# Draft Performance Standards for the BG1Luc Estrogen Receptor (ER) Transactivation *in Vitro* Assay to Detect ER Antagonists for TG 457

## INTRODUCTION

1. These Performance Standards (PS) accompany the BG1Luc Estrogen Receptor (ER) Transactivation (TA) Test Method for Identifying ER Agonists and Antagonists (TG 457) (1). Prior to the acceptance of a new test method for regulatory testing applications, validation studies are conducted to assess its reliability (i.e., the extent of intra- and inter-laboratory reproducibility over time when performed using the standardized protocol) and its relevance (i.e., the ability of the test method to correctly predict or measure the biological effect of interest) (2) (3) (4) (5). The purpose of performance standards is to communicate the basis by which new proprietary (i.e. copyrighted, trademarked, registered) and non-proprietary test methods have been determined to have sufficient accuracy (i.e., agreement between a test method result and an accepted reference value) and reliability (i.e., extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol) for a specific testing purpose. New test methods (i.e., "me-too" tests) can be added to TG 457. The Mutual Acceptance of Data will only be guaranteed if any proposed new or updated similar test method, developed according to these Performance Standards (for estrogen antagonist), and to the Performance Standards developed for TG 455 (for estrogen agonist), has been reviewed and adopted by the OECD.

2. Performance standards are based on an adequately validated test method(s) and provide a basis for evaluating the comparability of a proposed test method that is functionally and mechanistically similar (2) (3). The three elements of performance standards are:

- Essential test method components: These consist of essential structural, functional, and procedural elements of a validated test method. They should be included in the protocol of a proposed test method that is functionally and mechanistically similar to the validated method. Essential test method components include unique characteristics of the test method, critical procedural details, and quality control measures.

- A minimum list of Reference Chemicals: Reference Chemicals are used to assess the accuracy and reliability of a proposed functionally and mechanistically similar test method. These substances are a representative subset of those used to demonstrate the accuracy and reliability of the validated test method, and are the minimum number that should be used to evaluate the performance of a proposed mechanistically and functionally similar test method.

- Test method performance and reliability values: These are the standards for performance (i.e., accuracy, sensitivity, specificity, positive/negative predictivity) and reliability (i.e., degree to which the test method can be performed reproducibly within and among laboratories over time) that the proposed test method should meet or exceed when evaluated using the minimum list of Reference Chemicals.

3. The fully validated reference test method that provides the basis for this PS is the BG1Luc ER TA Test Method for Identifying ER Antagonists. This assay uses the BG1 Luc4E2 cell line which predominately expresses hERα with some contribution from hERß (1) (6) (7).

## ESSENTIAL TEST METHOD COMPONENTS AND OTHER VALIDATION CONSIDERATIONS

4. The BG1Luc ER TA test method uses an ER-responsive reporter gene (luc) in the human ovarian adenocarcinoma cell line BG-1 to detect substances with *in vitro* ER antagonist activity. The primary objective of this test method is to provide a qualitative assessment of *in vitro* anti-estrogenic activity (i.e., whether a substance is positive or negative for anti-estrogenic activity). Quantitative analysis is also performed to provide additional information on the potency of test substances. For example, quantitative analysis can determine the half-maximal inhibitory concentration ($IC_{50}$). Separate protocols are used to identify substances that possess ER agonist or antagonist activity, although the two protocols share most major components. Performance standards for the BG1Luc ER TA for detecting ER agonists are included as part of the PS for Stably Transfected Transactivation *In Vitro* Assays to Detect Estrogen Agonists (ER TA Methods) accompanying the Performance Based Test Guideline for TG 455 (8).

5. Certain principles are important in delineating the essential test method components that determine whether a modified or new test method is functionally and mechanistically similar to the BG1Luc ER TA test method. *In vitro* ER TA assays are designed to identify substances that might interfere with ER-mediated cellular processes *in vivo*. The interaction of estrogens with cellular ERs initiates a cascade of events leading to the expression of specific genes in multiple target tissues.

6. The following test method components may vary, so this PS applies to test methods that may differ in:

- cell type (e.g. mammalian, fish, yeast)
- cell line (tissue type)
- characteristics of the cell line including presence of other receptors and metabolism
- culture conditions
- plating density
- plate layout (including how controls are incorporated)
- ERα characteristics (full length or partial, species of origin); if other ER proteins are present, ERα should predominate and the relative expression of each receptor should be known
- reporter gene construct (promoter, receptor binding elements, reporter)
- method of determining cytotoxicity

These elements should be clearly described in the test method, and may be helpful for explaining any possible deviations from the BG1Luc ER TA (antagonist) TG.

7. Essential test method components for *in vitro* ER TA (antagonist) protocols should include:

- A full concentration response curve using a strong reference anti-estrogen, (e.g., raloxifene HCl) should be used in each experiment to demonstrate the adequacy of

the test method for detecting ER antagonists. At a minimum, a threefold reduction in the reference anti-estrogenic standard response should be demonstrated.

- A weak positive antagonist control (e.g., tamoxifen) that has an $IC_{50}$ slightly below 10 µM should be included to provide another quality control measure by which to judge the acceptability of the method for detecting a weak antagonist, and by which to evaluate the reproducibility of the test method. In addition, ER TA antagonist studies should include a concurrent control using the reference estrogen (e.g., E2) to establish a baseline level of induction (~80% of E2 maximum) against which the antagonistic activity of test substances can be assessed.

- A vehicle control (e.g., DMSO, EtOH, or $H_2O$) that is miscible with cell culture media at concentrations that are not cytotoxic and do not otherwise interfere with the test system.

- For initial range-finding, at least seven concentrations spaced at decadic logarithmic (log10) intervals should be tested up to the maximum concentration (see below). Based on these range-finding experiments, a suitable concentration range should then be used for testing the chemical in view of generating data on the possible potency of the substance and to derive categorical predictions (e.g., Positive, Negative).

- A qualitative or quantitative evaluation of cytotoxicity and how it is applied to the test method should be included in each study. Concentrations of test substances that clearly reduce viability should not be considered in the analysis of the data.

- All concentrations of the controls (e.g., vehicle, weak positive(s), or negative(s)), the reference estrogen, the reference anti-estrogen, and the test substance should be tested in more than one replicate well.

8. No standardized statistical methods for analyzing data obtained from *in vitro* ER TA antagonist assays have been developed. Each test method should establish a well-defined method for classifying a positive and a negative response. Positive results should be characterized by both the magnitude of the effect and the concentration at which the effect occurs (e.g., an $IC_{50}$, % max, etc.) when possible.

9. To ensure that a proposed *in vitro* ER TA test method possesses characteristics similar to other validated test methods, the Reference Chemicals for testing ER antagonists listed in Table 1 should be used to demonstrate the reliability and accuracy of the new test method. The 10 recommended Reference Chemicals, representing chemical classes commonly associated with ER activity, have been classified as ER antagonists or negatives based upon published reports, including *in vitro* assays for ER binding and TA (9) (10) (11). If a reference chemical is no longer commercially available, a substance with the same classification and comparable potency, mode of action, and chemical class can be used. Supplementary information including the full listings of chemicals tested in the BG1Luc ER TA (antagonist) is provided in Annex 2 (Table 1). Additional chemicals not included in the reference chemical list may be used to demonstrate an improvement (e.g., improved reproducibility and/or accuracy with regard to accepted reference data) of the new test method as compared with the fully validated test methods.

**Table 1. Reference Chemicals (10) for the Evaluation of Test Method Performance and Reliability for *In vitro* ER TA Assays to Detect ER Antagonists**

| Chemical[a] | CASRN | ICCVAM Consensus[b] | BG1Luc ER TA Consensus | BG1Luc ER TA Mean $IC_{50}$ (M)[c] | MeSH Chemical Class[d] | Product Class[d] |
|---|---|---|---|---|---|---|
| Tamoxifen | 10540-29-1 | POS | POS | $8.17 \times 10^{-7}$ | Hydrocarbon (Cyclic) | Pharmaceutical |
| 4-Hydroxytamoxifen | 68047-06-3 | POS | POS | $2.08 \times 10^{-7}$ | Hydrocarbon (Cyclic) | Pharmaceutical |
| Raloxifene HCl | 82640-04-8 | POS | POS | $1.19 \times 10^{-9}$ | Hydrocarbon (Cyclic) | Pharmaceutical |
| 17∝-Ethinyl estradiol | 57-63-6 | NEG | NEG | - | Steroid | Pharmaceutical, Veterinary Agent |
| Apigenin | 520-36-5 | NEG | NEG | - | Heterocyclic Compound | Dye, Natural Product, Pharmaceutical Intermediate |
| Chrysin | 480-40-0 | NEG | NEG | - | Flavonoid, Heterocyclic Compound | Natural Product |
| Coumestrol | 479-13-0 | NEG | NEG | - | Heterocyclic Compound | Natural Product |
| Genistein | 446-72-0 | NEG | NEG | - | Flavonoid, Heterocyclic Compound | Natural Product, Pharmaceutical |
| Kaempferol | 520-18-3 | NEG | NEG | - | Flavonoid, Heterocyclic Compound | Natural Product |
| Resveratrol | 501-36-0 | NEG | NEG | - | Hydrocarbon (Cyclic) | Natural Product |

Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; CASRN = CAS Registry Number (American Chemical Society); $IC_{50}$ = half-maximal inhibitory concentration; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; M = molar; MeSH = Medical Subject Headings (U.S. National Library of Medicine); NEG = negative; POS = positive.

[a]Chemicals are listed in order based upon $IC_{50}$ values.

[b]BG1Luc ER TA consensus classification represents the majority classification among the three validation laboratories.

[c]Mean $IC_{50}$ values were calculated with values reported by the laboratories of the BG1Luc ER TA validation study (XDS, ECVAM, and Hiyoshi) (9).

[d]Chemicals were assigned to one or more chemical classes using the U.S. National Library of Medicine's Medical Subject Headings (MeSH), an internationally recognized standardized classification scheme (available at http://www.nlm.nih.gov/mesh).

[e]Chemicals were assigned to one or more product classes using the U.S. National Library of Medicine's Hazardous Substances Data Bank [available at http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB].

11. New similar test methods should not be developed on the basis of the 10 Reference Chemicals, but rather on a sufficiently large test development set. Reference Chemicals should be preferentially used to determine equivalence of performance compared to the validated reference test methods.

12 All substances should be tested in a coded/blinded manner. When evaluated using these Reference Chemicals, the reliability and test method performance (i.e., sensitivity, specificity,

positive/negative predictivity) of the proposed ER TA test method should approximate the following defined reliability and accuracy values.

## DEFINED RELIABILITY AND ACCURACY PERFORMANCE VALUES

13. For the purposes of establishing the reliability and accuracy of the proposed test methods when transferred between laboratories, all 10 Reference Chemicals (Table 1) should be tested in two or (preferably) three laboratories. In each laboratory, all 10 References Chemicals should be tested in three runs.

### Within-laboratory (Intra-laboratory) reproducibility

14. For the assessment of within-laboratory reproducibility, the concordance of classifications (positive/negative) obtained in three independent consecutive test runs should be 100% for each laboratory for each of the 10 Reference Chemicals (Table 1). Three independent consecutive runs are required to fulfill the criteria for acceptance. If, for example, runs 2 and 3 are inconsistent with run 1, one additional run (run 4) will be sufficient to show within-lab reproducibility if run 4 is consistent with runs 2 and 3. If run 4 is consistent with run 1 instead, then at least two additional consecutive runs (runs 5 and 6) showing consistency with run 4 will be required to fulfill the requirement for three consecutive independent runs that have 100% concordance of classifications.

### Between-laboratory (Intra- laboratory) reproducibility

15. Between-laboratory reproducibility should be assessed using the 10 Reference Chemicals (Table 1). The concordance of classifications (positive/negative) in at least two, but preferably three, laboratories for the 10 Reference Chemicals should be 100% for the three positive substances, and at least 86% for the seven negative substances.

### Predictive capacity

16. The performance of the proposed test method (i.e., accuracy, sensitivity, specificity, positive/negative predictivity) should be comparable to that demonstrated for the fully validated BG1Luc ER TA (antagonist) test method (1) (9) when evaluating the 10 Reference Chemicals. Based upon the performance values of the validated reference method, the accuracy of the proposed ER TA test method should approximate those of the validated ER TA test method should be at least 90%.

17. Although it is not realistic to expect test methods to perform identically, discordant results should be addressed in terms of the ability of the test method to accurately classify other substances with similar potencies and from similar chemical classes as demonstrated by the fully validated test method (9).

LITERATURE


1. NIH Publication No. 11-7850: ICCVAM Test Method Evaluation Report, The LUMI-CELL® ER (BG1Luc ER TA) Test Method: An In Vitro Assay for Identifying Human Estrogen Receptor Agonist and Antagonist Activity of Chemicals [http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-TMER.htm]. Appendix B, BG1 Antagonist Protocol [*http://iccvam.niehs.nih.gov/docs/endo_docs/ERTA-TMER/AppxB2-AntagProtocol.pdf*]

2. OECD (2005), *Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment*, Series on Testing and Assessment No.34,,OECD, Paris. Available at:
   [http://www.oecd.org/officialdocuments/displaydocumentpdf?cote=ENV/JM/MONO(2005)14&doclanguage=en ]

3. ICCVAM (1997), Validation and Regulatory Acceptance of Toxicological Test Methods: A Report of the ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), NIH Publication No. 97-3981, National Institute of Environmental Health Sciences, Research Triangle Park, NC.

4. ICCVAM (2003), ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods, NIH Publication No. 03-4508, Research Triangle Park, NC, National Institute of Environmental Health Sciences

5. OECD(1996), Final Report of the OECD Workshop on Harmonization of Validation and Acceptance Criteria for Alternative Toxicological Test Methods, ENV/MC/CHEM/TG(96)9, OECD, Paris. Available at:
   [http://www.oecd.org/document/24/0,3746,en_2649_37465_47858904_1_1_1_37465,00.html ]

6. Pujol, P., *et al., (1998), Differential expression of estrogen re*$^{c}$*eptor-alpha and -beta messenger RNAs as a potential marker of ovarian carcinogenesis. Cancer* Res. 58(23): p. 5367-73.

7. Rogers, J.M. and M.S. Denison (2000),*Recombinant cell bioassays for endocrine disruptors: development of a stably transfected human ovarian cell line for the detection of estrogenic and anti-estrogenic chemicals*. *In Vitro* and Molecular Toxicology: Journal of Basic and Applied Research, 13(1): p. 67-82.

8. OECD (2009), *Test No. 455: The Stably Transfected Human Estrogen Receptor-alpha Transcriptional Activation Assay for Detection of Estrogenic Agonist Activity of Chemicals, (updated 2012): Performance-Based Test Guideline for The Stably Transfected Transactivation In Vitro Assays to Detect Estrogen Receptor Agonists*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing. [doi: 10.1787/9789264076372-en ].

9. ICCVAM (2011), ICCVAM Test Method Evaluation Report (TMER) on the LUMI-CELL® ER (BG1Luc ER TA) Test Method an *In Vitro* Method for Identifying ER Agonists and Antagonists, National Institute of Environmental Health Sciences, Research Triangle Park, NC. NIH Publication No. 11-785 [http://iccvam.niehs.nih.gov/docs/endo_docs/ERTA-TMER/BG1LucER-TA-TMER-Combined.pdf].

10. ICCVAM (2003), ICCVAM Evaluation of In Vitro Test Methods for Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and

Transcriptional Activation Assays,
Available at [http://iccvam.niehs.nih.gov/docs/endo_docs/edfinalrpt0503/edfinrpt.pdf]

11. ICCVAM (2002), Background Review Document: Estrogen Receptor Binding, Appendiz D, Substances Tested in the ER Binding Assay, NIH Publication Report No. 03-4504. Available at: [http://iccvam.niehs.nih.gov/docs/endo_docs/final1002/erbndbrd/ERBd034504.pdf]

# ANNEX 1

## Definitions and Abbreviations

**Acceptability criteria:** Minimum standards for the performance of experimental controls and reference standards. All acceptability criteria must be met for an experiment to be considered valid.

**Accuracy:** The closeness of agreement between a test method results and accepted reference values. It is a measure of test method performance and one aspect of relevance (2). Accuracy is determined by using the number of Reference Chemicals as denominator with number of correct responses in numerator normally expressed as a percent.

**Agonist:** A substance that produces a response, e.g., transcription, when it binds to a specific receptor.

**Antagonist:** A substance that inhibits an agonist response, e.g., transcription.

**Anti-estrogenic activity:** The capability of a chemical to inhibit 17β-estradiol or other estrogens in their ability to bind to and activate estrogen receptors. hERα-mediated estrogenic activity can be detected with the PBTG.

**BG-1:** An immortalized adenocarcinoma cell that endogenously express estrogen receptor.

**BG1Luc4E2**: The BG1Luc4E2 cell line was derived from BG-1 immortalized human-derived adenocarcinoma cells that endogenously express both forms of the estrogen receptor (ERα and ERβ) and have been stably transfected with the plasmid pGudLuc7.ERE. This plasmid contains four copies of a synthetic oligonucleotide containing the estrogen response element upstream of the mouse mammary tumor viral (MMTV) promoter and the firefly luciferase gene.

**Cytotoxicity:** Harmful effects to cell structure or function that can ultimately cause cell death and can be reflected by a reduction in the number of cells present in the well at the end of the exposure period or a reduction of the capacity for a measure of cellular function when compared to the concurrent vehicle control.

**E2:** 17β-estradiol

**ER:** Estrogen receptor

**hERα**: Human estrogen receptor alpha

**hERß**: Human estrogen receptor beta

**$IC_{50}$:** The half maximal inhibitory concentration of a test substance.

**Between-laboratory (Inter-laboratory) reproducibility:** A measure of the extent to which different qualified laboratories, using the same protocol and testing the same substances, can produce qualitatively and quantitatively similar results. Inter-laboratory reproducibility is determined during the prevalidation and validation processes, and indicates the extent to which a

test method can be successfully transferred between laboratories, also referred to as between-laboratory reproducibility (2).

**Within-laboratory (Intra-laboratory) reproducibility:** A determination of the extent that qualified people within the same laboratory can successfully replicate results using a specific protocol at different times (2).

**Me-too test:** A colloquial expression for a test method that is structurally and functionally similar to a validated and accepted reference test method. Interchangeably used with similar test method.

**PBTG:** Performance-Based Test Guideline.

**Performance standards:** Standards, based on a validated test method, that provide a basis for evaluating the comparability of a proposed test method that is mechanistically and functionally similar. Included are (1) essential test method components; (2) a minimum list of Reference Chemicals selected from among the chemicals used to demonstrate the acceptable performance of the validated test method; and (3) the comparable levels of accuracy and reliability, based on what was obtained for the validated test method, that the proposed test method should demonstrate when evaluated using the minimum list of Reference Chemicals (2).

**Predictivity (negative):** The proportion of correct negative responses among substances testing negative by a test method. It is an indicator of test method accuracy. Negative predictivity is a function of the sensitivity of the test method and the prevalence of negatives among the substances tested.

**Predictivity (positive):** The proportion of correct positive responses among substances testing positive by a test method. It is one indicator of test method accuracy. Positive predictivity is a function of the sensitivity of the test method and the prevalence of positives among the substances tested.

**Proficiency chemicals (substances):** Reference Chemicals included in the Performance Standards that can be used by laboratories to demonstrate technical competence with a standardized test method. Selection criteria for these substances typically include that they represent the range of responses, are commercially available, and have high quality reference data available.

**Proficiency:** The demonstrated ability to properly conduct a test method prior to testing unknown substances.

**Reference Chemicals (substances**): A set of chemicals to be used to demonstrate the ability of a new test method to meet the acceptability criteria demonstrated by the validated reference test method(s). These chemicals should be representative of the classes of chemicals for which the test method is expected to be used, and should represent the full range of responses that may be expected from the chemicals for which it may be used, from strong, to weak, to negative.

**Reference anti-estrogen:** Raloxifene HCl (Ral, CASRN 82640-04-8).

**Reference estrogen:** 17ß-estradiol (E2, CASRN 50-28-2).

**Relevance:** Description of relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method (2).

**Reliability:** Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility (2).

**SD:** Standard deviation

**Sensitivity**: The proportion of all positive/active substances that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results, and is an important consideration in assessing the relevance of a test method (2).

**Specificity**: The proportion of all negative/inactive substances that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results, and is an important consideration in assessing the relevance of attest method (2).

**Stable transfection:** When DNA is transfected into cultured cells in such a way that it is stably integrated into the cells genome, resulting in the stable expression of transfected genes. Clones of stably transfected cells are selected by stable markers (e.g., resistance to G418).

**Substance:** Used in the context of the UN GHS as chemical elements and their compounds in the natural state or obtained by any production process, including any additive necessary to preserve the stability of the product and any impurities deriving from the process used, but excluding any solvent which may be separated without affecting the stability of the substance or changing its composition.

**TA:** Transactivation.

**Transcription:** mRNA synthesis

**Transcriptional activation:** The initiation of mRNA synthesis in response to a specific chemical signal, such as a binding of an estrogen to the estrogen receptor.

**Validation:** The process by which the reliability and relevance of a particular approach, method, process or assessment is established for a defined purpose.

**VC:** Vehicle control, the solvent that is used to dissolve test and control chemicals is tested solely as vehicle without dissolved chemical.

**Weak positive control:** A weakly active substance selected from the Reference Chemicals list that is included in all tests to help ensure proper functioning of the assay.

# ANNEX 2

## Supplementary Information

## for the

## The Estrogen Receptor (BG1Luc ER TA) Transactivation Test Method for Identifying ER Antagonists

**Table 1: Substances Tested for ER Antagonist Activity During the BG1Luc ER TA Validation Study (9)**

| | Chemical[1] | CASRN | BG1Luc ER TA Classification[3] |
|---|---|---|---|
| 1 | 4-hydroxytamoxifen | 68047-06-3 | POS |
| 2 | Actinomycin D[2] | 50-76-0 | POS |
| 3 | Apomorphine | 58-00-4 | POS |
| 4 | Cycloheximide[2] | 66-81-9 | POS |
| 5 | Dibenzo[*a.h*] anthracene | 53-70-3 | POS |
| 6 | Ketoconazole | 65277-42-1 | POS |
| 7 | Medroxy-progesterone acetate | 71-58-9 | POS |
| 8 | Raloxifene HCl | 82640-04-8 | POS |
| 9 | Tamoxifen | 10540-29-1 | POS |
| 10 | 12 – *O* –tetradecanoyl-phorbol-13-acetate | 16561-29-8 | NEG |
| 11 | 17-α ethinyl estradiol | 57-63-6 | NEG |
| 12 | 17-β estradiol | 50-28-2 | NEG |
| 13 | 17β-trenbolone | 10161-33-8 | NEG |
| 14 | 19-nortestosterone | 434-22-0 | NEG |
| 15 | 2-*sec*-butylphenol | 89-72-5 | NEG |
| 16 | 2,4,5-trichlorophenoxy-acetic acid | 93-76-5 | NEG |
| 17 | 4-androstenedione | 63-05-8 | NEG |
| 18 | 4-cumylphenol | 599-64-4 | NEG |
| 19 | 4-hydroxyandrostenedione | 566-48-3 | NEG |
| 20 | Apigenin | 520-36-5 | NEG |
| 21 | 4-*tert*-octylphenol | 140-66-9 | NEG |
| 22 | 5α-dihydrotestosterone | 521-18-6 | NEG |
| 23 | Ammonium perchlorate | 7790-98-9 | NEG |
| 24 | Chrysin | 480-40-0 | NEG |
| 25 | Atrazine | 1912-24-9 | NEG |
| 26 | Bicalutamide | 90357-06-5 | NEG |
| 27 | Bisphenol A | 80-05-7 | NEG |
| 28 | Bisphenol B | 77-40-7 | NEG |
| 29 | Butylbenzyl phthalate | 85-68-7 | NEG |

| | Chemical[1] | CASRN | BG1Luc ER TA Classification[3] |
|---|---|---|---|
| 30 | Coumestrol | 479-13-0 | NEG |
| 31 | Corticosterone | 50-22-6 | NEG |
| 32 | Genistein | 446-72-0 | NEG |
| 33 | Cyproterone acetate | 427-51-0 | NEG |
| 34 | Daidzein | 486-66-8 | NEG |
| 35 | Dexamethasone | 50-02-2 | NEG |
| 36 | Di-*n*-butyl phthalate | 84-74-2 | NEG |
| 37 | Dicofol | 115-32-2 | NEG |
| 38 | Diethylhexyl phthalate | 117-81-7 | NEG |
| 39 | Diethylstilbestrol | 56-53-1 | NEG |
| 40 | Estrone | 53-16-7 | NEG |
| 41 | Ethyl paraben | 120-47-8 | NEG |
| 42 | Fenarimol | 60168-88-9 | NEG |
| 43 | Finasteride | 98319-26-7 | NEG |
| 44 | Flavone | 525-82-6 | NEG |
| 45 | Fluoranthene | 206-44-0 | NEG |
| 46 | Fluoxymestrone | 76-43-7 | NEG |
| 47 | Flutamide | 13311-84-7 | NEG |
| 48 | Kaempferol | 520-18-3 | NEG |
| 49 | Haloperidol | 52-86-8 | NEG |
| 50 | Hydroxyflutamide | 52806-53-8 | NEG |
| 51 | Resveratrol | 501-36-0 | NEG |
| 52 | Kepone | 143-50-0 | NEG |
| 53 | L-thyroxine | 51-48-9 | NEG |
| 54 | Linuron | 330-55-2 | NEG |
| 55 | *meso*-hexestrol | 84-16-2 | NEG |
| 56 | Methyl testosterone | 58-18-4 | NEG |
| 57 | Mifepristone | 84371-65-3 | NEG |
| 58 | Morin | 480-16-0 | NEG |
| 59 | Nilutamide | 63612-50-0 | NEG |
| 60 | Norethynodrel | 68-23-5 | NEG |

| | Chemical[1] | CASRN | BG1Luc ER TA Classification[3] |
|---|---|---|---|
| 61 | *o.p'*-DDT | 789-02-6 | NEG |
| 62 | Oxazepam | 604-75-1 | NEG |
| 63 | *p*-n-nonylphenol | 104-40-5 | NEG |
| 64 | *p.p'*-DDE | 72-55-9 | NEG |
| 65 | *p.p'*-methoxychlor | 72-43-5 | NEG |
| 66 | Phenobarbital | 50-06-6 | NEG |
| 67 | Phenolphthalin | 81-90-3 | NEG |
| 68 | Pimozide | 2062-78-4 | NEG |
| 69 | Procymidone | 32809-16-8 | NEG |
| 70 | Progesterone | 57-83-0 | NEG |
| 71 | Propylthiouracil | 51-52-5 | NEG |
| 72 | Sodium azide | 26628-22-8 | NEG |
| 73 | Spironolactone | 52-01-7 | NEG |
| 74 | Testosterone | 58-22-0 | NEG |
| 75 | Vinclozolin | 50471-44-8 | NEG |

[1]Table is sorted by classification and then alphabetically by chemical name. Only substances for which a definitive POS/NEG call could be made were included in the table.

[2]Actinomycin D and cycloheximide, inhibit protein biosynthesis, and should not be considered as antagonists.

[3]Classification based upon results reported in the ICCVAM Test Method Evaluation Report (TMER) on the LUMI-CELL® ER (BG1Luc ER TA) Test Method an *In Vitro* Method for Identifying ER Agonists and Antagonists [9].

## Summary of the Accuracy and Reliability Values Obtained During the Validation Studies for the BG1Luc ER TA (Antagonists)

The ICCVAM Test Methods Evaluation Report on the BG1Luc ER TA Test Methods (9) provides a comprehensive description of the data used to develop the accuracy and reliability values for the antagonist assay. The following is a summary of the test method performance and intra- and inter-laboratory reproducibility for the validated test method.

I. **Intra-laboratory (within-laboratory) reproducibility:** The closeness of agreement between test results obtained within a single laboratory when the procedure is performed using the same substance under identical conditions within a given period of time.

   a. The intra-laboratory reproducibility of the BG1Luc ER TA Antagonist test method was evaluated using 12 substances (2 positive, 10 negative), that were each tested three times on three separate days at each laboratory. There was 100% agreement within each laboratory for each of the three repeat tests of these Reference Chemicals (Phase 2 of the antagonist validation study).

**Table 2**: Intra-laboratory reproducibility for the BG1Luc ER TA Antagonist Assay (9)

| Activity per Test | XDS | ECVAM | Hiyoshi |
|---|---|---|---|
| Agreement Within Laboratory | 12/12 (100%) | 12/12 (100%) | 12/12 (100%) |
| +++ | 2/12 | 2/12 | 2/12 |
| --- | 10/12 | 10/12 | 10/12 |
| Discordance Within Laboratory | 0/12 (0%) | 0/12 (0%) | 0/12 (0%) |
| ++- | 0/12 | 0/12 | 0/12 |
| +-- | 0/12 | 0/12 | 0/12 |

Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods;
XDS = Xenobiotic Detection Systems, Inc.

\+ Denotes a positive test result.

\- Denotes a negative test result.

+++ Indicates that each of three replicate tests within each laboratory had a classification as positive.

--- Indicates that each of three replicate tests within each laboratory had a classification as negative.

++- Indicates that a test substance was classified as positive in two of three replicate tests. The substance was classified as negative in a third replicate test.

+-- Indicates that the test substance was classified as positive in one of three replicate tests. The substance was classified as negative in the remaining two tests.

II. **Inter-laboratory (Between-laboratory) reproducibility:** A measure of the extent to which different qualified laboratories using the same protocol and testing the same substances can produce qualitatively and quantitatively similar results. Inter-laboratory reproducibility is determined during the validation process and indicates the extent to which a test method can be transferred successfully among laboratories.

a. Inter-laboratory reproducibility was assessed using 53 chemicals that were tested at least once in each of 3 laboratories. There was 94% (50/53) agreement on the classifications for these chemicals among the laboratories (9). Two chemicals (2/53) had inadequate overall classifications (i.e., 1 positive, 1 negative and 1 inadequate call). The agreement between the laboratories is shown in Table 3.

**Table 3**: Inter-laboratory reproducibility for the BG1Luc ER TA Antagonist Assay (9)

| **Results Among Laboratories** | **Percent Agreement** |
|---|---|
| Agreement Among Laboratories | 50/53 (94%) |
| +++ | 4/53 (8%) |
| --- | 43/53 (81%) |
| ++I | 1/53 (2%) |
| --I | 2/53 (4%) |
| Discordance Among Laboratories | 3/53 (6%) |
| ++- | 0/53 (0%) |
| +-- | 1/53 (2%) |
| +-I | 2/53 (4%) |

Abbreviations: I = inadequate data (i.e., Data are classified as inadequate if, because of major qualitative or quantitative limitations, they cannot be interpreted as valid for showing either the presence or absence of activity. Inadequate data typically result from some type of systemic error, such as high background across the test plate or failure of a multi-tip pipette to dispense liquid in numerous wells.
[a]Only those substances that produced a definitive result in at least two of the three laboratories were used in this evaluation.
[b]Substances that produced an inadequate result in two laboratories during agonist testing were not included in this table.
\+ Denotes a positive test result.
\- Denotes a negative test result.
+++ Indicates that the substance was classified as positive at all three laboratories.
--- Indicates that the substance was classified as negative at all three laboratories.
++I Indicates that the substance was classified as positive at two of three laboratories but had inadequate data in the third.
--I Indicates that the substance was classified as negative at two of three laboratories but had inadequate data in the third.
+-I Indicates that the substance was classified as positive at one laboratory, negative at one laboratory, and inadequate at the third laboratory

III. **Predictive Capacity**: Measures of test method performance (i.e., accuracy, sensitivity, specificity, positive and negative predictivity), and overall accuracy provide a quantitative assessment of the closeness of agreement (e.g, the proportion of correct outcomes) between test method results and the values obtained from Reference Chemicals.

a. The predictive capacity was assessed using 25 Reference Chemicals (3 positive, 22 negative) that produced definitive results in the BG1Luc ER TA assay for antagonist activity (See Section 3.4 in reference 9).

**Table 4**: Predictive Capacity for the BG1 Luc ER TA (Antagonist Assay) (9)

| | BG1Luc ER TA | | |
|---|---|---|---|
| Negative | 0 | 22 | 22 |
| Total | 3 | 22 | 25 |

| | | |
|---|---|---|
| Accuracy | 100% | 25/25 |
| Sensitivity | 100% | 3/3 |
| Specificity | 100% | 22/22 |
| Positive predictivity | 100% | 3/3 |
| Negative predictivity | 100% | 22/22 |

**Table 5**: Template for Accuracy Analysis

| | | New Test Outcome | | |
|---|---|---|---|---|
| | | Positive | Negative | Total |
| Reference Test Classification | Positive | a | c | a + c |
| | Negative | b | d | b + d |
| | Total | a + b | c + d | a+b+c+d |

a = positive in both new assay and by reference test classification
b = positive in new assay and negative by reference test classification
c = negative in new assay and positive by reference test classification
d = negative in both new assay and by reference test classification
Accuracy = ([a+d]/[a+b+c+d])
Sensitivity = (a/[a+c])
Specificity = (d/[b+d])
Positive Predictivity = (a/[a+b])
Negative Predictivity = (d/[c+d])