table 9). Although it seems from table 9 that this did not increase the false positive rate, some of the data provides evidence that the use of PC10 values may actually create problems.</p> <p>Actinomycin was recommended by ICCVAM as a control for substances that indirectly interfere with the reporter gene activity since it's the only RNA synthesis inhibitor in the ICCVAM list. Following Table 8, Actinomycin was considered positive in the agonist assay, because it induced a very slight increase in luciferase activity at concentrations that might be cytotoxic.</p> <p>As mentioned before, Fluoranthene stimulate a positive response in the agonist assay at concentrations that also stimulated Renilla activity in the antagonist assay (10-5M) and therefore it cannot be excluded that this effect might be due to ARE-independent activity.</p> <p>In addition, to classify substances as negative in the agonist assay because they never induce the luciferase activity strongly enough although EC50 values can be calculated might be problematic, since low inductive activities at low concentrations could still be biologically significant. This should be discussed in more detail.</p> <p>For the antagonist assay lin.IC30 as used in the validation study. The reason for using IC30 is not clear. In particular Table 11 shows no differences between lin.IC50 and lin.IC30. In general, the false negative rate is quite high in both cases (almost 17%).</p> <p>Some information about the use of the analysis software should be included. Can additional software be used besides</p>
--	---