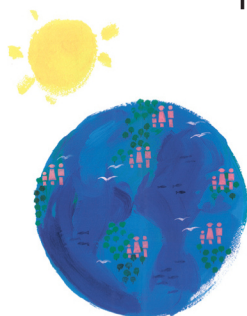


ICCVAM Test Method Evaluation Report

The LUMI-CELL® ER (BG1Luc ER TA) Test Method: An *In Vitro* Assay for Identifying Human Estrogen Receptor Agonist and Antagonist Activity of Chemicals

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

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



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

About the Interagency Coordinating Committee on the Validation of Alternative Methods and The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods

In 1997, the National Institute of Environmental Health Sciences (NIEHS), one of the National Institutes of Health, established the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) to:

- Coordinate interagency technical reviews of new and revised toxicological test methods, including alternative test methods that reduce, refine, or replace the use of animals
- Coordinate cross-agency issues relating to validation, acceptance, and national and international harmonization of new, modified, and alternative toxicological test methods

On December 19, 2000, the ICCVAM Authorization Act (42 U.S.C. 285l-3) established ICCVAM as a permanent interagency committee of NIEHS under the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM).

ICCVAM conducts technical evaluations of new, revised, and alternative toxicological and safety testing methods with regulatory applicability. ICCVAM also promotes the scientific validation and regulatory acceptance of toxicological test methods that more accurately assess the safety or hazards of chemicals and products and that reduce, refine (decrease or eliminate pain and distress), or replace animal use. NICEATM administers ICCVAM and provides scientific and operational support for ICCVAM-related activities. More information about NICEATM and ICCVAM can be found on the NICEATM-ICCVAM website (<http://iccvam.niehs.nih.gov>) or obtained by contacting NICEATM (telephone: [919] 541-2384, e-mail: niceatm@niehs.nih.gov).

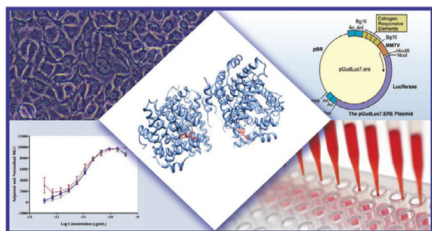
ICCVAM consists of representatives from the following 15 U.S. Federal regulatory and research member agencies that require, use, generate, or disseminate toxicological information:*

- *Consumer Product Safety Commission*
- *Department of Agriculture*
- *Department of Defense*
- *Department of Energy*
- Department of Health and Human Services
 - Centers for Disease Control and Prevention
 - *Agency for Toxic Substances and Disease Registry*
 - *National Institute for Occupational Safety and Health*
 - *Food and Drug Administration*
 - *National Institutes of Health*
 - *National Cancer Institute*
 - *National Institute of Environmental Health Sciences*
 - *National Library of Medicine*
- *Department of the Interior*
- Department of Labor
 - *Occupational Safety and Health Administration*
- *Department of Transportation*
- *Environmental Protection Agency*

*Italics indicate those agencies represented on ICCVAM, as specified in the ICCVAM Authorization Act.



The NICEATM-ICCVAM earth-and-sun graphic symbolizes the important role of new and alternative toxicological and safety testing methods in protecting and advancing the health of people, animals, and our environment.



On the cover: This collage of pictures includes a drawing representing estrogen receptor molecules bound to 17 β -estradiol surrounded by (clockwise from top left): cells used for the BG1Luc ER TA test method; drawing of a plasmid contained in the cells used for the BG1Luc ER TA test method which includes an estrogen response element and a luciferase reporter gene; pipettor and assay plate representing lab automation; and an example of data from a positive control substance in the BG1Luc ER TA agonist assay.

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

**Interagency Coordinating Committee on the
Validation of Alternative Methods**

**National Toxicology Program Interagency Center for the
Evaluation of Alternative Toxicological Methods**

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

2011

NIH Publication Number 11-7850

**National Toxicology Program
P.O. Box 12233
Research Triangle Park, NC 27709**

This document is available electronically at

<http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-TMER.htm>

When referencing this document, please cite as follows:

Interagency Coordinating Committee on the Validation of Alternative Methods. 2011. 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. NIH Publication No. 11-7850. Research Triangle Park, NC: National Institute of Environmental Health Sciences.

Table of Contents

List of Tables	vi
List of Figures.....	vii
List of Abbreviations and Acronyms.....	viii
Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM): Agency Representatives.....	x
Acknowledgements.....	xii
Preface.....	xvii
Executive Summary	xix
1.0 Introduction.....	1
2.0 ICCVAM Recommendations: Usefulness and Limitations of the BG1Luc ER TA Test Method.....	3
2.1 Background and Introduction.....	3
2.2 ICCVAM Recommendations: Test Method Usefulness and Limitations.....	3
2.2.1 Evaluation as a Screening Test to Identify Substances with Estrogen Receptor Agonist Activity.....	3
2.2.2 Evaluation as a Screening Test to Identify Substances with Estrogen Receptor Antagonist Activity.....	4
2.3 ICCVAM Recommendations: Test Method Protocol for the BG1Luc ER TA Test Method	5
2.4 ICCVAM Recommendations: Future Studies for the BG1Luc ER TA Test Method.....	6
2.5 ICCVAM Recommendations: Performance Standards for the BG1Luc ER TA Test Method	7
3.0 Validation Status for Use of the BG1Luc ER TA Test Method as a Screening Test to Identify <i>In Vitro</i> ER Agonists and Antagonists.....	8
3.1 Test Method Description.....	8
3.2 General Test Method Procedures	8
3.3 Validation Database	8
3.4 Test Method Accuracy	9
3.5 Test Method Reliability.....	13
3.6 Animal Welfare Considerations: Reduction, Refinement, and Replacement	16

4.0	BG1Luc ER TA Test Method Performance Standards	18
4.1	Elements of ICCVAM Performance Standards	18
4.2	LUMI-CELL (BG1Luc ER TA) Test Method Performance Standards	18
4.2.1	Background	18
4.2.2	BG1Luc ER TA Essential Test Method Components and Other Validation Considerations	19
4.2.3	Reference Substances for <i>In Vitro</i> ER TA Test Methods	24
4.2.4	Accuracy and Reliability Performance Values	28
5.0	ICCVAM Consideration of Public Comments	29
5.1	ICCVAM Consideration of Public and SACATM Comments	29
5.1.1	Public Comments in Response to 69 FR 21564 (April 21, 2004)	30
5.1.2	Public Comments in Response to 71 FR 13597 (March 16, 2006)	30
5.1.3	Public Comments in Response to 74 FR 62317 (November 27, 2009)	30
5.1.4	Public Comments in Response to 76 FR 4113 (January 24, 2011)	30
5.1.5	Public Comments in Response to 76 FR 23323 (April 26, 2011)	31
5.1.6	Public Comments in Response to 76 FR 28781 (May 18, 2011)	32
6.0	References	33
Appendix A	ICCVAM Evaluation of the BG1Luc ER TA (LUMI-CELL®) Test Method – Timeline	A-1
Appendix B	Updated ICCVAM-Recommended Protocols: BG1Luc ER TA	B-1
B1	BG1Luc ER TA – Agonist Protocol	B-3
B2	BG1Luc ER TA – Antagonist Protocol	B-37
Appendix C	Final Background Review Document: Validation Study of the BG1Luc4E2 Estrogen Receptor (ER) Transcriptional Activation (TA) Test Method	C-1
Appendix D	Independent Scientific Peer Review Panel Assessment	D-1
D1	Summary Minutes from the Independent Scientific Peer Review Panel Meeting on March 29-30, 2011	D-3
D2	Independent Scientific Peer Review Panel Report: Evaluation of the LUMI-CELL® ER (BG1Luc ER TA) Test Method	D-35

Appendix E	Federal Register Notices and Public Comments.....	E-1
E1	69 FR 21564 - <i>In Vitro</i> Endocrine Disruptor Test Methods: Request for Comments and Nominations (No public comments received).....	E-3
E2	71 FR 13597 - Notice of Availability of a Revised List of Recommended Reference Substances for Validation of <i>In Vitro</i> Estrogen and Androgen Receptor Binding and Transcriptional Activation Assays: Request for Comments and Submission of <i>In Vivo</i> and <i>In Vitro</i> Data (No public comments received).....	E-7
E3	74 FR 62317 - Evaluation of <i>In Vitro</i> Estrogen Receptor Transcriptional Activation and <i>In Vitro</i> Cell Proliferation Assays for Endocrine Disruptor Chemical Screening: Request for Nominations for an Independent Expert Peer Review Panel and Submission of Relevant <i>In Vitro</i> and <i>In Vivo</i> Data.....	E-11
E4	76 FR 4113 - Announcement of an Independent Scientific Peer Review Panel Meeting on an <i>In Vitro</i> Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening; Availability of Draft Background Review Document (BRD); Request for Comments	E-23
E5	76 FR 23323 - Meeting of the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM)	E-37
E6	76 FR 28781 - Independent Scientific Peer Review Panel Report: Evaluation of the Validation Status of an <i>In Vitro</i> Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening: Notice of Availability and Request for Public Comments (No public comments received).....	E-55
E7	Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) Comments: SACATM Meeting on June 16-18, 2011	E-59
Appendix F	Relevant Endocrine Disruptor Regulations and Testing Guidelines	F-1

List of Tables

Table 3-1	42 ICCVAM-Recommended Substances Used to Evaluate ER Agonist Accuracy	10
Table 3-2	25 ICCVAM-Recommended Substances Used to Evaluate ER Antagonist Accuracy	12
Table 3-3	Accuracy of the BG1Luc ER TA Agonist Data.....	13
Table 3-4	Accuracy of the BG1Luc ER TA Antagonist Data.....	13
Table 3-5	Agonist Within-Plate DMSO Control Data	14
Table 3-6	Antagonist Within-Plate DMSO Control Data	15
Table 3-7	Antagonist E2 Control Values	15
Table 3-8	Intralaboratory Agreement for Multiple Testing of the 12 Phase 2 Agonist Substances Tested Independently at Least Three Times at Each Laboratory	16
Table 4-1	34 Reference Substances for Evaluation of ER Agonist Accuracy	25
Table 4-2	10 Reference Substances for Evaluation of ER Antagonist Accuracy	27
Table 5-1	Opportunities for Public Comments	29

List of Figures

Figure 4-1	Example Concentration–Response Curve for an ER TA Agonist	21
Figure 4-2	Example Concentration–Response Curve for an ER TA Antagonist	22

List of Abbreviations and Acronyms

AR	Androgen receptor
BRD	Background review document
CASRN	CAS Registry Number [®] (a trademark of the American Chemical Society)
CERI	Chemicals Evaluation and Research Institute, Japan
CV	Coefficient of variation
DMSO	Dimethyl sulfoxide
EAC	Endocrine-active compound
EC ₅₀	Half-maximal effective concentration
ECVAM	European Centre for the Validation of Alternative Methods
ED	Endocrine disruptor
EDSP	Endocrine Disruptor Screening Program (U.S. EPA)
EDSTAC	Endocrine Disruptor Screening and Testing Advisory Committee (U.S. EPA)
EDWG	ICCVAM Interagency Endocrine Disruptor Working Group
EPA	U.S. Environmental Protection Agency
ER	Estrogen receptor
FDA	U.S. Food and Drug Administration
FR	<i>Federal Register</i>
FW	Formula weight
GLP	Good Laboratory Practice
I	Inadequate
IC ₅₀	Half-maximal inhibitory concentration
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
ILS	Integrated Laboratory Systems, Inc.
ISO	International Organization for Standardization
JaCVAM	Japanese Center for the Validation of Alternative Methods
KoCVAM	Korean Center for the Validation of Alternative Methods
M	Molar
Max	Maximum
MeSH [®]	Medical Subject Headings (U.S. National Library of Medicine)
N	Number; negative
NEG	Negative
NICEATM	National Toxicology Program Center for the Evaluation of Alternative Toxicological Methods

NIEHS	U.S. National Institute of Environmental Health Sciences
NIH	U.S. National Institutes of Health
NT	Not tested
NTP	U.S. National Toxicology Program (U.S. NIH)
OECD	Organisation for Economic Co-operation and Development
OPPTS	Office of Prevention, Pesticides and Toxic Substances (U.S. EPA)
PN	Presumed negative
POS	Positive
PP	Presumed positive
RLU	Relative light unit
SACATM	Scientific Advisory Committee on Alternative Toxicological Methods
SD	Standard deviation
SMT	Study Management Team
STTA	Stably transfected human estrogen receptor- α transcriptional activation
TA	Transcriptional activation
TG	Test Guideline
U.S.C.	United States Code
XDS	Xenobiotic Detection Systems, Inc.

Trademarks are the property of their respective owners.

Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM): Agency Representatives

Agency for Toxic Substances and Disease

Registry

- * Moiz Mumtaz, PhD
- Edward Murray, PhD
- Eric Sampson, PhD
- Bruce Fowler, PhD (to June 2011)

Consumer Product Safety Commission

- * Joanna Matheson, PhD (Vice-Chair)
- + Kristina Hatlelid, PhD
- Marilyn L. Wind, PhD (Chair, to July 2010)

Department of Agriculture

- * Jodie Kulpa-Eddy, DVM (Chair)
- + Elizabeth Goldentyer, DVM

Department of Defense

- * Patrick Mason, PhD
- + Terry Besch, DVM, MS, DACLAM, DACVPM
- + Patty Decot
- David Honey, PhD (to July 2011)
- Robert E. Foster, PhD (to Aug. 2010)
- Harry Salem, PhD (to Aug. 2010)
- Peter J. Schultheiss, DVM, DACLAM (to Aug. 2010)

Department of Energy

- * Michael Kuperberg, PhD
- Marvin Stodolsky, PhD (retired 2011)

Department of the Interior

- * Barnett A. Rattner, PhD
- Sarah Gerould, PhD (to Feb. 2009)

Department of Transportation

- + Steve Hwang, PhD
- George Cushmac, PhD (retired 2010)

Environmental Protection Agency

Office of Pesticide Programs

- * John R. "Jack" Fowle III, PhD, DABT
- + Vicki Dellarco, PhD
- Christine Augustyniak, PhD (*U.S. Coordinator, OECD Test Guidelines Program*) (to Sept. 2011)
- Tina Levine, PhD (to Sept. 2011)
- Deborah McCall (to Sept. 2011)

Office of Pollution Prevention and Toxics

- Jerry Smrcek, PhD (*U.S. Coordinator, OECD Test Guidelines Program*) (to July 2009)

Environmental Protection Agency (cont.)

Office of Research and Development

- Stephanie Padilla, PhD (to July 2009)
 - Julian Preston, PhD (to July 2009)
 - Suzanne McMaster, PhD (to Dec. 2008)
- #### Office of Science Coordination and Policy
- Karen Hamernik, PhD (to July 2009)

Food and Drug Administration

Office of the Commissioner

- * Suzanne Fitzpatrick, PhD, DABT

Center for Biologics Evaluation and Research

- Ying Huang, PhD
- Richard McFarland, PhD, MD

Center for Devices and Radiological Health

- Vasant G. Malshet, PhD, DABT

Center for Drug Evaluation and Research

- + Abigail C. Jacobs, PhD
- Paul C. Brown, PhD

Center for Food Safety and Applied Nutrition

- Robert Bronaugh, PhD
- David G. Hattan, PhD
- Diego Rua, PhD
- Neil Wilcox, DVM, MPH (to May 2011)

Center for Veterinary Medicine

- M. Cecilia Aguila, DVM
- Li You, PhD
- Devaraya Jagannath, PhD (to April 2011)

National Center for Toxicological Research

- Paul Howard, PhD
- Donna L. Mendrick, PhD
- William T. Allaben, PhD (to Jan. 2009)

Office of Regulatory Affairs

- Lawrence D'Hoostelaere, PhD

National Cancer Institute

- * T. Kevin Howcroft, PhD
- + Chand Khanna, DVM, PhD
- Alan Poland, MD (to Oct. 2008)

continued

- * Principal agency representative
- + Alternate principal agency representative

National Institute of Environmental Health Sciences

- * William S. Stokes, DVM, DACLAM
- + Warren Casey, PhD, DABT
- Rajendra S. Chhabra, PhD, DABT
- Jerrold J. Heindel, PhD
- Raymond R. Tice, PhD (to June 2010)

National Institute for Occupational Safety and Health

- * Paul Nicolaysen, VMD

National Institutes of Health

- * Margaret D. Snyder, PhD

National Library of Medicine

- * Pertti (Bert) Hakkinen, PhD
- + Jeanne C. Goshorn, MS

Occupational Safety and Health Administration

- * Surender Ahir, PhD
- * Principal agency representative
- + Alternate principal agency representative

Acknowledgements

ICCVAM Interagency Endocrine Disruptor Working Group (EDWG)

U.S. Consumer Product Safety Commission

Kent Carlson, PhD

U.S. Department of the Interior

Catherine Richter, PhD

Donald Tillitt, PhD

U.S. Environmental Protection Agency

Don Bergfelt, PhD

John R. “Jack” Fowle, III, PhD, DABT

Jesudoss Rowland

U.S. Food and Drug Administration

Center for Devices and Radiological Health

Thomas Umbreit, PhD

Center for Drug Evaluation and Research

Jeffrey Bray, PhD

Paul C. Brown, PhD

Karen L. Davis-Bruno, PhD

Ross Filice, MD

Abigail (Abby) Jacobs, PhD

Leslie McKinney, PhD

Center for Food Safety and Nutrition

Michael Bolger, PhD, DABT

David G. Hattan, PhD (Chair)

Center for Veterinary Medicine

M. Cecilia Aguila, DVM

Charles Eirkson, PhD

Kevin Gaido, PhD

Devaraya Jagannath, PhD

Annette McCarthy, PhD

Li You, PhD

National Center for Toxicological Research

Kenneth Delclos, PhD

Huixiao Hong, PhD

Jon Wilkes, PhD

Office of the Commissioner

Suzanne Fitzpatrick, PhD, DABT

National Institute of Environmental Health Sciences

Warren Casey, PhD, DABT

Jerrold Heindel, PhD

William Stokes, DVM, DACLAM

Julius Thigpen, PhD

Occupational Safety and Health Administration

Surender Ahir, PhD

European Centre for the Validation of Alternative Methods – Liaison

Susanne Bremer, PhD

Elise Grignard, PhD

Valerie Zuang, PhD

Japanese Center for the Validation of Alternative Methods – Liaison

Hajime Kojima, PhD

Atsushi Ono, PhD

Study Management Team

NICEATM

William Stokes, DVM, DACLAM (NIEHS)
Director, NICEATM

Warren Casey, PhD (NIEHS)
Deputy Director, NICEATM

Ray Tice, PhD (NIEHS)
Deputy Director, NICEATM
(On team until June 2005)

David Allen, PhD (ILS, Inc.)
Principal Investigator

Jeff Charles, PhD (ILS, Inc.)
Principal Investigator
(On team until April 2007)

Frank Deal, MS (ILS, Inc., until March 2011)
Staff Toxicologist

Patricia Ceger, MS (ILS, Inc.)
Project Coordinator

JaCVAM

Hajime Kojima, PhD
Director, JaCVAM

Atsushi Ono, PhD
National Institute of Health Sciences
(On team until January 2007)

KoCVAM

Soon Young Han, PhD
Director, KoCVAM

ECVAM

Susanne Bremer, PhD

Elise Grignard, PhD

Thomas Hartung, PhD
Head of ECVAM
(On team until April 2008)

Miriam Jacobs, PhD
(On team until August 2007)

Participants in the BG1Luc ER TA Validation Study

NICEATM gratefully acknowledges the generous contributions of the individuals who directly participated in the NICEATM/ECVAM/JaCVAM/KoCVAM validation study. Their time and efforts are greatly appreciated.

Xenobiotic Detection Systems, Inc. (XDS) (Lead Laboratory)

George Clark, PhD – Study Director from 2010
John Gordon, PhD – Study Director until 2008
Jun Liu, PhD – Study Director until 2009
Ruth McMillan, PhD – Technical Lead from 2010
Andrew Chu – Laboratory Manager until 2009
Ying Gu – Lead Chemist

ECVAM (In Vitro Methods Unit) (Participating Laboratory)

Miriam Jacobs, PhD – Study Director until 2008
Patricia Pazos, PhD – Study Director from 2009
Jan de Lange – Technical Lead

Hiyoshi Corporation (Participating Laboratory)

Masafumi Nakamura – Study Director
Hiyorhi Handa – Technical Lead
Clevita Eldrida Caesar Mascalenhas – Translator

NIEHS (Chemistry Resources Group)

Cynthia Smith, PhD – Branch Chief
Beby Jayaram, PhD – Chemist

Statistical Consultant

Joseph Haseman, PhD

**2011 LUMI-CELL[®] (BG1Luc ER TA) Test Method
Independent Scientific Peer Review Panel
(March 29 – 30, 2011)**

A. John Bailer, PhD

Distinguished Professor and Chair
Department of Statistics
Miami University
Oxford, Ohio

Christopher Borgert, PhD

President and Principal Scientist
Applied Pharmacology and Toxicology, Inc.
Gainesville, FL

Grantley Charles, PhD, DABT

Senior Scientist, Toxicology
Drug Safety Evaluation
Allergan Inc.
Irvine, CA

Daniel Desaulniers, PhD

Research Scientist
Health Canada
Hazard Identification Division
Ottawa, ON, Canada

J. Charles Eldridge, PhD

Professor of Physiology & Pharmacology
Wake Forest University Health Sciences
Winston-Salem, NC

William R. Kelce, PhD

Principal
Exponent[®]
Cary, NC

Hyung Kim, PhD

Associate Professor
College of Pharmacy, Pusan National
University
Busan, Republic of Korea

Steven L. Levine, PhD

Science Fellow and Product Safety Manager
Monsanto Company
St. Louis, MO

Alberto Mantovani, DVM

Head of Food and Veterinary Toxicology Unit
Italian National Health Institute
(ISS - Istituto Superiore di Sanità)
Rome, Italy

Ellen Mihaich, PhD, DABT

President and Principal Scientist
Environmental and Regulatory Resources,
LLC
Durham, NC

Hiroshi Ono, PhD

Scientific Advisor
Hatano Research Institute, Food and Drug
Safety Center
Kanagawa, Japan

John G. Vandenberg, PhD

Professor Emeritus
Department of Biology
North Carolina State University
Raleigh, NC

Sherry L. Ward, PhD, MBA

Consultant
BioTred Solutions
New Market, MD

Marc Weimer, PhD

Biostatistician
German Cancer Research Center
Department of Biostatistics
Heidelberg, Germany

James Wittliff, PhD, MD *hc*, FACB

Professor, Department of Biochemistry &
Molecular Biology
Director, Hormone Receptor Laboratory
Director, Institute for Molecular Diversity &
Drug Design
University of Louisville
Louisville, KY

James Yager, PhD

Professor and Director of Toxicology
Johns Hopkins University Bloomberg School
of Public Health
Baltimore, MD

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

National Institute of Environmental Health Sciences

William Stokes, DVM, DACLAM
Director (NICEATM); Project Officer

Warren Casey, PhD, DABT
Deputy Director

Deborah McCarley
Special Assistant; Assistant Project Officer

NICEATM Support Contract Staff (Integrated Laboratory Systems [ILS], Inc.)

David Allen, PhD
Patricia Ceger, MS
Frank Deal, MS (until March 2011)
Jonathan Hamm, PhD
Linda Litchfield
Steven Morefield, MD
Michael Paris
Catherine Sprankle
Linda Wilson

Statistical Consultant for ILS, Inc.

Joseph Haseman, PhD

Preface

Endocrine-active compounds (EACs) are both naturally occurring and synthetic substances. Some may, depending on the dose, interfere with the normal function of hormones in the endocrine system. Public health concerns have resulted largely from studies indicating that animal populations exposed to high levels of these substances, sometimes referred to as endocrine disruptors (EDs), have an increased incidence of reproductive and developmental abnormalities (EPA 1997; NRC 1999). In response to growing concerns about possible adverse health effects in humans exposed to such substances, the U.S. Congress enacted relevant provisions to safeguard public health in the Federal Food, Drug, and Cosmetic Act (21 U.S.C. 301 et seq.); the Food Quality Protection Act (7 U.S.C. 136); and the 1996 Amendments to the Safe Drinking Water Act (110 Stat 1613). The U.S. Environmental Protection Agency (EPA) was required to develop and validate a screening and testing program to identify substances with endocrine-disrupting activity. The EPA subsequently established the Endocrine Disruptor Screening Program (EDSP) and initiated efforts to standardize and validate test methods for inclusion in the EDSP (66 FR 23022). Validation is necessary to assess the usefulness and limitations of a test method for a specific proposed purpose and to characterize the extent to which test methods are sufficiently accurate and reproducible for their intended use (ICCVAM 1997).

In April 2000, the EPA nominated four types of *in vitro* test methods for detecting substances with potential endocrine-disrupting activity for review by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM). These included *in vitro* estrogen receptor (ER) and androgen receptor (AR) binding and ER and AR transcriptional activation (TA) test methods. The EPA also asked ICCVAM to develop performance standards that could be used to define acceptable *in vitro* ER and AR binding and TA assays. It was envisioned that these standards would be based on the performance of adequately validated *in vitro* ER- and AR-based assays.

In 2002, the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) prepared background review documents (BRDs) that included all available information on each of the four types of test methods (ICCVAM 2002d, 2002a, 2002c, 2002b). In a public meeting, an independent international expert panel (Panel) reviewed the information on the 137 assays described in the BRDs and concluded that there were no adequately validated *in vitro* ER- or AR-based test methods (ICCVAM 2002e). Based on recommendations from the Panel, ICCVAM published the *ICCVAM Evaluation of In Vitro Test Methods for Detecting Potential Endocrine Disruptors*, which included a list of reference substances that should be used to validate each of the four types of *in vitro* test methods (ICCVAM 2003a). It also identified essential test method components that should be included in each of the standardized test method protocols used for future validation studies. ICCVAM recommended that future performance standards for these methods be based on test methods that have undergone adequate validation studies using the recommended accuracy chemicals and essential test method components.

In January 2004, Xenobiotic Detection Systems, Inc. (XDS; Durham, NC), nominated the LUMI-CELL® BG1Luc4E2 ER TA test method (BG1Luc ER TA test method) for an interlaboratory validation study. ICCVAM and the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) recommended that the BG1Luc ER TA test method be considered a high priority for interlaboratory validation studies due to the lack of adequately validated test methods and the regulatory and public health need for such test methods. NICEATM subsequently led and coordinated an international validation study with its counterparts in Japan (JaCVAM) and Europe (ECVAM), using laboratories sponsored by each validation organization. NICEATM organized a validation Study Management Team (SMT) to oversee the scientific aspects of the

validation study and coordinate the day-to-day activities among the participating laboratories. A representative from the recently established Korean Center for the Validation of Alternative Methods (KoCVAM) joined the SMT in 2010.

ICCVAM reviewed the validation status of the BG1Luc ER TA test method for identification of substances with ER agonist or antagonist activity. NICEATM and the ICCVAM Interagency Endocrine Disruptor Working Group (EDWG) prepared a draft BRD that provided a comprehensive description and the data from the validation study used to assess the accuracy and reliability of the BG1Luc ER TA test method.

NICEATM convened an independent international scientific peer review panel (Panel) that met in public on March 29–30, 2011. The Panel was charged with reviewing the draft BRD for completeness, assessing the extent that established validation and acceptance criteria were adequately addressed, and determining the extent to which the data and information supported draft ICCVAM test method recommendations on the usefulness and limitations of the BG1Luc ER TA test method. The Panel also evaluated the proposed performance standards. The Panel included expert scientists nominated by ECVAM, JaCVAM, and KoCVAM.

ICCVAM considered the conclusions and recommendations of the Panel, along with comments from the public and SACATM, and then finalized the BRD and test method recommendations, which are provided in this test method evaluation report. As required by the ICCVAM Authorization Act (42 U.S.C. 285I-3), ICCVAM forwarded this report and recommendations to Federal agencies for their consideration and acceptance decisions where appropriate. The BG1Luc ER TA test method protocol and performance standards were also forwarded to the Organisation for Economic Co-operation and Development (OECD) Test Guidelines Programme for consideration and adoption as international testing guidelines.

We gratefully acknowledge the organizations and scientists who generated and provided data and information for this document, especially the staff at the participating validation laboratories: XDS, Inc., in Durham, North Carolina; Hiyoshi Corporation in Japan; and the In Vitro Methods Unit at ECVAM in Italy. We would also like to recognize the efforts of the individuals who contributed to its preparation, review, and revision. We thank Dr. David Hattan (U.S. Food and Drug Administration) for serving as Chair of the EDWG, as well as the members of the EDWG and ICCVAM representatives who subsequently reviewed and provided comments throughout the process leading to this test method evaluation report. We also want to thank Dr. Warren Casey, Deputy Director of NICEATM, for his excellent leadership and extensive efforts on this project.

Staff from the NICEATM support contractor, Integrated Laboratory Systems, Inc., are acknowledged for their excellent scientific and operational support, including Drs. David Allen, Jon Hamm, and Steven Morefield; Patricia Ceger, Frank Deal (until March 2011), Linda Litchfield, Michael Paris, Catherine Sprankle, and Linda Wilson. Finally, we want to thank Drs. Susanne Bremer and Elise Grignard, the EDWG liaisons from ECVAM, and Drs. Hajime Kojima and Atsushi Ono, the EDWG liaisons from JaCVAM, for their participation and support.

RADM William S. Stokes, DVM, DACLAM
Rear Admiral/Assistant Surgeon General, U.S. Public Health Service
Director, NICEATM
Executive Director, ICCVAM

Jodie A. Kulpa-Eddy, DVM
VS-NCIE-Agricultural Select Agent Program
U.S. Department of Agriculture
Chair, ICCVAM

Executive Summary

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) has completed its evaluation of the validation status of the LUMI-CELL® BG1Luc4E2 estrogen receptor (ER) transcriptional activation (TA) test method (hereafter BG1Luc ER TA test method) as a screening test to identify substances with *in vitro* ER agonist and antagonist activity. The BG1Luc ER TA test method uses BG-1 cells, a human ovarian adenocarcinoma cell line that is stably transfected with an estrogen-responsive luminescence (luciferase reporter) gene, to measure whether and how much a substance induces (agonist) or inhibits (antagonist) TA activity via ER-mediated pathways. Such substances could interfere with the normal function of hormones in the endocrine system (i.e., endocrine disruptors), which may lead to abnormal growth, development, or reproduction.

This test method evaluation report provides ICCVAM's recommendations for the BG1Luc ER TA test method based on the results of an international validation study and the demonstrated validity (usefulness and limitations). The report also includes (1) recommendations for future studies, (2) performance standards to evaluate functionally and mechanistically similar test methods, (3) protocols recommended by ICCVAM for future data collection and evaluation of the BG1Luc ER TA test method, and (4) a final background review document (BRD) describing the validation status of this test method.

In 2004, Xenobiotic Detection Systems, Inc. (XDS; Durham, NC), nominated the LUMI-CELL ER test method to ICCVAM for an interlaboratory validation study. ICCVAM and the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) recommended that the BG1Luc ER TA test method be considered a high priority for interlaboratory validation studies based on the lack of adequately validated test methods and the regulatory and public health need for such test methods.

When the BG1Luc ER TA validation study was initiated, no *in vitro* ER TA test methods were considered adequately valid for regulatory use. Today, only one *in vitro* ER TA test method is considered adequately validated by national and international agencies, the Organisation for Economic Co-operation and Development (OECD) Stably Transfected Human Estrogen Receptor- α Transcriptional Activation (STTA) Assay for the Detection of Estrogenic Agonist-Activity, described in OECD Chemicals Test Guideline (TG) 455 (OECD 2009). Validated by the Chemicals Evaluation and Research Institute (CERI, Japan), this method has been adopted by the U.S. Environmental Protection Agency (EPA) as OPPTS 890.1300: Estrogen Receptor Transcriptional Activation (Human Cell Line [HeLa-9903]) (EPA 2009).

After recommendation by ICCVAM and SACATM, the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) led and coordinated an international validation study with its counterparts in Europe (the European Centre for the Validation of Alternative Methods [ECVAM]) and Japan (the Japanese Center for the Evaluation of Alternative Methods [JaCVAM]) to assess the accuracy and reliability of the BG1Luc ER TA test method for the qualitative detection of substances with *in vitro* ER agonist or antagonist activity. The BG1Luc ER TA test method was evaluated using laboratories in the United States (XDS), Europe (ECVAM), and Japan (Hiyoshi Corporation).

The validation study proceeded in four phases. During Phase 1, each of the three participating centers (NICEATM, ECVAM, and JaCVAM) selected validation laboratories. The protocols were reviewed, and the laboratories demonstrated proficiency with the test method by successfully completing 10 replicate agonist and 10 replicate antagonist tests. In Phases 2 through 4, the protocols were evaluated and refined, and 78 ICCVAM reference substances that

should be used to standardize and validate *in vitro* ER and androgen receptor binding and TA test methods were tested.

After this study was completed, NICEATM, ICCVAM, and the ICCVAM Interagency Endocrine Disruptor Working Group (EDWG) prepared a draft BRD and draft test method recommendations. The drafts were provided to an independent international scientific peer review panel (hereafter Panel) and to the public for comment. The Panel met in public session on March 29–30, 2011, to discuss its peer review of the ICCVAM draft BRD and to provide conclusions and recommendations regarding the validation status of the BG1Luc ER TA test method. The Panel also reviewed how well the information contained in the draft BRD supported ICCVAM's draft test method recommendations.

In finalizing this test method evaluation report and the BRD, which is included here as an appendix, ICCVAM considered (1) the conclusions and recommendations of the Panel, (2) comments from SACATM, and (3) public comments.

ICCVAM Recommendations: Test Method Usefulness and Limitations

ICCVAM concludes that the accuracy and reliability of the BG1Luc ER TA test method support its use to screen substances for *in vitro* ER agonist and/or antagonist activity. This determination is based on an evaluation of data from the validation study and the corresponding accuracy and reliability. ICCVAM concludes that the accuracy of this assay is at least equivalent to that of the current ER TA test method included in regulatory testing guidance (EPA OPPTS 890.1300) (EPA 2009).

ICCVAM Recommendations: BG1Luc ER TA Test Method Protocol

For use of the BG1Luc ER TA test method to screen substances for *in vitro* ER agonist and/or antagonist activity, ICCVAM recommends using the ICCVAM BG1Luc ER TA protocols (included here as **Appendices B1** and **B2**). All future studies intended to further characterize the usefulness and limitations of the BG1Luc ER TA test method should use these protocols.

ICCVAM Recommendations: Future Studies

ICCVAM considers the BG1Luc ER TA test method to be valid as described. However, ICCVAM recommends the following for interested parties to further characterize and potentially improve the usefulness and applicability of the BG1Luc ER TA test method:

- Additional validation studies may be performed to determine whether the BG1Luc ER TA test method or other similar assays could replace the rat uterine cytosol ER binding assay.
- Further work may be carried out to determine if the BG1Luc ER TA test method could be combined with other methods (to include *in vitro* metabolic activation) in a weight-of-evidence approach to replace the uterotrophic bioassay.
- Additional studies/evaluations may be conducted to more completely characterize the ratio of ER α and ER β in the BG-1 cell line and the extent to which these receptor subtypes contribute to the overall performance of the BG1Luc ER TA test method.
- Additional studies/evaluations may be conducted to determine the feasibility of testing volatile substances using CO₂-permeable plastic film or other methods to seal the test plates.
- Additional studies/evaluations may be conducted to determine if substances that are not soluble in dimethyl sulfoxide (DMSO) could be tested in another vehicle that would more adequately dissolve the substance in culture media.
- Additional studies may be conducted to account for metabolic activation that could expand the utility of this and other ER TA test methods.
- As ER antagonists are identified, additional studies/evaluations may be conducted to expand the database of positive substances tested and thereby better characterize the usefulness and limitations of the BG1Luc ER TA test method as a screening test to identify substances with ER antagonist activity.

ICCVAM encourages users to provide to ICCVAM all data that are generated from future studies. These data could be used to further characterize the usefulness and limitations of the BG1Luc ER TA test method as a screening test to identify substances with ER agonist or antagonist activity.

Validation Status of the BG1Luc ER TA Test Method

ICCVAM evaluated the BG1Luc ER TA test method for its ability to correctly identify *in vitro* ER agonists and antagonists. For this analysis, test substance classification (positive or negative for ER agonist/antagonist activity) obtained during the validation study was compared to the ICCVAM reference classification of the same substance, which was based on a preponderance of available data.

The BG1Luc ER TA test method accuracy was evaluated based on several different analyses, but the primary evaluation was based on two comparisons: (1) the extent to which the result of the test method corresponds to the ICCVAM reference classification for each substance and (2) the accuracy of the BG1Luc ER TA test method compared to that of the EPA OPPTS 890.1300/OECD TG 455 (EPA 2009; OECD 2009)¹ assay.

Test Method Accuracy – Agonist Assay

Thirty-five substances (28 positive, 7 negative) were used to evaluate the accuracy of the BG1Luc ER TA agonist assay. The consensus classification obtained from all BG1Luc ER TA tests for these 35 substances yielded the following statistics: concordance of 97% (34/35), sensitivity of 96% (27/28), specificity of 100% (7/7), a false positive rate of 0% (0/7), and a false negative rate of 4% (1/28). Similar results were obtained when the results from each laboratory were used instead of the consensus classification.

EPA OPPTS 890.1300/OECD TG 455 is the only test guideline published by a U.S. regulatory agency for generating ER TA data. Therefore, BG1Luc ER TA test method concordance with EPA OPPTS 890.1300/OECD TG 455 was also evaluated using the 26 reference substances for which data are available from both BG1Luc ER TA and EPA OPPTS 890.1300/OECD TG 455 assays. Accuracy statistics for the two test methods were identical: concordance of 96% (25/26), sensitivity of 95% (21/22), specificity of 100% (4/4), a false positive rate of 0% (0/4), and a false negative rate of 5% (1/22).

Test Method Accuracy – Antagonist Assay

To evaluate the accuracy of the BG1Luc ER TA antagonist assay, 25 substances (3 positive, 22 negative) were used. The consensus classification obtained from all BG1Luc ER TA tests for these 25 substances yielded the following statistics: concordance of 100% (25/25), sensitivity of 100% (3/3), specificity of 100% (22/22), a false positive rate of 0% (0/22), and a false negative rate of 0% (0/3). Similar results were obtained when the results from each laboratory were used instead of the consensus classification.

Because there currently is no valid EPA OPPTS 890.1300/OECD TG 455 antagonist protocol, no comparison with the BG1Luc ER TA antagonist results was conducted.

Concordance with Other Endocrine Disruptor Assays

Although the primary goal of the BG1Luc ER TA test method is to provide a qualitative assessment of estrogenic/anti-estrogenic activity, quantitative measures of activity are usually obtained for positive results. The values obtained from BG1Luc ER TA test results (half-maximal effective concentration [EC₅₀] and half-maximal inhibitory concentration [IC₅₀]), were compared to median values from other ER TA test methods reported in the literature. This comparison

¹ The EPA OPPTS 890.1300/OECD TG 455 (OECD 2009) assay uses the hER α -HeLa-9903 human cervical cancer cell line to detect estrogen agonist activity mediated through human ER alpha (hER α).

found a high correlation. There was 97% (33/34) concordance between the BG1Luc ER TA test method and ER binding data. The only discordant substance (medroxyprogesterone acetate) was positive in the BG1Luc ER TA test method and negative based on ER binding data. Similarly, based on a comparison with available data in the *in vivo* uterotrophic assay, there was 92% (12/13) concordance between the BG1Luc ER TA test method and ER binding data. The only discordant substance (butylbenzyl phthalate) was positive in the BG1Luc ER TA test method and negative based on uterotrophic data.

Test Method Reliability

Intralaboratory reproducibility (whether multiple tests of the same substance at a single laboratory produce the same results) of the BG1Luc ER TA agonist and antagonist test methods was assessed by comparing (1) reference standard and control results for all plates tested within each laboratory during the course of the validation study and (2) results from Phase 2 testing, during which 12 substances were tested in at least three independent experiments in each of the three laboratories. Intralaboratory agreement for agonist and antagonist classification was determined for the 12 substances that were tested at least three times at each laboratory.

In the agonist testing, mean induction in each laboratory ranged from 4.6 to 7.8 fold, and 17 β -estradiol (E2) reference standard EC₅₀ values ranged from 8.0×10^{-12} to 1.2×10^{-11} M. There was 100% agreement within each laboratory for each of the three repeat tests, although the agonist classifications for some of the 12 test substances differed among the different laboratories.

In the antagonist testing, mean reduction ranged from 8.0 to 9.9 fold, and raloxifene reference standard IC₅₀ values ranged from 1.1×10^{-9} to 1.3×10^{-9} M. There was 100% agreement within each laboratory for each of the three repeat tests, although the antagonist classifications for some of the 12 test substances differed among the different laboratories.

Interlaboratory reproducibility (whether tests of a single substance run at different laboratories produce the same results) of the BG1Luc ER TA agonist and antagonist test methods was determined for the 12 substances that were tested at least three times for agonist and antagonist activity during Phase 2 at each of the three laboratories. The three laboratories agreed on 67% (8/12) of the substances tested for agonist activity and on 100% (12/12) of the substances tested for antagonist activity.

Interlaboratory reproducibility was also determined for 41 substances that were tested once for agonist and antagonist activity during Phase 3 testing at each of the three laboratories. Five of the 41 substances produced inadequate results for agonist activity and could not be considered in the evaluation. Among the 36 remaining substances that produced a definitive test result in at least two laboratories, there was 100% agreement. All 41 substances produced definitive results for antagonist activity. The three laboratories agreed on 93% (38/41) of these substances.

ICCVAM Recommendations: Performance Standards

Based on the results of this study, NICEATM and the EDWG developed performance standards applicable to methods that are functionally and mechanistically similar to the BG1Luc ER TA test method. These performance standards can also be used by laboratories with no experience with the BG1Luc ER TA test method to demonstrate technical proficiency.

Essential Test Method Components

In order to be considered functionally and mechanistically similar to the BG1Luc ER TA test method, a modified ER TA test method protocol must include the following components to ensure that the same biological effect is being measured:

- The test method should be based on a cell line that endogenously expresses ER.
- Reference standards, controls, and test substances should be dissolved in a solvent that mixes well with cell culture media at concentrations that are noncytotoxic and that do not otherwise interfere with the test system.
- The maximum test substance concentration should be 1 mM for ER TA agonist testing and 10 μ M for ER TA antagonist testing unless otherwise limited by solubility, cytotoxicity, or other mechanisms that interfere with assay performance.
- At least seven concentrations spaced at logarithmic (\log_{10}) intervals, up to the limit concentration, should be tested.
- An evaluation of cytotoxicity should be included, and only data from concentrations at or above 80% viability should be used for data analyses.
- A reference estrogen and a reference anti-estrogen should be used to demonstrate the adequacy of the test method for detecting ER TA agonist and antagonist activity.
- The ability of the reference estrogen to induce ER TA activity and the ability of the reference anti-estrogen to inhibit ER TA activity should be demonstrated by generating a full concentration–response curve in each experiment that provides a minimum threefold estrogenic induction and a minimum threefold anti-estrogenic reduction.
- A set of concurrent controls should be included. For agonist assays, this would include the vehicle control and a weak agonist. For antagonist assays, this would include the vehicle control, weak antagonist, and reference estrogen.
- Test substances that are positive for ER agonist activity should have a concentration–response curve consisting of a baseline, followed by a positive slope, with a response peak of at least 20% of the average maximal value of the reference estrogen response.
- Test substances are negative for agonist activity if all data points are below 20% of the average maximal value of the reference estrogen response.
- Test substances that are positive for ER antagonist activity should have a concentration–response curve consisting of a baseline, followed by a negative slope, with a response decrease to at least 80% of the average maximal value of the reference estrogen response.
- Test substances are negative for ER antagonist activity if all data points are above 80% of the average maximal value of the reference estrogen response.

Test method protocols should incorporate the essential components listed above. Modifications should be detailed and scientifically justified, and the modified test method should perform as well as or better than the BG1Luc ER TA test method.

Reference Substances

ICCVAM recommends for test method validation a subset of those substances that were definitively classified as positive or negative for ER TA activity in the scientific literature and that were tested in the BG1Luc ER TA validation study. The reference substances include a range of chemical and product classes commonly associated with endocrine disruption.

Test Method Accuracy and Reliability

When evaluated using this minimum list of recommended reference substances, a proposed ER TA test method should have accuracy (i.e., sensitivity, specificity, false positive rates, and false negative rates) and reliability characteristics equal to or better than those of the BG1Luc ER TA test method. Any misclassified reference substances should be addressed in terms of the test method's ability to accurately classify other substances with similar potencies and from the same chemical/product classes.

Using the Performance Standards

Test method developers are encouraged to consult directly with ICCVAM before using these performance standards to conduct a validation study for a proposed test method. Developers are also encouraged to submit results of validation studies to ICCVAM for an evaluation of the validation status. Upon completing its evaluation in accordance with the ICCVAM Authorization Act (42 U.S.C. 285l-3), ICCVAM will forward recommendations to ICCVAM agencies regarding the usefulness and limitations of the test method.

ICCVAM Consideration of the Independent Peer Review Panel Report and Other Comments

The ICCVAM evaluation process incorporates scientific peer review and a high level of transparency. The evaluation process for the BG1Luc ER TA test method included a public review meeting by an independent scientific peer review panel, multiple opportunities for public comments, and comments from SACATM. ICCVAM and the EDWG considered the Panel report, SACATM comments, and all public comments before finalizing the ICCVAM test method evaluation report and final BRD for the BG1Luc ER TA test method.

1.0 Introduction

In vitro estrogen receptor (ER) transcriptional activation (TA) assays are designed to identify agonist or antagonist substances that might interfere with estrogen activity *in vivo*. Unlike receptor binding assays, TA assays can distinguish between agonist and antagonist activity. The BG1Luc ER TA test method utilizes an ER-responsive reporter gene (*luc*) in the human ovarian adenocarcinoma cell line BG-1 to detect substances with *in vitro* ER agonist or antagonist activity. ER-mediated transcription of the *luc* gene results in the production of luciferase, the activity of which is quantified using a luminometer. A concentration–response curve can be established to provide qualitative and quantitative information regarding the *in vitro* estrogenic activity of a test substance (Rogers and Denison 2000).

The Federal Food, Drug, and Cosmetic Act; the Food Quality Protection Act; and the Safe Drinking Water Act all aim to identify potential endocrine disruptors and thereby protect humans and animals (7 U.S.C. 136; 21 U.S.C. 301 et seq.; 110 Stat 1613). The U.S. Environmental Protection Agency (EPA) was specifically required to “develop a screening program, using appropriate validated test systems and other scientifically relevant information, to determine whether certain substances may have an effect in humans that is similar to an effect produced by a naturally occurring estrogen, or such other endocrine effect as the Administrator may designate” (21 U.S.C. 346a[p][1]). In 1996, the EPA formed the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC), a committee of scientists and stakeholders that was charged by the EPA to provide recommendations on how to implement its Endocrine Disruptor Screening Program (EDSP). The EDSP is described in detail at <http://www.epa.gov/scipoly/oscpendo/>.

The EPA accepted EDSTAC’s recommendations for a two-tier screening program as proposed in the *Federal Register* (63 FR 71542). The purpose of Tier 1, which consists of *in vivo* and *in vitro* test methods, is to identify the potential of chemicals to interact with the estrogen, androgen, or thyroid hormonal systems. Tier 1 currently includes EPA OPPTS 890.1300: Estrogen Receptor Transcriptional Activation (Human Cell Line [HeLa-9903]) (EPA 2009). EPA OPPTS 890.1300 is an ER TA test method validated for the detection of *in vitro* ER agonists.

In 2004, Xenobiotic Detection Systems, Inc. (XDS), nominated their LUMI-CELL® ER test method (hereafter BG1Luc ER TA test method) to the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) for validation. ICCVAM and the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) recommended that the BG1Luc ER TA test method should be considered a high priority for interlaboratory validation studies based on the lack of adequately validated test methods and the regulatory and public health need for such test methods.

The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) led and coordinated an international validation study with its counterparts in Europe (the European Centre for the Validation of Alternative Methods [ECVAM]) and Japan (the Japanese Center for the Evaluation of Alternative Methods [JaCVAM]) using laboratories sponsored by each validation organization. NICEATM organized a Study Management Team (SMT) to oversee the scientific aspects of the validation study and coordinate the day-to-day activities among the participating laboratories (XDS, ECVAM, and Hiyoshi). A representative from the recently established Korean Center for the Validation of Alternative Methods (KoCVAM) joined the SMT in 2010.

The validation study proceeded in four phases. During Phase 1, each of the three participating centers (ICCVAM, ECVAM, and JaCVAM) selected validation laboratories. The protocols were reviewed, and the laboratories demonstrated proficiency with the test method by successfully completing 10 replicate agonist and 10 replicate antagonist tests. In Phases 2 through 4, the

protocols were evaluated and refined, and 78 ICCVAM reference substances that should be used to standardize and validate *in vitro* ER and AR binding and TA test methods were tested.

Based on the results of this study, ICCVAM reviewed the validation status of the BG1Luc ER TA test method for identification of substances with *in vitro* ER agonist or antagonist activity. NICEATM and the ICCVAM Interagency Endocrine Disruptor Working Group (EDWG) prepared a draft background review document (BRD) that provides a comprehensive description and the data from the validation study used to assess the accuracy and reliability of the BG1Luc ER TA test method.

On January 24, 2011, ICCVAM announced the availability of the draft BRD to the public and a public Panel meeting to review the validation status of the BG1Luc ER TA test method as a screening test to identify *in vitro* ER agonists and antagonists (76 FR 4113²). All of the information provided to the Panel, including the draft BRD, ICCVAM draft test method recommendations, and all public comments received before the Panel meeting, were made publicly available via the NICEATM–ICCVAM website (<http://iccvam.niehs.nih.gov/>).

The public Panel meeting was held on March 29–30, 2011. The Panel evaluated (1) the extent to which the draft BRD addressed established validation and acceptance criteria and (2) the extent to which the draft BRD supported ICCVAM’s draft test method recommendations. Interested stakeholders from the public were provided opportunities to comment at the Panel meeting. After considering all public comments, the Panel agreed with the ICCVAM draft recommendation that the BG1Luc ER TA test method can be used as a screening test to identify substances with *in vitro* ER agonist and antagonist activity. On May 18, 2011, ICCVAM posted a report of the Panel’s recommendations³ (see **Appendix D**) on the NICEATM–ICCVAM website for public review and comment (announced in 76 FR 28781).

ICCVAM provided SACATM with the draft BRD and test method recommendations, the Panel report, and all public comments for discussion at their meeting on June 16–17, 2011, where public stakeholders were given another opportunity to comment.

ICCVAM and the EDWG considered the SACATM comments, the Panel report, and all public comments before finalizing ICCVAM test method recommendations for use of the BG1Luc ER TA test method as a screening test to identify substances with *in vitro* ER agonist and antagonist activity. ICCVAM’s recommendations (see **Section 2.0**) and the final BRD (see **Appendix C**) are incorporated in this test method evaluation report. As required by the ICCVAM Authorization Act of 2000 (42 U.S.C. 285l-3), ICCVAM will forward this report and its recommendations to U.S. Federal agencies for consideration. Federal agencies must respond to ICCVAM within 180 days after receiving ICCVAM test method recommendations. ICCVAM recommendations are available to the public on the NICEATM–ICCVAM website (<http://iccvam.niehs.nih.gov/>), and agency responses will also be made available as they are received.

² *Federal Register* notices published by NICEATM–ICCVAM during evaluation of the BG1Luc ER TA test method are available in Appendix E and from the NICEATM–ICCVAM website (<http://iccvam.niehs.nih.gov/>).

³ http://iccvam.niehs.nih.gov/docs/endo_docs/EDPRPrept2011.pdf

2.0 ICCVAM Recommendations: Usefulness and Limitations of the BG1Luc ER TA Test Method

2.1 Background and Introduction

ICCVAM has completed its evaluation of the validation status of the BG1Luc ER TA test method, an *in vitro* method proposed to identify potential agonist or antagonist substances that might interfere with normal estrogen activity. NICEATM and ICCVAM prepared a comprehensive BRD that includes the data and information available to characterize the validity of this proposed use of the BG1Luc ER TA test method. The information included in the BRD (**Appendix C**) is based on an international validation study that utilized 78 reference substances that should be used to standardize and validate *in vitro* ER and androgen receptor (AR) binding and TA test methods. Based on the results of this study, ICCVAM developed these draft test method recommendations on the usefulness and limitations of the BG1Luc ER TA test method for identifying potential ER agonists or antagonists. ICCVAM also developed draft recommendations for standardized test method protocols, future studies, and performance standards.

2.2 ICCVAM Recommendations: Test Method Usefulness and Limitations

2.2.1 Evaluation as a Screening Test to Identify Substances with Estrogen Receptor Agonist Activity

ICCVAM concludes that the BG1Luc ER TA test method can be used as a screening test to identify substances with *in vitro* ER agonist activity. This recommendation is based on an evaluation of available validation study data and corresponding accuracy and reliability. ICCVAM concludes that the accuracy of this assay is at least equivalent to that of EPA OPPTS 890.1300, part of the EDSP Tier 1 screening battery. The supporting accuracy analysis used 35 ICCVAM reference substances, which produced the following definitive results in agonist testing when compared with existing reference data from other *in vitro* ER TA assays:

- Concordance of 97% (34/35)
- Sensitivity of 96% (27/28)
- Specificity of 100% (7/7)
- False positive rate of 0% (0/7)
- False negative rate of 4% (1/28)

Only L-thyroxine was false negative in the BG1Luc ER TA test method when compared to the ICCVAM reference classification. This reference substance is classified as positive (2/3) based on two reports of positive agonist activity and one report of no agonist activity. The two positive results were in GH3 cells (rat pituitary adenoma) and HeLa cells (human cervical carcinoma), whereas MCF-7 cells (human breast adenocarcinoma) showed no estrogenic response when exposed to L-thyroxine. These results indicate a possible tissue-specific response to L-thyroxine, which may explain the lack of ER agonist activity observed in this experiment with BG-1 cells (human ovarian carcinoma).

During Phase 1, 12 substances were tested in each of the three laboratories (XDS, ECVAM, and Hiyoshi) to evaluate intralaboratory reproducibility. Although the classifications for some of the test substances differed among the laboratories, there was 100% agreement within each laboratory for each of the three repeat tests. When results were compared *across* laboratories for these 12 substances, all three laboratories agreed on 67% (8/12) of the substances. An additional 36 substances tested for agonist activity once in each laboratory produced a definitive result in at

least two laboratories. There was 100% agreement among the laboratories for 83% (30/36) of these substances.

Only one *in vitro* ER TA test method is currently accepted to assess ER α agonist activity of test substances. This test method was validated by the Chemicals Evaluation and Research Institute (CERI) and is described in Organisation for Economic Co-operation and Development (OECD) Test Guideline (TG) 455: the Stably Transfected Human Estrogen Receptor- α Transcriptional Activation (STTA) Assay for the Detection of Estrogenic Agonist-Activity (OECD 2009). Adopted by the EPA as OPPTS 890.1300: Estrogen Receptor Transcriptional Activation (Human Cell Line [HeLa-9903]) (EPA 2009), it is considered adequately validated by national and international regulatory agencies.

Because the BG1Luc ER TA test method is another STTA assay that could be considered for regulatory use, a comparison of test method accuracy between these two test methods was conducted based on a list of ICCVAM-recommended agonist reference substances for which definitive classifications have been produced in both methods. These results show identical levels of accuracy when both methods tested the same agonist reference chemicals: concordance of 95% (24/25), sensitivity of 95% (21/22), and specificity of 100% (4/4). Overall, these data indicate that the BG1Luc ER TA test method is equivalent to the EPA OPPTS 890.1300/OECD TG 455 method for assessing ER α agonist activity.

Based on these results, the BG1Luc ER TA agonist test method can be applied to a wide range of substances, provided they (1) can be dissolved in dimethyl sulfoxide (DMSO), (2) do not react with DMSO or the cell culture medium, and (3) are not toxic to the cells. Although this method may apply to mixtures, none was evaluated in this validation study. Volatile substances may yield acceptable results if CO₂-permeable plastic film is used to seal the test plates, but no volatile substances were evaluated in this validation study. Although relatively few are known, substances with endogenous luminescence or that naturally inhibit luciferase activity cannot be used in this or any other luciferase-based test method. The demonstrated performance of the BG1Luc ER TA agonist test method suggests that data generated with this test method could be routinely considered for prioritization of substances for further testing.

Independent Peer Review Panel Conclusions and Recommendations

The Panel concluded that the available data and test method performance support the ICCVAM draft recommendation that the BG1Luc ER TA test method can be used as a screening test to identify substances with *in vitro* ER agonist activity. However, the Panel emphasized that, because there has been no clear regulatory guidance on how ER TA test methods will be used in the EPA EDSP Program, the use of the BG1Luc ER TA test method in the overall strategy of hazard identification or safety assessment of endocrine-disruptive chemicals is unclear.

2.2.2 Evaluation as a Screening Test to Identify Substances with Estrogen Receptor Antagonist Activity

Based on an evaluation of available data and corresponding performance (accuracy and reliability), ICCVAM recommends that the BG1Luc ER TA test method can be used as a screening test to identify substances with ER antagonist activity. The accuracy analysis, conducted with 25 reference substances, produced the following definitive results in antagonist testing:

- Accuracy of 100% (25/25)
- Sensitivity of 100% (3/3)
- Specificity of 100% (22/22)
- False positive rate of 0% (0/22)
- False negative rate of 0% (0/3)

Intralaboratory reproducibility of the BG1Luc ER TA agonist and antagonist test methods was assessed by comparing (1) reference standard and control results for all plates tested within each laboratory during the course of the validation study and (2) results from Phases 2a and 2b testing, during which 12 substances were tested in at least three independent experiments in each of three laboratories. Although the classifications for some of the test substances differed among the laboratories, there was 100% agreement within each laboratory for each of the three repeat tests.

When results were compared *across* laboratories for these 12 substances, there was 100% agreement among the three laboratories for all 12 substances. An additional 41 substances tested once in each laboratory for antagonist activity during Phase 3 produced a definitive result in at least two laboratories. There was 100% agreement among the laboratories for 93% (38/41) of the 41 substances.

Based on these results, the limitations of the BG1Luc ER TA antagonist test method appear to be the same as those identified for the agonist test method described above. Although the validation database is somewhat limited in number ($n = 25$), the demonstrated performance of the BG1Luc ER TA antagonist test method suggests that data generated with this test method could be routinely considered for prioritization of substances for further testing. This is further supported by the fact that so few ER antagonists have been definitively identified, and all three tested in the BG1Luc ER TA antagonist test method were correctly identified.

Independent Peer Review Panel Conclusions and Recommendations

The Panel concluded that the available data and test method performance support the ICCVAM draft recommendation that the BG1Luc ER TA test method can be used as a screening test to identify substances with *in vitro* ER antagonist activity. The Panel further concluded that, based upon support of the ICCVAM draft recommendation, the BG1Luc ER TA test method could be considered as a replacement for the currently accepted ER TA assay (EPA OPPTS 890.1300/OECD TG 455) and the rat uterine cytosol binding assays. However, the Panel noted that additional analysis may be necessary to further support this recommendation, particularly regarding the rat uterine cytosol ER binding assay.

2.3 ICCVAM Recommendations: Test Method Protocol for the BG1Luc ER TA Test Method

For use of the BG1Luc ER TA test method as a screening test to identify substances with *in vitro* ER agonist or antagonist activity, ICCVAM recommends using the ICCVAM BG1Luc ER TA agonist and antagonist test method protocols (**Appendix B**). In addition, all future studies intended to further characterize the usefulness and limitations of the BG1Luc ER TA agonist and antagonist test methods should be conducted using these recommended protocols.

Independent Peer Review Panel Conclusions and Recommendations

The Panel concluded that the BG1Luc ER TA test method protocols are complete and adequate in detail for a laboratory to conduct the study (see **Appendix D**). The Panel noted several advantages provided by this assay over the currently accepted test method (EPA OPPTS 890.1300/OECD TG 455). The BG1Luc ER TA test method:

- Has more detailed and complete test method protocols than those provided in EPA OPPTS 890.1300/OECD TG 455
- Is validated for testing up to 1 mM per EPA requirements. EPA OPPTS 890.1300/OECD TG 455 is only validated up to a limit dose of 10 μ M.
- Has a more restrictive set of classification criteria for determination of a positive response, which will reduce the number of false positive results, resulting in fewer follow-up tests conducted in animal studies
- Can detect substances with *in vitro* anti-estrogenic activity

- Endogenously expresses both hER α and hER β , whereas the HeLa-9903 cell line used in EPA OPPTS 890.1300/OECD TG 455 was transfected only with hER α

2.4 ICCVAM Recommendations: Future Studies for the BG1Luc ER TA Test Method

ICCVAM promotes the scientific validation and regulatory acceptance of new methods that reduce, refine, or replace animal use where scientifically feasible. The rat uterine cytosol ER binding assay, currently listed as part of the EDSP Tier 1 screening battery, requires the use of animals as a source of ERs. Results from the BG1Luc ER TA test method were examined for concordance with published reports of ER binding for 34 reference substances. There was 97% (33/34) concordance between the BG1Luc ER TA test method and ER binding data from the literature, and 100% sensitivity (no false negatives). In light of the excellent degree of agreement between ER binding and BG1Luc ER TA data, it appears that evaluating results from BG1Luc ER TA agonist and antagonist testing may provide a viable alternative to conducting ER binding studies. This cannot currently be accomplished with EPA OPPTS 890.1300/OECD TG 455 due to the inability of this method to assess ER antagonist activity. ICCVAM recommends that additional validation studies could be performed to determine whether or not the BG1Luc ER TA method could replace the rat uterine cytosol ER binding assay.

Results from the BG1Luc ER TA test method were examined for concordance with published data from the uterotrophic bioassay (n = 13 reference substances), which is currently listed as part of the EDSP Tier 1 screening battery. There was 92% (12/13) concordance between the BG1Luc ER TA test method and the uterotrophic bioassay data, and 100% specificity (no false negatives). These data indicate that the BG1Luc ER TA agonist test method has very good agreement with the *in vivo* results obtained with the uterotrophic bioassay. Accordingly, ICCVAM recommends that further work be carried out to determine if the BG1Luc ER TA test method could be used in combination with other methods (to include *in vitro* metabolic activation) in a weight-of-evidence approach to replace the uterotrophic bioassay.

To further characterize the BG1Luc ER TA test method, ICCVAM identified additional studies that may be considered by interested parties:

- Additional studies/evaluations may be conducted to more completely characterize the ratio of ER α and ER β in the BG-1 cell line and the extent to which these receptor subtypes contribute to the overall performance of the BG1Luc ER TA test method.
- Additional studies/evaluations may be conducted to determine the feasibility of testing volatile substances using CO₂-permeable plastic film or other methods to seal the test plates.
- Additional studies/evaluations may be conducted to determine if substances that are not soluble in DMSO could be tested in another vehicle that would more adequately solubilize the substance in culture media.
- As ER antagonists are identified, additional studies/evaluations may be conducted to expand the database of positive substances tested and thereby better characterize the usefulness and limitations of the BG1Luc ER TA test method as a screening test to identify substances with ER antagonist activity.
- ICCVAM encourages users to provide all data that are generated from future studies to ICCVAM so that they may be used to further characterize the usefulness and limitations of the BG1Luc ER TA test method as a screening test to identify substances with *in vitro* ER agonist or antagonist activity.

Independent Peer Review Panel Conclusions and Recommendations

The Panel concluded that the available data support the draft ICCVAM-recommended future studies. The Panel encouraged additional studies and evaluations to assess the utility of the current visual assessment of cytotoxicity evaluation for chemicals, as well as efforts to identify a

quantitative cytotoxicity method. The Panel also recommended future studies to account for metabolic activation that could expand the utility of this and other ER TA methods. The Panel further recommended an effort to expand the reference substance list and associated BG1Luc ER TA database with additional negative agonist and positive antagonist test substances as they are identified.

2.5 ICCVAM Recommendations: Performance Standards for the BG1Luc ER TA Test Method

ICCVAM has developed test method performance standards so that modified versions of the BG1Luc ER TA test method that are mechanistically and functionally similar can be effectively and efficiently evaluated for their validity by national and international validation organizations (e.g., ICCVAM, ECVAM, and JaCVAM) or other organizations. The ICCVAM-recommended BG1Luc ER TA agonist and antagonist test method protocols are the key references used to establish these performance standards.

Independent Peer Review Panel Conclusions and Recommendations

The Panel concluded that the draft ICCVAM performance standards are adequate, but they proposed modifications that could expand the performance standards' applicability. The Panel suggested that the specific tissue source, type, and species used for the cell system in ER TA test methods may not be critical but recommended that the appropriate cellular machinery be included. The Panel also recommended that, ideally, more negatives should be included. They recognized, however, that data on such substances are not currently available. The Panel also suggested that reference substance classification be based upon reports that have been ranked with a method that focuses on the reliability of the published data (e.g., Klimisch criteria) (Klimisch et al. 1997).

Classification of reference substances was based on the following published guidance from ICCVAM (ICCVAM 2003a, 2006):

- A substance was classified as “positive” if it was reported as positive in >50% of referenced ER TA studies.
- A substance was classified as “presumed positive” if it was positive in 50% or less of referenced ER TA studies.

Prior to the BG1Luc ER TA test method validation study, L-thyroxine was classified as positive because two of three literature citations described estrogenic activity for this compound. Because the BG-1 validation study will be considered a published study, and L-thyroxine was negative in the study, the updated database will reflect that this compound is reported as positive in two of four studies (50%), changing its classification from positive to presumed positive per the guidelines given above. Because only those compounds with definitive classifications (positive or negative) are used as reference substances, L-thyroxine will not be used as a reference substance in future studies.

3.0 Validation Status for Use of the BG1Luc ER TA Test Method as a Screening Test to Identify *In Vitro* ER Agonists and Antagonists

The ICCVAM BRD (see **Appendix C**) provides a comprehensive review of the current validation status of the BG1Luc ER TA test method, including its accuracy and reliability, the substances tested, the rationale for the standardized test method protocol used for the validation study, and all available data supporting its validity. This section provides a brief description and summary of the validation status of the BG1Luc ER TA test method.

3.1 Test Method Description

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 agonist or antagonist activity. ER-mediated transcription of the *luc* gene results in the production of luciferase, the activity of which is quantified using a luminometer. A concentration–response curve can be established to provide qualitative and quantitative information regarding the *in vitro* estrogenic activity of a test substance.

3.2 General Test Method Procedures

ICCVAM previously recommended minimum essential test method components for *in vitro* ER TA assays (ICCVAM 2003a), and these components are incorporated into the ICCVAM-recommended BG1Luc ER TA protocols (see **Appendices B1** and **B2**). These protocols include three sequential phases: solubility, range finder, and comprehensive testing. During solubility testing, the maximum test substance concentration that is soluble in 100% DMSO is established in order to set the starting concentration for range finder testing. The test substance concentration range to be included in comprehensive testing is established during range finder testing. Results from comprehensive testing are used to determine the extent to which a test substance influences ER-mediated luciferase transcription as a correlate to *in vitro* ER TA activity. These data can then be used to classify a test substance based on its *in vitro* ER agonist or antagonist activity.

3.3 Validation Database

The validation database used to evaluate the BG1Luc ER TA test method is based upon the list of 78 substances that ICCVAM recommended for use in validation studies for *in vitro* ER and AR binding and TA test methods (ICCVAM 2003a, 2006). The purpose of this list is to ensure that the usefulness and limitations of *in vitro* ER and AR binding and TA assays can be adequately characterized across a broad range of chemical classes and responses. These substances were selected based on information contained in the ICCVAM BRDs for ER and AR binding and TA test methods (ICCVAM 2002d, 2002a, 2002c, 2002b), as well as information obtained from publications reviewed or published after completion of the ICCVAM BRDs. The complete list of substances and their respective reference classifications for agonist and antagonist activity based on available reference data is provided in Section 3-2 of the BG1Luc ER TA BRD (**Appendix C**).

Only those substances that could be definitively classified as positive (POS) or negative (NEG) were used to assess accuracy, resulting in 48 unique substances used to assess accuracy. (Substances classified as presumed positive [PP] or presumed negative [PN] were not considered when evaluating test method accuracy.) Separate lists were generated for evaluating accuracy based on agonist (42 substances: 33 positive, 9 negative) activity and antagonist (25 substances: 3 positive, 22 negative) activity. Nineteen substances appeared on both reference lists. The 42 reference substances used to assess accuracy based on ER agonist activity are provided in **Table 3-1**, and the 25 reference substances used to assess accuracy based on ER antagonist

activity are provided in **Table 3-2**. These tables also include the BG1Luc ER TA results from each of the participating laboratories.

3.4 Test Method Accuracy

Thirty-five substances (28 positive, 7 negative) had definitive results and were used to evaluate test method accuracy for ER agonist activity. The remaining seven (17%) of the 42 substances used to evaluate test method accuracy had inadequate (I) testing results and were therefore excluded from the analysis. 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 agonist activity. The following seven substances had inadequate BG1Luc ER TA agonist test method data:

- Clomiphene citrate
- *p,p'*-DDE
- 5 α -Dihydrotestosterone
- Flutamide
- Procymidone
- Resveratrol
- Tamoxifen

It should be emphasized that the “inadequate” classification is usually a result of poor data quality and would normally require retesting. However, the classification system was revised after testing to include positive, negative, and inadequate classifications. Retesting of these substances was therefore not possible.

These seven substances (clomiphene citrate, *p,p'*-DDE, 5 α -dihydrotestosterone, flutamide, procymidone, resveratrol, and tamoxifen) represent eight chemical classes (two cyclic hydrocarbons, and one each of an amide, amine, carboxylic acid, halogenated hydrocarbon, heterocyclic compound, polycyclic compound, and steroid) and five product classes (four pharmaceuticals and one each of a fungicide, natural product, pesticide intermediate, and veterinary agent). The diversity of chemical and product classes indicates that no one category or class is overrepresented with inadequate data.

Table 3-1 42 ICCVAM-Recommended Substances Used to Evaluate ER Agonist Accuracy

Substance	CASRN	Classification ^a				
		ICCVAM Consensus	BG1Luc ER TA Consensus ^b	XDS	ECVAM	Hiyoshi
17 α -Estradiol	57-91-0	POS	POS	POS (1/1)	POS (3/3)	POS (2/2)
17 α -Ethinyl estradiol	57-63-6	POS	POS	POS (3/3)	POS (3/3)	POS (3/3)
17 β -Estradiol	50-28-2	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
19-Nortestosterone	434-22-0	POS	POS	POS (1/1)	NT	NT
4-Cumylphenol	599-64-4	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
4- <i>tert</i> -Octylphenol	140-66-9	POS	POS	I (1/1)	POS (1/1)	POS (2/2)
5 α -Dihydrotestosterone	521-18-6	POS	I	I (1/1)	I (1/1)	POS (1/1)
Apigenin	520-36-5	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
Atrazine	1912-24-9	NEG	NEG	NEG (3/3)	POS (3/3)	NEG (3/3)
Bicalutamide	90357-06-5	NEG	NEG	NEG (1/1)	NT	NT
Bisphenol A	80-05-7	POS	POS	POS (3/3)	POS (3/3)	POS (3/3)
Bisphenol B	77-40-7	POS	POS	POS (3/3)	POS (3/3)	POS (3/3)
Butylbenzyl phthalate	85-68-7	POS	POS	POS (3/3)	POS (3/3)	POS (3/3)
Chrysin	480-40-0	POS	POS	POS (2/2)	NT	NT
Clomiphene citrate	50-41-9	POS	I	I (1/1)	NEG (1/1)	POS (1/1)
Corticosterone	50-22-6	NEG	NEG	NEG (3/3)	POS (3/3)	NEG (4/4)
Coumestrol	479-13-0	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
Daidzein	486-66-8	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
Dicofol	115-32-2	POS	POS	POS (1/1)	NEG (1/1)	POS (1/1)
Diethylstilbestrol	56-53-1	POS	POS	POS (3/3)	POS (3/3)	POS (3/3)
Estrone	53-16-7	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
Ethyl paraben	120-47-8	POS	POS	I (1)	POS (1/1)	POS (1/1)
Fenarimol	60168-88-9	POS	POS	POS (1/1)	NT	NT
Flutamide	13311-84-7	NEG	I	I (1)	NT	NT
Genistein	446-72-0	POS	POS	POS (3/3)	POS (3/3)	POS (4/4)
Hydroxyflutamide	52806-53-8	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Kaempferol	520-18-3	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
Kepone	143-50-0	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
L-Thyroxine	51-48-9	POS	NEG	NEG (1/1)	NT	NT
Linuron	330-55-2	NEG	NEG	NEG (1/1)	NT	NT
<i>meso</i> -Hexestrol	84-16-2	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
Methyl testosterone	58-18-4	POS	POS	POS (3/3)	POS (1/1)	POS (2/2)

Substance	CASRN	Classification ^a				
		ICCVAM Consensus	BG1Luc ER TA Consensus ^b	XDS	ECVAM	Hiyoshi
Norethynodrel	68-23-5	POS	POS	POS (2/2)	POS (1/1)	POS (2/2)
<i>o,p'</i> -DDT	789-02-6	POS	POS	POS (3/3)	POS (3/3)	POS (3/3)
<i>p</i> -n-Nonylphenol	104-40-5	POS	POS	POS (3/3)	POS (3/3)	POS (3/3)
<i>p,p'</i> -DDE	72-55-9	POS	I	I (1/1)	I (1/1)	NEG (1/1)
<i>p,p'</i> -Methoxychlor	72-43-5	POS	POS	POS (1/1)	POS (1/1)	POS (2/2)
Phenobarbital	50-06-6	NEG	NEG	NEG (1/1)	NEG (1/1)	NT
Procymidone	32809-16-8	NEG	I	I (1/1)	NT	NT
Resveratrol	501-36-0	POS	I	POS (1/1)	I (1/1)	NEG (2/3)
Spironolactone	52-01-7	NEG	NEG	NEG (1/1)	NT	NT
Tamoxifen	10540-29-1	POS	I	I (1/1)	I (1/1)	POS (1/1)

Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; CASRN = CAS Registry Number (American Chemical Society); ECVAM = European Centre for the Validation of Alternative Methods; I = inadequate (positive or negative classification could not be determined because of poor-quality data); NEG = negative; NT = not tested; POS = positive; XDS = Xenobiotic Detection Systems, Inc.

^a Number in parentheses represents test results (POS, NEG, or I) over the total number of trials that met test plate acceptance criteria.

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

Definitive classifications (positive or negative) were obtained for all 25 substances used to evaluate test method accuracy for ER antagonist activity, allowing all 25 substances to be used to assess antagonist accuracy.

Table 3-2 25 ICCVAM-Recommended Substances Used to Evaluate ER Antagonist Accuracy

Substance	CASRN	Classification ^a				
		ICCVAM Consensus	BG1Luc ER TA Consensus ^b	XDS	ECVAM	Hiyoshi
17 α -Ethinyl estradiol	57-63-6	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
4-Hydroxytamoxifen	68047-06-3	POS	POS	POS (1/1)	I (2/2)	POS (1/1)
5 α -Dihydrotestosterone	521-18-6	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Apigenin	520-36-5	NEG	NEG	NEG (3/3)	NEG (3/3)	NEG (4/4)
Bisphenol A	80-05-7	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Butylbenzyl phthalate	85-68-7	NEG	NEG	NEG (3/3)	NEG (3/3)	NEG (4/4)
Chrysin	480-40-0	NEG	NEG	NEG (1/1)	NT	NT
Coumestrol	479-13-0	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Daidzein	486-66-8	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Di- <i>n</i> -butyl phthalate	84-74-2	NEG	NEG	NEG (2/2)	NEG (1/1)	NEG (1/1)
Dicofol	115-32-2	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Diethylhexyl phthalate	117-81-7	NEG	NEG	NEG (1/1)	NEG (2/2)	NEG (1/1)
Diethylstilbestrol	56-53-1	NEG	NEG	NEG (1/1)	NEG (1/1)	POS (1/1)
Genistein	446-72-0	NEG	NEG	NEG (3/3)	NEG (3/3)	NEG (3/3)
Kaempferol	520-18-3	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Kepone	143-50-0	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Mifepristone	84371-65-3	NEG	NEG	NEG (1/1)	NT	NT
Norethynodrel	68-23-5	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
<i>o,p'</i> -DDT	789-02-6	NEG	NEG	NEG (3/3)	NEG (3/3)	NEG (4/4)
<i>p</i> -n-Nonylphenol	104-40-5	NEG	NEG	NEG (3/3)	NEG (3/3)	NEG (3/3)
<i>p,p'</i> -DDE	72-55-9	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Progesterone	57-83-0	NEG	NEG	NEG (3/3)	NEG (3/3)	NEG (3/3)
Raloxifene HCl	82640-04-8	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
Resveratrol	501-36-0	NEG	NEG	NEG (3/3)	NEG (3/3)	NEG (3/3)
Tamoxifen	10540-29-1	POS	POS	POS (4/4)	POS (3/3)	POS (3/3)

Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; CASRN = CAS Registry Number (American Chemical Society); ECVAM = European Centre for the Validation of Alternative Methods; I = inadequate (positive or negative classification could not be determined because of poor-quality data); NEG = negative; NT = not tested; POS = positive; XDS = Xenobiotic Detection Systems, Inc.

^a Number in parentheses represents test results (POS, NEG, or I) over the total number of trials that met test plate acceptance criteria.

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

The accuracy analysis using the 35 ICCVAM reference substances that produced a definitive BG1Luc ER TA result in agonist testing indicated accuracy of 97% (34/35), sensitivity of 96% (27/28), specificity of 100% (7/7), false positive rate of 0% (0/7), and false negative rate of 4% (1/28) (**Table 3-3**). Analysis of accuracy using individual laboratory results indicated accuracy ranging from 86% (25/29) to 97% (33/34), sensitivity from 92% (23/25) to 96% (27/28),

specificity from 50% (2/4) to 100% (6/6), false positive rates from 0% (0/6) to 50% (2/4), and false negative rates from 4% (1/28) to 8% (2/25).

Table 3-3 Accuracy of the BG1Luc ER TA Agonist Data

Laboratory	N	Accuracy	Sensitivity	Specificity	False Positive Rate	False Negative Rate
Combined	35 ^a	97% (34/35)	96% (27/28)	100% (7/7)	0% (0/7)	4% (1/28)
XDS	34	97% (33/34)	96% (27/28)	100% (6/6)	0% (0/6)	4% (1/28)
ECVAM	29	86% (25/29)	92% (23/25)	50% (2/4)	50% (2/4)	8% (2/25)
Hiyoshi	32	94% (30/32)	93% (27/29)	100% (3/3)	0% (0/3)	7% (2/29)

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

^a A total of 42 substances were evaluated in the BG1Luc ER TA agonist test method. Seven substances did not produce a consensus classification and were omitted, leaving 35 substances for analysis.

The antagonist accuracy analysis indicated an overall accuracy of 100% (25/25), sensitivity of 100% (3/3), specificity of 100% (22/22), false positive rate of 0% (0/22), and false negative rate of 0% (0/3) (**Table 3-4**). Similarly, individual laboratory results indicated accuracy ranging from 96% (22/23) to 100% (25/25), sensitivity of 100% (3/3), and specificity of 95% (19/20) to 100% (22/22).

Table 3-4 Accuracy of the BG1Luc ER TA Antagonist Data

Laboratory	N	Accuracy	Sensitivity	Specificity	False Positive Rate	False Negative Rate
Combined	25	100% (25/25)	100% (3/3)	100% (22/22)	0% (0/22)	0% (0/3)
XDS	25	100% (25/25)	100% (3/3)	100% (22/22)	0% (0/22)	0% (0/3)
ECVAM	23	100% (23/23)	100% (3/3)	100% (20/20)	0% (0/20)	0% (0/3)
Hiyoshi	23	96% (22/23)	100% (3/3)	95% (19/20)	5% (1/20)	0% (0/3)

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

3.5 Test Method Reliability

Intralaboratory reproducibility of the BG1Luc ER TA agonist and antagonist test methods was assessed quantitatively by comparing the following:

- Relative light unit (RLU) values for the agonist and antagonist DMSO control and the antagonist E2 control for all plates tested within each laboratory during the course of the validation study

- Results from Phases 2a and 2b testing, during which 12 substances were tested in at least three independent experiments in each of the three laboratories

Because DMSO control RLU values are not normalized, they vary considerably between test plates and across time. Therefore, intralaboratory reproducibility was evaluated by comparing the within-plate variability of the four replicate DMSO control RLU values for all test plates that passed acceptance criteria (i.e., coefficient of variation [CV] associated with within-plate DMSO control RLU values). The range of means and CV values for within-plate DMSO control RLU values are provided in **Table 3-5**. Mean plate DMSO RLU values ranged from a low of 511 to a high of 9885, with a mean of 3749. However, within-plate variability of DMSO RLU control values between replicate DMSO wells was low. Coefficients of variation ranged from 1% to 43%, with a mean of 8%. Of the 218 agonist test plates that passed acceptance criteria, only six plates had within-plate CV values greater than 20%.

Table 3-5 Agonist Within-Plate DMSO Control Data

Laboratory	Mean and Range of DMSO Control RLU Values	Mean and Range of CV (%)	N
Combined	3749 (511-9885)	8 (1-43)	218
XDS	2800 (511-9885)	8 (1-43)	93
ECVAM	3379 (828-7306)	8 (1-33)	60
Hiyoshi	5465 (1362-9383)	6 (1-24)	65

Abbreviations: CV = coefficient of variation; DMSO = dimethyl sulfoxide; ECVAM = European Centre for the Validation of Alternative Methods; N = number of plates that passed acceptance criteria; RLU = relative light unit; XDS = Xenobiotic Detection Systems, Inc.

For the antagonist assay, although mean plate DMSO RLU values ranged from a low of 132 to a high of 8451 (mean = 3299), within-plate variability of DMSO RLU control values between replicate DMSO wells was low, with CV values ranging from 1% to 52% (mean = 8%) (**Table 3-6**). Of the 194 antagonist test plates that passed acceptance criteria, only eight plates had within-plate CV values greater than 20%.

Table 3-6 Antagonist Within-Plate DMSO Control Data

Laboratory	Mean and Range of DMSO Control RLU Values	Mean and Range of CV (%)	N
Combined	3299 (132-8451)	8 (1-52)	194
XDS	2230 (132-6860)	9 (1-52)	79
ECVAM	3622 (1352-7333)	9 (1-37)	62
Hiyoshi	4030 (1625-8451)	6 (1-20)	53

Abbreviations: CV = coefficient of variation; DMSO = dimethyl sulfoxide; ECVAM = European Centre for the Validation of Alternative Methods; N = number of plates that passed acceptance criteria; RLU = relative light unit; XDS = Xenobiotic Detection Systems, Inc.

Normalized and adjusted antagonist E2 control RLU values were used as acceptance criteria throughout the validation study. The mean, standard deviation (SD), and CV values calculated for the E2 control RLU value from all antagonist test plates that passed acceptance criteria are provided in **Table 3-7**. Mean E2 control RLU values ranged from 5793 at Hiyoshi to 9246 at ECVAM. Variability was low, with associated CV values ranging from 9% at ECVAM to 19% at XDS.

Table 3-7 Antagonist E2 Control Values

Laboratory	Mean RLU	SD	CV (%)	N
XDS	7524	1443	19	79
ECVAM	9246	805	9	62
Hiyoshi	5793	791	14	53

Abbreviations: CV = coefficient of variation; ECVAM = European Centre for the Validation of Alternative Methods; N = number of plates that passed acceptance criteria; RLU = relative light unit; SD = standard deviation; XDS = Xenobiotic Detection Systems, Inc.

Test substances are classified as positive or negative for agonist activity based on a specific set of criteria. The resulting classifications for each of the 12 substances that were tested at least three times at each laboratory were used to evaluate the extent of intralaboratory agreement (see **Table 3-8**). Although the classifications for some of the test substances differed among the laboratories, there was 100% agreement within each laboratory for each of the three repeat tests. There were no “inadequate” data generated at any laboratory during this phase of the validation study.

Table 3-8 Intralaboratory Agreement for Multiple Testing of the 12 Phase 2 Agonist Substances Tested Independently at Least Three Times at Each Laboratory

Activity per Test	XDS	ECVAM	Hiyoshi
Agreement within laboratory	12/12 (100%)	12/12 (100%)	12/12 (100%)
+++	8/12	12/12	9/12
---	4/12	0/12	3/12
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 in two of three replicate tests, a test substance was classified as positive. The substance was classified as negative in a third replicate test.

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

3.6 Animal Welfare Considerations: Reduction, Refinement, and Replacement

The BG1Luc ER TA test method utilizes cultured human ovary adenocarcinoma cells that endogenously express human ER and contain an estrogen-inducible gene expression system. Except for the fetal bovine sera used as part of the cell culture media, the test method does not require the use of animals.

The BG1Luc ER TA test method is being proposed as an independent part of a weight-of-evidence approach to prioritize potentially endocrine-active substances for further testing. Therefore, like the EPA OPPTS 890.1300/OECD TG 455 method, the test does not directly reduce, refine, or replace animal use. However, regulators currently use the following three *in vivo* methods to assess the estrogenic potential of substances: (1) rat uterotrophic assay, (2) rat pubertal female assay, and (3) fish short-term reproduction assay. In addition, the “*in vitro*” rat uterine cytosol ER binding assay also requires the use of animals as a source of ER.

Results from the BG1Luc ER TA test method were examined for concordance with published reports of ER binding. There was 97% (33/34) concordance between the BG1Luc ER TA test method and ER binding data. In light of the excellent degree of agreement between ER binding and BG1Luc ER TA test method results (with no false negative results), it appears that evaluating results from BG1Luc ER TA agonist and antagonist testing may provide a viable alternative to conducting ER binding studies, which use animals as a source of ER. This cannot currently be accomplished with the only accepted ER TA method because of the inability of the EPA OPPTS 890.1300/OECD TG 455 method to assess ER antagonist activity.

Results from the BG1Luc ER TA test method were examined for concordance with published data from the uterotrophic assay. Based on a comparison with the *in vivo* uterotrophic assay classification, the 13 substances with data from the uterotrophic assay and conclusive test results in the BG1Luc ER TA agonist test method produced overall concordance of 92% (12/13). All

substances found positive in the uterotrophic assay were also positive in the BG1Luc ER TA test method. The only discordant substance, butylbenzyl phthalate, was positive for ER agonist activity in the BG1Luc ER TA agonist test method and negative in the uterotrophic assay. These data indicate that the BG1Luc ER TA agonist test method had very good agreement with the *in vivo* results obtained with the uterotrophic assay, with no false negative results.

The development of a battery of *in vitro* and *in silico* methods that can replace animal testing for detecting potential EDs is a biologically complex challenge. The experience derived from validating and using the *in vitro* BG1Luc ER TA test method is expected to contribute to our knowledge and promote progress toward this goal.

4.0 BG1Luc ER TA Test Method Performance Standards

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 interlaboratory reproducibility) and its relevance (i.e., the ability of the test method to correctly predict or measure the biological effect of interest) (ICCVAM 1997, 2003b; OECD 1996, 2005). The purpose of performance standards is to communicate the basis by which new proprietary and nonproprietary test methods have been determined to have sufficient accuracy and reliability for a specific testing purpose. These performance standards can then be used to evaluate the accuracy and reliability of other proposed test methods that are considered functionally and mechanistically similar to the accepted test method.

4.1 Elements of ICCVAM Performance Standards

Performance standards are based on an adequately validated test method and provide a basis for evaluating the comparability of a proposed test method that is functionally and mechanistically similar (ICCVAM 2003b). The three elements of performance standards are the following:

- 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 substances: Reference substances 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.
- Accuracy and reliability values: These are the standards for accuracy and reliability that the proposed test method should meet or exceed when evaluated using the minimum list of reference substances.

4.2 LUMI-CELL (BG1Luc ER TA) Test Method Performance Standards

4.2.1 Background

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 agonist or antagonist activity. The primary objective of this test method is to provide a qualitative assessment of *in vitro* estrogenic activity (i.e., whether a substance is positive or negative for estrogenic activity). Quantitative analysis is also performed to provide additional information on the estrogenic potency of test substances. For example, quantitative analysis can determine the half-maximal effective concentration (EC₅₀) or the half-maximal inhibitory concentration (IC₅₀). Separate protocols are used to identify substances that possess ER agonist or antagonist activity, although the two protocols share most major components (see **Appendices B1 and B2**).

NICEATM coordinated and led an international validation study of the BG1Luc ER TA test method with ECVAM and JaCVAM. The study proceeded in four phases, during which 78 reference substances were tested (see **Appendix C**). Results from this validation study served as the basis for the BG1Luc ER TA test method performance standards, which are applicable for assessing the validity of methods that are functionally and mechanistically similar to the BG1Luc ER TA test method. These performance standards can also be used by naïve laboratories to demonstrate technical proficiency in performing the BG1Luc ER TA test method. The

performance standards consist of (1) essential test method components, (2) reference substances, and (3) an assessment of accuracy and reliability.

4.2.2 BG1Luc ER TA Essential Test Method Components and Other Validation Considerations

Certain principles are important in delineating the essential test method components that determine whether a modified test 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 estrogenic homeostasis *in vivo*. The interaction of estrogens with cellular ERs initiates a cascade of events. A number of *in vitro* endpoints can be used to assess ER–ligand interactions, including receptor binding, cellular proliferation, and transcriptional activation (reporter gene). Unlike receptor binding assays, TA assays can identify whether ligand–receptor association potentiates (agonist) or inhibits (antagonist) estrogenic signaling (Davenport and Russell 1996).

In the BG1Luc ER TA test method, ER-mediated transcription of the *luc* gene results in the production of luciferase, the activity of which is quantified using a luminometer. A concentration–response curve can be established to provide qualitative and quantitative information regarding the *in vitro* estrogenic activity of a test substance.

4.2.2.1 Essential Test Method Components

ICCVAM previously recommended minimum essential test method components for *in vitro* ER TA test method protocols (ICCVAM 2003a). These components were incorporated into the BG1Luc ER TA test method protocols during a protocol standardization study. During the protocol standardization study, protocols were developed for use in the international validation study (see **Appendices B1** and **B2**). During the multiphase validation study, the protocols were refined, ultimately resulting in optimized protocols for agonist and antagonist testing. In order to be considered functionally and mechanistically similar to the BG1Luc ER TA test method, a modified ER TA test method protocol must include the following components, which are based on the optimized test method protocols, to ensure that the same biological effect is being measured. If any of these criteria are not met, then these performance standards cannot be used for validation of the modified test method.

Cell Line

The BG1Luc ER TA test method is based on a human ovarian adenocarcinoma cell line that endogenously expresses ER α (90%) and ER β (10%) (Pujol et al. 1998) and uses a stably transfected luciferase-based reporter gene system. Other cell lines that endogenously express human ERs and are stably transfected with a reporter gene system may be appropriate for validation using these performance standards.

Solvent

Reference standards, controls, and test substances should be dissolved in a solvent (e.g., 1% DMSO) that is miscible with cell culture media at concentrations that are not cytotoxic and that do not otherwise interfere with the test system.

Limit Concentration and Cytotoxicity

The maximum test substance concentration should be 1 mM for ER TA agonist testing and 10 μ M for ER TA antagonist testing unless otherwise limited by solubility, cytotoxicity, or other mechanisms that interfere with assay performance. A minimum of seven concentrations spaced at logarithmic (log10) intervals, up to the limit concentration, should be tested. An evaluation of cytotoxicity and how it is applied to the test method should be included in each study. Any

concentration of test substance that reduces viability by greater than 20% should not be considered in the analysis of the data.

Reference Standards

A reference estrogen (e.g., 17 β -estradiol [E2]) and a reference anti-estrogen (e.g., raloxifene HCl) should be used as reference standards to demonstrate the adequacy of the test method for detecting ER TA agonist and antagonist activity, respectively. The ability of the reference estrogen to induce ER TA activity and the reference anti-estrogen to inhibit ER TA activity should be demonstrated by generating a full concentration–response curve in each experiment. At a minimum, the E2 reference standard should provide a threefold induction relative to the solvent control. For antagonist testing, a minimum threefold reduction in the reference anti-estrogenic standard response (e.g., raloxifene HCl) should be demonstrated.

Controls

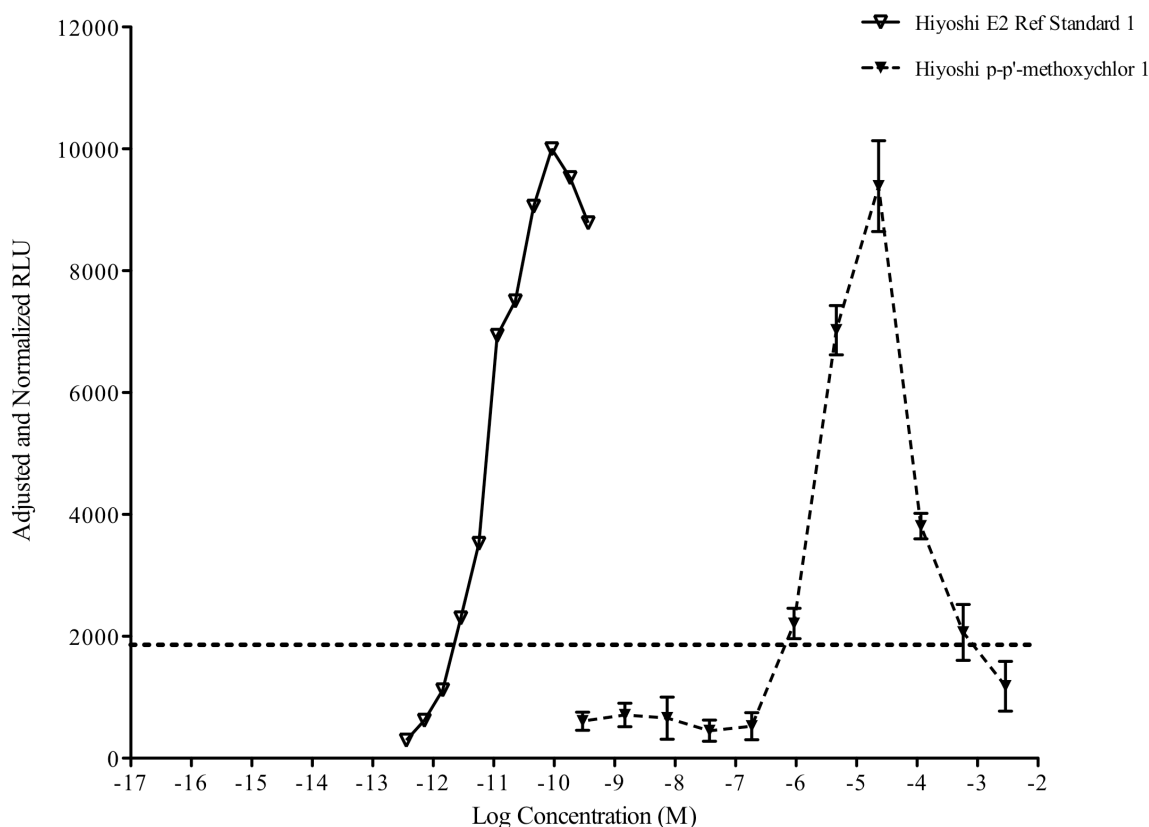
A set of concurrent controls (i.e., solvent, cell culture media) should be included in each experiment to provide a measure of ER TA activity in the absence of reference standards or test substances. A weak positive agonist control (e.g., *p.p'*-methoxychlor) with an EC₅₀ five to six orders of magnitude higher than the reference estrogen should be included in each ER TA agonist study to demonstrate that the test method is functioning properly and is sufficiently sensitive to detect weak ER agonists. A weak positive antagonist control (e.g., tamoxifen) that demonstrates ER TA antagonist activity slightly below the 10 μ M limit concentration should be included in each ER TA antagonist study to demonstrate that the test method is functioning properly and is sufficiently sensitive to detect weak ER antagonists. 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 antagonistic activity of test substances can be assessed.

Interpretation of Results

For ER TA agonist testing:

- All test substances classified as positive for ER TA agonist activity should have a concentration–response curve consisting of a baseline followed by a positive slope, concluding in a plateau or peak. In some cases, only two of these characteristics (baseline–slope or slope–peak) may be defined.
- The line defining the positive slope must contain at least three points with nonoverlapping error bars (mean \pm SD). Points forming the baseline are excluded, but the linear portion of the curve may include the peak or first point of the plateau.
- A positive classification requires a response amplitude, the difference between baseline and peak, of at least 20% of the average maximal value of the reference estrogen, e.g., 2000 RLUs when the maximal response value of the reference estrogen is adjusted to 10,000 RLUs. (See **Figure 4-1** for an example of a concentration–response curve for a substance that is positive for ER TA agonist activity.)
- If possible, an EC₅₀ value should be calculated for each positive substance.
- For all concentration–response curves that fail to meet the criteria for a positive response, test substances are classified as negative for agonist activity if all data points are below 20% of the maximal value for the reference estrogen, e.g., 2000 RLUs when the maximal response value of the reference estrogen is adjusted to 10,000 RLUs.

Figure 4-1 Example Concentration–Response Curve for an ER TA Agonist



Abbreviations: E2 = 17 β -estradiol; M = molar; RLU = relative light unit.

Horizontal dotted line represents 20% of the maximum response of the E2 reference standard.

Test substance shown is *p,p'*-methoxychlor.

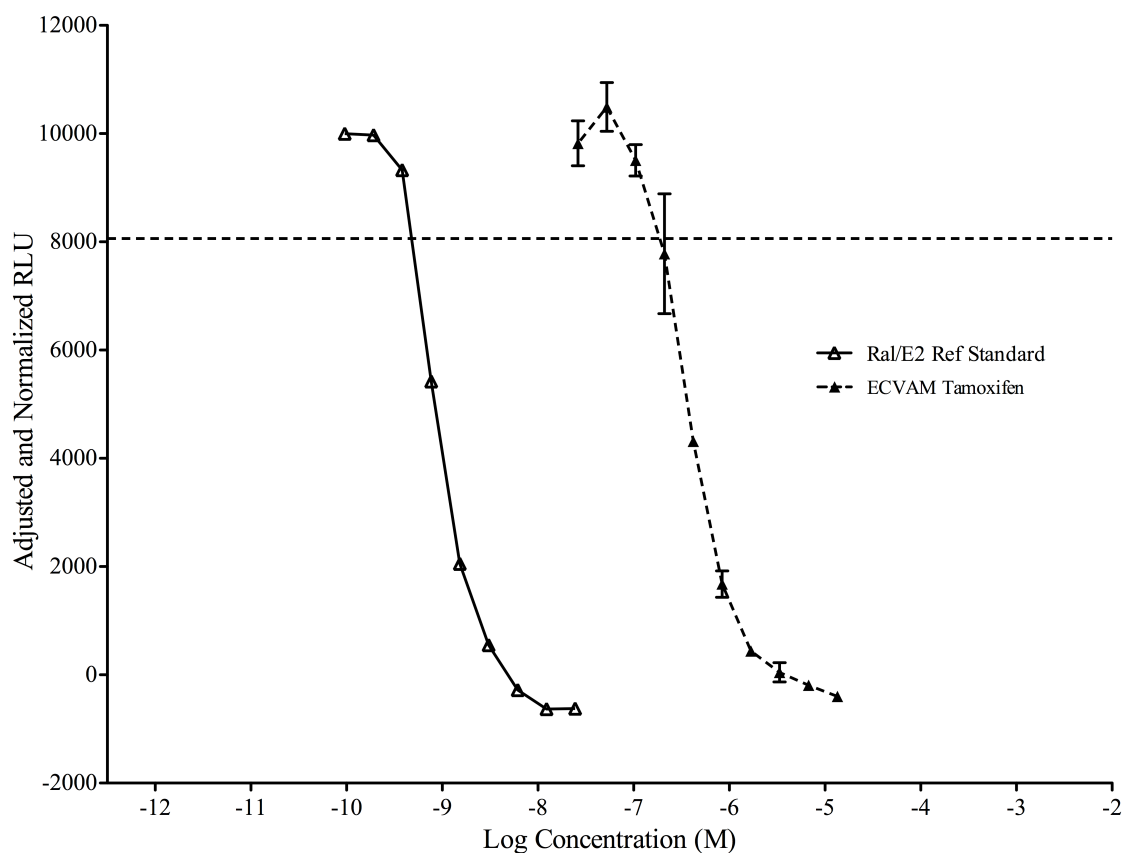
E2 reference standard data is presented as the mean value of duplicate wells.

p,p'-Methoxychlor data are presented as the mean and SD values of three replicate wells.

For ER TA antagonist testing:

- All substances classified as positive for ER antagonist activity should have a concentration–response curve consisting of a baseline followed by a negative slope.
- The line defining the negative slope must contain at least three points with nonoverlapping error bars (representative of means \pm SDs). Points forming the baseline are excluded, but the linear portion of the curve may include the first point of the plateau.
- A positive classification requires a response amplitude of less than 80% of the value for the reference estrogen. The response amplitude is defined as the difference between the baseline, established by the reference estrogen, and the bottom of the dose–response curve.
- The highest noncytotoxic concentrations of the test substance should be less than or equal to 10 μ M. (See **Figure 4-2** for an example of a concentration–response curve for a substance that is positive for ER TA antagonist activity.)
- Test substances are classified as negative for ER antagonist activity if all data points are above 80% of the reference estrogen response, or 8000 RLU.

Figure 4-2 Example Concentration–Response Curve for an ER TA Antagonist



Abbreviations: M = molar; RLU = relative light unit.

Horizontal dotted line represents 80% of the response of the 17β-estradiol reference estrogen.

Test substance shown is tamoxifen.

Ral/E2 reference standard data are presented as the mean value of duplicate wells.

Tamoxifen data are presented as the mean and SD values of three replicate wells.

Data and Reporting

The validation report should include the following information:

—Reporter Plasmid (if different than that used in BG1Luc ER TA test method)

- Type and structure of ER response elements
- Description of promoter region
- Name, identification, and source of original plasmid used to make construct
- Description and methodology used to make the transfected plasmid
- Nomenclature and genetic components comprising the reporter construct

—Cell Line

- Source and nomenclature of the cell line and protocol for its maintenance before and after transfection
- Source of cell culture media, materials, and supplies
- Passage number of subcultures used in the study
- Methods for maintaining stably transfected cell line

- Methods used to monitor the stability of the cell line used for testing
- Rationale, based on data, for deciding on the number of passages a cell line can undergo without a decrease in activity
- Details regarding selection requirements needed to maintain stable cell lines
- If known, details regarding the relative amounts of ER α and ER β

—*Test Method Conditions*

- Composition of media and reagents used
- Incubation volume, duration, and temperature
- Method used to measure ER TA activity
- Methods used to evaluate data, determine response, and calculate EC₅₀ or IC₅₀ values

—*Reference Standards, Controls, and Test Substances*

- Name, chemical structure, CAS Registry Number (CASRN), purity, and supplier
- Physicochemical properties relevant to the study (e.g., solubility, pH, stability, volatility)
- Concentrations and volumes used

—*Solvent*

- Name, CASRN, purity, and supplier
- Justification for choice of solvent
- Information on the solubility of test substances in solvent used
- Information to demonstrate that the solvent, at the maximum volume used, is not cytotoxic and does not otherwise interfere with the study

—*Criteria for an Acceptable Test*

- Concurrent reference standard and control data
- Laboratory-specific historical ranges of reference standard and control data
- Definition of exclusion criteria and description of the impact of any excluded data

—*Results*

- Reference standard and control results
- Test substance solubility results
- Test substance cell viability results
- Calculated reference standard and test substance EC₅₀ and IC₅₀ values
- Graphically presented reference standard, control, and test substance results

—*Discussion of Results*

- Impact of solubility and cytotoxicity on test results
- Reproducibility of reference standard and control data

—*Conclusion*

- Classification of test substances with regard to *in vitro* ER TA agonist or antagonist activity

Other Validation Considerations

The following additional points should be considered during the validation of test methods that are functionally and mechanistically similar to the BG1Luc ER TA test method:

- Appropriate quality assurance systems (i.e., in accordance with Good Laboratory Practice guidelines (EPA 2006b, 2006a; FDA 2009; OECD 1998) are required.
- The study should be conducted according to U.S. (ICCVAM 1997) and international validation principles (OECD Guidance Document 34) (OECD 2005).

4.2.3 Reference Substances for *In Vitro* ER TA Test Methods

To ensure that a proposed *in vitro* ER TA test method possesses reliability and accuracy characteristics similar to those of the validated test method (in this case the BG1Luc ER TA test method), the proposed test method should use at least the agonist reference substances listed in **Table 4-1** and the antagonist reference substances listed in **Table 4-2**. All substances should be tested in a coded/blinded manner. When evaluated using these reference substances, the accuracy (i.e., sensitivity, specificity, false positive rates, and false negative rates) and reliability of the proposed ER TA test method should approximate those of the validated ER TA test method, as detailed in **Section 4.2.4**. Although it is not realistic to expect test methods to perform identically, discordant results should be addressed in terms of the test method's ability to accurately classify other substances with similar potencies and from the same chemical/product classes.

4.2.3.1 Criteria for Selection of Reference Substances

ICCVAM previously compiled and recommended a list of 78 substances for use in validation studies for *in vitro* ER and AR binding and TA test methods (ICCVAM 2003a, 2006). These substances were selected based on information contained in the ICCVAM BRDs for AR and ER binding and TA test methods (ICCVAM 2002d, 2002a, 2002c, 2002b), as well as information obtained from publications reviewed or published after completion of the ICCVAM BRDs.

Factors and criteria considered necessary for selecting reference substances included:

- A well-defined chemical structure
- Comparatively low systemic toxicity
- Good availability from commercial sources
- A concentration–response range that could be measured or predicted by the test method
- Minimal disposal cost

Because the BG1Luc ER TA test method is used only to detect substances with *in vitro* ER TA agonist or antagonist activity, the following criteria were used to classify each reference substance with respect to ER TA agonist and antagonist activity:

- A substance was classified as POS if it was reported as positive in >50% of referenced ER TA studies.
- A substance was classified as NEG if it was reported as negative in all referenced ER TA studies (at least two studies were required for negative classification).
- A substance was classified as PP (presumed positive) if it was positive in 50% or fewer referenced ER TA studies, or if it was reported positive in the single study conducted.
- A substance was classified as PN (presumed negative) if it was reported negative in a single ER TA study.
- Substances without data were classified as PP or PN based on other available information, including their known mechanism of action or their responses in other ER assays.

Only those substances that could be definitively classified as POS or NEG were used to assess accuracy (substances classified as PP or PN were not considered when evaluating test method accuracy). Accordingly, this subset of substances was used to select the final list of reference substances listed in **Tables 4-1** and **4-2**. Recognizing that the number of available reference substances that are definitively negative for agonist activity (**Table 4-1**) or definitively positive for antagonist activity (**Table 4-2**) is limited, these lists may be updated as additional substances with these characteristics are identified. Accordingly, users should be aware that the reference substance list could be revised based on any additional studies that are conducted in the future. ICCVAM recommends that users consult the NICEATM–ICCVAM website (<http://iccvam.niehs.nih.gov/>) to ensure use of the most current reference substance list.

Table 4-1 34 Reference Substances for Evaluation of ER Agonist Accuracy

Substance ^a	CASRN	ICCVAM Consensus	BG1Luc ER TA Consensus ^b	BG1Luc ER TA Mean EC ₅₀ (M) ^c	MeSH Chemical Class ^d	Product Class ^e
Ethyl paraben	120-47-8	POS	POS	2.48×10^{-5}	Carboxylic Acid, Phenol	Pharmaceutical, Preservative
Fenarimol	60168-88-9	POS	POS	4.59×10^{-6}	Heterocyclic Compound, Pyrimidine	Fungicide
Kaempferol	520-18-3	POS	POS	3.99×10^{-6}	Flavonoid, Heterocyclic Compound	Natural Product
Methyl testosterone	58-18-4	POS	POS	3.29×10^{-6}	Steroid	Pharmaceutical, Veterinary Agent
Chrysin	480-40-0	POS	POS	3.20×10^{-6}	Flavonoid, Heterocyclic Compound	Natural Product
<i>p</i> -n-Nonylphenol	104-40-5	POS	POS	3.06×10^{-6}	Phenol	Chemical Intermediate
Dicofol	115-32-2	POS	POS	2.22×10^{-6}	Hydrocarbon (Cyclic), Hydrocarbon (Halogenated)	Pesticide
Butylbenzyl phthalate	85-68-7	POS	POS	1.98×10^{-6}	Carboxylic Acid, Ester, Phthalic Acid	Plasticizer, Industrial Chemical
<i>p,p'</i> -Methoxychlor	72-43-5	POS	POS	1.92×10^{-6}	Hydrocarbon (Halogenated)	Pesticide, Veterinary Agent
Apigenin	520-36-5	POS	POS	1.85×10^{-6}	Heterocyclic Compound	Dye, Natural Product, Pharmaceutical Intermediate
19-Nortestosterone	434-22-0	POS	POS	1.80×10^{-6}	Steroid	Pharmaceutical, Veterinary Agent
Daidzein	486-66-8	POS	POS	8.71×10^{-7}	Flavonoid, Heterocyclic Compound	Natural Product
Bisphenol A	80-05-7	POS	POS	5.33×10^{-7}	Phenol	Chemical Intermediate, Flame Retardant, Fungicide
Kepone	143-50-0	POS	POS	4.91×10^{-7}	Hydrocarbon (Halogenated)	Pesticide
<i>o,p'</i> -DDT	789-02-6	POS	POS	3.94×10^{-7}	Hydrocarbon (Halogenated)	Pesticide
4-Cumylphenol	599-64-4	POS	POS	3.20×10^{-7}	Phenol	Chemical Intermediate
Genistein	446-72-0	POS	POS	2.71×10^{-7}	Flavonoid, Heterocyclic Compound	Natural Product, Pharmaceutical

Substance ^a	CASRN	ICCVAM Consensus	BG1Luc ER TA Consensus ^b	BG1Luc ER TA Mean EC ₅₀ (M) ^c	MeSH Chemical Class ^d	Product Class ^e
Bisphenol B	77-40-7	POS	POS	1.67×10^{-7}	Phenol	Chemical Intermediate, Flame Retardant, Fungicide
Coumestrol	479-13-0	POS	POS	8.77×10^{-8}	Heterocyclic Compound	Natural Product
4- <i>tert</i> -Octylphenol	140-66-9	POS	POS	3.19×10^{-8}	Phenol	Chemical Intermediate, Pharmaceutical Intermediate
17 α -Estradiol	57-91-0	POS	POS	1.54×10^{-9}	Steroid	Pharmaceutical, Veterinary Agent
Norethynodrel	68-23-5	POS	POS	9.39×10^{-10}	Steroid	Pharmaceutical
Estrone	53-16-7	POS	POS	2.57×10^{-10}	Steroid	Pharmaceutical, Veterinary Agent
Diethylstilbestrol	56-53-1	POS	POS	3.34×10^{-11}	Hydrocarbon (Cyclic)	Pharmaceutical, Veterinary Agent
<i>meso</i> -Hexestrol	84-16-2	POS	POS	1.65×10^{-11}	Steroid	Pharmaceutical, Veterinary Agent
17 β -Estradiol	50-28-2	POS	POS	8.37×10^{-12}	Steroid	Pharmaceutical, Veterinary Agent
17 α -Ethinyl estradiol	57-63-6	POS	POS	7.31×10^{-12}	Steroid	Pharmaceutical, Veterinary Agent
Atrazine	1912-24-9	NEG	NEG	-	Heterocyclic Compound	Herbicide
Bicalutamide	90357-06-5	NEG	NEG	-	Amide	Pharmaceutical
Corticosterone	50-22-6	NEG	NEG	-	Steroid	Pharmaceutical
Hydroxyflutamide	52806-53-8	NEG	NEG	-	Amide	Pharmaceutical
Linuron	330-55-2	NEG	NEG	-	Urea	Herbicide
Phenobarbital	50-06-6	NEG	NEG	-	Heterocyclic Compound, Pyrimidine	Pharmaceutical, Veterinary Agent
Spironolactone	52-01-7	NEG	NEG	-	Lactone, Steroid	Pharmaceutical

Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; CASRN = CAS Registry Number (American Chemical Society); EC₅₀ = half-maximal effective concentration of a test substance; 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 Substances are listed in order based upon EC₅₀ values.

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

^c Mean EC₅₀ values were calculated with values reported by the laboratories of the BG1Luc ER TA validation study (XDS, ECVAM, and Hiyoshi).

^d Substances were assigned to one or more chemical or product 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 Substances 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>).

Table 4-2 10 Reference Substances for Evaluation of ER Antagonist Accuracy

Substance ^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	8.17×10^{-7}	Hydrocarbon (Cyclic)	Pharmaceutical
4-Hydroxytamoxifen	68047-06-3	POS	POS	2.08×10^{-7}	Hydrocarbon (Cyclic)	Pharmaceutical
Raloxifene HCl	82640-04-8	POS	POS	1.19×10^{-9}	Hydrocarbon (Cyclic)	Pharmaceutical
17 α -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); IC₅₀ = 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 Substances are listed in order based upon IC₅₀ values.

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

^c Mean IC₅₀ values were calculated with values reported by the laboratories of the BG1Luc ER TA validation study (XDS, ECVAM, and Hiyoshi).

^d Substances 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 Substances 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>).

4.2.3.2 Characteristics of Selected Reference Substances

The reference substances include a range of chemical and product classes representative of the classes commonly associated with endocrine disruption.

Agonist and antagonist test method intralaboratory reproducibility was evaluated using nine substances and four substances, respectively, that were each tested three times on three separate days at each laboratory. Agonist and antagonist test method interlaboratory reproducibility was

evaluated using 27 and 8 substances, respectively, that were tested at least once in each laboratory during the validation study.

4.2.4 Accuracy and Reliability Performance Values

The final elements of performance standards are the accuracy and reliability values (i.e., test method performance) that should be met or exceeded by the proposed test method when evaluated with the reference substances. *Accuracy* is defined as the closeness of agreement between a test method result and an accepted reference value. *Reliability* is the degree to which a test method can be performed reproducibly within and among laboratories over time (ICCVAM 2003b). For these performance standards, the proposed test method should have accuracy and reliability characteristics that approximate those of the validated ER TA test method, which are detailed below. Although it is not realistic to expect test methods to perform identically, discordant results should be addressed in terms of the test method's ability to accurately classify other substances with similar potencies and from the same chemical/product classes.

4.2.4.1 Test Method Accuracy

The analysis of agonist activity for the 34 substances in **Table 4-1** indicated an overall accuracy of 100% (34/34), sensitivity of 100% (27/27), specificity of 100% (7/7), false positive rate of 0% (0/7), and false negative rate of 0% (0/27).

The analysis of antagonist activity for the 10 substances in **Table 4-2** indicated an overall accuracy of 100% (10/10), sensitivity of 100% (3/3), specificity of 100% (7/7), false positive rate of 0% (0/7), and false negative rate of 0% (0/3).

4.2.4.2 Test Method Reliability

For the BG1Luc ER TA agonist test method, there was 100% agreement within each laboratory for each of the three repeat tests for nine reference substances tested in Phase 2 of the agonist validation study. When results were compared across laboratories for these nine substances, there was 78% (7/9) agreement among the three laboratories for the substances. An additional 17 substances tested once in each laboratory for agonist activity produced a definitive result in at least two laboratories. There was agreement among the laboratories for 82% (14/17) of these substances.

For the BG1Luc ER TA antagonist test method, there was 100% agreement within each laboratory for each of the three repeat tests for four reference substances tested in Phase 2 of the antagonist validation study. When results were compared across laboratories for these four substances, there was 100% agreement among the three laboratories for all four substances. An additional five substances tested once in each laboratory for antagonist activity produced a definitive result in at least two laboratories. There was agreement among the laboratories for 80% (4/5) of these substances.

5.0 ICCVAM Consideration of Public Comments

The ICCVAM evaluation process incorporates a high level of scientific peer review and transparency. The evaluation process on the use of the BG1Luc ER TA test method as a screening method to identify *in vitro* ER agonists and antagonists included one public review meeting by an independent scientific peer review panel, multiple opportunities for public comments, and comments from SACATM. ICCVAM and the EDWG considered the Panel report, SACATM comments, and all public comments before finalizing the ICCVAM test method evaluation report and BRD for the use of the BG1Luc ER TA test method. This section summarizes the ICCVAM consideration of public comments (see **Appendix E**).

5.1 ICCVAM Consideration of Public and SACATM Comments

Six opportunities for public comment were provided during the ICCVAM evaluation of the BG1Luc ER TA test method (**Table 5-1**). A total of nine comments were submitted. *Federal Register* notices published by NICEATM–ICCVAM during evaluation of the BG1Luc ER TA test method are available in **Appendix E** and from the NICEATM–ICCVAM website (<http://iccvam.niehs.nih.gov/>). Comments received in response to or related to the *Federal Register* notices are available on the NICEATM–ICCVAM website.⁴ The following sections, delineated by *Federal Register* notice and public meeting, briefly discuss the public comments received.

Table 5-1 Opportunities for Public Comments

Opportunity for Public Comment	Date	Number of Public Comments Received
69 FR 21564 - <i>In Vitro</i> Endocrine Disruptor Test Methods: Request for Comments and Nominations	April 21, 2004	0
71 FR 13597 - Notice of Availability of a Revised List of Recommended Reference Substances for Validation of <i>In Vitro</i> Estrogen and Androgen Receptor Binding and Transcriptional Activation Assays: Request for Comments and Submission of <i>In Vivo</i> and <i>In Vitro</i> Data	March 16, 2006	0
74 FR 62317 - Evaluation of <i>In Vitro</i> Estrogen Receptor Transcriptional Activation and <i>In Vitro</i> Cell Proliferation Assays for Endocrine Disruptor Chemical Screening: Request for Nominations for an Independent Expert Peer Review Panel and Submission of Relevant <i>In Vitro</i> and <i>In Vivo</i> Data	November 27, 2009	6
76 FR 4113 - Announcement of an Independent Scientific Peer Review Panel Meeting on an <i>In Vitro</i> Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening; Availability of Draft Background Review Document (BRD); Request for Comments	January 24, 2011	1
76 FR 23323 - Meeting of the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM)	April 26, 2011	2
76 FR 28781 - Independent Scientific Peer Review Panel Report: Evaluation of the Validation Status of an <i>In Vitro</i> Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening: Notice of Availability and Request for Public Comments	May 18, 2011	0

⁴ <http://ntp-apps.niehs.nih.gov/iccvampb/searchPubCom.cfm>

5.1.1 Public Comments in Response to 69 FR 21564 (April 21, 2004)

***In Vitro* Endocrine Disruptor Test Methods: Request for Comments and Nominations**

NICEATM requested nomination of ER and AR binding and TA test methods for validation studies.

No public comments were received in response to this *Federal Register* notice.

5.1.2 Public Comments in Response to 71 FR 13597 (March 16, 2006)

Notice of Availability of a Revised List of Recommended Reference Substances for Validation of *In Vitro* Estrogen and Androgen Receptor Binding and Transcriptional Activation Assays: Request for Comments and Submission of *In Vivo* and *In Vitro* Data

NICEATM announced the availability of an addendum (ICCVAM 2006) to the *ICCVAM Evaluation of In Vitro Test Methods for Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays* (ICCVAM 2003a). The addendum describes the rationale for proposed revisions to the original list of recommended reference substances for validation of *in vitro* ER and AR binding and TA assays. NICEATM requested public comments on the substances proposed as substitutes for six of the 78 substances in the original list. Data were also requested from *in vitro* and *in vivo* studies evaluating the estrogenic and androgenic activity of the 78 substances in the revised list of reference substances.

No public comments were received in response to this *Federal Register* notice.

5.1.3 Public Comments in Response to 74 FR 62317 (November 27, 2009)

Evaluation of *In Vitro* Estrogen Receptor Transcriptional Activation and *In Vitro* Cell Proliferation Assays for Endocrine Disruptor Chemical Screening: Request for Nominations for an Independent Expert Peer Review Panel and Submission of Relevant *In Vitro* and *In Vivo* Data

NICEATM requested:

- Nominations of expert scientists for consideration as potential Panel members
- Submission of existing data from the LUMI-CELL ER and the CertiChem MCF-7 cell proliferation assays
- Submission of data from *in vivo* or other *in vitro* assessments for the 78 reference substances recommended by ICCVAM for the validation of *in vitro* ER and AR binding and TA test methods

NICEATM received six public comments in which nine potential panelists were nominated for consideration. The nominees were included in the database of experts from which the Panel was selected.

5.1.4 Public Comments in Response to 76 FR 4113 (January 24, 2011)

Announcement of an Independent Scientific Peer Review Panel Meeting on an *In Vitro* Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening; Availability of Draft Background Review Document (BRD); Request for Comments

NICEATM invited public comments on the draft BRD and draft ICCVAM test method recommendations. One public comment was received that included a number of suggestions.

The commenter proposed assigning a level of confidence ranking to the reference data. Substances for which there is a low degree of confidence in the reference data should be deleted from the reference list and omitted from validation studies. With regard to specific test

substances, the commenter stated that the discordant results among laboratories for atrazine, corticosterone, and dicofol were not fully explained.

ICCVAM Response

The independent scientific peer review panel concluded that it is reasonable to use the majority classification criteria among published study results (i.e., >50%) to establish the consensus reference classification for each reference substance. The Panel suggested that this approach could be improved by a ranking method, such as Klimisch criteria (Klimisch et al. 1997), that focuses primarily on the reliability of the data. Such a method would clarify the relative quality of the reference data and strengthen the classification. ICCVAM concurred that additional review and ranking of the published reports would strengthen the utility of literature citations for classifying the reference substances and agreed to take this into consideration in future evaluations.

The commenter questioned the use of flavone as the weak positive control in the antagonist protocol. The commenter further stated that differences among the laboratories in range finder starting concentrations were not fully explained.

ICCVAM Response

During protocol standardization, a number of substances were evaluated for use as the weak antagonist control. Flavone produced a dose response and an $IC_{50} = 4.3 \times 10^{-7} M$, which was consistent with the single literature reference for this compound (reported $IC_{50} = \sim 15 \mu M$) and was two times below that of raloxifene. Based on these results, flavone was chosen as the weak antagonist control for the validation study. However, after review of the data from the completed study, it was apparent that the vast majority of test substances classified as “negative” or “presumed negative” produced a “positive” response at concentrations above $\sim 10 \mu M$. Use of flavone as a weak antagonist control was therefore reconsidered.

The commenter suggested including quantitative comparison of test substances (such as EC_{50} values) and indicated that it would be helpful to include data presented as a relative potency index (the EC_{50} of the positive control divided by the EC_{50} of the test substances, multiplied by 100).

ICCVAM Response

Quantitative measures of activity (i.e., EC_{50} and IC_{50} values) were generated and presented in the BRD. The independent scientific peer review panel considered the descriptive approach for evaluating test method reliability acceptable but also suggested additional statistical analyses that could be performed to better characterize and clarify variability. The Panel suggested that a quantitative measure of activity should be included in each future study report, and the uncertainty associated with these estimates should also be reported.

5.1.5 Public Comments in Response to 76 FR 23323 (April 26, 2011)

Meeting of the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM)

NICEATM announced the SACATM meeting and requested written and public oral comments on the agenda topics. Two public comments were received.

One commenter supported the validation of the BG1Luc ER TA test method and recommended modifications of the protocol that would allow for the implementation of a liquid handling system. The commenter felt that the use of a liquid handling system would greatly increase sample throughput.

ICCVAM Response

The use of a liquid handling system represents a potential improvement to the protocol that could increase throughput. Use of a liquid handling system at the lead laboratory was considered during

the initial phases of the validation study. However, because of difficulties experienced with the system that was acquired at the outset of the study, a decision was made to focus on the “benchtop” version of the assay and perhaps reconsider incorporating automated procedures into the assay at a later time.

A second commenter also supported the validation of the BG1Luc ER TA test method and recommended improvements. The commenter recommended a quantitative comparison of the BG1Luc ER TA data to EPA OPPTS 890.1300/OECD TG 455 data and development of a relative potency index for the reference substances.

ICCVAM Response

As stated above, ICCVAM concurred that additional review and ranking of the published reports would strengthen the utility of literature citations for classifying the reference substances. A comparison of median EC₅₀ and IC₅₀ values from the BG1Luc ER TA test method and literature references is provided in the BRD (**Appendix C**). A relative potency index for the reference substances has not been calculated; however, data provided in the current review permit calculation of such an index.

5.1.6 Public Comments in Response to 76 FR 28781 (May 18, 2011)

Independent Scientific Peer Review Panel Report: Evaluation of the Validation Status of an *In Vitro* Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening: Notice of Availability and Request for Public Comments

NICEATM requested submission of written public comments on the *Peer Review Panel Report: Evaluation of the LUMI-CELL ER[®] (BG1Luc ER TA) Test Method* (**Appendix D2**). No comments were received in response to this request.

6.0 References

- Food Quality Protection Act of 1996. 7 U.S.C. 136. Public Law 104-170.
- Federal Food, Drug, and Cosmetic Act of 1938 (as amended). 21 U.S.C. 301 et seq. Public Law 75-717.
- ICCVAM Authorization Act of 2000. 42 U.S.C. 285f-3. Public Law 106-545. Available: http://frwebgate.access.gpo.gov/cgi-bin/useftp.cgi?IPaddress=162.140.64.183&filename=publ545.pdf&directory=/disks/wais/data/106_cong_public_laws [accessed 17 September 2009].
- Safe Drinking Water Act Amendments of 1996. 110 Stat 1613. Public Law 104-182.
- Davenport AP, Russell FD. 1996. Radioligand Binding Assays: Theory and Practice. In: Current Directions in Radiopharmaceutical Research and Development. (Mather SJ, ed). Boston:Kluwer.
- EPA. 1997. Special Report on Environmental Endocrine Disruption: An Effects Assessment and Analysis. 630/R-96/012. Washington, DC:U.S. Environmental Protection Agency. Available: <http://www.epa.gov/raf/publications/pdfs/ENDOCRINE.PDF>.
- EPA. 2006a. Good Laboratory Practice Standards. Federal Insecticide, Fungicide and Rodenticide Act. 40 CFR 160. Available: http://www.access.gpo.gov/nara/cfr/waisidx_06/40cfr160_06.html [accessed 1 November 2010].
- EPA. 2006b. Good Laboratory Practice Standards. Toxic Substances Control Act. 40 CFR 792. Available: http://www.access.gpo.gov/nara/cfr/waisidx_06/40cfr792_06.html [accessed 1 November 2010].
- EPA. 2009. Endocrine Disruptor Screening Program Test Guidelines: OPPTS 890.1300 - Estrogen Receptor Transcriptional Activation (Human Cell Line [HeLa-9903]). EPA 740-C-09-006. Washington, DC:U.S. Environmental Protection Agency. Available: http://www.epa.gov/ocspp/pubs/frs/publications/Test_Guidelines/series890.htm.
- FDA. 2009. Good Laboratory Practice for Nonclinical Laboratory Studies. 21 CFR 58. Available: http://www.access.gpo.gov/nara/cfr/waisidx_08/21cfr58_08.html [accessed 1 November 2010].
- 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. Research Triangle Park, NC:National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/docs/about_docs/validate.pdf [accessed 23 December 2010].
- ICCVAM. 2002a. Background Review Document. Current Status of Test Methods for Detecting Endocrine Disruptors: In Vitro Androgen Receptor Binding Assays. NIH Publication No. 03-4506. Research Triangle Park, NC:National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/docs/endo_docs/final1002/arbndbrd/ARBd034506.pdf.
- ICCVAM. 2002b. Background Review Document. Current Status of Test Methods for Detecting Endocrine Disruptors: In Vitro Androgen Receptor Transcriptional Activation Assays. NIH Publication No. 03-4507. Research Triangle Park, NC:National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/docs/endo_docs/final1002/arta_brd/ARTA034507.pdf.
- ICCVAM. 2002c. Background Review Document. Current Status of Test Methods for Detecting Endocrine Disruptors: In Vitro Estrogen Receptor Binding Assays. NIH Publication No. 03-4504. Research Triangle Park, NC:National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/docs/endo_docs/final1002/erbndbrd/ERBd034504.pdf.

ICCVAM. 2002d. Background Review Document. Current Status of Test Methods for Detecting Endocrine Disruptors: In Vitro Estrogen Receptor Transcriptional Activation Assays. NIH Publication No. 03-4505. Research Triangle Park, NC:National Institute of Environmental Health Sciences. Available:

http://iccvam.niehs.nih.gov/docs/endo_docs/final1002/erta_brd/ERTA034505.pdf.

ICCVAM. 2002e. Expert Panel Evaluation of the Validation Status of In Vitro Test Methods for Detecting Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays - Expert Panel Final Report. Research Triangle Park, NC:National Institute of Environmental Health Sciences. Available:

http://iccvam.niehs.nih.gov/docs/endo_docs/expertpanfinalrpt/panelrpt1102.pdf.

ICCVAM. 2003a. ICCVAM Evaluation of In Vitro Test Methods for Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays. NIH Publication No. 03-4503. Research Triangle Park, NC:National Institute of Environmental Health Sciences. Available:

http://iccvam.niehs.nih.gov/docs/endo_docs/edfinalrpt0503/edfinrpt.pdf.

ICCVAM. 2003b. 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. Available:

http://iccvam.niehs.nih.gov/SuppDocs/SubGuidelines/SD_subg034508.pdf [accessed 4 November 2010].

ICCVAM. 2006. Finalized Addendum to ICCVAM Evaluation of In Vitro Test Methods for Detecting Potential Endocrine Disruptors. Research Triangle Park, NC:National Institute of Environmental Health Sciences. Available:

http://iccvam.niehs.nih.gov/docs/endo_docs/EDAddendFinal.pdf.

Klimisch HJ, Andreae M, Tillmann U. 1997. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. *Regul Toxicol Pharmacol* 25(1): 1-5.

Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0031080387&partnerID=40&md5=208e321978d83c1922f194ed33f0bda3>.

NRC. 1999. *Hormonally Active Agents in the Environment*. Washington, DC:National Academy Press. Available: <http://www.nap.edu/openbook.php?isbn=0309064198>.

OECD. 1996. *Final Report of the OECD Workshop on Harmonization of Validation and Acceptance Criteria for Alternative Toxicological Test Methods*. Paris:OECD Publishing.

OECD. 1998. *OECD Series on Good Laboratory Practice and Compliance Monitoring No. 1: OECD Principles on Good Laboratory Practice [as revised in 1997]*. Paris:OECD Publishing. Available: http://www.oecd.org/document/63/0,3343,en_2649_34381_2346175_1_1_1_1,00.html [accessed 30 July 2010].

OECD. 2005. *Guidance Document No. 34: Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment*. Paris:OECD Publishing.

OECD. 2009. Test No. 455. *Stably Transfected Human Estrogen Receptor- α Transcriptional Activation Assay for Detection of Estrogenic Agonist-Activity of Chemicals [adopted 7 September 2009]*. In: *OECD Guidelines for the Testing of Chemicals, Section 4: Health Effects*. Paris:OECD Publishing. Available: <http://dx.doi.org/10.1787/9789264076372-en>.

Pujol P, Rey JM, Nirde P, Roger P, Gastaldi M, Laffargue F, et al. 1998. Differential expression of estrogen receptor- α and - β messenger RNAs as a potential marker of ovarian

carcinogenesis. *Cancer Res* 58(23): 5367-5373. Available:
<http://cancerres.aacrjournals.org/content/58/23/5367.full.pdf+html>.

Rogers JM, Denison MS. 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 Mol Toxicol* 13(1): 67-82. Available:
<http://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-0033936762&partnerID=40&rel=R6.0.0>.

This page intentionally left blank

Appendix A

ICCVAM Evaluation of the BG1Luc ER TA (LUMI-CELL[®]) Test Method – Timeline

This page intentionally left blank

ICCVAM Evaluation Timeline

January 2004	Xenobiotic Detection Systems, Inc. (XDS, Durham, NC) nominates their LUMI-CELL® BG1Luc ER TA Test Method for an interlaboratory validation study
April 21, 2004	<i>Federal Register</i> Notice (69 FR 21564) – <i>In Vitro</i> Endocrine Disruptor Test Methods: Request for Comments and Nominations
March 16, 2006	<i>Federal Register</i> Notice (71 FR 13597) - Notice of Availability of a Revised List of Recommended Reference Substances for Validation of <i>In Vitro</i> Estrogen and Androgen Receptor Binding and Transcriptional Activation Assays: Request for Comments and Submission of <i>In Vivo</i> and <i>In Vitro</i> Data
November 27, 2009	<i>Federal Register</i> Notice (74 FR 62317) – Evaluation of <i>In Vitro</i> Estrogen Receptor Transcriptional Activation and <i>In Vitro</i> Cell Proliferation Assays for Endocrine Disruptor Chemical Screening: Request for Nominations for an Independent Expert Peer Review Panel and Submission of Relevant <i>In Vitro</i> and <i>In Vivo</i> Data
January 24, 2011	<i>Federal Register</i> Notice (76 FR 4113) – Announcement of an Independent Scientific Peer Review Panel Meeting on an <i>In Vitro</i> Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening; Availability of Draft Background Review Document (BRD); Request for Comments
March 29-30, 2011	Independent Scientific Peer Review Panel holds a public meeting, with opportunity for public comments, at Natcher Conference Center in Bethesda, MD. The Panel was charged with reviewing the current validation status of the BG1Luc ER TA test method and commenting on the extent to which the information in the draft BRD supported the draft ICCVAM test method recommendations.
April 26, 2011	<i>Federal Register</i> Notice (76 FR 23323) – Meeting of the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM)
May 18, 2011	<i>Federal Register</i> Notice (76 FR 28781) – Independent Scientific Peer Review Panel Report: Evaluation of the Validation Status of an <i>In Vitro</i> Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening: Notice of Availability and Request for Public Comments
June 16-17, 2011	SACATM public meeting, SACATM and public comments on the draft Panel conclusions and recommendations
August 24, 2011	ICCVAM approves forwarding the test method evaluation report to Federal agencies for consideration.
Fall 2011	Transmittal of ICCVAM recommendations to Federal agencies

This page intentionally left blank

Appendix B

Updated ICCVAM-Recommended Protocols:

BG1Luc ER TA

B1	BG1Luc ER TA – Agonist Protocol	B-3
B2	BG1Luc ER TA – Antagonist Protocol.....	B-37

This page intentionally left blank

Appendix B1

BG1Luc ER TA – Agonist Protocol

This page intentionally left blank

**BG1Luc ER TA TEST METHOD
AGONIST PROTOCOL**

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

List of Acronyms and Abbreviations

13 mm test tube	13 x 100 mm glass test tubes
DMEM	Dulbecco's Modification of Eagle's Medium
DMSO	Dimethyl Sulfoxide
DMSO control	1% v/v dilution of DMSO in tissue culture media used as a vehicle control
E2	17 β -estradiol
E2 reference standard	11 Point Serial Dilution of 17 β -estradiol reference standard for the BG1Luc ER TA agonist assay
EC ₅₀ value	Concentration that produces a half-maximal response as calculated using the four parameter Hill function.
ER	Estrogen Receptor
Estrogen-free DMEM	DMEM (phenol red free) supplemented with 1% Penicillin/Streptomycin, 2% L-Glutamine, and 5% Charcoal-dextran treated FBS
FBS	Fetal Bovine Serum
G418	Gentamycin
Methoxychlor	<i>p,p'</i> -Methoxychlor
Methoxychlor control	3.13 μ g/mL Methoxychlor Weak Positive Control for the BG1Luc ER TA Agonist Assay
RPMI	RPMI 1640 growth medium
TA	Transcriptional Activation
T25	25 cm ² tissue culture flask
T75	75 cm ² tissue culture flask
T150	150 cm ² tissue culture flask

1.0 Purpose

This protocol is designed to evaluate coded test substances for potential estrogen receptor (ER) agonist activity using the BG1Luc ER TA test method.

2.0 Sponsor

(As Appropriate)

3.0 Definitions

Dosing Solution: The test substance, control substance, or reference standard solution, which is to be placed into the tissue culture wells for experimentation.

Raw Data: Raw data includes information that has been collected but not formatted or analyzed, and consists of the following:

- Data recorded in the Study Notebook
- Computer printout of initial luminometer data
- Other data collected as part of GLP compliance, e.g.:
 - Equipment logs and calibration records
 - Test substance and tissue culture media preparation logs
 - Cryogenic freezer inventory logs

Soluble: Test substance exists in a clear solution without visible cloudiness or precipitate.

Study Notebook: The study notebook contains recordings of all activities related to the conduct of the BG1Luc ER TA agonist assay.

Test Substances: Substances supplied to the testing laboratories that are coded and distributed such that only the Project Officer, Study Management Team (SMT), and the Substance Inventory and Distribution Management have knowledge of their true identity. The test substances will be purchased, aliquoted, coded, and distributed by the Supplier under the guidance of the Project Officer and the SMT.

4.0 Testing Facility and Key Personnel

4.1 Testing Facility

(As Appropriate)

4.2 Key Personnel

Study Director: (As Appropriate)

Quality Assurance Director: (As Appropriate)

5.0 Identification of Reference Standard and control substances

5.1 Controls

Controls for the ER agonist protocol are as follows:

Vehicle control (dimethyl sulfoxide [DMSO]): 1% (v/v) DMSO (CASRN 67-68-5) diluted in tissue culture media

Positive control (p,p'-Methoxychlor [methoxychlor]): Methoxychlor (CASRN 72-43-5), 3.13 µg/mL in tissue culture media, used as a weak positive control

5.2 Reference Standard

Reference standard (17β-estradiol [E2]): Three concentrations of E2 (CASRN 50-28-2) in duplicate for range finder testing and a serial dilution consisting of 11 concentrations of E2 in duplicate for comprehensive testing

6.0 Overview of General Procedures for Agonist Testing

All experimental procedures are to be carried out under aseptic conditions and all solutions, glassware, plastic ware, pipettes, etc., shall be sterile. All methods and procedures shall be documented in the study notebook.

Agonist range finder testing is conducted on 96-well plates using four concentrations of E2 (5.00×10^{-5} , 1.25×10^{-5} , 3.13×10^{-6} , and 7.83×10^{-7} µg/mL) in duplicate as the reference standard and four replicate wells for the DMSO control. Range finder testing uses all wells of the 96-well plate to test six substances as seven point 1:10 serial dilutions in duplicate.

Comprehensive testing is conducted on 96-well plates using 11 concentrations of E2 in duplicate as the reference standard (**Table 6-1**). Four replicate wells for the DMSO control and four replicate wells for the methoxychlor control are included on each plate. Comprehensive testing uses all wells of the 96-well plate to test 2 substances as 11 point serial dilutions in triplicate.

Table 6-1 Concentrations of E2 Reference Standard Used in Comprehensive Testing

E2 Concentrations ¹		
1.00×10^{-4}	6.25×10^{-6}	3.92×10^{-7}
5.00×10^{-5}	3.13×10^{-6}	1.95×10^{-7}
2.50×10^{-5}	1.56×10^{-6}	9.78×10^{-8}
1.25×10^{-5}	7.83×10^{-7}	

¹ Concentrations are presented in µg/mL.

Visual observations for cell viability are conducted for all experimental plates just prior to luminescence measurements, as outlined in **Section 11.2**.

Luminescence data, measured in relative light units (RLUs), is corrected for background luminescence by subtracting the mean RLU value of the vehicle control (DMSO) wells from the RLU measurements for each of the other wells of the 96-well plate. Data is then transferred into Excel® data management spreadsheets and GraphPad Prism® statistical software, graphed, and evaluated as follows:

- A response is considered positive for agonist activity when the average adjusted RLU for a given concentration is greater than the mean RLU value plus three times the standard deviation for the vehicle control.
- Any response below this threshold is considered negative for agonist activity.

For substances that are positive at one or more concentrations, the concentration that causes a half-maximal response (EC₅₀) is calculated using a Hill function analysis. The Hill function is a four-parameter logistic mathematical model relating the substance concentration to the response (typically following a sigmoidal curve) using the equation below:

$$Y = \text{Bottom} + \frac{\text{Top} - \text{Bottom}}{1 + 10^{(\log \text{EC}_{50} - X) \text{HillSlope}}}$$

where Y = response (i.e., relative light units); X = the logarithm of concentration; Bottom = the minimum response; Top = the maximum response; $\log EC_{50}$ = the logarithm of X as the response midway between Top and Bottom; and HillSlope describes the steepness of the curve. The model calculates the best fit for the Top, Bottom, HillSlope, and EC_{50} parameters. See **Section 11.6.5** for more details.

Acceptance or rejection of a test is based on evaluation of reference standard and control results from each experiment conducted on a 96-well plate. Results for these controls are compared to historical results compiled in the historical database, as seen in **Section 14.0**.

6.1 Range Finder Testing

Agonist range finding for coded substances consists of a seven point, 1:10 serial dilution using duplicate wells per concentration. Concentrations for comprehensive testing are selected based on the response observed in range finder testing. If necessary, a second range finder test can be conducted to clarify the optimal concentration range to test (see **Section 12.0**).

6.2 Comprehensive Testing

Comprehensive agonist testing for coded substances consists of 11 point, serial dilutions, with each concentration tested in triplicate wells of the 96-well plate. Three separate experiments are conducted for comprehensive testing on three separate days, except during Phases III and IV of the validation effort, in which comprehensive testing experiments are conducted once (see **Section 13.0**).

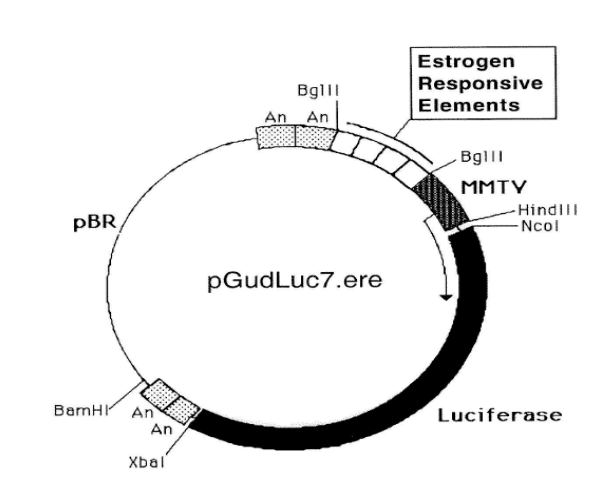
7.0 Materials for BG1Luc ER TA Agonist Testing

This section provides the materials needed to conduct BG1Luc ER TA testing, with associated brand names/vendors¹ in brackets.

7.1 BG1Luc4E2 Cells

Human ovarian cancer cell line stably transfected with a plasmid containing an estrogen response element pGudLuc7.0 (**Figure 7-1**). The BG1Luc4E2 cell line is available upon request from the University of California, Davis, California, USA, and is also available from Xenobiotic Detection Systems Inc., Durham, North Carolina, USA.

Figure 7-1 pGudLuc7.ERE Plasmid



¹ Brand names and vendors should not be considered an endorsement by the U.S. Government or any member of the U.S. Government; such information is provided as an example.

7.2 Technical Equipment

All technical equipment may be obtained from Fisher Scientific International, Inc. (Liberty Lane Hampton, NH, USA 03842). Equivalent technical equipment from another commercial source can be used.

- Analytical balance (Cat. No. 01-910-320)
- Berthold Orion 1 Microplate Luminometer [Berthold CatNo.: Orion 1 MPL3] or equivalent and dedicated computer
- Biological safety hood, class II, and stand (Cat. No. 16-108-99)
- Centrifuge (low speed, tabletop with swinging bucket rotor) (Cat. No. 04-978-50 centrifuge, and 05-103B rotor)
- Combustion test kit (CO₂ monitoring) (Cat. No. 10-884-1)
- Drummond diaphragm pipetter (Cat. No. 13-681-15)
- Freezers, –20°C (Cat. No. 13-986-150), and –70°C (Cat. No. 13-990-86)
- Hand tally counter (Cat. No. 07905-6)
- Hemocytometer, cell counter (Cat. No. 02-671-5)
- Light microscope, inverted (Cat. No. 12-561-INV)
- Light microscope, upright (Cat. No. 12-561-3M)
- Liquid nitrogen flask (Cat. No. 11-675-92)
- Micropipetter, repeating (Cat. No. 21-380-9)
- Pipetters, air displacement, single channel (0.5 –10 µL (Cat. No. 21-377-191), 2 –20 µL (Cat. No. 21-377-287), 20 – 200 µl (Cat. No. 21-377-298), 200 - 1000 µL (Cat. No. 21-377-195))
- Refrigerator/freezer (Cat. No. 13-986-106A)
- Shaker for 96-well plates (Cat. No. 14-271-9)
- Sodium hydroxide (Cat. No. 5318-500)
- Sonicating water bath (Cat. No. 15-335-30)
- Tissue culture incubator with CO₂ and temperature control (Cat. No. 11-689-4)
- Vacuum pump with liquid trap (side arm Erlenmeyer) (Cat. No. 01-092-29)
- Vortex mixer (Cat. No. 12-814)

Equipment should be maintained and calibrated as per GLP guidelines and individual laboratory SOPs.

7.3 Reference Standard, Controls, and Tissue Culture Supplies

All tissue culture reagents must be labeled to indicate source, identity, storage conditions and expiration dates. Tissue culture solutions must be labeled to indicate concentration, stability (where known), and preparation and expiration dates.

Equivalent tissue culture media and sera from another commercial source can be used, but must first be tested as described in **Section 15.0** to determine suitability for use in this test method.

The following are the necessary tissue culture reagents and possible commercial sources (in brackets) based on their use in the pre-validation studies:

- BackSeal-96/384, white adhesive bottom seal for 96-well and 384-well microplate [Perkin-Elmer, Cat. No. 6005199]
- 17 β-estradiol (CAS RN: 50-28-2) [Sigma-Aldrich, Cat. No. E8875]
- Cryovial, 2 mL (Corning Costar) [Fisher Scientific Cat. No. 03-374-21]
- Culture tube 13 x 100mm (case) [Thomas Scientific Cat. No.: 10009186R38]²
- Culture tube, 50 mL conical (Corning Costar) [Fisher Scientific Cat. No. 05-526C]

² If glass tubes cannot be obtained from Thomas Scientific, the preference is for flint glass, then lime glass, then borosilicate glass.

- DMSO, U.S.P. analytical grade. [Sigma-Aldrich, Cat. No. 34869-100ML]
- Dulbecco's Modification of Eagle's Medium (DMEM), containing 4.5 g/L glucose, with sodium pyruvate, without phenol red or L-glutamine [Mediatech/Cellgro, Cat. No. 17-205-CV]
- Fetal Bovine Serum [Mediatech/Cellgro Cat. No. MT 35-010-CV]
- Fetal Bovine Serum, charcoal/dextran treated, triple 0.1 µm sterile filtered [Hyclone, Cat. No. SH30068.03]
- Gentamycin Sulfate (G418), 50 mg/mL [Mediatech/Cellgro Cat. No. 30-234-CR]
- L-glutamine, 29.2 mg/mL [Cellgro, Cat. No. 25005-CI]
- Luciferase Assay System (10-Pack) [Promega Cat. No. E1501]
- Lysis Solution 5X [Promega, Cat. No. E1531]
- Methoxychlor (CAS RN: 72-43-5) [Sigma-Aldrich, Cat. No. 49054]
- Penicillin/streptomycin solution, 5000 I.U. penicillin, 5000 µg/mL streptomycin [Cellgro, Cat. No. 30-001-CI].
- Phosphate buffered saline (PBS, 1X) without calcium and magnesium [Cellgro, Cat. No. 21-040-CV]
- Pipettes, serological: 2.0 mL [Sigma-Aldrich, Cat. No. P1736], 5.0 mL [Sigma-Aldrich, Cat. No. P1986], 25 mL [Sigma-Aldrich, Cat. No. P2486]
- RPMI 1640 medium, containing L-glutamine [Mediatech, Cat. No. 10-040-CV]
- Tissue culture flasks (Corning-Costar): 25 cm² (T25) [Fisher Cat. No. 10-126-28]; 75 cm² (T75) [Fisher Cat. No. 10-126-37]; and 150 cm² (T150) [Fisher Cat. No. 10-126-34]
- Tissue culture plates (Corning-Costar): 96-well [Thomas Scientific Cat. No. 6916A05]
- Trypsin (10X), 2.5% in Hank's balanced salt solution (HBSS), without calcium and magnesium, without phenol red [Cellgro, Cat. No. 25-054-CI].

All reagent lot numbers and expiration dates must be recorded in the study notebook.

8.0 Preparation of Tissue Culture Media and Solutions

All tissue culture media and media supplements must be quality tested before use in experiment (see **Section 15.0**).

8.1 RPMI 1640 Growth Medium (RPMI)

RPMI 1640 is supplemented with 0.9% Pen-Strep and 8.0% FBS to make RPMI growth medium (RPMI).

Procedure for one 549 mL bottle:

1. Remove FBS from -70°C freezer, and Pen-Strep from -20°C freezer and allow to equilibrate to room temperature.
2. Add 44 mL of FBS and 5 mL Pen-Strep to the bottle of RPMI 1640.
3. Label RPMI bottle as indicated in **Section 7.3**

Store at 2-8°C for no longer than six months or until the shortest expiration date of any media component.

8.2 Estrogen-Free DMEM Medium

DMEM is supplemented to contain 4.5% charcoal/dextran treated FBS, 1.9% L-glutamine, 0.9% Pen-Strep.

Procedure for one 539 mL bottle:

1. Remove charcoal/dextran treated FBS from -70°C freezer, and L-glutamine and Pen-Strep from -20°C freezer and allow to equilibrate to room temperature.
2. Add 24 mL of charcoal/dextran treated FBS, 10 mL L-glutamine, and 5 mL Pen-Strep to one 500 mL bottle of DMEM.

3. Label estrogen-free DMEM bottle as indicated in Section 7.3

Store at 2-8 °C for no longer than six months or until the shortest expiration date of any media component..

8.3 1X Trypsin Solution

1X Trypsin solution is prepared by dilution from a 10X premixed stock solution. The 10X stock solution should be stored in 10 mL aliquots in a -20°C freezer.

Procedure for making 100 mL of 1X trypsin:

1. Remove a 10 mL aliquot of 10X trypsin from -20°C freezer and allow to equilibrate to room temperature.
2. Aliquot 1 mL Trypsin (10X) along with 9 mL of 1X PBS into ten 15 mL sterile centrifuge tubes.
3. Label 1X trypsin aliquots as indicated in Section 7.3

1X Trypsin should be stored at -20°C.

8.4 1X Lysis Solution

Lysis solution is prepared by dilution from a 5X premixed stock solution. Both the 5X and 1X solutions can be repeatedly freeze-thawed.

The procedure for making 10 mL of 1X lysis solution:

1. Thaw the 5X Promega Lysis solution and allow it to reach room temperature.
2. Remove 2 mL of 5X solution and place it in a 15 mL conical centrifuge tube.
3. Add 8 mL of distilled, de-ionized water to the conical tube.
4. Cap and shake gently until solutions are mixed.

Store at -20 °C for no longer than 1 year from receipt.

8.5 Reconstituted Luciferase Reagent

Luciferase reagent consists of two components, luciferase buffer and lyophilized luciferase substrate.

For long term storage, unopened containers of the luciferase buffer and lyophilized luciferase substrate can be stored at -70°C for up to one year.

To reconstitute luciferase reagent:

1. Remove luciferase buffer and luciferase substrate from -70°C freezer, and allow them to equilibrate to room temperature.
2. Add 10 mL of luciferase buffer solution to luciferase substrate container and swirl or vortex gently to mix; the Luciferase substrate should readily go into solution.
3. After solutions are mixed, aliquot to a 15mL centrifuge tube.
4. Store complete solution at -20°C.

Reconstituted luciferase reagent is stable for up to 1 month at - 20°C.

9.0 Overview of Propagation and Experimental Plating of BG1Luc4E2 cells

BG1Luc4E2 cells are based on a continuous ovarian carcinoma cell line (BG-1 cells) that endogenously express ER α and ER β and have been stably transfected with an ER responsive reporter gene (*luc*).

Although the cell line has proved to be stable over long-term passage *in vitro*, careful handling and the use of quality cell culture materials is required to maintain the stability and integrity of the cell line.

Procedures specified in the Guidance on Good Cell Culture Practice (Coecke 2005) should be followed to

assure the quality of all materials and methods in order to maintain the integrity, validity, and reproducibility of any work conducted.

The BG1Luc4E2 cells are stored in liquid nitrogen in 2 mL cryovials. BG-1 cells are grown as a monolayer in tissue culture flasks in a dedicated tissue culture incubator at $37^{\circ}\text{C} \pm 1^{\circ}\text{C}$, $90\% \pm 5\%$ humidity, and $5.0\% \pm 1\%$ CO_2/air . The cells should be examined, on a daily basis during working days, under an inverted phase contrast microscope and any changes in morphology and/or adhesive properties must be noted in the study notebook.

Two T150 flasks containing cells at 80 to 90% confluence will usually yield a sufficient number of cells to fill three 96-well plates for use in experiments.

9.1 Procedures for Thawing Cells and Establishing Tissue Cultures

Warm all of the tissue culture media and solutions to room temperature by placing them under the tissue culture hood several hours before use.

All tissue culture media, media supplements, and tissue culture plasticware must be quality tested before use in experiments (**Section 15.0**).

9.1.1 Thawing Cells

1. Remove a cryovial of frozen BG-1 cells from the liquid nitrogen flask.
2. Facilitate rapid thawing by loosening the top slightly (do not remove top) to release trapped gasses and retightening it. Roll vial between palms.
3. Use a micropipette to transfer cells to a 50 mL conical centrifuge tube.
4. Rinse cryovial twice with 1X PBS and add PBS rinse material to the conical tube.
5. Add 20 mL of RPMI to the conical tube.
6. Centrifuge at $1000 \times g$ for eight min. If a pellet of cells has not formed, centrifuge for an additional 5 minutes.
7. Aspirate media from pellet and re-suspend it in 5 mL RPMI, drawing the pellet repeatedly through a 1.0 mL serological pipette to break up any clumps of cells.
8. Transfer cells to a T25 flask, place them in an incubator (see conditions in Section 9.0) and grow to 80% to 90% confluence (approximately 48 to 72 hrs).

9.1.2 Establishing Tissue Cultures

Once cells have reached 80% to 90% confluence, transfer the cells to a T75 flask by performing, for example, the following steps:

1. Remove the T25 flask from the incubator.
2. Aspirate the RPMI, then add 5 mL 1X PBS, making sure that the cells are coated with PBS.
3. Aspirate 1X PBS, then add 1 to 2 mL 1X trypsin to the T25 flask, gently swirling the flask to coat all cells with the trypsin.
4. Place the flask in an incubator (see conditions in **Section 9.0**) for 5 to 10 min.
5. Detach cells by hitting the side of the flask sharply against the palm or heel of the hand.
6. Confirm cell detachment by examination under an inverted microscope. If cells have not detached, return the flask to the incubator for an additional 2 minutes, then hit the flask again.
7. After cells have detached, add 5 mL PBS, and transfer the suspended cells to a 50 mL centrifuge tube. Wash the flask one additional time with 5 mL PBS.
8. Immediately add 20 mL RPMI to the conical tube to inhibit further cellular digestion by residual trypsin.
9. Pellet the cells by centrifugation, as described in **Section 9.1.1**, and re-suspend the cells in 10 mL RPMI medium.
10. Draw the pellet repeatedly through a 25 mL serological pipette to break up clumps of cells.

11. Transfer cells to a T75 flask, then place the flask in an incubator (see conditions in Section 9.0) and grow to 80% to 90% confluence (approximately 48 to 72 hrs).

When cells have reached 80% to 90% confluency, transfer them into a T150 flask by performing, for example, the following steps:

12. Remove the T75 flask from the incubator, aspirate the old media and add 5 mL 1X PBS.
13. Aspirate 1X PBS, add 2 mL of 1X trypsin to the flask, and place it in an incubator (see conditions in **Section 9.0**) for 5 to 10 min.
14. Repeat steps **5** through **11** in **Section 9.1.2**, re-suspending the pellet in 20 mL of RPMI.
15. Transfer cells to a T150 flask and place it in the incubator (see conditions in **Section 9.0**) and grow to 80% to 90% confluence (approximately 48 to 72 hrs).
16. Remove the T150 flask from the incubator.
17. Aspirate the RPMI and add 5 mL 1X PBS.
18. Aspirate 1X PBS and add 3 mL 1X trypsin to the T150 flask, making sure that the cells are coated with the trypsin.
19. Incubate cells in an incubator (see conditions in **Section 9.0**) for 5 to 10 min.
20. Detach cells by hitting the side of the flask sharply against the palm or heel of the hand.
21. Confirm cell detachment by examination under an inverted microscope. If cells have not detached, return the flask to the incubator for an additional 2 minutes, then hit the flask again.
22. After cells have detached, add 5 mL 1X PBS and transfer the suspended cells from the T150 flask to a 50 mL conical tube. Add an additional 5 mL PBS to the flask, swirl around the flask, and then transfer the PBS to the 50 mL conical tube.
23. Immediately add 20 mL RPMI to the conical tube to inhibit further cellular digestion by residual trypsin.
24. Centrifuge at 1000 x g for eight min. If a pellet of cells has not formed, centrifuge for an additional 5 minutes.
25. Aspirate the media from the pellet and re-suspend it in 40 mL RPMI, drawing the pellet repeatedly through a 25 mL serological pipette to break up any clumps of cells.
26. Transfer 20 mL of cell suspension to each of two T150 flasks, place them in an incubator (see conditions in Section 9.0) and grow to 80% to 90% confluence (approximately 48 to 72 hrs).

9.2 Ongoing Tissue Culture Maintenance, Conditioning in Estrogen-free Medium, and Plating Cells for Experimentation

The following procedure is used to condition the BG1Luc4E2 cells to an estrogen-free environment prior to plating the cells in 96-well plates for analysis of estrogen dependent induction of luciferase activity.

To start the tissue culture maintenance and estrogen-free conditioning, split the two T150 culture flasks into four T150 flasks. Two of these flasks will be used for continuing tissue culture and will use the RPMI media mentioned above. The other two flasks will be cultured in estrogen-free DMEM for experimental use. Extra care must be taken to avoid contaminating the estrogen-free cells with RPMI.

1. Remove both T150 flasks from the incubator.
2. Aspirate the medium and rinse the cells with 5 mL 1X PBS.
3. Aspirate 1X PBS, then add 3 mL 1X trypsin to the flasks, gently swirling the flask to coat all cells with the trypsin.
4. Incubate cells in the incubator (see conditions in **Section 9.0**) for 5 to 10 min.
5. Detach cells by hitting the side of the flask sharply against the palm or heel of the hand.
6. Confirm cell detachment by examination under an inverted microscope. If cells have not detached, return the flask to the incubator for an additional 2 minutes, then hit the flask again.
7. After cells have detached, add 5 mL 1X PBS to the first T150 flask and transfer the suspended cells to the second T150 flask.

8. Transfer the contents of both flasks to a 50 mL conical tube. Repeat step 7 with an additional 5 mL 1X PBS and transfer to the 50 mL conical tube.
9. Immediately add 20 mL estrogen-free DMEM to the 50 mL conical tube to inhibit further cellular digestion by residual trypsin.
10. Centrifuge at 1000 x g for eight min. If a pellet of cells has not formed, centrifuge for an additional 5 minutes.
11. Aspirate media from pellet and re-suspend it in 4 mL estrogen-free DMEM, drawing the pellet repeatedly through a 1 mL serological pipette to break up clumps of cells.

At this point, cells are ready to be divided into the ongoing tissue culture and estrogen-free conditioning groups.

9.2.1 Ongoing Tissue Culture Maintenance

1. Add 20 mL RPMI to two T150 flasks.
2. Add 220 μ L G418 to the RPMI in the T150 flasks
3. Add 1 mL of cell suspension from **9.2 step 11** to each flask.
4. Place T150 flasks in tissue culture incubator (see conditions in **Section 9.0**) and grow to 80% to 90% confluence (approximately 48 to 72 hrs).
5. Tissue culture medium may need to be changed 24 hours after addition of G418 to remove cells that have died because they do not express reporter plasmid.
6. G418 does not need to be added to the flasks a second time.
7. Repeat Section 9.2 steps 1-11 for ongoing tissue culture maintenance.

9.2.2 Conditioning in Estrogen-free Medium

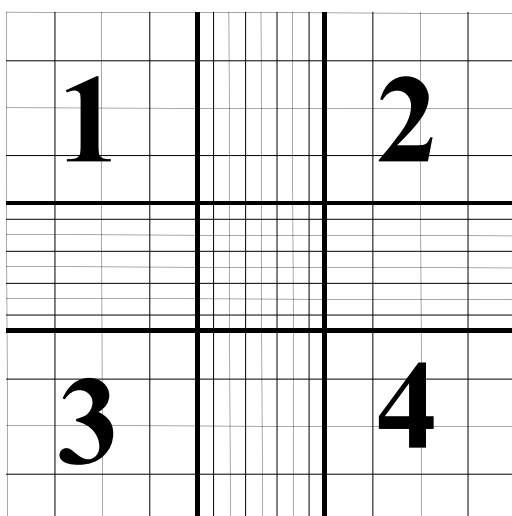
1. Add 20 mL estrogen-free DMEM to two T150 flasks.
2. Add 150 μ L G418 to the estrogen-free DMEM in the T150 flasks.
3. Add 1 mL of cell suspension from **Section 9.2 step 11** to each flask.
4. Tissue culture medium may need to be changed 24 hours after addition of G418 to remove cells that have died because they do not express reporter plasmid.
5. G418 does not need to be added to the flasks a second time.
6. Place the T150 flasks in the incubator (see conditions in Section 9.0) and grow to 80% to 90% confluence (approximately 48 to 72 hrs).

9.2.3 Plating Cells Grown in Estrogen-free DMEM for Experimentation

1. Remove the T150 flasks that have been conditioned in estrogen-free DMEM for 48 to 72 hours from the incubator.
2. Aspirate the medium, then rinse the cells with 5 mL 1X PBS.
3. Aspirate 1X PBS, then add 3 mL 1X trypsin to the flasks, gently swirling the flask to coat all cells with the trypsin.
4. Place the flasks in an incubator (see conditions in **Section 9.0**) for 5 to 10 min.
5. Detach cells by hitting the side of the flask sharply against the palm or the heel of the hand.
6. Confirm cell detachment by examination under an inverted microscope. If cells have not detached, return the flask to the incubator for 2 additional minutes, then hit the flask again.
7. After cells have detached, add 5 mL 1X PBS and transfer the suspended cells from the T150 flask to a 50 mL conical tube. Add an additional 5 mL PBS to the flask, gently swirl around the flask, and then transfer to the 50 mL conical tube.
8. Immediately add 20 mL estrogen-free DMEM to each conical tube to inhibit further cellular digestion by residual trypsin.
9. Centrifuge at 1000 x g for eight min. If a pellet of cells has not formed, centrifuge for an additional 5 minutes.

10. Aspirate the media from the pellet and re-suspend it in 20 mL DMEM, drawing the pellet repeatedly through a 25 mL serological pipette to break up any clumps of cells.
11. Pipette 15 μ L of the cell suspension into the “v” shaped slot on the hemocytometer. Ensure that the solution covers the entire surface area of the hemocytometer grid, and allow cells to settle before counting.
12. Using 100x magnification, view the counting grid.
13. The counting grid on the hemocytometer consists of nine sections, four of which are counted (upper left, upper right, lower left, and lower right, see Figure 9-1). Each section counted consists of four by four grids. Starting at the top left and moving clockwise, count all cells in each of the four by four grids. Some cells will be touching the outside borders of the square, but only count those that touch the top and right borders of the square. This value is then used in the calculation below to get to the desired concentration of 200,000 cells/mL.

Figure 9-1 Hemocytometer Counting Grid



The volume of each square is 10^{-4} mL, therefore:

Cells/mL=(average number per grid) x 10^{-4} mL x 1/(starting dilution).

Starting dilution: 20 mL (for T150 flasks)

Harvested cells for a T150 flask are suspended in 20 mL of estrogen-free DMEM and sampled for determination of concentration of cells/mL.

Example Calculation:

- Grids 1, 2, 3, and 4 are counted and provide the following data:
 - 50, 51, 49, and 50: average number of cells per grid is equal to 50.

Cells/mL = 50 cells per grid $\div 10^{-4}$ mL volume of grid = 50×10^4 cells/mL (or 500,000 cells/mL)

Total # of Cells Harvested = 500,000 cells/mL x 20 mL

Desired Concentration (or Concentration_{Final})= 200,000 cells/mL

Formula: (Concentration_{Final} x Volume_{Final} = Concentration_{Initial} x Volume_{Initial})

Concentration_{Final} = 200,000 cells/mL

Concentration_{Initial} = 500,000 cells/mL

Volume_{Initial} = 20 mL

Volume_{Final} – to be solved for.

Therefore: 200,000 cells/mL x Volume_{Final} = 500,000 cells/mL x 20 mL

Solving for Volume_{Final} we find = 50 mL

Therefore, add 30 mL of estrogen-free DMEM to the cell suspension for a total volume of 50 mL, which will yield the desired concentration of 200,000 cells/mL for plating.

14. This dilution scheme will give a concentration of 200,000 cells/mL. 200 µL of this cell suspension is used for each well of a 96-well plate (i.e., 40,000 cells per well).
15. Remove a 96-well plate from its sterile packaging. Use a repeater pipetter to pipette 200 µL of cell suspension into each well for to be used for the testing of coded substances, reference standard and controls (**note**: add 200 µL of estrogen-free DMEM only to any wells not being used for testing).
16. Incubate plate(s) in an incubator (see conditions in Section 9.0) for a minimum of 24 hours, but no longer than 48 hours before dosing.

Two T150 flasks containing cells at 80% to 90% confluence will typically yield sufficient cells to fill four 96-well plates.

10.0 Preparation of Test Substances

The solvent used for dissolution of test substances is 100% DMSO. All test substances should be allowed to equilibrate to room temperature before being dissolved and diluted. Test substance solutions (except for reference standards and controls) should not be prepared in bulk for use in subsequent tests. Test substances are to be used within 24 hours of preparation. Solutions should not have noticeable precipitate or cloudiness.

All information on weighing, solubility testing, and calculation of final concentrations for test substances, reference standards and controls is to be recorded in the study notebook.

10.1 Determination of Test Substance Solubility

1. Prepare a 100 mg/mL solution of the test substance in 100% DMSO in a 4 mL conical tube.
2. Vortex to mix.
3. If the test substance does not dissolve at 100 mg/mL, prepare a 10 mg/mL solution and vortex as above.
4. If the test substance does not dissolve at 10 mg/mL solution, prepare a 1 mg/mL solution in a 4 mL conical tube and vortex as above.
5. If the test substance does not dissolve at 1 mg/mL, prepare a 0.1 mg/mL solution in a 4 mL conical tube and vortex as above.
6. Continue testing, using 1/10 less substance in each subsequent attempt until test substance is solubilized in DMSO.

Once the test substance has fully dissolved in 100% DMSO, the test substance is ready to be used for BGI_{Luc} ER TA testing.

10.2 Preparation of Reference Standards, Control and Test Substances

All “dosing solutions” of test substance concentrations are to be expressed as µg/mL in the study notebook and in all laboratory reports.

All information on preparation of test substances, reference standards and controls is to be recorded in the study notebook.

10.2.1 Preparation of Reference Standard and Positive Control Stock Solutions

Stock solutions of E2 and methoxychlor are prepared in 100% DMSO and stored at room temperature for up to three years or until the expiration date listed in the certificate of analysis for that substance.

10.2.1.1 E2 Stock Solution

The final concentration of the E2 stock solution is 1.0×10^{-2} µg/mL. Prepare the E2 stock as shown in **Table 10-1**.

Table 10-1 Preparation of E2 Stock Solution

Step #	Action	DMSO	E2 Concentration
1	Make a 10 mg/mL stock solution in 100% DMSO in a 4mL vial.	-	10 mg/mL
2	Transfer 10 µL E2 solution from Step #1 to a new 4 mL vial.	Add 990 µL of 100% DMSO. Vortex to mix.	100 µg/mL
3	Transfer 10 µL E2 solution from Step #2 to a new 4mL vial.	Add 990 µL of 100% DMSO. Vortex to mix.	1 µg/mL
4	Transfer 10 µL E2 solution from Step #3 to a 13 mm test tube to create the working solution.	Add 990 µL of 100% DMSO. Vortex to mix.	1.0×10^{-2} µg/mL

10.2.1.2 Methoxychlor Stock Solution

The final concentration of the methoxychlor stock solution is 313 µg/mL.

To prepare the methoxychlor stock solution, proceed as follows:

1. Make a 10 mg/mL stock solution of methoxychlor in 100% DMSO in a 4 mL vial.
2. Remove 94 µL of the methoxychlor solution and place it in a new 4 mL vial.
3. Add 2.906 mL of 100% DMSO to the 4mL vial and gently vortex to mix.

10.2.2 Preparation of Reference Standard and DMSO Control for Range Finder Testing

Range finder testing is conducted on 96-well plates using four concentrations of E2 in duplicate as the reference standard. Four replicate wells are used for the DMSO control. All wells on the 96 well plate are used during range finder testing.

Store dosing solutions at room temperature. Use within 24 hours of preparation.

10.2.2.1 Preparation of E2 Reference Standard for Range Finder Testing

To make E2 dosing solutions:

1. Label four 4 mL conical tubes with numbers 1 through 4 and place them in a tube rack.
2. Label four 13 mm glass test tubes with numbers 1 through 4, place them in a tube rack, and add 600 µL of estrogen-free DMEM to each tube.

Prepare dilutions to give final concentrations of the E2 as shown in **Table 10-2**.

Table 10-2 Preparation of E2 Reference Standard Dosing Solution for Range Finder Testing

Tube Number	100% DMSO	E2 ¹	Estrogen-free DMEM ²	Final Volume	E2 Concentration
1	6 µL	6 µL of 1.0×10^{-2} µg/mL working solution	600 µL	606 µL	5.00×10^{-5} µL
2	18 µL	6 µL of 1.0×10^{-2} µg/mL working solution	600 µL	606 µL	1.25×10^{-5} µL
3	18 µL	6 µL from conical tube #2	600 µL	606 µL	3.13×10^{-6} µL
4	18 µL	6 µL from conical tube #3	600 µL	606 µL	7.83×10^{-7} µL

¹ Add specified volume of 100% DMSO and 6 µL of the specified E2 solution to labeled 4 mL conical tubes, and vortex.

² Transfer 6 µL of DMSO/E2 solution from 4 mL conical tube to labeled 13 mL glass tubes containing DMEM and vortex.

10.2.2.2 Preparation of DMSO Control for Range Finder Testing

1. Add 10 µL of 100% DMSO to a 13 mm glass tube.
2. Add 1000 µL of estrogen-free DMEM to tube and vortex.

10.2.3 Preparation of Test Substance Dosing Solutions for Range Finder Testing

Range finder experiments are used to determine the concentrations of test substance to be used during comprehensive testing. Agonist range finding for coded substances consists of seven point, 1:10 serial dilutions run in duplicate.

To make dosing solutions for coded substances:

1. Label seven 4 mL conical tubes with numbers 1 through 7 and place them in a tube rack
2. Label seven 13 mm glass test tubes with numbers 1 through 7, place them in a tube rack and add 600 µL of estrogen-free DMEM to each tube

Prepare dilutions as shown in **Table 10-3**.

Table 10-3 Preparation of Test Substance Dosing Solutions for Range Finder Testing

Tube Number	100% DMSO	Test Substance ¹	Transfer	Estrogen-free DMEM	Final Volume
1	-	6 µL of test substance solution from Section 10.1 step 10	6 µL	600 µL	606 µL
2	90 µL	10 µL of test substance solution from Section 10.1 step 10	6 µL	600 µL	606 µL
3	90 µL	10 µL from conical tube #2	6 µL	600 µL	606 µL
4	90 µL	10 µL from conical tube #3	6 µL	600 µL	606 µL
5	90 µL	10 µL from conical tube #4	6 µL	600 µL	606 µL
6	90 µL	10 µL from conical tube #5	6 µL	600 µL	606 µL
7	90 µL	10 µL from conical tube #6	6 µL	600 µL	606 µL

¹ Add specified volume of 100% DMSO and test substance solution to labeled 4 mL conical tubes, and vortex.

² Transfer 6 µL of DMSO/E2 solution from 4 mL conical tube to labeled 13 mL glass tubes containing DMEM and vortex.

Determination of whether a substance is positive in range finder testing and selection of starting concentrations for comprehensive testing will be discussed in **Section 12.0**.

10.2.4 Preparation of Reference Standard and Positive Control Dosing Solutions for Comprehensive Testing

Comprehensive testing is conducted on 96-well plates using 11 concentrations of E2 in duplicate as the reference standard. Four replicate wells for the DMSO control and three replicate wells for the methoxychlor control are included on each plate.

Store dosing solutions at room temperature. Use within 24 hours of preparation.

10.2.4.1 Preparation of E2 Reference Standard for Comprehensive Testing

To make E2 dosing solutions:

1. Label 11 4 mL conical tubes with numbers 1 through 11 and place them in a tube rack
2. Label 11 13 mm glass test tubes with numbers 1 through 11, place them in a tube rack and add 600 µL of DMEM to each tube

Prepare dilutions to give final concentrations of E2 as shown in **Table 10-4**.

Table 10-4 Preparation of E2 Reference Standard Dosing Solution for Comprehensive Testing

Tube Number	100% DMSO	E2 ¹	Estrogen-free DMEM ²	Final Volume	E2 Concentration
1	-	6 µL of 1.0×10^{-2} µg/mL working solution	600 µL	606 µL	1.00×10^{-4} µL
2	6 µL	6 µL of 1.0×10^{-2} µg/mL working solution	600 µL	606 µL	5.00×10^{-5} µL
3	6 µL	6 µL from conical tube #2	600 µL	606 µL	2.50×10^{-5} µL
4	6 µL	6 µL from conical tube #3	600 µL	606 µL	1.25×10^{-5} µL
5	6 µL	6 µL from conical tube #4	600 µL	606 µL	6.25×10^{-6} µL
6	6 µL	6 µL from conical tube #5	600 µL	606 µL	3.13×10^{-6} µL
7	6 µL	6 µL from conical tube #6	600 µL	606 µL	1.56×10^{-6} µL
8	6 µL	6 µL from conical tube #7	600 µL	606 µL	7.83×10^{-7} µL
9	6 µL	6 µL from conical tube #8	600 µL	606 µL	3.92×10^{-7} µL
10	6 µL	6 µL from conical tube #9	600 µL	606 µL	1.95×10^{-7} µL
11	6 µL	6 µL from conical tube #10	600 µL	606 µL	9.78×10^{-8} µL

¹ Add specified volume of 100% DMSO and 6 µL of the specified E2 solution to labeled 4 mL conical tubes, and vortex.

² Transfer 6 µL of DMSO/E2 solution from 4 mL conical tube to labeled 13 mL glass tubes containing DMEM and vortex.

10.2.4.2 Preparation of Methoxychlor Control Dosing Solution for Comprehensive Testing

1. Add 10 µL of the 313 µg/mL methoxychlor to a 13 mm glass tube.
2. Add 1000 µL of estrogen-free DMEM to the tube and vortex.

10.2.4.3 Preparation of DMSO Control Dosing Solution for Comprehensive Testing

1. Add 10 µL of 100% DMSO to four 13 mm tubes (solvent/negative controls).
2. Add 1000 µL of estrogen-free DMEM to the tube and vortex.

10.2.5 Preparation of Test Substance Dosing Solutions for Comprehensive Testing

Comprehensive testing experiments are used to determine whether a substance possesses ER agonist activity in the BG1Luc ER TA test method. Agonist comprehensive testing for coded substances consists of either an 11 point 1:2 serial dilution or an 11 point 1:5 serial dilution, depending on the results from

range finder testing (see **Section 12.0**) with each concentration tested in triplicate wells of the 96-well plate.

10.2.5.1 Preparation of Test Substance 1:2 Serial Dilutions for Comprehensive Testing

Start the 11-point serial dilution according to criteria in **Section 12.0**.

To make test substance 1:2 serial dilutions for comprehensive testing:

1. Label eleven 4 mL conical tubes with numbers 1 through 11 and place them in a tube rack.
2. Label eleven 13 mm glass test tubes with numbers 1 through 11, place them in a tube rack and add 800 μ L of estrogen-free DMEM to each tube.

Prepare dilution of test substance as shown in **Table 10-6**.

Table 10-5 Preparation of Test Substance 1:2 Serial Dilutions for Comprehensive Testing

Tube Number	100% DMSO	Test Substance ¹	Transfer	Estrogen-free DMEM	Final Volume
1	-	8 μ L of highest concentration of test substance solution	8 μ L	800 μ L	808 μ L
2	8 μ L	8 μ L of highest concentration of test substance solution	8 μ L	800 μ L	808 μ L
3	8 μ L	8 μ L from conical tube #2	8 μ L	800 μ L	808 μ L
4	8 μ L	8 μ L from conical tube #3	8 μ L	800 μ L	808 μ L
5	8 μ L	8 μ L from conical tube #4	8 μ L	800 μ L	808 μ L
6	8 μ L	8 μ L from conical tube #5	8 μ L	800 μ L	808 μ L
7	8 μ L	8 μ L from conical tube #6	8 μ L	800 μ L	808 μ L
8	8 μ L	8 μ L from conical tube #7	8 μ L	800 μ L	808 μ L
9	8 μ L	8 μ L from conical tube #8	8 μ L	800 μ L	808 μ L
10	8 μ L	8 μ L from conical tube #9	8 μ L	800 μ L	808 μ L
11	8 μ L	8 μ L from conical tube #10	8 μ L	800 μ L	808 μ L

¹Add specified volume of 100% DMSO and test substance solution to labeled 4 mL conical tubes, and vortex.

10.2.5.2 Preparation of Test Substance 1:5 Serial Dilutions for Comprehensive Testing

Start the 11-point serial dilution according to criteria in **Section 12.0**.

To make test substance 1:5 serial dilutions for comprehensive testing:

1. Label eleven 4 mL conical tubes with numbers 1 through 11 and place them in a tube rack.
2. Label eleven 13 mm glass test tubes with numbers 1 through 11, place them in a tube rack and add 800 μ L of estrogen-free DMEM to each tube.

Prepare dilution of test substance as shown in **Table 10-6**.

Table 10-6 Preparation of Test Substance 1:5 Serial Dilutions for Comprehensive Testing

Tube Number	100% DMSO	Test Substance ¹	Transfer	Estrogen-free DMEM	Final Volume
1	-	8 μ L of highest concentration of test substance solution	8 μ L	800 μ L	808 μ L
2	16 μ L	4 μ L of highest concentration of test substance solution	8 μ L	800 μ L	808 μ L
3	16 μ L	4 μ L from conical tube #2	8 μ L	800 μ L	808 μ L
4	16 μ L	4 μ L from conical tube #3	8 μ L	800 μ L	808 μ L
5	16 μ L	4 μ L from conical tube #4	8 μ L	800 μ L	808 μ L

Tube Number	100% DMSO	Test Substance ¹	Transfer	Estrogen-free DMEM	Final Volume
6	16 µL	4 µL from conical tube #5	8 µL	800 µL	808 µL
7	16 µL	4 µL from conical tube #6	8 µL	800 µL	808 µL
8	16 µL	4 µL from conical tube #7	8 µL	800 µL	808 µL
9	16 µL	4 µL from conical tube #8	8 µL	800 µL	808 µL
10	16 µL	4 µL from conical tube #9	8 µL	800 µL	808 µL
11	16 µL	4 µL from conical tube #10	8 µL	800 µL	808 µL

¹ Add specified volume of 100% DMSO and test substance solution to labeled 4 mL conical tubes, and vortex.

11.0 General Procedures for the Testing of Coded Substances

Range finder experiments are used to determine the concentrations of test substance to be used during comprehensive testing. Comprehensive testing experiments are used to determine whether a substance possesses ER agonist activity in the BG1Luc ER TA assay.

General procedures for range finder and comprehensive are similar. For specific details (such as plate layout) of range finder testing see **Section 12.0**. For specific details of comprehensive testing, see **Section 13.0**.

11.1 Application of Reference Standard, Controls, and Test Substances

1. Remove seeded 96-well plates from the incubator, inspect them using an inverted microscope. Only use plates in which the cells in all wells giving a score of 1 according to **Table 11-1**.
2. Remove medium by inverting the plate onto blotter paper. Gently tap plate against the bench surface to remove residual liquid trapped in the wells.
3. Add 200 µL of reference standard, control, or test substance to each well (see **Sections 12.0** and **13.0** for specific plate layouts).
4. Return plates to incubator and incubate (see **Section 9.0** for details) for 19 to 24 hours to allow maximal induction of luciferase activity in the cells.

11.1.1 Preparation of Microsoft Excel® Data Analysis Template For Range Finder Testing

1. In Excel, open a new “AgRFTemplate” and save it with the appropriate project name as indicated in the NICEATM Style Guide.
2. Fill out the table at the top of the “Raw Data” worksheet with information regarding the Microplate reader used, Reading Direction, No. of Intervals, Tot. Meas. Time/Well (s), etc. (**note:** this information can be permanently added to the default template “AgRFTemplate” on a laboratory specific basis).
3. Add the following information regarding the assay to the “Compound Tracking” worksheet.
 - Plate # - Enter the experiment ID or plate number into cell E1
 - Cell Lot # - Enter the passage or lot number of the cells used for this experiment into cell B5
 - DMSO and Media Lot #'s – Enter the lot numbers for the DMSO and Media in cells B6 and B7
 - Test Substance Code – Enter the test substance codes into cells C13 to C18
 - Name: Enter the experimenter name into cell G6
 - Date: Enter the experiment date in the format day/month/year into cell G10
 - Comments: - Enter any comments about the experiment in this box (e.g., plate contaminated)
4. Enter the following substance testing information to the “List” page:
 - Concentration – Type in the test substance concentration in µg/mL in descending order.
 - Also add any replicate-specific comments on this page (e.g, spilled tube, etc.), in the comments section
 - All of the remaining cells on the List tab should populate automatically.

- The “Template”, “Compound Mixing” and “Visual Inspection” tabs should automatically populate with the information entered into the Compound Tracking and List tabs.
5. Save the newly named project file.
 6. Print out either the “List” or “Template” page for help with dosing the 96-well plate. Sign and date the print out and store in study notebook.

11.1.2 Preparation of Microsoft Excel Data Analysis Template for Comprehensive Testing

1. In Excel, open a new “AgCTTemplate” and save it with the appropriate project name as indicated in the NICEATM Style Guide.
2. Fill out the table at the top of the “Raw Data” worksheet with information regarding the Microplate reader used, Reading Direction, No. of Intervals, Tot. Meas. Time/Well (s), etc. (**note:** this information can be permanently added to the default template “AgCTTemplate” on a laboratory specific basis).
3. On the “Compound Tracking” tab, enter the following information:
 - Plate # - Enter the experiment ID or plate number into cell E1
 - Cell Lot # - Enter the passage or lot number of the cells used for this experiment into cell C5
 - DMSO and Media Lot #'s – Enter the lot numbers for the DMSO and Media in cells C6 and C7
 - Test Substance Code – Enter the test substance codes into cells C15 and C16. Enter the test substance dilution into cells E25 and E26.
 - Name: Enter the experimenter name into cell G6
 - Date: Enter the experiment date in the format day\month\year into cell G10
 - Comments: - Enter any comments about the experiment in this box (e.g., plate contaminated)
4. Enter substance testing concentrations to the “List” page. Also add any replicate-specific comments on this page (e.g, spilled tube, etc.).
5. Save the newly named project file.
6. Print out either the “List” or “Template” page for help with dosing the 96-well plate. Sign and date the print out and store in study notebook.

11.2 Visual Evaluation of Cell Viability

1. Nineteen to 24 hours after dosing the plate, remove the plate from the incubator and remove the media from the wells by inverting the plate onto blotter paper. Gently tap plate against the bench surface to remove residual liquid trapped in the wells.
2. Use a repeat pipetter to add 50 μ L 1X PBS to all wells. Immediately remove PBS by inversion.
3. Using an inverted microscope, inspect all of the wells used in the 96-well plate and record the visual observations using the scores in **Table 11-1**.

Table 11-1 Visual Observation Scoring

Viability Score	Brief Description ¹
1	Normal Cell Morphology and Cell Density
2	Altered Cell Morphology and/or Small Gaps between Cells
3	Altered Cell Morphology and/or Large Gaps between Cells
4	Few (or no) Visible Cells
P	Wells containing precipitation are to be noted with “P”

¹ Reference photomicrographs are provided in the BG1Luc ER TA Validation Study “Visual Observation Cell Viability Manual.”

11.3 Lysis of Cells for BG1Luc ER TA

1. Apply the reflective white backing tape to the bottom of the 96-well plate (this will increase the effectiveness of the luminometer).
2. Add 30 µL 1X lysis reagent to the assay wells and place the 96-well plate on an orbital shaker for one minute.
3. Remove plate from shaker and measure luminescence (as described in Section 11.4).

11.4 Measurement of Luminescence

Luminescence is measured in the range of 300 to 650 nm, using an injecting luminometer and with software that controls the injection volume and measurement interval. Light emission from each well is expressed as RLU per well. The luminometer output is saved as raw data in an Excel® spreadsheet. A hard copy of the luminometer raw data should be signed, dated and stored in the study notebook.

11.5 Data Analysis

BG1Luc ER TA uses a Microsoft Excel spreadsheet to collect and adjust the RLU values obtained from the luminometer and a GraphPad Prism® template to analyze and graph data. The Excel spreadsheet subtracts background luminescence (average DMSO solvent control RLU value) from test substance, reference standard and control RLU values. Plate induction is calculated using these corrected RLU values. Test substance, reference standard, and control RLU values are then adjusted relative to the highest E2 reference standard RLU value, which is set to 10,000. After adjustment, values are transferred to GraphPad Prism for data analysis and graphing.

11.5.1 Collection and Adjustment of Luminometer Data for Range Finder Testing

The following steps describe the procedures required to populate the Excel® spreadsheet that has been configured to collect and adjust the RLU values obtained from the luminometer.

1. Open the raw data file and the corresponding experimental Excel spreadsheet from **Section 11.1.1**.
2. Copy the raw data using the Excel copy function, then paste the copied data into cell B19 of the “RAW DATA” tab in the experimental Excel® spreadsheet using the **Paste Special – Values** command. This position corresponds to position A1 in the table labeled Table 1 in this tab.
3. Examine the DMSO data in Table 1 of the Excel spreadsheet to determine whether there are any potential outliers. See **Section 11.6.2** for further explanation of outlier determinations.
4. If an outlier is identified, perform the following steps to remove the outlier from calculations:
 - correct the equation used to calculate DMSO background in Table 1 [e.g., if outlier is located in cell F26, adjust the calculation in cell H40 to read =AVERAGE(G26:I26)]
 - then correct the equation used to calculate the average DMSO value in Table 2 [e.g., following the above example, adjust cell M42 to read =AVERAGE(G26:I26)]
 - then correct the equation used to calculate the standard deviation of the DMSO value in Table 2 [e.g., following the above example, adjust cell M43 to read =STDEV(G36:I36)]
5. Excel will automatically subtract the background (the average DMSO control value) from all of the RLU values in Table 1 and populate Table 2 with these adjusted values.
6. To calculate plate induction, identify the cell containing the E2a replicate in Table 1, plate row H that has the highest RLU value (i.e., cell B26, C26, D26, or E26).
7. Click into cell D14 and enter the cell number from the previous step into the numerator.
8. Identify the cell containing the E2b replicate in Table 1, plate row H that has the highest RLU value (i.e., cell J26, K26, L26, or M26).
9. Click into cell E14 and enter the cell number from the previous step into the numerator.
10. Click on the “ER Agonist Report” worksheet.

11. The data for the E2 reference standard, methoxychlor, and DMSO replicates populate the left portion (columns A – F) of the spreadsheet. The data is automatically placed in an Excel® graph.
12. To set the highest RLU value for the reference standard to 10,000 RLU, go to cell D2 of “ER Agonist Report” tab and check the formula contained within that cell. The divisor should be the cell number of the cell containing the highest Mean E2 RLU value (i.e., cell A16, A17, A18, or A19).
13. Open the “Visual Observation Scoring” worksheet. Enter the visual observation scores for each well on the 96-well plate. This data will be linked to the “ER Agonist Report” worksheet.
14. After the testing results have been evaluated and reviewed for quality control, enter the following information into the Compound Tracking worksheet:
 - Enter pass/fail results for plate reference standard and control parameters into the Plate Pass/Fail Table
 - Enter information from the testing of coded substances into the Testing Results Table
 - Reviewer Name – Enter the name of the person who Reviewed\QC’ed the data into cell A34
 - Date – Enter the date on which the data was reviewed into cell D34

11.5.2 Collection and Adjustment of Luminometer Data for Comprehensive Testing

The following steps describe the procedures required to populate the Excel spreadsheet that has been configured to collect and adjust the RLU values obtained from the luminometer.

1. Open the raw data file and the corresponding experimental Excel spreadsheet from **Section 11.1.2**.
2. Copy the raw data using the Excel copy function, then paste the copied data into cell B16 of the “RAW DATA” worksheet in the experimental Excel spreadsheet using the **Paste Special – Values** command. This position corresponds to position A1 in the table labeled Table 1 in this worksheet.
3. Fill out the table at the top of the “Raw Data” worksheet with information regarding the Microplate reader used, Reading Direction, No. of Intervals, Tot. Meas. Time/Well (s), etc. If desired, this information can be added to the Laboratory Template File.
4. Examine the DMSO data in Table 1 of the Excel spreadsheet to determine whether there are any potential outliers. See **Section 11.6.2** for further explanation of outlier determinations.
5. If an outlier is identified, perform the following steps to remove the outlier from calculations:
 - correct the equation used to calculate DMSO background in Table 1 [e.g., if outlier is located in cell M17, adjust the calculation in cell H37 to read =AVERAGE(M16,M18:M19)]
 - then correct the equation used to calculate the DMSO mean and SD values [e.g., following the above example, adjust cell M39 to read =AVERAGE(M28,M30:M31), and adjust cell M40 to read =STDEV(M28,M30:M31)]
6. Excel will automatically subtract the background (the average DMSO control value) from all of the RLU values in Table 1 and populate Table 2 with these adjusted values.
7. To calculate plate induction, identify the cell in containing the E2 replicate in Table 1, plate row G that has the highest RLU value.
8. Click into cell D11 and enter the cell number from the previous step into the numerator.
9. Identify the cell containing the E2 replicate in plate row H that has the highest RLU value.
10. Click into cell E11 and enter the cell number from the previous step into the numerator.
11. Open the “ER Agonist Report” worksheet.
12. The data for the E2 reference standard, methoxychlor, and DMSO replicates populate the left portion (columns A – E) of the spreadsheet. The data is automatically placed in an Excel graph.
13. To set the highest RLU value for the reference standard to 10,000 RLU, go to cell E2 of “ER Agonist Report” tab and check the formula contained within that cell. The divisor should be the cell number of the cell containing the highest Avg E2 RLU value (cells A16 through A26).
14. Open the “Visual Observation Scoring” worksheet. Enter the visual observation scores for each well on the 96-well plate. This data will be linked to the “ER Agonist Report” worksheet.
15. Copy the data from the “ER Agonist Report” worksheet into GraphPad Prism for the calculation of EC₅₀ values and to graph experimental results as indicated in the NICEATM Prism Users Guide.

16. After the testing results have been evaluated and reviewed for quality control, enter the following information into the Compound Tracking worksheet:
- Enter pass/fail results for plate reference standard and control parameters into the Plate Pass/Fail Table
 - Enter information from the testing of coded substances into the Testing Results Table
 - Reviewer Name – Enter the name of the person who Reviewed\QC’ed the data into cell A32
 - Date – Enter the date on which the data was reviewed into cell D32

11.5.3 Determination of Outliers

The Study Director will use good statistical judgment for determining “unusable” wells that will be excluded from the data analysis and will provide an explanation in the study notebook for any excluded data. This judgment for data acceptance will include Q-test analysis.

The formula for the Q test is:

Outlier – Nearest Neighbor

Range (Highest – Lowest)

where the outlier is the value proposed for exclusion, the nearest neighbor is the value closest to the outlier, and the range is the range of the three values (Q values for samples sizes from 3 to 10 are provided in **Table 11-2**). For example, if the value of this ratio is greater than 0.94 (the Q value for the 90% confidence interval for a sample size of three) or 0.76 (the Q value for the 90% confidence interval for a sample size of four), the outlier may be excluded from data analysis.

Table 11-2 Q Test Values

Number Of Observations	Q Value
2	-
3	0.94
4	0.76
5	0.64
6	0.56
7	0.51
8	0.47
9	0.44
10	0.41

For E2 reference standard replicates (sample size of two), any adjusted RLU value for a replicate at a given concentration of E2 is considered and outlier if its value is more than 20% above or below the adjusted RLU value for that concentration in the historical database.

11.5.4 Acceptance Criteria

11.5.4.1 Range Finder Testing

Acceptance or rejection of a test is based on evaluation of reference standard and control results from each experiment conducted on a 96-well plate. Results are compared to quality controls (QC) for these parameters derived from the historical database, which are summarized below.

- Induction: Plate induction, as measured by dividing the averaged highest E2 reference standard RLU value by the averaged DMSO control RLU value, must be greater than three-fold.

- DMSO control results: Solvent control RLU values must be within 2.5 times the standard deviation of the historical solvent control mean RLU value.

An experiment that fails either acceptance criterion will be discarded and repeated.

11.5.4.2 Comprehensive Testing

Acceptance or rejection of a test is based on evaluation of reference standard and control results from each experiment conducted on a 96-well plate. Results are compared to quality controls (QC) for these parameters derived from the historical database, which are summarized below.

- Induction: Plate induction, as measured by dividing the averaged highest E2 reference standard RLU value by the averaged DMSO control RLU value, must be greater than three-fold.
- Reference standard results: The E2 reference standard concentration-response curve should be sigmoidal in shape and have at least three values within the linear portion of the concentration-response curve.
- DMSO control results: DMSO control RLU values must be within 2.5 times the standard deviation of the historical solvent control mean RLU value.
- Positive control results: Methoxychlor control RLU values must be above the line representing the DMSO mean plus three times the standard deviation from the DMSO mean.

An experiment that fails any single acceptance criterion will be discarded and repeated.

12.0 Range Finder Testing

Agonist range finding for coded substances consists of seven point, 1:10 serial dilutions, with each concentration tested in duplicate wells of the 96-well plate. **Figure 12-1** contains a template for the plate layout to be used in agonist range finder testing.

Figure 12-1 Agonist Range Finder Test Plate Layout

TS1-1	TS1-1	TS2-1	TS2-1	TS3-1	TS3-1	TS4-1	TS4-1	TS5-1	TS5-1	TS6-1	TS6-1
TS1-2	TS1-2	TS2-2	TS2-2	TS3-2	TS3-2	TS4-2	TS4-2	TS5-2	TS5-2	TS6-2	TS6-2
TS1-3	TS1-3	TS2-3	TS2-3	TS3-3	TS3-3	TS4-3	TS4-3	TS5-3	TS5-3	TS6-3	TS6-3
TS1-4	TS1-4	TS2-4	TS2-4	TS3-4	TS3-4	TS4-4	TS4-4	TS5-4	TS5-4	TS6-4	TS6-4
TS1-5	TS1-5	TS2-5	TS2-5	TS3-5	TS3-5	TS4-5	TS4-5	TS5-5	TS5-5	TS6-5	TS6-5
TS1-6	TS1-6	TS2-6	TS2-6	TS3-6	TS3-6	TS4-6	TS4-6	TS5-6	TS5-6	TS6-6	TS6-6
TS1-7	TS1-7	TS2-7	TS2-7	TS3-7	TS3-7	TS4-7	TS4-7	TS5-7	TS5-7	TS6-7	TS6-7
E2-1	E2-2	E2-3	E2-4	VC	VC	VC	VC	E2-1	E2-2	E2-3	E2-4

Abbreviations: E2 = E2 control; TS = test substance; VC = vehicle control (DMSO [1% v/v EFM]).

E2-1 to E2-4 = concentrations of the E2 reference standard (from high to low)

TS1-1 to TS1-7 = concentrations (from high to low) of test substance 1 (TS1)

TS2-1 to TS2-7 = concentrations (from high to low) of test substance 2 (TS2)

TS3-1 to TS3-7 = concentrations (from high to low) of test substance 3 (TS3)

TS4-1 to TS4-7 = concentrations (from high to low) of test substance 4 (TS4)

TS5-1 to TS5-7 = concentrations (from high to low) of test substance 5 (TS5)

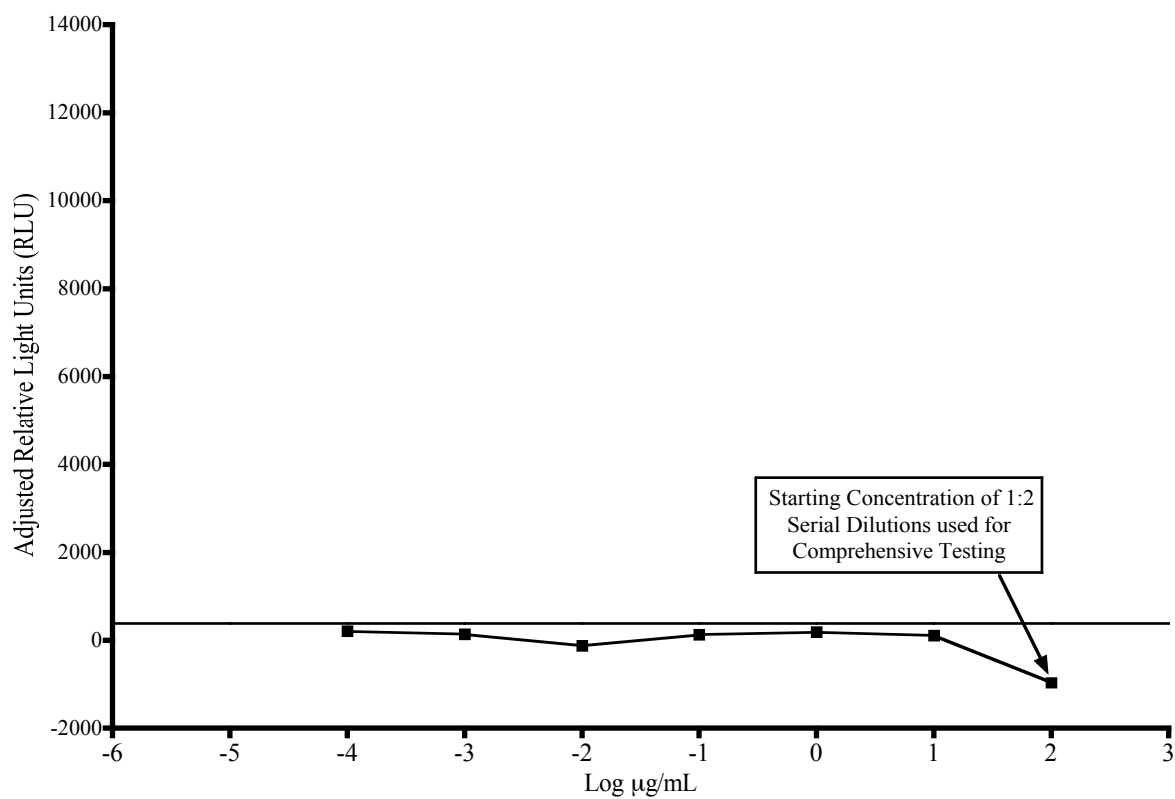
TS6-1 to TS6-7 = concentrations (from high to low) of test substance 6 (TS6)

Evaluate whether range finder experiments have met the acceptance criteria (see **Section 11.5.4.1**).

To determine starting concentrations for comprehensive testing use the following criteria:

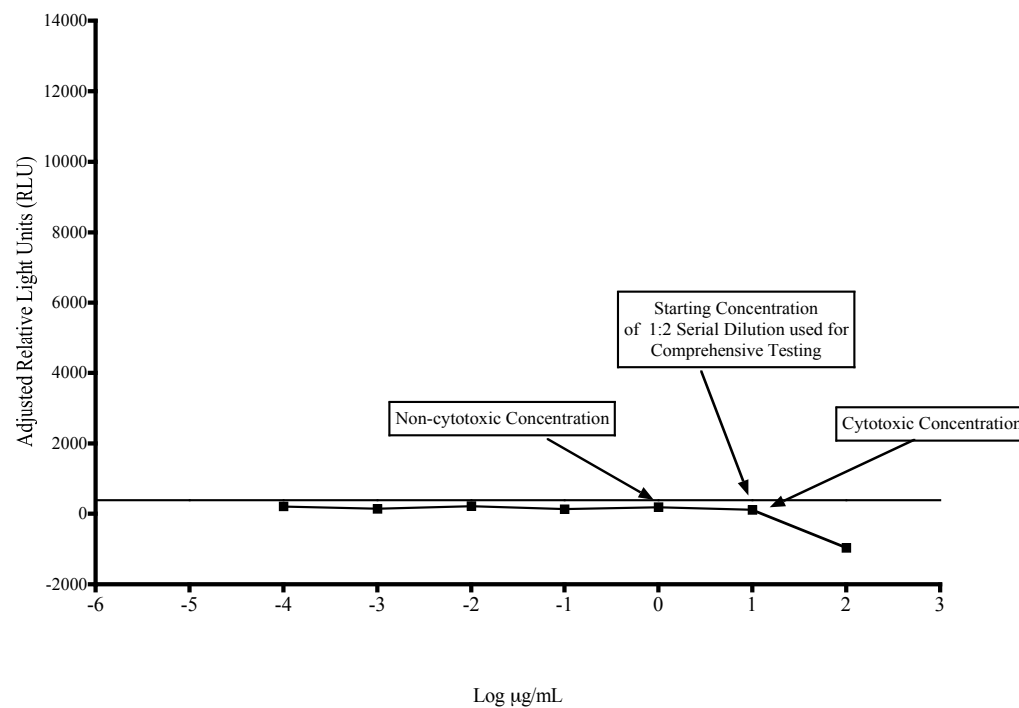
- If results in the range finder test suggest that the test substance is negative for agonist activity (i.e., if there are no points on the test substance concentration curve that are greater than the mean plus three times the standard deviation of the DMSO control, see **Figure 12-2**), comprehensive testing will be conducted using an 11 point 1:2 serial dilution starting at the maximum soluble concentration.
- If results in the range finder test suggest that the test substance is negative for agonist activity (i.e., if there are no points on the test substance concentration curve that are greater than the mean plus three times the standard deviation of the DMSO control), and the higher concentrations in the range finder are cytotoxic, comprehensive testing will be conducted using an 11 point 1:2 serial dilution with the lowest cytotoxic concentration as the starting concentration (see **Figure 12-3**).
- If results in the range finder test suggest that the test substance is positive for agonist activity (i.e., if there are points on the test substance concentration curve that are greater than the mean plus three times the standard deviation of the DMSO control), the starting concentration to be used for the 11-point dilution scheme in comprehensive testing should be one log higher than the concentration giving the highest adjusted RLU value in the range finder. The 11-point dilution scheme will be based on either 1:2 or 1:5 dilutions according to the following criteria:
 - An 11-point 1:2 serial dilution should be used if the resulting concentration range (note: an 11-point 1:2 serial dilution will cover a range of concentrations over approximately three orders of magnitude [three logs]) will encompass the full range of responses based on the concentration response curve generated in the range finder test (see **Figure 12-4**).
 - If the concentration range that would be generated with the 1:2 serial dilution will not encompass the full range of responses based on the concentration response curve in the range finder test (see **Figure 12-5** and **12-6**), an 11-point 1:5 serial dilution should be used instead.
- If a substance exhibits a biphasic concentration response curve in the range finder test, both phases should also be resolved in comprehensive testing. In order to resolve both curves, the starting concentration should be based on the peak associated with the higher concentration and should be one log higher than the concentration giving the highest adjusted RLU value in the range finder. As an example, an 11-point 1:5 serial dilution should be used based on the range finder results presented in **Figure 12-7**.

Figure 12-2 Agonist Range Finder (Example 1)



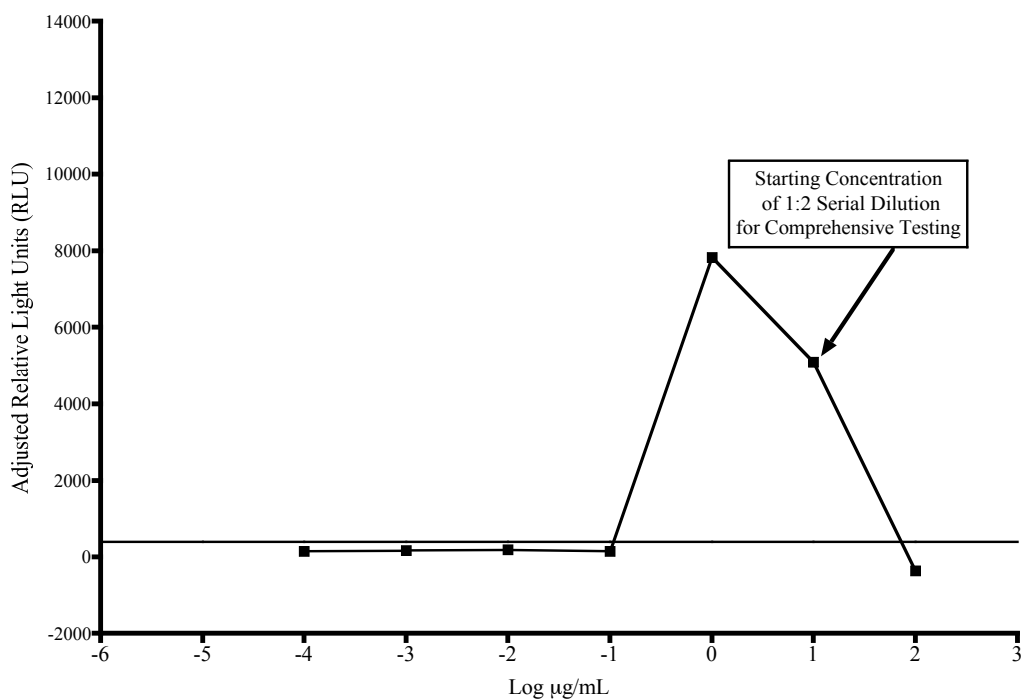
The solid horizontal line represents the mean plus three times the standard deviation of the DMSO control.

Figure 12-3 Agonist Range Finder (Example 2)

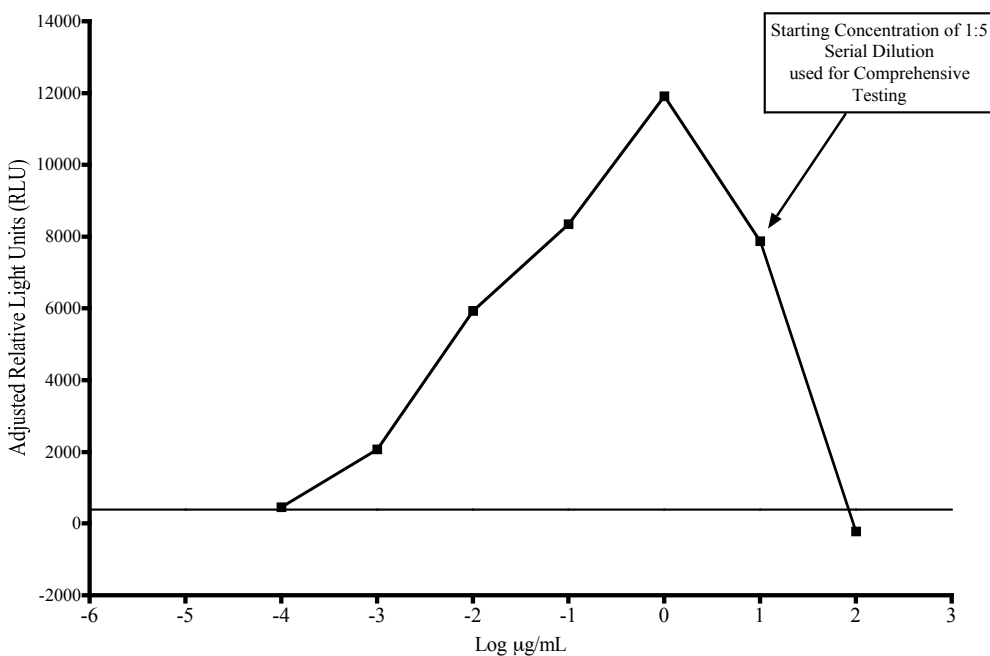


The solid horizontal line represents the mean plus three times the standard deviation of the DMSO control.

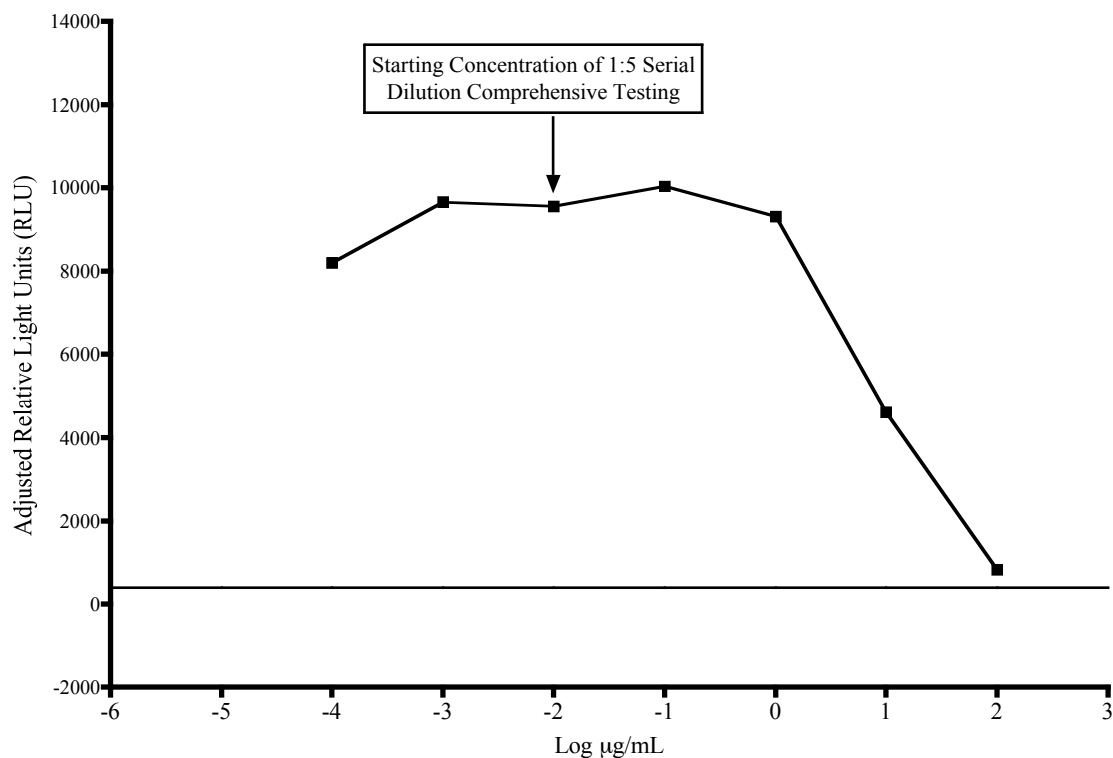
Figure 12-4 Agonist Range Finder (Example 3)



The solid horizontal line represents the mean plus three times the standard deviation of the DMSO control.

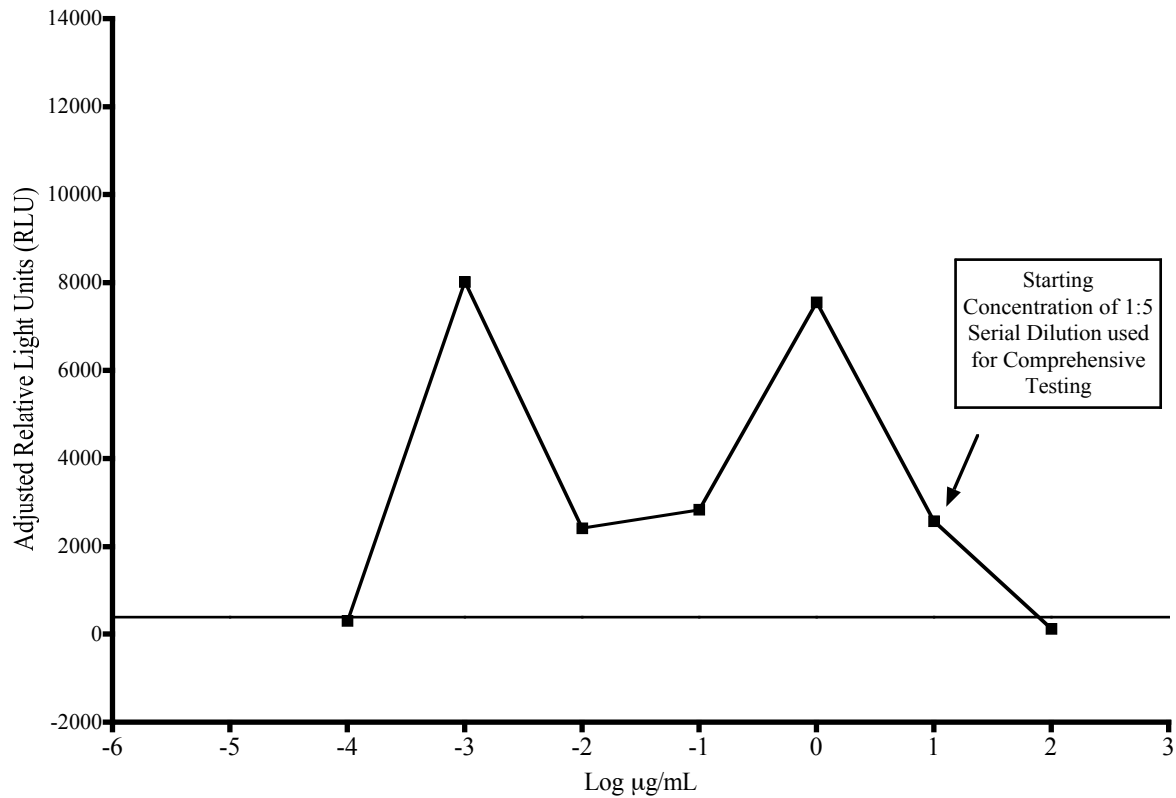
Figure 12-5 Agonist Range Finder (Example 4)

The solid horizontal line represents the mean plus three times the standard deviation of the DMSO control.

Figure 12-6 Agonist Range Finder (Example 5)

The solid horizontal line represents the mean plus three times the standard deviation of the DMSO control.

Figure 12-7 Agonist Range Finder (Example 6)



The solid horizontal line represents the mean plus three times the standard deviation of the DMSO control.

13.0 Comprehensive Testing

Agonist comprehensive testing for coded substances consists of 11 point serial dilutions (either 1:2 or 1:5 serial dilutions based on the starting concentration for comprehensive testing criteria in **Section 12.0**) with each concentration tested in triplicate wells of the 96-well plate. **Figure 13-1** contains a template for the plate layout to be used in agonist comprehensive testing.

Figure 13-1 Agonist Comprehensive Test Plate Layout

TS1-1	TS1-2	TS1-3	TS1-4	TS1-5	TS1-6	TS1-7	TS1-8	TS1-9	TS1-10	TS1-11	VC
TS1-1	TS1-2	TS1-3	TS1-4	TS1-5	TS1-6	TS1-7	TS1-8	TS1-9	TS1-10	TS1-11	VC
TS1-1	TS1-2	TS1-3	TS1-4	TS1-5	TS1-6	TS1-7	TS1-8	TS1-9	TS1-10	TS1-11	VC
TS2-1	TS2-2	TS2-3	TS2-4	TS2-5	TS2-6	TS2-7	TS2-8	TS2-9	TS2-10	TS2-11	VC
TS2-1	TS2-2	TS2-3	TS2-4	TS2-5	TS2-6	TS2-7	TS2-8	TS2-9	TS2-10	TS2-11	Met
TS2-1	TS2-2	TS2-3	TS2-4	TS2-5	TS2-6	TS2-7	TS2-8	TS2-9	TS2-10	TS2-11	Met
E2-1	E2-2	E2-3	E2-4	E2-5	E2-6	E2-7	E2-8	E2-9	E2-10	E2-11	Met
E2-1	E2-2	E2-3	E2-4	E2-5	E2-6	E2-7	E2-8	E2-9	E2-10	E2-11	Met

Abbreviations: E2 = E2 control; Met = *p,p'*-methoxychlor weak positive control; TS = test substance; VC = vehicle control (DMSO [1% v/v EFM]).

TS1-1 to TS1-11 = concentrations (from high to low) of test substance 1

TS2-1 to TS2-11 = concentrations (from high to low) of test substance 2

Evaluate whether comprehensive experiments have met acceptance criteria (see **Section 11.6.4**) and graph the data as described in the NICEATM Prism® users guide. Then evaluate the test substance results.

13.1 Test Substance Positive and Negative Criteria

Positive Classification:

- All test substances classified as positive for ER agonist activity should have a concentration–response curve consisting of a baseline, followed by a positive slope, and concluding in a plateau or peak. In some cases, only two of these characteristics (baseline–slope or slope–peak) may be defined.
- The line defining the positive slope must contain at least three points with nonoverlapping error bars (mean ± SD). Points forming the baseline are excluded, but the linear portion of the curve may include the peak or first point of the plateau.
- A positive classification requires a response amplitude, the difference between baseline and peak, of at least 20% of the maximal value for the reference estrogen (i.e., 2000 RLUs when the maximal response value of the reference estrogen is adjusted to 10,000 RLUs).
- If possible, an EC₅₀ value should be calculated for each positive substance.

Negative Classification:

- For all concentration–response curves that fail to meet the criteria for a positive response, test substances are classified as negative for agonist activity if all data points are below 20% of the maximal value for the reference estrogen (i.e., 2000 RLUs when the maximal response value of the reference estrogen is adjusted to 10,000 RLUs).

Inadequate Classification:

- 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.

14.0 Compilation of the Historical Quality Control Database

Historical databases are maintained in order to ensure that the assay is functioning properly. Historical databases are compiled using Excel spreadsheets and are separate from the spreadsheets used to collect the data for individual test plates. Reference standard and control data are used to develop and maintain the historical database and are used as quality controls to determine acceptance of individual test plates.

The sources of the data needed to compile the historical database for the DMSO control are the experiment specific Excel data collection and analysis spreadsheets used for BG1Luc ER TA agonist and antagonist testing (see **Section 11.5.2** of the BG1Luc ER TA agonist protocol and **Section 13.5.2** in the BG1Luc ER TA antagonist protocol).

14.1 DMSO Control

Open the combined agonist and antagonist BG1Luc ER TA historical database Excel spreadsheet (LUMI_AgandAntQC.xls) and save under a new name using the Excel “Save As” function, adding the laboratory designator to the file name (e.g., for Laboratory H, the new name would be HLUMI_AgandAntQC.xls). Enter the date and experiment name into worksheet columns A and B respectively. Enter the experimental mean DMSO control value (from cell H37 in the RAW DATA worksheet of the agonist and antagonist Excel data collection and analysis spreadsheet) into worksheet column C. Acceptance or rejection of the plate DMSO control data for range finding and comprehensive testing is based on whether the mean plate DMSO RLU value falls within 2.5 times the standard deviation of the DMSO value in the historical database (columns G and H in the DMSO worksheet).

15.0 Quality Testing of Materials

All information pertaining to the preparation and testing of media, media supplements, and other materials should be recorded in the Study Notebook.

15.1 Tissue Culture Media

Each lot of tissue culture medium must be tested in a single growth flask of cells before use in ongoing tissue culture or experimentation (**note:** each bottle within a given lot of charcoal/dextran treated FBS must be tested separately).

1. Every new lot of media (RPMI and DMEM) and media components (FBS, charcoal/dextran treated FBS, and L-glutamine) must first be tested on the BG1Luc ER TA prior to being used in any GLP acceptable assays.
2. Add 4 μ L of DMSO (previously tested) into four separate 13 mm tubes.
3. Add 400 μ L media (to be tested) to the same tubes.
4. Dose an experimental plate as in **Section 12.0**, treating the media being tested as a test substance.
5. Analyze 96-well plate as described in **Section 12.0**, comparing the data from the DMSO controls made using previously tested tissue culture media to the new media being tested.
6. Use the agonist historical database to determine if the new media with DMSO lies within 2.5 standard deviations of the mean for the media. If the RLU values for the new media with DMSO lie within 2.5 standard deviation of the mean for the historical data on DMSO, the new lot of media is acceptable. If the RLU values for the new media with DMSO do not lie within 2.5 standard deviations of the DMSO mean from historical database, the new lot may not be used in the assay.
7. Note date and lot number in study notebook.

8. If the new bottle passes quality testing as described in Section 15.1 step 6, apply the media to a single flask of cells and observe cell growth and morphology over the following 2 – 3 days. If there is no change in growth or morphology, the new media is acceptable for use.

15.2 G418

1. New lots of G418 must first be tested on the BG1Luc ER TA prior to being used in any GLP acceptable assays.
2. Add 220 μ L of G418 (previously tested) to a single flask containing cells growing in RPMI.
3. Add 220 μ L of G418 (to be tested) to a different flask containing cells growing in RPMI.
4. Observe cellular growth and morphology in both tissue culture flasks over a 48 to 72 hour period. If there are no differences in observed growth rate and morphology between the two flasks, the new G418 lot is acceptable.
5. If cellular growth is decreased, or the cells exhibit abnormal morphology, the new lot of G418 is not acceptable.
6. Note date and lot number in study book.

15.3 DMSO

1. Every new bottle of DMSO must be tested on the BG1Luc ER TA prior to use in any GLP acceptable assays.
2. Add 4 μ L of DMSO (to be tested) into four separate 13 mm tubes.
3. Add 400 μ L media (previously tested) to the same tubes.
4. Dose an experimental plate as in **Section 12.0**, treating the DMSO containing media being tested as a test substance.
5. Analyze 96-well plate as described in **Section 12.0**, comparing the data from the DMSO controls made using previously tested tissue culture media.
6. Use the agonist historical database to determine if media with new DMSO lies within 2.5 standard deviations of the DMSO mean from historical database. If the RLU values for the media with new DMSO lie within 2.5 standard deviations of the DMSO mean from the historical database, the new lot of DMSO is acceptable. If the RLU values for media with new DMSO do not lie within 2.5 standard deviations of the DMSO mean from historical database, the new lot may not be used in the assay.
7. Note the date, lot number, and bottle number in study book.
8. If no DMSO has been previously tested, test several bottles as described in Section 15.3, and determine whether any of the bottles of DMSO have a lower average RLU than the other bottle(s) tested. Use the DMSO with the lowest average RLU for official experiments.

15.4 Plastic Tissue Culture Materials

1. Grow one set of cells, plate them for experiments on plastic ware from the new lot and one set of cells in the plastic ware from a previous lot, and dose them with E2 reference standard and controls.
2. Perform the BG1Luc ER TA experiment with both sets of cells.
3. If all of the analysis falls within acceptable QC criteria, then the new manufacturer's products may be used.

16.0 References

Coecke, S., Guidance on Good Cell Culture Practice. 2005. ATLA 33, 261-287.

Eli Lilly and Company and National Institutes of Health Chemical Genomics Center. 2005. Assay Guidance Manual Version 4.1. Bethesda, MD: National Institutes of Health. Available: http://www.ncgc.nih.gov/guidance/manual_toc.html [accessed 05 September 2006]

ICCVAM. 2001. Guidance Document on Using *In Vitro* Data to Estimate *In Vivo* Starting Doses for Acute Toxicity. NIH Pub. No. 01-4500. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/methods/invidocs/guidance/iv_guide.pdf [accessed 31 August 2006]

Appendix B2

BG1Luc ER TA – Antagonist Protocol

This page intentionally left blank

**BG1Luc ER TA TEST METHOD
ANTAGONIST PROTOCOL**

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

List of Acronyms and Abbreviations

13 mm test tube	13 x 100 mm glass test tubes
DMEM	Dulbecco's Modification of Eagle's Medium
DMSO	Dimethyl Sulfoxide
DMSO Control	1% v/v dilution of DMSO in tissue culture media used as a vehicle control
E2	17 β -estradiol
E2 Control	2.5 x 10 ⁻⁵ μ g/mL E2 used as a control.
IC ₅₀ Value	Concentration that produces a half-maximal response as calculated using the four parameter Hill function.
ER	Estrogen Receptor
Estrogen-free DMEM	DMEM (phenol red free), supplemented with 1 % Penicillin/Streptomycin, 2 % L-Glutamine, and 5% Charcoal-dextran treated FBS
FBS	Fetal Bovine Serum
TAM/E2 Control	1.26 μ g/mL tamoxifen + 2.5 x 10 ⁻⁵ μ g/mL E2, used as a weak positive control.
G418	Gentamycin
Ral/E2 Reference Standard	Nine point dilution of raloxifene HCl + 2.5 x 10 ⁻⁵ 17 β -estradiol reference standard for the BG1Luc ER TA antagonist assay
RPMI	RPMI 1640 growth medium
TA	Transcriptional Activation
T25	25 cm ² tissue culture flask
T75	75 cm ² tissue culture flask
T150	150 cm ² tissue culture flask

1.0 Purpose

This protocol is designed to evaluate coded test substances for potential estrogen receptor (ER) antagonist activity using the BG1Luc ER TA test method.

2.0 Sponsor

(As Appropriate)

3.0 Definitions

- **Dosing Solution:** The test substance, control substance, or reference standard solution which is to be placed into the tissue culture wells for experimentation.
- **Raw Data:** Raw data includes information that has been collected but not formatted or analyzed, and consists of the following:
 - Data recorded in the Study Notebook
 - Computer printout of initial luminometer data
 - Other data collected as part of GLP compliance, e.g.:
 - Equipment logs and calibration records
 - Test substance and tissue culture media preparation logs
 - Cryogenic freezer inventory logs
- **Soluble:** Test substance exists in a clear solution without visible cloudiness or precipitate.
- **Study Notebook:** The study notebook contains recordings of all activities related to the conduct of the BG1Luc ER TA TA antagonist test method.
- **Test Substances:** Substances supplied to the testing laboratories that are coded and distributed such that only the Project Officer, Study Management Team (SMT), and the Substance Inventory and Distribution Management have knowledge of their true identity. The test substances will be purchased, aliquoted, coded, and distributed by the Supplier under the guidance of the NIEHS/NTP Project Officer and the SMT.

4.0 Testing Facility and Key Personnel

4.1 Testing Facility

(As Appropriate)

4.2 Key Personnel

- Study Director: (As Appropriate)
- Quality Assurance Director: (As Appropriate)

5.0 Identification of Reference Standard and Control Substances

5.1 Controls

Controls for the ER antagonist protocol are as follows:

Vehicle control (dimethyl sulfoxide [DMSO]): 1% v/v dilution of DMSO (CASRN 67-68-5) diluted in tissue culture media

E2 control: 17 β -estradiol, 2.5×10^{-5} μ g/mL E2 in tissue culture media used as a base line negative control

TAM/E2 Control: Tamoxifen (TAM), CASRN 10540-29-1, 1.26 µg/mL, with 2.5×10^{-5} µg/mL E2 in tissue culture media used as a weak positive control

5.2 Reference Standard

Ral/E2 reference standard for range finder testing: Three concentrations (1.56×10^{-3} , 3.91×10^{-4} , and 9.77×10^{-5} µg/mL) of raloxifene HCl (Ral), CASRN 84449-90-1, plus a fixed concentration (2.5×10^{-5} µg/mL) of 17β-estradiol (E2), CASRN: 50-28-2, in duplicate wells.

Ral/E2 reference standard for comprehensive testing: A serial dilution of Ral plus a fixed concentration (2.5×10^{-5} µg/mL) of E2 consisting of nine concentrations of Ral/E2 in duplicate wells.

6.0 Overview of General Procedures for Antagonist Testing

All experimental procedures are to be carried out under aseptic conditions and all solutions, glassware, plastic ware, pipettes, etc., shall be sterile. All methods and procedures shall be documented in the study notebook.

Antagonist range finder testing is conducted on 96-well plates using three concentrations of Ral/E2 (1.56×10^{-3} , 3.91×10^{-4} , and 9.77×10^{-5} µg/mL Ral) with 2.50×10^{-5} µg/mL E2 in duplicate as the reference standard, with three replicate wells for the E2 and DMSO controls.

Comprehensive testing is conducted on 96-well plates using nine concentrations of Ral/E2 in duplicate as the reference standard (**Table 6-1**). Four replicate wells for the DMSO control, TAM/E2 and E2 controls are included on each plate.

Table 6-1 Concentrations of Ral/E2 Reference Standard Used for Comprehensive Testing

Raloxifene Concentrations ¹	E2 Concentrations
1.25×10^{-2}	2.5×10^{-5}
6.25×10^{-3}	2.5×10^{-5}
3.13×10^{-3}	2.5×10^{-5}
1.56×10^{-3}	2.5×10^{-5}
7.81×10^{-4}	2.5×10^{-5}
3.91×10^{-4}	2.5×10^{-5}
1.95×10^{-4}	2.5×10^{-5}
9.77×10^{-5}	2.5×10^{-5}
4.88×10^{-5}	2.5×10^{-5}

¹ Concentrations are presented in µg/mL.

Visual observations for cell viability are conducted for all experimental plates just prior to luminescence measurements, as outlined in **Section 11.4**.

Luminescence data, measured in relative light units (RLUs), is corrected for background luminescence by subtracting the mean RLU value of the vehicle control (DMSO) wells from the RLU measurements for each of the other wells of the 96-well plate. Data is then transferred into Excel® data management spreadsheets and GraphPad PRISM® statistical software, graphed, and evaluated for a positive or negative response as follows:

- A response is considered positive for antagonist activity when the average adjusted RLU for a given concentration is less than the mean RLU value minus three times the standard deviation for the E2 control.

- Any luminescence at or above this threshold is considered a negative response.

For substances that are positive at one or more concentrations, the concentration of test substance that causes a half-maximal response (the relative IC₅₀) is calculated using a Hill function analysis. The Hill function is a four-parameter logistic mathematical model relating the substance concentration to the response (typically following a sigmoidal curve) using the equation below

$$Y = \text{Bottom} + \frac{\text{Top} - \text{Bottom}}{1 + 10^{(\log \text{IC}_{50} - X)\text{HillSlope}}}$$

where Y = response (i.e., relative light units); X = the logarithm of concentration; Bottom = the minimum response; Top = the maximum response; log IC₅₀ = the logarithm of X as the response midway between Top and Bottom; and HillSlope describes the steepness of the curve. The model calculates the best fit for the Top, Bottom, HillSlope, and IC₅₀ parameters. See **Section 13.6.5** for more details.

Acceptance or rejection of a test is based on evaluation of reference standard and control results from each experiment conducted on a 96-well plate. Results for these controls are compared to historical results compiled in the historical database, as seen in **Section 16.0**.

6.1 Range Finder Testing

Antagonist range finding for coded substances consists of a seven-point 1:10 serial dilution using duplicate wells per concentration. Concentrations for comprehensive testing are selected based on the response observed in range finder testing. If necessary, a second range finder test can be conducted to clarify the optimal concentration range to test (see **Section 14.0**).

6.2 Comprehensive Testing

Comprehensive antagonist testing for coded substances consists of 11-point serial dilutions, with each concentration tested in triplicate wells of the 96-well plate. Three separate experiments are conducted for comprehensive testing on three separate days, except during Phases III and IV of the validation effort, in which comprehensive testing experiments are conducted once (see **Section 15.0**).

7.0 Materials for BG1Luc ER TA Antagonist Testing

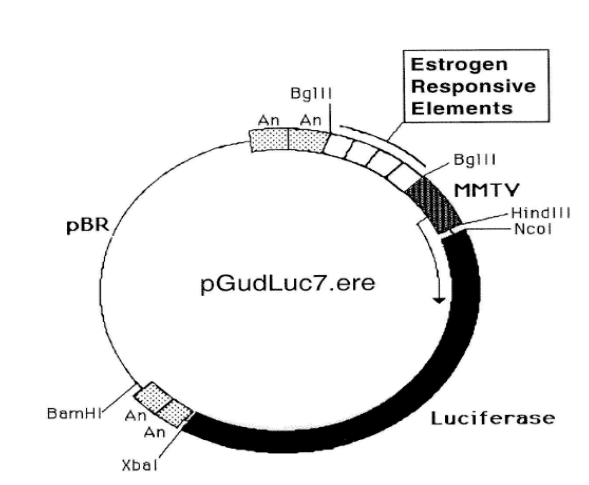
This section provides the materials needed to conduct BG1Luc ER TA testing, with associated brand names/vendors¹ in brackets.

7.1 BG1Luc4E2 Cells

Human ovarian cancer cell line stably transfected with a plasmid containing an estrogen response element (**Figure 7-1**). The BG1Luc4E2 cell line is available upon request from the University of California, Davis, California, USA, and is also available from Xenobiotic Detection Systems Inc., Durham, North Carolina, USA.

¹ Brand names and vendors should not be considered an endorsement by the U.S. Government or any member of the U.S. Government; such information is provided as an example.

Figure 7-1 pGudLuc7.ERE Plasmid



7.2 Technical Equipment:

All technical equipment may be obtained from Fisher Scientific International, Inc. (Liberty Lane Hampton, NH, USA 03842). Equivalent technical equipment from another commercial source can be used.

- Analytical balance (Cat. No. 01-910-320)
- Berthold Orion 1 Microplate Luminometer [Berthold CatNo.: Orion 1 MPL3] or equivalent and dedicated computer
- Biological safety hood, class II, and stand (Cat. No. 16-108-99)
- Centrifuge (low speed, tabletop with swinging bucket rotor) (Cat. No. 04-978-50 centrifuge, and 05-103B rotor)
- Combustion test kit (CO₂ monitoring) (Cat. No. 10-884-1)
- Drummond diaphragm pipetter (Cat. No. 13-681-15)
- Freezers, -20°C (Cat. No. 13-986-150), and -70°C (Cat. No. 13-990-86)
- Hand tally counter (Cat. No. 07905-6)
- Hemocytometer, cell counter (Cat. No. 02-671-5)
- Light microscope, inverted (Cat. No. 12-561-INV)
- Light microscope, upright (Cat. No. 12-561-3M)
- Liquid nitrogen flask (Cat. No. 11-675-92)
- Micropipetter, repeating (Cat. No. 21-380-9)
- Pipettors, air displacement, single channel (0.5 –10µL (Cat. No. 21-377-191), 2 –20 µL (Cat. No. 21-377-287), 20 – 200µL (Cat. No. 21-377-298), 200 - 1000 µL (Cat. No. 21-377-195))
- Refrigerator/freezer (Cat. No. 13-986-106A)
- Shaker for 96-well plates (Cat. No. 14-271-9)
- Sodium hydroxide (Cat. No. 5318-500)
- Sonicating water bath (Cat. No. 15-335-30)
- Tissue culture incubator with CO₂ and temperature control (Cat. No. 11-689-4)
- Vacuum pump with liquid trap (side arm Erlenmeyer) (Cat. No. 01-092-29)
- Vortex mixer (Cat. No. 12-814)

Equipment should be maintained and calibrated as per GLP guidelines and individual laboratory SOPs.

7.3 Reference Standard, Controls, and Tissue Culture Supplies

All tissue culture reagents must be labeled to indicate source, identity, storage conditions and expiration dates. Tissue culture solutions must be labeled to indicate concentration, stability (where known), and preparation and expiration dates.

Equivalent tissue culture media and sera from another commercial source can be used, but must first be tested as described in **Section 17.0** to determine suitability for use in this test method.

The following are the necessary tissue culture reagents and possible sources based on their use in the pre-validation studies:

- BackSeal-96/384, white adhesive bottom seal for 96-well and 384-well microplate [Perkin-Elmer, Cat. No. 6005199]
- 17 β -estradiol (CAS RN: 50-28-2) [Sigma-Aldrich, Cat. No. E8875]
- CellTiter-Glo[®] Luminescent Cell Viability Assay [Promega Cat. No. G7572]
- Cryovial, 2 mL (Corning Costar) [Fisher Scientific Cat. No. 03-374-21]
- Culture tube 13 x 100mm (case) [Thomas Scientific Cat. No.: 10009186R38]²
- Culture tube, 50 mL conical (Corning Costar) [Fisher Scientific Cat. No. 05-526C]
- DMSO, U.S.P. analytical grade. [Sigma-Aldrich, Cat. No. 34869-100ML]
- Dulbecco's Modification of Eagle's Medium (DMEM), containing 4.5 g/L glucose, with sodium pyruvate, without phenol red or L-glutamine [Mediatech/Cellgro, Cat. No. 17-205-CV]
- Fetal Bovine Serum [Mediatech/Cellgro Cat. No. MT 35-010-CV]
- Fetal Bovine Serum, charcoal/dextran treated, triple 0.1 μ m sterile filtered [Hyclone, Cat. No. SH30068.03]
- Tamoxifen (CASRN: 10540-29-1) [Sigma-Aldrich, Cat. No. T5648]
- Gentamycin Sulfate (G418), 50 mg/mL [Mediatech/Cellgro Cat. No. 30-234-CR]
- L-glutamine, 29.2 mg/mL [Cellgro, Cat. No. 25005-CI]
- Luciferase Assay System (10-Pack) [Promega Cat. No. E1501]
- Lysis Solution 5X [Promega, Cat. No. E1531]
- Penicillin/streptomycin solution, 5000 I.U. penicillin, 5000 μ g/mL streptomycin [Cellgro, Cat. No. 30-001-CI].
- Phosphate buffered saline (PBS, 1X) without calcium and magnesium [Cellgro, Cat. No. 21-040-CV]
- Pipettes, serological: 2.0 mL [Sigma-Aldrich, Cat. No. P1736], 5.0 mL [Sigma-Aldrich, Cat. No. P1986], 25 mL [Sigma-Aldrich, Cat. No. P2486]
- Raloxifene (CASRN 84449-90-1) [Sigma-Aldrich Cat. No. R1402]
- RPMI 1640 medium, containing L-glutamine [Mediatech, Cat. No. 10-040-CV]
- Tissue culture flasks (Corning-Costar): 25 cm² (T25) [Fisher Cat. No. 10-126-28]; 75 cm² (T75) [Fisher Cat. No. 10-126-37]; and 150 cm² (T150) [Fisher Cat. No. 10-126-34]
- Tissue culture plates (Corning-Costar): 96-well [Thomas Scientific Cat. No. 6916A05]
- Trypsin (10X), 2.5% in Hank's balanced salt solution (HBSS), without calcium and magnesium, without phenol red [Cellgro, Cat. No. 25-054-CI].

All reagent lot numbers and expiration dates must be recorded in the study notebook.

8.0 Preparation of Tissue Culture Media and Solutions

All tissue culture media and media supplements must be quality tested before use in experiments (see **Section 15.0**).

² If glass tubes cannot be obtained from Thomas Scientific, the preference is for flint glass, then lime glass, then borosilicate glass.

8.1 RPMI 1640 Growth Medium (RPMI)

RPMI 1640 is supplemented with 0.9% Pen-Strep and 8.0% FBS to make RPMI growth medium (RPMI).

Procedure for one 549 mL bottle:

1. Remove FBS from -70°C freezer, and Pen-Strep from -20°C freezer and allow to equilibrate to room temperature.
2. Add 44 mL of FBS and 5 mL Pen-Strep to the bottle of RPMI 1640.
3. Label RPMI bottle as indicated in **Section 7.3**

Store at 2-8 °C for no longer than six months or until the shortest expiration date of any media component.

8.2 Estrogen-Free DMEM Medium

DMEM is supplemented to contain 4.5% charcoal/dextran treated FBS, 1.9% L-glutamine, 0.9% Pen-Strep.

Procedure for one 539 mL bottle:

1. Remove charcoal/dextran treated FBS from -70°C freezer, and L-glutamine and Pen-Strep from -20°C freezer and allow to equilibrate to room temperature.
2. Add 24 mL of charcoal/dextran treated FBS, 10 mL L-glutamine, and 5 mL Pen-Strep to one 500 mL bottle of DMEM.
3. Label estrogen-free DMEM bottle as indicated in **Section 7.3**

Store at 2-8 °C for no longer than six months or until the shortest expiration date of any media component.

8.3 1X Trypsin Solution

1X Trypsin solution is prepared by dilution from a 10X premixed stock solution. The 10X stock solution should be stored in 10 mL aliquots in a -20°C freezer.

Procedure for making 100 mL of 1X trypsin:

1. Remove a 10mL aliquot of 10X trypsin from -20°C freezer and allow to equilibrate to room temperature.
2. Aliquot 1 mL Trypsin (10X) along with 9 mL of 1X PBS into ten 15 mL centrifuge tubes.
3. Label 1X trypsin aliquots as indicated in **Section 7.3**

1X Trypsin should be stored at -20°C.

8.4 1X Lysis Solution

Lysis solution is prepared by dilution from a 5X premixed stock solution. Both the 5X and 1X solutions can be repeatedly freeze-thawed.

The procedure for making 10 mL of 1X lysis solution:

1. Thaw the 5X Promega Lysis solution and allow it to reach room temperature.
2. Remove 2 mL of 5X solution and place it in a 15 mL conical centrifuge tube.
3. Add 8 mL of distilled, de-ionized water to the conical tube.

4. Cap and shake gently until solutions are mixed.

Store at -20°C for no longer than 1 year from receipt.

8.5 Reconstituted Luciferase Reagent

Luciferase reagent consists of two components, luciferase buffer and lyophilized luciferase substrate.

For long-term storage, unopened containers of the luciferase buffer and lyophilized luciferase substrate can be stored at -70°C for up to six months.

To reconstitute luciferase reagent:

1. Remove luciferase buffer and luciferase substrate from -70°C freezer and allow them to equilibrate to room temperature.
2. Add 10 mL of luciferase buffer solution to luciferase substrate container and swirl or vortex to mix, the luciferase substrate should readily go into solution.
3. Luciferase substrate should readily go into solution.
4. After solutions are mixed aliquot to a 15mL centrifuge tube.
5. Store complete solution at -20°C.

Reconstituted luciferase reagent is stable for 1 month at -20°C.

9.0 Overview of Propagation and Experimental Plating of BG1Luc4E2 Cells

BG1Luc4E2 cells are based on a continuous ovarian carcinoma cell line (BG-1 cells) that endogenously express ER α and ER β and have been stably transfected with an ER responsive reporter gene (*luc*).

Although the cell line has proved to be stable over long-term passage *in vitro*, careful handling and the use of quality cell culture materials is required to maintain the stability and integrity of the cell line.

Procedures specified in the Guidance on Good Cell Culture Practice (Coecke 2005) should be followed to assure the quality of all materials and methods in order to maintain the integrity, validity, and reproducibility of any work conducted.

BG1Luc4E2 cells are stored in liquid nitrogen in 2 mL cryovials. BG-1 cells are grown as a monolayer in tissue culture flasks in a dedicated tissue culture incubator at 37°C \pm 1°C, 90% \pm 5% humidity, and 5.0% \pm 1% CO₂/air. The cells should be examined on a daily basis during working days under an inverted phase contrast microscope, and any changes in morphology and adhesive properties must be noted in the study notebook.

Two T150 flasks containing cells at 80% to 90% confluence will usually yield a sufficient number of cells to fill three 96-well plates for use in experiments.

9.1 Procedures for Thawing Cells and Establishing Tissue Cultures

Warm all tissue culture media and solutions to room temperature by placing them under the tissue culture hood several hours before use.

All tissue culture media, media supplements, and tissue culture plasticware must be quality tested before use in experiments (**Section 17.0**).

9.1.1 Thawing Cells

1. Remove a cryovial of frozen BG-1 cells from the liquid nitrogen flask.
2. Facilitate rapid thawing by loosening the top slightly (do not remove top) to release trapped gasses and retightening it. Roll vial between palms.
3. Use a micropipette to transfer cells to a 50 mL conical centrifuge tube.
4. Rinse cryovial twice with 1X PBS and add PBS rinse material to the conical tube.
5. Add 20 mL of RPMI to the conical tube.
6. Centrifuge at 1000 x g for eight min. If a pellet of cells has not formed, centrifuge for an additional 5 minutes.
7. Aspirate media from pellet and re-suspend it in 5 mL RPMI, drawing the pellet repeatedly through a 1.0 mL serological pipette to break up any clumps of cells.
8. Transfer cells to a T25 flask, place them in an incubator (see conditions in **Section 9.0**) and grow to 80% to 90% confluence (approximately 48 to 72 hrs).

9.1.2 Establishing Tissue Cultures

Once cells have reached 80% to 90% confluence, transfer the cells to a T75 flask by performing, for example, the following steps:

1. Remove the T25 flask from the incubator.
2. Aspirate the RPMI, then add 5 mL 1X PBS, making sure that the cells are coated with PBS.
3. Aspirate 1X PBS, then add 1 to 2 mL 1X trypsin to the T25 flask, gently swirling the flask to coat all cells with the trypsin.
4. Place the flask in an incubator (see conditions in **Section 9.0**) for 5 to 10 min.
5. Detach cells by hitting the side of the flask sharply against the palm or heel of the hand.
6. Confirm cell detachment by examination under an inverted microscope. If cells have not detached, return the flask to the incubator for an additional 2 minutes, then hit the flask again.
7. After cells have detached, add 5 mL PBS, and transfer the suspended cells to a 50 mL centrifuge tube. Wash the flask one additional time with 5 mL PBS.
8. Immediately add 20 mL RPMI to the conical tube to inhibit further cellular digestion by residual trypsin.
9. Pellet the cells by centrifugation, as described in **Section 9.1.1**, and re-suspend the cells in 10 mL RPMI medium.
10. Draw the pellet repeatedly through a 25 mL serological pipette to break up clumps of cells
11. Transfer cells to a T75 flask, then place the flask in an incubator (see conditions in **Section 9.0**) and grow to 80% to 90% confluence (approximately 48 to 72 hrs).

When cells have reached 80% to 90% confluency, transfer them into a T150 flask by performing, for example, the following steps:

12. Remove the T75 flask from the incubator, aspirate the old media and add 5 mL 1X PBS.
13. Aspirate 1X PBS, add 2 mL of 1X trypsin to the flask, and place it in an incubator (see conditions in **Section 9.0**) for 5 to 10 min.

14. Repeat **steps 5** through **11** in **Section 9.1.2**, re-suspending the pellet in 20 mL of RPMI.
15. Transfer cells to a T150 flask and place it in the incubator (see conditions in **Section 9.0**) and grow to 80% to 90% confluence (approximately 48 to 72 hrs).
16. Remove the T150 flask from the incubator.
17. Aspirate the RPMI and add 5 mL 1X PBS.
18. Aspirate 1X PBS and add 3 mL 1X trypsin to the T150 flask, making sure that the cells are coated with the trypsin.
19. Incubate cells in an incubator (see conditions in **Section 9.0**) for 5 to 10 min.
20. Detach cells by hitting the side of the flask sharply against the palm or heel of the hand.
21. Confirm cell detachment by examination under an inverted microscope. If cells have not detached, return the flask to the incubator for an additional 2 minutes, then hit the flask again.
22. After cells have detached, add 5mL 1X PBS and transfer the suspended cells from the T150 flask to a 50 mL conical tube. Add an additional 5 mL PBS to the flask, then transfer to the 50 mL conical tube.
23. Immediately add 20 mL RPMI to the conical tube to inhibit further cellular digestion by residual trypsin.
24. Centrifuge at 1000 x g for eight minutes. If a pellet of cells has not formed, centrifuge for an additional 5 minutes.
25. Aspirate the media from the pellet and re-suspend it in 40 mL RPMI, drawing the pellet repeatedly through a 25 mL serological pipette to break up any clumps of cells.
26. Transfer 20 mL of cell suspension to each of two T150 flasks, place them in an incubator (see conditions in **Section 9.0**) and grow to 80% to 90% confluence (approximately 48 to 72 hrs).

9.2 Ongoing Tissue Culture Maintenance, Conditioning in Estrogen-free Medium, and Plating Cells for Experimentation

The following procedure is used to condition the BG1Luc4E2 cells to an estrogen-free environment prior to plating the cells in 96-well plates for analysis of estrogen dependent induction of luciferase activity.

To start the tissue culture maintenance and estrogen-free conditioning, split the two T150 culture flasks into four T150 flasks. Two of these flasks will be used for continuing tissue culture and will use the RPMI media mentioned above. The other two flasks will be cultured in estrogen-free DMEM for experimental use. Extra care must be taken to avoid contaminating the estrogen-free cells with RPMI.

1. Remove both T150 flasks from the incubator.
2. Aspirate the medium and rinse the cells with 5 mL 1X PBS.
3. Aspirate 1X PBS, then add 3 mL 1X trypsin to the flasks, gently swirling the flask to coat all cells with the trypsin.
4. Incubate cells in the incubator (see conditions in **Section 9.0**) for 5 to 10 min.
5. Detach cells by hitting the side of the flask sharply against the palm or heel of the hand.
6. Confirm cell detachment by examination under an inverted microscope. If cells have not detached, return the flask to the incubator for an additional 2 minutes, then hit the flask again.
7. After cells have detached, add 5 mL 1X PBS to the first T150 flask and transfer the suspended cells to the second T150 flask.

8. Transfer the contents of both flasks to a 50 mL conical tube. Repeat step 7 with an additional 5 mL 1X PBS and transfer to the 50 mL conical tube.
9. Immediately add 20 mL estrogen-free DMEM to the 50 mL conical tube to inhibit further cellular digestion by residual trypsin.
10. Centrifuge at 1000 x g for eight minutes. If a pellet of cells has not formed, centrifuge for an additional 5 minutes.
11. Aspirate media from pellet and re-suspend it in 4 mL estrogen-free DMEM, drawing the pellet repeatedly through a 1 mL serological pipette to break up clumps of cells.

At this point, cells are ready to be divided into the ongoing tissue culture and estrogen-free conditioning groups.

9.2.1 Ongoing Tissue Culture Maintenance

1. Add 20 mL RPMI to two T150 flasks.
2. Add 220 µL G418 to the RPMI in the T150 flasks
3. Add 1 mL of cell suspension from **Section 9.2 step 11** to each flask.
4. Place T150 flasks in tissue culture incubator (see conditions in **Section 9.0**) and grow to 80% to 90% confluence (approximately 48 to 72 hrs).
5. Tissue culture medium may need to be changed 24 hours after addition of G418 to remove cells that have died because they do not express reporter plasmid.
6. G418 does not need to be added to the flasks a second time.
7. Repeat **Section 9.2 steps 1-11** for ongoing tissue culture maintenance.

9.2.2 Conditioning in Estrogen-free Medium

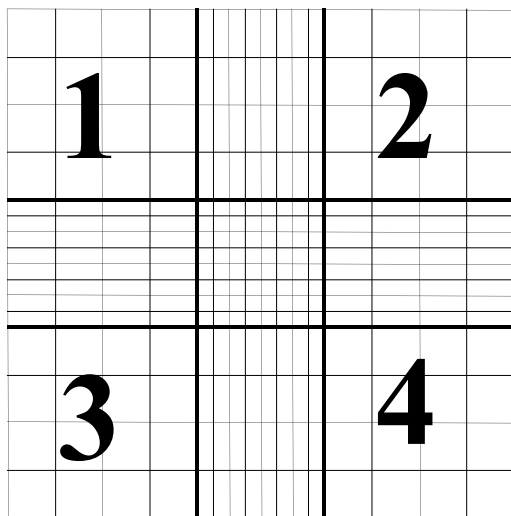
1. Add 20 mL estrogen-free DMEM to two T150 flasks.
2. Add 150 µL G418 to the estrogen-free DMEM in the T150 flasks.
3. Add 1 mL of cell suspension from **Section 9.2 step 11** to each flask.
4. Tissue culture medium may need to be changed 24 hours after addition of G418 to remove cells that have died because they do not express reporter plasmid.
5. G418 does not need to be added to the flasks a second time.
6. Place the T150 flasks in the incubator (see conditions in **Section 9.0**) and grow to 80% to 90% confluence (approximately 48 to 72 hrs).

9.2.3 Plating Cells Grown in Estrogen-free DMEM for Experimentation

1. Remove the T150 flasks that have been conditioned in estrogen-free DMEM for 48 to 72 hours from the incubator.
2. Aspirate the medium, then rinse the cells with 5 mL 1X PBS.
3. Aspirate 1X PBS, then add 3 mL 1X trypsin to the flasks, gently swirling the flask to coat all cells with the trypsin.
4. Place the flasks in an incubator (see conditions in **Section 9.0**) for 5 to 10 min.
5. Detach cells by hitting the side of the flask sharply against the palm or the heel of the hand.

6. Confirm cell detachment by examination under an inverted microscope. If cells have not detached, return the flask to the incubator for 2 additional minutes, then hit the flask again.
7. After cells have detached, add 5 mL 1X PBS and transfer the suspended cells from the T150 flask to a 50 mL conical tube. Add an additional 5 mL PBS to the flask, then transfer to the 50 mL conical tube.
8. Immediately add 20 mL estrogen-free DMEM to each conical tube to inhibit further cellular digestion by residual trypsin.
9. Centrifuge at 1000 x g for eight minutes. If a pellet of cells has not formed, centrifuge for an additional 5 minutes.
10. Aspirate off the media from the pellet and re-suspend it in 20 mL DMEM, drawing the pellet repeatedly through a 25 mL serological pipette to break up any clumps of cells.
11. Pipette 15 μ L of the cell suspension into the “v” shaped slot on the hemocytometer. Ensure that the solution covers the entire surface area of the hemocytometer grid, and allow cells to settle before counting.
12. Using 100x magnification, view the counting grid.
13. The counting grid on the hemocytometer consists of nine sections, four of which are counted (upper left, upper right, lower left, and lower right, see **Figure 9-1**). Each section counted consists of four by four grids. Starting at the top left and moving clockwise, count all cells in each of the four by four grids. Some cells will be touching the outside borders of the square, but only count those that touch the top and right borders of the square. This value is then used in the calculation below to get to the desired concentration of 200,000 cells/mL.

Figure 9-1 Hemocytometer Counting Grid



The volume of each square is 10^{-4} mL, therefore:

Cells/mL = (average number per grid) x 10^{-4} mL. x 1/(starting dilution).

Starting dilution: 20mL (for T150 flasks)

Harvested cells for a T150 flask are suspended in 20 mL of estrogen-free DMEM and sampled for determination of concentration of cells/mL.

Example Calculation:

- Grids 1, 2, 3, and 4 are counted and provide the following data:

- 50, 51, 49, and 50: average number of cells per grid is equal to 50.

Cells/mL = 50 cells per grid $\div 10^{-4}$ mL volume of grid = 50×10^4 cells/mL (or 500,000 cells/mL)

Total # of Cells Harvested = 500,000 cells/mL \times 20 mL

Desired Concentration (or Concentration_{Final}) = 200,000 cells/mL

Formula: (Concentration_{Final} \times Volume_{Final} = Concentration_{Initial} \times Volume_{Initial})

Concentration_{Final} = 200,000 cells/mL

Concentration_{Initial} = 500,000 cells/mL

Volume_{Initial} = 20 mL

Volume_{Final} – to be solved for.

Therefore: 200,000 cells/mL \times Volume_{Final} = 500,000 cells/mL \times 20 mL

Solving for Volume_{Final} we find = 50 mL

Therefore, add 30 mL of estrogen-free DMEM to the cell suspension for a total volume of 50 mL, which will yield the desired concentration of 200,000 cells/mL for plating.

14. This dilution scheme will give a concentration of 200,000 cells/mL. 200 mL of this cell suspension is used for each well of a 96-well plate (i.e., 40,000 cells per well).
15. Remove a 96-well plate from its sterile packaging. Use a repeater pipetter to pipette 200 μ L of cell suspension into each well to be used for the testing of coded substances, reference standard and controls (**note**: add 200 μ L of estrogen-free DMEM only to any wells not being used for testing).
16. Incubate plate(s) in an incubator (see conditions in **Section 9.0**) for a minimum of 24 hours, but no longer than 48 hours before dosing.

Two T150 flasks containing cells at 80% to 90% confluence will typically yield sufficient cells to fill four 96-well plates (not including the perimeter wells).

10.0 Preparation of Test Substances

The solvent used for dissolution of test substances is 100% DMSO. All test substances should be allowed to equilibrate to room temperature before being dissolved and diluted. Test substance solutions (except for reference standards and controls) should not be prepared in bulk for use in subsequent tests. Test substances are to be used within 24 hours of preparation. Solutions should not have noticeable precipitate or cloudiness.

All information on weighing, solubility testing, and calculation of final concentrations for test substances, reference standards and controls is to be recorded in the study notebook.

10.1 Determination of Test Substance Solubility

1. Prepare a 200 mg/mL solution of the test substance in 100% DMSO in a 4 mL conical tube.
2. Vortex to mix.
3. If the test substance does not dissolve at 200 mg/mL, prepare a 20 mg/mL solution and vortex as above.

4. If the test substance does not dissolve at 20 mg/mL solution, prepare a 2 mg/mL solution in a 4 mL conical tube and vortex as above.
5. If the test substance does not dissolve at 2 mg/mL, prepare a 0.2 mg/mL solution in a 4 mL conical tube and vortex as above.
6. Continue testing, using 1/10 less substance in each subsequent attempt until test substance is solubilized in DMSO.

Once the test substance has fully dissolved in 100% DMSO, the test substance is ready to be used for BG1Luc ER TA testing.

11.0 Preparation of Reference Standard, Control and Test Substance Stock Solutions for Range Finder and Comprehensive Testing

All information on preparation of test substances, reference standards and controls is to be recorded in the study notebook.

11.1 Preparation of Ral/E2 Stock Solutions

E2 and raloxifene stocks are prepared separately and then combined into Ral/E2 stocks, which are then used to prepare dosing solutions in **Section 12**.

11.1.1 E2 Stock Solution

The final concentration of the E2 stock solution is 5.0×10^{-3} µg/mL. Prepare the E2 stock as shown in **Table 11-1**.

Table 11-1 Preparation of E2 Stock Solution

Step #	Action	DMSO	E2 Concentration
1	Make a 10 mg/mL stock solution in 100% DMSO in a 4mL vial.	-	10 mg/mL
2	Transfer 10 µL E2 solution from Step #1 to a new 4 mL vial.	Add 990 µL of 100% DMSO. Vortex to mix.	100 µg/mL
3	Transfer 10 µL E2 solution from Step #2 to a new 4mL vial.	Add 990 µL of 100% DMSO. Vortex to mix.	1 µg/mL
4	Transfer 100 µL E2 solution from Step #3 to a new glass container large enough to hold 15 mL.	Add 9.90 mL of 100% DMSO. Vortex to mix.	1.0×10^{-2} µg/mL
5	Transfer 5 mL E2 solution from Step #4 to a new glass container large enough to hold 15 mL	Add 5 mL of 100% DMSO. Vortex to mix.	5.0×10^{-3} µg/mL

11.1.2 Raloxifene Stock Solution

Prepare a 2.5 µg/mL raloxifene working stock solution as shown in **Table 11-2**.

Table 11-2 Preparation of Raloxifene Stock Solution

Step #	Action	DMSO	Raloxifene Concentration
1	Make a 10 mg/mL solution of raloxifene in a 4 mL glass vial.	-	1.0×10^4 µg/mL
2	Transfer 10 µL raloxifene solution from Step #1 to a new 4 mL vial.	Add 990 µL of 100% DMSO. Vortex to mix.	100 µg/mL
3	Transfer 150 µL raloxifene solution from Step #2 to a new 4 mL vial.	Add 2.850 mL of 100% DMSO. Vortex to mix.	5 µg/mL
4	Transfer 1.5 mL raloxifene solution from Step #3 to a new 13 mm test tube.	Add 1.5 mL of 100% DMSO. Vortex to mix.	2.5 µg/mL

11.2 Ral/E2 Range Finder Testing Stock

11.2.1 Raloxifene Dilutions

Number three 4 mL vials with the numbers 1 to 3 and use the raloxifene solution prepared in **Section 11.1.2** to make raloxifene dilutions as shown **Table 11-3**.

Table 11-3 Preparation of Raloxifene Dilutions for Range Finder Testing

Step #	Action	DMSO	Raloxifene Concentration
1	Transfer 250 µL of the 2.5 µg/mL raloxifene working stock solution to a 4 mL tube	Add 750 µL of 100% DMSO and vortex	6.25×10^{-1} µg/mL
2	Transfer 500 µL of the 6.25×10^{-1} µg/mL raloxifene solution to a 4 mL tube	Add 500 µL of 100% DMSO and vortex	3.13×10^{-1} µg/mL
3	Transfer 250 µL of the 3.13×10^{-1} µg/mL raloxifene solution to a 4 mL tube	Add 750 µL of 100% DMSO and vortex	7.81×10^{-2} µg/mL
4	Transfer 125 µL of the 7.81×10^{-2} µg/mL raloxifene solution to a 4 mL tube	Add 375 µL of 100% DMSO and vortex	1.95×10^{-2} µg/mL

11.2.2 Preparation of Ral/E2 Range Finder Working Stocks

Label three 4 mL conical tubes with numbers 1 through 3 and add 500 µL of the 5×10^{-3} µg/mL E2 solution prepared in **Section 11.1.1** to each tube. Add 500 µL of the 3.13×10^{-1} , 7.81×10^{-2} , and 1.95×10^{-2} µg/mL raloxifene solutions prepared in **Section 11.2.1** to tubes 1, 2, and 3 respectively. Vortex each tube to mix. The final concentrations for raloxifene and E2 are listed in **Table 11-4**.

Table 11-4 Concentrations of Raloxifene and E2 in the Ral/E2 Range Finder Working Stocks

Tube #	Raloxifene (µg/mL)	E2 (µg/mL)
1	1.56×10^{-1}	2.5×10^{-3}
2	3.91×10^{-2}	2.5×10^{-3}
3	9.77×10^{-3}	2.5×10^{-3}

11.3 Ral/E2 Comprehensive Testing Stock

11.3.1 Raloxifene Dilutions

Use the raloxifene solution prepared in **Section 11.1.2** to make a nine-point serial dilution of raloxifene as shown **Table 11-5**.

Table 11-5 Preparation of Raloxifene Dilutions for Comprehensive Testing

Step #	Action	DMSO	Discard	Raloxifene Concentration
1	Transfer 500 µL of the raloxifene working stock solution to a new 4 mL vial.	-	-	2.5 µg/mL
2	Transfer 500 µL of the raloxifene working stock solution to a new 4 mL vial.	Add 500 µL of 100% DMSO. Vortex to mix.	-	1.25 µg/mL
3	Transfer 500 µL raloxifene solution from Step #2 to a new 4 mL vial.	Add 500 µL of 100% DMSO. Vortex to mix.	-	6.25×10^{-1} µg/mL
4	Transfer 500 µL raloxifene solution from Step #3 to a new 4 mL vial.	Add 500 µL of 100% DMSO. Vortex to mix.	-	3.13×10^{-1} µg/mL
5	Transfer 500 µL raloxifene solution from Step #4 to a new 4 mL vial.	Add 500 µL of 100% DMSO. Vortex to mix.	-	1.56×10^{-1} µg/mL
6	Transfer 500 µL raloxifene solution from Step #5 to a new 4 mL vial.	Add 500 µL of 100% DMSO. Vortex to mix.	-	7.81×10^{-2} µg/mL
7	Transfer 500 µL raloxifene solution from Step #6 to a new 4 mL vial.	Add 500 µL of 100% DMSO. Vortex to mix.	-	3.91×10^{-2} µg/mL
8	Transfer 500 µL raloxifene solution from Step #7 to a new 4 mL vial.	Add 500 µL of 100% DMSO. Vortex to mix.		1.95×10^{-2} µg/mL
9	Transfer 500 µL raloxifene solution from Step #8 to a new 4 mL vial.	Add 500 µL of 100% DMSO. Vortex to mix.	Discard 500 µL from Tube #9	9.77×10^{-3} µg/mL

11.3.2 Preparation of Ral/E2 Comprehensive Testing Working Stocks:

Add 500 µL of the 5×10^{-3} µg/mL E2 solution prepared in **Section 11.1.1** to each of the 9 raloxifene dilution vials (including the working stock solution in Tube #1). Vortex each tube to mix. The final concentrations for raloxifene and E2 are listed in **Table 11-6**.

Table 11-6 Concentrations of Raloxifene and E2 in the Ral/E2 Working Stocks

Tube #	Raloxifene (µg/mL)	E2 (µg/mL)
1	1.25	2.5×10^{-3}
2	6.25×10^{-1}	2.5×10^{-3}
3	3.13×10^{-1}	2.5×10^{-3}
4	1.56×10^{-1}	2.5×10^{-3}
5	7.81×10^{-2}	2.5×10^{-3}
6	3.91×10^{-2}	2.5×10^{-3}
7	1.95×10^{-2}	2.5×10^{-3}
8	9.77×10^{-3}	2.5×10^{-3}
9	4.88×10^{-3}	2.5×10^{-3}

11.4 TAM/E2 Stock Solution

To prepare the TAM/E2 stock solution, proceed as follows:

1. Prepare 1 mL of 2.52×10^{-1} µg/mL tamoxifen
2. Add 1 mL of 5×10^{-3} µg/mL E2 (prepared as in **Section 11.1.1**) to the 2.52×10^{-1} µg/mL tamoxifen. This will make a working solution of 1.26×10^{-1} µg/mL tamoxifen with 2.5×10^{-3} µg/mL E2.

12.0 Preparation of Reference Standard, Control and Test Substance Dosing Solutions for Range Finder and Comprehensive Testing**12.1 Preparation of Reference Standard and Control Dosing Solutions - Range Finder Testing**

Range finder testing is conducted on 96-well plates using three concentrations of Ral/E2 in duplicate as the reference standard. Three replicate wells for the DMSO, and E2 controls are included on each plate.

All “dosing solutions” of test substance concentrations are to be expressed as µg/mL in the study notebook and in all laboratory reports. Dosing solutions are to be used within 24 hours of preparation.

12.1.1 Preparation of Ral/E2 Reference Standard Range Finder Dosing Solutions

1. Label three 13 mm glass tubes with the numbers 1 to 3.
2. Add 6 µL Ral/E2 stock from tube #1 (**Section 11.2.2**) to 13 mm glass test tube #1.
3. Add 6 µL of Ral/E2 stock from tube #2 from **Section 11.2.2** to the 13 mm glass test tube labeled #2. Repeat for tube #3.
4. Add 600 µL of estrogen-free DMEM to each tube and vortex.

12.1.2 Preparation of DMSO Control Range Finder Dosing Solution

1. Add 8 µL of 100% DMSO to a 13 mm glass test tube.
2. Add 800 µL of estrogen-free DMEM to each tube and vortex.

12.1.3 Preparation of E2 Control Range Finder Dosing Solution

1. Add 4 µL of the E2 stock from **Section 11.1.1** to a 13 mm glass test tube.
2. Add 4 µL of 100% DMSO to the tube.
3. Add 800 µL of estrogen-free DMEM to the tube and vortex to mix.

12.2 Preparation of Test Substance Dosing Solutions for Range Finder Testing

Range finder experiments are used to determine the concentrations of test substance to be used during comprehensive testing. Antagonist range finding for coded substances consists of seven-point 1:10 serial dilutions in duplicate.

To prepare test substance dosing solutions:

1. Label two sets of seven glass 13 mm test tubes with the numbers 1 through 7 and place them in a test tube rack. Perform a serial dilution of test substance as shown in **Table 12-1** using one set of tubes.

Table 12-1 Preparation of Test Substance Serial Dilution for Range Finder Testing

Tube #	100% DMSO	Test Substance ¹	Final Volume
1	-	100 µL of test substance solution from Section 10.1	100 µL
2	90 µL	10 µL of test substance solution from Section 10.1	100 µL
3	90 µL	10 µL from Tube #2	100 µL
4	90 µL	10 µL from Tube #3	100 µL
5	90 µL	10 µL from Tube #4	100 µL
6	90 µL	10 µL from Tube #5	100 µL
7	90 µL	10 µL from Tube #6	100 µL

¹ Vortex tubes #2 through 6 before removing test substance/DMSO solution to place in the next tube in the series.

2. Transfer test substance/DMSO solutions to the second set of labeled tubes and add E2 as shown in **Table 12-2**.

Table 12-2 Addition of E2 to Test Substance Serial Dilution for Range Finder Testing

Tube Number	Test Substance	E2	Estrogen-free DMEM ³	Final Volume
1	Transfer 4 μ L of test substance from Tube #1 in Section 12.2 step 1 to a new tube	Add 4 μ L of the 5×10^{-3} μ g/mL E2 solution prepared in Section 11.1.1 . Vortex to mix.	800 μ L	808 μ L
2	Transfer 4 μ L of test substance from Tube #2 to a new tube	Add 4 μ L of the 5×10^{-3} μ g/mL E2 solution prepared in Section 11.1.1 . Vortex to mix.	800 μ L	808 μ L
3	Transfer 4 μ L of test substance from Tube #3 to a new tube	Add 4 μ L of the 5×10^{-3} μ g/mL E2 solution prepared in Section 11.1.1 . Vortex to mix.	800 μ L	808 μ L
4	Transfer 4 μ L of test substance from Tube #4 to a new tube	Add 4 μ L of the 5×10^{-3} μ g/mL E2 solution prepared in Section 11.1.1 . Vortex to mix.	800 μ L	808 μ L
5	Transfer 4 μ L of test substance from Tube #5 to a new tube	Add 4 μ L of the 5×10^{-3} μ g/mL E2 solution prepared in Section 11.1.1 . Vortex to mix.	800 μ L	808 μ L
6	Transfer 4 μ L of test substance from Tube #6 to a new tube	Add 4 μ L of the 5×10^{-3} μ g/mL E2 solution prepared in Section 11.1.1 . Vortex to mix.	800 μ L	808 μ L
7	Transfer 4 μ L of test substance from Tube #7 to a new tube	Add 4 μ L of the 5×10^{-3} μ g/mL E2 solution prepared in Section 11.1.1 . Vortex to mix.	800 μ L	808 μ L

Determination of whether a substance is positive in range finder testing and selection of starting concentrations for comprehensive testing will be discussed in **Section 14.0**.

12.3 Preparation of Reference Standard and Control Dosing Solutions for Comprehensive Testing

Comprehensive testing is conducted on 96-well plates using nine concentrations of Ral/E2 in duplicate as the reference standard. Four replicate wells for the DMSO, E2 and TAM/E2 controls are included on each plate.

All “dosing solutions” of test substance concentrations are to be expressed as μ g/mL in the study notebook and in all laboratory reports.

Store dosing solutions at room temperature. Use within 24 hours of preparation.

12.3.1 Preparation of Ral/E2 Reference Standard Dosing Solutions for Comprehensive Testing

In preparation for making Ral/E2 1:2 serial dilutions, label two sets of nine glass 13 mm test tubes with the numbers 1 through 9 and place them in a test tube rack. Tube number 1 will contain the highest concentration of raloxifene (**Table 12-3**).

Table 12-3 Preparation of Ral/E2 Reference Standard Dosing Solution for Comprehensive Testing

Tube Number	Ral/E2 Stock	Estrogen-free DMEM	Final Volume
1	6 μ L of Tube #1 from Section 11.3.2	600 μ L	606 μ L
2	6 μ L of Tube #2 from Section 11.3.2	600 μ L	606 μ L
3	6 μ L of Tube #3 from Section 11.3.2	600 μ L	606 μ L
4	6 μ L of Tube #4 from Section 11.3.2	600 μ L	606 μ L
5	6 μ L of Tube #5 from Section 11.3.2	600 μ L	606 μ L
6	6 μ L of Tube #6 from Section 11.3.2	600 μ L	606 μ L

Tube Number	Ral/E2 Stock	Estrogen-free DMEM	Final Volume
7	6 µL of Tube #7 from Section 11.3.2	600 µL	606 µL
8	6 µL of Tube #8 from Section 11.3.2	600 µL	606 µL
9	6 µL of Tube #9 from Section 11.3.2	600 µL	606 µL

12.3.2 Preparation of DMSO Control Comprehensive Testing Dosing Solution

1. Add 10 µL of 100% DMSO to a 13 mm glass test tube.
2. Add 1000 µL of estrogen-free DMEM to the tube and vortex to mix.

12.3.3 Preparation of E2 Control Comprehensive Testing Dosing Solution

1. Add 5 µL of the E2 stock from **Section 11.1.1** to a 13 mm glass test tube.
2. Add 5 µL of 100% DMSO to the tube.
3. Add 1000 µL of estrogen-free DMEM to the tube and vortex to mix.

12.3.4 Preparation of TAM/E2 Control Comprehensive Dosing Solution

1. Add 10 µL of TAM/E2 from **Section 11.4** to a 13 mm glass test tube.
2. Add 1000 µL of estrogen-free DMEM to the tube and vortex to mix.

12.4 Preparation of Test Substance Dosing Solutions for Comprehensive Testing

Comprehensive testing experiments are used to determine whether a substance possesses ER antagonist activity in the BG1Luc ER TA test method. Antagonist comprehensive testing for coded substances consists of either an 11-point 1:2 serial dilution, or an 11-point 1:5 serial dilution with each concentration tested in triplicate wells of the 96-well plate.

12.4.1 Preparation of Test Substance 1:2 Serial Dilutions for Comprehensive Testing

Start the 11-point serial dilution according to criteria in **Section 14.0**.

To make test substance 1:2 serial dilutions for comprehensive testing:

1. Label eleven 4 mL conical tubes with numbers 1 through 11 and place them in a tube rack.
2. Label eleven 13 mm glass test tubes with numbers 1 through 11, place them in a tube rack, and add 800 µL of estrogen-free DMEM to each tube.

Prepare dilution of test substance as shown in **Table 12-4**.

Table 12-4 Preparation of Test Substance 1:2 Serial Dilutions for Comprehensive Testing

Tube Number	100% DMSO	Test Substance ¹	Discard	E2 Testing Stock	Estrogen-free DMEM ²	Final Volume
1	-	4 µL of test substance solution from Section 10.2.4 step 1	-	4 µL	800 µL	808 µL
2	4 µL	4 µL of test substance solution from Section 10.2.4 step 1	-	4 µL	800 µL	808 µL
3	4 µL	4 µL from Tube #2	-	4 µL	800 µL	808 µL
4	4 µL	4 µL from Tube #3	-	4 µL	800 µL	808 µL
5	4 µL	4 µL from Tube #4	-	4 µL	800 µL	808 µL
6	4 µL	4 µL from Tube #5	-	4 µL	800 µL	808 µL

Tube Number	100% DMSO	Test Substance ¹	Discard	E2 Testing Stock	Estrogen-free DMEM ²	Final Volume
7	4 µL	4 µL from Tube #6	-	4 µL	800 µL	808 µL
8	4 µL	4 µL from Tube #7	-	4 µL	800 µL	808 µL
9	4 µL	4 µL from Tube #8	-	4 µL	800 µL	808 µL
10	4 µL	4 µL from Tube #9	-	4 µL	800 µL	808 µL
11	4 µL	4 µL from Tube #10	4 µL	4 µL	800 µL	808 µL

¹ Vortex tubes #2 through 10 before removing test substance/DMSO solution to place in the next tube in the series.

² Vortex all tubes to mix media, test substance, and E2.

12.4.2 Preparation of Test Substance 1:5 Serial Dilutions for Comprehensive Testing

Start the 11-point serial dilution according to criteria in **Section 14.0**.

To make test substance 1:5 serial dilutions for comprehensive testing:

1. Label eleven 4 mL conical tubes with numbers 1 through 11 and place them in a tube rack.
2. Label eleven 13 mm glass test tubes with numbers 1 through 11, place them in a tube rack, and add 800 µL of estrogen-free DMEM to each tube.

Prepare dilution of test substance as shown in **Table 12-5**.

Table 12-5 Preparation of Test Substance 1:5 Dilutions for Comprehensive Testing

Tube Number	100% DMSO	Test Substance ¹	Discard	E2 Testing Stock	Estrogen-free DMEM ²	Final Volume
1	-	4 µL of test substance solution from Section 10.2.4 step 1	-	4 µL	800 µL	808 µL
2	16 µL	4 µL of test substance solution from Section 10.2.4 step 1	-	4 µL	800 µL	808 µL
3	16 µL	4 µL from Tube #2	-	4 µL	800 µL	808 µL
4	16 µL	4 µL from Tube #3	-	4 µL	800 µL	808 µL
5	16 µL	4 µL from Tube #4	-	4 µL	800 µL	808 µL
6	16 µL	4 µL from Tube #5	-	4 µL	800 µL	808 µL
7	16 µL	4 µL from Tube #6	-	4 µL	800 µL	808 µL
8	16 µL	4 µL from Tube #7	-	4 µL	800 µL	808 µL
9	16 µL	4 µL from Tube #8	-	4 µL	800 µL	808 µL
10	16 µL	4 µL from Tube #9	-	4 µL	800 µL	808 µL
11	16 µL	4 µL from Tube #10	20 µL	4 µL	800 µL	808 µL

¹ Vortex tubes #2 through 10 before removing test substance/DMSO solution to place in the next tube in the series.

² Vortex all tubes to mix media, test substance, and E2.

13.0 General Procedures for the Testing of Coded Substances

Range finder experiments are used to determine the concentrations of test substance to be used during comprehensive testing. Comprehensive testing experiments are used to determine whether a substance possesses ER antagonist activity in the BG1Luc ER TA test method.

General procedures for range finder and comprehensive testing are nearly identical. For specific details (such as plate layout) of range finder testing see **Section 14.0**. For specific details of comprehensive testing, see **Section 15.0**.

13.1 Application of Reference Standard, Control and Test Substances

1. Remove the 96-well plates (from **Section 9.2.3 step 18**) from the incubator; inspect them using an inverted microscope. Only use plates in which the cells in all wells receive a score of 1 according to **Table 11-1**.
2. Remove medium by inverting the plate onto blotter paper. Gently tap plate against the bench surface to remove residual liquid trapped in the wells.
3. Add 200 μ L of medium, reference standard, control or test substance to each well (see **Sections 14.0** and **15.0** for specific plate layouts).
4. Return plates to incubator (see **Section 9.0** for details) for 19 to 24 hours to allow maximal induction of luciferase activity in the cells.

13.1.1 Preparation of Microsoft Excel® Data Analysis Template For Range Finder Testing

1. In Microsoft Excel, open a new “AntRFTemplate” and save it with the appropriate project name as indicated in the NICEATM Style Guide.
2. Fill out the table at the top of the “Raw Data” worksheet with information regarding the Microplate reader used, Reading Direction, No. of Intervals, Tot. Meas. Time/Well (s), etc. (**Note:** this information can be permanently added to the default template “AntRFTemplate” on a laboratory specific basis).
3. Add the following information regarding the assay to the “Compound Tracking” worksheet.
 - Plate # - Enter the experiment ID or plate number into cell E1
 - Cell Lot # - Enter the passage or lot number of the cells used for this experiment into cell B5
 - DMSO and Media Lot #'s – Enter the lot numbers for the DMSO and Media in cells B6 and B7
 - Test Substance Code – Enter the test substance codes into cells C14 to C19
 - Name: Enter the experimenter name into cell G6
 - Date: Enter the experiment date in the format day\month\year into cell G10
 - Comments: - Enter any comments about the experiment in this box (e.g., plate contaminated)
4. Enter the following substance testing information to the “List” worksheet:
 - Concentration – Type in the test substance concentration in μ g/mL in descending order.
 - Any specific comments about the test substance or condition of the wells should be entered into this sheet, in the comments section
 - All of the remaining cells on the “List” worksheet should populate automatically.
 - The “Template”, “Compound Mixing” and “Visual Inspection” worksheet should automatically populate with the information entered into the “Compound Tracking” and “List” worksheet.
5. Save the newly named project file.
6. Print out either the “List” or “Template” worksheet for help with dosing the 96-well plate. Sign and date the print out and store in study notebook.

13.1.2 Preparation of Excel Data Analysis Template for Comprehensive Testing

1. In Excel, open a new “AntCTTemplate” and save it with the appropriate project name as indicated in the NICEATM Style Guide.
2. Fill out the table at the top of the “Raw Data” worksheet with information regarding the Microplate reader used, Reading Direction, No. of Intervals, Tot. Meas. Time/Well (s), etc. (**Note:** this

information can be permanently added to the default template “AntCTTemplate” on a laboratory-specific basis).

3. On the “Compound Tracking” worksheet, enter the following information:
 - Plate # - Enter the experiment ID or plate number into cell E1
 - Cell Lot # - Enter the passage or lot number of the cells used for this experiment into cell C5
 - DMSO and Media Lot #'s – Enter the lot numbers for the DMSO and Media in cells C6 and C7
 - Test Substance Code – Enter the test substance codes into cells C15 and C16. Enter the test substance dilution into cells D15 and D16.
 - Name: Enter the experimenter name into cell F6
 - Date: Enter the experiment date in the format day\month\year into cell G10
 - Comments: - Enter any comments about the experiment in this box (e.g., plate contaminated)
4. Enter the following substance testing information to the “List” worksheet:
 - Concentration – Type in the test substance concentration in µg/mL in descending order.
 - Any specific comments about the test substance or condition of the wells should be entered into this sheet, in the comments section
 - All of the remaining cells on the “List” worksheet should populate automatically.
 - The “Template”, “Compound Mixing” and “Visual Inspection” worksheet should automatically populate with the information entered into the “Compound Tracking” and “List” worksheet.
5. Save the newly named project file.
6. Print out either the “List” or “Template” worksheet for help with dosing the 96-well plate. Sign and date the print out and store in study notebook.

13.2 Visual Evaluation of Cell Viability

1. Nineteen to 24 hours after dosing the plate, remove the plate from the incubator and remove the media from the wells by inverting the plate onto blotter paper. Gently tap plate against the bench surface to remove residual liquid trapped in the wells.
2. Use a repeat pipetter to add 50 µL 1X PBS to all wells. Immediately remove PBS by inversion.
3. Using an inverted microscope, inspect all of the wells used in the 96-well plate and record the visual observations using the scores in **Table 13-1**.

Table 13-1 Visual Observation Scoring

Viability Score	Brief Description ¹
1	Normal Cell Morphology and Cell Density
2	Altered Cell Morphology and/or Small Gaps between Cells
3	Altered Cell Morphology and/or Large Gaps between Cells
4	Few (or no) Visible Cells
P	Wells containing precipitation are to be noted with “P”

¹ Reference photomicrographs are provided in the BG1LUC ER TA Validation Study “Visual Observation Cell Viability Manual.”

13.3 Lysis of Cells for BG1Luc ER TA

1. Apply the reflective white backing tape to the bottom of the 96-well plate (this will increase the effectiveness of the luminometer).
2. Add 30 μ L 1X lysis reagent to the assay wells and place the 96-well plate on an orbital shaker for one minute.
3. Remove plate from shaker and measure luminescence (as described in **Section 13.4**).

13.4 Measurement of Luminescence

Luminescence is measured in the range of 300 to 650 nm, using an injecting luminometer and with software that controls the injection volume and measurement interval. Light emission from each well is expressed as relative light units (RLU) per well. The luminometer output is saved as raw data in an Excel spread sheet. A hard copy of the luminometer raw data should be signed, dated and stored in the study notebook.

13.5 Data Analysis

BG1Luc ER TA uses an Excel spreadsheet to collect and adjust the RLU values obtained from the luminometer and a GraphPad Prism[®] template to analyze and graph data. Plate reduction is calculated using unadjusted RLU values.

The Excel spreadsheet subtracts background luminescence (average DMSO solvent control RLU value) from test substance, reference standard and control RLU values. Test substance, reference standard, and control RLU values are then adjusted relative to the highest Ral/E2 reference standard RLU value, which is set to 10,000. After adjustment, values are transferred to GraphPad Prism for data analysis and graphing.

13.5.1 Collection and Adjustment of Luminometer Data for Range Finder Testing

The following steps describe the procedures required to populate the Excel spreadsheet that has been configured to collect and adjust the RLU values obtained from the luminometer.

1. Open the raw data file and the corresponding experimental Excel[®] spreadsheet from **Section 13.1.1**.
2. Copy the raw data using the Excel copy function, then paste the copied data into cell B19 of the "RAW DATA" tab in the experimental Excel spreadsheet using the **Paste Special – Values** command. This position corresponds to position A1 in the table labeled Table 1 in this tab.
3. Examine the DMSO data in Table 1 of the Excel spreadsheet to determine whether there are any potential outliers. See **Section 13.5.3** for further explanation of outlier determinations.
4. If an outlier is identified, perform the following steps to remove the outlier from calculations:
 - Correct the equation used to calculate DMSO background in Table 1 [e.g., if outlier is located in cell F26, adjust the calculation in cell H40 to read =AVERAGE(E26,G26)]
 - Then correct the equation used to calculate the average DMSO value in Table 2 [e.g., following the above example, adjust cell M42 to read =AVERAGE(E38,G38)]
 - Then correct the equation used to calculate the standard deviation of the DMSO value in Table 2 [e.g., following the above example, adjust cell M43 to read =STDEV(E38,G38)]
5. Excel will automatically subtract the background (the average DMSO control value) from all of the RLU values in Table 1 and populate Table 2 with these adjusted values.
6. To calculate plate reduction, identify the cell containing the Ral/E2a replicate in Table 1, plate row H that has the lowest RLU value (i.e., cell B26, C26, or D26).

7. Identify the cell containing the Ral/E2a replicate in Table 1, plate row H that has the highest RLU value (i.e., cell B26, C26, or D26).
8. Click into cell D14 and enter the cell number from **Section 13.5.1 step 7** into the numerator and the cell number from **step 6** into the denominator.
9. Identify the cell containing the Ral/E2b replicate in Table 1, plate row H that has the lowest RLU value (i.e., cell K26, L26, or M26).
10. Identify the cell containing the Ral/E2b replicate in Table 1, plate row H that has the highest RLU value (i.e., cell K26, L26, or M26).
11. Click into cell E14 and enter the cell number from **Section 13.5.1 step 10** into the numerator and the cell number from **step 9** into the denominator.
12. Click on the “ER Antagonist Report” worksheet.
13. The data for the Ral/E2 reference standard, DMSO, and E2, replicates populate the left portion (columns A-F) of the spreadsheet. The data is automatically placed into an Excel® graph.
14. To set the highest RLU value for the reference standard to 10,000 RLU, go to cell C2 of “ER Antagonist Report” worksheet and check the formula contained within that cell. The divisor should be the cell number of the cell containing the highest averaged Ral/E2 RLU value (column A).
15. Open the “Visual Observation Scoring” worksheet. Enter the visual observation scores for each well on the 96-well plate. This data will be linked to the “ER Antagonist Report” worksheet.
16. After the testing results have been evaluated and reviewed for quality control, enter the following information into the Compound Tracking worksheet:
 - Enter pass/fail results for plate reference standard and control parameters into the Plate Pass/Fail Table
 - Enter information from the testing of coded substances into the Testing Results Table
 - Reviewer Name – Enter the name of the person who Reviewed\QC’ed the data into cell A34
 - Date – Enter the date on which the data was reviewed into cell D34

13.5.2 Collection and Adjustment of Luminometer Data for Comprehensive Testing

The following steps describe the procedures required to populate the Excel® spreadsheet that has been configured to collect and adjust the RLU values obtained from the luminometer.

1. Open the raw data file and the corresponding experimental Excel® spreadsheet from **Section 13.1.2**.
2. Copy the raw data using the Excel copy function, then paste the copied data into cell B14 of the “RAW DATA” tab in the experimental Excel® spreadsheet using the **Paste Special – Values** command. This position corresponds to position A1 in the table labeled Table 1 in this tab.
3. Examine the DMSO data in Table 1 of the Excel® spreadsheet to determine whether there are any potential outliers. See **Section 13.5.3** for further explanation of outlier determinations.
4. If an outlier is identified, perform the following steps to remove the outlier from calculations:
 - Correct the equation used to calculate DMSO background in Table 1 [e.g., if outlier is located in cell M14, adjust the calculation in cell H40 to read =AVERAGE(M15:M17)]
 - Then correct the equation used to calculate the average DMSO value in Table 2 [e.g., following the above example, adjust cell M35 to read =AVERAGE(M25:M27)]
 - Then correct the equation used to calculate the standard deviation of the DMSO value in Table 2 [e.g., following the above example, adjust cell M36 to read =STDEV(M25:M27)]

5. Excel will automatically subtract the background (the average DMSO control value) from all of the RLU values in Table 1 and populate Table 2 with these adjusted values.
6. To calculate plate reduction, identify the cell containing the Ral/E2 replicate in plate row G that has the lowest RLU value.
7. Identify the cell containing the Ral/E2 replicate in plate row G that has the highest RLU value.
8. Click into cell D14 and enter the cell number from **Section 13.5.2 step 7** into the numerator and the cell number from **step 6** into the denominator.
9. Identify the cell containing the Ral/E2 replicate in plate row H that has the lowest RLU value.
10. Identify the cell containing the Ral/E2 replicate in plate row H that has the highest RLU value.
11. Click into cell E14 and enter the cell number from **Section 13.5.2 step 10** into the numerator and the cell number from **step 9** into the denominator.
12. Click on the “ER Antagonist Report” worksheet.
13. The data for the Ral/E2 reference standard, DMSO, E2, and TAM/E2 replicates populate the left portion (columns A-E) of the spreadsheet. The data is automatically placed into an Excel® graph.
14. To set the highest RLU value for the reference standard to 10,000 RLU, go to cell D2 of “ER Antagonist Report” worksheet and check the formula contained within that cell. The divisor should be the cell number of the cell containing the highest averaged Ral/E2 RLU value (column A).
15. Open the “Visual Observation Scoring” worksheet. Enter the visual observation scores for each well on the 96-well plate. This data will be linked to the “ER Antagonist Report” worksheet.
16. Copy the data into GraphPad Prism® for the calculation of IC₅₀ values and to graph experimental results as indicated in the NICEATM Prism® Users Guide.
17. After the testing results have been evaluated and reviewed for quality control, enter the following information into the Compound Tracking worksheet:
 - Enter pass/fail results for plate reference standard and control parameters into the Plate Pass/Fail Table
 - Enter information from the testing of coded substances into the Testing Results Table
 - Reviewer Name – Enter the name of the person who Reviewed\QC’ed the data into cell A34
 - Date – Enter the date on which the data was reviewed into cell D32

13.5.3 Determination of Outliers

The Study Director will use good statistical judgment for determining “unusable” wells that will be excluded from the data analysis and will provide an explanation in the study notebook for any excluded data. This judgment for data acceptance will include Q-test analysis.

The formula for the Q test is:

Outlier – Nearest Neighbor

Range (Highest – Lowest)

where the outlier is the value proposed for exclusion, the nearest neighbor is the value closest to the outlier, and the range is the range of the three values (Q values for samples sizes from 3 to 10 are provided in **Table 13-2**). For example, if the value of this ratio is greater than 0.94 (the Q value for the 90% confidence interval for a sample size of three) or 0.76 (the Q value for the 90% confidence interval for a sample size of four), the outlier may be excluded from data analysis.

Table 13-2 Q Test Values

Number Of Observations	Q Value
2	-
3	0.94
4	0.76
5	0.64
6	0.56
7	0.51
8	0.47
9	0.44
10	0.41

For E2 reference standard replicates (sample size of two), any adjusted RLU value for a replicate at a given concentration of E2 is considered an outlier if its value is more than 20% above or below the adjusted RLU value for that concentration in the historical database.

13.5.4 Acceptance Criteria

13.5.4.1 Range Finder Testing

Acceptance or rejection of a range finder test is based on reference standard and solvent control results from each experiment conducted on a 96-well plate.

- Reduction: Plate reduction, as measured by dividing the averaged highest Ral/E2 reference standard RLU value by the averaged DMSO control RLU value, must be greater than three-fold.
- E2 control results: E2 control RLU values must be within 2.5 times the standard deviation of the historical E2 control mean RLU value (**See Section 16.1**).
- DMSO control results: DMSO control RLU values must be within 2.5 times the standard deviation of the historical solvent control mean RLU value (see **Section 16.2**).

An experiment that fails either acceptance criterion will be discarded and repeated.

13.5.4.2 Comprehensive Testing

Acceptance or rejection of a test is based on evaluation of reference standard and control results from each experiment conducted on a 96-well plate. Results are compared to quality controls (QC) for these parameters derived from the historical database (see **Section 16.5**), which are summarized below.

- Reduction: Plate reduction, as measured by dividing the averaged highest Ral/E2 reference standard RLU value by the averaged lowest Ral/E2 control RLU value, must be greater than three-fold.
- DMSO control results: DMSO control RLU values must be within 2.5 times the standard deviation of the historical solvent control mean RLU value (see **Section 16.5**).
- Reference standard results: The Ral/2 reference standard concentration-response curve should be sigmoidal in shape and have at least three values within the linear portion of the concentration-response curve.
- E2 control results: E2 control RLU values must be within 2.5 times the standard deviation of the historical E2 control mean RLU value.
- Positive control results: TAM/E2 control RLU values must be less than the E2 control mean minus three times the standard deviation from the E2 control mean.

An experiment that fails any single acceptance criterion will be discarded and repeated.

14.0 Range Finder Testing

Antagonist range finding for coded substances consists of seven point, 1:10 serial dilutions tested in duplicate wells of the 96-well plate. **Figure 14-1** contains a template for the plate layout used in antagonist range finder testing.

Figure 14-1 Antagonist Range Finder Plate Layout

TS1-1	TS1-1	TS2-1	TS2-1	TS3-1	TS3-1	TS4-1	TS4-1	TS5-1	TS5-1	TS6-1	TS6-1
TS1-2	TS1-2	TS2-2	TS2-2	TS3-2	TS3-2	TS4-2	TS4-2	TS5-2	TS5-2	TS6-2	TS6-2
TS1-3	TS1-3	TS2-3	TS2-3	TS3-3	TS3-3	TS4-3	TS4-3	TS5-3	TS5-3	TS6-3	TS6-3
TS1-4	TS1-4	TS2-4	TS2-4	TS3-4	TS3-4	TS4-4	TS4-4	TS5-4	TS5-4	TS6-4	TS6-4
TS1-5	TS1-5	TS2-5	TS2-5	TS3-5	TS3-5	TS4-5	TS4-5	TS5-5	TS5-5	TS6-5	TS6-5
TS1-6	TS1-6	TS2-6	TS2-6	TS3-6	TS3-6	TS4-6	TS4-6	TS5-6	TS5-6	TS6-6	TS6-6
TS1-7	TS1-7	TS2-7	TS2-7	TS3-7	TS3-7	TS4-7	TS4-7	TS5-7	TS5-7	TS6-7	TS6-7
Ral-1	Ral-2	Ral-3	VC	VC	VC	E2	E2	E2	Ral-1	Ral-2	Ral-3

Abbreviations: E2 = E2 control; Ral = raloxifene; TS = test substance; VC = vehicle control (DMSO [1% v/v EFM]).

Ral-1 to Ral-3 = concentrations of the raloxifene/E2 reference standard (from high to low)

TS1-1 to TS1-7 = concentrations (from high to low) of test substance 1 (TS1)

TS2-1 to TS2-7 = concentrations (from high to low) of test substance 2 (TS2)

TS3-1 to TS3-7 = concentrations (from high to low) of test substance 3 (TS3)

TS4-1 to TS4-7 = concentrations (from high to low) of test substance 4 (TS4)

TS5-1 to TS5-7 = concentrations (from high to low) of test substance 5 (TS5)

TS6-1 to TS6-7 = concentrations (from high to low) of test substance 6 (TS6)

Evaluate whether range finder experiments have met acceptance criteria (see **Section 13.6.3**).

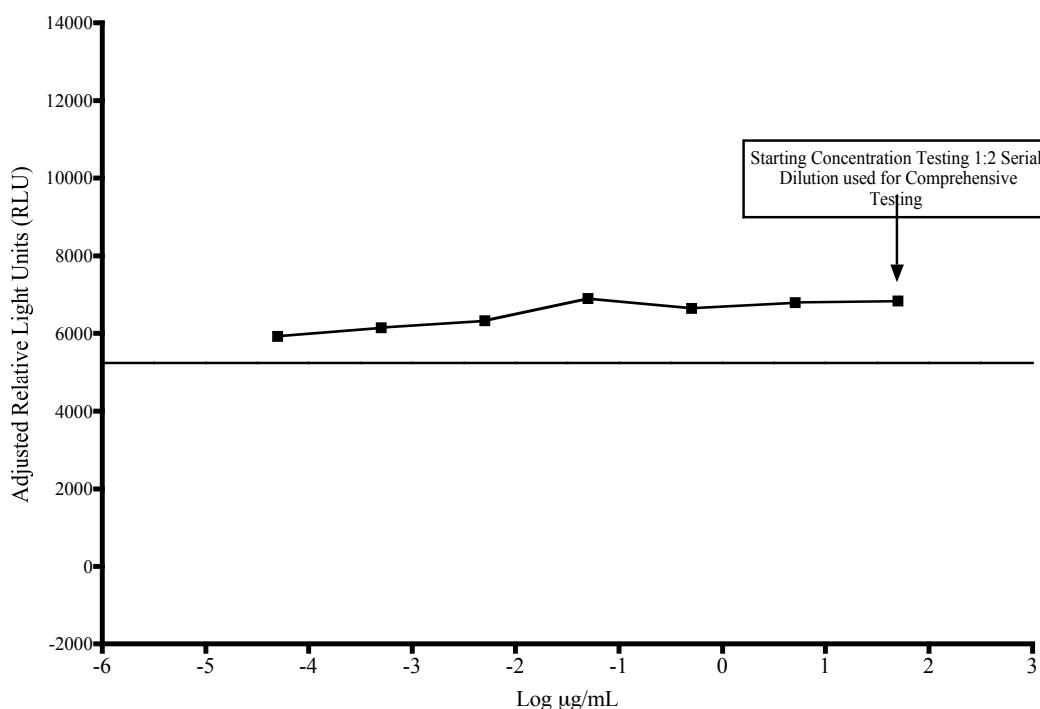
To determine starting concentrations for comprehensive testing use the following criteria:

- If results in the range finder test suggest that the test substance is negative for antagonist activity (i.e., if there are no points on the test substance concentration curve that are less than the mean minus three times the standard deviation of the E2 control, see **Figure 14-2**), comprehensive testing will be conducted using an 11-point 1:2 serial dilution using the maximum soluble concentration of test substance as the with the limit dose as the starting concentration.
- If results in the range finder test suggest that the test substance is negative for antagonist activity (i.e., if there are no points on the test substance concentration curve that are than the mean minus three times the standard deviation of the E2 control), and the higher concentrations in the range finder are cytotoxic, comprehensive testing will be conducted using an 11 point 1:2 serial dilution with the lowest cytotoxic concentration as the starting concentration (see **Figure 14-3**).
- If results in the range finder test suggest that the test substance is positive for antagonist activity (i.e., if there are points on the test substance concentration curve that are less than the mean minus three times the standard deviation of the E2 control), the top concentration to be used for the 11-point dilution scheme in comprehensive testing should be one of the following:
 - The concentration giving the lowest adjusted RLU value in the range finder
 - The maximum soluble concentration (See **Figure 14-2**)
 - The lowest cytotoxic concentration (See **Figure 14-3** for a related example).

The 11-point dilution scheme will be based on either a 1:2 or 1:5 serial or dilution according to the following criteria:

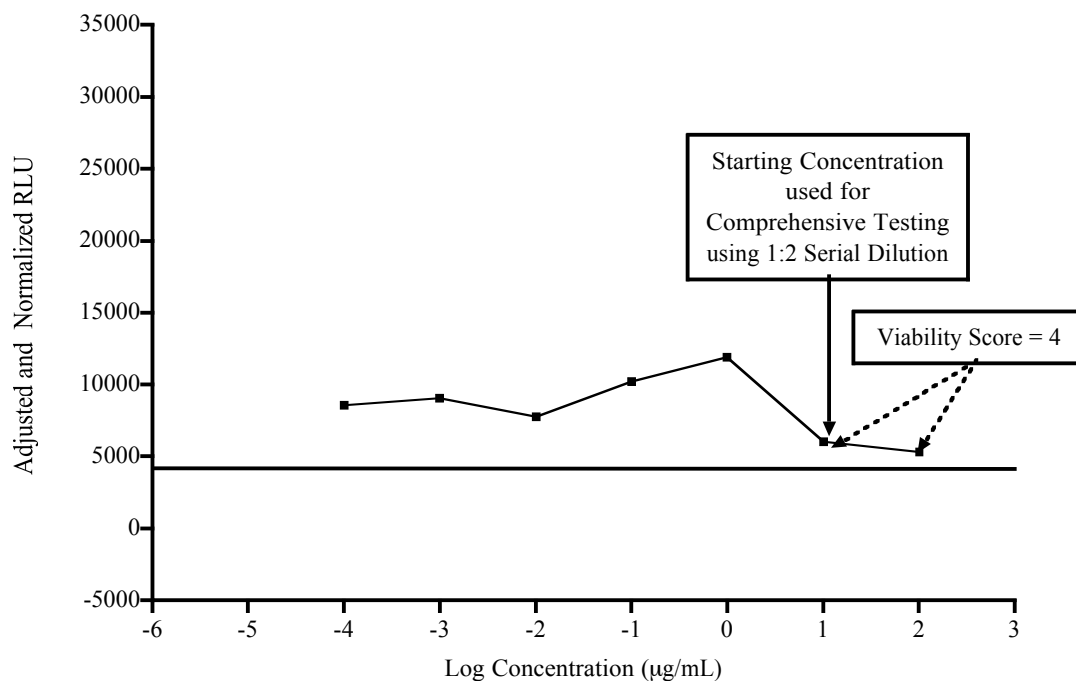
- An 11-point 1:2 serial dilution should be used if the resulting concentration range (note: an 11-point 1:2 serial dilution will cover a range of concentrations over approximately three orders of magnitude [three logs]) will encompass the full range of responses based on the concentration response curve generated in the range finder test (see **Figure 14-4**).
- If the concentration range that would be generated with the 1:2 serial dilution will not encompass the full range of responses based on the concentration response curve in the range finder test (see **Figure 14-5**), an 11-point 1:5 serial dilution should be used instead.
- If a substance exhibits a biphasic concentration response curve in the range finder test (see **Figure 14-6**), both phases should also be resolved in comprehensive testing. In this case, two peaks could potentially be used to identify the top concentration to be used for the 11-point dilution scheme in comprehensive testing. In order to resolve both curves, the top concentration should be based on the peak associated with the higher concentration and the top dose one log concentration higher than the concentration giving the lowest adjusted RLU value in the range finder. An 11-point 1:5 serial dilution should be used.

Figure 14-2 Antagonist Range Finder (Example 1)



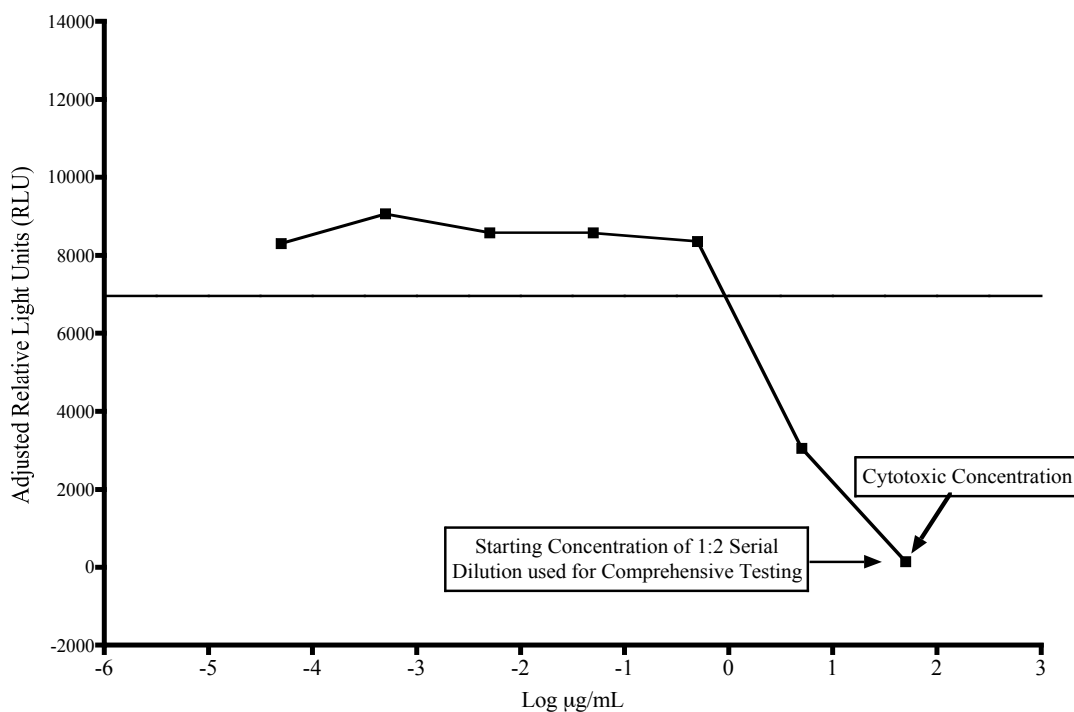
The solid horizontal line represents the mean minus three times the standard deviation of the E2 control.

Figure 14-3 Antagonist Range Finder (Example 2)



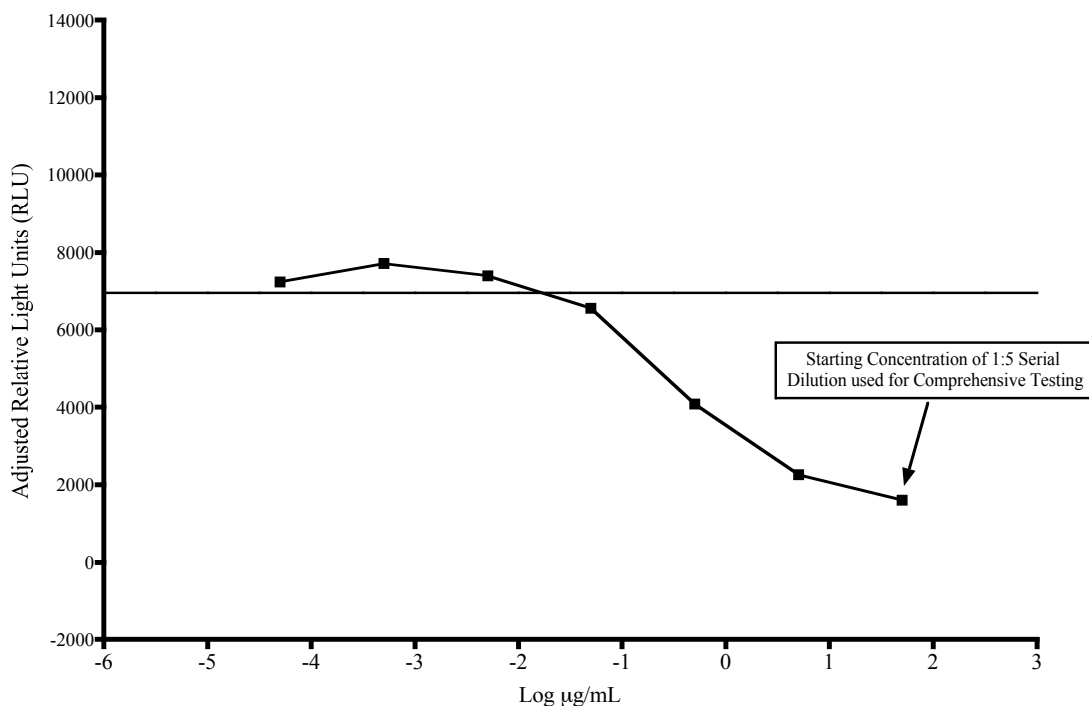
The solid horizontal line represents the mean minus three times the standard deviation of the E2 control.

Figure 14-4 Antagonist Range Finder (Example 3)



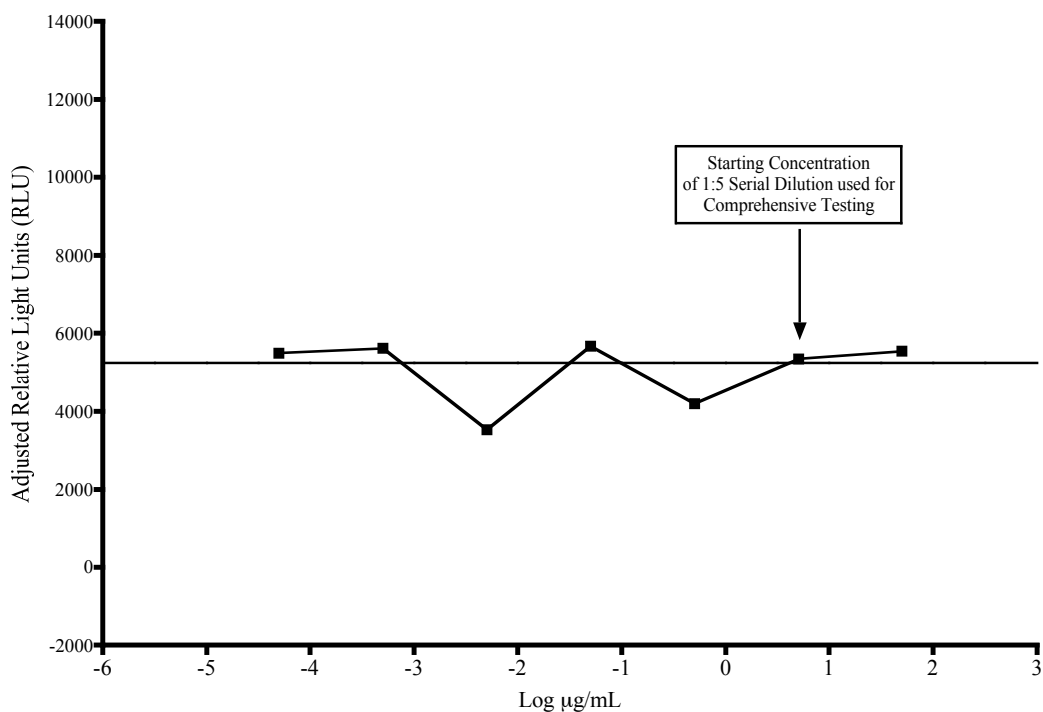
The solid horizontal line represents the mean minus three times the standard deviation of the E2 control.

Figure 14-5 Antagonist Range Finder (Example 4)



The solid horizontal line represents the mean minus three times the standard deviation of the E2 control.

Figure 14-6 Antagonist Range Finder (example 5)



The solid horizontal line represents the mean minus three times the standard deviation of the E2 control.

15.0 Comprehensive Testing

Antagonist comprehensive testing for coded substances consists of 11 point, 1:2 serial dilutions, with each concentration tested in triplicate wells of the 96-well plate. **Figure 15-1** contains a template for the plate layout to be used in antagonist comprehensive testing.

Figure 15-1 Antagonist Comprehensive Test Plate Layout

TS1-1	TS1-2	TS1-3	TS1-4	TS1-5	TS1-6	TS1-7	TS1-8	TS1-9	TS1-10	TS1-11	VC
TS1-1	TS1-2	TS1-3	TS1-4	TS1-5	TS1-6	TS1-7	TS1-8	TS1-9	TS1-10	TS1-11	VC
TS1-1	TS1-2	TS1-3	TS1-4	TS1-5	TS1-6	TS1-7	TS1-8	TS1-9	TS1-10	TS1-11	VC
TS2-1	TS2-2	TS2-3	TS2-4	TS2-5	TS2-6	TS2-7	TS2-8	TS2-9	TS2-10	TS2-11	VC
TS2-1	TS2-2	TS2-3	TS2-4	TS2-5	TS2-6	TS2-7	TS2-8	TS2-9	TS2-10	TS2-11	Tam
TS2-1	TS2-2	TS2-3	TS2-4	TS2-5	TS2-6	TS2-7	TS2-8	TS2-9	TS2-10	TS2-11	Tam
Ral-1	Ral-2	Ral-3	Ral-4	Ral-5	Ral-6	Ral-7	Ral-8	Ral-9	E2	E2	Tam
Ral-1	Ral-2	Ral-3	Ral-4	Ral-5	Ral-6	Ral-7	Ral-8	Ral-9	E2	E2	Tam

Note: All reference and test wells contain a fixed concentration of E2 (4.90×10^{-11} M).

Abbreviations: E2 = E2 control; Tam = tamoxifen/E2 weak positive control; VC = vehicle control (DMSO [1% v/v EFM]).

Ral-1 to Ral-9 = concentrations of the raloxifene/E2 reference standard (from high to low)

TS1-1 to TS1-11 = concentrations (from high to low) of test substance 1 (TS1)

TS2-1 to TS2-11 = concentrations (from high to low) of test substance 2 (TS2)

Evaluate whether comprehensive experiments have met acceptance criteria (see **Section 13.6.3**) and graph the data as described in the NICEATM Prism[®] users guide. Then evaluate the test substance results.

15.1 Test Substance Positive and Negative Criteria

Positive Classification:

- All test substances classified as positive for ER antagonist activity should have a concentration–response curve consisting of a baseline followed by a negative slope.
- The line defining the negative slope must contain at least three points with nonoverlapping error bars. Points forming the baseline are excluded, but the linear portion of the curve may include the first point of the plateau.
- A positive classification requires a response amplitude, the difference between baseline and bottom, of at least 20% of the maximal value for the reference estrogen (i.e., 2000 RLUs when the maximal response value of the reference estrogen is adjusted to 10,000 RLUs).
- The highest noncytotoxic concentrations of the test substance should be less than or equal to 1×10^{-5} M (approximately 3.13×10^0 µg/µL).
- If possible, an IC_{50} value should be calculated for each positive substance

Negative Classification:

- Test substances are classified as negative for antagonist activity if all data points are above the EC_{80} value (80% of the E2 response, or 8000 RLUs).

Inadequate Classification:

- 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.

16.0 Compilation of the Historical Quality Control Database

Historical databases are maintained in order to ensure that the assay is functioning properly. Historical databases are compiled using Excel spreadsheets and are separate from the spreadsheets used to collect the data for individual test plates. Reference standard and control data is used to develop and maintain the historical database and are used as quality controls to determine acceptance of individual test plates.

The sources of data needed to compile the historical database for the E2 control and TAM/E2 control values are the experiment specific Excel data collection and analysis spreadsheets (see **Section 13.5.2**) used for BG1Luc ER TA antagonist testing. The sources of the data needed to compile the historical database for the DMSO control are the experiment specific Excel data collection and analysis spreadsheets used for BG1Luc ER TA antagonist and agonist testing (see **Section 13.5.2** of the BG1Luc ER TA antagonist protocol and **Section 11.5.2** in the BG1Luc ER TA agonist protocol).

16.1 E2 Control

Open the BG1Luc ER TA antagonist specific historical database Excel spreadsheet (LUMI_AgandAntQC.xls) and save under a new name using the Excel “Save As” function, adding the laboratory designator to the file name (e.g., for Laboratory H, new name = HLUMI_AgandAntQC.xls). Open the E2 Control worksheet and enter the date and experiment name into worksheet columns A and B respectively. Enter the experimental mean adjusted E2 control value (from cell D37 in the ER Antagonist Report worksheet of the Excel® data collection and analysis spreadsheet) into the Antagonist E2 control worksheet, column C. Acceptance or rejection of plate E2 control data for comprehensive testing is based on whether the mean plate E2 RLU value falls within 2.5 times the standard deviation of the E2 value in the historical database (columns G and H in the E2 Control worksheet).

16.2 DMSO

Open the combined agonist and antagonist BG1Luc ER TA historical database Excel® spreadsheet (LUMI_AgandAntQC.xls) and save under a new name using the Excel® “Save As” function, adding the laboratory designator to the file name (e.g., for Laboratory H, new name = HLUMI_AgandAntQC.xls). Enter the date and experiment name into worksheet columns A and B respectively. Enter the experimental mean DMSO control value (from cell H37 in the RAW DATA worksheet of the agonist and antagonist Excel® data collection and analysis spreadsheet) into worksheet column C. Acceptance or rejection of the plate DMSO control data for range finding and comprehensive testing is based on whether the mean plate DMSO RLU value falls within 2.5 times the standard deviation of the DMSO value in the historical database (columns G and H in the DMSO worksheet).

17.0 Quality Testing of Materials

All information pertaining to the preparation and testing of media, media supplements, and other materials should be recorded in the Study Notebook.

17.1 Tissue Culture Media

Each lot of tissue culture medium must be tested in a single growth flask of cells before use in ongoing tissue culture or experimentation (**note:** each bottle within a given lot of charcoal/dextran treated FBS must be tested separately).

1. Every new lot of media (RPMI and DMEM) and media components (FBS, charcoal/dextran treated FBS, and L-glutamine) must first be tested on the BG1Luc ER TA prior to being used in any GLP acceptable assays.
2. Add 4 μ L of DMSO (previously tested) into four separate 13 mm tubes.
3. Add 400 mL media (to be tested) to 13 mm tube.
4. Dose an experimental plate as in **Section 12.0**, treating the media being tested as a test substance.
5. Analyze 96-well plate as described in **Section 12.0**, comparing the data from the DMSO controls made using previously tested tissue culture media to the new media being tested.
6. Use the agonist historical database to determine if the new media with DMSO lies within 2.5 standard deviations of the mean for the media. If the RLU values for the new media with DMSO lie within 2.5 standard deviations of the DMSO mean from the historical database, the new lot of media is acceptable. If the RLU values for the new media with DMSO do not lie within 2.5 standard deviations of the DMSO mean from the historical database, the new lot may not be used in the assay.
7. Note date and lot number in study notebook.
8. If the new bottle passes quality testing as described in **Section 15.1 step 6**, apply the media to a single flask cells and observe the cells growth and morphology over the following 2 to 3 days. If there is no change in growth or morphology, the new media is acceptable for use.

17.2 G418

1. New lots of G418 must first be tested on the BG1Luc ER TA prior to being used in any GLP acceptable assays.
2. Add 220 μ L of G418 (previously tested) to a single flask containing cells growing in RPMI.
3. Add 220 μ L of G418 (to be tested) to a different flask containing cells growing in RPMI.
4. Observe cellular growth and morphology in both tissue culture flasks over a 48 to 72 hour period. If there are no differences in observed growth rate and morphology between the two flasks, the new G418 lot is acceptable.
5. If cellular growth is decreased, or the cells exhibit abnormal morphology, the new lot of G418 is not acceptable.
6. Note date and lot number in study book.

17.3 DMSO

1. Every new bottle of DMSO must be tested on the BG1Luc ER TA prior to use in any GLP acceptable assays.
2. Add 4 μ L of DMSO (to be tested) into four separate 13 mm tubes.
3. Add 400 mL media (previously tested) the same tubes.
4. Dose an experimental plate as in **Section 15.0**, treating the media being tested as a test substance.

5. Analyze 96-well plate as described in **Section 15.0**, comparing the data from the DMSO controls made using previously tested tissue culture media to the new media being tested.
6. Use the agonist historical database to determine if media with new DMSO lies within 2.5 standard deviations of the DMSO mean from historical database. If the RLU values for the media with new DMSO lie within 2.5 standard deviations of the DMSO mean from the historical database, the new lot of DMSO is acceptable. If the RLU values for media with new DMSO do not lie within 2.5 standard deviations of the DMSO mean from historical database, the new lot may not be used in the assay.
7. Note the date, lot number, and bottle number in study book.
8. If no DMSO has been previously tested, test several bottles as described in **Section 15.3**, and determine whether any of the bottles of DMSO have a higher average RLU than the other bottle(s) tested. Use the DMSO with the lowest average RLU for official experiments.

17.4 Plastic Tissue Culture Materials

1. Grow one set of cells, plate them for experiments on plastic ware from the new lot and one set of cells in the plastic ware from a previous lot, and dose them with E2 reference standard and controls.
2. Perform the BG1Luc ER TA experiment with both sets of cells.
3. If all of the analysis falls within acceptable QC criteria, then the new manufacturer's products may be used.

18.0 References

Coecke S. Guidance on Good Cell Culture Practice. 2005. ATLA 33, 261-287.

Eli Lilly and Company and National Institutes of Health Chemical Genomics Center. 2005. Assay Guidance Manual Version 4.1. Bethesda, MD: National Institutes of Health. Available: http://www.ncgc.nih.gov/guidance/manual_toc.html [accessed 05 September 2006]

ICCVAM. 2001. Guidance Document on Using *In Vitro* Data to Estimate *In Vivo* Starting Doses for Acute Toxicity. NIH Pub. No. 01-4500. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/methods/invidocs/guidance/iv_guide.pdf [accessed 31 August 2006]

This page intentionally left blank

Appendix C

Final Background Review Document

Validation Study of the BG1Luc4E2 Estrogen Receptor (ER) Transcriptional Activation (TA) Test Method

This page intentionally left blank

Final Background Review Document
Validation Study of the BG1Luc4E2 Estrogen Receptor (ER)
Transcriptional Activation (TA) Test Method

**Interagency Coordinating Committee on the
Validation of Alternative Methods**

**National Toxicology Program Interagency Center for the
Evaluation of Alternative Toxicological Methods**

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

2011

**National Toxicology Program
P.O. Box 12233
Research Triangle Park, NC 27709**

This page intentionally left blank

Table of Contents

List of Tables.....	C-10
List of Figures	C-13
List of Abbreviations and Acronyms.....	C-14
Interagency Coordinating Committee on the Validation of Alternative Methods:	
Agency Representatives	C-17
Acknowledgements	C-19
Preface	C-23
Executive Summary.....	C-25
1.0 Introduction	C-29
1.1 Objective	C-29
1.2 Public Health Perspective	C-29
1.3 Historical Background	C-29
1.4 Nomination and Pre-Screen Evaluation of the BG1Luc4E2 ER TA Test Method.....	C-30
1.5 Basis for High Priority for Validation Studies.....	C-31
1.5.1 Criterion 1	C-31
1.5.2 Criterion 2	C-32
1.5.3 Criterion 3	C-33
1.5.4 Criterion 4	C-33
1.5.5 Criterion 5	C-34
1.6 BG1Luc ER TA Test Method Protocol Standardization Study	C-34
1.7 Interlaboratory BG1Luc ER TA Validation Study	C-34
1.8 Scientific Basis for the BG1Luc ER TA Test Method.....	C-36
1.9 Range of Substances Amenable to the BG1Luc ER TA Test Method	C-37
2.0 BG1Luc ER TA Test Method Protocol Components.....	C-38
2.1 Overview	C-38
2.1.1 General Procedural Overview	C-39
2.2 Materials.....	C-39
2.2.1 BG1Luc4E2 Cells	C-39
2.2.2 Cell Culture Reagents and Supplies	C-39
2.2.3 Equipment	C-40
2.3 Cell Culture	C-40

2.4	Reference Standards and Controls.....	C-41
2.4.1	Vehicle Control.....	C-41
2.4.2	Estrogenic Reference Standard.....	C-41
2.4.3	Weak Agonist Control	C-42
2.4.4	Anti-Estrogenic Reference Standard	C-42
2.4.5	Weak Antagonist Control	C-43
2.5	Test Substance Concentrations.....	C-44
2.5.1	Solubility Testing.....	C-44
2.5.2	Cytotoxicity Testing	C-44
2.6	Range Finder Testing.....	C-45
2.6.1	Agonist Range Finder Testing	C-46
2.6.2	Antagonist Range Finder Testing	C-47
2.7	Comprehensive Testing	C-49
2.7.1	Comprehensive Agonist Testing.....	C-49
2.7.2	Comprehensive Antagonist Testing.....	C-52
3.0	Substances Used for the Validation of the BG1Luc ER TA Test Method.....	C-55
3.1	Development of the List of 78 ICCVAM-Recommended Test Substances	C-55
3.2	Substances Used to Evaluate Test Method Accuracy	C-60
3.3	Substances Used to Evaluate Concordance with Other Accepted Methods.....	C-67
3.3.1	Substances Used to Evaluate BG1 Luc ER TA Concordance with the CERi-STTA	C-67
3.3.2	Substances Used to Evaluate BG1Luc ER TA Concordance with <i>In Vitro</i> ER Binding Test Methods	C-67
3.3.3	Substances Used to Evaluate BG1Luc ER TA Concordance with the <i>In Vivo</i> Uterotrophic Bioassay	C-68
3.4	Substances Tested in Each Phase of Validation	C-68
3.5	Substances Used to Assess Intralaboratory Reproducibility	C-80
3.6	Substances Used to Assess Interlaboratory Reproducibility	C-80
3.7	Chemical Classes Represented in the List of Substances.....	C-80
3.8	Product Classes Represented in the List of Substances.....	C-81
3.9	Test Substance Procurement, Coding, and Distribution.....	C-82
4.0	Test Method Data And Results.....	C-83
4.1	Availability of Original Data Used to Evaluate Test Method Performance.....	C-83
4.2	BG1Luc ER TA Agonist and Antagonist Reference Standard and Control Data.....	C-83
4.2.1	Agonist E2 Reference Standard.....	C-83

4.2.2	Agonist DMSO Control Values	C-84
4.2.3	Maximum Fold Induction of E2 Response During Agonist Testing	C-86
4.2.4	Weak Agonist Positive Control: Methoxychlor	C-86
4.2.5	Antagonist Raloxifene Reference Standard	C-87
4.2.6	Antagonist DMSO Control Values	C-88
4.2.7	Antagonist E2 Reference Standard	C-90
4.2.8	Maximum Fold Reduction of E2 Response During Antagonist Testing.....	C-90
4.2.9	Weak Antagonist Positive Control: Flavone	C-91
4.3	Solubility Test Results	C-92
4.4	Test Results for Coded Test Substances	C-98
4.4.1	Cell Viability Assessment	C-98
4.4.2	BG1Luc ER TA Agonist and Antagonist Data	C-99
5.0	Accuracy of the BG1Luc ER TA Test Method.....	C-110
5.1	Substances Used for Accuracy Analysis.....	C-111
5.2	Accuracy Analysis of the BG1Luc ER TA Agonist Test Method.....	C-115
5.2.1	Discordant Results for Agonist Analysis	C-115
5.3	Accuracy Analysis of the BG1Luc ER TA Antagonist Test Method	C-115
5.4	Comparison of BG1Luc ER TA Results with CERi-STTA (U.S. EPA OPPTS 890.1300) Results.....	C-116
5.5	Comparison of BG1Luc ER TA EC ₅₀ and IC ₅₀ Values with Values from ICCVAM Reference Data	C-118
5.6	Concordance of BG1Luc ER TA Results with Estrogen Receptor Binding Results	C-121
5.7	Comparison of BG1Luc ER TA Results with Uterotrophic Assay Results.....	C-122
6.0	Test Method Reliability	C-125
6.1	Intralaboratory Reproducibility.....	C-125
6.1.1	Agonist DMSO Control	C-125
6.1.2	Agonist E2 Reference Standard EC ₅₀ and Methoxychlor Control	C-128
6.1.3	Intralaboratory Reproducibility of Phase 2 Agonist Reference Substances	C-130
6.1.4	Antagonist DMSO Control.....	C-131
6.1.5	Antagonist E2 Control.....	C-134
6.1.6	Antagonist Raloxifene Reference Standard IC ₅₀ and Flavone Control Values.....	C-135
6.1.7	Intralaboratory Reproducibility of Phase 2 Antagonist Reference Substances	C-138

6.2	Interlaboratory Reproducibility	C-139
6.2.1	Interlaboratory Reproducibility of Phase 2 Reference Substances.....	C-139
6.2.2	Interlaboratory Reproducibility of Phase 3 Agonist Reference Substances.....	C-140
6.2.3	Interlaboratory Reproducibility of Phase 3 Antagonist Reference Substances.....	C-143
7.0	BG1Luc ER TA Data Quality.....	C-145
7.1	Compliance with GLP Regulations	C-145
7.2	QA Audit Results.....	C-145
7.3	Test Plate Failure Rates	C-146
7.3.1	Phase 2a	C-147
7.3.2	Phases 2b, 3, and 4 Failure Rates.....	C-150
7.4	Inadequate Results	C-151
7.5	Availability of Laboratory Notebooks or Other Records	C-153
8.0	Other Scientific Reports.....	C-154
8.1	Summaries of Available Data from Studies Using the BG1Luc ER TA Test Method	C-154
8.1.1	Rogers and Denison (2000)	C-154
8.1.2	Jefferson et al. (2002)	C-154
8.1.3	Gordon et al. (2003).....	C-154
8.1.4	Gordon et al. (2004).....	C-155
8.1.5	Gordon et al. (2005).....	C-155
8.1.6	Clark et al. (2007)	C-155
9.0	Animal Welfare Considerations (Reduction, Refinement, and Replacement)	C-157
9.1	Reduction, Refinement, and Replacement Considerations	C-157
9.2	Use of Animals in the BG1Luc ER TA Test Method.....	C-158
10.0	Practical Considerations	C-159
10.1	Transferability of the BG1Luc ER TA Test Method.....	C-159
10.1.1	Facilities and Major Equipment.....	C-159
10.1.2	General Availability of Other Necessary Equipment and Supplies.....	C-159
10.1.3	BG1Luc4E2 Cell Line	C-159
10.2	BG1Luc ER TA Test Method Training Considerations	C-160
10.3	Time and Cost Considerations.....	C-160
11.0	References.....	C-161
12.0	Glossary	C-171

The following annexes are available on the enclosed CD-ROM or at <http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-BRDannex.htm>:

Annex A	Submission of XDS's LUMI-CELL™ ER High-Throughput System for Screening Estrogen-Like Chemicals for Review by ICCVAM
Annex B	ICCVAM/NICEATM Evaluation – BG1Luc ER TA Submission
Annex C	NICEATM Report on the XDS Protocol Standardization Study Agonist and Antagonist Protocols
Annex D	Addendum to ICCVAM Evaluation of <i>In Vitro</i> Test Methods for Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays (NIH Publication No. 03-4503)
Annex E	ICCVAM/NICEATM BG1Luc ER TA – Agonist Protocol
Annex F	ICCVAM/NICEATM BG1Luc ER TA – Antagonist Protocol
Annex G	Materials Relating to Cell Viability
Annex H	ICCVAM/NICEATM BG1Luc4E2 ER TA – Validation Work Plan
Annex I	Substances Used for the Validation of the BG1Luc ER TA Test Method
Annex J	ICCVAM/NICEATM BG1Luc ER TA – Prism Files: Prism Files for All Validation Study Phases
Annex K	Experimental Details
Annex L	Within-Plate DMSO Control Values for BG1Luc ER TA Agonist and Antagonist Assays
Annex M	ICCVAM/NICEATM BG1Luc4E2 ER TA – Plate Redesign and Compilation of a Historical Database
Annex N	Literature Update Files

List of Tables

Table 1-1	Study Management Team for the BG1Luc ER TA Validation Study.....	C-35
Table 2-1	E2 Concentrations Tested for Agonist and Antagonist Methods.....	C-42
Table 2-2	Raloxifene Standard Concentrations Tested in the Antagonist Assay.....	C-43
Table 2-3	Visual Observation Scoring Table for Cell Viability.....	C-45
Table 3-1	Reference Substances Tested for ER TA Activity.....	C-55
Table 3-2	Reference Data Summary for ER Agonist and Antagonist TA Assays.....	C-61
Table 3-3	Test Substances Used for Statistical Determination of ER TA Agonist Assay Accuracy	C-64
Table 3-4	Test Substances Used for Statistical Determination of ER TA Antagonist Assay Accuracy	C-66
Table 3-5	Agonist Substances by Study Phase.....	C-69
Table 3-6	Antagonist Substances by Study Phase.....	C-74
Table 3-7	Distribution of Reference Substances by Chemical Class	C-80
Table 3-8	Distribution of Reference Substances by Product Class	C-81
Table 4-1	Summary of Agonist E2 Reference Standard EC ₅₀ Data by Study Phase.....	C-84
Table 4-2	Summary of Agonist Within-Plate DMSO Control Data by Laboratory and Study Phase	C-85
Table 4-3	Summary of Agonist Maximum Fold Induction Data by Laboratory and Study Phase	C-86
Table 4-4	Summary of Agonist Methoxychlor Control Data by Laboratory and Study Phase.....	C-87
Table 4-5	Summary of Antagonist Raloxifene Reference Standard IC ₅₀ Data by Laboratory and Study Phase	C-88
Table 4-6	Summary of Antagonist Within-Plate DMSO Control Data by Study Phase.....	C-89
Table 4-7	Summary of Antagonist E2 Control Data by Study Phase.....	C-90
Table 4-8	Summary of Antagonist Maximum Fold Reduction Data by Laboratory and Study Phase	C-91
Table 4-9	Summary of Antagonist Flavone Control Data by Study Phase	C-92
Table 4-10	Agonist Range Finder Starting Concentrations in Culture Medium.....	C-93
Table 4-11	Antagonist Range Finder Starting Concentrations in Culture Medium	C-96
Table 4-12	Agonist Summary Data for Phases 2a, 2b, and 3	C-99
Table 4-13	Antagonist Summary Data for Phases 2a, 2b, and 3	C-104
Table 4-14	Phase 4 Results from XDS.....	C-108
Table 5-1	Template for Concordance Analysis.....	C-110
Table 5-2	42 ICCVAM-Recommended Substances Used to Evaluate ER Agonist Accuracy ..	C-112

Table 5-3	25 ICCVAM-Recommended Substances Used to Evaluate ER Antagonist Accuracy	C-113
Table 5-4	Accuracy of the BG1Luc ER TA Agonist Test Method.....	C-115
Table 5-5	Discordant Substance in the BG1Luc ER TA Agonist Test Method	C-115
Table 5-6	Accuracy of the BG1Luc ER TA Antagonist Test Method.....	C-116
Table 5-7	Substances Used in the Evaluation of Accuracy of the BG1Luc ER TA and CERI-STTA Test Method Results	C-116
Table 5-8	Accuracy of the BG1Luc ER TA Test Method Assessed Using Agonist Reference Chemicals Listed in Table 5-7	C-117
Table 5-9	Accuracy of the CERI-STTA Method Assessed Using Agonist Reference Chemicals Listed in Table 5-7.....	C-118
Table 5-10	Median EC ₅₀ Values for Substances Used to Generate EC ₅₀ Linear Regression	C-118
Table 5-11	Median IC ₅₀ Values for Substances Used to Generate IC ₅₀ Linear Regression	C-119
Table 5-12	Substances Used to Assess BG1Luc ER TA Concordance with ER Binding Data ..	C-121
Table 5-13	Concordance of BG1Luc ER TA Results and ER Binding Results	C-122
Table 5-14	Substances Used in the Comparison of BG1Luc ER TA Agonist Classification and <i>In Vivo</i> Uterotrophic Assay Data.....	C-123
Table 5-15	Concordance of BG1Luc ER TA Agonist Classification and <i>In Vivo</i> Uterotrophic Assay Data.....	C-124
Table 6-1	Agonist Within-Plate DMSO Control Data.....	C-126
Table 6-2	Agonist E2 EC ₅₀ and Methoxychlor Control Values.....	C-128
Table 6-3	Intralaboratory Agreement for Multiple Testing of 12 Phase 2 Agonist Substances Tested Independently Three Times at Each Laboratory.....	C-131
Table 6-4	Antagonist DMSO Control Values.....	C-132
Table 6-5	Antagonist E2 Control Values	C-134
Table 6-6	Antagonist Raloxifene IC ₅₀ and Flavone Control Values.....	C-135
Table 6-7	Intralaboratory Agreement for Multiple Testing of 12 Phase 2 Antagonist Substances Tested Independently Three Times at Each Laboratory	C-139
Table 6-8	Interlaboratory Agreement for Phase 2 Test Substances	C-140
Table 6-9	Interlaboratory Agreement for Phase 3 Substances Tested Once at Each Laboratory	C-141
Table 7-1	Quality Assurance Review Dates	C-146
Table 7-2	Phase 2a Test Substance ER TA Agonist Activity for Plates That Passed or Failed Acceptance Criteria	C-148
Table 7-3	Phase 2a Test Substance ER TA Antagonist Activity for Plates That Passed or Failed Acceptance Criteria	C-148
Table 7-4	Comparison of Phase 2a Test Substance EC ₅₀ /IC ₅₀ Values for Plates That Passed or Failed Acceptance Criteria	C-149

Table 7-5	Test Plate Failure Rates for Agonists: Phases 2b–4	C-150
Table 7-6	Test Plate Failure Rates for Antagonists: Phases 2b–4	C-151
Table 7-7	Summary of Test Results Classified as Inadequate	C-153
Table 8-1	Visual Observation Scoring Table to Assess Cell Viability	C-156
Table 10-1	Example Suppliers and Costs of Major Equipment for the BG1Luc ER TA Test Method	C-159

List of Figures

Figure 1-1	NICEATM/ECVAM/JaCVAM Validation Study Phases.....	C-36
Figure 2-1	96-Well Test Plate Layout for Agonist Range Finder Testing.....	C-46
Figure 2-2	96-Well Test Plate Layout for Antagonist Range Finder Testing.....	C-48
Figure 2-3	96-Well Test Plate Layout for Comprehensive Agonist Testing	C-49
Figure 2-4	96-Well Test Plate Layout for Comprehensive Antagonist Testing.....	C-52
Figure 5-1	Relationship of EC ₅₀ Values Obtained in the BG1Luc ER TA Test Method and EC ₅₀ Values from ICCVAM Reference Data	C-120
Figure 5- 2	Relationship of IC ₅₀ Values Obtained in the BG1Luc ER TA Test Method and IC ₅₀ Values from ICCVAM Reference Data.....	C-120
Figure 6-1	Agonist DMSO Control Within-Plate RLU Values During Phase 1 at XDS.....	C-126
Figure 6-2	Agonist DMSO Control Within-Plate RLU Values During Phase 1 at ECVAM	C-127
Figure 6-3	Agonist DMSO Control Within-Plate RLU Values During Phase 1 at Hiyoshi	C-127
Figure 6-4	Agonist E2 Reference Standard EC ₅₀ Values	C-129
Figure 6-5	Agonist Methoxychlor Control Values.....	C-130
Figure 6-6	Antagonist DMSO Control Within-Plate RLU Values During Phase 1 at XDS.....	C-132
Figure 6-7	Antagonist DMSO Control Within-Plate RLU Values During Phase 1 at ECVAM	C-133
Figure 6-8	Antagonist DMSO Control Within-Plate RLU Values During Phase 1 at Hiyoshi ..	C-134
Figure 6-9	Antagonist Raloxifene Reference Standard IC ₅₀ Values	C-136
Figure 6-10	Antagonist Flavone Control Values	C-137
Figure 6-11	Antagonist E2 Control Values	C-138
Figure 6-12	Fluoranthene Results at All Three Laboratories: Impact of Differences in Solubility on Comprehensive Test Results	C-142
Figure 6-13	Resveratrol Results at XDS and Hiyoshi: Impact of Selecting the Incorrect Starting Concentration Based on Range Finder Results	C-143
Figure 7-1	Inadequate Test Results: Tamoxifen Tested at XDS and ECVAM.....	C-152

List of Abbreviations and Acronyms

AR	Androgen receptor
ATP	Adenosine triphosphate
BPA	Bisphenol A
BPB	Bisphenol B
BRD	Background review document
CASRN	CAS Registry Number [®] (a trademark of the American Chemical Society)
CERI	Chemicals Evaluation and Research Institute, Japan
CV	Coefficient of variation
DES	Diethylstilbestrol
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
E2	17 β -Estradiol
EAC	Endocrine-active compound
EC ₅₀	Half-maximal effective concentration
ECVAM	European Centre for the Validation of Alternative Methods
ED	Endocrine disruptor
EDSP	Endocrine Disruptor Screening Program (U.S. EPA)
EDSTAC	EPA Endocrine Disruptor Screening and Testing Advisory Committee (U.S. EPA)
EDTA	Endocrine Disruptors Testing and Assessment (OECD)
EDWG	ICCVAM Interagency Endocrine Disruptor Working Group
EEC	European Economic Community
EFM	Estrogen-free media
EPA	U.S. Environmental Protection Agency
ER	Estrogen receptor
ERE	Estrogen response element
FBS	Fetal bovine serum
FDA	U.S. Food and Drug Administration
FR	<i>Federal Register</i>
G418	Gentamycin
GLP	Good Laboratory Practice
I	Inadequate
IC	Inconclusive
IC ₅₀	Half-maximal inhibitory concentration
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
ILS	Integrated Laboratory Systems, Inc.
ISO	International Organization for Standardization
JaCVAM	Japanese Center for the Validation of Alternative Methods

KoCVAM	Korean Center for the Validation of Alternative Methods
LEC	Lowest effective concentration
M	Molar
MEM	Minimum essential medium
MeSH®	Medical Subject Headings (U.S. National Library of Medicine)
Met	Methoxychlor
MMTV	Mouse mammary tumor virus
MPA	Medroxyprogesterone acetate
MSDS	Material Safety Data Sheet
MTD	Maximum tolerated dose
N	Negative; number
NA	Not applicable
NC	Not calculated
NCGC	NIH Chemical Genomics Center (U.S. National Institutes of Health)
NEG	Negative
NICEATM	National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods
NIEHS	U.S. National Institute of Environmental Health Sciences
NIH	U.S. National Institutes of Health
NR	Not reviewed
NT	Not tested
NTP	U.S. National Toxicology Program (U.S. NIH)
NTPSI	National Toxicology Program Substances Inventory
OECD	Organisation for Economic Co-operation and Development
OPPTS	Office of Prevention, Pesticides and Toxic Substances (U.S. EPA)
P	Positive
PBS	Phosphate-buffered saline
Pen-Strep	Penicillin-streptomycin
PN	Presumed negative
POS	Positive
PP	Presumed positive
QA	Quality assurance
Ral	Raloxifene
REACH	Registration, Evaluation, Authorisation and Restriction of Chemicals
RLU	Relative light unit
SACATM	Scientific Advisory Committee on Alternative Toxicological Methods
SD	Standard deviation
SMT	Study Management Team
SOP	Standard operating procedure

STTA	Stably transfected human estrogen receptor- α transcriptional activation
TA	Transcriptional activation
Tam/TAM	Tamoxifen
TG	Test Guideline
TS	Test substance
U.S.C.	United States Code
WHO	World Health Organization
XDS	Xenobiotic Detection Systems, Inc.
VC	Vehicle control

Trademarks are the property of their respective owners.

Interagency Coordinating Committee on the Validation of Alternative Methods: Agency Representatives

Agency for Toxic Substances and Disease

Registry

- * Moiz Mumtaz, PhD
- Edward Murray, PhD
- Eric Sampson, PhD
- Bruce Fowler, PhD (to June 2011)

Consumer Product Safety Commission

- * Joanna Matheson, PhD (Vice-Chair)
- + Kristina Hatlelid, PhD
- Marilyn L. Wind, PhD (Chair, to July 2010)

Department of Agriculture

- * Jodie Kulpa-Eddy, DVM (Chair)
- + Elizabeth Goldentyer, DVM

Department of Defense

- * Patrick Mason, PhD
- + Terry Besch, DVM, MS, DACLAM, DACVPM
- + Patty Decot
- David Honey, PhD (to July 2011)
- Robert E. Foster, PhD (to Aug. 2010)
- Harry Salem, PhD (to Aug. 2010)
- Peter J. Schultheiss, DVM, DACLAM (to Aug. 2010)

Department of Energy

- * Michael Kuperberg, PhD
- Marvin Stodolsky, PhD (retired 2011)

Department of the Interior

- * Barnett A. Rattner, PhD
- Sarah Gerould, PhD (to Feb. 2009)

Department of Transportation

- + Steve Hwang, PhD
- George Cushmac, PhD (retired 2010)

Environmental Protection Agency

Office of Pesticide Programs

- * John R. “Jack” Fowle III, PhD, DABT
- + Vicki Dellarco, PhD
- Christine Augustyniak, PhD (*U.S. Coordinator, OECD Test Guidelines Program*) (to Sept. 2011)
- Tina Levine, PhD (to Sept. 2011)
- Deborah McCall (to Sept. 2011)

Office of Pollution Prevention and Toxics

- Jerry Smrchek, PhD (*U.S. Coordinator, OECD Test Guidelines Program*) (to July 2009)

Environmental Protection Agency (cont.)

Office of Research and Development

- Stephanie Padilla, PhD (to July 2009)
- Julian Preston, PhD (to July 2009)
- Suzanne McMaster, PhD (to Dec. 2008)

Office of Science Coordination and Policy

- Karen Hamernik, PhD (to July 2009)

Food and Drug Administration

Office of the Commissioner

- * Suzanne Fitzpatrick, PhD, DABT

Center for Biologics Evaluation and Research

- Ying Huang, PhD
- Richard McFarland, PhD, MD

Center for Devices and Radiological Health

- Vasant G. Malshet, PhD, DABT

Center for Drug Evaluation and Research

- + Abigail C. Jacobs, PhD
- Paul C. Brown, PhD

Center for Food Safety and Applied Nutrition

- Robert Bronaugh, PhD
- David G. Hattan, PhD
- Diego Rua, PhD
- Neil Wilcox, DVM, MPH (to May 2011)

Center for Veterinary Medicine

- M. Cecilia Aguila, DVM
- Li You, PhD
- Devaraya Jagannath, PhD (to April 2011)

National Center for Toxicological Research

- Paul Howard, PhD
- Donna L. Mendrick, PhD
- William T. Allaben, PhD (to Jan. 2009)

Office of Regulatory Affairs

- Lawrence D’Hoostelaere, PhD

National Cancer Institute

- * T. Kevin Howcroft, PhD
- + Chand Khanna, DVM, PhD
- Alan Poland, MD (to Oct. 2008)

continued

- * Principal agency representative
- + Alternate principal agency representative

National Institute of Environmental Health Sciences

- * William S. Stokes, DVM, DACLAM
- + Warren Casey, PhD, DABT
- Rajendra S. Chhabra, PhD, DABT
- Jerrold J. Heindel, PhD
- Raymond R. Tice, PhD (to June 2010)

National Institute for Occupational Safety and Health

- * Paul Nicolaysen, VMD

National Institutes of Health

- * Margaret D. Snyder, PhD

National Library of Medicine

- * Pertti (Bert) Hakkinen, PhD
- + Jeanne C. Goshorn, MS

Occupational Safety and Health Administration

- * Surender Ahir, PhD
- * Principal agency representative
- + Alternate principal agency representative

Acknowledgements

ICCVAM Interagency Endocrine Disruptor Working Group (EDWG)

U.S. Consumer Product Safety Commission

Kent Carlson, PhD

U.S. Department of the Interior

Catherine Richter, PhD

Donald Tillitt, PhD

U.S. Environmental Protection Agency

Don Bergfelt, PhD

John R. “Jack” Fowle, III, PhD, DABT

Jesudoss Rowland

U.S. Food and Drug Administration

Center for Devices and Radiological Health

Thomas Umbreit, PhD

Center for Drug Evaluation and Research

Jeffrey Bray, PhD

Paul C. Brown, PhD

Karen L. Davis-Bruno, PhD

Ross Filice, MD

Abigail (Abby) Jacobs, PhD

Leslie McKinney, PhD

Center for Food Safety and Nutrition

Michael Bolger, PhD, DABT

David G. Hattan, PhD (Chair)

Center for Veterinary Medicine

M. Cecilia Aguila, DVM

Charles Eirkson, PhD

Kevin Gaido, PhD

Devaraya Jagannath, PhD

Annette McCarthy, PhD

Li You, PhD

National Center for Toxicological Research

Kenneth Delclos, PhD

Huixiao Hong, PhD

Jon Wilkes, PhD

Office of the Commissioner

Suzanne Fitzpatrick, PhD, DABT

National Institute of Environmental Health Sciences

Warren Casey, PhD, DABT

Jerrold Heindel, PhD

William Stokes, DVM, DACLAM

Julius Thigpen, PhD

Occupational Safety and Health Administration

Surender Ahir, PhD

European Centre for the Validation of Alternative Methods – Liaison

Susanne Bremer, PhD

Elise Grignard, PhD

Valerie Zuang, PhD

Japanese Center for the Validation of Alternative Methods – Liaison

Hajime Kojima, PhD

Atsushi Ono, PhD

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

National Institute of Environmental Health Sciences

William Stokes, DVM, DACLAM
Director (NICEATM); Project Officer

Warren Casey, PhD, DABT
Deputy Director

Deborah McCarley
Special Assistant; Assistant Project Officer

NICEATM Support Contract Staff (Integrated Laboratory Systems [ILS], Inc.)

David Allen, PhD

Patricia Ceger, MS

Frank Deal, MS (until March 2011)

Jonathan Hamm, PhD

Linda Litchfield

Steven Morefield, MD

Michael Paris

Catherine Sprankle

Linda Wilson

Statistical Consultant for ILS, Inc.

Joseph Haseman, PhD

Participants in the BG1Luc ER TA Validation Study

NICEATM gratefully acknowledges the generous contributions of the individuals who directly participated in the NICEATM/ECVAM/JaCVAM/KoCVAM Validation Study. Their time and efforts are greatly appreciated.

Xenobiotic Detection Systems, Inc. (XDS)

(Lead Laboratory)

George Clark, PhD – Study Director from 2010
John Gordon, PhD – Study Director until 2008
Jun Liu, PhD – Study Director until 2009
Ruth McMillan, PhD- Technical Lead from 2010
Andrew Chu – Laboratory Manager until 2009
Ying Gu – Lead Chemist

ECVAM Laboratory (In Vitro Methods Unit) **(Participating Laboratory)**

Miriam Jacobs, PhD – Study Director until 2008
Patricia Pazos, PhD – Study Director from 2009
Jan de Lange – Technical Lead

Hiyoshi Corporation

(Participating Laboratory)

Masafumi Nakamura – Study Director
Hiyorhi Handa – Technical Lead
Clevita Eldrida Caesar Mascalenhas – Translator

Statistical Consultant

Joseph Haseman, PhD

NIEHS (Chemistry Resources Group)

Cynthia Smith, PhD – Branch Chief
Beby Jayaram, PhD – Chemist

Study Management Team (SMT)

NICEATM

William Stokes, DVM, DACLAM (NIEHS) –
Director, NICEATM
Warren Casey, PhD (NIEHS) – Deputy Director,
NICEATM
David Allen, PhD (ILS, Inc.) – Principal
Investigator
Frank Deal, MS (ILS, Inc.) – Staff Toxicologist
until March 2011
Patricia Ceger, MS (ILS, Inc.) – Project
Coordinator

ECVAM

Elise Grignard, PhD
Susanne Bremer, PhD

JaCVAM

Hajime Kojima, PhD – Director, JaCVAM
Atsushi Ono, PhD – National Institute of Health
Sciences

KoCVAM

Soon Young Han, PhD – Director, KoCVAM

This page intentionally left blank

Preface

Endocrine-active compounds (EACs) are both naturally occurring and synthetic substances. Some may, depending on the dose, interfere with the normal function of hormones in the endocrine system. In response to growing concerns about possible adverse health effects in humans exposed to such substances, sometimes referred to as endocrine disruptors (EDs), the U.S. Congress enacted relevant provisions in the Food Quality Protection Act of 1996 (7 U.S.C. 136) and the 1996 Amendments to the Safe Drinking Water Act (110 Stat 1613). In 1998, the U.S. Environmental Protection Agency (EPA) established the Endocrine Disruptor Screening Program (EDSP), a screening and testing program to identify substances with endocrine-disrupting activity.

In 2000, the EPA requested that the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) review the validation status of four types of test methods:

- Estrogen receptor (ER) binding test methods
- Androgen receptor (AR) binding test methods
- Estrogen receptor transcriptional activation (ER TA) test methods
- Androgen receptor transcriptional activation (AR TA) test methods

In 2002, NICEATM prepared background review documents (BRDs) that included all available information on each of the four types of test methods (ICCVAM 2002d, 2002a, 2002c, 2002b). In a public meeting, an independent international expert panel (Panel) reviewed the 137 individual assays identified in the BRDs and concluded that there were no adequately validated *in vitro* ER- or AR-based test methods.

In 2004, Xenobiotic Detection Systems, Inc. (XDS; Durham, NC), nominated their LUMI-CELL® ER test method (BG1Luc ER TA test method) to ICCVAM for an interlaboratory validation study. This *in vitro* test method uses BG-1 cells, a human ovarian carcinoma cell line that is stably transfected with an estrogen-responsive luciferase reporter gene, to measure whether and to what extent a substance induces or inhibits TA activity via ER-mediated pathways. ICCVAM considered the nomination a high priority. NICEATM subsequently coordinated an international validation study for the BG1Luc ER TA test method. Scientists from the European Centre for the Validation of Alternative Methods (ECVAM) and the Japanese Center for the Validation of Alternative Methods (JaCVAM) served as liaisons to the ICCVAM Interagency Endocrine Disruptor Working Group (EDWG).

Following completion of the validation study, NICEATM, ICCVAM, and the EDWG prepared (1) a draft BRD that detailed the results of the validation study and described the validation status of the BG1Luc ER TA test method and (2) draft test method recommendations for usefulness and limitations, standardized protocols, future studies, and performance standards. ICCVAM released these documents to the public for comment prior to a meeting of an independent international scientific peer review panel (Panel). The Panel met in public session on March 29–30, 2011, and later prepared a report summarizing its conclusions and recommendations (ICCVAM 2011). The Panel report was provided to the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) along with the draft BRD, draft test method recommendations, and all public comments. A detailed timeline of the evaluation is included with this report.

ICCVAM solicited and considered public comments and stakeholder involvement throughout the test method evaluation process. ICCVAM considered the SACATM comments, the conclusions of the Panel, and all public comments before finalizing the ICCVAM test method recommendations. The recommendations and performance standards are incorporated in this ICCVAM test method evaluation report, and the BRD is provided as an appendix.

This test method evaluation report provides ICCVAM's recommendations regarding the usefulness and limitations of the BG1Luc ER TA test method for identifying potential agonist or antagonist substances that might interfere with normal estrogen activity. The report also summarizes the validation status of the BG1Luc ER TA test method and provides the ICCVAM-recommended protocols, future studies, and performance standards. As required by the ICCVAM Authorization Act (42 U.S.C. 285f-3), ICCVAM will forward its recommendations to U.S. Federal agencies for their consideration. Federal agencies must respond to ICCVAM within 180 days after receiving the ICCVAM test method recommendations. The ICCVAM report and recommendations are available to the public on the NICEATM-ICCVAM website (<http://iccvam.niehs.nih.gov/>). Agency responses will be made available on the website as they are received.

We gratefully acknowledge the organizations and scientists who generated and provided data and information for this document, including the staff at the participating validation laboratories: XDS, Inc.; Hiyoshi Corporation; and the In Vitro Methods Unit at ECVAM. We especially recognize the Panel members for their thoughtful evaluations and generous contributions of time and effort. Special thanks are extended to Dr. John Vandenberg for serving as the Panel Chair and to Drs. Christopher Borgert, William Kelce, Steven Levine, and Ellen Mihaich for their service as Evaluation Group Chairs for the Panel. We thank the EDWG for assuring a meaningful and comprehensive review, especially Dr. David Hattan (U.S. Food and Drug Administration Center for Food Safety and Nutrition) for serving as Chair of the EDWG. Integrated Laboratory Systems, Inc., the NICEATM support contractor, provided excellent scientific support, for which we thank Drs. David Allen, Jon Hamm, and Steven Morefield; Patricia Ceger; Frank Deal (until March 2011); Linda Litchfield; Michael Paris; Catherine Sprinkle; and Linda Wilson. We thank Drs. Susanne Bremer and Elise Grignard, the EDWG liaisons from ECVAM, and Drs. Hajime Kojima and Atsushi Ono, the EDWG liaisons from JaCVAM, for their participation and support. Finally, we thank Dr. Warren Casey for leading the analysis of the validation study results and coordinating the ICCVAM test method evaluation.

RADM William S. Stokes, DVM, DACLAM
Rear Admiral/Assistant Surgeon General, U.S. Public Health Service
Director, NICEATM
Executive Director, ICCVAM

Jodie A. Kulpa-Eddy, DVM
VS-NCIE-Agricultural Select Agent Program
U.S. Department of Agriculture
Chair, ICCVAM

Executive Summary

Background

In April 2000, the U.S. Environmental Protection Agency (EPA) nominated four types of *in vitro* test methods for detecting substances with potential to interfere with the normal function of hormones in the endocrine system (i.e., endocrine disruptors [EDs]) for review by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM). ICCVAM then recommended that these methods undergo independent scientific peer review based on their potential interagency applicability and public health significance. The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) compiled available data and information on the four types of *in vitro* test methods: estrogen receptor (ER) binding, androgen receptor (AR) binding, ER transcriptional activation (TA), and AR TA test methods. ICCVAM, the ICCVAM Interagency Endocrine Disruptor Working Group (EDWG), and NICEATM prepared four background review documents (BRDs) that detailed the available data and information needed to evaluate the current validation status of each of the four types of test methods (ICCVAM 2002d, 2002a, 2002c, 2002b).

In collaboration with ICCVAM and the EDWG, NICEATM organized an independent evaluation of these *in vitro* test methods. ICCVAM considered the international panel's conclusions and recommendations, as well as public comments. ICCVAM then developed test method recommendations that included minimum procedural standards and a list of 78 reference substances that should be used to standardize and validate *in vitro* ER and AR binding and TA test methods (ICCVAM 2003a).

In January 2004, Xenobiotic Detection Systems, Inc. (XDS; Durham, NC), nominated their LUMI-CELL® BG1Luc4E2 ER TA test method (hereafter BG1Luc ER TA test method) for an interlaboratory validation study. This method uses BG-1 cells (a human ovarian adenocarcinoma cell line) that are stably transfected with an estrogen-responsive luminescence (luciferase reporter) gene to measure whether and how much a substance induces (agonist) or inhibits (antagonist) TA activity via ER-mediated pathways (Rogers and Denison 2000; Rogers and Denison 2002). XDS included test results for 56 of the 78 ICCVAM reference substances for agonist activity and 6 of the 78 ICCVAM reference substances for antagonist activity. These studies were funded primarily by a Small Business Innovation Research grant (SBIR43ES010533-01) from the National Institute of Environmental Health Sciences (NIEHS).

ICCVAM considered the BG1Luc ER TA test method to be a high priority for interlaboratory validation studies, and the NIEHS agreed to support the effort. NICEATM led and coordinated an international interlaboratory validation study with its counterparts at the European Centre for the Validation of Alternative Methods (ECVAM) and the Japanese Center for the Validation of Alternative Methods. The BG1Luc ER TA test method was evaluated in four phases, during which the 78 ICCVAM-recommended substances were tested at laboratories in the United States (XDS), Europe (ECVAM), and Japan (Hiyoshi Corporation).

NICEATM, in conjunction with the EDWG, prepared this BRD, which summarizes the available data and information regarding the current validation status of the BG1Luc ER TA test method.

BG1Luc ER TA Test Method Protocol

The BG1Luc ER TA test method uses an ER-responsive luminescence (luciferase reporter) gene (*luc*) in the human ovarian adenocarcinoma cell line BG-1 to detect substances with *in vitro* ER agonist or antagonist activity. To help define the upper limit for test substance concentrations, scores for cell viability are assigned using visual observation of numbers (density) and shapes (morphology) of cells. ER-mediated transcription of the *luc* gene produces the luminescence enzyme luciferase, which

catalyzes the production of light from luciferin. The light is measured using a luminometer. In accordance with earlier ICCVAM recommendations, 17 β -estradiol (E2) is used as the estrogen reference standard for agonist tests, and raloxifene is used as the anti-estrogenic reference standard for antagonist tests to demonstrate the adequacy of the BG1Luc ER TA test method. To provide qualitative and quantitative information regarding the *in vitro* estrogenic activity of a test substance, a concentration–response curve is established. To determine if a test substance is positive or negative for ER agonism or antagonism, criteria associated with the concentration–response curve are used. The advantages of using a luciferase reporter gene system are low background, high sensitivity, speed, and a wide dynamic range.

Substances Used in the Validation Study

To assess the performance of four different test methods (ER TA and AR TA agonist and antagonist assays), ICCVAM developed a list of 78 recommended reference substances based on a review of the literature. Only those substances that could be definitively classified as positive or negative for ER TA activity (48 unique substances) were used to assess overall accuracy of the test method. Separate lists were generated to evaluate accuracy of the test methods for activities of agonists (42 substances: 33 positive, 9 negative) and antagonists (25 substances: 3 positive, 22 negative).

BG1Luc ER TA Test Method Accuracy

The BG1Luc ER TA test method was evaluated for its ability to correctly identify ER agonists and antagonists. The BG1Luc ER TA test method was evaluated for accuracy based on a number of analyses, but the primary evaluation was based on two comparisons: (1) the extent to which the BG1Luc ER TA result corresponded to the ICCVAM reference classification for each substance and (2) the accuracy of the BG1Luc ER TA test method compared with the CERI-STTA (Chemicals Evaluation and Research Institute Stably Transfected Human Estrogen Receptor- α Transcriptional Activation) assay (OECD 2009).¹ The positive or negative classification of the BG1Luc ER TA test result for individual substances was assigned based on the majority result from the three participating laboratories (XDS, ECVAM, and Hiyoshi).

Of the 42 substances used to evaluate agonist accuracy, 17% (7/42) had “inadequate” testing results in the BG1Luc ER TA test method and were therefore excluded from the analysis. The remaining 35 substances (28 positive, 7 negative) were used for evaluation. The BG1Luc ER TA test method produced the following results compared to the reference classifications for these 35 substances: concordance of 97% (34/35), sensitivity of 96% (27/28), specificity of 100% (7/7), a false positive rate of 0% (0/7), and a false negative rate of 4% (1/28).

The CERI-STTA assay is the only *in vitro* ER TA test method currently accepted by U.S. regulatory agencies for ER agonist testing. (No ER antagonist test methods are currently accepted by U.S. regulatory agencies.) BG1Luc ER TA and CERI-STTA data show identical levels of accuracy when the same 26 agonist reference chemicals were tested: concordance of 96% (25/26), sensitivity of 95% (21/22), specificity of 100% (4/4), a false positive rate of 0% (0/4), and a false negative rate of 5% (1/22).

All 25 of the antagonist reference substances produced definitive results in the BG1Luc ER TA test method and yielded an overall concordance of 100% (25/25), sensitivity of 100% (3/3), specificity of 100% (22/22), a false positive rate of 0% (0/22), and a false negative rate of 0% (0/3).

Although the primary goal of the BG1Luc ER TA test method is to provide a qualitative assessment of estrogenic/anti-estrogenic activity, quantitative measures of activity (i.e., half-maximal effective [EC₅₀] and half-maximal inhibitory concentration [IC₅₀] values) are usually obtained for positive

¹ The CERI-STTA assay (OECD 2009) uses the hER α -HeLa-9903 human cervical cancer cell line to detect estrogenic agonist activity mediated through human estrogen receptor alpha (hER α).

results. EC₅₀ and IC₅₀ values obtained from BG1Luc ER TA test results were highly correlated with median values from other ER TA test methods reported in the literature. BG1Luc ER TA test results also showed 97% (33/34) concordance with the ICCVAM reference classifications. The only discordant substance was positive in the BG1Luc ER TA test method and negative based on ER binding data. Similarly, BG1Luc ER TA agonist test results showed 92% (12/13) concordance with available data from the *in vivo* uterotrophic assay. The only discordant substance was positive in the BG1Luc ER TA test method and negative based on uterotrophic data.

BG1Luc ER TA Test Method Reliability

Intralaboratory Reproducibility

Intralaboratory reproducibility (whether multiple tests of the same substance at a single laboratory produce the same results) of the BG1Luc ER TA agonist and antagonist test methods was assessed by comparing reference standard and control results for all plates tested within each laboratory during the course of the validation study.

In the agonist test method, mean induction in each laboratory ranged from 4.6 to 7.8 fold, and E2 reference standard EC₅₀ values ranged from 8.0×10^{-12} to 1.1×10^{-11} M. In the antagonist test method, mean reduction ranged from 8.0 to 9.9 fold, and raloxifene reference standard IC₅₀ values ranged from 1.1×10^{-9} to 1.3×10^{-9} M.

Intralaboratory reproducibility for positive or negative classification was determined for each of the 12 substances that were tested at least three times for agonist and antagonist activity during Phase 2 at each of the three laboratories. There was 100% agreement within each laboratory for each of the three repeat tests, for both agonists and antagonists, although the classifications for some of the test substances differed among the different laboratories.

Interlaboratory Reproducibility

Interlaboratory reproducibility (whether tests of a single substance run at different laboratories produce the same results) was determined using results from Phase 2 testing, during which 12 substances were tested in at least three independent experiments for agonist and antagonist activity in each of the three laboratories. The three laboratories had 67% agreement (8/12) for agonist activity and 100% agreement (12/12) for antagonist activity.

Interlaboratory reproducibility was also determined for 41 substances that were tested once for agonist and antagonist activity during Phase 3 testing at each of the three laboratories. Five of the 41 substances produced inadequate results for agonist activity and could not be considered in the evaluation. Among the 36 remaining substances that produced a definitive test result in at least two laboratories, there was 100% agreement. All 41 substances produced definitive results for antagonist activity. The three laboratories agreed on 93% (38/41) of these substances.

Animal Welfare Considerations

The BG1Luc ER TA test method may be applicable for addressing the ER TA component of the U.S. EPA Endocrine Disruptor Screening Program (EDSP) Tier 1 screening battery. Although the EDSP currently includes an *in vitro* ER TA test method for ER agonist testing (i.e., the CER1-STTA method), currently no *in vitro* test methods are accepted for ER antagonist testing. Therefore, the BG1Luc ER TA test method provides an opportunity to reduce animal use in ED testing by identifying both ER agonist and antagonist substances. This information can be used as part of a weight-of-evidence approach to prioritize substances for additional investigation of ED activity in test methods that require animals.

Regulators currently use the following three *in vivo* methods to assess the estrogenic potential of substances: rat uterotrophic assay, rat pubertal female assay, and fish short-term reproduction assay.

An additional test, the “*in vitro*” rat uterine cytosol ER binding assay, also requires the use of animals as a source of ER. Although the BG1Luc ER TA test method is not proposed as a direct replacement for any of these existing methods, it could be incorporated as part of a weight-of-evidence approach to reduce or eliminate the need to use animals for identifying substances with potential estrogenic or anti-estrogenic activity.

Test Method Transferability

Transferability of the BG1Luc ER TA test method was demonstrated based on results of the interlaboratory validation study detailed above. The primary practical considerations associated with the BG1Luc ER TA test method are the availability of the requisite cell line and the standard laboratory equipment necessary to conduct sterile cell culture procedures. The BG1Luc4E2 cell line is available upon request from Dr. Michael S. Denison, Department of Environmental Toxicology, University of California, Davis. The level of training, expertise, and time needed to conduct the BG1Luc ER TA test method should be similar to that needed for the currently accepted CERI-STTA method.

ICCVAM Test Method Recommendations

ICCVAM considered the data and analysis provided in this BRD and developed recommendations on the usefulness and limitations of the BG1Luc ER TA test method as a screening test to identify substances with ER agonist and antagonist activity. ICCVAM also developed recommendations for a standardized test method protocol, proposed future studies, and performance standards.

1.0 Introduction

1.1 Objective

The objective of this validation study was to assess the accuracy and reliability of the BG1Luc4E2 estrogen receptor (ER) transcriptional activation (TA) test method (hereafter BG1Luc ER TA test method) for the qualitative detection of substances with ER agonist or antagonist activity.

1.2 Public Health Perspective

Endocrine disruptors (EDs) interfere with the function of hormones in the endocrine system, which can lead to abnormal growth, development, or reproduction (Ankley et al. 1998; Baker 2001; Brown et al. 2001; Combes 2000; Greim 2004; Kavlock 1999). Potential EDs are widespread in our environment and include both synthetic (e.g., pesticides, pharmaceuticals, industrial chemicals) and naturally occurring (e.g., plant products known as *phytoestrogens*) substances. A number of studies have indicated that animal populations exposed to high levels of these substances have an increased incidence of reproductive and developmental abnormalities (Guillette and Gunderson 2001; Kelly et al. 2009; Rozman et al. 2006; Segner 2005; Soin and Smagghe 2007; Sormo et al. 2006; Tyler et al. 1998).

Exposure of humans to EDs is also linked to adverse health outcomes such as altered reproduction and immune function, increased incidence of cancer, and increased incidence of obesity and associated complications such as cardiovascular disease and type 2 diabetes (Kavlock et al. 2006; Rozman et al. 2006; Tsai 2006; Whitten and Naftolin 1992; Whitten et al. 1992; Whitten et al. 1995; Whitten and Naftolin 1998; Whitten and Patisaul 2001). In light of the growing concern surrounding this important issue, the accurate and timely identification of potential EDs by the BG1Luc ER TA test method is an important aspect of protecting public health.

1.3 Historical Background

The Federal Food, Drug, and Cosmetic Act (21 U.S.C. 301 et seq.); the Food Quality Protection Act (7 U.S.C. 136); and the Safe Drinking Water Act (110 Stat 1613) all aim to identify potential endocrine disruptors and thereby protect humans and animals. The U.S. Environmental Protection Agency (EPA) was specifically required to “develop a screening program, using appropriate validated test systems and other scientifically relevant information, to determine whether certain substances may have an effect in humans that is similar to an effect produced by a naturally occurring estrogen, or such other endocrine effect as the Administrator may designate” (21 U.S.C. 346a[p][1]). In 1996, the EPA formed the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC). This committee of scientists and stakeholders was charged by the EPA to provide recommendations on how to implement the EPA’s Endocrine Disruptor Screening Program (EDSP).

In 1998, EDSTAC proposed a two-tier screening program (63 FR 71542), and the EPA accepted the recommendation. Tier 1 consists of *in vivo* and *in vitro* test methods. Its purpose is to identify the potential of chemicals to interact with the estrogen, androgen, or thyroid hormonal systems. A negative result in Tier 1 can signify minimal potential to cause endocrine disruption. A positive result necessitates further testing using *in vivo* methods in Tier 2. The purpose of Tier 2 is to more definitively identify and characterize the potential hazard to the endocrine system. Results from Tier 2 testing can also be used in required risk assessment to further evaluate the potential for adverse health effects from exposure to the chemicals. The EPA describes the EDSP in detail on their website at <http://www.epa.gov/scipoly/oscpendo/>.

In April 2000, the EPA nominated four types of *in vitro* test methods for review by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM):

- Estrogen receptor (ER) binding test methods
- Androgen receptor (AR) binding test methods
- Estrogen receptor transcriptional activation (ER TA) test methods
- Androgen receptor transcriptional activation (AR TA) test methods

These types of test methods detect substances that may cause endocrine disruption (Combes 2000). In 2001, ICCVAM recommended that these test methods should undergo independent scientific peer review based on their potential interagency applicability and public health significance. In response, the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) compiled four separate comprehensive background review documents (BRDs) that included all available information on each of the four types of test methods (ICCVAM 2002d, 2002a, 2002c, 2002b).

In 2001, NICEATM collaborated with ICCVAM and the ICCVAM Interagency Endocrine Disruptor Working Group (EDWG) to organize an independent international peer review panel (Panel) meeting to assess the suitability of the 137 individual *in vitro* test methods identified in the BRDs. The Panel reviewed the information and draft ICCVAM recommendations and concluded that there were no adequately validated *in vitro* ER- or AR-based test methods. The Panel detailed their conclusions and recommendations in a final report (ICCVAM 2002e).

ICCVAM considered the Panel's conclusions and recommendations and all comments received.² ICCVAM then published test method recommendations for minimum essential test method components, along with a list of 78 ICCVAM reference substances that should be used to standardize and validate *in vitro* ER and AR binding and TA test methods (ICCVAM 2003a, 2006). Based on the lack of adequately validated test methods, coupled with the public health issues identified above, ICCVAM and the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) recommended the validation of *in vitro* endocrine disruptor screening methods as a high-priority activity (69 FR 54298).

1.4 Nomination and Pre-Screen Evaluation of the BG1Luc4E2 ER TA Test Method

In January 2004, Xenobiotic Detection Systems, Inc. (XDS), nominated the LUMI-CELL® BG1Luc ER TA test method for an interlaboratory validation study (**Annex A**). This method uses BG-1 cells, a human ovarian carcinoma cell line that is stably transfected with an estrogen-responsive luciferase reporter gene, to measure whether and to what extent a substance induces or inhibits TA activity via ER-mediated pathways (Rogers and Denison 2000; Rogers and Denison 2002). The nomination package included test results from XDS for 56 of the 78 ICCVAM reference substances for agonist activity and 16 of the 78 ICCVAM reference substances for antagonist activity. These studies were funded primarily by a Small Business Innovation Research (SBIR) grant (SBIR43ES010533-01) from the National Institute of Environmental Health Sciences (NIEHS).

In accordance with the ICCVAM nomination process (ICCVAM 2003b), NICEATM conducted a pre-screen evaluation of the nomination package (**Annex B**) to determine (1) the extent to which it addressed the ICCVAM prioritization criteria (**Section 1.5**) and (2) how well it adhered to the ICCVAM recommendations for the standardization and validation of *in vitro* endocrine disruptor test methods (ICCVAM 2003b). Based on this evaluation, ICCVAM recommended the following:

² Text of comments available at <http://ntp-apps.niehs.nih.gov/iccvampb/searchPubCom.cfm?ftitle=02-26733>.

- The BG1Luc ER TA test method should be considered a high priority for interlaboratory validation studies as an *in vitro* test method to detect test substances with ER agonist and antagonist activity.
- Validation studies should include coordination and collaboration with the European Centre for the Validation of Alternative Methods (ECVAM) and the Japanese Center for the Validation of Alternative Methods (JaCVAM). Studies should include one laboratory in each of the three respective geographic regions (United States, Europe, Japan).
- In preparation for the interlaboratory validation study, XDS should conduct protocol standardization studies with an emphasis on filling data gaps in the antagonist protocol for the BG1Luc ER TA test method.

The mission of the National Toxicology Program (NTP) includes the development and validation of improved testing methods. As one of three NTP agencies, the NIEHS agreed to support the validation study.

1.5 Basis for High Priority for Validation Studies

NICEATM performs preliminary evaluations of all test method submissions and nominations and summarizes the extent to which the test methods meet the five ICCVAM prioritization criteria (ICCVAM 2003b). As noted in **Section 1.4**, ICCVAM assigned a high priority to conducting an interlaboratory validation study for the BG1Luc ER TA test method. This section details the rationale for this prioritization and summarizes more-recent national and international developments that further emphasize the need to develop and validate *in vitro* ER TA test methods like the BG1Luc ER TA test method.

1.5.1 Criterion 1

In keeping with ICCVAM's prioritization criteria, NICEATM evaluates and summarizes the extent to which test methods are applicable to (1) regulatory testing needs and (2) multiple agencies/programs (ICCVAM 2003b).

The EPA EDSP Tier 1 screening battery currently includes an ER TA test method, OPPTS 890.1300: Estrogen Receptor Transcriptional Activation (Human Cell Line [HeLa-9903]) (EPA 2009). The screening battery also provides for the use of other scientifically valid methods. Therefore, the BG1Luc ER TA test method may be applicable for addressing the ER TA component of the EPA EDSP Tier 1 screening battery.

The NTP conducted the major health review of bisphenol A (BPA) that prompted widespread reconsideration of its use by industry and the introduction of such alternative products as the BPA-free water bottle. Over the past decade, the NIEHS has made a substantial investment in research that focuses on endocrine disruptors. Endocrine disruption continues to be a focal point in NIEHS studies of commercial products that are in wide use, such as flame retardants and pesticides.

The high-throughput evaluation of chemicals, which allows rapid screening of many substances, is an important aspect of many research and testing programs within government and industry. The BG1Luc ER TA test method is currently being evaluated by the National Institutes of Health (NIH) Chemical Genomics Center (NCGC) for its adaptability to a high-throughput screening format, which could be used to support high-throughput screening and testing programs.

In response to requests by the U.S. House of Representatives and Senate Appropriations Committee, NICEATM and ICCVAM published a Five-Year Plan to do the following:

- Research, develop, translate, and validate new and revised non-animal and other alternative assays for integration of relevant and reliable methods into Federal agency testing programs

- Identify areas of high priority for new and revised non-animal and alternative assays or batteries of those assays to create a path forward for the reduction, replacement, and refinement of animal tests when this is scientifically valid and appropriate (ICCVAM 2008; Poland et al. 2008; Stokes and Wind 2009)

The evaluation of test methods for identifying endocrine-disrupting chemicals was identified as one of the priority activities for NICEATM–ICCVAM.

The Organisation for Economic Co-operation and Development (OECD) has also made a substantial investment in research focusing on endocrine disruptors. In June 2002, the OECD Task Force on Endocrine Disruptors Testing and Assessment (EDTA) developed a conceptual framework for the testing and assessment of potential endocrine-disrupting substances (Gelbke et al. 2004; Hass et al. 2004; OECD 2002a). Several international efforts are underway that include using weight-of-evidence approaches to assess the endocrine-disrupting potential of commercial chemicals, as described in the conceptual framework. The following are prominent examples:

- European Commission Registration, Evaluation, Authorisation and Restriction of Chemicals [REACH] Regulation (EC) 1907/2006 and Directive 2006/121/EC corrigendum
- European Economic Community (EEC) Cosmetics Directive 76/768/EEC
- EEC Plant Protection Products Directive 91/414/EEC
- Japanese Extended Tasks on Endocrine Disruption [EXTEND 2010] program

The BG1Luc ER TA test method could be used as part of a weight-of-evidence approach in such programs.

It should be noted that individual U.S. and international agencies and programs must sanction the adoption of any test method. Discussion of the potential applicability of the BG1Luc ER TA test method in this BRD does not imply acceptance or adoption by any agency or program.

1.5.2 Criterion 2

ICCVAM's second prioritization criterion gauges the extent to which proposed test methods are warranted based on (1) the extent of expected use or application and (2) the impact on human, animal, or ecological health (ICCVAM 2003b).

Endocrine disruptors appear in a variety of products, including drugs (e.g., diethylstilbesterol), natural chemicals (e.g., genistein), and industrial chemicals (e.g., bisphenol A). Because of their ubiquitous uses, EDs are widespread in the environment. The association between exposure to EDs and adverse health effects in human and wildlife populations has led to worldwide concern. Health effects that have led to this concern include the following:

- Global increases in endometriosis and hormone-responsive cancers (e.g., testicular and breast cancers)
- Regional declines in sperm counts
- Increased prevalence of obesity
- Alterations to the onset of puberty
- Increases in altered sex ratios in wildlife populations (IPCS 2002; Latendresse et al. 2009; Newbold 2008; Newbold et al. 2008; Newbold et al. 2009; Newbold 2010; vom Saal et al. 2007)

Knowledge of these potential effects may reduce use of endocrine-disrupting chemicals, thereby decreasing the prevalence of associated reproductive and developmental issues. Several national and international programs are working to identify chemicals with endocrine-disrupting potential (**Section 1.5.1**), and the BG1Luc ER TA test method may be applicable to these programs. Depending on how it is used, an appropriate screen such as the BG1Luc ER TA test method may limit human and ecological exposure to EDs by identifying which chemicals are potential endocrine disruptors.

1.5.3 Criterion 3

As part of ICCVAM's third criterion, NICEATM evaluates the potential for the test method to reduce, refine, or replace animal use compared to current test methods accepted by regulatory agencies (ICCVAM 2003b).

The BG1Luc ER TA test method does not directly reduce, refine, or replace animal use compared to the current *in vitro* OPPTS 890.1300 test method (EPA 2009). To assess the estrogenic potential of substances, regulators commonly use the following three *in vivo* test methods: (1) rat uterotrophic assay, (2) rat pubertal female assay, and (3) fish short-term reproduction assay. In addition, animals must be used in the "*in vitro*" rat uterine cytosol ER binding assay as sources of ER. Although the BG1Luc ER TA test method will not directly replace any of these existing methods, it could be incorporated as part of a weight-of-evidence approach to reduce or eliminate the need for testing in these animal models.

1.5.4 Criterion 4

ICCVAM prioritizes proposed test methods for review and evaluation based upon their potential to better predict adverse health or environmental effects compared to current test methods accepted by regulatory agencies (ICCVAM 2003b).

When the BG1Luc ER TA validation study was initiated, no *in vitro* ER TA test methods were considered adequately valid for regulatory use. Today, only one *in vitro* ER TA test method is considered adequately validated by national and international agencies: the OECD Stably Transfected Human Estrogen Receptor- α Transcriptional Activation (STTA) Assay for the Detection of Estrogenic Agonist-Activity, described in OECD Chemicals Test Guideline (TG) 455 (OECD 2009). This method has been adopted by the EPA as part of the EDSP Tier 1 screening battery as OPPTS 890.1300 (EPA 2009).

The ER TA test method described in TG 455 uses HeLa-9903 cells, a human cervical carcinoma cell line in which human ER α and a reporter gene have been stably transfected. HeLa-9903 cells do not express endogenous ER α or ER β . The BG1Luc ER TA test method may improve prediction of adverse health effects in humans because it uses a human cell line (BG-1) that endogenously expresses both ER α and ER β (Park et al. 2009; Pujol et al. 1998; Rogers and Denison 2000; Zhou et al. 2005). BG-1 cells also express cofactors that may not be present in cells that do not express estrogen receptors (Marsaud et al. 2003; Shang et al. 2000; Webb et al. 1995).

The biological significance of two ER subtypes is still being explored, but there is mounting evidence for a role of ER β in a number of normal and abnormal physiologic processes (Brown et al. 2009; Foryst-Ludwig and Kintscher 2010; Harris 2007; Hayashi et al. 2003; Skliris et al. 2008; Weiser et al. 2008). Although there are presently no known naturally occurring ER β -specific substances, it is known that a number of substance types (e.g., isoflavones) are ER β selective (Chrzan and Bradford 2007; Escande et al. 2006; Kuiper et al. 1998; Mohler et al. 2010), with more potent responses through ER β than through ER α (Chrzan and Bradford 2007; Kuiper et al. 1998). The BG1Luc ER TA test method, using cells that express both ER α and ER β , allows for the potential detection of a wider range of substances than test methods that use cells expressing only the ER α receptor.

The BG1Luc ER TA test method also differs from the STTA assay in its ability to identify substances that possess ER antagonist activity. This is important because ER antagonists have a number of potential clinical uses, such as the treatment of osteoporosis and breast cancers (Ball et al. 2009; Bowers et al. 2000; Komm et al. 2005; Mohler et al. 2010). In addition, there is concern that any environmental anti-estrogens could have a detrimental influence on development and reproductive capacity of wildlife (Chamness et al. 1979; Fry and Toone 1981; Jones and Hajek 1995; Morris et al. 1967).

1.5.5 Criterion 5

Preliminary evaluations also summarize the extent to which a test method provides other advantages (e.g., reduced cost and time to perform) compared to current methods (ICCVAM 2003b).

The BG1Luc ER TA test method is a rapid *in vitro* method that can identify ER agonists and antagonists within approximately four days at a cost of a few thousand dollars per substance (**Section 10.3**). The test method also provides concentration–response activity and information on the potency of a substance relative to a reference estrogen or anti-estrogen. *In vivo* methods require 30 to 60 days for completion and may cost many thousands of dollars (**Section 10.3**). The BG1Luc ER TA test method, used as a potential screen, may lead to cost and time savings compared to an *in vivo* test and could alleviate the ethical concerns raised by the use of animals. In contrast, the STTA test method provides a concentration response and information on the potency of a substance relative to a reference estrogen only. The uterotrophic assay provides a concentration response but is not generally used to determine relative potency.

1.6 BG1Luc ER TA Test Method Protocol Standardization Study

As a result of the high priority of validation studies, NICEATM initiated and managed the ICCVAM-recommended study to standardize the BG1Luc ER TA test method agonist and antagonist protocols. The following essential test method components recommended by ICCVAM were incorporated in the protocols (ICCVAM 2003a):

- Reference estrogen and associated TA response
- Preparation of test substances and the volume of the administered solvent
- Concentration range of test substances that should be tested
- Solvent and positive controls
- Number of within-test replicates
- Methods for data analysis
- Experiment acceptance criteria
- Interpretation of results

The agonist and antagonist protocols were then standardized.

NICEATM evaluated the intralaboratory reproducibility and accuracy of the standardized protocols by testing a representative subset of the ICCVAM reference substances. Results of the protocol standardization study are provided in **Annex C**.

1.7 Interlaboratory BG1Luc ER TA Validation Study

NICEATM, which carries out independent validation studies consistent with the NTP mission, coordinated and led the international validation study with its counterparts in Europe (ECVAM) and Japan (JaCVAM). In 2009, NICEATM organized a Study Management Team (SMT) to oversee the scientific aspects of the validation study (**Table 1-1**). The SMT also directly coordinated the day-to-day activities of the validation study with the assistance of the NICEATM support contractor.

The BG1Luc ER TA test method was evaluated using laboratories in the United States (XDS), Europe (ECVAM), and Japan (Hiyoshi Corporation). The study proceeded in four phases (**Figure 1-1**). During Phase 1 of the validation, each of the three participating centers (ICCVAM, ECVAM, and JaCVAM) selected validation laboratories. The protocols were reviewed, and the laboratories demonstrated proficiency with the test method by successfully completing 10 replicate agonist and 10 replicate antagonist tests. In Phases 2 through 4 (**Figure 1-1**), the protocols were evaluated and refined, and 78 ICCVAM reference substances were tested (**Section 3.0**). Throughout the study, the SMT and NICEATM interacted to do the following:

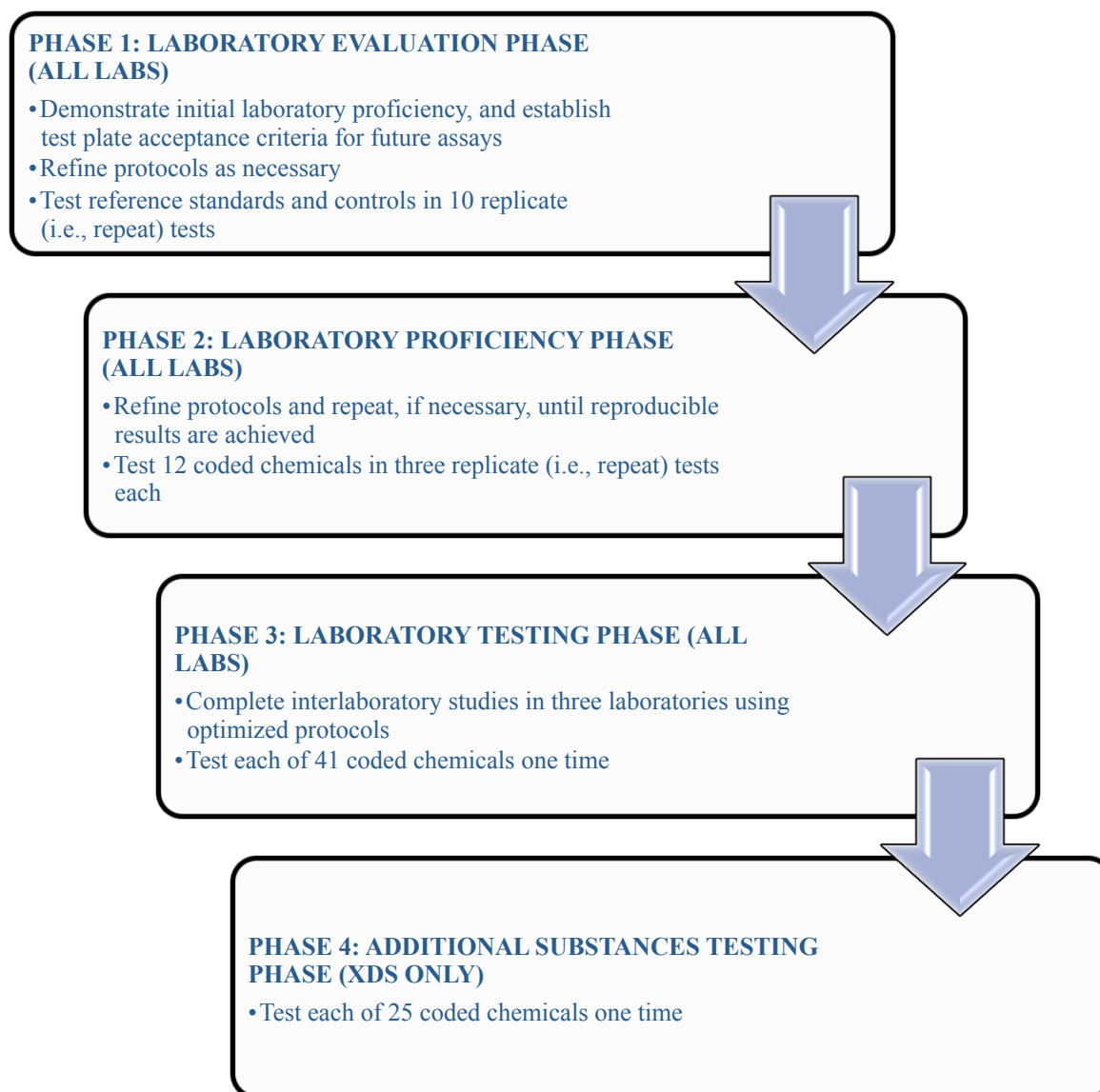
- Ensure that the study adhered to the principles stated in OECD Guidance Document Number 34 for prospective validation studies (OECD 2005)
- Develop a Statement of Work for the laboratories
- Determine timelines and deliverables
- Arrange for purchasing, coding, and distributing test substances to the laboratories
- Collect data from the laboratories and initiate statistical analyses
- Evaluate the reproducibility of results at each phase and refine the protocols, if necessary, before proceeding to the next phase
- Guide the study to conclusion and prepare documentation of the study

Table 1-1 Study Management Team for the BG1Luc ER TA Validation Study

Study Management Team Member	Affiliation
Dr. William Stokes	NIEHS/NICEATM
Dr. Warren Casey	NIEHS/NICEATM
Dr. Susanne Bremer	ECVAM
Dr. Elise Grignard	ECVAM
Dr. Hajime Kojima	JaCVAM
Dr. Atsushi Ono	JaCVAM
Dr. Soon Young Han	KoCVAM
Dr. David Allen	ILS/NICEATM
Ms. Patricia Ceger	ILS/NICEATM
Mr. Frank Deal (until March 2011)	ILS/NICEATM

Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; ILS = Integrated Laboratory Systems, Inc. (contract support staff for NICEATM); JaCVAM = Japanese Center for the Validation of Alternative Methods; KoCVAM = Korean Center for the Validation of Alternative Methods; NICEATM = National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods; NIEHS = National Institute of Environmental Health Sciences.

Figure 1-1 NICEATM/ECVAM/JaCVAM Validation Study Phases



1.8 Scientific Basis for the BG1Luc ER TA Test Method

The scientific basis of ER TA assays has been reviewed previously (Huet 2000; ICCVAM 2002d; OECD 2002b). Briefly, *in vitro* ER TA assays are designed to identify agonist or antagonist substances that might interfere with estrogen activity *in vivo*. Unlike receptor binding assays, TA assays can distinguish between agonist and antagonist activity. *In vitro* ER TA assays that are used to evaluate agonist activity are generally performed by quantifying the induction of a reporter gene product in response to activation of the ER by the test substance. *In vitro* ER TA assays that evaluate antagonist activity measure the ability of a test substance to inhibit the induction of the reporter gene product by a reference estrogenic substance.

The interaction of estrogens with the ER in a cell initiates a cascade of events. A number of endpoints can be used to measure endocrine activity at the cellular level, including receptor binding, cellular proliferation, and transcriptional activation. Upon ligand binding, the ER undergoes a conformational

change that allows dissociation of co-repressor proteins and the recruitment of co-activator proteins. *In vitro* binding assays such as the rat uterine cytosol binding assay measure the ability of a test substance to displace estradiol from the ER. The ligand-bound ER complex dimerizes and binds to an estrogen response element (ERE) located upstream of genes under estrogen control. Binding alters the transcription of estrogen-controlled genes, which leads to the initiation or inhibition of cellular processes, including those necessary for cell proliferation, normal fetal development, and adult homeostasis.

Transcriptional activation assays have several advantages over binding assays, including the following:

- Performance at physiologically relevant temperatures
- Measurement of biological response to receptor binding (i.e., RNA transcription and translation)
- The ability to distinguish between an agonist and an antagonist
- Detection of substances that initiate a transcriptional response in an indirect manner (Hall et al. 2001; Tremblay et al. 1999)

The BG1Luc ER TA test method is a transcriptional activation assay that uses a human cell line (BG-1) that endogenously expresses both ER α and ER β . An ERE that is coupled to a luciferase reporter gene has been stably transfected into BG1 cells. Stable transfection is desirable for transcriptional activation assays (Carey et al. 2009) because:

- The reporter gene is usually in a more stable configuration.
- The reporter gene is usually present in a more natural copy number.
- Cells that express the reporter gene have been selected for and clonally expanded, leading to increased reporter efficiency.
- Stably transfected cells do not need to be transfected each time the assay is performed.

Activation of the ER in response to estrogenic compounds drives transcription of the luciferase reporter, which is then quantified using a luminometer.

The BG1Luc4E2 cell line is suitable for ED testing on account of several properties:

- Endogenous expression of ERs and appropriate transcription machinery for hormone responsiveness
- Large number of ERs (Baldwin et al. 1998)
- High responsiveness to estrogens *in vitro* (Baldwin et al. 1998)
- Low background activity of the reporter gene in estrogen-free medium (Rogers and Denison 2002)
- Estrogen receptor specificity (Rogers and Denison 2002)

1.9 Range of Substances Amenable to the BG1Luc ER TA Test Method

The BG1Luc ER TA test method can be applied to a wide range of substances, provided they can be dissolved in dimethyl sulfoxide (DMSO) and are not toxic to BG1Luc4E2 cells at concentrations of 10 μ M or less. Although other solvents may be used for this test method, DMSO was the solvent of choice for this validation study. The BG1Luc ER TA test method may be applicable to chemical mixtures. No mixtures, however, were evaluated in this validation study. Volatile substances may yield acceptable results if CO₂-permeable plastic film is used to seal the test plates. No volatile substances were evaluated in this validation study. Substances with endogenous luminescence (Evans and Diepenhorst 1926), or those that naturally inhibit luciferase activity, cannot be used in this luciferase-based test method.

2.0 BG1Luc ER TA Test Method Protocol Components

2.1 Overview

The BG1Luc ER TA test method uses an estrogen-responsive reporter gene (*luc*) in the human ovarian adenocarcinoma cell line, BG-1, to detect substances with *in vitro* ER agonist or antagonist activity. Estrogen receptor-mediated transcription of the *luc* gene results in the production of luciferase, the activity of which is quantified using a luminometer (see **Section 2.2.1**). A concentration–response curve can be established to provide qualitative and quantitative information regarding the *in vitro* estrogenic activity of a test substance. The advantages of using a luciferase reporter gene system are low background, high sensitivity, speed, and a wide dynamic range.

The primary objective of this test method is to provide a qualitative assessment of *in vitro* estrogenic activity (i.e., whether a substance is positive or negative for estrogenic activity). Quantitative analysis is also performed to provide additional information on the estrogenic potency of test substances. For example, quantitative analysis can determine the half-maximal effective concentration (EC₅₀) or the half-maximal inhibitory concentration (IC₅₀). Separate protocols are used to identify substances that possess ER agonist or antagonist activity, although the two protocols share most major components.

In a 2003 evaluation, ICCVAM recommended minimum essential test method components for *in vitro* ER TA test method protocols (ICCVAM 2003a), which included the following considerations:

- A reference standard should be included to demonstrate the adequacy of the test method for detecting ER agonists or antagonists.
- Each study should include a set of concurrent solvent controls.
- Each study should include an evaluation of cytotoxicity.
- A weak positive agonist control with an EC₅₀ two to three orders of magnitude higher than the reference estrogen should be included in each study to demonstrate that the test method is functioning properly and is sufficiently sensitive to detect weak estrogen agonists.
- To demonstrate that the test method is functioning properly and is sufficiently sensitive to detect weak estrogen antagonists, each study should include a weak positive antagonist control that inhibits the reference estrogen response by 50% (IC₅₀) at a concentration two to three orders of magnitude higher than the reference anti-estrogen.
- The maximum test substance concentration should be 1 mM unless otherwise limited by solubility or cytotoxicity.
- At least seven concentrations spaced at logarithmic (log₁₀) intervals, up to the limit concentration, should be tested.
- EC₅₀ or IC₅₀ values should be calculated for all positive substances when possible.
- Protocols should contain established test plate acceptance criteria.

The ICCVAM-recommended test method components were incorporated into the BG1Luc ER TA test method protocols during a protocol standardization study coordinated by NICEATM and conducted at XDS (**Annex C**). The goal of the standardization study, in which eight agonists and eight antagonists were tested, was to develop protocols for use in the ICCVAM-sponsored international validation study. Once the multiphase validation study was initiated, the protocols continued to be refined after each phase resulting in optimized protocols for agonist and antagonist testing (see **Annexes E** and **F**, respectively). The remainder of this section provides details on the essential test method components and the rationale for their inclusion in the optimized protocols.

2.1.1 General Procedural Overview

Agonist and antagonist testing in the BG1Luc ER TA test method is conducted in three steps:

1. Solubilization and dilution of test substances
2. Range finder testing and selection of starting concentrations and dilution factors for test substances to be used in comprehensive testing
3. Comprehensive testing, qualitative assessment of *in vitro* estrogenic activity, and, where appropriate, quantitative analysis to assess estrogenic or anti-estrogenic potency

2.2 Materials

2.2.1 BG1Luc4E2 Cells

The BG-1 cell line, developed by Rogers and Denison (2000), is derived from immortalized human ovarian adenocarcinoma cells. The cell line has been used extensively to examine the estrogenic effects of chemicals (Baldwin et al. 1998; Park et al. 2009; Pujol et al. 1998; Rogers and Denison 2000; Rogers and Denison 2002; Zhou et al. 2005; Zimniski et al. 1989). BG-1 cells endogenously express both human ER α and ER β (Wong and Matsumura 2006), although ER α is the predominant isoform (90%) (Monje and Boland 2001; Pujol et al. 1998; Welshons et al. 1988). Rogers and Denison stably transfected BG-1 cells with a plasmid containing a firefly luciferase reporter gene under control of four estrogen response elements placed upstream of the mouse mammary tumor virus (MMTV) promoter. The resulting BG1Luc4E2 cell line expresses luciferase activity in response to estrogen and estrogen-like substances. While the MMTV promoter sequence used for the BG-1 plasmid construct lacks the glucocorticoid-responsive elements normally present in this region (Garrison et al. 1996; Lee et al. 1984; Rogers and Denison 2000), the BG-1 developers examined possible cross-reactivity with other steroid and nonsteroid hormones.

Progesterone, testosterone, all-trans retinoic acid, and thyroid hormone did not induce luciferase activity. Dihydrotestosterone (AR ligand) and dexamethasone (glucocorticoid receptor ligand) induced only a small degree of luciferase activity (Rogers and Denison 2000). Together, these results indicate that the BG1Luc4E2 cells exhibit only minor cross-reactivity with other ligands for members of the nuclear hormone receptor superfamily.

XDS provided cryopreserved BG1Luc4E2 cells from their cell bank to ECVAM and Hiyoshi for the validation study. ECVAM and Hiyoshi propagated and cryopreserved multiple ampoules of cells to establish their working cell banks for use throughout the study.

2.2.2 Cell Culture Reagents and Supplies

The BG1Luc ER TA test method requires general cell culture materials, reagents, and supplies (see **Annexes E and F** [protocols] for formulations, and concentrations of solutions and media). The participating laboratories independently acquired cell culture materials, reagents, and supplies.

The following reagents are used for cell culture procedures in the BG1Luc ER TA test method:

- DMSO
- Luciferase reagent
- Phosphate-buffered saline (PBS)
- Trypsin (2.5% v/v in PBS as a cell dissociation agent)
- Gentamycin (G418)
- Penicillin-streptomycin (Pen-Strep)
- L-glutamine
- Fetal bovine serum (FBS)

- Charcoal/dextran treated FBS
- RPMI 1640 media containing L-glutamine
- Dulbecco's modified Eagle's medium (DMEM) containing high glucose (4.5 g/L) and sodium pyruvate, without L-glutamine or phenol red.

RPMI 1640, supplemented with Pen-Strep and FBS, is used for the routine maintenance of cell culture, for freezing cells, and for thawing cells.

DMEM supplemented with charcoal/dextran treated FBS (to remove free hormones from sera), Pen-Strep, and L-glutamine is designated as estrogen-free media (EFM). Cells are transferred from RPMI to EFM before testing.

The BG1Luc ER TA test method requires the following laboratory cell culture supplies:

- Cryogenic ampoules (cryovials)
- Plastic culture tubes (e.g., 50-mL conical tubes)
- Hemocytometer
- Pipettes, pipettors, repeat pipettors, pipette tips
- Sterile, disposable tissue culture plasticware (e.g., 25-cm² and 75-cm² tissue culture flasks, 96-well microtiter plates)

2.2.3 Equipment

Performance of the BG1Luc ER TA test method requires a laboratory equipped with a designated cell culture area. Equipment required for the conduct of the test method includes the following:

- Analytical balance
- Biological safety hood, Class II or higher, with HEPA filter
- Centrifuge (capable of 1000 x g)
- 4°C refrigerator
- Freezers, -20°C and -70°C
- Incubator (37°C ± 1°C, 90% ± 5% humidity, and 5% ± 1% CO₂/air)
- Liquid nitrogen cryostorage
- Microplate, auto-injecting luminometer
- Shaker for 96-well plates
- Vortex mixer

2.3 Cell Culture

The primary objective of any tissue culture operation is to maintain consistency in the cultures used. To do this, strict control of culture conditions (i.e., growth and maintenance media, culture schedules, culture flasks and plates, substrate type, seeding conditions, dissociation methods) must be maintained. Strict control must also be taken to properly freeze, maintain, and thaw cultures in a systematic manner because cryopreservation techniques can affect subsequent culture growth and performance. All pertinent information about cell culture reagents and supplies (e.g., lot number, manufacturer, product code, certificates of analysis) should be collected and maintained in log books and reports.

Cryopreserved BG1Luc4E2 cells are thawed, resuspended in RPMI media, transferred into 25-cm² tissue culture flasks, and incubated at 37°C ± 1°C, 90% ± 5% humidity, and 5% ± 1% CO₂/air for 48 to 72 hours (see **Annexes E and F** [protocols] for cell culture specifics). When cells reach 80% to 90% confluence (as estimated from a visual inspection of cell density), they are removed from the flask by trypsinization. A dissociated single-cell suspension is added to new flasks for propagation, and the cells are passaged/subcultured at least twice before conditioning in EFM. Forty-eight to 72 hours after the second subculture, cells are trypsinized and pelleted. The RPMI media are

removed. Cells are then resuspended in EFM, and the cell suspension is added to new flasks for conditioning. At this time, G418 is added to the EFM to select BG1Luc4E2 cells containing the G418-resistant reporter plasmid. When cells are 80% to 90% confluent, they are trypsinized, counted, and seeded into 96-well plates for testing.

2.4 Reference Standards and Controls

ICCVAM recommends the use of appropriate reference standards and controls for ER TA test methods in order to maximize test method intra- and interlaboratory reproducibility and minimize the likelihood of erroneous results (ICCVAM 2003a).

2.4.1 Vehicle Control

- 1% DMSO in EFM is used as the concurrent vehicle control for all testing in agonist and antagonist protocols.

A concurrent vehicle control in ER TA agonist and antagonist test methods provides a measure of the extent of TA in the absence of the reference standard, control, or test substances. For ER TA test methods, the concurrent vehicle control is the baseline against which the extent of TA induction is determined. (In the BG1Luc ER TA test method, vehicle control relative light units [RLUs] for each test plate are averaged and then subtracted from test substance, reference standard, and control RLU values.) XDS tested several solvents when developing the BG1Luc ER TA test method and selected a solution of 1% DMSO (v/v) in EFM because of its ability to solubilize a wide range of both hydrophilic and hydrophobic substances and to achieve relatively high concentrations of test substance without reducing cell viability.

2.4.2 Estrogenic Reference Standard

- In accordance with the ICCVAM recommendations, 17 β -estradiol (E2, CAS Registry Number [CASRN] 50-28-2) is used as the reference estrogen to demonstrate the adequacy of the ER TA test method. In the BG1Luc ER TA test method, adequacy is based on the ability of the E2 reference standard to induce ER TA activity.

Table 2-1 provides the concentrations of E2 used in different phases of testing. In ER agonist range finder testing, a 4-point dilution was used in validation testing to broadly define the E2 curve response in terms of bottom, slope, and top. An 11-point dilution of E2 was then used in comprehensive ER agonist testing to more fully define the E2 response curve, thereby allowing an EC₅₀ to be calculated. Antagonist testing used an E2 concentration of 9.18×10^{-11} M in ER antagonist range finder and comprehensive testing in order to provide a level of induction against which antagonistic effects of test substances could be assessed.

Table 2-1 E2 Concentrations Tested for Agonist and Antagonist Methods

E2 Concentration (M)	Agonist Test Method		Antagonist Test Method	
	Comprehensive Testing	Range Finder Testing	Comprehensive Testing	Range Finder Testing
3.67×10^{-10}	X	-	-	-
1.84×10^{-10}	X	X	-	-
9.18×10^{-11}	X	-	X	X
4.59×10^{-11}	X	X	-	-
2.29×10^{-11}	X	-	-	-
1.15×10^{-11}	X	X	-	-
5.73×10^{-12}	X	-	-	-
2.87×10^{-12}	X	X	-	-
1.44×10^{-12}	X	-	-	-
7.16×10^{-13}	X	-	-	-
3.59×10^{-13}	X	-	-	-

X = tested; - = not tested

Abbreviations: E2 = 17 β -estradiol; M = molar.

2.4.3 Weak Agonist Control

- p,p'*-Methoxychlor (Met, CASRN 72-43-5) is used as the weak positive control in agonist comprehensive testing. A weak positive control is not used during agonist range finder testing.

ICCVAM recommends that a positive control with an EC₅₀ two to three orders of magnitude higher than E2 (EC₅₀ = 3×10^{-12} M) be included in each study to demonstrate that the test method is functioning properly and is sufficiently sensitive for detecting weak estrogen agonists (2003a). However, given that the range of responses expected to be assessed with this method during the protocol standardization study was greater than six orders of magnitude during the protocol standardization study, the SMT concluded that a positive control with a higher EC₅₀ multiple would be more appropriate. During protocol standardization, a number of substances were evaluated for use as the weak agonist control (**Annex C**). Based on this evaluation, Met was considered the most appropriate control because it produced the most consistent ER TA response curves in the desired range (EC₅₀ = 6 μ M), approximately six orders of magnitude higher than E2 (EC₅₀ = 3×10^{-12} M in the BG1Luc ER TA test method). A Met concentration of 9.06×10^{-6} M was selected because it was the lowest concentration that gave the maximum response.

2.4.4 Anti-Estrogenic Reference Standard

- Raloxifene HCl (Ral, CASRN 82640-04-8) is used as the anti-estrogenic reference standard.

Although ICCVAM recommends ICI 182,780 as a reference standard in ER TA antagonist assays, this substance has limited commercial availability (ICCVAM 2006). As an alternative, the more commercially available Ral was evaluated for use as the reference standard during the protocol standardization study. Ral is a strong estrogen antagonist also recommended by ICCVAM as a reference standard (ICCVAM 2006). Ral consistently produced full concentration–response curves with a mean IC₅₀ value of 2.24×10^{-9} M in the BG1Luc ER TA test method (**Annex C**). Therefore, the SMT selected Ral as the anti-estrogenic reference standard for the validation study.

The concentrations of Ral used in ER antagonist range finder and comprehensive testing are provided in **Table 2-2**. A 3-point dilution was used in ER antagonist range finder testing to broadly define the top, slope, and bottom of the Ral response curve. A 9-point dilution of Ral was then used in comprehensive ER antagonist testing to more fully define the Ral response curve, thereby allowing the calculation of an IC₅₀.

Table 2-2 Raloxifene Standard Concentrations Tested in the Antagonist Assay

Raloxifene Concentration (M)	Antagonist Comprehensive Testing	Antagonist Range Finder Testing
2.45×10^{-8}	X	-
1.23×10^{-8}	X	-
6.14×10^{-9}	X	-
3.06×10^{-9}	X	X
1.53×10^{-9}	X	-
7.67×10^{-10}	X	X
3.82×10^{-10}	X	-
1.92×10^{-10}	X	X
9.57×10^{-11}	X	-

X = tested; - = not tested

Abbreviations: M = molar.

2.4.5 Weak Antagonist Control

- Tamoxifen (Tam, CASRN 10540-29-1) is used as the weak positive control for antagonist comprehensive testing. A weak positive control is not used for antagonist range finder testing.

The use of a weak antagonist as a concurrent control in ER TA antagonist test methods provides a measure of the range of responses that can be detected by the test. ICCVAM recommends using a weak positive control with an IC₅₀ at least three orders of magnitude higher than the reference antagonist (2003a), Ral (IC₅₀ = 2.24×10^{-9} M). During protocol standardization (see **Annex C**), a number of substances were evaluated for use as the weak antagonist control. Flavone produced a dose response and an IC₅₀ = 4.3×10^{-7} M, which was consistent with the single literature reference for this compound (reported IC₅₀ = ~15 µM) (Collins-Burow et al. 2000) and was two times below that of Ral. Based on these results, flavone was chosen as the weak antagonist control for the validation study. However, data from the completed study showed that the vast majority of test substances classified as “negative” or “presumed negative” produced a “positive” response at concentrations above ~10 µM (see **Annex K** for ER TA antagonist testing results). Consequently, the use of flavone as a weak antagonist control was reconsidered, as discussed below.

The antagonist method is a “loss-of-function” method, in which a positive result is based on a *decrease* in luciferase activity. This is in contrast to the agonist method, in which an *increase* in luciferase activity (i.e., “gain of function”) indicates a positive response. Consequently, any substance that disturbs cellular homeostasis or causes cytotoxicity will produce an apparent positive response (i.e., dead cells produce no signal and, therefore, produce the maximum response). To account for this, an assessment of cell viability is included in the agonist and antagonist test method protocols (**Section 2.6**). Data from antagonist validation testing were reviewed to determine whether the observed decrease in luciferase activity (positive response) correlated with a loss in cellular viability. In many cases, there was no observed decrease in cellular viability at the highest concentration tested.

In cases where a loss of viability was observed, a decrease in luciferase activity usually *preceded* a loss of cellular viability, sometimes at concentrations up to two or three log dilutions lower than the cytotoxic concentration.

These findings indicate that cellular viability cannot be reliably used as an indicator of test substance interference with the BG1Luc ER TA, and that it is impossible to distinguish true positives from false positives at concentrations above ~10 μ M. In addition, NICEATM–ICCVAM could not identify in the literature any substances classified as positive for ER antagonism with an $IC_{50} > 10 \mu$ M.

Therefore, the SMT established 10 μ M as the upper limit of utility for determining antagonist activity in the BG1Luc ER TA test method. Because the 10 μ M would preclude the use of flavone as a weak antagonist control ($IC_{50} = 15 \mu$ M), the SMT selected Tam as a weak antagonist control because it has been conclusively shown to bind to both ERs (46/46 studies; **Table 3-2**) and act as an ER antagonist in most ER TA studies (20/22 studies; **Table 3-2**). The mean IC_{50} for Tam in ER TA studies is 7.20×10^{-7} M, which is twofold above that of Ral yet below the 10 μ M upper limit of the assay.

2.5 Test Substance Concentrations

- For agonist testing, the highest soluble, noncytotoxic concentration should be tested up to a limit of 1 mM.
- For antagonist testing, the highest soluble, noncytotoxic concentration should be tested up to a limit of 10 μ M.

ICCVAM recommends that the maximum test substance concentration should be 1 mM unless otherwise limited by solubility or cytotoxicity (2003a). (**Note:** Reference substances were coded in order to conduct the validation study in a blinded manner; therefore, the participating laboratories were instructed to use 100 mg/mL as the limit concentration.) However, as outlined in **Section 2.4.5**, the BG1Luc ER TA test method validation study data indicate that concentrations above 10 μ M in the antagonist assay consistently produce false positive responses in this loss-of-function assay. Consequently, the SMT established 10 μ M as the upper limit of utility for determining antagonist activity in the BG1Luc ER TA test method.

2.5.1 Solubility Testing

- The starting concentration for range finder testing is established by determining the maximum test substance solubility in 100% DMSO.

ICCVAM recommends that the maximum test substance concentration be 1 mM unless limited by solubility or cytotoxicity (2003a). Procedures used to assess solubility are described in this section, and procedures used to assess cytotoxicity are described in **Section 2.5.2**.

During Phase 1 and Phase 2 testing, maximum test substance solubilities were determined at log intervals up to 1 mg/mL (v/v in 1% DMSO/cell culture media). Following Phase 2 of the validation study, a high degree of variability was noted in solubility assessment performed on the same substance at different laboratories. Problems associated with log scale dilutions in the 1% DMSO medium were believed to be causing the variability. To reduce differences in solubility estimates between laboratories, protocols were modified to use test substance solubility in 100% DMSO as the starting concentration for range finder testing. This protocol modification was used for Phase 3 and Phase 4 testing. Test substance solubility data are provided in **Section 4**.

2.5.2 Cytotoxicity Testing

A qualitative visual observation method that assesses viability on a scale of 1 (normal) to 4 (significant loss of viability) is used to assess cell viability in the BG1Luc ER TA test method. Viability scores of 2 or greater are classified as cytotoxic.

The assessment of cytotoxicity is an integral part of agonist and antagonist range finder and comprehensive testing and data analysis. Cytotoxicity results play an additional role in the interpretation of agonist and antagonist range finder data, as described below in **Sections 2.6.1** and **2.6.2**, respectively.

The peer review panel (ICCVAM 2002d, 2002e) recommended an assessment of cell viability to help define the upper limit for test substance concentrations, similar to the maximum tolerated dose (MTD) approach used in *in vivo* studies. During the protocol standardization study for the BG1Luc ER TA test method (**Annex C**), XDS used the CellTiter-Glo[®] (Promega Corporation) quantitative cell viability assay to assess the viability of BG1Luc4E2 cells following exposure to increasing concentrations of test substance. CellTiter-Glo measures cell viability via a luminescent signal that is proportional to the amount of adenosine triphosphate (ATP) in viable cells. Results indicated that the ER TA activity of the fixed amount of E2 used in antagonist testing was significantly reduced when the reduction in ATP level per well exceeded 20%. Based on these results, concentrations of substance that reduced cell viability more than 20% were classified as cytotoxic. However, like the BG1Luc ER TA test method, the CellTiter-Glo assay is based on a luminescent endpoint (ER TA luciferase vs. ATP luminescence). For this reason, the use of parallel plates is necessary because ATP luminescence cannot be delineated from ER TA-associated luciferase activity.

Therefore, an alternative qualitative method to assess cell viability was developed by XDS during the protocol standardization study (**Annex C**). This method relies on visual observation of cell density and morphology to assign cell viability scores using criteria listed in **Table 2-3**. Test substance concentrations of 2 or greater are considered to be cytotoxic.

A direct comparison of the CellTiter-Glo assay and visual observation methods indicated that CellTiter-Glo values of 80% or greater corresponded to a viability score of 1 in the visual observation method study (**Annex C**). Therefore, the visual observation method was considered adequate for assessing cell viability in the BG1Luc ER TA test method, thereby precluding the need for parallel test plates.

Table 2-3 Visual Observation Scoring Table for Cell Viability

Viability Score	Brief Description
1	Normal cell morphology and cell density
2	Altered cell morphology and/or small gaps between cells
3	Altered cell morphology and/or large gaps between cells
4	Few (or no) visible cells
P	Wells containing precipitation are to be noted with “P”

2.6 Range Finder Testing

The purpose of range finder testing is to establish the concentration range of a test substance to be included in comprehensive testing. This involves identifying both an appropriate starting concentration and a dilution scheme. The starting concentration of a test substance is based on the highest soluble concentration that is not cytotoxic, as described in **Section 2.5**. Results from range finder testing are used to select a 1:5 or 1:2 dilution scheme for comprehensive testing. A 1:5 dilution covers a wider concentration range (7.5 log dilutions), while a 1:2 dilution provides higher resolution over a smaller range (3.5 log dilutions). Procedures for range finder testing, along with the criteria used to determine the appropriate testing range, are provided below.

2.6.1 Agonist Range Finder Testing

Reference Standard and Control Concentrations Used for Agonist Range Finder Testing

- E2, the reference estrogen, is run in duplicate at four concentrations (1.84×10^{-10} , 4.59×10^{-11} , 1.15×10^{-11} , and 2.87×10^{-12} M).
- The vehicle control (1% DMSO v/v in EFM) is run in quadruplicate.

Agonist Range Finder Plate Design

- All 96 wells of the test plate are used during range finder testing. A maximum of six substances can be tested at seven concentrations in duplicate on each range finder plate. Starting concentrations are determined during solubility testing. Plate design for agonist testing is provided below in **Figure 2-1**.

In Phase 1 of the validation study, the lead laboratory (XDS) conducted studies to optimize the plate design in order to improve the statistical power and allow all 96 wells to be used (**Annex M**). Results demonstrated that, although there were statistically significant differences in values between outside and inside wells, the differences did not affect the selection of the appropriate starting concentrations for comprehensive testing (see **Annex M**). Therefore, the design of agonist and antagonist range finder plates was modified to use all 96 wells of the test plate, with six test substances being tested at seven concentrations in duplicate on each range finder plate.

Figure 2-1 96-Well Test Plate Layout for Agonist Range Finder Testing

TS1-1	TS1-1	TS2-1	TS2-1	TS3-1	TS3-1	TS4-1	TS4-1	TS5-1	TS5-1	TS6-1	TS6-1
TS1-2	TS1-2	TS2-2	TS2-2	TS3-2	TS3-2	TS4-2	TS4-2	TS5-2	TS5-2	TS6-2	TS6-2
TS1-3	TS1-3	TS2-3	TS2-3	TS3-3	TS3-3	TS4-3	TS4-3	TS5-3	TS5-3	TS6-3	TS6-3
TS1-4	TS1-4	TS2-4	TS2-4	TS3-4	TS3-4	TS4-4	TS4-4	TS5-4	TS5-4	TS6-4	TS6-4
TS1-5	TS1-5	TS2-5	TS2-5	TS3-5	TS3-5	TS4-5	TS4-5	TS5-5	TS5-5	TS6-5	TS6-5
TS1-6	TS1-6	TS2-6	TS2-6	TS3-6	TS3-6	TS4-6	TS4-6	TS5-6	TS5-6	TS6-6	TS6-6
TS1-7	TS1-7	TS2-7	TS2-7	TS3-7	TS3-7	TS4-7	TS4-7	TS5-7	TS5-7	TS6-7	TS6-7
E2-1	E2-2	E2-3	E2-4	VC	VC	VC	VC	E2-1	E2-2	E2-3	E2-4

Abbreviations: E2 = E2 control; TS = test substance; VC = vehicle control (DMSO [1% v/v EFM]).

E2-1 to E2-4 = concentrations of the E2 reference standard (from high to low)

TS1-1 to TS1-7 = concentrations (from high to low) of test substance 1 (TS1)

TS2-1 to TS2-7 = concentrations (from high to low) of test substance 2 (TS2)

TS3-1 to TS3-7 = concentrations (from high to low) of test substance 3 (TS3)

TS4-1 to TS4-7 = concentrations (from high to low) of test substance 4 (TS4)

TS5-1 to TS5-7 = concentrations (from high to low) of test substance 5 (TS5)

TS6-1 to TS6-7 = concentrations (from high to low) of test substance 6 (TS6)

Agonist Range Finder Plate Acceptance Criteria

- The mean DMSO control RLU values must be within 2.5 times the standard deviation of the historical DMSO control mean RLU value.
- E2 induction must be greater than threefold. Induction is calculated by averaging the highest E2 reference RLU value from both E2 concentration curves and then dividing this by the average DMSO control RLU value.

Data from plates that fail any acceptance criterion should be discarded and the experiment repeated.

Interpretation of Results from Agonist Range Finder Testing

- If no points on the test substance concentration curve are greater than the DMSO control mean plus three times its standard deviation (SD), comprehensive testing for ER agonist activity should be conducted using the highest noncytotoxic concentration tested.
- If any points on the test substance concentration curve are greater than the DMSO control mean plus three times its SD, testing should use a starting concentration one log higher than the concentration giving the highest adjusted RLU value.
- An 11-point 1:2 serial dilution (covering approximately three orders of magnitude) should be used if the resulting concentration range will resolve the full dose response curve of the test substance, as estimated from the range finder data. Otherwise, an 11-point 1:5 dilution should be used.
- An 11-point 1:5 serial dilution (covering approximately seven orders of magnitude) should be used if a substance exhibits a biphasic, hormetic, or U-shaped (Calabrese and Baldwin 2001) concentration–response curve not associated with cytotoxicity in the range finder test. For hormetic or U-shaped curves, the dilution starts at a concentration one log higher than the concentration associated with the peak of activity (maximum RLU). In a biphasic curve, the starting concentration is one log higher than the concentration associated with the maximum RLU of the peak at the highest end of the concentration–response curve.

2.6.2 Antagonist Range Finder Testing*Reference Standard and Control Concentrations Used for Antagonist Range Finder Testing*

- A single concentration of E2 (9.18×10^{-11} M), intended to provide 80% of the maximum E2 induction, is run in triplicate.
- Three concentrations of the reference anti-estrogen, raloxifene HCl, are each run in duplicate (3.06×10^{-9} , 7.67×10^{-10} , and 1.92×10^{-10} M).
- The vehicle control (1% DMSO v/v in EFM) is run in triplicate.
- All reference anti-estrogen and test wells must contain a fixed concentration of E2 (9.18×10^{-11} M), intended to provide 80% of the maximum E2 induction.

Antagonist Range Finder Plate Design

All 96 wells of the test plate are used during range finder testing. A maximum of six substances can be tested at seven concentrations in duplicate on each range finder plate. Starting concentrations are determined using starting concentrations that were determined during solubility testing. The plate design for antagonist testing is provided in **Figure 2-2**.

Figure 2-2 96-Well Test Plate Layout for Antagonist Range Finder Testing

TS1-1	TS1-1	TS2-1	TS2-1	TS3-1	TS3-1	TS4-1	TS4-1	TS5-1	TS5-1	TS6-1	TS6-1
TS1-2	TS1-2	TS2-2	TS2-2	TS3-2	TS3-2	TS4-2	TS4-2	TS5-2	TS5-2	TS6-2	TS6-2
TS1-3	TS1-3	TS2-3	TS2-3	TS3-3	TS3-3	TS4-3	TS4-3	TS5-3	TS5-3	TS6-3	TS6-3
TS1-4	TS1-4	TS2-4	TS2-4	TS3-4	TS3-4	TS4-4	TS4-4	TS5-4	TS5-4	TS6-4	TS6-4
TS1-5	TS1-5	TS2-5	TS2-5	TS3-5	TS3-5	TS4-5	TS4-5	TS5-5	TS5-5	TS6-5	TS6-5
TS1-6	TS1-6	TS2-6	TS2-6	TS3-6	TS3-6	TS4-6	TS4-6	TS5-6	TS5-6	TS6-6	TS6-6
TS1-7	TS1-7	TS2-7	TS2-7	TS3-7	TS3-7	TS4-7	TS4-7	TS5-7	TS5-7	TS6-7	TS6-7
Ral-1	Ral-2	Ral-3	VC	VC	VC	E2	E2	E2	Ral-1	Ral-2	Ral-3

Abbreviations: E2 = E2 control; Ral = raloxifene; TS = test substance; VC = vehicle control (DMSO [1% v/v EFM]).

Ral-1 to Ral-3 = concentrations of the raloxifene/E2 reference standard (from high to low)

TS1-1 to TS1-7 = concentrations (from high to low) of test substance 1 (TS1)

TS2-1 to TS2-7 = concentrations (from high to low) of test substance 2 (TS2)

TS3-1 to TS3-7 = concentrations (from high to low) of test substance 3 (TS3)

TS4-1 to TS4-7 = concentrations (from high to low) of test substance 4 (TS4)

TS5-1 to TS5-7 = concentrations (from high to low) of test substance 5 (TS5)

TS6-1 to TS6-7 = concentrations (from high to low) of test substance 6 (TS6)

Antagonist Range Finder Plate Acceptance Criteria

- The mean DMSO control RLU value for each plate must be within 2.5 times the SD of the historical DMSO control mean RLU value.
- Test plate E2 control RLU values must be within 2.5 times the SD of the historical E2 control mean RLU value.
- Plate reduction must be greater than threefold. Reduction is calculated by averaging the highest Ral reference RLU values, then dividing by the averaged lowest Ral RLU value.

Data from plates that fail any acceptance criterion should be discarded and the experiment repeated.

Interpretation of Results from Antagonist Range Finder Testing

- If no points on the test substance concentration curve are less than the mean of the E2 control minus three times the SD, comprehensive testing for ER antagonist activity should be conducted using the highest noncytotoxic concentration tested.
- If any points on the test substance concentration curve are less than the E2 control mean minus three times the SD, testing should use a starting concentration one log higher than the concentration giving the lowest adjusted RLU value.
- An 11-point 1:2 serial dilution (covering approximately three orders of magnitude) should be used if the resulting concentration range will resolve the full concentration–response curve of the test substance as estimated from the range finder data. Otherwise, an 11-point 1:5 dilution should be used.
- An 11-point 1:5 serial dilution (covering approximately seven orders of magnitude) should be used if a substance exhibits a biphasic, hormetic, or U-shaped (Calabrese and Baldwin 2001) concentration–response curve not associated with cytotoxicity in the range finder test. For hormetic or U-shaped curves, the dilution starts at a concentration one log higher than the

concentration associated with the peak of activity (maximum RLU). In a biphasic curve, the starting concentration is one log higher than the concentration associated with the maximum RLU of the peak at the highest end of the concentration–response curve.

2.7 Comprehensive Testing

2.7.1 Comprehensive Agonist Testing

Reference Standard and Control Concentrations Used for Agonist Comprehensive Testing

- E2, the reference estrogen, is run in duplicate at eleven concentrations (see **Table 2-1**).
- Met, the weak positive control, is run in quadruplicate at a single concentration of 9.06×10^{-6} M.
- The vehicle control (1% DMSO in EFM) is run in quadruplicate.

Plate Design

All 96 wells of the test plate are used during comprehensive agonist testing. Two substances can be tested at eleven concentrations, in triplicate, on each plate. Starting concentrations and dilution factors are determined based on range finder results (**Section 2.6**). Plate design for comprehensive agonist testing is provided below in **Figure 2-3**.

To evaluate the effect of using outer test plate wells on comprehensive testing, EC₅₀ values from serial dilutions of BPA derived from replicates using outside wells were compared to EC₅₀ values derived from replicates using inside wells. The comparisons showed no significant differences between EC₅₀ values derived from replicates using outside wells and those derived from using inside wells (see **Annex M**).

Figure 2-3 96-Well Test Plate Layout for Comprehensive Agonist Testing

TS1-1	TS1-2	TS1-3	TS1-4	TS1-5	TS1-6	TS1-7	TS1-8	TS1-9	TS1-10	TS1-11	VC
TS1-1	TS1-2	TS1-3	TS1-4	TS1-5	TS1-6	TS1-7	TS1-8	TS1-9	TS1-10	TS1-11	VC
TS1-1	TS1-2	TS1-3	TS1-4	TS1-5	TS1-6	TS1-7	TS1-8	TS1-9	TS1-10	TS1-11	VC
TS2-1	TS2-2	TS2-3	TS2-4	TS2-5	TS2-6	TS2-7	TS2-8	TS2-9	TS2-10	TS2-11	VC
TS2-1	TS2-2	TS2-3	TS2-4	TS2-5	TS2-6	TS2-7	TS2-8	TS2-9	TS2-10	TS2-11	Met
TS2-1	TS2-2	TS2-3	TS2-4	TS2-5	TS2-6	TS2-7	TS2-8	TS2-9	TS2-10	TS2-11	Met
E2-1	E2-2	E2-3	E2-4	E2-5	E2-6	E2-7	E2-8	E2-9	E2-10	E2-11	Met
E2-1	E2-2	E2-3	E2-4	E2-5	E2-6	E2-7	E2-8	E2-9	E2-10	E2-11	Met

Abbreviations: E2 = E2 control; Met = *p,p'*-methoxychlor weak positive control; TS = test substance; VC = vehicle control (DMSO [1% v/v EFM]).

TS1-1 to TS1-11 = concentrations (from high to low) of test substance 1

TS2-1 to TS2-11 = concentrations (from high to low) of test substance 2

E2-1 to E2-11 = concentrations of the E2 reference standard (from high to low)

Plate Acceptance Criteria for Comprehensive Agonist Testing

- The mean DMSO control RLU value for each plate must be within 2.5 times the SD of the historical DMSO control mean RLU value.
- E2 induction must be greater than threefold. Induction is calculated by averaging the highest E2 reference RLU value from each E2 concentration curve and then dividing this by the average DMSO control RLU value.
- The E2 reference standard curve should be sigmoidal in shape and have at least three values within the linear portion of the curve.
- The mean plate Met control RLU value must be greater than the mean DMSO control RLU value plus three times the SD.

Modification of Plate Acceptance Criteria for Comprehensive Agonist Testing

Following Phase 2a of the validation study, NICEATM–ICCVAM evaluated the failure rates of plates used to comprehensively test four agonist substances. The percentage of agonist test plates that failed test plate acceptance criteria across the participating laboratories was 61% (33/54) (see **Section 7, Table 7-4**). NICEATM–ICCVAM reviewed the data to determine whether changes to test plate acceptance criteria could reduce the failure rates of comprehensive test plates without compromising the ability of the test method to detect and quantify test substance agonist or antagonist activity. A comparison was made between (1) qualitative and quantitative outcomes for test plates that met all acceptance criteria and (2) those that failed to meet one or more acceptance criteria (see **Section 7, Tables 7-5 and 7-6**). Qualitative outcomes are the positive or negative agonist classifications, and quantitative outcomes are EC₅₀ values.

Test plate acceptance criteria based on the DMSO control RLU values and E2 reference standard minimum fold induction values were not considered for modification because they are essential for monitoring background activity and reference estrogen performance.

Therefore, the test plate acceptance criteria that were considered for modification were the E2 EC₅₀ and Meth RLU control values. Based on this evaluation, it was determined that agonist test plate acceptance criteria could be modified without compromising the ability of the test method to detect and quantify test substance agonist activity. Following are the modifications:

- The requirement for the mean plate E2 reference standard EC₅₀ value to be within 2.5 times the SD of the historical mean EC₅₀ value was eliminated.
- The requirement for the mean plate Met control RLU value was changed from within 2.5 times the SD of the historical mean Met control RLU value to within 3 times the SD of the historical Met control RLU.

Changes to the agonist test plate acceptance criteria described above were used for Phases 2b, 3, and 4 testing.

Interpretation of Results from Comprehensive Agonist Testing

Positive classification—

- All test substances classified as positive for ER agonist activity should have a concentration–response curve consisting of a baseline, followed by a positive slope, and concluding in a plateau or peak. In some cases, only two of these characteristics (baseline–slope or slope–peak) may be defined.
- The line defining the positive slope must contain at least three points with nonoverlapping error bars (mean \pm SD). Points forming the baseline are excluded, but the linear portion of the curve may include the peak or first point of the plateau.

- A positive classification requires a response amplitude, the difference between baseline and peak, of at least 20% of the maximal value for the reference estrogen (i.e., 2000 RLUs when the maximal response value of the reference estrogen is adjusted to 10,000 RLUs).
- If possible, an EC₅₀ value should be calculated for each positive substance (**Section 4**).

Negative classification—

For all concentration–response curves that fail to meet the criteria for a positive response, test substances are classified as negative for agonist activity if all data points are below 20% of the maximal value for the reference estrogen (i.e., 2000 RLUs when the maximal response value of the reference estrogen is adjusted to 10,000 RLUs).

Inadequate—

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 agonist activity.

New Classification Scheme

The BG1Luc ER TA test method is intended as part of a weight-of-evidence approach to help prioritize substances for ED testing *in vivo*. Part of this prioritization procedure will be the classification of the test substance as positive or negative for either ER agonist or antagonist activity. There currently are no universally accepted standards for determining whether a substance is positive for ER agonist or antagonist activity. A common approach for the classification of substances as positive is to determine the lowest effective concentration (LEC), i.e., the concentration that is significantly different from the concurrent negative control (Judson et al. 2010; Martin et al. 2010). For the protocol standardization study and all phases of testing in the BG1Luc ER TA test method validation study, an LEC method was used to determine whether a test substance was positive or negative. Specifically:

- A substance is considered positive for agonist activity when the average adjusted RLU for a given concentration is greater than the mean DMSO control RLU value plus three times its SD (3X-SD).
- A substance is considered negative for agonist activity if the average adjusted RLU for a given concentration is at or below the mean DMSO control RLU value plus three times its SD.

Because this classification system appeared to work well during the protocol standardization study and the early phases of testing (Phase 1, Phase 2a, Phase 2b), it was used for Phase 3 and Phase 4 testing also. However, the data indicated that this classification scheme was resulting in an unacceptable level of false positives (71 out of 78 test substances were classified as positive) in the agonist assay. The contributing factors appeared to be as follows:

- The binary nature of the classification system, in which all substances will be classified as positive or negative, was too restrictive.
- Classification was based on individual values (not a curve shape) and did not accommodate high background levels or variability in test data. Consequently, single data points often exceed the 3X-SD DMSO control line because of the variability of the test, causing substances to be classified as positive.
- Many test substances caused a significant increase in background RLUs, resulting in a baseline that was near or above the 3X-SD DMSO control and thereby causing the substances to be classified as positive.
- No allowances were made for poor-quality test data; only plate acceptance criteria were considered for quality control purposes.

In light of the above, the SMT agreed on a new classification scheme that addressed each of these deficiencies. These new classification criteria were applied retrospectively to all test data for the assessment of test method accuracy (**Section 5**).

2.7.2 Comprehensive Antagonist Testing

Reference Standard and Control Concentrations Used for Antagonist Comprehensive Testing

- Ral, the anti-estrogenic reference standard, is plated in a serial dilution consisting of nine concentrations of Ral in duplicate (see **Table 2-2**).
- A single concentration of E2 (9.18×10^{-11} M), intended to provide 80% of the maximum E2 induction, is run in quadruplicate.
- The vehicle control (1% DMSO in EFM) is run in quadruplicate.
- All reference anti-estrogen and test wells contain a fixed concentration of E2 (9.18×10^{-11} M), intended to provide 80% of the maximum E2 induction.
- Tam, a weak antagonist reference standard, is plated in quadruplicate at 3.36×10^{-6} M.

Plate Design

- All 96 wells of the test plate are used during comprehensive testing. Two substances can be tested at eleven concentrations, in triplicate, on each plate. Starting concentrations and dilution factors are determined based on range finder results (**Section 2.6.2**). The plate design for comprehensive antagonist testing is provided in **Figure 2-4**.

Figure 2-4 96-Well Test Plate Layout for Comprehensive Antagonist Testing

TS1-1	TS1-2	TS1-3	TS1-4	TS1-5	TS1-6	TS1-7	TS1-8	TS1-9	TS1-10	TS1-11	VC
TS1-1	TS1-2	TS1-3	TS1-4	TS1-5	TS1-6	TS1-7	TS1-8	TS1-9	TS1-10	TS1-11	VC
TS1-1	TS1-2	TS1-3	TS1-4	TS1-5	TS1-6	TS1-7	TS1-8	TS1-9	TS1-10	TS1-11	VC
TS2-1	TS2-2	TS2-3	TS2-4	TS2-5	TS2-6	TS2-7	TS2-8	TS2-9	TS2-10	TS2-11	VC
TS2-1	TS2-2	TS2-3	TS2-4	TS2-5	TS2-6	TS2-7	TS2-8	TS2-9	TS2-10	TS2-11	Tam
TS2-1	TS2-2	TS2-3	TS2-4	TS2-5	TS2-6	TS2-7	TS2-8	TS2-9	TS2-10	TS2-11	Tam
Ral-1	Ral-2	Ral-3	Ral-4	Ral-5	Ral-6	Ral-7	Ral-8	Ral-9	E2	E2	Tam
Ral-1	Ral-2	Ral-3	Ral-4	Ral-5	Ral-6	Ral-7	Ral-8	Ral-9	E2	E2	Tam

Note: All reference and test wells contain a fixed concentration of E2 (4.90×10^{-11} M).

Abbreviations: E2 = E2 control; Ral = raloxifene; Tam = tamoxifen/E2 weak positive control; TS = test substance;

VC = vehicle control (DMSO [1% v/v EFM]).

Ral-1 to Ral-9 = concentrations of the raloxifene/E2 reference standard (from high to low)

TS1-1 to TS1-11 = concentrations (from high to low) of test substance 1 (TS1)

TS2-1 to TS2-11 = concentrations (from high to low) of test substance 2 (TS2)

Plate Acceptance Criteria for Comprehensive Antagonist Testing

- The mean DMSO control RLU values must be within 2.5 times the standard deviation of the historical DMSO control mean RLU value.
- Ral reduction must be greater than threefold. Reduction is calculated by dividing the averaged highest Ral reference RLU value by the averaged lowest Ral RLU value.

- The Ral reference standard curve should be sigmoidal in shape and have at least three values within the linear portion of the curve.
- The averaged E2 control RLU value must be within 2.5 times the standard deviation of the historical E2 control mean RLU value.
- The mean plate Tam control RLU value must be less than the mean E2 control RLU value minus three times the SD.

Following Phase 2a of the validation study, NICEATM–ICCVAM evaluated the failure rates of plates used to comprehensively test four antagonist substances. The percentage of antagonist test plates that failed test plate acceptance criteria across the participating laboratories was 38% (13/34) (see **Section 7.3.1**). NICEATM–ICCVAM reviewed the data to determine whether changes to test plate acceptance criteria could reduce the failure rates of comprehensive test plates without compromising the ability of the test method to detect and quantify test substance antagonist activity. A comparison was made of (1) qualitative and quantitative outcomes for test plates that met all acceptance criteria and (2) those that failed to meet one or more acceptance criteria (see **Section 7, Tables 7-2 and 7-3**). Qualitative outcomes are positive or negative antagonist classifications, and quantitative outcomes constitute IC₅₀ values.

Test plate acceptance criteria based on the DMSO control RLU values and the Ral reference standard minimum fold reduction values were not considered for modification because they are essential for monitoring background activity and reference antagonist performance. In addition, the E2 control test plate acceptance criterion was not considered for modification because it is essential for determining test substance antagonist activity.

Therefore, the test plate acceptance criteria that were considered for modification were the Ral IC₅₀ and flavone control RLU values. Based on this evaluation, it was determined that antagonist test plate acceptance criteria could be modified without compromising the ability of the test method to detect and quantify test substance agonist or antagonist activity. These modifications were as follows:

- The requirement that the mean plate Ral reference standard IC₅₀ value must be within 2.5 times the SD of the historical mean IC₅₀ value was eliminated. It was replaced with a requirement that the Ral reference standard curve should be sigmoidal in shape and have at least three values within the linear portion of the curve.
- The requirement that the mean plate flavone control RLU value must be within 2.5 times the SD of the historical mean flavone control RLU value was changed. The flavone control RLU value must now be less than three times the SD of the mean plate RLU value of the flavone control.

Changes to the antagonist test plate acceptance criteria described above were used for Phases 2b, 3, and 4 testing. However, as detailed in **Section 2.4.5**, further evaluation of the data after the study was completed led to the replacement of flavone with Tam as the weak positive control for ER antagonism.

Interpretation of Results from Comprehensive Antagonist Testing

As described above, criteria used to classify substances as positive or negative for ER agonism or antagonism were modified following a retrospective analysis of the data. These new classification criteria, provided above, were applied to all test data to assess test method accuracy (**Section 5**).

Positive classification—

- All test substances classified as positive for ER antagonist activity should have a concentration–response curve consisting of a baseline followed by a negative slope.
- The line defining the negative slope must contain at least three points with nonoverlapping error bars. Points forming the baseline are excluded, but the linear portion of the curve may include the first point of the plateau.

- A positive classification requires a response amplitude, the difference between baseline and bottom, of at least 20% of the maximal value for the reference estrogen (i.e., 8000 RLUs when the maximal response value of the reference estrogen is adjusted to 10,000 RLUs).
- The highest noncytotoxic concentrations of the test substance should be less than or equal to 1×10^{-5} M.

Negative classification—

- Test substances are classified as negative for antagonist activity if all data points are above the EC_{80} value (80% of the E2 response, or 8000 RLUs).

Inadequate—

- 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.

3.0 Substances Used for the Validation of the BG1Luc ER TA Test Method

3.1 Development of the List of 78 ICCVAM-Recommended Test Substances

ICCVAM previously recommended a list of 78 substances for use in validation studies of *in vitro* ER and AR binding and TA test methods (ICCVAM 2003a, 2006). The purpose of this list is to ensure that the usefulness and limitations of *in vitro* ER and AR binding and TA assays can be adequately characterized across a broad range of chemical classes and responses. These substances were selected based on information contained in the corresponding ICCVAM BRDs (ICCVAM 2002d, 2002a, 2002c, 2002b), as well as information obtained from publications reviewed or published after completion of the ICCVAM BRDs (**Annex N**). ICCVAM considered the following factors and criteria in compiling the list:

- Published or submitted data demonstrating reproducible positive or negative responses in multiple studies and/or test methods
- The extent to which these substances covered the range of responses (negative, weakly positive to strongly positive)
- Representative distribution of the proposed substances among chemical and product classes

To better evaluate test method specificity, approximately 25% of the total number of substances should be negative for the endpoint being measured. Substances that might interfere with transcriptional activation by altering metabolic pathways, such as RNA and protein synthesis, should be included.

The 78 ICCVAM-recommended substances used in the BG1Luc ER TA validation study are listed in **Table 3-1**. Physicochemical properties, including chemical structures, for each of the recommended substances are provided in **Annex I**.

Table 3-1 Reference Substances Tested for ER TA Activity

Substance	CASRN	MeSH Chemical Class ^a	Product Class ^b	Purity (%)	Manufacturer
12- <i>O</i> -Tetradecanoylphorbol-13-acetate	16561-29-8	Hydrocarbon (Cyclic)	Laboratory Chemical	>99.5	LC Laboratories
17 α -Estradiol	57-91-0	Steroid	Pharmaceutical, Veterinary Agent	99.5	Sigma-Aldrich Corporation
17 α -Ethinyl estradiol	57-63-6	Steroid	Pharmaceutical, Veterinary Agent	\geq 98.0	Sigma-Aldrich Corporation
17 β -Estradiol	50-28-2	Steroid	Pharmaceutical, Veterinary Agent	98.0	Sigma-Aldrich Corporation
17 β -Trenbolone	10161-33-8	Steroid	Pharmaceutical	96.6	Spectrum Chemicals & Laboratory Products
19-Nortestosterone	434-22-0	Steroid	Pharmaceutical, Veterinary Agent	98.0	Toronto Research Chemicals, Inc. (TRC)
2- <i>sec</i> -Butylphenol	89-72-5	Phenol	Chemical Intermediate, Pesticide Intermediate	98.0	Sigma-Aldrich Corporation

Substance	CASRN	MeSH Chemical Class ^a	Product Class ^b	Purity (%)	Manufacturer
2,4,5-Trichlorophenoxyacetic acid	93-76-5	Carboxylic Acid	Herbicide	99.3	Sigma-Aldrich Corporation
4-Androstenedione	63-05-8	Steroid	Pharmaceutical	98.6	Sigma-Aldrich Corporation/ Hiyoshi-International Laboratory USA
4-Cumylphenol	599-64-4	Phenol	Chemical Intermediate	99.9	Sigma-Aldrich Corporation
4-Hydroxytamoxifen	68047-06-3	Hydrocarbon (Cyclic)	Pharmaceutical	99.5	Sigma-Aldrich Corporation
4-Hydroxyandrostenedione	566-48-3	Steroid	Pharmaceutical	99.6	Sigma-Aldrich Corporation
4- <i>tert</i> -Octylphenol	140-66-9	Phenol	Chemical Intermediate, Pharmaceutical Intermediate	99.3	Chem Service, Inc.
5 α -Dihydrotestosterone	521-18-6	Steroid	Pharmaceutical	≥ 97.5	Sigma-Aldrich Corporation
Actinomycin D	50-76-0	Heterocyclic Compound, Polycyclic Compound	Laboratory Chemical, Pharmaceutical, Veterinary Agent	99.7	USB Corporation
Ammonium perchlorate	7790-98-9	Amine, Onium Compound	Industrial Chemical, Laboratory Chemical, Pharmaceutical	100.0	Sigma-Aldrich Corporation
Apigenin	520-36-5	Heterocyclic Compound	Dye, Natural Product, Pharmaceutical Intermediate	>99.0	Sigma-Aldrich Corporation
Apomorphine	58-00-4	Heterocyclic Compound	Pharmaceutical, Veterinary Agent	99.8	Sigma-Aldrich Corporation
Atrazine	1912-24-9	Heterocyclic Compound	Herbicide	98.0	Chem Service, Inc.
Bicalutamide	90357-06-5	Amide	Pharmaceutical	>99.5	LKT Laboratories, Inc.
Bisphenol A	80-05-7	Phenol	Chemical Intermediate, Flame Retardant, Fungicide	97.0	Sigma-Aldrich Corporation
Bisphenol B	77-40-7	Phenol	Chemical Intermediate, Flame Retardant, Fungicide	97.4	City Chemical LLC

Substance	CASRN	MeSH Chemical Class ^a	Product Class ^b	Purity (%)	Manufacturer
Butylbenzyl phthalate	85-68-7	Carboxylic Acid, Ester, Phthalic Acid	Plasticizer, Industrial Chemical	98.0	Sigma-Aldrich Corporation
Chrysin	480-40-0	Flavonoid, Heterocyclic Compound	Natural Product	99.8	Sigma-Aldrich Corporation
Clomiphene citrate	50-41-9	Amine, Carboxylic Acid, Heterocyclic Compound	Pharmaceutical	100.0	Sigma-Aldrich Corporation
Corticosterone	50-22-6	Steroid	Pharmaceutical	99.0	Sigma-Aldrich Corporation
Coumestrol	479-13-0	Heterocyclic Compound	Natural Product	98.0	BIOMOL International, Inc.
Cycloheximide	66-81-9	Heterocyclic Compound	Fungicide, Pharmaceutical, Veterinary Agent	99.0	Sigma-Aldrich Corporation
Cyproterone acetate	427-51-0	Steroid	Pharmaceutical	99.6	Sigma-Aldrich Corporation
Daidzein	486-66-8	Flavonoid, Heterocyclic Compound	Natural Product	≥97.5	Alfa Aesar GmbH
Dexamethasone	50-02-2	Steroid	Pharmaceutical, Veterinary Agent	99.0	Sigma-Aldrich Corporation
Di- <i>n</i> -butyl phthalate	84-74-2	Ester, Phthalic Acid	Cosmetic Ingredient, Industrial Chemical, Plasticizer	≥98.0	City Chemical LLC
Dibenzo[<i>a,h</i>]anthracene	53-70-3	Polycyclic Compound	Laboratory Chemical, Natural Product	99.9	Supelco Analytical
Dicofol	115-32-2	Hydrocarbon (Cyclic), Hydrocarbon (Halogenated)	Pesticide	98.0	Chem Service, Inc.
Diethylhexyl phthalate	117-81-7	Phthalic Acid	Pesticide Intermediate, Plasticizer	98.0	Alfa Aesar GmbH
Diethylstilbestrol	56-53-1	Hydrocarbon (Cyclic)	Pharmaceutical, Veterinary Agent	≥99.0	Sigma-Aldrich Corporation
Estrone	53-16-7	Steroid	Pharmaceutical, Veterinary Agent	99.0	Sigma-Aldrich Corporation
Ethyl paraben	120-47-8	Carboxylic Acid, Phenol	Pharmaceutical, Preservative	99.0	Sigma-Aldrich Corporation

Substance	CASRN	MeSH Chemical Class ^a	Product Class ^b	Purity (%)	Manufacturer
Fenarimol	60168-88-9	Heterocyclic Compound, Pyrimidine	Fungicide	99.5	Chem Service, Inc.
Finasteride	98319-26-7	Steroid	Pharmaceutical	>99.0	Sigma-Aldrich Corporation
Flavone	525-82-6	Flavonoid, Heterocyclic Compound	Natural Product, Pharmaceutical	99.7	Sigma-Aldrich Corporation
Fluoranthene	206-44-0	Polycyclic Compound	Industrial Chemical, Laboratory Chemical, Pharmaceutical Intermediate	99.6	Sigma-Aldrich Corporation
Fluoxymestrone	76-43-7	Steroid	Pharmaceutical	>99.0	Sigma-Aldrich Corporation
Flutamide	13311-84-7	Amide	Pharmaceutical, Veterinary Agent	100.0	Sigma-Aldrich Corporation
Genistein	446-72-0	Flavonoid, Heterocyclic Compound	Natural Product, Pharmaceutical	98.8	Sigma-Aldrich Corporation
Haloperidol	52-86-8	Ketone	Pharmaceutical, Veterinary Agent	>99.0	Sigma-Aldrich Corporation
Hydroxyflutamide	52806-53-8	Amide	Pharmaceutical	99.4	LKT Laboratories, Inc.
Kaempferol	520-18-3	Flavonoid, Heterocyclic Compound	Natural Product	99.0	INDOFINE Chemical Company, Inc.
Kepone	143-50-0	Hydrocarbon (Halogenated)	Pesticide	>99.9	Supelco Analytical
Ketoconazole	65277-42-1	Heterocyclic Compound	Pharmaceutical	>99.0	Sigma-Aldrich Corporation
L-Thyroxine	51-48-9	Amino Acid	Pharmaceutical, Veterinary Agent	98.0	Sigma-Aldrich Corporation
Linuron	330-55-2	Urea	Herbicide	99.5	Chem Service, Inc.
Medroxyprogesterone acetate	71-58-9	Steroid	Pharmaceutical	99.0	Sigma-Aldrich Corporation
meso-Hexestrol	84-16-2	Steroid	Pharmaceutical, Veterinary Agent	99.3	City Chemical LLC
Methyl testosterone	58-18-4	Steroid	Pharmaceutical, Veterinary Agent	99.0	Sigma-Aldrich Corporation
Mifepristone	84371-65-3	Steroid	Pharmaceutical	99.1	Sigma-Aldrich Corporation

Substance	CASRN	MeSH Chemical Class ^a	Product Class ^b	Purity (%)	Manufacturer
Morin	480-16-0	Flavonoid, Heterocyclic Compound	Dye, Natural Product, Pharmaceutical Intermediate	95.3	TCI America
Nilutamide	63612-50-0	Heterocyclic Compound, Imidazole	Pharmaceutical	100.0	Sigma-Aldrich Corporation
Norethynodrel	68-23-5	Steroid	Pharmaceutical	≥95.0	Research Plus Inc.
<i>o,p'</i> -DDT	789-02-6	Hydrocarbon (Halogenated)	Pesticide	98.9	Chem Service, Inc.
Oxazepam	604-75-1	Heterocyclic Compound	Pharmaceutical, Veterinary Agent	99.5	Sigma-Aldrich Corporation
<i>p</i> -n-Nonylphenol	104-40-5	Phenol	Chemical Intermediate	99.6	Alfa Aesar GmbH
<i>p,p'</i> -Methoxychlor	72-43-5	Hydrocarbon (Halogenated)	Pesticide, Veterinary Agent	99.1	Chem Service, Inc.
<i>p,p'</i> -DDE	72-55-9	Hydrocarbon (Halogenated)	Pesticide Intermediate	99.0	Sigma-Aldrich Corporation
Phenobarbital	50-06-6	Heterocyclic Compound, Pyrimidine	Pharmaceutical, Veterinary Agent	100.0	Spectrum Chemical Manufacturing Corp.
Phenolphthalin	81-90-3	Carboxylic Acid, Phenol	Dye, Laboratory Chemical	95.0	Sigma-Aldrich Corporation
Pimozide	2062-78-4	Heterocyclic Compound	Pharmaceutical	>99.0	Sigma-Aldrich Corporation
Procymidone	32809-16-8	Polycyclic Compound	Fungicide	99.0	Chem Service, Inc.
Progesterone	57-83-0	Steroid	Pharmaceutical, Veterinary Agent	≥99.0	Sigma-Aldrich Corporation
Propylthiouracil	51-52-5	Heterocyclic Compound, Pyrimidine	Pharmaceutical, Veterinary Agent	100.0	Sigma-Aldrich Corporation
Raloxifene HCl	82640-04-8	Hydrocarbon (Cyclic)	Pharmaceutical	100.0	Sigma-Aldrich Corporation
Reserpine	50-55-5	Heterocyclic Compound, Indole	Pharmaceutical, Veterinary Agent	98.0	Sigma-Aldrich Corporation
Resveratrol	501-36-0	Hydrocarbon (Cyclic)	Natural Product	≥99.0	Sigma-Aldrich Corporation
Sodium azide	26628-22-8	Azide, Salt (Inorganic)	Chemical Intermediate, Fungicide, Herbicide	99.7	Sigma-Aldrich Corporation

Substance	CASRN	MeSH Chemical Class ^a	Product Class ^b	Purity (%)	Manufacturer
Spironolactone	52-01-7	Lactone, Steroid	Pharmaceutical	99.7	Sigma-Aldrich Corporation
Tamoxifen	10540-29-1	Hydrocarbon (Cyclic)	Pharmaceutical	≥99.0	Sigma-Aldrich Corporation
Testosterone	58-22-0	Steroid	Pharmaceutical, Veterinary Agent	>99.0	Sigma-Aldrich Corporation
Vinclozolin	50471-44-8	Heterocyclic Compound	Fungicide	99.5	Chem Service, Inc.

Abbreviations: CASRN = CAS Registry Number (American Chemical Society); MeSH = Medical Subject Headings (U.S. National Library of Medicine).

^a Substances 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>).

^b Substances 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>).

The following sections describe the subsets of this list that were used to evaluate BG1Luc ER TA test method accuracy and reproducibility, as well as the rationale for selection of each subset. The data and rationale used to establish a reference classification for each of the 78 substances are also discussed.

3.2 Substances Used to Evaluate Test Method Accuracy

Accuracy is the closeness of agreement between a test method result and an accepted reference value, the extent to which a test method obtains the “correct” answer. It is a measure of test method performance. The ICCVAM list of 78 recommended reference substances was developed to assess test method performance of four different test methods: (1) ER binding, (2) ER TA, (3) AR binding, and (4) AR TA. Each test method has its own unique set of “correct” classifications for these substances. However, this validation study focused only on the ER TA and on the ability of the BG1Luc ER TA test method to detect substances with *in vitro* ER TA agonist or antagonist activity. Therefore, each of the 78 reference substances was assigned a classification specific for ER TA agonist and ER TA antagonist activity based on a preponderance of evidence found in a review of the scientific literature.

NICEATM conducted a broad literature search using online sources including Scopus[®] (Elsevier BV), PubMed[®] (U.S. National Library of Medicine), and Web of ScienceSM (Thomson Reuters). Publically available information from U.S. government agencies and the OECD was also considered. This search strategy yielded 103 publications with relevant ER TA data. The following information was extracted from each reference and is provided in **Annex N**:

- Name and purity of the substance being tested
- Characteristics of cell line (e.g., name of cell line, tissue of origin)
- Reporter gene construct (e.g., ER source, reporter vector, endpoint measured, whether cell toxicity measurements were made, and transfection method [i.e., whether stable or transient])
- Assay type (i.e., agonism or antagonism)
- Any relevant quantitative information (e.g., IC₅₀, EC₅₀)

There was considerable disparity in the number of ER TA references applicable to each substance. Therefore, the following criteria were used to classify each reference substance with respect to ER TA agonist and antagonist activity:

- A substance was classified as positive (POS) if it was reported as positive in >50% of referenced ER TA studies.
- A substance was classified as negative (NEG) if it was reported as negative in all referenced ER TA studies (at least two studies were required for negative classification).
- A substance was classified as presumed positive (PP) if it was positive in 50% or fewer referenced ER TA studies or if it was reported as positive in the single study conducted.
- A substance was classified as presumed negative (PN) if it was reported as negative in a single ER TA study.
- Substances without data were classified as PP or PN based on other available information, including their known mechanism of action or their responses in other ER assays.

Table 3-2 provides the following information:

- A summary of the literature findings
- Results from the validation studies of the CERI-STTA and OECD uterotrophic methods for all 78 ICCVAM reference substances. (See **Section 3.3** for discussion of the CERI-STTA, an ER TA test method adopted in the United States as EPA OPPTS 890.1300 and internationally as OECD TG 455.)
- Resulting ER TA classifications for agonist and antagonist activity based on the criteria provided above

Table 3-2 Reference Data Summary for ER Agonist and Antagonist TA Assays

ICCVAM Reference Substance	CASRN	ER TA Agonist Activity ^a	ER TA Antagonist Activity ^b	ER Binding Activity ^c	CERI ER TA Activity ^d	Uterotrophic Activity ^e
12- <i>O</i> -Tetradecanoylphorbol-13-acetate	16561-29-8	PN (NT)	PN (NT)	PN (NT)	NT	NT
17 α Estradiol	57-91-0	POS (10/10)	PN (0/1)	POS (15/15)	POS	POS (NT/+)
17 α Ethinyl estradiol	57-63-6	POS (21/21)	NEG (0/9)	POS (32/32)	POS	POS (+/+)
17 β Estradiol	50-28-2	POS (226/226)	PN (0/1)	POS (160/160)	POS	NT
17 β -Trenbolone	10161-33-8	PP (1/1)	PN (NT)	PN (NT)	POS	NT
19-Nortestosterone*	434-22-0	POS (3/3)	PN (NT)	PP (1/7)	NT	NT
2- <i>sec</i> -Butylphenol	89-72-5	PN (0/1)	PN (NT)	POS (2/2)	NEG	NT
2,4,5-Trichlorophenoxyacetic acid	93-76-5	PP (1/3)	PN (0/2)	PP (1/3)	NT	NT
4-Androstenedione	63-05-8	PP (1/1)	PN (0/1)	PP (1/5)	NEG	NT
4-Cumylphenol	599-64-4	POS (4/4)	PN (NT)	POS (3/3)	POS	NT
4-Hydroxyandrostenedione*	566-48-3	PP (1/2)	PN (NT)	PP (NT)	NT	NT
4-Hydroxytamoxifen	68047-06-3	PP (17/56)	POS (27/27)	POS (36/36)	NT	NT
4- <i>tert</i> -Octylphenol	140-66-9	POS (20/23)	PN (NT)	POS (20/20)	POS	POS (NT/+)
5 α -Dihydrotestosterone	521-18-6	POS (15/17)	NEG (0/3)	POS (17/18)	NT	POS (NT/+)
Actinomycin D	50-76-0	PN (NT)	PN (NT)	PN (NT)	NT	NT

ICCVAM Reference Substance	CASRN	ER TA Agonist Activity ^a	ER TA Antagonist Activity ^b	ER Binding Activity ^c	CERI ER TA Activity ^d	Uterotrophic Activity ^e
Ammonium perchlorate	7790-98-9	PN (NT)	PN (NT)	PN (NT)	NT	NT
Apigenin	520-36-5	POS (25/25)	NEG (0/11)	POS	POS	NT
Apomorphine	58-00-4	PN (NT)	PN (NT)	PN (NT)	NT	NT
Atrazine	1912-24-9	NEG (0/29)	PN (0/1)	PP (2/19)	NEG	NT
Bicalutamide	90357-06-5	NEG (0/5)	PN (NT)	PN (NT)	NT	NT
Bisphenol A	80-05-7	POS (64/64)	NEG (0/12)	POS (46/47)	POS	POS (+/+)
Bisphenol B	77-40-7	POS (5/5)	PN (0/1)	POS (2/2)	POS	POS (NT/+)
Butylbenzyl phthalate	85-68-7	POS (11/13)	NEG (0/3)	POS (10/19)	POS	NEG (-/-)
Chrysin*	480-40-0	POS (6/9)	NEG (0/4)	PP (2/10)	NT	NT
Clomiphene citrate	50-41-9	POS (3/4)	PP (1/1)	POS (8/8)	POS	NT
Corticosterone	50-22-6	NEG (0/5)	PN (0/2)	NEG (0/6)	NEG	NT
Coumestrol	479-13-0	POS (29/29)	NEG (0/8)	POS (38/38)	POS	NT
Cycloheximide	66-81-9	PN (NT)	PP (NT)	PN (NT)	NT	NT
Cyproterone acetate	427-51-0	PP (1/6)	PN (0/1)	PP (1/2)	NT	NT
Daidzein	486-66-8	POS (38/38)	NEG (0/6)	POS (32/35)	POS	POS (NT/+)
Dexamethasone	50-02-2	PP (2/6)	PN (NT)	PP (1/4)	NT	NT
Di- <i>n</i> -butyl phthalate	84-74-2	PP (5/10)	NEG (0/3)	POS (7/13)	NT	NEG (-/-)
Dibenzo[<i>a,h</i>] anthracene	53-70-3	PP (1/2)	PP (NT)	PN (0/1)	NT	NT
Dicofol*	115-32-2	POS (4/6)	NEG (0/2)	POS (2/2)	NT	NT
Diethylhexyl phthalate	117-81-7	PP (4/9)	NEG (0/3)	PP (4/8)	NEG	NEG (NT/-)
Diethylstilbestrol	56-53-1	POS (41/41)	NEG (0/2)	POS (52/52)	POS	NT
Estrone	53-16-7	POS (25/27)	PN (0/1)	POS (29/29)	POS	POS (NT/+)
Ethyl paraben	120-47-8	POS (5/5)	PN (NT)	POS (4/5)	POS	NT
Fenarimol	60168-88-9	POS (5/6)	PN (0/1)	POS (2/2)	NT	NT
Finasteride	98319-26-7	PN (NT)	PN (0/1)	PN (0/1)	NT	NT
Flavone	525-82-6	PP (2/5)	PN (0/1)	PP (3/13)	NT	NT
Fluoranthene	206-44-0	PN (NT)	PN (NT)	PN (0/1)	NT	NT
Fluoxymestrone	76-43-7	PN (NT)	PN (NT)	PN (0/1)	NT	NT
Flutamide	13311-84-7	NEG (0/5)	PN (0/1)	NEG (0/2)	NT	NT
Genistein	446-72-0	POS (99/101)	NEG (0/13)	POS (64/64)	POS	POS (+/+)
Haloperidol	52-86-8	PN (0/1)	PN (NT)	PN (0/1)	NT	NT
Hydroxyflutamide	52806-53-8	NEG (0/2)	PN (NT)	PP (1/4)	NT	NT
Kaempferol	520-18-3	POS (22/22)	NEG (0/9)	POS (19/19)	POS	NT
Kepone	143-50-0	POS (13/17)	NEG (0/2)	POS (14/15)	POS	NT
Ketoconazole	65277-42-1	PN (0/1)	PN (NT)	PN (0/1)	NEG	NT
L-Thyroxine	51-48-9	POS (2/3)	PN (NT)	POS (2/2)	NT	NT

ICCVAM Reference Substance	CASRN	ER TA Agonist Activity ^a	ER TA Antagonist Activity ^b	ER Binding Activity ^c	CERI ER TA Activity ^d	Uterotrophic Activity ^e
Linuron	330-55-2	NEG (0/7)	PN (NT)	POS (2/3)	NEG	NT
Medroxyprogesterone acetate	71-58-9	PP (1/2)	PN (0/1)	POS (2/2)	NEG	NT
meso-Hexestrol	84-16-2	POS (3/3)	PN (NT)	POS (11/11)	NT	NT
Methyl testosterone	58-18-4	POS (4/5)	PN (0/1)	POS (2/3)	POS	NT
Mifepristone	84371-65-3	PP (3/6)	NEG (0/3)	POS (4/6)	NEG	NT
Morin	480-16-0	PP (1/1)	PN (NT)	POS (3/3)	POS	NT
Nilutamide	63612-50-0	PN (NT)	PN (NT)	PN (NT)	NT	NT
Norethynodrel	68-23-5	POS (4/4)	NEG (0/2)	POS (7/7)	POS	NT
<i>o,p'</i> -DDT	789-02-6	POS (24/25)	NEG (0/3)	POS (20/22)	NT	POS (+/NT)
Oxazepam	604-75-1	PN (NT)	PN (NT)	PN (NT)	NT	NT
<i>p</i> -n-Nonylphenol	104-40-5	POS (9/9)	NEG (0/2)	POS (21/21)	NEG	IC (+/-)
<i>p,p'</i> -DDE	72-55-9	POS (5/7)	NEG (2/2)	PP (5/15)	NT	NT
<i>p,p'</i> -Methoxychlor	72-43-5	POS (23/26)	PP (1/5)	POS (16/26)	POS	IC (+/-)
Phenobarbital	50-06-6	NEG (0/2)	PN (NT)	PN (0/1)	NT	NT
Phenolphthalin	81-90-3	PN (0/1)	PN (NT)	POS (2/2)	NEG	NT
Pimozide	2062-78-4	PN (NT)	PN (NT)	PN (NT)	NT	NT
Procymidone	32809-16-8	NEG (0/4)	PN (NT)	PP (2/5)	NT	NT
Progesterone	57-83-0	PP (3/15)	NEG (0/2)	PP (2/20)	NEG	NT
Propylthiouracil	51-52-5	PN (NT)	PN (NT)	PN (NT)	NT	NT
Raloxifene HCl*	82640-04-8	PP (7/31)	POS (13/13)	POS (16/16)	NEG	NT
Reserpine	50-55-5	PN (0/1)	PN (NT)	PN (0/1)	NEG	NT
Resveratrol*	501-36-0	POS (24/37)	NEG (0/16)	POS (9/12)	NT	NT
Sodium azide	26628-22-8	PN (0/1)	PN (NT)	PN (NT)	NT	NT
Spironolactone	52-01-7	NEG (0/3)	PN (NT)	PN (0/1)	NEG	NT
Tamoxifen	10540-29-1	POS (15/22)	POS (20/22)	POS (46/46)	POS	NT
Testosterone	58-22-0	PP (4/9)	PN (0/1)	PP (5/12)	POS	NT
Vinclozolin	50471-44-8	PP (6/13)	PN (0/1)	POS (3/5)	POS	NT

Abbreviations: CASRN = CAS Registry Number (American Chemical Society); CERI = Chemicals Evaluation and Research Institute, Japan; ER = estrogen receptor; IC = inconclusive; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; NEG = negative; NT = not tested; PN = presumed negative; POS = positive; PP = presumed positive; TA = transcriptional activation.

* Substance in original list (ICCVAM 2003a) was replaced in the *Finalized Addendum to ICCVAM Evaluation of In Vitro Test Methods for Detecting Potential Endocrine Disruptors* (ICCVAM 2006) due to excessive cost or limited availability.

^a Values in parentheses are the number of positive ER TA agonist studies/total number of studies identified in the 2010 literature update.

^b Values in parentheses are the number of positive ER TA antagonist studies/total number of studies (2010).

^c Values in parentheses are the number of positive binding studies/total number of studies (2010).

^d Chemicals Evaluation and Research Institute (CERI), Japan, evaluated substances using the OECD Stably Transfected Human Estrogen Receptor- α Transcriptional Activation Assay for Detection of Estrogenic Agonist-Activity, described in OECD Test Guideline (TG) 455 (OECD 2009; Takeyoshi 2006).

^e Values in parentheses are the *in vivo* uterotrophic classifications using OECD study data/CERI study data (Kanno et al. 2003a, 2003b; Takeyoshi 2006). A consensus *in vivo* uterotrophic classification was made when OECD and CERI data were in agreement. When *in vivo* uterotrophic data from OECD and CERI provided conflicting classifications, the overall classification was “inconclusive” (IC).

Of the 78 substances listed in **Table 3-2**, only those substances that could be definitively classified as POS or NEG were used to assess test method accuracy (substances classified as PP or PN were not considered). This resulted in the use of 48 unique substances to assess accuracy. Separate lists were generated for evaluating accuracy based on agonist (42 substances: 33 positive, 9 negative) and antagonist (25 substances: 3 positive, 22 negative) activity. Nineteen substances were common to both reference lists. The 42 reference substances used to assess accuracy based on ER agonist activity are listed in **Table 3-3**, and the 25 reference substances used to assess accuracy based on ER antagonist activity are provided in **Table 3-4**. Substances that were classified as positive but for which EC₅₀ values were not reported are shown as “not calculated” (NC).

Table 3-3 Test Substances Used for Statistical Determination of ER TA Agonist Assay Accuracy

Substance	CASRN	ICCVAM Consensus Classification ^a	Mean EC ₅₀ (M) ^b
17α-Estradiol	57-91-0	POS	1.92 × 10 ⁻⁷
17α-Ethinyl estradiol	57-63-6	POS	2.44 × 10 ⁻⁹
17β-Estradiol	50-28-2	POS	1.33 × 10 ⁻⁸
19-Nortestosterone	434-22-0	POS	1.30 × 10 ⁻⁷
4-Cumylphenol	599-64-4	POS	3.22 × 10 ⁻⁷
4-tert-Octylphenol	140-66-9	POS	4.54 × 10 ⁻⁶
5α-Dihydrotestosterone	521-18-6	POS	2.50 × 10 ⁻⁷
Apigenin	520-36-5	POS	7.64 × 10 ⁻⁷
Atrazine	1912-24-9	NEG	NA
Bicalutamide	90357-06-5	NEG	NA
Bisphenol A	80-05-7	POS	3.69 × 10 ⁻⁶
Bisphenol B	77-40-7	POS	4.18 × 10 ⁻⁵
Butylbenzyl phthalate	85-68-7	POS	5.10 × 10 ⁻⁶
Chrysin	480-40-0	POS	NC
Clomiphene citrate	50-41-9	POS	5.00 × 10 ⁻⁹
Corticosterone	50-22-6	NEG	NA
Coumestrol	479-13-0	POS	2.00 × 10 ⁻⁷
Daidzein	486-66-8	POS	3.05 × 10 ⁻⁶
Dicofol	115-32-2	POS	7.05 × 10 ⁻⁶
Diethylstilbestrol	56-53-1	POS	1.29 × 10 ⁻⁷
Estrone	53-16-7	POS	8.33 × 10 ⁻⁸

Substance	CASRN	ICCVAM Consensus Classification ^a	Mean EC ₅₀ (M) ^b
Ethyl paraben	120-47-8	POS	5.00×10^{-5}
Fenarimol	60168-88-9	POS	7.00×10^{-6}
Flutamide	13311-84-7	NEG	NA
Genistein	446-72-0	POS	1.66×10^{-5}
Hydroxyflutamide	52806-53-8	NEG	NA
Kaempferol	520-18-3	POS	1.60×10^{-7}
Kepone	143-50-0	POS	NC
L-Thyroxine	51-48-9	POS	5.00×10^{-9}
Linuron	330-55-2	NEG	NA
<i>meso</i> -Hexestrol	84-16-2	POS	1.13×10^{-10}
Methyl testosterone	58-18-4	POS	1.38×10^{-6}
Norethynodrel	68-23-5	POS	6.59×10^{-8}
<i>o,p'</i> -DDT	789-02-6	POS	1.67×10^{-4}
<i>p</i> -n-Nonylphenol	104-40-5	POS	1.59×10^{-6}
<i>p,p'</i> -Methoxychlor	72-43-5	POS	1.56×10^{-4}
<i>p,p'</i> -DDE	72-55-9	POS	3.00×10^{-6}
Phenobarbital	50-06-6	NEG	NA
Procymidone	32809-16-8	NEG	NA
Resveratrol	501-36-0	POS	7.86×10^{-6}
Spironolactone	52-01-7	NEG	NA
Tamoxifen	10540-29-1	POS	1.35×10^{-6}

Abbreviations: CASRN = CAS Registry Number (American Chemical Society); EC₅₀ = half-maximal effective concentration; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; M = molar; NA = not applicable; NC = not calculated; NEG = negative; POS = positive.

^a Estrogenic activity based on a literature review for effects on ER binding, ER TA based on CER1, and uterotrophic response.

^b Mean EC₅₀ calculated from values reported in the literature.

Table 3-4 Test Substances Used for Statistical Determination of ER TA Antagonist Assay Accuracy

Substance	CASRN	ICCVAM Consensus Classification ^a	Mean IC ₅₀ (M) ^b
17 α -Ethinyl estradiol	57-63-6	NEG	NA
4-Hydroxytamoxifen	68047-06-3	POS	1.93 $\times 10^{-8}$
5 α -Dihydrotestosterone	521-18-6	NEG	NA
Apigenin	520-36-5	NEG	NA
Bisphenol A	80-05-7	NEG	NA
Butylbenzyl phthalate	85-68-7	NEG	NA
Chrysin	480-40-0	NEG	NA
Coumestrol	479-13-0	NEG	NA
Daidzein	486-66-8	NEG	NA
Di- <i>n</i> -butyl phthalate	84-74-2	NEG	NA
Dicofol	115-32-2	NEG	NA
Diethylhexyl phthalate	117-81-7	NEG	NA
Diethylstilbestrol	56-53-1	NEG	NA
Genistein	446-72-0	NEG	NA
Kaempferol	520-18-3	NEG	NA
Kepone	143-50-0	NEG	NA
Mifepristone	84371-65-3	NEG	NA
Norethynodrel	68-23-5	NEG	NA
<i>o,p'</i> -DDT	789-02-6	NEG	NA
<i>p</i> -n-Nonylphenol	104-40-5	NEG	NA
<i>p,p'</i> -DDE	72-55-9	NEG	NA
Progesterone	57-83-0	NEG	NA
Raloxifene HCl	82640-04-8	POS	6.23 $\times 10^{-8}$
Resveratrol	501-36-0	NEG	NA
Tamoxifen	10540-29-1	POS	1.26 $\times 10^{-6}$

Abbreviations: CASRN = CAS Registry Number (American Chemical Society); IC₅₀ = half-maximal inhibitory concentration; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; M = molar; NA = not applicable; NEG = negative; POS = positive.

^a Estrogenic activity based on a literature review for effects on ER binding, ER TA based on CER1, and uterotrophic response.

^b Mean IC₅₀ calculated from values reported in the literature.

3.3 Substances Used to Evaluate Concordance with Other Accepted Methods

The primary accuracy evaluation described in **Section 5.0** compares the test substance classification by the BG1Luc ER TA test method to the ICCVAM reference classification of that same substance, as outlined in **Section 3.2**. However, this evaluation also considered concordance with other methods currently accepted by regulators to evaluate estrogenic activity. Following are the most commonly used methods:

- *In vitro* stably transfected transactivation assay (STTA) by the Japanese Chemicals Evaluation and Research Institute (CERI) using the hER α - HeLa-9903 cell line (CERI-STTA) for ER agonists
- *In vitro* ER binding assays
- *In vivo* rodent uterotrophic bioassay

The substances used in the concordance analyses with each of these methods, and the rationale for their selection, are detailed in **Sections 3.3.1 to 3.3.3**.

3.3.1 Substances Used to Evaluate BG1 Luc ER TA Concordance with the CERI-STTA

The *in vitro* assessment of ER TA activity is included in Tier 1 of the EPA's EDSP screening battery and has been incorporated into the OECD Conceptual Framework for the Testing and Assessment of Endocrine Disrupting Chemicals as Level 2 assays to provide mechanistic information to prioritize testing. At present, the CERI-STTA is the only *in vitro* ER TA test method that has been adopted by regulatory agencies for identifying substances with potential ER agonist activity. This test method has recently been adopted in the United States as OPPTS 890.1300: Estrogen Receptor Transcriptional Activation (Human Cell Line [HeLa-9903]) (EPA 2009). Internationally, it has been adopted as OECD TG 455 (OECD 2009). The hER α -HeLa-9903 cell line is derived from a human cervical tumor with two stably inserted constructs: (1) the hER α expression construct (encoding the full-length human receptor) and (2) a firefly luciferase reporter construct bearing five tandem repeats of a vitellogenin ERE driven by a mouse metallothionein promoter TATA element (OECD 2009; Takeyoshi 2006).

There were 41 substances common to both the BG1Luc ER TA and CERI-STTA validation studies. CERI-STTA data (ER TA agonist classifications) for these 41 reference substances are included in **Table 3-2**. Using these data, ICCVAM compared concordance between agonist classifications from the BG1Luc ER TA and CERI-STTA validation studies (**Section 5**).

3.3.2 Substances Used to Evaluate BG1Luc ER TA Concordance with *In Vitro* ER Binding Test Methods

The *in vitro* assessment of ER binding activity is included in Tier 1 of the EPA's EDSP screening battery and has been incorporated into the OECD Conceptual Framework as Level 2 assays to provide mechanistic information to prioritize testing. *In vitro* ER binding assays identify substances that can bind to the ER, whereas *in vitro* ER TA assays measure the ability of a test substance to activate or inhibit the transactivation of a reporter gene via ER-mediated pathways. Accordingly, the ability of a test substance to bind to the ER *in vitro* suggests (but does not demonstrate) the ability of the substance to activate or inhibit *in vitro* ER-mediated transactivation. In order to determine the extent of agreement between the BG1Luc ER TA and ER binding data, ICCVAM evaluated concordance using data from the BG1Luc ER TA validation study and published ER binding data (**Section 5**).

Classification of the reference substances with respect to *in vitro* ER binding was based on a preponderance of evidence found in a review of the scientific literature, as described for BG1Luc ER TA assays in **Section 3.2**. Relevant information from 67 publications describing *in vitro* ER binding data was extracted and is provided in **Annex N**.

A summary of the ER binding literature data for all ICCVAM reference substances is provided in **Table 3-2**, along with the resulting ER binding classifications.

3.3.3 Substances Used to Evaluate BG1Luc ER TA Concordance with the *In Vivo* Uterotrophic Bioassay

As stated in OECD TG 440 (Uterotrophic Bioassay in Rodents), the uterotrophic bioassay is a short-term screening test that evaluates the ability of a substance to elicit estrogenic activity (Kanno et al. 2003a, 2003b; OECD 2007; Owens and Koeter 2003). In this *in vivo* test method, the uterus responds to estrogens initially with an increase in weight resulting from water imbibition, followed by further weight gain due to increased tissue growth. The uterotrophic bioassay is included in Level 3 of the OECD Conceptual Framework as an *in vivo* assay providing data about estrogenicity. The rat uterotrophic bioassay is also included as one of the *in vivo* methods in the EPA's EDSP Tier 1 screening battery. In order to determine the extent of agreement between the BG1Luc ER TA test method and the rat uterotrophic bioassay, ICCVAM conducted a concordance evaluation using data from the BG1Luc ER TA validation study and published uterotrophic bioassay data (**Section 5**).

Classification of the reference substances with respect to *in vivo* rodent uterotrophic activity was based on data from studies sponsored by the OECD (OECD 2007) and studies that were conducted in conjunction with CERI (ER TA assay validation studies) (Kanno et al. 2003a, 2003b). Combined, these studies tested 15 substances from the list of 78 ICCVAM reference substances. The *in vivo* uterotrophic data used to compare BG1Luc ER TA validation study agonist results were selected using the following criteria:

- Substances that tested positive in both the OECD and CERI studies (three substances)
- Substances that tested negative in both the OECD and CERI studies (two substances)
- Substances that tested positive or negative in at least one OECD or CERI study but that were not tested in both studies (seven positive and one negative)

Two substances were positive in one study but negative in the other. These substances were defined as “inconclusive” and were not used in the comparison.

Classification of the 15 reference substances with respect to uterotrophic activity is provided in **Table 3-2**.

3.4 Substances Tested in Each Phase of Validation

As described in **Section 1.0**, the test method validation was conducted in four consecutive phases in order to identify and resolve sources of variation early in the validation process. During Phase 1 of the validation, the three participating centers (ICCVAM, ECVAM, and JaCVAM) generated historical databases. In Phases 2 through 4, the 78 ICCVAM reference substances were tested. Substances used in each phase of the agonist and antagonist testing are listed in **Table 3-5** and **Table 3-6**, respectively.

Table 3-5 Agonist Substances by Study Phase

Study Phase	Substance	CASRN	MeSH Chemical Class^a	Product Class^b	ICCVAM Consensus Classification^c
1	17 β -Estradiol	50-28-2	Steroid	Pharmaceutical, Veterinary Agent	POS
1	<i>p,p'</i> -Methoxychlor	72-43-5	Hydrocarbon (Halogenated)	Pesticide, Veterinary Agent	POS
2a	Bisphenol A	80-05-7	Phenol	Chemical Intermediate, Flame Retardant, Fungicide	POS
2a	Bisphenol B	77-40-7	Phenol	Chemical Intermediate, Flame Retardant, Fungicide	POS
2a	Corticosterone	50-22-6	Steroid	Pharmaceutical	NEG
2a	Diethylstilbestrol	56-53-1	Hydrocarbon (Cyclic)	Pharmaceutical, Veterinary Agent	POS
2b	17 α -Ethinyl estradiol	57-63-6	Steroid	Pharmaceutical, Veterinary Agent	POS
2b	Atrazine	1912-24-9	Heterocyclic Compound	Herbicide	NEG
2b	Butylbenzyl phthalate	85-68-7	Carboxylic Acid, Phthalic Acid	Pharmaceutical, Veterinary Agent	POS
2b	Flavone	525-82-6	Flavonoid, Heterocyclic Compound	Natural Product, Pharmaceutical	PP
2b	Genistein	446-72-0	Flavonoid, Heterocyclic Compound	Natural Product, Pharmaceutical	POS
2b	<i>o,p'</i> -DDT	789-02-6	Hydrocarbon (Halogenated)	Pesticide	POS
2b	<i>p</i> -n-Nonylphenol	104-40-5	Phenol	Chemical Intermediate	POS
2b	Vinclozolin	50471-44-8	Heterocyclic Compound	Fungicide	PP
3	12- <i>O</i> -Tetradecanoylphorbol-13-acetate	16561-29-8	Hydrocarbon (Cyclic)	Laboratory Chemical	PN

Study Phase	Substance	CASRN	MeSH Chemical Class ^a	Product Class ^b	ICCVAM Consensus Classification ^c
3	17 α -Estradiol	57-91-0	Steroid	Pharmaceutical, Veterinary Agent	POS
3	17 β -Estradiol	50-28-2	Steroid	Pharmaceutical, Veterinary Agent	POS
3	2- <i>sec</i> -Butylphenol	89-72-5	Phenol	Chemical Intermediate, Pesticide Intermediate, Plasticizer	PN
3	2,4,5-Trichlorophenoxyacetic acid	93-76-5	Carboxylic Acid	Herbicide	PP
3	4-Androstenedione	63-05-8	Steroid	Pharmaceutical	PP
3	4-Cumylphenol	599-64-4	Phenol	Chemical Intermediate	POS
3	4-Hydroxytamoxifen	68047-06-3	Hydrocarbon (Cyclic)	Pharmaceutical	PP
3	4- <i>tert</i> -Octylphenol	140-66-9	Phenol	Chemical Intermediate, Pharmaceutical Intermediate	POS
3	5 α -Dihydrotestosterone	521-18-6	Steroid	Pharmaceutical	POS
3	Actinomycin D	50-76-0	Heterocyclic Compound, Polycyclic Compound	Laboratory Chemical, Pharmaceutical, Veterinary Agent	PN
3	Apigenin	520-36-5	Heterocyclic Compound	Dye, Natural Product, Pharmaceutical Intermediate	POS
3	Clomiphene citrate	50-41-9	Amine, Carboxylic Acid, Heterocyclic Compound	Pharmaceutical	POS
3	Coumestrol	479-13-0	Heterocyclic Compound	Natural Product	POS
3	Daidzein	486-66-8	Flavonoid, Heterocyclic Compound	Natural Product	POS

Study Phase	Substance	CASRN	MeSH Chemical Class ^a	Product Class ^b	ICCVAM Consensus Classification ^c
3	Dexamethasone	50-02-2	Steroid	Pharmaceutical, Veterinary Agent	PP
3	Di- <i>n</i> -butyl phthalate	84-74-2	Ester, Phthalic Acid	Cosmetic Ingredient, Industrial Chemical, Plasticizer	PP
3	Dibenzo[<i>a,h</i>]anthracene	53-70-3	Polycyclic Compound	Laboratory Chemical, Natural Product	PP
3	Dicofol	115-32-2	Hydrocarbon (Cyclic), Hydrocarbon (Halogenated)	Pesticide	POS
3	Diethylhexyl phthalate	117-81-7	Phthalic Acid	Pesticide Intermediate, Plasticizer	PP
3	Estrone	53-16-7	Steroid	Pharmaceutical, Veterinary Agent	POS
3	Ethyl paraben	120-47-8	Carboxylic Acid, Phenol	Pharmaceutical, Preservative	POS
3	Fluoranthene	206-44-0	Polycyclic Compound	Industrial Chemical, Laboratory Chemical, Pharmaceutical Intermediate	PN
3	Hydroxyflutamide	52806-53-8	Amide	Pharmaceutical	NEG
3	Kaempferol	520-18-3	Flavonoid, Heterocyclic Compound	Natural Product	POS
3	Kepone	143-50-0	Hydrocarbon (Halogenated)	Pesticide	POS
3	<i>meso</i> -Hexestrol	84-16-2	Steroid	Pharmaceutical, Veterinary Agent	POS
3	Methyl testosterone	58-18-4	Steroid	Pharmaceutical, Veterinary Agent	POS

Study Phase	Substance	CASRN	MeSH Chemical Class ^a	Product Class ^b	ICCVAM Consensus Classification ^c
3	Morin	480-16-0	Flavonoid, Heterocyclic Compound	Dye, Natural Product, Pharmaceutical Intermediate	PP
3	Norethynodrel	68-23-5	Steroid	Pharmaceutical	POS
3	<i>p,p'</i> -Methoxychlor	72-43-5	Hydrocarbon (Halogenated)	Pesticide, Veterinary Agent	POS
3	<i>p,p'</i> -DDE	72-55-9	Hydrocarbon (Halogenated)	Pesticide Intermediate	POS
3	Phenobarbital	50-06-6	Heterocyclic Compound, Pyrimidine	Pharmaceutical, Veterinary Agent	NEG
3	Phenolphthalin	81-90-3	Carboxylic Acid, Phenol	Dye, Laboratory Chemical	PN
3	Progesterone	57-83-0	Steroid	Pharmaceutical, Veterinary Agent	PP
3	Propylthiouracil	51-52-5	Heterocyclic Compound, Pyrimidine	Pharmaceutical, Veterinary Agent	PN
3	Raloxifene HCl	82640-04-8	Hydrocarbon (Cyclic)	Pharmaceutical	PP
3	Resveratrol	501-36-0	Hydrocarbon (Cyclic)	Natural Product	POS
3	Sodium azide	26628-22-8	Azide, Salt (Inorganic)	Chemical Intermediate, Fungicide, Herbicide	PN
3	Tamoxifen	10540-29-1	Hydrocarbon (Cyclic)	Pharmaceutical	POS
3	Testosterone	58-22-0	Steroid	Pharmaceutical, Veterinary Agent	PP
4	17 β -Trenbolone	10161-33-8	Steroid	Pharmaceutical	PP
4	19-Nortestosterone	434-22-0	Steroid	Pharmaceutical, Veterinary Agent	POS
4	4-Hydroxyandrostenedione	566-48-3	Steroid	Pharmaceutical	PP

Study Phase	Substance	CASRN	MeSH Chemical Class ^a	Product Class ^b	ICCVAM Consensus Classification ^c
4	Ammonium perchlorate	7790-98-9	Amine, Onium Compound	Industrial Chemical, Laboratory Chemical, Pharmaceutical	PN
4	Apomorphine	58-00-4	Heterocyclic Compound	Pharmaceutical, Veterinary Agent	PN
4	Bicalutamide	90357-06-5	Amide	Pharmaceutical	NEG
4	Chrysin	480-40-0	Flavonoid, Heterocyclic Compound	Natural Product	POS
4	Cycloheximide	66-81-9	Heterocyclic Compound	Fungicide, Pharmaceutical, Veterinary Agent	PN
4	Cyproterone acetate	427-51-0	Steroid	Pharmaceutical	PP
4	Fenarimol	60168-88-9	Heterocyclic Compound, Pyrimidine	Fungicide	POS
4	Finasteride	98319-26-7	Steroid	Pharmaceutical	PN
4	Fluoxymestron	76-43-7	Steroid	Pharmaceutical	PN
4	Flutamide	13311-84-7	Amide	Pharmaceutical, Veterinary Agent	NEG
4	Haloperidol	52-86-8	Ketone	Pharmaceutical, Veterinary Agent	PN
4	Ketoconazole	65277-42-1	Heterocyclic Compound	Pharmaceutical	PN
4	L-Thyroxine	51-48-9	Amino Acid	Pharmaceutical, Veterinary Agent	POS
4	Linuron	330-55-2	Urea	Herbicide	NEG
4	Medroxyprogesterone acetate	71-58-9	Steroid	Pharmaceutical	PP
4	Mifepristone	84371-65-3	Steroid	Pharmaceutical	PP
4	Nilutamide	63612-50-0	Heterocyclic Compound, Imidazole	Pharmaceutical	PN

Study Phase	Substance	CASRN	MeSH Chemical Class ^a	Product Class ^b	ICCVAM Consensus Classification ^c
4	Oxazepam	604-75-1	Heterocyclic Compound	Pharmaceutical, Veterinary Agent	PN
4	Pimozide	2062-78-4	Heterocyclic Compound	Pharmaceutical	PN
4	Procymidone	32809-16-8	Polycyclic Compound	Fungicide	NEG
4	Reserpine	50-55-5	Heterocyclic Compound, Indole	Pharmaceutical, Veterinary Agent	PN
4	Spironolactone	52-01-7	Lactone, Steroid	Pharmaceutical	NEG

Abbreviations: CASRN = CAS Registry Number (American Chemical Society); ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; MeSH= Medical Subject Headings (National Library of Medicine); NEG = negative; PN = presumed negative; POS = positive; PP = presumed positive.

^a Substances 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>).

^b Substances 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>).

^c Estrogenic activity based on a literature review for effects on ER binding, ER TA based on CER1, and uterotrophic response.

Table 3-6 Antagonist Substances by Study Phase

Study Phase	Substance	CASRN	MeSH Chemical Class ^a	Product Class ^b	ICCVAM Consensus Classification ^c
1	17 β -Estradiol	50-28-2	Steroid	Pharmaceutical, Veterinary Agent	PN
1	Flavone	525-82-6	Flavonoid, Heterocyclic Compound	Natural Product, Pharmaceutical	PN
1	Raloxifene HCl	82640-04-8	Hydrocarbon (Cyclic)	Pharmaceutical	POS
2a	Dibenzo[<i>a,h</i>]anthracene	53-70-3	Polycyclic Compound	Laboratory Chemical, Natural Product	PP
2a	<i>p</i> -n-Nonylphenol	104-40-5	Phenol	Chemical Intermediate	NEG
2a	Progesterone	57-83-0	Steroid	Pharmaceutical, Veterinary Agent	NEG

Study Phase	Substance	CASRN	MeSH Chemical Class ^a	Product Class ^b	ICCVAM Consensus Classification ^c
2a	Tamoxifen	10540-29-1	Hydrocarbon (Cyclic)	Pharmaceutical	POS
2b	Apigenin	520-36-5	Heterocyclic Compound	Dye, Natural Product, Pharmaceutical Intermediate	NEG
2b	Atrazine	1912-24-9	Heterocyclic Compound	Herbicide	PN
2b	Butylbenzyl phthalate	85-68-7	Carboxylic Acid, Phthalic Acid	Pharmaceutical, Veterinary Agent	NEG
2b	Corticosterone	50-22-6	Steroid	Pharmaceutical	PN
2b	Flavone	525-82-6	Flavonoid, Heterocyclic Compound	Natural Product, Pharmaceutical	PN
2b	Genistein	446-72-0	Flavonoid, Heterocyclic Compound	Natural Product, Pharmaceutical	NEG
2b	<i>o,p'</i> -DDT	789-02-6	Hydrocarbon (Halogenated)	Pesticide	NEG
2b	Resveratrol	501-36-0	Hydrocarbon (Cyclic)	Natural Product	NEG
3	12- <i>O</i> -Tetradecanoylphorbol-13-acetate	16561-29-8	Hydrocarbon (Cyclic)	Laboratory Chemical	PN
3	17 α -Estradiol	57-91-0	Steroid	Pharmaceutical, Veterinary Agent	PN
3	17 α -Ethinyl estradiol	57-63-6	Steroid	Pharmaceutical, Veterinary Agent	NEG
3	17 β -Estradiol	50-28-2	Steroid	Pharmaceutical, Veterinary Agent	PN
3	2- <i>sec</i> -Butylphenol	89-72-5	Phenol	Chemical Intermediate, Pesticide Intermediate, Plasticizer	PN
3	2,4,5-Trichlorophenoxyacetic acid	93-76-5	Carboxylic Acid	Herbicide	PN

Study Phase	Substance	CASRN	MeSH Chemical Class ^a	Product Class ^b	ICCVAM Consensus Classification ^c
3	4-Androstenedione	63-05-8	Steroid	Pharmaceutical	PN
3	4-Cumylphenol	599-64-4	Phenol	Chemical Intermediate	PN
3	4-Hydroxytamoxifen	68047-06-3	Hydrocarbon (Cyclic)	Pharmaceutical	POS
3	4- <i>tert</i> -Octylphenol	140-66-9	Phenol	Chemical Intermediate, Pharmaceutical Intermediate	PN
3	5 α -Dihydrotestosterone	521-18-6	Steroid	Pharmaceutical	NEG
3	Actinomycin D	50-76-0	Heterocyclic Compound, Polycyclic Compound	Laboratory Chemical, Pharmaceutical, Veterinary Agent	PN
3	Bisphenol A	80-05-7	Phenol	Chemical Intermediate, Flame Retardant, Fungicide	NEG
3	Bisphenol B	77-40-7	Phenol	Chemical Intermediate, Flame Retardant, Fungicide	PN
3	Clomiphene citrate	50-41-9	Amine, Carboxylic Acid, Heterocyclic Compound	Pharmaceutical	PP
3	Coumestrol	479-13-0	Heterocyclic Compound	Natural Product	NEG
3	Daidzein	486-66-8	Flavonoid, Heterocyclic Compound	Natural Product	NEG
3	Dexamethasone	50-02-2	Steroid	Pharmaceutical, Veterinary Agent	PN
3	Di- <i>n</i> -butyl phthalate	84-74-2	Ester, Phthalic Acid	Cosmetic Ingredient, Industrial Chemical, Plasticizer	NEG

Study Phase	Substance	CASRN	MeSH Chemical Class ^a	Product Class ^b	ICCVAM Consensus Classification ^c
3	Dicofol	115-32-2	Hydrocarbon (Cyclic), Hydrocarbon (Halogenated)	Pesticide	NEG
3	Diethylhexyl phthalate	117-81-7	Phthalic Acid	Pesticide Intermediate, Plasticizer	NEG
3	Diethylstilbestrol	56-53-1	Hydrocarbon (Cyclic)	Pharmaceutical, Veterinary Agent	NEG
3	Estrone	53-16-7	Steroid	Pharmaceutical, Veterinary Agent	PN
3	Ethyl paraben	120-47-8	Carboxylic Acid, Phenol	Pharmaceutical, Preservative	PN
3	Fluoranthene	206-44-0	Polycyclic Compound	Industrial Chemical, Laboratory Chemical, Pharmaceutical Intermediate	PN
3	Hydroxyflutamide	52806-53-8	Amide	Pharmaceutical	PN
3	Kaempferol	520-18-3	Flavonoid, Heterocyclic Compound	Natural Product	NEG
3	Kepone	143-50-0	Hydrocarbon (Halogenated)	Pesticide	NEG
3	<i>meso</i> -Hexestrol	84-16-2	Steroid	Pharmaceutical, Veterinary Agent	PN
3	Methyl testosterone	58-18-4	Steroid	Pharmaceutical, Veterinary Agent	PN
3	Morin	480-16-0	Flavonoid, Heterocyclic Compound	Dye, Natural Product, Pharmaceutical Intermediate	PN
3	Norethynodrel	68-23-5	Steroid	Pharmaceutical	NEG
3	<i>p,p'</i> -DDE	72-55-9	Hydrocarbon (Halogenated)	Pesticide Intermediate	NEG
3	<i>p,p'</i> -Methoxychlor	72-43-5	Hydrocarbon (Halogenated)	Pesticide, Veterinary Agent	PP

Study Phase	Substance	CASRN	MeSH Chemical Class ^a	Product Class ^b	ICCVAM Consensus Classification ^c
3	Phenobarbital	50-06-6	Heterocyclic Compound, Pyrimidine	Pharmaceutical, Veterinary Agent	PN
3	Phenolphthalin	81-90-3	Carboxylic Acid, Phenol	Dye, Laboratory Chemical	PN
3	Propylthiouracil	51-52-5	Heterocyclic Compound, Pyrimidine	Pharmaceutical, Veterinary Agent	PN
3	Raloxifene HCl	82640-04-8	Hydrocarbon (Cyclic)	Pharmaceutical	POS
3	Sodium azide	26628-22-8	Azide, Salt (Inorganic)	Chemical Intermediate, Fungicide, Herbicide	PN
3	Testosterone	58-22-0	Steroid	Pharmaceutical, Veterinary Agent	PN
3	Vinclozolin	50471-44-8	Heterocyclic Compound	Fungicide	PN
4	17β-Trenbolone	10161-33-8	Steroid	Pharmaceutical	PN
4	19-Nortestosterone	434-22-0	Steroid	Pharmaceutical, Veterinary Agent	PN
4	4-Hydroxyandrostenedione	566-48-3	Steroid	Pharmaceutical	PN
4	Ammonium perchlorate	7790-98-9	Amine, Onium Compound	Industrial Chemical, Laboratory Chemical, Pharmaceutical	PN
4	Apomorphine	58-00-4	Heterocyclic Compound	Pharmaceutical, Veterinary Agent	PN
4	Bicalutamide	90357-06-5	Amide	Pharmaceutical	PN
4	Chrysin	480-40-0	Flavonoid, Heterocyclic Compound	Natural Product	NEG
4	Cycloheximide	66-81-9	Heterocyclic Compound	Fungicide, Pharmaceutical, Veterinary Agent	PP

Study Phase	Substance	CASRN	MeSH Chemical Class^a	Product Class^b	ICCVAM Consensus Classification^c
4	Cyproterone acetate	427-51-0	Steroid	Pharmaceutical	PN
4	Fenarimol	60168-88-9	Heterocyclic Compound, Pyrimidine	Fungicide	PN
4	Finasteride	98319-26-7	Steroid	Pharmaceutical	PN
4	Fluoxymestrone	76-43-7	Steroid	Pharmaceutical	PN
4	Flutamide	13311-84-7	Amide	Pharmaceutical, Veterinary Agent	PN
4	Haloperidol	52-86-8	Ketone	Pharmaceutical, Veterinary Agent	PN
4	Ketoconazole	65277-42-1	Heterocyclic Compound	Pharmaceutical	PN
4	L-Thyroxine	51-48-9	Amino Acid	Pharmaceutical, Veterinary Agent	PN
4	Linuron	330-55-2	Urea	Herbicide	PN
4	Medroxyprogesterone acetate	71-58-9	Steroid	Pharmaceutical	PN
4	Mifepristone	84371-65-3	Steroid	Pharmaceutical	NEG
4	Nilutamide	63612-50-0	Heterocyclic Compound, Imidazole	Pharmaceutical	PN
4	Oxazepam	604-75-1	Heterocyclic Compound	Pharmaceutical, Veterinary Agent	PN
4	Pimozide	2062-78-4	Heterocyclic Compound	Pharmaceutical	PN
4	Procymidone	32809-16-8	Polycyclic Compound	Fungicide	PN
4	Reserpine	50-55-5	Heterocyclic Compound, Indole	Pharmaceutical, Veterinary Agent	PN
4	Spirolactone	52-01-7	Lactone, Steroid	Pharmaceutical	PN

Abbreviations: CASRN = CAS Registry Number (American Chemical Society); Interagency Coordinating Committee on the Validation of Alternative Methods; MeSH = Medical Subject Headings (National Library of Medicine); NEG = negative; PN = presumed negative; POS = positive; PP = presumed positive.

^a Substances 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>).

^b Substances 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>).

^c Estrogenic activity based on a literature review for effects on ER binding, ER TA based on CERI, and uterotrophic response.

3.5 Substances Used to Assess Intralaboratory Reproducibility

Intralaboratory reproducibility was assessed using data generated by testing the 12 coded reference substances in Phase 2. Each of the 12 was tested three times on three separate days. The substances tested in Phase 2 that were used to assess intralaboratory reproducibility of the agonist test methods are listed in **Table 3-5**. Those that were used to assess intralaboratory reproducibility of antagonist test methods are listed in **Table 3-6**.

3.6 Substances Used to Assess Interlaboratory Reproducibility

Because this validation study was conducted in four phases, not all substances were tested in all laboratories. Consequently, only those coded substances tested in all three laboratories (Phase 2 and Phase 3) could be used to assess interlaboratory reproducibility. The 53 substances tested in Phase 2 and Phase 3 that were used to assess interlaboratory reproducibility of the agonist and antagonist test methods are listed in **Table 3-5** and **Table 3-6**, respectively.

3.7 Chemical Classes Represented in the List of Substances

The chemical classes shown for each of the 78 reference substances were assigned by the U.S. National Library of Medicine's Medical Subject Headings (MeSH[®]; available at <http://www.nlm.nih.gov/mesh>), an internationally recognized standardized classification scheme. The distribution of substances by chemical class is provided in **Table 3-7**.

Table 3-7 Distribution of Reference Substances by Chemical Class

MeSH Chemical Class ^a	All Substances	Substances Used for Agonist Accuracy	Substances Used for Antagonist Accuracy
Amides	3	3	1
Amines	2	1	0
Amino Acids	1	1	0
Azides	1	0	0
Carboxylic Acids	5	4	1
Esters	2	0	0
Flavonoids	8	7	1
Heterocyclic Compounds	22	12	3
Hydrocarbons (Cyclic)	7	4	2
Hydrocarbons (Halogenated)	5	5	3
Imidazoles	1	0	0
Indoles	1	0	0
Ketones	1	1	0
Lactones	1	1	0

MeSH Chemical Class ^a	All Substances	Substances Used for Agonist Accuracy	Substances Used for Antagonist Accuracy
Onium Compounds	1	0	0
Phenols	8	8	1
Phthalic Acids	3	1	1
Polycyclic Compounds	4	1	0
Pyrimindines	3	2	1
Salts (Inorganic)	1	1	0
Steroids	22	12	5
Ureas	1	1	0

Abbreviations: MeSH = Medical Subject Headings (National Library of Medicine).

^a Substances 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>).

3.8 Product Classes Represented in the List of Substances

The product classes assigned to each reference substance are based on information obtained from the U.S. National Library of Medicine's Hazardous Substances Data Bank (available at <http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB>). For *in vitro* ER test methods, the distribution of substances by product class is provided in **Table 3-8**.

Table 3-8 Distribution of Reference Substances by Product Class

Product Class ^a	All Substances	Substances Used for Agonist Accuracy	Substances Used for Antagonist Accuracy
Chemical Intermediate	7	7	1
Cosmetic Ingredient	1	0	0
Dye	3	3	0
Flame Retardant	2	2	1
Fungicide	7	5	2
Herbicide	4	3	1
Industrial Chemical	4	0	0
Laboratory Chemical	6	1	0
Natural Product	10	8	1
Pesticide	4	4	2
Pesticide Intermediate	3	2	1
Pharmaceutical	46	25	10
Pharmaceutical Intermediate	4	3	0

Product Class ^a	All Substances	Substances Used for Agonist Accuracy	Substances Used for Antagonist Accuracy
Plasticizer	4	1	0
Preservative	1	1	0
Veterinary Agent	22	13	3

^a Substances 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>).

3.9 Test Substance Procurement, Coding, and Distribution

On behalf of NICEATM, the National Toxicology Program Substances Inventory (NTPSI) procured and distributed all reference standards and controls to the participating laboratories, with the exception of some that were classified as controlled substances (i.e., 4-androstenedione, 5 α -dihydrotestosterone, methyl testosterone, testosterone, and phenobarbital). To avoid the extensive amount of documentation required (and associated delays) to import controlled substances, ECVAM and JaCVAM made efforts to procure these specific substances from their regional suppliers. ECVAM procured methyl testosterone and phenobarbital from EU-based suppliers but not 4-androstenedione, 5 α -dihydrotestosterone, or testosterone. Therefore, ECVAM obtained the required EU regulatory permissions for the importation of 4-androstenedione, 5 α -dihydrotestosterone, and testosterone, which were subsequently procured by the NTPSI and exported to the ECVAM laboratory accordingly. JaCVAM was able to procure 4-androstenedione, 5 α -dihydrotestosterone, methyl testosterone, and testosterone from Japanese suppliers. However, phenobarbital, classified as a Schedule IV controlled substance according to the U.S. Drug Enforcement Administration, was not procured because the JaCVAM-sponsored Hiyoshi Laboratory did not have an appropriate license for handling Schedule IV substances.

Reference substances were coded with unique laboratory-specific identifiers (see **Annex I** for laboratory-specific reference substance codes), and aliquots were sent in coded vials to participating laboratories. (**Note:** The NTPSI also provided empty coded vials to ECVAM and JaCVAM for the controlled substances that were procured from regional distributors as detailed above.) Material Safety Data Sheets (MSDSs) were provided along with the reference substances and controls. Coded reference substances were provided with a sealed health and safety packet containing the identity of each test substance, as well as its MSDS, to be opened in the event of an accident (e.g., chemical spill). The NTPSI, ECVAM, and/or JaCVAM also obtained Certificates of Analysis for reference standards, controls, and reference substances.

Procedures for shipping substances to the participating laboratories were the same regardless of whether NTPSI, ECVAM, or JaCVAM was the responsible party. Substances were packaged so as to minimize damage during transit and shipped under appropriate storage conditions and according to the appropriate regulatory transportation procedures. The NICEATM validation study project manager maintained Certificates of Analysis for all test substances. The participating laboratories were notified upon shipment in order to prepare for receipt. Test substance shipments were delivered to each participating laboratory. Information regarding weight or volume and storage conditions for each coded reference substance was also provided to each laboratory well before shipment. The shipment included the following instructions for the participating laboratories:

- Contact the NTPSI and the NICEATM validation study project manager upon receipt of test substances.
- Contact the validation study project manager if test facility personnel opened the health and safety packet at any time, for any reason, during the study.

4.0 Test Method Data And Results

This section summarizes the results from testing of 53 coded reference substances in the three participating laboratories (XDS, ECVAM, and Hiyoshi) and an additional 25 coded reference substances tested in the lead laboratory (XDS) using the agonist and antagonist protocols for the BG1Luc ER TA test method.

4.1 Availability of Original Data Used to Evaluate Test Method Performance

All data were provided to the validation study project coordinator at NICEATM as electronic Microsoft Excel[®] and GraphPad Prism[®] files. Data files and laboratory reports are available upon request from NICEATM. Requests can be made by mail, fax, or e-mail to Dr. William S. Stokes, NICEATM, NIEHS, P.O. Box 12233, MD K2-16, Research Triangle Park, NC, 27709, (phone) 919-541-2384, (fax) 919-541-0947, (e-mail) niceatm@niehs.nih.gov.

4.2 BG1Luc ER TA Agonist and Antagonist Reference Standard and Control Data

During Phase 1, each laboratory established a historical database for the control and reference substances. The database was used to calculate acceptance criteria using reference standards and controls for use in subsequent study phases. Although E2 reference standard EC₅₀, Ral reference standard IC₅₀, Met RLU, and flavone RLU values were not used for plate acceptance after Phase 2a of the validation study (see **Sections 2.7.1** and **2.7.2**), these values were collected throughout the study for information purposes (see **Tables 4-1** through **4-7**). Because the RLU values for the agonist and antagonist DMSO control and the antagonist E2 control were used for acceptance criteria throughout the study, they were used in the evaluation of intra- and interlaboratory reproducibility (see **Section 6**). The reported data represent only plates that passed test plate acceptance criteria. The total number of plates that were run (combination of number of acceptable plates and plates that failed one or more acceptance criteria) are also reported. Details of the rationale for any plate failures, along with their impact on intralaboratory reproducibility, are discussed in **Section 6**.

4.2.1 Agonist E2 Reference Standard

As shown in **Table 4-1**, the historical E2 EC₅₀ data collected by each laboratory in Phase 1 ranged from 8.47×10^{-12} to 1.13×10^{-11} M on the 10 acceptable plates required to generate the historical database at XDS and Hiyoshi. XDS successfully generated their historical database in 10 consecutive experiments. ECVAM generated data on 18 consecutive experiments due to a concern that a portion of the plates might not meet the acceptance criteria. However, none of these 18 plates failed acceptance; therefore, the ECVAM historical database is based on a total of 18 plates. Hiyoshi required two additional experiments because two plates failed the fold induction acceptance criterion. E2 EC₅₀ values collected by each laboratory in subsequent phases of the validation study ranged from 6.15×10^{-12} to 1.74×10^{-11} M (see **Table 4-1**).

Table 4-1 Summary of Agonist E2 Reference Standard EC₅₀ Data by Study Phase

Laboratory	Study Phase	Mean EC ₅₀ (M) ^a	SD	N
XDS	1	8.47×10^{-12}	1.66×10^{-12}	10/10
ECVAM	1	8.34×10^{-12}	3.10×10^{-12}	18/18
Hiyoshi	1	1.13×10^{-11}	2.91×10^{-12}	10/12
XDS	2a	9.95×10^{-12}	1.53×10^{-12}	7/15
ECVAM	2a	1.16×10^{-11}	4.07×10^{-12}	6/30
Hiyoshi	2a	8.54×10^{-12}	1.73×10^{-12}	8/9
XDS	2b	9.97×10^{-12}	2.88×10^{-12}	13/13
ECVAM	2b	7.82×10^{-12}	4.80×10^{-12}	12/16
Hiyoshi	2b	1.02×10^{-11}	1.94×10^{-12}	13/16
XDS	3	1.36×10^{-11}	1.28×10^{-11}	34/47
ECVAM	3	1.48×10^{-11}	3.02×10^{-11}	24/35
Hiyoshi	3	6.15×10^{-12}	1.31×10^{-12}	34/34
XDS	4	1.74×10^{-11}	2.66×10^{-11}	29/41

Abbreviations: EC₅₀ = half-maximal effective concentration; ECVAM = European Centre for the Validation of Alternative Methods; M = molar; N = number of plates that passed acceptance criteria/total number of plates; SD = standard deviation; XDS = Xenobiotic Detection Systems, Inc.

^a This value was used as a test plate acceptance criterion during Phase 2a of the validation study only. After Phase 2a, this value was monitored but was no longer used to determine whether test plates passed acceptance criteria.

4.2.2 Agonist DMSO Control Values

Because DMSO control RLU values are not normalized, they can vary considerably between test plates. DMSO RLU values at all laboratories during all validation study phases ranged from a low of 511 (Phase 3 at XDS) to a high of 9885 (Phase 1 at XDS), with a mean of 3749 (see **Table 4-2**). However, within-plate variability of DMSO control RLU values between replicate DMSO wells was low, with associated coefficient of variation (CV) values ranging from 1% to 43% and a mean of 8% (see **Table 4-2**). Of the 218 agonist test plates that met acceptance criteria, only six plates had within-plate CV values greater than 20%. (See **Annex L** for individual test plate mean DMSO control RLU values and associated CV values.)

Table 4-2 Summary of Agonist Within-Plate DMSO Control Data by Laboratory and Study Phase

Laboratory	Study Phase	Mean and Range of DMSO Control RLU Values	Mean and Range of CV (%)	N
XDS	1	5362 (2031-9885)	7 (5-9)	10/10
ECVAM	1	3519 (1379-6342)	8 (2-14)	18/18
Hiyoshi	1	4213 (2323-6087)	7 (4-15)	10/12
XDS	2a	2271 (636-5114)	10 (3-21)	7/15
ECVAM	2a	2900 (828-5017)	8 (1-17)	6/30
Hiyoshi	2a	4199 (2023-6314)	5 (1-9)	8/9
XDS	2b	2084 (628-4094)	5 (2-10)	13/13
ECVAM	2b	4291 (3256-6209)	6 (3-11)	12/16
Hiyoshi	2b	6291 (4330-8078)	5 (1-10)	13/16
XDS	3	2314 (511-6826)	10 (1-43)	34/47
ECVAM	3	2938 (1097-7306)	10 (3-33)	24/35
Hiyoshi	3	5760 (1362-9383)	6 (1-24)	34/34
XDS	4	2943 (913-5987)	8 (1-17)	29/41
All	All	3749 (511-9885)	8 (1-43)	218/286

Abbreviations: CV = coefficient of variation; DMSO = dimethyl sulfoxide; ECVAM = European Centre for the Validation of Alternative Methods; N = number of plates that passed acceptance criteria/total number of plates; XDS = Xenobiotic Detection Systems, Inc.

4.2.3 Maximum Fold Induction of E2 Response During Agonist Testing

As shown in **Table 4-3**, mean fold induction across the three laboratories throughout the validation study was 5.72 ± 1.82 . With the exception of Phase 2b, ECVAM consistently reported the highest mean fold induction. ECVAM's highest mean value (9.2) was observed during Phase 3. Hiyoshi reported the lowest values in all study phases except Phase 3. During Phase 3, XDS and Hiyoshi reported similar values (4.3 and 4.9, respectively). The lowest mean fold induction reported during the validation study was 4.0, which was observed at both Hiyoshi (Phase 2b) and XDS (Phase 4).

Table 4-3 Summary of Agonist Maximum Fold Induction Data by Laboratory and Study Phase

Laboratory	Study Phase	Mean Fold Induction ^{a,b}	SD	N
XDS	1	4.7	0.7	10/10
ECVAM	1	8.1	0.9	18/18
Hiyoshi	1	4.5	0.9	10/12
XDS	2a	6.4	2.7	7/15
ECVAM	2a	8.0	1.9	6/30
Hiyoshi	2a	4.4	0.7	8/9
XDS	2b	7.3	2.0	13/13
ECVAM	2b	4.6	0.9	12/16
Hiyoshi	2b	4.0	0.7	13/16
XDS	3	4.3	1.0	34/47
ECVAM	3	9.2	3.0	24/35
Hiyoshi	3	4.9	1.0	34/34
XDS	4	4.0	1.3	29/41
All	All	5.72	1.82	13/13

Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; N = number of plates that passed acceptance criteria/total number of plates; SD = standard deviation; XDS = Xenobiotic Detection Systems, Inc.

^a Fold induction is measured by dividing the test plate averaged highest E2 reference standard RLU value by the averaged DMSO control mean RLU value (see **Section 2.7.1**).

^b Test plate acceptance criteria for maximum fold induction state that fold induction must be greater than 3.

4.2.4 Weak Agonist Positive Control: Methoxychlor

During the development of the historical Met control databases, the normalized and adjusted response was highest at Hiyoshi and lowest at ECVAM (**Table 4-4**). Variability was low in all three laboratories ($CV \leq 17\%$). Variability remained low throughout subsequent phases of the validation study ($CV \leq 23\%$).

Table 4-4 Summary of Agonist Methoxychlor Control Data by Laboratory and Study Phase

Laboratory	Study Phase	Mean Adjusted RLU Value ^{a,b}	SD	N
XDS	1	5709	974	10/10
ECVAM	1	4494	590	18/18
Hiyoshi	1	7917	430	10/12
XDS	2a	5494	981	7/15
ECVAM	2a	5199	508	6/30
Hiyoshi	2a	8500	424	8/9
XDS	2b	6126	941	13/13
ECVAM	2b	8117	789	12/16
Hiyoshi	2b	7861	854	13/16
XDS	3	6420	1475	35/47
ECVAM	3	6885	1043	24/35
Hiyoshi	3	8029	1579	34/34
XDS	4	5902	1275	29/41

Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; N = number of plates that passed acceptance criteria/total number of plates; RLU = relative light unit; SD = standard deviation; XDS = Xenobiotic Detection Systems, Inc.

^a Agonist test plate data are adjusted by subtracting the mean DMSO control RLU value from the RLU value for each agonist test plate well. The data are then normalized by setting the maximum E2 response to 10,000 RLU and adjusting all other RLU values relative to the maximum E2 response.

^b This value was used as a test plate acceptance criterion during Phase 2a of the validation study only. After Phase 2a, test plate acceptance criteria were modified to state that this value must be greater than the DMSO mean plus three times the standard deviation from that mean.

4.2.5 Antagonist Raloxifene Reference Standard

As shown in **Table 4-5**, the historical Ral IC₅₀ values obtained by each laboratory ranged from 8.43×10^{-10} to 1.23×10^{-9} M. As in the agonist testing, the laboratories were instructed to generate historical reference standard and control databases based on data generated from at least 10 acceptable test plates. All three laboratories generated data on more than 10 acceptable test plates due to concerns that a portion of the plates might not pass the acceptance criterion (i.e., fold induction ≥ 3) which required a >3-fold reduction in E2 control values. The historical databases at XDS, ECVAM, and Hiyoshi were based on 14, 18, and 12 plates, respectively. None of the runs at ECVAM or Hiyoshi failed the acceptance criterion, and XDS had a single plate failure. The calculated CV of the Ral IC₅₀ values was within 33% for all laboratories, with the exception of XDS during Phase 3, when a CV value of 60% was observed.

Table 4-5 Summary of Antagonist Raloxifene Reference Standard IC₅₀ Data by Laboratory and Study Phase

Laboratory	Study Phase	Mean IC ₅₀ (M) ^a	SD	N
XDS	1	8.35×10^{-10}	1.76×10^{-10}	14/15
ECVAM	1	8.43×10^{-10}	1.54×10^{-10}	18/18
Hiyoshi	1	1.23×10^{-9}	2.53×10^{-10}	12/12
XDS	2a	7.43×10^{-10}	2.44×10^{-10}	8/14
ECVAM	2a	8.39×10^{-10}	1.56×10^{-10}	7/14
Hiyoshi	2a	1.23×10^{-9}	3.31×10^{-10}	6/6
XDS	2b	1.06×10^{-9}	1.88×10^{-10}	12/12
ECVAM	2b	1.15×10^{-9}	2.32×10^{-10}	12/18
Hiyoshi	2b	1.48×10^{-9}	1.95×10^{-10}	14/14
XDS	3	1.25×10^{-9}	7.49×10^{-10}	30/59
ECVAM	3	1.84×10^{-9}	4.67×10^{-10}	25/36
Hiyoshi	3	9.94×10^{-10}	1.76×10^{-10}	21/24
XDS	4	5.76×10^{-10}	1.19×10^{-10}	15/23

Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; IC₅₀ = half-maximal inhibitory concentration; M = molar; N = number of plates that passed acceptance criteria/total number of plates; SD = standard deviation; XDS = Xenobiotic Detection Systems, Inc.

^a This value was used as a test plate acceptance criterion during Phase 2a of the validation study only. After Phase 2a, this value was monitored but was no longer used to determine whether test plates met acceptance criteria.

4.2.6 Antagonist DMSO Control Values

Because DMSO control RLU values are not normalized, they can vary considerably between test plates; therefore, mean plate DMSO RLU values ranged from a low of 132 at XDS during Phase 1 to a high of 8451 at Hiyoshi during Phase 3, with a mean of 3299 for plates that passed acceptance criteria at all laboratories (see **Table 4-6**). However, within-plate variability of DMSO RLU control values between replicate DMSO wells was low, with associated CV values ranging from 1% to 52% and a mean of 8% (see **Table 4-6**). Of the 194 antagonist test plates that passed acceptance criteria, only eight plates had within-plate CV values greater than 20%. (See **Annex L** for individual test plate mean DMSO control RLU values and associated CV values.)

Table 4-6 Summary of Antagonist Within-Plate DMSO Control Data by Study Phase

Laboratory	Study Phase	Mean and Range of DMSO Control RLU Values	Mean and Range of CV (%)	N
XDS	1	499 (132-1331)	9 (3-18)	14/15
ECVAM	1	3783 (1490-7333)	8 (3-17)	18/18
Hiyoshi	1	4048 (1625-6541)	5 (3-9)	12/12
XDS	2a	1378 (271-2073)	10 (2-14)	8/14
ECVAM	2a	2154 (1352-5102)	11 (1-23)	7/14
Hiyoshi	2a	4915 (2846-7221)	5 (1-12)	6/6
XDS	2b	1910 (930-2773)	4 (2-9)	12/12
ECVAM	2b	4128 (2522-5102)	7 (1-18)	12/18
Hiyoshi	2b	6280 (4633-7992)	7 (1-20)	14/14
XDS	3	2746 (415-6860)	8 (2-52)	30/59
ECVAM	3	3852 (2615-5498)	12 (4-37)	25/36
Hiyoshi	3	4030 (2018-8451)	7 (1-20)	21/24
XDS	4	3742 (2498-6482)	8 (1-15)	15/23
All	All	3299 (132-8451)	8 (1-52)	194/251

Abbreviations: CV = coefficient of variation; DMSO = dimethyl sulfoxide; ECVAM = European Centre for the Validation of Alternative Methods; N = number of plates that passed acceptance criteria/total number of plates; RLU = relative light unit; XDS = Xenobiotic Detection Systems, Inc.

4.2.7 Antagonist E2 Reference Standard

Using the historical data developed by each laboratory during Phase 1, XDS and ECVAM reported similar normalized E2 responses (8284–8881 mean adjusted RLU), while Hiyoshi was considerably lower (5728 mean adjusted RLU) (**Table 4-7**). With the exception of Phase 1 testing at Hiyoshi (CV = 21%), the calculated CV was no more than 14% at any of the laboratories throughout the study.

Table 4-7 Summary of Antagonist E2 Control Data by Study Phase

Laboratory	Study Phase	Mean Adjusted RLU ^{a,b}	SD	N
XDS	1	8284	744	14/15
ECVAM	1	8881	640	18/18
Hiyoshi	1	5728	1221	12/12
XDS	2a	8646	783	8/14
ECVAM	2a	9106	554	7/14
Hiyoshi	2a	5767	347	6/6
XDS	2b	8259	711	12/12
ECVAM	2b	9175	725	12/18
Hiyoshi	2b	5270	478	14/14
XDS	3	7851	1065	30/49
ECVAM	3	9584	901	25/36
Hiyoshi	3	6185	521	21/24
XDS	4	7428	662	15/23

Abbreviations: E2 = 17 β -estradiol; ECVAM = European Centre for the Validation of Alternative Methods; N = number of plates that passed acceptance criteria/total number of plates; RLU = relative light unit; SD = standard deviation; XDS = Xenobiotic Detection Systems, Inc.

^a Antagonist test plate data are adjusted by subtracting the DMSO control RLU values from the RLU value for each antagonist test plate well. The data are then normalized by setting the maximum Ral response to 10,000 RLU and adjusting all other RLU values relative to the maximum Ral response.

^b The mean E2 control RLU value must be within the mean plus or minus 2.5 times the SD of the historical mean RLU value for the E2 control.

4.2.8 Maximum Fold Reduction of E2 Response During Antagonist Testing

As shown in **Table 4-8**, mean fold reduction of E2 response across the three laboratories throughout the validation study was 9.56 ± 2.47 . Both the highest (14.2 in Phase 1) and lowest (6.5 in Phase 3) values reported were from XDS. There was no consistency as to which laboratory reported the highest value in each phase.

Table 4-8 Summary of Antagonist Maximum Fold Reduction Data by Laboratory and Study Phase

Laboratory	Study Phase	Mean Fold Reduction ^{a,b}	SD	N
XDS	1	14.2	2.4	14/15
ECVAM	1	8.0	0.7	18/18
Hiyoshi	1	7.9	2.3	12/12
XDS	2a	11.1	2.7	8/14
ECVAM	2a	12.1	1.7	7/14
Hiyoshi	2a	11.4	3.2	6/6
XDS	2b	11.4	2.4	12/12
ECVAM	2b	6.6	0.6	12/18
Hiyoshi	2b	10.9	1.6	14/14
XDS	3	6.5	2.5	30/59
ECVAM	3	7.5	1.2	25/36
Hiyoshi	3	9.8	2.1	21/24
XDS	4	7.0	2.3	15/23
All	All	9.56	2.47	13/13

Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; N = number of plates that passed acceptance criteria/total number of plates; SD = standard deviation; XDS = Xenobiotic Detection Systems, Inc.

^a Reduction for comprehensive test plates is measured by dividing the averaged highest Ral reference standard RLU value by the lowest averaged Ral reference standard RLU value (see **Section 2.7.2**).

^b Test plate acceptance criteria for mean fold reduction state that fold reduction must be greater than 3.

4.2.9 Weak Antagonist Positive Control: Flavone

During the development of the historical flavone control databases, the normalized response was highest at XDS (3583), where the lowest CV (30%) was also observed. The response was lowest at ECVAM (644), where the highest CV (71%) was also observed (**Table 4-9**). Variability was lowest at XDS, but high CVs were seen in all laboratories during Phases 2 through 4 of the study (CVs ranged from 40% to 217%).

Table 4-9 Summary of Antagonist Flavone Control Data by Study Phase

Laboratory	Study Phase	Mean Adjusted RLU Value ^{a,b}	SD	N
XDS	1	3583	1089	14/15
ECVAM	1	644	458	18/18
Hiyoshi	1	1226	723	12/12
XDS	2a	3620	753	8/14
ECVAM	2a	733	521	7/14
Hiyoshi	2a	497	203	6/6
XDS	2b	3164	1272	12/12
ECVAM	2b	801	580	12/18
Hiyoshi	2b	87	188	14/14
XDS	3	3081	1627	30/59
ECVAM	3	431	361	25/36
Hiyoshi	3	1302	697	21/24
XDS	4	1444	870	15/23

Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; N = number of plates that passed acceptance criteria/total number of plates; RLU = relative light unit; SD = standard deviation; XDS = Xenobiotic Detection Systems, Inc.

^a Antagonist test plate data are adjusted by subtracting the DMSO control mean RLU values from the RLU value for each antagonist test plate well. The data are then normalized by setting the maximum Ral response to 10,000 RLU and adjusting all other RLU values relative to the maximum Ral response.

^b This value was used as a test plate acceptance criterion during Phase 2a of the validation study only. After Phase 2a, test plate acceptance criteria were modified to state that this value must be less than the E2 control mean minus three times the standard deviation from that mean (i.e., the flavone control must be positive).

4.3 Solubility Test Results

As indicated in **Section 2.5.1**, starting concentrations for range finder testing during Phases 2a and 2b were established by determining the maximum soluble test substance concentration at log intervals up to the 1 mg/mL (v/v in 1% DMSO/EFM) limit concentration. Following Phase 2b comprehensive testing, differences in ER TA antagonist activity were noted across laboratories for two substances (flavone and genistein). The differences in antagonist activity were attributed to differences in solubility. At XDS and ECVAM, 100 µg/mL was considered the maximum soluble concentration for these two substances and was therefore used as the starting concentration for range finder testing. Both ultimately tested positive for antagonist activity at concentrations above 10 µg/mL.³ In contrast, Hiyoshi considered 10 µg/mL to be the maximum soluble concentration for these two substances, which was then used as the starting concentration for range finder testing. Both substances were

³ ER TA antagonist activity classifications for Phase 2 did not limit the evaluation of concentrations above 10 µM (see **Section 2.4.5**).

negative for antagonist activity and were subsequently retested at Hiyoshi up to 100 µg/mL, at which point both were positive.

To maximize the likelihood of detecting weak agonists and antagonists, protocols were modified to determine test substance solubility in 100% DMSO as the starting concentration for range finder testing. This protocol modification was used for Phases 3 and 4 range finder testing. Recognizing that this could result in range finder testing concentrations of substances that precipitate out when added to EFM, the SMT concluded that there would be enough sufficiently soluble concentrations within the 7-point log serial dilution to effectively determine the starting concentrations for comprehensive testing. However, differences in the maximum starting concentrations in 100% DMSO were still observed across laboratories (see **Tables 4-10** and **4-11**).

Where these differences occurred, comprehensive test results were evaluated to determine if lower starting concentrations were responsible for discordances among the laboratories. This occurred for only three agonist substances: 4-androstenedione, 2-*sec*-butylphenol, and fluoranthene. With a starting concentration of 10 µg/mL, 4-androstenedione was negative at ECVAM. It was positive at Hiyoshi with a starting concentration of 100 µg/mL. With a starting concentration of 100 µg/mL, 2-*sec*-butylphenol was negative at ECVAM. It was positive at Hiyoshi at this concentration and positive at XDS with a starting concentration of 1000 µg/mL. Fluoranthene was negative at ECVAM with a starting concentration of 100 µg/mL but positive at Hiyoshi and XDS with a starting concentration of 1000 µg/mL. (See **Table 4-12** for ER TA agonist testing results.)

Table 4-10 Agonist Range Finder Starting Concentrations in Culture Medium

Chemical Name	Study Phase	FW	XDS Max Concentration Tested		ECVAM Max Concentration Tested		Hiyoshi Max Concentration Tested	
			µg/mL	M	µg/mL	M	µg/mL	M
Bisphenol A	2a	228.3	100	4.38×10^{-4}	1000	4.38×10^{-3}	1000	4.38×10^{-3}
Bisphenol B	2a	242.3	1000	4.13×10^{-3}	100	4.13×10^{-4}	100	4.13×10^{-4}
Corticosterone	2a	346.5	1000	2.89×10^{-3}	1000	2.89×10^{-3}	1000	2.89×10^{-3}
Diethylstilbestrol	2a	268.4	100	3.73×10^{-4}	100	3.73×10^{-4}	10	3.73×10^{-5}
17 α -Ethinyl estradiol	2b	296.4	100	3.37×10^{-4}	100	3.37×10^{-4}	10	3.37×10^{-5}
Atrazine	2b	215.7	100	4.64×10^{-4}	100	4.64×10^{-4}	100	4.64×10^{-4}
Butylbenzyl phthalate	2b	312.4	100	3.20×10^{-4}	10	3.20×10^{-5}	10	3.20×10^{-5}
Flavone	2b	222.2	100	4.50×10^{-4}	100	4.50×10^{-4}	100	4.50×10^{-4}
Genistein	2b	270.2	100	3.70×10^{-4}	100	3.70×10^{-4}	100	3.70×10^{-4}
<i>o,p'</i> -DDT	2b	354.5	100	2.82×10^{-4}	100	2.82×10^{-4}	10	2.82×10^{-5}
<i>p</i> -n Nonylphenol	2b	220.4	100	4.54×10^{-4}	10	4.54×10^{-5}	100	4.54×10^{-4}
Vinclozolin	2b	286.1	100	3.50×10^{-4}	10	3.50×10^{-5}	100	3.50×10^{-4}
12- <i>O</i> - Tetradecanoylphorbol-13-acetate	3	616.8	1000	1.62×10^{-3}	100	1.62×10^{-4}	10	1.62×10^{-5}
17 α -Estradiol	3	272.4	1000	3.67×10^{-3}	1000	3.67×10^{-3}	10	3.67×10^{-5}
17 β -Estradiol	3	272.4	1000	3.67×10^{-3}	1000	3.67×10^{-3}	10	3.67×10^{-5}

Chemical Name	Study Phase	FW	XDS Max Concentration Tested		ECVAM Max Concentration Tested		Hiyoshi Max Concentration Tested	
			µg/mL	M	µg/mL	M	µg/mL	M
2-sec-Butylphenol	3	150.2	1000	6.66×10^{-3}	100	6.66×10^{-4}	100	6.66×10^{-4}
2,4,5-Trichloro-phenoxyacetic acid	3	255.5	1000	3.91×10^{-3}	1000	3.91×10^{-3}	1000	3.91×10^{-3}
4-Androstenedione	3	286.4	100	3.49×10^{-4}	10	3.49×10^{-5}	100	3.49×10^{-4}
4-Cumylphenol	3	212.3	1000	4.71×10^{-3}	1000	4.71×10^{-3}	100	4.71×10^{-4}
4-Hydroxytamoxifen	3	387.5	1000	2.58×10^{-3}	100	2.58×10^{-4}	10	2.58×10^{-5}
4-tert-Octylphenol	3	206.3	1000	4.85×10^{-3}	100	4.85×10^{-4}	10	4.85×10^{-5}
5α-Dihydrotestosterone	3	290.4	1000	3.44×10^{-3}	10	3.44×10^{-5}	10	3.44×10^{-5}
Actinomycin D	3	1255.4	1000	7.97×10^{-4}	100	7.97×10^{-5}	100	7.97×10^{-5}
Apigenin	3	270.2	1000	3.70×10^{-3}	1000	3.70×10^{-3}	100	3.70×10^{-4}
Clomiphene citrate	3	598.1	1000	1.67×10^{-3}	100	1.67×10^{-4}	10	1.67×10^{-5}
Coumestrol	3	268.2	1000	3.73×10^{-3}	100	3.73×10^{-4}	10	3.73×10^{-5}
Daidzein	3	254.2	1000	3.93×10^{-3}	100	3.93×10^{-4}	100	3.93×10^{-4}
Dexamethasone	3	392.5	1000	2.55×10^{-3}	1000	2.55×10^{-3}	10	2.55×10^{-5}
Di - n -butyl phthalate	3	278.3	1000	3.59×10^{-3}	1000	3.59×10^{-3}	100	3.59×10^{-4}
Dibenzo[a,h]anthracene	3	278.4	10	3.59×10^{-5}	1	3.59×10^{-6}	10	3.59×10^{-5}
Dicofol	3	370.5	1000	2.70×10^{-3}	1000	2.70×10^{-3}	10	2.70×10^{-5}
Diethylhexyl phthalate	3	330.2	1000	3.03×10^{-3}	1000	3.03×10^{-3}	10	3.03×10^{-5}
Estrone	3	270.4	1000	3.70×10^{-3}	100	3.70×10^{-4}	10	3.70×10^{-5}
Ethyl paraben	3	166.2	1000	6.02×10^{-3}	1000	6.02×10^{-3}	100	6.02×10^{-4}
Fluoranthene	3	202.3	1000	4.94×10^{-3}	100	4.94×10^{-4}	1000	4.94×10^{-3}
Hydroxyflutamide	3	292.2	1000	3.42×10^{-3}	100	3.42×10^{-4}	100	3.42×10^{-4}
Kaempferol	3	286.2	1000	3.49×10^{-3}	100	3.49×10^{-4}	100	3.49×10^{-4}
Kepone	3	490.6	1000	2.04×10^{-3}	1000	2.04×10^{-3}	10	2.04×10^{-5}
meso-Hexestrol	3	270.4	1000	3.70×10^{-3}	1000	3.70×10^{-3}	100	3.70×10^{-4}
Methyl testosterone	3	302.5	1000	3.31×10^{-3}	100	3.31×10^{-4}	100	3.31×10^{-4}
Morin	3	302.2	1000	3.31×10^{-3}	100	3.31×10^{-4}	1000	3.31×10^{-3}
Norethynodrel	3	298.4	1000	3.35×10^{-3}	100	3.35×10^{-4}	100	3.35×10^{-4}
p,p'-DDE	3	318.0	1000	3.14×10^{-3}	1000	3.14×10^{-3}	10	3.14×10^{-5}
p,p'-Methoxychlor	3	345.7	1000	2.89×10^{-3}	1000	2.89×10^{-3}	10	2.89×10^{-5}
Phenobarbital	3	232.2	1000	4.31×10^{-3}	100	4.31×10^{-4}	NT	NT
Phenolphthalin	3	320.3	1000	3.12×10^{-3}	1000	3.12×10^{-3}	1000	3.12×10^{-3}
Progesterone	3	314.5	100	3.18×10^{-4}	100	3.18×10^{-4}	10	3.18×10^{-5}
Propylthiouracil	3	170.2	1000	5.87×10^{-3}	1000	5.87×10^{-3}	1000	5.87×10^{-3}

Chemical Name	Study Phase	FW	XDS Max Concentration Tested		ECVAM Max Concentration Tested		Hiyoshi Max Concentration Tested	
			µg/mL	M	µg/mL	M	µg/mL	M
Raloxifene HCl	3	510.1	1000	1.96×10^{-3}	100	1.96×10^{-4}	10	1.96×10^{-5}
Resveratrol	3	228.2	1000	4.38×10^{-3}	100	4.38×10^{-4}	100	4.38×10^{-4}
Sodium azide	3	65.0	100	1.54×10^{-3}	100	1.54×10^{-3}	100	1.54×10^{-3}
Tamoxifen	3	371.5	100	2.69×10^{-4}	100	2.69×10^{-4}	10	2.69×10^{-5}
Testosterone	3	288.4	1000	3.47×10^{-3}	100	3.47×10^{-4}	100	3.47×10^{-4}
17β-Trenbolone	4	270.4	1000	3.70×10^{-3}	NT	NT	NT	NT
19-Nortestosterone	4	274.4	1000	3.64×10^{-3}	NT	NT	NT	NT
4-Hydroxyandrostenedione	4	302.4	1000	3.31×10^{-3}	NT	NT	NT	NT
Ammonium perchlorate	4	117.5	1000	8.51×10^{-3}	NT	NT	NT	NT
Apomorphine	4	267.3	1000	3.74×10^{-3}	NT	NT	NT	NT
Bicalutamide	4	430.4	1000	2.32×10^{-3}	NT	NT	NT	NT
Chrysin	4	254.2	1000	3.93×10^{-3}	NT	NT	NT	NT
Cycloheximide	4	281.4	1000	3.55×10^{-3}	NT	NT	NT	NT
Cyproterone acetate	4	416.9	1000	2.40×10^{-3}	NT	NT	NT	NT
Fenarimol	4	331.2	1000	3.02×10^{-3}	NT	NT	NT	NT
Finasteride	4	372.5	1000	2.68×10^{-3}	NT	NT	NT	NT
Fluoxymestrone	4	336.4	1000	2.97×10^{-3}	NT	NT	NT	NT
Flutamide	4	276.2	1000	3.62×10^{-3}	NT	NT	NT	NT
Haloperidol	4	375.9	100	2.66×10^{-4}	NT	NT	NT	NT
Ketoconazole	4	531.4	10	9.41×10^{-5}	NT	NT	NT	NT
L-Thyroxine	4	776.9	1000	1.29×10^{-3}	NT	NT	NT	NT
Linuron	4	249.1	1000	4.01×10^{-3}	NT	NT	NT	NT
Medroxyprogesterone acetate	4	386.5	100	2.59×10^{-4}	NT	NT	NT	NT
Mifepristone	4	429.6	1000	2.33×10^{-3}	NT	NT	NT	NT
Nilutamide	4	317.2	1000	3.15×10^{-3}	NT	NT	NT	NT
Oxazepam	4	286.7	1000	3.49×10^{-3}	NT	NT	NT	NT
Pimozide	4	461.6	100	2.17×10^{-4}	NT	NT	NT	NT
Procymidone	4	284.1	100	3.52×10^{-4}	NT	NT	NT	NT
Reserpine	4	608.7	1000	1.64×10^{-3}	NT	NT	NT	NT
Spironolactone	4	416.6	1000	2.40×10^{-3}	NT	NT	NT	NT

Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; FW = formula weight; M = molar; Max = maximum; NT = not tested; XDS = Xenobiotic Detection Systems, Inc.

Table 4-11 Antagonist Range Finder Starting Concentrations in Culture Medium

Chemical Name	Study Phase	FW	XDS Max Concentration Tested		ECVAM Max Concentration Tested		Hiyoshi Max Concentration Tested	
			µg/mL	M	µg/mL	M	µg/mL	M
Dibenzo[<i>a,h</i>]anthracene	2a	278.4	10	3.59×10^{-5}	10	3.59×10^{-5}	10	3.59×10^{-5}
<i>p</i> -n -Nonylphenol	2a	220.4	1	4.54×10^{-6}	100	4.54×10^{-4}	10	4.54×10^{-5}
Progesterone	2a	314.5	100	3.18×10^{-4}	100	3.18×10^{-4}	10	3.18×10^{-5}
Tamoxifen	2a	371.5	10	2.69×10^{-5}	100	2.69×10^{-4}	10	2.69×10^{-5}
Apigenin	2b	270.2	100	3.70×10^{-4}	10	3.70×10^{-5}	10	3.70×10^{-5}
Atrazine	2b	215.7	100	4.64×10^{-4}	100	4.64×10^{-4}	100	4.64×10^{-4}
Butylbenzyl phthalate	2b	312.4	100	3.20×10^{-4}	10	3.20×10^{-5}	10	3.20×10^{-5}
Corticosterone	2b	346.5	1000	2.89×10^{-3}	1000	2.89×10^{-3}	100	2.89×10^{-4}
Flavone	2b	222.2	100	4.50×10^{-4}	100	4.50×10^{-4}	10	4.50×10^{-5}
Genistein	2b	270.2	100	3.70×10^{-4}	100	3.70×10^{-4}	10	3.70×10^{-5}
<i>o,p'</i> -DDT	2b	354.5	100	2.82×10^{-4}	NA	NA	10	2.82×10^{-5}
Resveratrol	2b	228.2	100	4.38×10^{-4}	100	4.38×10^{-4}	100	4.38×10^{-4}
12- <i>O</i> -Tetradecanoylphorbol-13-acetate	3	616.8	1000	1.62×10^{-3}	1000	1.62×10^{-3}	10	1.62×10^{-5}
17 α -Estradiol	3	272.4	1000	3.67×10^{-3}	100	3.67×10^{-4}	10	3.67×10^{-5}
17 α -Ethinyl estradiol	3	296.4	100	3.37×10^{-4}	10	3.37×10^{-5}	10	3.37×10^{-5}
17 β -Estradiol	3	272.4	1000	3.67×10^{-3}	100	3.67×10^{-4}	10	3.67×10^{-5}
2- <i>sec</i> -Butylphenol	3	150.2	1000	6.66×10^{-3}	1000	6.66×10^{-3}	100	6.66×10^{-4}
2,4,5-Trichlorophenoxyacetic acid	3	255.5	1000	3.91×10^{-3}	100	3.91×10^{-4}	1000	3.91×10^{-3}
4-Androstenedione	3	286.4	100	3.49×10^{-4}	100	3.49×10^{-4}	100	3.49×10^{-4}
4-Cumylphenol	3	212.3	100	4.71×10^{-4}	1000	4.71×10^{-3}	10	4.71×10^{-5}
4-Hydroxytamoxifen	3	387.5	100	2.58×10^{-4}	100	2.58×10^{-4}	10	2.58×10^{-5}
4- <i>tert</i> -Octylphenol	3	206.3	1000	4.85×10^{-3}	100	4.85×10^{-4}	10	4.85×10^{-5}
5 α -Dihydrotestosterone	3	290.4	1000	3.44×10^{-3}	100	3.44×10^{-4}	10	3.44×10^{-5}
Actinomycin D	3	1255.4	1000	7.97×10^{-4}	100	7.97×10^{-5}	100	7.97×10^{-5}
Bisphenol A	3	228.3	1000	4.38×10^{-3}	100	4.38×10^{-4}	100	4.38×10^{-4}
Bisphenol B	3	242.3	1000	4.13×10^{-3}	100	4.13×10^{-4}	100	4.13×10^{-4}
Clomiphene citrate	3	598.1	100	1.67×10^{-4}	100	1.67×10^{-4}	10	1.67×10^{-5}
Coumestrol	3	268.2	1000	3.73×10^{-3}	100	3.73×10^{-4}	10	3.73×10^{-5}
Daidzein	3	254.2	1000	3.93×10^{-3}	100	3.93×10^{-4}	10	3.93×10^{-5}

Chemical Name	Study Phase	FW	XDS Max Concentration Tested		ECVAM Max Concentration Tested		Hiyoshi Max Concentration Tested	
			µg/mL	M	µg/mL	M	µg/mL	M
Dexamethasone	3	392.5	100	2.55×10^{-4}	100	2.55×10^{-4}	100	2.55×10^{-4}
Di- <i>n</i> -butyl phthalate	3	278.3	1000	3.59×10^{-3}	1000	3.59×10^{-3}	10	3.59×10^{-5}
Dicofol	3	370.5	10	2.70×10^{-5}	1000	2.70×10^{-3}	10	2.70×10^{-5}
Diethylhexyl phthalate	3	330.2	100	3.03×10^{-4}	1000	3.03×10^{-3}	10	3.03×10^{-5}
Diethylstilbestrol	3	268.4	100	3.73×10^{-4}	100	3.73×10^{-4}	10	3.73×10^{-5}
Estrone	3	270.4	100	3.70×10^{-4}	10	3.70×10^{-5}	10	3.70×10^{-5}
Ethyl paraben	3	166.2	1000	6.02×10^{-3}	1000	6.02×10^{-3}	1000	6.02×10^{-3}
Fluoranthene	3	202.3	1000	4.94×10^{-3}	100	4.94×10^{-4}	10	4.94×10^{-5}
Hydroxyflutamide	3	292.2	1000	3.42×10^{-3}	1000	3.42×10^{-3}	100	3.42×10^{-4}
Kaempferol	3	286.2	100	3.49×10^{-4}	100	3.49×10^{-4}	10	3.49×10^{-5}
Kepone	3	490.6	1000	2.04×10^{-3}	1000	2.04×10^{-3}	10	2.04×10^{-5}
<i>meso</i> -Hexestrol	3	270.4	100	3.70×10^{-4}	100	3.70×10^{-4}	10	3.70×10^{-5}
Methyl testosterone	3	302.5	1000	3.31×10^{-3}	1000	3.31×10^{-3}	100	3.31×10^{-4}
Morin	3	302.2	1000	3.31×10^{-3}	100	3.31×10^{-4}	100	3.31×10^{-4}
Norethynodrel	3	298.4	1000	3.35×10^{-3}	1000	3.35×10^{-3}	10	3.35×10^{-5}
<i>p,p'</i> -DDE	3	318.0	1000	3.14×10^{-3}	100	3.14×10^{-4}	10	3.14×10^{-5}
<i>p,p'</i> -Methoxychlor	3	345.7	10	2.89×10^{-5}	1000	2.89×10^{-3}	10	2.89×10^{-5}
Phenobarbital	3	232.2	1000	4.31×10^{-3}	1000	4.31×10^{-3}	NT	NT
Phenolphthalin	3	320.3	1000	3.12×10^{-3}	1000	3.12×10^{-3}	1000	3.12×10^{-3}
Propylthiouracil	3	170.2	1000	5.87×10^{-3}	1000	5.87×10^{-3}	100	5.87×10^{-4}
Raloxifene HCl	3	510.1	10	1.96×10^{-5}	100	1.96×10^{-4}	10	1.96×10^{-5}
Sodium azide	3	65.0	100	1.54×10^{-3}	100	1.54×10^{-3}	100	1.54×10^{-3}
Testosterone	3	288.4	1000	3.47×10^{-3}	1000	3.47×10^{-3}	100	3.47×10^{-4}
Vinclozolin	3	286.1	1000	3.50×10^{-3}	100	3.50×10^{-4}	10	3.50×10^{-5}
17β-Trenbolone	4	270.4	1000	3.70×10^{-3}	NT	NT	NT	NT
19-Nortestosterone	4	274.4	1000	3.64×10^{-3}	NT	NT	NT	NT
4-Hydroxyandrostenedione	4	302.4	100	3.31×10^{-4}	NT	NT	NT	NT
Ammonium perchlorate	4	117.5	1000	8.51×10^{-3}	NT	NT	NT	NT
Apomorphine	4	267.3	1000	3.74×10^{-3}	NT	NT	NT	NT
Bicalutamide	4	430.4	1000	2.32×10^{-3}	NT	NT	NT	NT
Chrysin	4	254.2	1000	3.93×10^{-3}	NT	NT	NT	NT
Cycloheximide	4	281.4	1000	3.55×10^{-3}	NT	NT	NT	NT
Cyproterone acetate	4	416.9	1000	2.40×10^{-3}	NT	NT	NT	NT
Fenarimol	4	331.2	1000	3.02×10^{-3}	NT	NT	NT	NT

Chemical Name	Study Phase	FW	XDS Max Concentration Tested		ECVAM Max Concentration Tested		Hiyoshi Max Concentration Tested	
			µg/mL	M	µg/mL	M	µg/mL	M
Finasteride	4	372.5	100	2.68×10^{-4}	NT	NT	NT	NT
Fluoxymestron	4	336.4	100	2.97×10^{-4}	NT	NT	NT	NT
Flutamide	4	276.2	1000	3.62×10^{-3}	NT	NT	NT	NT
Haloperidol	4	375.9	100	2.66×10^{-4}	NT	NT	NT	NT
Ketoconazole	4	531.4	100	1.88×10^{-4}	NT	NT	NT	NT
L-Thyroxine	4	776.9	100	1.29×10^{-4}	NT	NT	NT	NT
Linuron	4	249.1	1000	4.01×10^{-3}	NT	NT	NT	NT
Medroxyprogesterone acetate	4	386.5	10	2.59×10^{-5}	NT	NT	NT	NT
Mifepristone	4	429.6	1000	2.33×10^{-3}	NT	NT	NT	NT
Nilutamide	4	317.2	1000	3.15×10^{-3}	NT	NT	NT	NT
Oxazepam	4	286.7	1000	3.49×10^{-3}	NT	NT	NT	NT
Pimozide	4	461.6	100	2.17×10^{-4}	NT	NT	NT	NT
Procymidone	4	284.1	100	3.52×10^{-4}	NT	NT	NT	NT
Reserpine	4	608.7	100	1.64×10^{-4}	NT	NT	NT	NT
Spironolactone	4	416.6	1000	2.40×10^{-3}	NT	NT	NT	NT

Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; FW = formula weight; M = molar; Max = maximum; NT = not tested; XDS = Xenobiotic Detection Systems, Inc.

4.4 Test Results for Coded Test Substances

4.4.1 Cell Viability Assessment

Cell viability was assessed to determine if reduction of ER TA activity is the result of cell loss. The visual observation method described in **Section 2.5.2** was used to assess cell viability in all wells of the test plates. Cell viability results from range finder testing were used to establish starting concentrations for comprehensive testing (see **Sections 2.6.1** and **2.6.2**) and to identify cytotoxic concentrations in comprehensive testing. This was of particular importance in antagonist testing because it is critical for distinguishing whether reduction of ER TA activity is caused by cell loss or ER antagonism. **Annex G3** lists the lowest concentrations that produced cell viability scores of 2 or greater for each substance evaluated in agonist and antagonist range finder and comprehensive testing.

Results were evaluated to determine if differences in cell viability were responsible for ER TA activity discordances among the laboratories. Ten substances were identified as discordant for ER TA agonist activity: 4-androstenedione, atrazine, 2-sec-butylphenol, clomiphene citrate, corticosterone, dicofol, flavone, fluoranthene, resveratrol, and vinclozolin (see **Table 4-12**). However, evaluation of range finder and comprehensive testing results indicated that the discordance was not due to differences in cell viability. Two substances were positive for ER TA antagonist activity at one laboratory but negative or inconclusive at the other two laboratories (17-α estradiol was positive at XDS but negative at ECVAM and inconclusive at Hiyoshi; clomiphene citrate was positive at Hiyoshi, negative at ECVAM, and inconclusive at XDS [see **Table 4-13**]). However, all cells for

these substances were viable below the 1.0×10^{-5} M limit concentration for determining ER TA antagonist activity, indicating that the discordance was not due to differences in cell viability.

4.4.2 BG1Luc ER TA Agonist and Antagonist Data

Test substances were evaluated in a phased approach as follows:

- Phase 2a. Four coded agonist and four coded antagonist substances were tested independently at least three times at each laboratory.
- Phase 2b. Eight coded agonist and eight coded antagonist substances were tested independently at least three times at each laboratory.
- Phase 3. Up to 41 coded agonist and 41 coded antagonist substances were tested at least once at each laboratory.
- Phase 4. The lead laboratory (XDS) tested 25 coded substances once each to further characterize the remainder of the 78 ICCVAM reference substances. Several of these substances had been assigned presumptive calls (ICCVAM 2003a, 2006; OECD 2007), but no ER TA data were available.

The results from Phases 2 and 3 are provided in **Table 4-12** (agonist) and **Table 4-13** (antagonist). **Table 4-14** provides the Phase 4 data generated by the lead laboratory.

Table 4-12 Agonist Summary Data for Phases 2a, 2b, and 3

Chemical	Study Phase	Laboratory	EC ₅₀ (M)	SD	CV (%)	# Plates for EC ₅₀ / # Plates Tested ^a	Classification ^b
Bisphenol A	2a	XDS	3.86×10^{-7}	3.27×10^{-8}	8	3/8	P (3/3)
		ECVAM	8.18×10^{-7}	2.53×10^{-8}	3	3/16	P (3/3)
		Hiyoshi	3.95×10^{-7}	1.86×10^{-8}	5	3/4	P (3/3)
Bisphenol B	2a	XDS	1.60×10^{-7}	2.56×10^{-8}	16	3/7	P (3/3)
		ECVAM	1.74×10^{-7}	5.25×10^{-8}	30	3/14	P (3/3)
		Hiyoshi	2.52×10^{-7}	7.44×10^{-9}	3	3/4	P (3/3)
Corticosterone	2a	XDS	-	-	-	0/8	N (3/3)
		ECVAM	NC	-	-	0/16	P (3/3)
		Hiyoshi	-	-	-	0/4	N (4/4)
Diethylstilbestrol	2a	XDS	4.87×10^{-11}	1.98×10^{-11}	41	3/9	P (3/3)
		ECVAM	3.60×10^{-11}	2.55×10^{-11}	71	2/14	P (3/3)
		Hiyoshi	2.07×10^{-11}	7.97×10^{-12}	39	4/4	P (4/4)
Atrazine	2b	XDS	-	-	-	4/6	N (4/4)
		ECVAM	7.43×10^{-5}	1.25×10^{-4}	168	3/11	P (3/3)
		Hiyoshi	-	-	-	0/4	N (3/3)
Butylbenzyl phthalate	2b	XDS	1.18×10^{-6}	3.57×10^{-7}	30	3/3	P (3/3)
		ECVAM	2.17×10^{-6}	9.92×10^{-7}	46	3/3	P (3/3)
		Hiyoshi	2.92×10^{-6}	3.69×10^{-7}	13	2/3	P (3/3)
o,p'-DDT	2b	XDS	6.12×10^{-8}	1.87×10^{-8}	30	3/3	P (3/3)
		ECVAM	4.22×10^{-7}	6.20×10^{-8}	15	3/5	P (3/3)
		Hiyoshi	6.98×10^{-7}	9.19×10^{-8}	13	3/3	P (3/3)

Chemical	Study Phase	Laboratory	EC ₅₀ (M)	SD	CV (%)	# Plates for EC ₅₀ / # Plates Tested ^a	Classification ^b
17- α Ethinyl estradiol	2b	XDS	7.60×10^{-12}	2.32×10^{-12}	31	4/7	P (4/4)
		ECVAM	5.85×10^{-12}	1.44×10^{-12}	25	3/3	P (3/3)
		Hiyoshi	8.38×10^{-12}	1.99×10^{-12}	24	3/4	P (3/3)
Flavone	2b	XDS	-	-	-	0/3	N (3/3)
		ECVAM	7.05×10^{-6}	8.82×10^{-7}	13	3/5	P (3/3)
		Hiyoshi	NC	-	-	0/4	P (3/3)
Genistein	2b	XDS	2.09×10^{-8}	6.01×10^{-9}	29	3/3	P (3/3)
		ECVAM	3.00×10^{-7}	3.24×10^{-8}	11	3/5	P (3/3)
		Hiyoshi	4.39×10^{-7}	1.76×10^{-7}	40	4/5	P (4/4)
<i>p</i> -n-Nonylphenol	2b	XDS	1.78×10^{-6}	6.95×10^{-8}	4	3/6	P (3/3)
		ECVAM	2.50×10^{-6}	1.06×10^{-6}	43	3/5	P (3/3)
		Hiyoshi	5.83×10^{-6}	2.89×10^{-7}	5	2/4	P (3/3)
Vinclozolin	2b	XDS	-	-	-	0/6	N (4/4)
		ECVAM	4.45×10^{-6}	3.57×10^{-6}	80	3/8	P (6/6)
		Hiyoshi	-	-	-	0/5	N (4/4)
Actinomycin D	3	XDS	-	-	-	0/3	I (1/1)
		ECVAM	-	-	-	0/2	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
4-Androstenedione	3	XDS	-	-	-	0/1	I (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	NC	-	-	0/1	P (1/1)
Apigenin	3	XDS	2.74×10^{-6}	-	-	1/1	P (1/1)
		ECVAM	1.63×10^{-6}	1.09×10^{-6}	67	3/4	P (3/3)
		Hiyoshi	1.62×10^{-6}	-	-	1/1	P (1/1)
Clomiphene citrate	3	XDS	-	-	-	0/1	I (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	4.38×10^{-8}	-	-	1/1	P (1/1)
Coumestrol	3	XDS	2.40×10^{-12}	-	-	1/3	P (1/1)
		ECVAM	2.58×10^{-7}	-	-	1/4	P (1/1)
		Hiyoshi	5.00×10^{-9}	-	-	1/1	P (1/1)
4-Cumylphenol	3	XDS	2.62×10^{-7}	-	-	1/1	P (1/1)
		ECVAM	3.03×10^{-7}	-	-	1/1	P (1/1)
		Hiyoshi	3.95×10^{-7}	-	-	1/1	P (1/1)
Daidzein	3	XDS	6.84×10^{-7}	-	-	1/1	P (1/1)
		ECVAM	1.19×10^{-6}	-	-	1/1	P (1/1)
		Hiyoshi	7.39×10^{-7}	-	-	1/1	P (1/1)
Dibenzo[<i>a,h</i>]anthracene	3	XDS	-	-	-	0/1	I (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/2	N (2/2)

Chemical	Study Phase	Laboratory	EC ₅₀ (M)	SD	CV (%)	# Plates for EC ₅₀ / # Plates Tested ^a	Classification ^b
Di- <i>n</i> -butyl phthalate	3	XDS	NC	-	-	0/1	P (1/1)
		ECVAM	1.91×10^{-7}	-	-	1/1	P (1/1)
		Hiyoshi	7.98×10^{-6}	6.60×10^{-7}	8	2/2	P (2/2)
<i>p,p'</i> -DDE	3	XDS	-	-	-	0/4	I (2/2)
		ECVAM	-	-	-	0/1	I (1/1)
		Hiyoshi	-	-	-	0/4	N (4/4)
Diethylhexyl phthalate	3	XDS	NC	-	-	0/1	P (1/1)
		ECVAM	-	-	-	0/1	I (1/1)
		Hiyoshi	-	-	-	0/1	I (1/1)
Dexamethasone	3	XDS	-	-	-	0/1	I (1/1)
		ECVAM	9.63×10^{-6}	-	-	1/1	P (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
5 α -Dihydrotestosterone	3	XDS	-	-	-	0/1	I (1/1)
		ECVAM	-	-	-	0/1	I (1/1)
		Hiyoshi	8.97×10^{-8}	2.56×10^{-8}	29	2/2	P (2/2)
Dicofol	3	XDS	2.22×10^{-6}	-	-	1/1	P (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	NC	-	-	0/1	P (1/1)
17- α Estradiol	3	XDS	4.85×10^{-12}	-	-	1/2	P (1/1)
		ECVAM	2.46×10^{-9}	3.53×10^{-9}	143	3/4	P (3/3)
		Hiyoshi	3.32×10^{-10}	-	-	1/1	P (1/1)
17- β Estradiol	3	XDS	1.34×10^{-11}	-	-	1/2	P (1/1)
		ECVAM	NC	-	-	0/2	P (1/1)
		Hiyoshi	3.37×10^{-12}	-	-	1/1	P (1/1)
Ethyl paraben	3	XDS	-	-	-	0/1	I (1/1)
		ECVAM	3.19×10^{-5}	-	-	1/2	P (1/1)
		Hiyoshi	2.12×10^{-5}	1.96×10^{-6}	9	2/2	P (1/1)
Estrone	3	XDS	3.52×10^{-10}	-	-	1/1	P (1/1)
		ECVAM	2.36×10^{-10}	-	-	1/2	P (1/1)
		Hiyoshi	1.82×10^{-10}	-	-	1/2	P (1/1)
Fluoranthene	3	XDS	2.03×10^{-5}	-	-	1/1	P (1/1)
		ECVAM	-	-	-	0/2	N (1/1)
		Hiyoshi	9.30×10^{-6}	-	-	1/1	P (1/1)
<i>meso</i> -Hexestrol	3	XDS	2.36×10^{-11}	-	-	1/2	P (1/1)
		ECVAM	1.16×10^{-11}	-	-	1/4	P (1/1)
		Hiyoshi	1.53×10^{-11}	3.77×10^{-12}	25	2/2	P (2/2)
Hydroxyflutamide	3	XDS	-	-	-	0/6	N (1/1)
		ECVAM	-	-	-	0/2	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)

Chemical	Study Phase	Laboratory	EC ₅₀ (M)	SD	CV (%)	# Plates for EC ₅₀ / # Plates Tested ^a	Classification ^b
Kepone	3	XDS	9.19×10^{-7}	-	-	1/2	P (1/1)
		ECVAM	1.23×10^{-7}	-	-	1/1	P (1/1)
		Hiyoshi	4.32×10^{-7}	-	-	1/1	P (1/1)
Kaempferol	3	XDS	7.65×10^{-6}	-	-	1/2	P (1/1)
		ECVAM	NC	-	-	0/1	P (1/1)
		Hiyoshi	3.35×10^{-7}	-	-	1/1	P (1/1)
<i>p,p'</i> -Methoxychlor	3	XDS	2.88×10^{-6}	-	-	1/4	P (2/2)
		ECVAM	1.22×10^{-6}	-	-	1/1	P (1/1)
		Hiyoshi	1.80×10^{-6}	1.09×10^{-6}	61	2/2	P (2/2)
Morin	3	XDS	2.62×10^{-5}	-	-	1/2	P (1/1)
		ECVAM	2.68×10^{-5}	-	-	1/1	P (1/1)
		Hiyoshi	4.80×10^{-5}	-	-	1/1	P (1/1)
Methyl testosterone	3	XDS	5.22×10^{-7}	4.50×10^{-7}	86	3/6	P (3/3)
		ECVAM	1.25×10^{-5}	-	-	1/1	P (1/1)
		Hiyoshi	2.36×10^{-6}	-	-	1/2	P (2/2)
Norethynodrel	3	XDS	1.39×10^{-9}	7.25×10^{-10}	52	2/4	P (2/2)
		ECVAM	3.65×10^{-10}	-	-	1/2	P (1/1)
		Hiyoshi	6.03×10^{-10}	-	-	1/2	P (2/2)
4- <i>tert</i> -Octylphenol	3	XDS	-	-	-	0/1	I (1/1)
		ECVAM	5.38×10^{-8}	-	-	1/1	P (1/1)
		Hiyoshi	1.01×10^{-8}	-	-	1/3	P (3/3)
4-Hydroxytamoxifen	3	XDS	-	-	-	0/1	I (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/3	N (3/3)
Phenobarbital	3	XDS	-	-	-	0/4	N (2/2)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	NT	NT	NT	0/0	NT
Phenolphthalein	3	XDS	2.40×10^{-5}	-	-	1/2	P (1/1)
		ECVAM	9.99×10^{-5}	-	-	1/1	P (1/1)
		Hiyoshi	8.33×10^{-5}	1.24×10^{-5}	15	2/2	P (2/2)
Progesterone	3	XDS	5.06×10^{-6}	-	-	1/4	P (2/2)
		ECVAM	1.27×10^{-6}	-	-	1/1	P (1/1)
		Hiyoshi	1.18×10^{-6}	5.08×10^{-7}	43	1/2	P (2/2)
Propylthiouracil	3	XDS	-	-	-	0/3	N (2/2)
		ECVAM	-	-	-	0/3	I (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
Raloxifene HCl	3	XDS	-	-	-	0/1	N (1/1)
		ECVAM	-	-	-	0/2	N (1/1)
		Hiyoshi	-	-	-	0/2	N (2/2)

Chemical	Study Phase	Laboratory	EC ₅₀ (M)	SD	CV (%)	# Plates for EC ₅₀ / # Plates Tested ^a	Classification ^b
Resveratrol	3	XDS	3.97×10^{-6}	-	-	1/2	P (1/1)
		ECVAM	-	-	-	0/1	I (1/1)
		Hiyoshi	-	-	-	0/3	N (3/3)
Sodium azide	3	XDS	-	-	-	0/4	N (3/3)
		ECVAM	-	-	-	0/3	I (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
2-sec-Butylphenol	3	XDS	1.18×10^{-9}	-	-	1/1	P (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	2.95×10^{-5}	8.24×10^{-6}	28	2/2	P (2/2)
Tamoxifen	3	XDS	-	-	-	0/2	I (1/1)
		ECVAM	-	-	-	0/1	I (1/1)
		Hiyoshi	6.73×10^{-8}	-	-	1/2	P (2/2)
2,4,5-Trichlorophenoxyacetic acid	3	XDS	-	-	-	0/1	I (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/2	N (2/2)
Testosterone	3	XDS	4.88×10^{-7}	5.77×10^{-7}	118	3/4	P (3/3)
		ECVAM	NC	-	-	0/1	P (1/1)
		Hiyoshi	9.95×10^{-5}	-	-	1/2	P (2/2)
12-O-Tetradecanoylphorbol-13-acetate	3	XDS	-	-	-	0/5	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)

Abbreviations: CV = coefficient of variation; ECVAM = European Centre for the Validation of Alternative Methods; EC₅₀ = half-maximal effective concentration; I = inadequate (positive or negative classification could not be determined because of poor-quality data); M = molar; N = negative; NC = not calculated; NT = not tested; P = positive; SD = standard deviation; XDS = Xenobiotic Detection Systems, Inc.

^a Values represent the number of acceptable plates used to determine the EC₅₀ value vs. the total number of plates tested (includes all acceptable and unacceptable plates).

^b Number in parentheses represents test results (P, N, or I) over the total number of acceptable trials.

Table 4-13 Antagonist Summary Data for Phases 2a, 2b, and 3

Chemical	Study Phase	Laboratory	IC ₅₀ (M)	SD	CV (%)	# Plates for IC ₅₀ / # Plates Tested ^a	Classification ^b
Dibenzo[<i>a,h</i>]anthracene	2a	XDS	NC	-	-	0/6	P (3/3)
		ECVAM	NC	-	-	0/4	P (3/3)
		Hiyoshi	NC	-	-	0/3	P (3/3)
Progesterone	2a	XDS	-	-	-	0/6	N (3/3)
		ECVAM	-	-	-	0/4	N (3/3)
		Hiyoshi	-	-	-	0/3	N (3/3)
<i>p</i> -n-Nonylphenol	2a	XDS	-	-	-	0/6	N (3/3)
		ECVAM	-	-	-	0/4	N (2/3)
		Hiyoshi	-	-	-	0/3	N (3/3)
Tamoxifen	2a	XDS	8.28×10^{-7}	2.36×10^{-7}	29	4/8	P (4/4)
		ECVAM	4.31×10^{-7}	2.69×10^{-7}	6	3/10	P (3/3)
		Hiyoshi	1.19×10^{-6}	3.67×10^{-6}	31	3/3	P (3/3)
Apigenin	2b	XDS	-	-	-	0/3	N (3/3)
		ECVAM	-	-	-	0/5	N (3/3)
		Hiyoshi	-	-	-	0/4	N (4/4)
Atrazine	2b	XDS	-	-	-	0/5	N (4/4)
		ECVAM	-	-	-	0/5	N (3/3)
		Hiyoshi	-	-	-	0/3	N (3/3)
Butylbenzyl phthalate	2b	XDS	-	-	-	0/3	N (3/3)
		ECVAM	-	-	-	0/4	N (3/3)
		Hiyoshi	-	-	-	0/4	N (4/4)
Corticosterone	2b	XDS	-	-	-	0/3	N (3/3)
		ECVAM	-	-	-	0/4	N (3/3)
		Hiyoshi	-	-	-	0/3	N (3/3)
<i>o,p'</i> -DDT	2b	XDS	-	-	-	0/3	N (3/3)
		ECVAM	-	-	-	0/4	N (3/3)
		Hiyoshi	-	-	-	0/4	N (4/4)
Flavone	2b	XDS	-	-	-	0/3	N (3/3)
		ECVAM	-	-	-	0/5	N (3/3)
		Hiyoshi	-	-	-	0/4	N (4/4)
Genistein	2b	XDS	-	-	-	0/3	N (3/3)
		ECVAM	-	-	-	0/4	N (3/3)
		Hiyoshi	-	-	-	0/3	N (3/3)

Chemical	Study Phase	Laboratory	IC ₅₀ (M)	SD	CV (%)	# Plates for IC ₅₀ / # Plates Tested ^a	Classification ^b
Resveratrol	2b	XDS	-	-	-	0/3	N (3/3)
		ECVAM	-	-	-	0/5	N (3/3)
		Hiyoshi	-	-	-	0/3	N (3/3)
Actinomycin D	3	XDS	2.67×10^{-7}	-	-	1/6	P (1/1)
		ECVAM	1.98×10^{-8}	-	-	1/3	P (1/1)
		Hiyoshi	NC	-	-	0/1	P (1/1)
Bisphenol A	3	XDS	-	-	-	0/5	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
Bisphenol B	3	XDS	-	-	-	0/1	I (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
Diethylstilbestrol	3	XDS	-	-	-	0/2	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	1.70×10^{-5}	-	-	1/1	P (1/1)
17- α Ethinyl estradiol	3	XDS	-	-	-	0/1	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
4-Androstenedione	3	XDS	-	-	-	0/2	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
Clomiphene citrate	3	XDS	-	-	-	0/2	I (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	NC	-	-	0/1	P (1/1)
Coumestrol	3	XDS	-	-	-	0/1	N (1/1)
		ECVAM	-	-	-	0/3	N (0/2)
		Hiyoshi	-	-	-	0/1	N (1/1)
4-Cumylphenol	3	XDS	-	-	-	0/2	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
Daidzein	3	XDS	-	-	-	0/1	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
Di- <i>n</i> -butyl phthalate	3	XDS	-	-	-	0/5	N (2/2)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)

Chemical	Study Phase	Laboratory	IC ₅₀ (M)	SD	CV (%)	# Plates for IC ₅₀ / # Plates Tested ^a	Classification ^b
<i>p,p'</i> -DDE	3	XDS	-	-	-	0/1	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
Diethylhexyl phthalate	3	XDS	-	-	-	0/3	N (1/1)
		ECVAM	-	-	-	0/3	N (2/2)
		Hiyoshi	-	-	-	0/2	N (1/1)
Dexamethasone	3	XDS	-	-	-	0/1	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
5 α -Dihydrotestosterone	3	XDS	-	-	-	0/6	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
Dicofol	3	XDS	-	-	-	0/2	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/2	N (1/1)
17 α Estradiol	3	XDS	4.26 \times 10 ⁻⁶	-	-	1/2	P (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	I (1/1)
17 β Estradiol	3	XDS	-	-	-	0/4	N (1/1)
		ECVAM	-	-	-	0/2	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
Ethyl paraben	3	XDS	-	-	-	0/2	N (1/1)
		ECVAM	-	-	-	0/3	N (2/2)
		Hiyoshi	-	-	-	0/2	N (1/1)
Estrone	3	XDS	-	-	-	0/2	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
Fluoranthene	3	XDS	-	-	-	0/6	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/2	N (1/1)
<i>meso</i> -Hexestrol	3	XDS	-	-	-	0/1	N (1/1)
		ECVAM	-	-	-	0/1	I (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
Hydroxyflutamide	3	XDS	-	-	-	0/1	N (1/1)
		ECVAM	-	-	-	0/3	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)

Chemical	Study Phase	Laboratory	IC ₅₀ (M)	SD	CV (%)	# Plates for IC ₅₀ / # Plates Tested ^a	Classification ^b
Kepone	3	XDS	-	-	-	0/1	N (1/1)
		ECVAM	-	-	-	0/2	N (1/1)
		Hiyoshi	-	-	-	0/2	N (1/1)
Kaempferol	3	XDS	-	-	-	0/1	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
<i>p,p'</i> -Methoxychlor	3	XDS	-	-	-	0/5	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
Morin	3	XDS	-	-	-	0/3	N (1/1)
		ECVAM	-	-	-	0/2	N (1/1)
		Hiyoshi	-	-	-	0/2	N (1/1)
Methyl testosterone	3	XDS	-	-	-	0/6	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
Norethynodrel	3	XDS	-	-	-	0/3	N (1/1)
		ECVAM	-	-	-	0/2	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
4- <i>tert</i> -Octylphenol	3	XDS	-	-	-	0/1	N (1/1)
		ECVAM	-	-	-	0/2	N (1/1)
		Hiyoshi	-	-	-	0/2	N (1/1)
4-Hydroxytamoxifen	3	XDS	4.13×10^{-7}	5.77×10^{-7}	140	2/3	P (3/3)
		ECVAM	-	-	-	0/2	I (1/1)
		Hiyoshi	3.87×10^{-9}	-	-	1/1	P (1/1)
Phenobarbital	3	XDS	-	-	-	0/1	N (1/1)
		ECVAM	-	-	-	0/2	N (1/1)
		Hiyoshi	NT	NT	NT	0/0	NT
Phenolphthalein	3	XDS	-	-	-	0/6	N (1/1)
		ECVAM	-	-	-	0/3	N (2/2)
		Hiyoshi	-	-	-	0/1	N (1/1)
Propylthiouracil	3	XDS	-	-	-	0/5	N (1/1)
		ECVAM	-	-	-	0/3	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
Raloxifene HCl	3	XDS	2.16×10^{-9}	-	-	1/1	P (1/1)
		ECVAM	5.41×10^{-10}	-	-	1/1	P (1/1)
		Hiyoshi	8.84×10^{-10}	-	-	1/1	P (1/1)

Chemical	Study Phase	Laboratory	IC ₅₀ (M)	SD	CV (%)	# Plates for IC ₅₀ / # Plates Tested ^a	Classification ^b
Sodium azide	3	XDS	-	-	-	0/4	N (1/1)
		ECVAM	-	-	-	0/4	N (1/1)
		Hiyoshi	-	-	-	0/2	N (1/1)
2-sec-Butylphenol	3	XDS	-	-	-	0/1	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
2,4,5-Trichlorophenoxyacetic acid	3	XDS	-	-	-	0/1	N (1/1)
		ECVAM	-	-	-	0/3	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
Testosterone	3	XDS	-	-	-	0/6	N (1/1)
		ECVAM	-	-	-	0/4	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
12-O-Tetradecanoylphorbol-13-acetate	3	XDS	-	-	-	0/4	N (2/2)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)
Vinclozolin	3	XDS	-	-	-	0/1	N (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	N (1/1)

Abbreviations: CV = coefficient of variation; ECVAM = European Centre for the Validation of Alternative Methods; I = inadequate (positive or negative classification could not be determined because of poor-quality data); IC₅₀ = half-maximal inhibitory concentration; M = molar; N = negative; NC = not calculated; NT = not tested; P = positive; SD = standard deviation; XDS = Xenobiotic Detection Systems, Inc.

^a Values represent the number of acceptable plates used to determine the IC₅₀ value vs. the total number of plates tested (includes all acceptable and unacceptable plates).

^b Number in parentheses represents test results (P, N, or I) over the total number of acceptable trials.

Table 4-14 Phase 4 Results from XDS

Chemical	Agonist			Antagonist		
	EC ₅₀ ^a (M)	Classification ^b	# Plates Tested	IC ₅₀ ^c (M)	Classification ^b	# Plates Tested
17β-Trenbolone	9.58 × 10 ⁻⁸	P (1/1)	2	-	N (2/2)	4
19-Nortestosterone	1.80 × 10 ⁻⁶	P (1/1)	1	-	N (1/1)	1
4-Hydroxyandrostenedione	3.91 × 10 ⁻⁵	P (1/1)	2	-	N (1/1)	1
Ammonium perchlorate	-	N (1/1)	3	-	N (1/1)	1
Apomorphine	-	N (2/2)	3	NC	P (1/1)	1
Bicalutamide	-	N (1/1)	2	-	N (1/1)	1
Chrysin	3.20 × 10 ⁻⁶	P (2/2)	3	-	N (1/1)	1
Cycloheximide	-	I (2/2)	1	9.67 × 10 ⁻⁷	P (1/1)	1

Chemical	Agonist			Antagonist		
	EC ₅₀ ^a (M)	Classification ^b	# Plates Tested	IC ₅₀ ^c (M)	Classification ^b	# Plates Tested
Cyproterone acetate	-	N (1/2)	4	-	N (1/1)	1
Fenarimol	4.59×10^{-6}	P (2/2)	6	-	N (1/1)	4
Finasteride	-	N (1/1)	3	-	N (1/1)	2
Fluoxymestron	2.22×10^{-5}	P (2/2)	4	-	N (1/1)	1
Flutamide	-	I (1/1)	3	-	N (1/1)	1
Haloperidol	-	N (1/1)	3	-	N (1/1)	1
Ketoconazole	-	N (1/1)	3	1.23×10^{-6}	P (1/1)	3
L-Thyroxine	-	N (2/2)	4	-	N (1/1)	1
Linuron	-	N (2/2)	5	-	N (1/1)	1
Medroxyprogesterone acetate	-	N (2/2)	5	NC	P (1/1)	1
Mifepristone	-	N (2/2)	2	-	N (1/1)	1
Nilutamide	NC	P (1/1)	2	-	N (2/2)	4
Oxazepam	-	I (1/1)	3	-	N (1/1)	1
Pimozide	-	N (1/1)	1	-	N (1/1)	1
Procymidone	-	I (1/1)	3	-	N (1/1)	3
Reserpine	-	N (2/2)	5	-	I (1/1)	1
Spironolactone	-	N (1/1)	2	-	N (1/1)	1

Abbreviations: EC₅₀ = half-maximal effective concentration; I = inadequate (positive or negative classification could not be determined because of poor-quality data); IC₅₀ = half-maximal inhibitory concentration; M = molar; N = negative; NC = not calculated; P = positive.

^a EC₅₀ values are from one test, except 4-hydroxyandrostenedione (mean value from two tests [SD = 3.91×10^{-5} ; coefficient of variation = 52%]).

^b Number in parentheses represents test results (P, N, or I) over the total number of acceptable trials.

^c IC₅₀ values are from one test.

5.0 Accuracy of the BG1Luc ER TA Test Method

This section discusses the accuracy of the BG1Luc ER TA test method in the multilaboratory validation effort. Accuracy is evaluated by assessing the following:

- **Concordance:** The proportion of all substances tested that are correctly classified as positive or negative. It is a measure of test method performance, and it is often used interchangeably with *accuracy*.
- **Sensitivity:** The proportion of all positive substances that are classified correctly as positive in a test method. It is a measure of test method accuracy.
- **Specificity:** The proportion of all negative substances that are classified correctly as negative in a test method. It is a measure of test method accuracy.
- **False positive rate:** The proportion of all negative (inactive) substances falsely identified as positive. It is a measure of test method performance.
- **False negative rate:** The proportion of all positive (active) substances falsely identified as negative. It is a measure of test method performance.

Each of these variables can be calculated as follows (**Table 5-1**):

Table 5-1 Template for Concordance Analysis

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

a = positive in assay and positive by reference test classification

b = positive in assay and negative by reference test classification

c = negative in assay and positive by reference test classification

d = negative in assay and negative by reference test classification

$$\text{Concordance} = ([a+d]/[a+b+c+d])$$

$$\text{Sensitivity} = (a/[a+c])$$

$$\text{Specificity} = (d/[b+d])$$

$$\text{False positive rate} = (b/[b+d])$$

$$\text{False negative rate} = (c/[a+c])$$

The BG1Luc ER TA test method was evaluated for its ability to correctly identify ER agonists and antagonists. For this analysis, test substance classification (positive or negative for ER agonist/antagonist activity) obtained during the validation study was compared to the classification of the same substance based on a preponderance of published data. Positive or negative classifications based on BG1Luc ER TA data were based on the majority classification assigned using results from each of the three participating laboratories (XDS, ECVAM, and Hiyoshi). For example, if a substance tested positive at one laboratory but negative in the other two, the overall classification would be negative for the accuracy calculations. Substances that failed to meet the decision criteria for either a positive or negative response, defined in **Section 2.7**, are considered inadequate for analysis. The classification of data as “inadequate” is due to poor data quality and would normally require retesting.

However, this classification system was developed after testing was complete; therefore, these substances were excluded from the accuracy analyses described here.

5.1 Substances Used for Accuracy Analysis

As detailed in **Section 3.2**, NICEATM completed a comprehensive literature review of available *in vitro* data to identify substances that could be considered unequivocally positive or negative for ER agonist or antagonist activity. A total of 48 unique reference substances were considered in the evaluation of test method accuracy. Separate lists were generated for evaluating accuracy based on agonist (42 substances: 33 positive, 9 negative) and antagonist (25 substances: 3 positive, 22 negative) activity. Nineteen substances appeared on both reference lists.

Table 5-2 lists the 42 reference substances used to evaluate test method accuracy for ER agonist activity. Of these 42 substances, seven (17%) had inadequate testing results and were therefore excluded from the analysis, leaving 35 (28 positive, 7 negative) substances for evaluation. The following seven substances had inadequate BG1Luc ER TA agonist test method data:

- Clomiphene citrate
- *p,p'*-DDE
- 5 α -Dihydrotestosterone
- Flutamide
- Procymidone
- Resveratrol
- Tamoxifen

These seven substances represent eight chemical classes (two cyclic hydrocarbons and one each of an amide, amine, carboxylic acid, halogenated hydrocarbon, heterocyclic compound, polycyclic compound, and steroid) and five product classes (four pharmaceuticals and one each of a fungicide, natural product, pesticide intermediate, and veterinary agent). The diversity of chemical and product classes indicates that no one category or class is overrepresented with inadequate data. Again, it should be emphasized that the “inadequate” classification is usually a result of poor data quality and would normally require retesting. However, this classification system was developed after testing was complete; therefore, retesting of these substances was not possible.

Table 5-3 lists the 25 reference substances used to evaluate test method accuracy for ER antagonist activity. Definitive classifications (positive or negative) were obtained for all 25 substances tested, allowing all substances to be used for the assessment of antagonist accuracy.

Table 5-2 42 ICCVAM-Recommended Substances Used to Evaluate ER Agonist Accuracy

Substance	CASRN	Classification ^a				
		ICCVAM Consensus	BG1Luc ER TA Consensus ^b	XDS	ECVAM	Hiyoshi
17 α -Estradiol	57-91-0	POS	POS	POS (1/1)	POS (3/3)	POS (2/2)
17 α -Ethinyl estradiol	57-63-6	POS	POS	POS (3/3)	POS (3/3)	POS (3/3)
17 β -Estradiol	50-28-2	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
19-Nortestosterone	434-22-0	POS	POS	POS (1/1)	NT	NT
4-Cumylphenol	599-64-4	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
4- <i>tert</i> -Octylphenol	140-66-9	POS	POS	I (1/1)	POS (1/1)	POS (2/2)
5 α -Dihydrotestosterone	521-18-6	POS	I	I (1/1)	I (1/1)	POS (1/1)
Apigenin	520-36-5	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
Atrazine	1912-24-9	NEG	NEG	NEG (3/3)	POS (3/3)	NEG (3/3)
Bicalutamide	90357-06-5	NEG	NEG	NEG (1/1)	NT	NT
Bisphenol A	80-05-7	POS	POS	POS (3/3)	POS (3/3)	POS (3/3)
Bisphenol B	77-40-7	POS	POS	POS (3/3)	POS (3/3)	POS (3/3)
Butylbenzyl phthalate	85-68-7	POS	POS	POS (3/3)	POS (3/3)	POS (3/3)
Chrysin	480-40-0	POS	POS	POS (2/2)	NT	NT
Clomiphene citrate	50-41-9	POS	I	I (1/1)	NEG (1/1)	POS (1/1)
Corticosterone	50-22-6	NEG	NEG	NEG (3/3)	POS (3/3)	NEG (3/3)
Coumestrol	479-13-0	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
Daidzein	486-66-8	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
Dicofol	115-32-2	POS	POS	POS (1/1)	NEG (1/1)	POS (1/1)
Diethylstilbestrol	56-53-1	POS	POS	POS (3/3)	POS (3/3)	POS (3/3)
Estrone	53-16-7	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
Ethyl paraben	120-47-8	POS	POS	I (1)	POS (1/1)	POS (1/1)
Fenarimol	60168-88-9	POS	POS	POS (1/1)	NT	NT
Flutamide	13311-84-7	NEG	I	I (1)	NT	NT
Genistein	446-72-0	POS	POS	POS (3/3)	POS (3/3)	POS (4/4)
Hydroxyflutamide	52806-53-8	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Kaempferol	520-18-3	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
Kepone	143-50-0	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
L-Thyroxine	51-48-9	POS	NEG	NEG (1/1)	NT	NT

Substance	CASRN	Classification ^a				
		ICCVAM Consensus	BG1Luc ER TA Consensus ^b	XDS	ECVAM	Hiyoshi
Linuron	330-55-2	NEG	NEG	NEG (1/1)	NT	NT
<i>meso</i> -Hexestrol	84-16-2	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
Methyl testosterone	58-18-4	POS	POS	POS (3/3)	POS (1/1)	POS (2/2)
Norethynodrel	68-23-5	POS	POS	POS (2/2)	POS (1/1)	POS (2/2)
<i>o,p'</i> -DDT	789-02-6	POS	POS	POS (3/3)	POS (3/3)	POS (3/3)
<i>p</i> -n-Nonylphenol	104-40-5	POS	POS	POS (3/3)	POS (3/3)	POS (2/3)
<i>p,p'</i> -DDE	72-55-9	POS	I	I (1/1)	I (1/1)	NEG (1/1)
<i>p,p'</i> -Methoxychlor	72-43-5	POS	POS	POS (1/1)	POS (1/1)	POS (2/2)
Phenobarbital	50-06-6	NEG	NEG	NEG (1/1)	NEG (1/1)	NT
Procymidone	32809-16-8	NEG	I	I (1/1)	NT	NT
Resveratrol	501-36-0	POS	I	POS (1/1)	I (1/1)	NEG (1/3)
Spironolactone	52-01-7	NEG	NEG	NEG (1/1)	NT	NT
Tamoxifen	10540-29-1	POS	I	I (1/1)	I (1/1)	POS (1/1)

Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; CASRN = CAS Registry Number (American Chemical Society); ECVAM = European Centre for the Validation of Alternative Methods; I = inadequate (positive or negative classification could not be determined because of poor-quality data); ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; NEG = negative; NT = not tested; POS = positive; XDS = Xenobiotic Detection Systems, Inc.

^a Numbers in parentheses represent test results (POS, NEG, or I) over the total number of trials that met test plate acceptance criteria.

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

Table 5-3 25 ICCVAM-Recommended Substances Used to Evaluate ER Antagonist Accuracy

Substance	CASRN	Classification ^a				
		ICCVAM Consensus	BG1Luc ER TA Consensus ^b	XDS	ECVAM	Hiyoshi
17 α -Ethinyl estradiol	57-63-6	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
4-Hydroxytamoxifen	68047-06-3	POS	POS	POS (1/1)	I (2/2)	POS (1/1)
5 α -Dihydrotestosterone	521-18-6	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Apigenin	520-36-5	NEG	NEG	NEG (3/3)	NEG (3/3)	NEG (4/4)
Bisphenol A	80-05-7	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Butylbenzyl phthalate	85-68-7	NEG	NEG	NEG (3/3)	NEG (3/3)	NEG (4/4)

Substance	CASRN	Classification ^a				
		ICCVAM Consensus	BG1Luc ER TA Consensus ^b	XDS	ECVAM	Hiyoshi
Chrysin	480-40-0	NEG	NEG	NEG (1/1)	NT	NT
Coumestrol	479-13-0	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Daidzein	486-66-8	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Di-n-butyl phthalate	84-74-2	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Dicofol	115-32-2	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Diethylhexyl phthalate	117-81-7	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Diethylstilbestrol	56-53-1	NEG	NEG	NEG (1/1)	NEG (1/1)	POS (1/1)
Genistein	446-72-0	NEG	NEG	NEG (3/3)	NEG (3/3)	NEG (3/3)
Kaempferol	520-18-3	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Kepone	143-50-0	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Mifepristone	84371-65-3	NEG	NEG	NEG (1/1)	NT	NT
Norethynodrel	68-23-5	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
<i>o,p'</i> -DDT	789-02-6	NEG	NEG	NEG (3/3)	NEG (3/3)	NEG (3/3)
<i>p</i> -n-Nonylphenol	104-40-5	NEG	NEG	NEG (3/3)	NEG (3/3)	NEG (3/3)
<i>p,p'</i> -DDE	72-55-9	NEG	NEG	NEG (1/1)	NEG (1/1)	NEG (1/1)
Progesterone	57-83-0	NEG	NEG	NEG (3/3)	NEG (3/3)	NEG (3/3)
Raloxifene HCl	82640-04-8	POS	POS	POS (1/1)	POS (1/1)	POS (1/1)
Resveratrol	501-36-0	NEG	NEG	NEG (3/3)	NEG (3/3)	NEG (3/3)
Tamoxifen	10540-29-1	POS	POS	POS (3/3)	POS (3/3)	POS (3/3)

Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; CASRN = CAS Registry Number (American Chemical Society); ECVAM = European Centre for the Validation of Alternative Methods; I = inadequate (positive or negative classification could not be determined because of poor-quality data); ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; NEG = negative; NT = not tested; POS = positive; XDS = Xenobiotic Detection Systems, Inc.

^a Numbers in parentheses represent test results (POS, NEG, or I) over the total number of trials that met test plate acceptance criteria.

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

5.2 Accuracy Analysis of the BG1Luc ER TA Agonist Test Method

The accuracy analysis using the 35 ICCVAM reference substances that produced a definitive BG1Luc ER TA result in agonist testing indicated accuracy of 97% (34/35), sensitivity of 96% (27/28), specificity of 100% (7/7), false positive rate of 0% (0/7), and false negative rate of 4% (1/28) (Table 5-4).

Table 5-4 Accuracy of the BG1Luc ER TA Agonist Test Method

N	Accuracy	Sensitivity	Specificity	False Positive Rate	False Negative Rate
35 ^a	97% (34/35)	96% (27/28)	100% (7/7)	0% (0/7)	4% (1/28)

Abbreviations: N = number.

^a A total 42 substances were evaluated in the BG1Luc ER TA agonist test method. Seven substances did not produce a consensus classification and were omitted, leaving 35 substances for analysis.

5.2.1 Discordant Results for Agonist Analysis

Among the 35 substances used to calculate accuracy statistics, only L-thyroxine was a false negative in the BG1Luc ER TA test method when compared to the ICCVAM reference classification (Table 5-5). This Phase 4 substance was tested once in one laboratory, XDS. This substance is classified as positive (2/3) by ICCVAM based on two reports of positive agonist activity and one report of no agonist activity. The two positive results were in GH3 cells (rat pituitary adenoma) (Fujimoto et al. 2004) and HeLa cells (human cervical carcinoma) (Takeyoshi 2006), whereas MCF-7 cells (human breast adenocarcinoma) (Fujimoto et al. 2004) showed no estrogenic response when exposed to L-thyroxine. These reports indicate a possible tissue-specific response to this chemical, which may explain the lack of ER agonist activity observed in this experiment with BG-1 cells (human ovarian carcinoma).

Table 5-5 Discordant Substance in the BG1Luc ER TA Agonist Test Method

Substance	CASRN	MeSH Chemical Class	Product Class	BG1Luc ER TA Classification	ICCVAM Reference Classification
L-Thyroxine	51-48-9	Amino Acid	Pharmaceutical, Veterinary Agent	NEG	POS

Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; CASRN = CAS Registry Number (American Chemical Society); ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; MeSH = Medical Subject Headings (National Library of Medicine); NEG = negative; N = number; POS = positive.

5.3 Accuracy Analysis of the BG1Luc ER TA Antagonist Test Method

Accuracy analysis conducted with the 25 reference substances that produced a definitive result in antagonist testing showed an overall accuracy of 100% (25/25), sensitivity of 100% (3/3), specificity of 100% (22/22), false positive rate of 0% (0/22), and false negative rate of 0% (0/3) (Table 5-6).

Table 5-6 Accuracy of the BG1Luc ER TA Antagonist Test Method

N	Accuracy	Sensitivity	Specificity	False Positive Rate	False Negative Rate
25	100% (25/25)	100% (3/3)	100% (22/22)	0% (0/22)	0% (0/3)

Abbreviations: N = number.

5.4 Comparison of BG1Luc ER TA Results with CERI-STTA (U.S. EPA OPPTS 890.1300) Results

The CERI-STTA (OECD 2009; Takeyoshi 2006) method for assessing ER α agonist activity of test substances is currently the only ER TA test method accepted by regulatory agencies. This test system utilizes the hER α -HeLa-9903 cell line, which is derived from a human cervical tumor, with two stably inserted constructs: the hER α expression construct (encoding the full-length human receptor) and a firefly luciferase reporter construct bearing five tandem repeats of a vitellogenin ERE driven by a mouse metallothionein promoter TATA element. Because the BG1Luc ER TA test method is another STTA that could be considered for regulatory use, a comparison of test method accuracy between these two test methods was conducted based on a list of ICCVAM-recommended agonist reference substances for which definitive classifications have been produced in both methods. These substances are listed in **Table 5-7**. The results show identical levels of accuracy when both methods tested the same agonist reference chemicals: concordance of 96% (25/26), sensitivity of 95% (21/22), and specificity of 100% (4/4) (**Table 5-8** and **Table 5-9**). The test methods differed only in the one false negative from each method: L-thyroxine was false negative in the BG1Luc ER TA test method, and *p*-n-nonylphenol was false negative in the CERI-STTA. Overall, these data suggest a very high level of agreement in the performance of these two assays.

Table 5-7 Substances Used in the Evaluation of Accuracy of the BG1Luc ER TA and CERI-STTA Test Method Results

Substance	CASRN	ICCVAM Reference Classification	BG1Luc ER TA Classification	CERI-STTA Classification ^a
17 α -Estradiol	57-91-0	POS	POS	POS
17 α -Ethinyl estradiol	57-63-6	POS	POS	POS
17 β -Estradiol	50-28-2	POS	POS	POS
4-Cumylphenol	599-64-4	POS	POS	POS
4- <i>tert</i> -Octylphenol	140-66-9	POS	POS	POS
Apigenin	520-36-5	POS	POS	POS
Atrazine	1912-24-9	NEG	NEG	NEG
Bisphenol A	80-05-7	POS	POS	POS
Bisphenol B	77-40-7	POS	POS	POS
Butylbenzyl phthalate	85-68-7	POS	POS	POS
Corticosterone	50-22-6	NEG	NEG	NEG

Substance	CASRN	ICCVAM Reference Classification	BG1Luc ER TA Classification	CERI-STTA Classification ^a
Coumestrol	479-13-0	POS	POS	POS
Daidzein	486-66-8	POS	POS	POS
Diethylstilbestrol	56-53-1	POS	POS	POS
Estrone	53-16-7	POS	POS	POS
Ethyl paraben	120-47-8	POS	POS	POS
Genistein	446-72-0	POS	POS	POS
Kaempferol	520-18-3	POS	POS	POS
Kepone	143-50-0	POS	POS	POS
Linuron	330-55-2	NEG	NEG	NEG
L-Thyroxine	51-48-9	POS	NEG	POS
Methyl testosterone	58-18-4	POS	POS	POS
Norethynodrel	68-23-5	POS	POS	POS
<i>p</i> -n-Nonylphenol	104-40-5	POS	POS	NEG
<i>p,p'</i> -Methoxychlor	72-43-5	POS	POS	POS
Spironolactone	52-01-7	NEG	NEG	NEG

Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; CASRN = CAS Registry Number (American Chemical Society); CERI = Chemicals Evaluation and Research Institute, Japan; I = inadequate (positive or negative classification could not be determined because of poor-quality data); ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; NEG = negative; OECD = Organisation for Economic Co-operation and Development; POS = positive; STTA = stably transfected transactivation assay.

^a Data published by the Chemicals Evaluation and Research Institute, Japan (CERI) (Takeyoshi 2006).

Table 5-8 Accuracy of the BG1Luc ER TA Test Method Assessed Using Agonist Reference Chemicals Listed in Table 5-7

		BG1Luc ER TA Agonist Classification		
		Positive	Negative	Total
ICCVAM Consensus Classification	Positive	21	1	22
	Negative	0	4	4
	Total	21	5	26

Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods.

Concordance 96% (25/26)

Sensitivity 95% (21/22)

Specificity 100% (4/4)

Table 5-9 Accuracy of the CERi-STTA Method Assessed Using Agonist Reference Chemicals Listed in Table 5-7

ICCVAM Consensus Classification		CERI-STTA Classification			
		Positive	Negative	Total	
		Positive	21	1	22
		Negative	0	4	4
Total		21	5	26	

Abbreviations: CERi = Chemicals Evaluation and Research Institute, Japan; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; STTA = stably transfected transactivation assay.

Concordance 96% (25/26)

Sensitivity 95% (21/22)

Specificity 100% (4/4)

5.5 Comparison of BG1Luc ER TA EC₅₀ and IC₅₀ Values with Values from ICCVAM Reference Data

Although the primary goal of the BG1Luc ER TA test method is to provide a qualitative assessment of estrogenic/anti-estrogenic activity, quantitative measures of activity (i.e., EC₅₀ and IC₅₀ values) are usually obtained for positive results. EC₅₀ and IC₅₀ values obtained from BG1Luc ER TA test results were compared to median values from other ER TA test methods reported in the literature. The substances used for these comparisons are listed in **Table 5-10** for EC₅₀ and **Table 5-11** for IC₅₀ comparisons. Regression analyses of these data are presented in **Figures 5-1** and **5-2**, respectively.

Based on EC₅₀ values obtained for 26 substances, the correlation coefficient between the log EC₅₀ for the BG1Luc ER TA agonist test method and that reported for other ER TA test methods in the literature was $R^2 = 0.839$. Although EC₅₀ values can differ by an order of magnitude between methods, this relatively high correlation indicates that the BG1Luc ER TA agonist test method might be considered for quantitative as well as qualitative assessment of estrogenic activity.

Likewise, based on IC₅₀ values obtained for three substances, the correlation coefficient between the log IC₅₀ for the BG1Luc ER TA antagonist test method and that reported for other ER TA test methods in the literature was $R^2 = 0.95$. Again, this high correlation suggests that the BG1Luc ER TA test method might also be considered for quantitative as well as qualitative assessment of anti-estrogenic activity. However, this conclusion is necessarily limited by the small number of substances (n = 3) upon which it is based.

Table 5-10 Median EC₅₀ Values for Substances Used to Generate EC₅₀ Linear Regression

Substance	BG1Luc ER TA Median EC ₅₀ (M)	ICCVAM Reference Data Median EC ₅₀ (M)
17 α -Estradiol	3.02×10^{-10}	5.20×10^{-09}
17 α -Ethinyl estradiol	7.09×10^{-12}	5.20×10^{-11}
17 β -Estradiol	3.37×10^{-12}	8.65×10^{-11}
19-Nortestosterone	1.65×10^{-06}	2.00×10^{-07}
4-Cumylphenol	3.03×10^{-07}	3.22×10^{-07}
4- <i>tert</i> -Octylphenol	2.08×10^{-08}	1.00×10^{-07}

Substance	BG1Luc ER TA Median EC ₅₀ (M)	ICCVAM Reference Data Median EC ₅₀ (M)
5 α -Dihydrotestosterone	8.97×10^{-08}	1.33×10^{-07}
Apigenin	1.40×10^{-06}	7.65×10^{-07}
Bisphenol A	3.95×10^{-07}	5.00×10^{-07}
Bisphenol B	2.36×10^{-07}	9.20×10^{-08}
Coumestrol	1.31×10^{-07}	1.60×10^{-08}
Daidzein	6.75×10^{-07}	4.90×10^{-07}
Dicofol	2.22×10^{-06}	7.05×10^{-06}
Diethylstilbestrol	2.08×10^{-11}	6.60×10^{-11}
Estrone	2.16×10^{-10}	2.10×10^{-09}
Fenarimol	9.15×10^{-06}	7.00×10^{-06}
Genistein	3.00×10^{-07}	6.75×10^{-08}
Kaempferol	2.55×10^{-07}	1.60×10^{-07}
<i>meso</i> -Hexestrol	1.62×10^{-11}	1.00×10^{-10}
Methyl testosterone	6.49×10^{-07}	1.58×10^{-08}
Norethynodrel	1.26×10^{-07}	6.40×10^{-09}
<i>o,p'</i> -DDT	4.22×10^{-07}	1.69×10^{-06}
<i>p</i> -n-Nonylphenol	2.50×10^{-06}	3.60×10^{-07}
<i>p,p'</i> -Methoxychlor	8.43×10^{-07}	5.25×10^{-06}
Tamoxifen	6.73×10^{-08}	5.30×10^{-07}
Testosterone	4.85×10^{-07}	2.00×10^{-07}

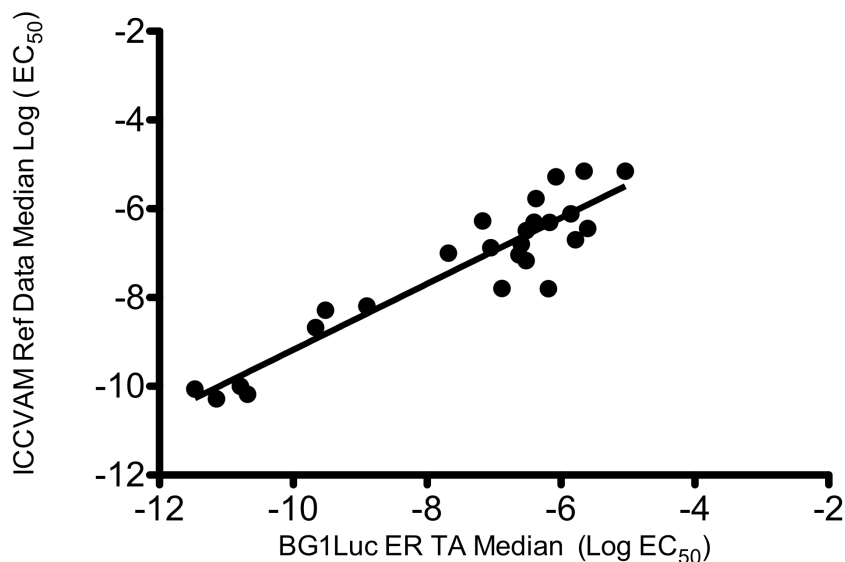
Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; EC₅₀ = half-maximal effective concentration; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; M = molar.

Table 5-11 Median IC₅₀ Values for Substances Used to Generate IC₅₀ Linear Regression

Substance Name	BG1Luc ER TA Median IC ₅₀ (M)	ICCVAM Reference Data Median IC ₅₀ (M)
4-Hydroxytamoxifen	4.94×10^{-09}	2.13×10^{-09}
Raloxifene HCl	1.24×10^{-09}	2.31×10^{-09}
Tamoxifen	7.12×10^{-07}	4.00×10^{-07}

Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; IC₅₀ = half-maximal inhibitory concentration; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; M = molar.

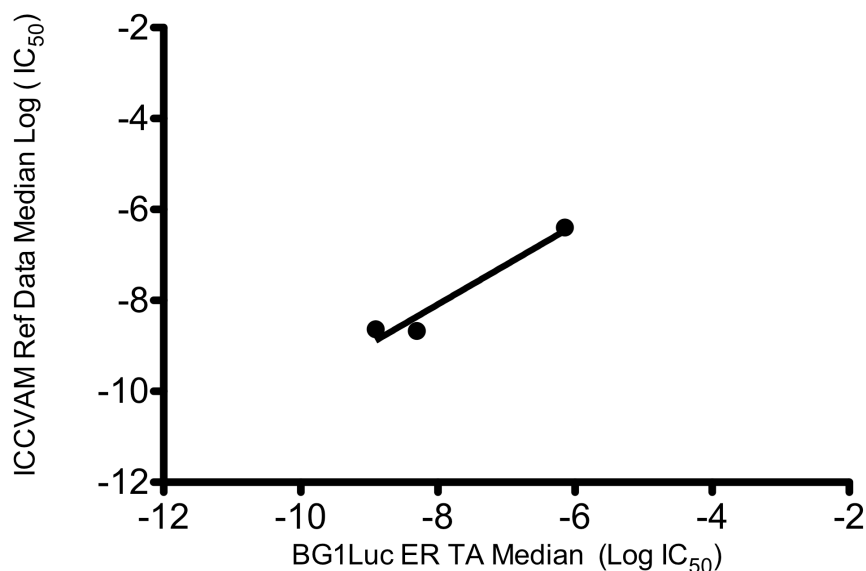
Figure 5-1 Relationship of EC₅₀ Values Obtained in the BG1Luc ER TA Test Method and EC₅₀ Values from ICCVAM Reference Data



Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; EC₅₀ = half-maximal effective concentration; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods.

Each point in this figure represents a median EC₅₀ value obtained in the BG1Luc ER TA test method compared with the median ICCVAM EC₅₀ value (from the literature reference data updated in 2010, discussed in **Section 3** and provided in **Annex N**).

Figure 5-2 Relationship of IC₅₀ Values Obtained in the BG1Luc ER TA Test Method and IC₅₀ Values from ICCVAM Reference Data



Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; IC₅₀ = half-maximal inhibitory concentration; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods.

Each point in this figure represents a median IC₅₀ value obtained in the BG1Luc ER TA test method compared with the median ICCVAM IC₅₀ value (from the literature reference data updated in 2010, discussed in **Section 3** and provided in **Annex N**).

5.6 Concordance of BG1Luc ER TA Results with Estrogen Receptor Binding Results

Results from the BG1Luc ER TA test method were examined for concordance with published reports of ER binding. ER binding results for the 34 reference substances used for this analysis, along with agonist and antagonist test results from the BG1Luc ER TA test method, are provided in **Table 5-12**. Because results in binding studies only indicate the ability to bind the ER receptor and therefore do not distinguish between agonist or antagonist activity, a positive result in the BG1Luc ER TA test method for either agonist or antagonist activity was considered positive in the concordance analysis provided in **Table 5-13**. There was 97% (33/34) concordance between the BG1Luc ER TA test method and ER binding data from the literature. The single discordant test substance was medroxyprogesterone acetate (MPA), which was positive in the BG1Luc ER TA antagonist assay but was reported in two published studies as negative for ER binding. MPA was tested once during Phase 4 at one participating laboratory. XDS reported an IC_{50} of 5.0×10^{-5} M. In light of the excellent degree of agreement between ER binding and the BG1Luc ER TA test method (with no false negative results), it appears that evaluating results from the BG1Luc ER TA agonist and antagonist testing would provide a viable alternative to conducting ER binding studies. This cannot currently be accomplished with the only accepted ER TA method due to the inability of the CERI-STTA to assess ER antagonist activity.

Table 5-12 Substances Used to Assess BG1Luc ER TA Concordance with ER Binding Data

Substance	CASRN	BG1 Agonist Classification	BG1 Antagonist Classification	Overall BG1 Classification	ER Binding Classification (Literature)
17 α -Estradiol	57-91-0	POS	I	POS	POS
17 α -Ethinyl estradiol	57-63-6	POS	NEG	POS	POS
17 β -Estradiol	50-28-2	POS	NEG	POS	POS
2-sec-Butylphenol	89-72-5	POS	NEG	POS	POS
4-Cumylphenol	599-64-4	POS	NEG	POS	POS
4-Hydroxytamoxifen	68047-06-3	NEG	POS	POS	POS
4-tert-Octylphenol	140-66-9	POS	NEG	POS	POS
Apigenin	520-36-5	POS	NEG	POS	POS
Bisphenol A	80-05-7	POS	NEG	POS	POS
Bisphenol B	77-40-7	POS	NEG	POS	POS
Butylbenzyl phthalate	85-68-7	POS	NEG	POS	POS
Corticosterone	50-22-6	NEG	NEG	NEG	NEG
Coumestrol	479-13-0	POS	NEG	POS	POS
Daidzein	486-66-8	POS	NEG	POS	POS
Dicofol	115-32-2	POS	NEG	POS	POS
Diethylstilbestrol	56-53-1	POS	NEG	POS	POS
Estrone	53-16-7	POS	NEG	POS	POS
Ethyl paraben	120-47-8	POS	NEG	POS	POS

Substance	CASRN	BG1 Agonist Classification	BG1 Antagonist Classification	Overall BG1 Classification	ER Binding Classification (Literature)
Fenarimol	60168-88-9	POS	NEG	POS	POS
Genistein	446-72-0	POS	NEG	POS	POS
Kaempferol	520-18-3	POS	NEG	POS	POS
Kepone	143-50-0	POS	NEG	POS	POS
L-Thyroxine	51-48-9	NEG	NEG	NEG	NEG
Medroxyprogesterone acetate	71-58-9	NEG	POS	POS	NEG
<i>meso</i> -Hexestrol	84-16-2	POS	NEG	POS	POS
Mifepristone	84371-65-3	NEG	NEG	POS	POS
Morin	480-16-0	POS	NEG	POS	POS
Norethynodrel	68-23-5	POS	NEG	POS	POS
<i>o,p'</i> -DDT	789-02-6	POS	NEG	POS	POS
<i>p</i> -n-Nonylphenol	104-40-5	POS	NEG	POS	POS
<i>p,p'</i> -Methoxychlor	72-43-5	POS	NEG	POS	POS
Phenolphthalin	81-90-3	POS	NEG	POS	POS
Raloxifene HCl	82640-04-8	NEG	POS	POS	POS
Tamoxifen	10540-29-1	I	POS	POS	POS

Abbreviations: BG1 = LUMI-CELL BG1Luc4E2 ER TA test method; CASRN = CAS Registry Number (American Chemical Society); ER = estrogen receptor; I = inadequate (positive or negative classification could not be determined because of poor-quality data); NEG = negative; POS = positive.

Table 5-13 Concordance of BG1Luc ER TA Results and ER Binding Results

		BG1Luc ER TA Classification		
		Positive	Negative	Total
ER Binding	Positive	31	0	31
	Negative	1	2	3
	Total	32	2	34

Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; ER = estrogen receptor.

Concordance 97% (33/34)

5.7 Comparison of BG1Luc ER TA Results with Uterotrophic Assay Results

Results from the BG1Luc ER TA test method were examined for concordance with published data from the uterotrophic assay (Owens and Ashby 2002; Owens and Koeter 2003). Data from the uterotrophic assay were available for 13 substances tested in the BG1Luc ER TA agonist test method (Table 5-14). Based on a comparison with the *in vivo* uterotrophic assay classification, the 13 substances with conclusive test results in the BG1Luc ER TA agonist test method produced overall

concordance of 92% (12/13) (**Table 5-15**). All substances found positive in the uterotrophic assay were also positive in the BG1Luc ER TA method. The only discordant substance, butylbenzyl phthalate, was positive for ER agonist activity in the BG1Luc ER TA agonist test method and negative in the uterotrophic assay. These data indicate that the BG1Luc ER TA agonist test method has very good agreement with the *in vivo* results obtained with the uterotrophic assay, with no false negative results.

Table 5-14 Substances Used in the Comparison of BG1Luc ER TA Agonist Classification and *In Vivo* Uterotrophic Assay Data

ICCVAM Reference Substance	CASRN	BG1Luc ER TA Agonist Classification	Overall Uterotrophic Assay Study Data	OECD Study Uterotrophic Assay Data ^a	CERI Study Uterotrophic Assay Data ^b
17 α Estradiol	57-91-0	POS	POS	NT	POS
17 α Ethinyl estradiol	57-63-6	POS	POS	POS	POS
4- <i>tert</i> -Octylphenol	140-66-9	POS	POS	NT	POS
4-Cumylphenol	599-64-4	POS	POS	NT	POS
Bisphenol A	80-05-7	POS	POS	POS	POS
Bisphenol B	77-40-7	POS	POS	NT	POS
Butylbenzyl phthalate	85-68-7	POS	NEG	NEG	NEG
Daidzein	486-66-8	POS	POS	NT	POS
Estrone	53-16-7	POS	POS	NT	POS
Genistein	446-72-0	POS	POS	POS	POS
Ketoconazole	65277-42-1	NEG	NEG	NT	NEG
Methyl testosterone	58-18-4	POS	POS	NT	POS
<i>o,p'</i> -DDT	789-02-6	POS	POS	POS	NT

Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; CASRN = CAS Registry Number (American Chemical Society); CERI = Chemicals Evaluation and Research Institute, Japan; NEG = negative; NT = not tested; OECD = Organisation for Economic Co-operation and Development; POS = positive.

^a Pooled data from the validation of the OECD uterotrophic bioassay (Kanno et al. 2003a, 2003b; Owens and Ashby 2002).

^b Data published by the Chemicals Evaluation and Research Institute, Japan (CERI), as part of a comparison database of ER TA and uterotrophic data (Takeyoshi 2006).

Table 5-15 **Concordance of BG1Luc ER TA Agonist Classification and *In Vivo* Uterotrophic Assay Data**

		BG1Luc ER TA Agonist Classification		
		Positive	Negative	Total
<i>In Vivo</i> Uterotrophic Data	Positive	11	0	11
	Negative	1	1	2
	Total	12	1	13

Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method.

Concordance 92% (12/13)

6.0 Test Method Reliability

An assessment of test method reliability (intra- and interlaboratory reproducibility) is an essential element of any evaluation of the performance of an alternative test method (ICCVAM 2003b). Intralaboratory reproducibility refers to the extent to which qualified personnel within the same laboratory can replicate results using a specific test protocol. Interlaboratory reproducibility refers to the extent to which different laboratories can replicate results using the same protocol and test substances. Interlaboratory reproducibility indicates the extent to which a test method can be transferred successfully among laboratories.

This section describes the reliability assessment for the BG1Luc ER TA test method, which was based on validation study results for substances tested multiple times within and across laboratories.

6.1 Intralaboratory Reproducibility

As discussed in **Section 4.2**, the agonist and antagonist DMSO control and antagonist E2 control RLU values were the only quantitative values used for acceptance criteria for agonist test plates throughout the study; therefore, intralaboratory reproducibility of the BG1Luc ER TA agonist and antagonist test methods was assessed by comparing (1) RLU values for the agonist and antagonist DMSO control and the antagonist E2 control for all plates tested within each laboratory during the course of the validation study and (2) results from Phases 2a and 2b testing, during which 12 substances were tested in at least three independent experiments in each of the three participating laboratories (XDS, ECVAM, and Hiyoshi).

6.1.1 Agonist DMSO Control

Because DMSO control RLU values are not normalized, they can vary considerably between test plates and across time. Therefore, intralaboratory reproducibility was evaluated by comparing the within-plate variability (CV) of the four replicate DMSO control RLU values for all test plates that passed acceptance criteria. The range of means and CV values for within-plate DMSO control RLU values are provided in **Table 6-1**. (See **Annex L** for the mean and CV values of individual agonist test plates.) Although mean plate DMSO RLU values ranged from a low of 511 to a high of 9885, with a mean of 3749, within-plate variability of DMSO control RLU values between replicate DMSO wells was low, with CV values ranging from 1% to 43% and a mean of 8%. Of the 218 agonist test plates that met acceptance criteria, only six plates had within-plate CV values greater than 20%. (See **Annex L** for individual test plate mean DMSO control RLU values and associated CV values.)

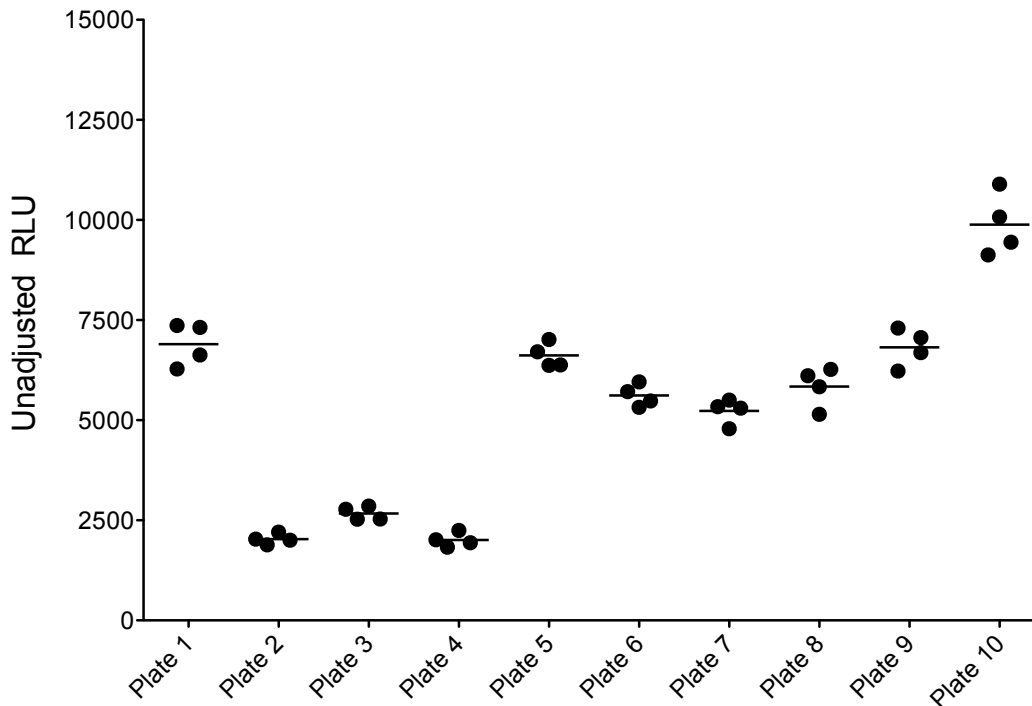
Table 6-1 Agonist Within-Plate DMSO Control Data

Laboratory	Mean and Range of DMSO Control RLU Values	Mean and Range of CV (%)	N
XDS	2800 (511–9885)	8 (1–43)	93
ECVAM	3379 (828–7306)	8 (1–33)	60
Hiyoshi	5465 (1362–9383)	6 (1–24)	65
All Laboratories	3749 (511–9885)	8 (1–43)	218

Abbreviations: CV = coefficient of variation; DMSO = dimethyl sulfoxide; ECVAM = European Centre for the Validation of Alternative Methods; N = number of plates that passed acceptance criteria; RLU = relative light unit; XDS = Xenobiotic Detection Systems, Inc.

Figures 6-1 through 6-3 show the within-plate agonist DMSO control RLU values for Phase 1 of the validation study as examples of the low variability for this parameter. As discussed above, within-plate CVs were low throughout the validation study.

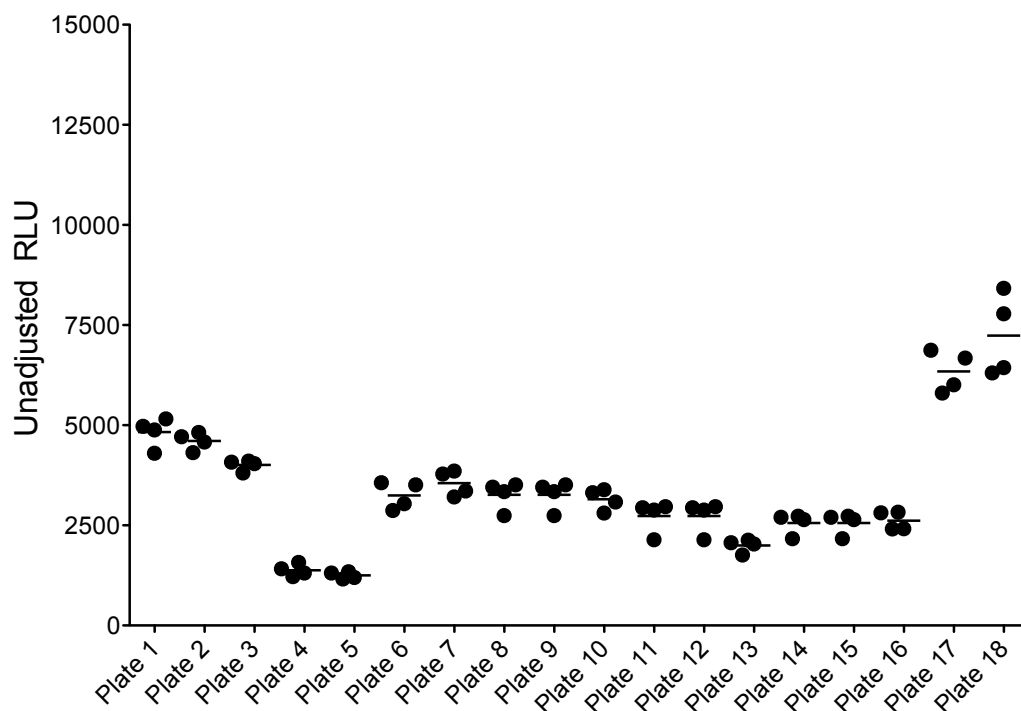
Figure 6-1 Agonist DMSO Control Within-Plate RLU Values During Phase 1 at XDS



Abbreviations: DMSO = dimethyl sulfoxide; RLU = relative light unit.

Each point represents the non-normalized DMSO value for a single well in a 96-well plate.

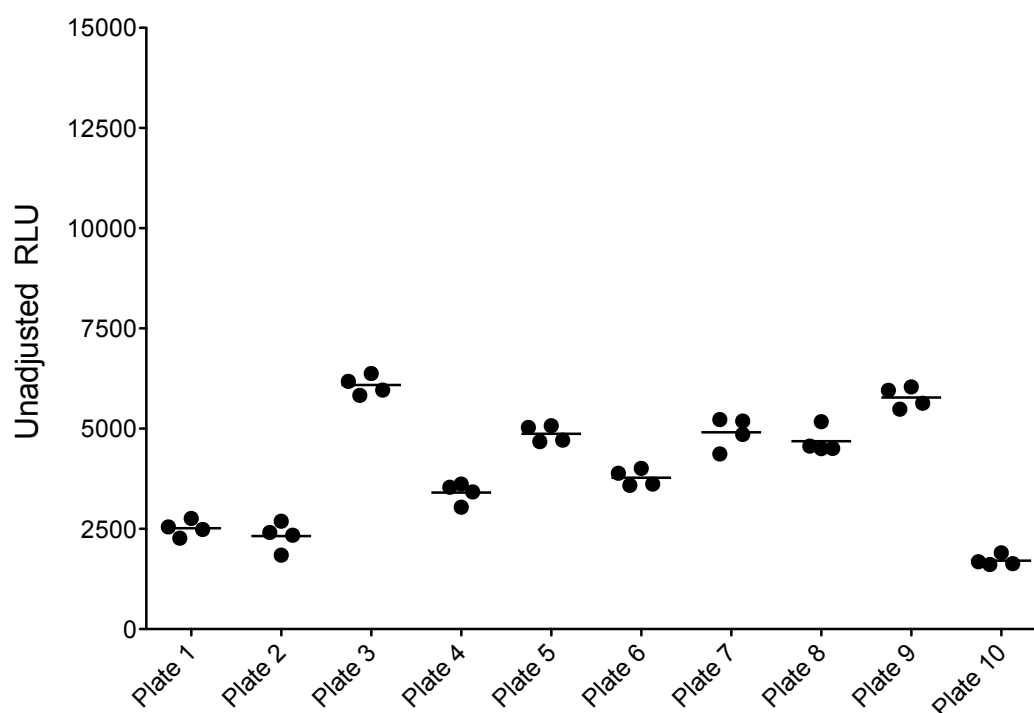
Within-plate DMSO variance at XDS during Phase 1 was fairly low, with coefficients of variation ranging from 5% to 9%.

Figure 6-2 Agonist DMSO Control Within-Plate RLU Values During Phase 1 at ECVAM

Abbreviations: DMSO = dimethyl sulfoxide; RLU = relative light unit.

Each point represents the non-normalized DMSO value for a single well in a 96-well plate.

Within-plate DMSO variance at ECVAM during Phase 1 was fairly low, with coefficients of variation ranging from 2% to 14%.

Figure 6-3 Agonist DMSO Control Within-Plate RLU Values During Phase 1 at Hiyoshi

Abbreviations: DMSO = dimethyl sulfoxide; RLU = relative light unit.

Each point represents the non-normalized DMSO value for a single well in a 96-well plate.

Within-plate DMSO variance at Hiyoshi during Phase 1 was fairly low, with coefficients of variation ranging from 4% to 15%.

6.1.2 Agonist E2 Reference Standard EC₅₀ and Methoxychlor Control

Although E2 reference standard EC₅₀ and Met control RLU values were not used for plate acceptance after Phase 2a of the validation study (see **Section 2.7.1**), these values were collected throughout the study for information purposes. The means and SDs for these parameters from all plates that passed acceptance criteria are provided in **Table 6-2**.

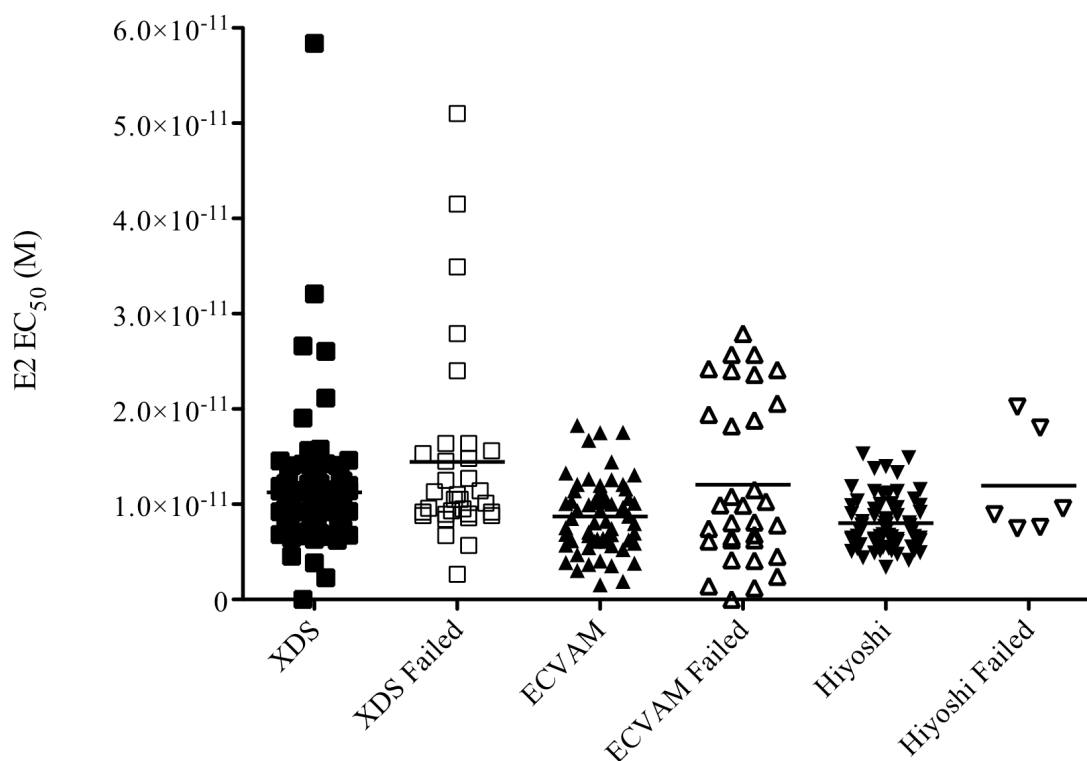
Table 6-2 Agonist E2 EC₅₀ and Methoxychlor Control Values

Laboratory	Mean	SD	N
E2 Reference Standard EC₅₀ (M)			
XDS	1.1×10^{-11}	6.7×10^{-12}	93
ECVAM	1.1×10^{-11}	1.9×10^{-11}	60
Hiyoshi	8.0×10^{-12}	2.8×10^{-12}	65
Methoxychlor (RLU)			
XDS	6075	1283	93
ECVAM	6246	1609	60
Hiyoshi	8029	1233	65

Abbreviations: EC₅₀ = half-maximal effective concentration; ECVAM = European Centre for the Validation of Alternative Methods; M = molar; N = number of plates that passed acceptance criteria; RLU = relative light unit; SD = standard deviation; XDS = Xenobiotic Detection Systems, Inc.

As shown in **Table 6-2**, mean E2 reference standard EC₅₀ values ranged from 8.0×10^{-12} to 1.1×10^{-11} M. Met control RLU values, which ranged from 6075 to 8029, were highest at Hiyoshi and lowest at XDS.

E2 reference standard EC₅₀ and Met control RLU values for all plates tested during the validation study are presented in **Figures 6-4** and **6-5**, respectively. The three laboratories were relatively consistent when data from only acceptable plates were considered. These data also indicated that the variability of each parameter is generally higher when only values obtained from plates that failed one or more acceptance criteria were considered. With the exception of E2 EC₅₀ at XDS, all outlier values among the parameters evaluated were associated with these failed plates.

Figure 6-4 Agonist E2 Reference Standard EC₅₀ Values

Abbreviations: EC₅₀ = half-maximal effective concentration; ECVAM = European Centre for the Validation of Alternative Methods; M = molar; XDS = Xenobiotic Detection Systems, Inc.

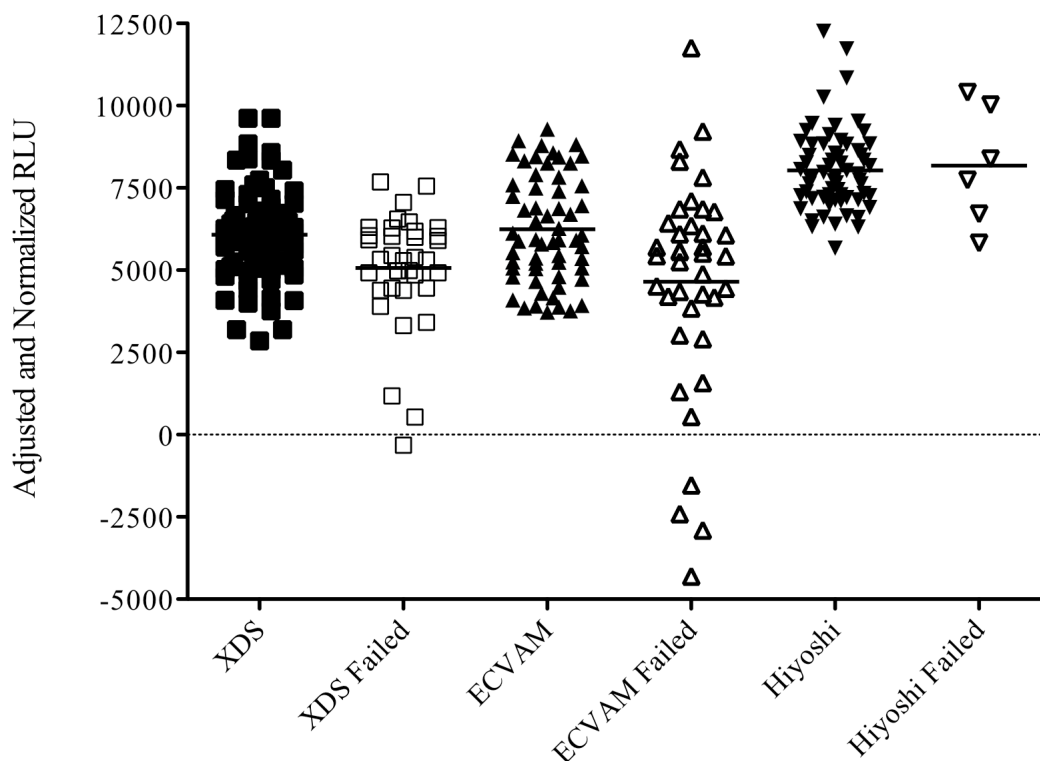
Each point represents a single plate.

An EC₅₀ value (1.18×10^{-9} M) from one experiment that failed acceptance criteria at XDS was excluded from the graph to minimize scale distortion.

EC₅₀ values (1.69×10^{-10} M and 7.78×10^{-11} M) from two experiments that passed acceptance criteria at XDS were excluded from the graph to minimize scale distortion.

An EC₅₀ value (1.56×10^{-10} M) from one experiment that passed acceptance criteria at ECVAM was excluded from the graph to minimize scale distortion.

Figure 6-5 Agonist Methoxychlor Control Values



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

Each point represents a single plate.

Methoxychlor control values (35581, -74511, and -6995) from three experiments that failed acceptance criteria at XDS were excluded from this graph to minimize scale distortion.

Methoxychlor control values (-127587 and -8464) from two experiments that failed acceptance criteria at ECVAM were excluded from the graph to minimize scale distortion.

6.1.3 Intralaboratory Reproducibility of Phase 2 Agonist Reference Substances

As described in **Section 2.0**, test substances were classified as positive or negative for agonist activity based on a specific set of criteria. The resulting classifications for each of the 12 substances that were tested at least three times at each laboratory were used to evaluate the extent of intralaboratory agreement (see **Table 6-3**). Although the classifications for some of the test substances differed among the laboratories, there was 100% agreement within each laboratory for each of the three repeat tests. No “inadequate” data were generated at any laboratory during this phase of the validation study.

Table 6-3 Intralaboratory Agreement for Multiple Testing of 12 Phase 2 Agonist Substances Tested Independently Three Times at Each Laboratory

Activity per Test	XDS	ECVAM	Hiyoshi
Agreement Within Laboratory	12/12 (100%)	12/12 (100%)	12/12 (100%)
+++	8/12	12/12	9/12
---	4/12	0/12	3/12
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.

6.1.4 Antagonist DMSO Control

Because DMSO control RLU values are not normalized, they can vary considerably between test plates and across time. Therefore, intralaboratory reproducibility was evaluated by comparing the within-plate variability (CV) of the DMSO control RLU values for all test plates that passed acceptance criteria. The range of means and CV values for within-plate DMSO control RLU values are provided in **Table 6-4**. (See **Annex L** for the mean and CV values of individual antagonist test plates.) Although mean plate DMSO RLU values ranged from 132 to 8451, with a mean of 3299, within-plate variability of DMSO control RLU values between replicate DMSO wells was low. Associated CV values ranged from 1% to 52%, with a mean of 8%. Of the 194 antagonist test plates that passed acceptance criteria, only eight plates had within-plate CV values greater than 20%. (See **Annex L** for individual test plate mean DMSO control RLU values and associated CV values.)

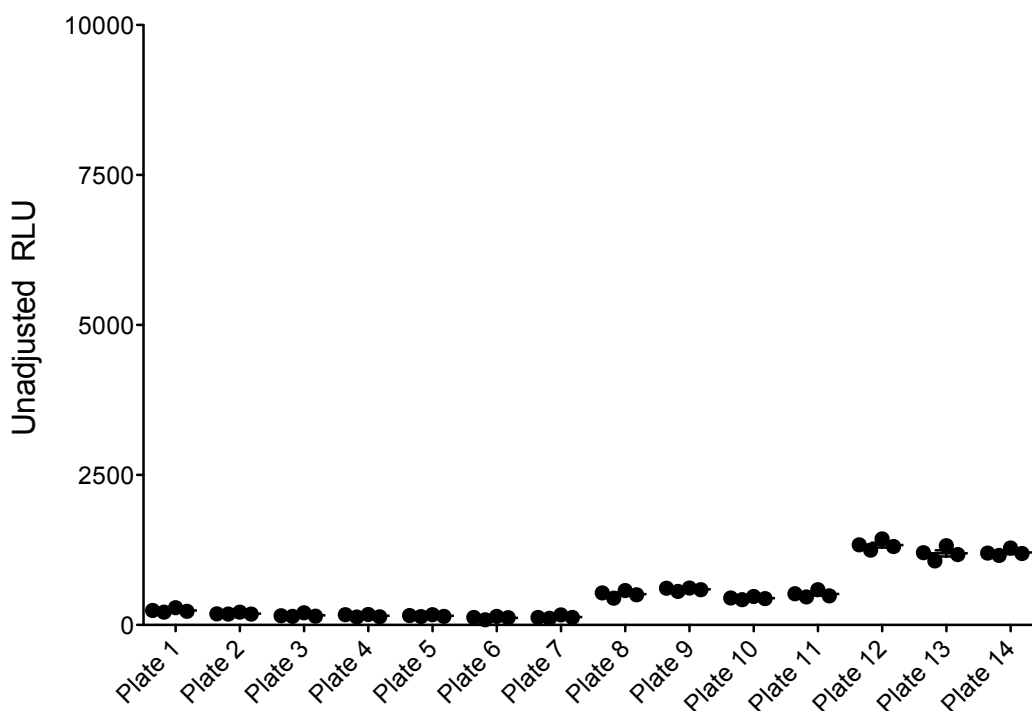
Table 6-4 Antagonist DMSO Control Values

Laboratory	Mean and Range of DMSO Control RLU Values	Mean and Range of CV (%)	N
XDS	2230 (132–6860)	9 (1–52)	79
ECVAM	3622 (1352–7333)	9 (1–37)	62
Hiyoshi	4030 (1625–8451)	6 (1–20)	53
All Laboratories	3299 (132–8451)	8 (1–52)	194

Abbreviations: CV = coefficient of variation; DMSO = dimethyl sulfoxide; ECVAM = European Centre for the Validation of Alternative Methods; N = number of plates that passed acceptance criteria; RLU = relative light unit; XDS = Xenobiotic Detection Systems, Inc.

Figures 6-6 through 6-8 show the within-plate agonist DMSO control RLU values for Phase 1 of the validation study as examples of the low variability for this parameter. As discussed above, within-plate CVs were low throughout the validation study.

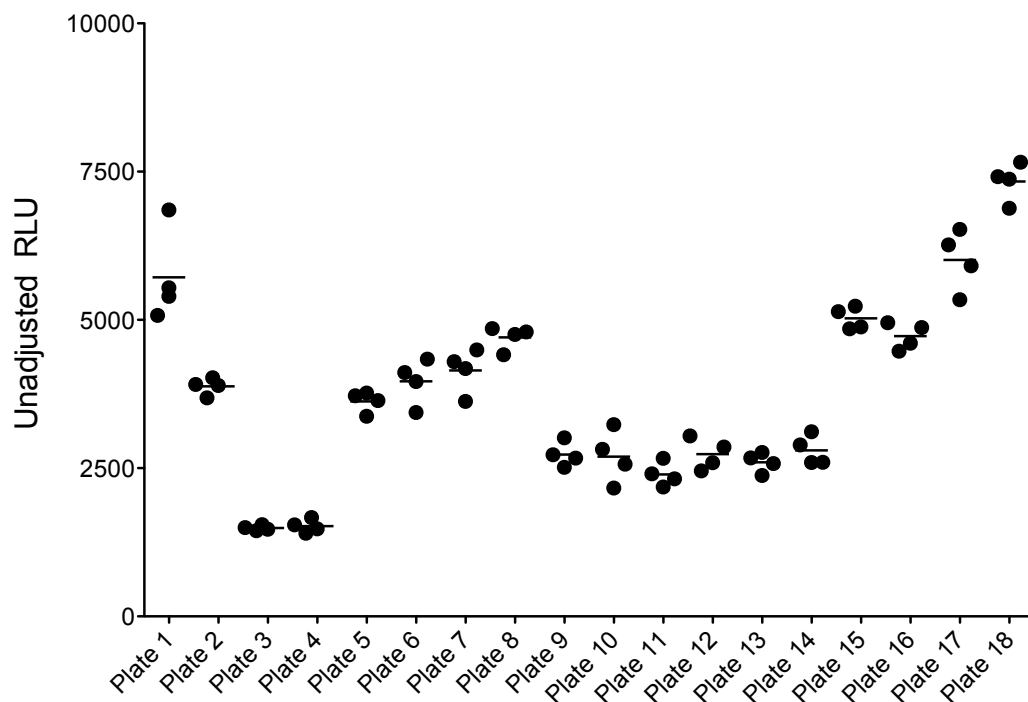
Figure 6-6 Antagonist DMSO Control Within-Plate RLU Values During Phase 1 at XDS



Abbreviations: DMSO= dimethyl sulfoxide; RLU = relative light unit; XDS = Xenobiotic Detection Systems, Inc. Each point represents the non-normalized DMSO value for a single well in a 96-well plate.

Within-plate DMSO variance at XDS during Phase 1 was fairly low, with coefficients of variation ranging from 3% to 18%.

Figure 6-7 Antagonist DMSO Control Within-Plate RLU Values During Phase 1 at ECVAM

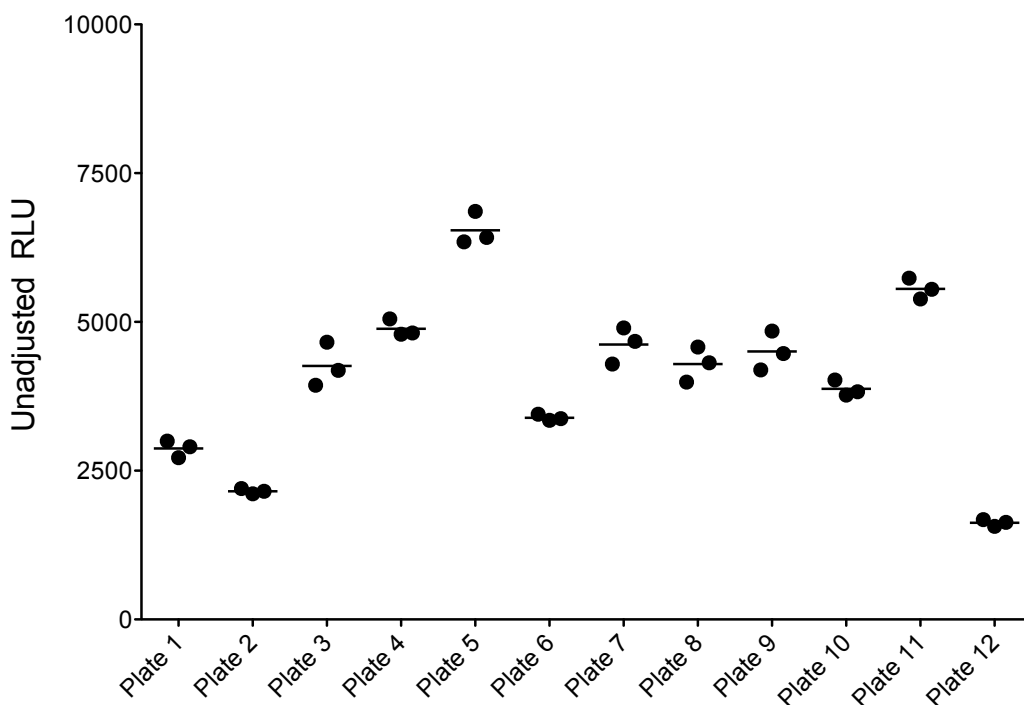


Abbreviations: DMSO = dimethyl sulfoxide; ECVAM = European Centre for the Validation of Alternative Methods;
RLU = relative light unit.

Each point represents the non-normalized DMSO value for a single well in a 96-well plate.

Within-plate DMSO variance at ECVAM during Phase 1 was fairly low, with coefficients of variation ranging from 3% to 17%.

Figure 6-8 Antagonist DMSO Control Within-Plate RLU Values During Phase 1 at Hiyoshi



Abbreviations: DMSO = dimethyl sulfoxide; RLU = relative light unit.

Each point represents the non-normalized DMSO value for a single well in a 96-well plate.

Within-plate DMSO variance at Hiyoshi during Phase 1 was fairly low, with coefficients of variation ranging from 3% to 9%.

6.1.5 Antagonist E2 Control

Normalized and adjusted antagonist E2 control RLU values were used as acceptance criteria throughout the validation study. The mean, SD, and CV values calculated for the E2 control RLU value from all antagonist test plates that passed acceptance criteria are provided in **Table 6-5**. Mean E2 control RLU values ranged from 5793 at Hiyoshi to 9246 at ECVAM. Variability was low, with associated CV values ranging from 9% at ECVAM to 19% at XDS.

Table 6-5 Antagonist E2 Control Values

Laboratory	Mean RLU	SD	CV (%)	N
XDS	7524	1443	19	79
ECVAM	9246	805	9	62
Hiyoshi	5793	791	14	53

Abbreviations: CV = coefficient of variation; ECVAM = European Centre for the Validation of Alternative Methods; N = number of plates that passed acceptance criteria; RLU = relative light unit; SD = standard deviation; XDS = Xenobiotic Detection Systems, Inc.

6.1.6 Antagonist Raloxifene Reference Standard IC₅₀ and Flavone Control Values

Although Ral reference standard IC₅₀ and flavone control RLU values were not used for plate acceptance after Phase 2a of the validation study (see **Section 2.7.2**), these values were collected throughout the study for information purposes. The means and SDs for these parameters from all plates that passed acceptance criteria are provided in **Table 6-6**.

Table 6-6 Antagonist Raloxifene IC₅₀ and Flavone Control Values

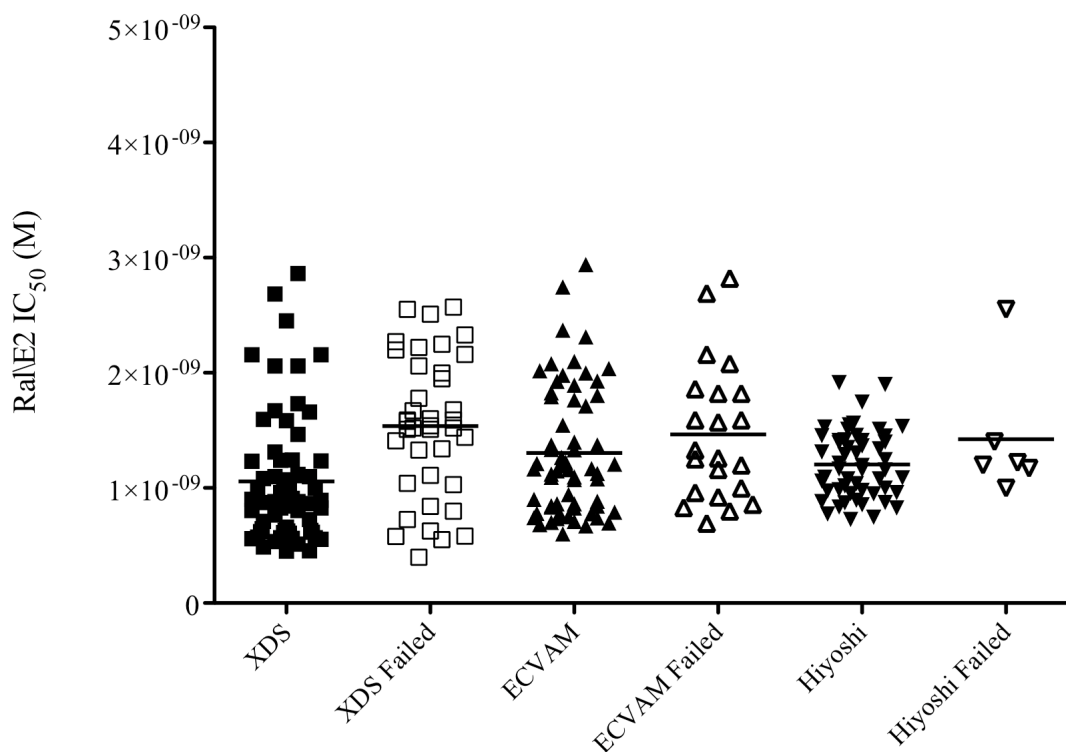
Laboratory	Mean	SD	N
Raloxifene Reference Standard IC₅₀ (M)			
XDS	1.1×10^{-9}	5.6×10^{-10}	79
ECVAM	1.3×10^{-9}	5.6×10^{-10}	62
Hiyoshi	1.2×10^{-9}	2.9×10^{-10}	53
Flavone (RLU)			
XDS	3774	1366	79
ECVAM	599	468	62
Hiyoshi	873	772	53

Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; IC₅₀ = half-maximal inhibitory concentration; M = molar; N = number of plates that passed acceptance criteria; RLU = relative light unit; SD = standard deviation; XDS = Xenobiotic Detection Systems, Inc.

As shown in **Table 6-6**, mean Ral reference standard IC₅₀ values ranged from 1.1×10^{-9} to 1.3×10^{-9} M. Mean flavone control RLU values ranged from 599 at ECVAM to 3774 at XDS.

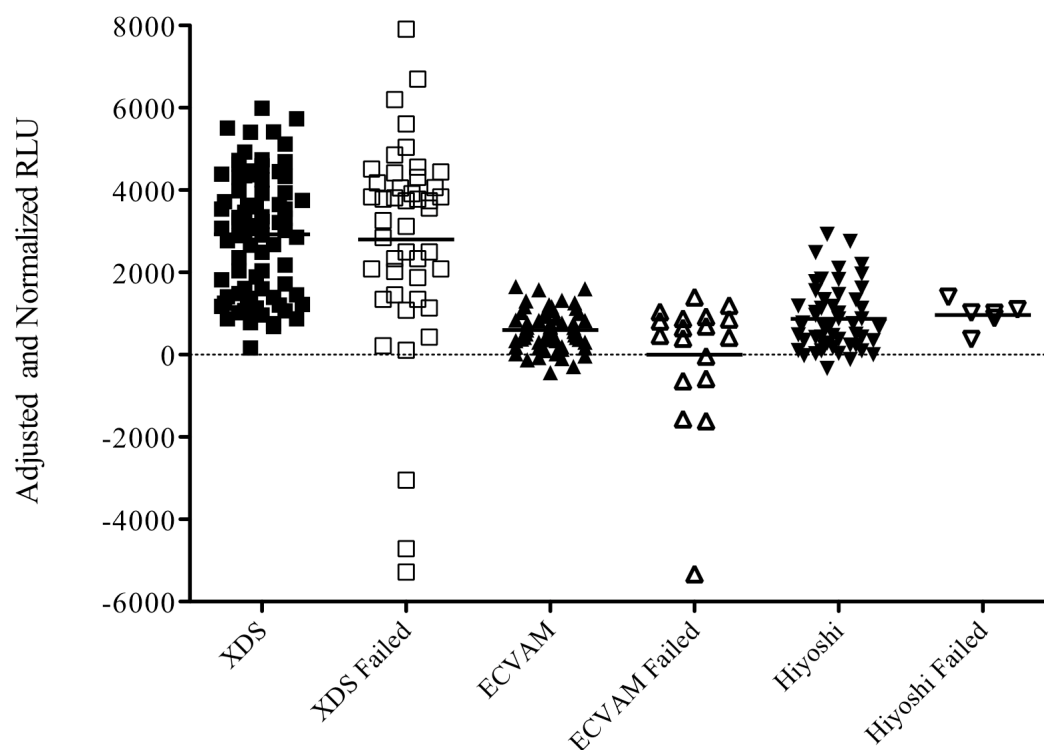
Ral reference standard IC₅₀, flavone control, and E2 control RLU values for all plates tested during the validation study are presented in **Figures 6-9** through **6-11**. The laboratories were relatively consistent when data from only acceptable plates were considered. These data also indicate that the variability of each parameter is generally higher when considering only values obtained from plates that failed one or more acceptance criteria. Additionally, any outlier values among the parameters evaluated were associated with these failed plates.

Figure 6-9 Antagonist Raloxifene Reference Standard IC₅₀ Values



Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; IC₅₀ = half-maximal inhibitory concentration; M = molar; Ral = raloxifene; XDS = Xenobiotic Detection Systems, Inc.

Each point represents a single plate.

Figure 6-10 Antagonist Flavone Control Values

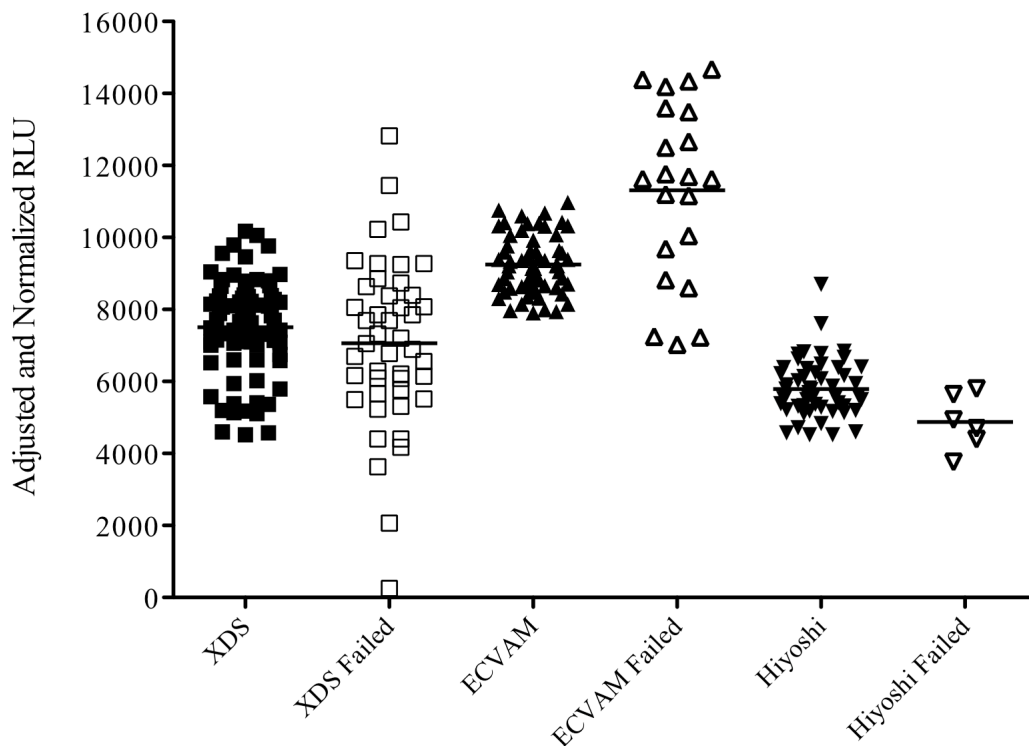
Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; RLU = relative light unit;
 XDS = Xenobiotic Detection Systems, Inc.

Each point represents a single plate.

Flavone control values from two experiments that passed acceptance criteria at XDS were excluded from the graph (237690 and 23164) to minimize scale distortion.

Flavone control values from four experiments that failed acceptance criteria at XDS were excluded from the graph (22676, -21568, -16714, and -8081) to minimize scale distortion.

Figure 6-11 Antagonist E2 Control Values



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

Each point represents a single plate.

E2 control values from two experiments that failed acceptance criteria at XDS were excluded from the graph (41227 and -3995) to minimize scale distortion.

A flavone control value from one experiment that failed acceptance criteria at ECVAM was excluded from the graph (20345) to minimize scale distortion.

6.1.7 Intralaboratory Reproducibility of Phase 2 Antagonist Reference Substances

As described in **Section 2.0**, test substances were classified as positive or negative for antagonist activity based on a specific set of criteria. The resulting classifications for each of the 12 substances that were tested at least three times at each laboratory were used to evaluate the extent of intralaboratory agreement (see **Table 6-7**). Although the classifications for some of the test substances differed among the laboratories, there was 100% agreement within each laboratory for each of the three repeat tests. No “inadequate” data were generated at any laboratory during Phase 2 of the validation study.

Table 6-7 Intralaboratory Agreement for Multiple Testing of 12 Phase 2 Antagonist Substances Tested Independently Three Times at Each Laboratory

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.

6.2 Interlaboratory Reproducibility

Similar to the intralaboratory analyses described in **Sections 6.1.3** and **6.1.7**, the classifications for each of the substances that were tested for agonist and antagonist activity during Phases 2 and 3 were also used to evaluate the extent of interlaboratory agreement as indicators of reproducibility among the laboratories.

6.2.1 Interlaboratory Reproducibility of Phase 2 Reference Substances

For each of the 12 substances that were tested at least three times for agonist and antagonist activity during Phase 2, agreement among the three laboratories was determined based on the consensus classification assigned by each laboratory for each of the 12 substances. (See **Tables 4-13** and **4-14** for agonist and antagonist results, respectively.) As previously noted, no “inadequate” data were generated at any laboratory during Phase 2 of the validation study.

As shown in **Table 6-8**, all three laboratories classified the same eight of twelve (67%) substances as agonists (positive). Among the remaining four substances, one (flavone) was identified as positive by 2/3 laboratories (ECVAM and Hiyoshi) but negative at XDS. Although the starting concentrations for flavone were identical at all three laboratories (100 µg/mL), all three tests at XDS were uniformly negative and there was no increasing concentration response noted. The other three substances that were discordant among the laboratories (atrazine, corticosterone, and vinclozolin) were identified as negative by 2/3 laboratories (XDS and Hiyoshi) but positive at ECVAM. Note that all three substances appeared to be negative for agonist activity during range finder testing at ECVAM, but all three were uniformly positive when comprehensively tested. Therefore, the positive agonist results observed for atrazine, corticosterone, and vinclozolin during comprehensive testing at ECVAM may be due to contamination of stocks after range finder testing.

Table 6-8 Interlaboratory Agreement for Phase 2 Test Substances

Results Among Laboratories	Agonist Testing	Antagonist Testing
Agreement Among Laboratories	8/12 (67%)	12/12 (100%)
+++	8/12	2/12
---	0/12	10/12
Discordance Among Laboratories	4/12 (33%)	0/12 (0%)
++-	1/12	0/12
+--	3/12	0/12

+ 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.

+- indicates that a test substance was classified as positive in two of three laboratories. The substance was classified as negative in the third laboratory.

+-- indicates that the test substance was classified as positive in one of three laboratories.

Among the substances tested for antagonist activity, there was 100% agreement among the three laboratories for all 12 substances. Two of these substances (dibenzo[*a,h*]anthracene and tamoxifen) were positive in all three laboratories. The other 10 substances were negative in all three laboratories (see **Table 6-8**).

6.2.2 Interlaboratory Reproducibility of Phase 3 Agonist Reference Substances

The classifications for each of the 41 substances that were tested once for agonist activity at all three laboratories during Phase 3 were also used to evaluate the extent of interlaboratory agreement. Unlike Phase 2, some of the substances tested in Phase 3 produced results that were considered inadequate (i.e., substances failed to meet the decision criteria for either a positive or negative response as defined in **Section 2.7.1**). While such results could not be used in the evaluation of test method accuracy detailed in **Section 5.0**, these results are tabulated in this section as an indication of how often one or more laboratories produced inadequate results. However, only those substances that produced a definitive result in at least two of the three laboratories were used to assess interlaboratory reproducibility.

Of the 41 substances tested in Phase 3, 88% (36/41) produced a definitive result in at least two laboratories and were therefore used for the assessment of reproducibility. A definitive result (i.e., determination of a positive or negative response) was not determined for the remaining 12% of substances. (In these cases, testing produced inadequate results for these substances in at least two laboratories, so the results were not used to assess interlaboratory reproducibility, as noted above.) Among the remaining 36 substances, the three laboratories agreed on 83% (30/36) of the substances tested for agonist activity (see **Table 6-9**). Of the 30 substances that had 100% agreement across laboratories, 20 were positive for ER agonist activity and 10 were negative for ER agonist activity. There was discordance among the laboratories for the remaining six substances, as indicated in the lower portion of **Table 6-9**. Three of these substances (2-*sec*-butylphenol, dicofol, and fluoranthene) were positive in 2/3 laboratories (XDS and Hiyoshi) but negative at ECVAM. The other three substances (4-androstenedione, clomiphene citrate, and resveratrol) were discordant between the two

laboratories that produced a definitive result. That is, a negative result was produced in one laboratory, a positive result in another laboratory, and an inadequate result in the third laboratory.

The discordance among the laboratories for at least four of the six substances listed above (4-androstenedione, 2-*sec*-butylphenol, fluoranthene, and resveratrol) appears to have resulted from differences in the concentration selected for comprehensive testing by the discordant laboratory. As detailed in **Section 2.0**, the starting concentrations for comprehensive testing were chosen based on data from range finder tests. The highest dose used for range finder tests is directly related to the highest soluble concentration. For one of these four substances (fluoranthene), the discordance among laboratories appears to be due to differing interpretations of test substance solubility, where the highest concentration used for comprehensive testing at ECVAM was at least an order of magnitude lower than the highest concentration selected at XDS or Hiyoshi (see **Figure 6-12**). For the remaining three substances (androstenedione, 2-*sec*-butylphenol, and resveratrol), the differences in starting concentrations for comprehensive testing appear to have resulted from incorrect interpretation of data during range finder experiments (see **Figure 6-13** as an example).

The discordance among the laboratories for the remaining two substances (clomiphene citrate and dicofol) was not based on either differences in solubility or interpretation of range finder results. Clomiphene citrate was clearly positive at Hiyoshi and clearly negative at ECVAM when comprehensively tested over the same concentration range. Although dicofol was positive when tested at Hiyoshi using a starting concentration an order of magnitude higher than those used by XDS and ECVAM, it was clearly positive at XDS and clearly negative at ECVAM when comprehensively tested over the same concentration range.

Table 6-9 Interlaboratory Agreement for Phase 3 Substances Tested Once at Each Laboratory

Results Among Laboratories	Agonist Testing	Antagonist Testing
Agreement Among Laboratories	30/36 (83%)	38/41 (93%)
+++	18/36	2/41
--- ^a	4/36	33/41
++I	2/36	1/41
--I	6/36	2/41
Discordance Among Laboratories	6/36 (17%)	3/41 (7%)
++-	3/36	0/41
+--	0/36	1/41
+--I	3/36	2/41

Abbreviations: I = inadequate data.

Only those substances that produced a definitive result in at least two of the three laboratories were used in this evaluation.

Five 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.

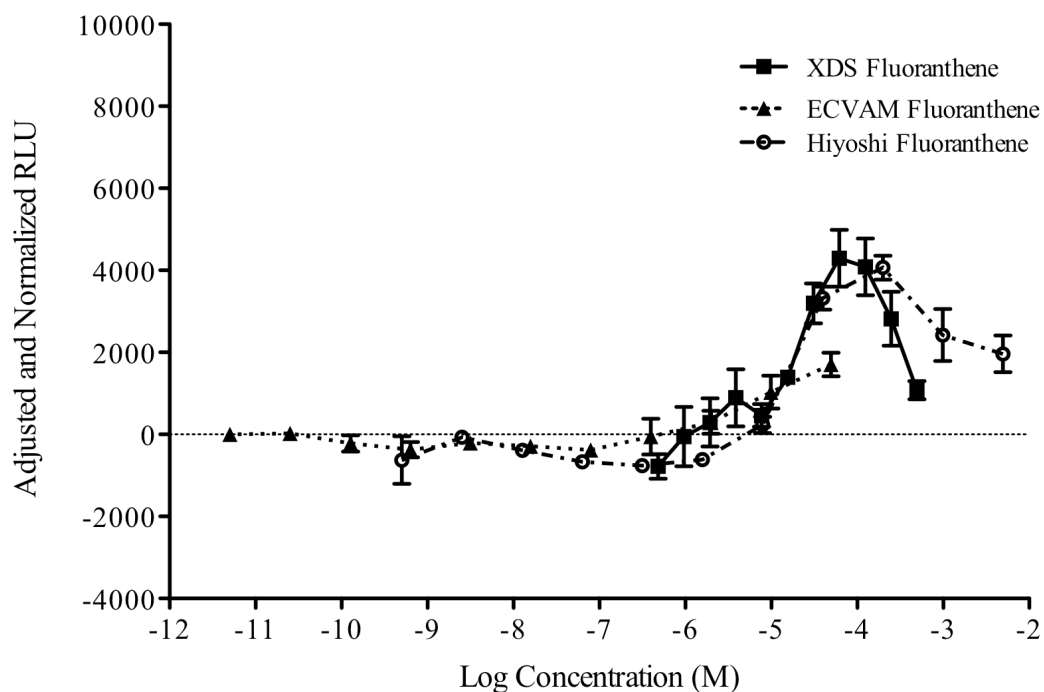
++I indicates that the substance was classified as positive at two of three laboratories but had inadequate data in the third.

--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.

^a Includes one substance (phenobarbital) that was tested in only two laboratories (XDS and ECVAM, see **Section 3.0**).

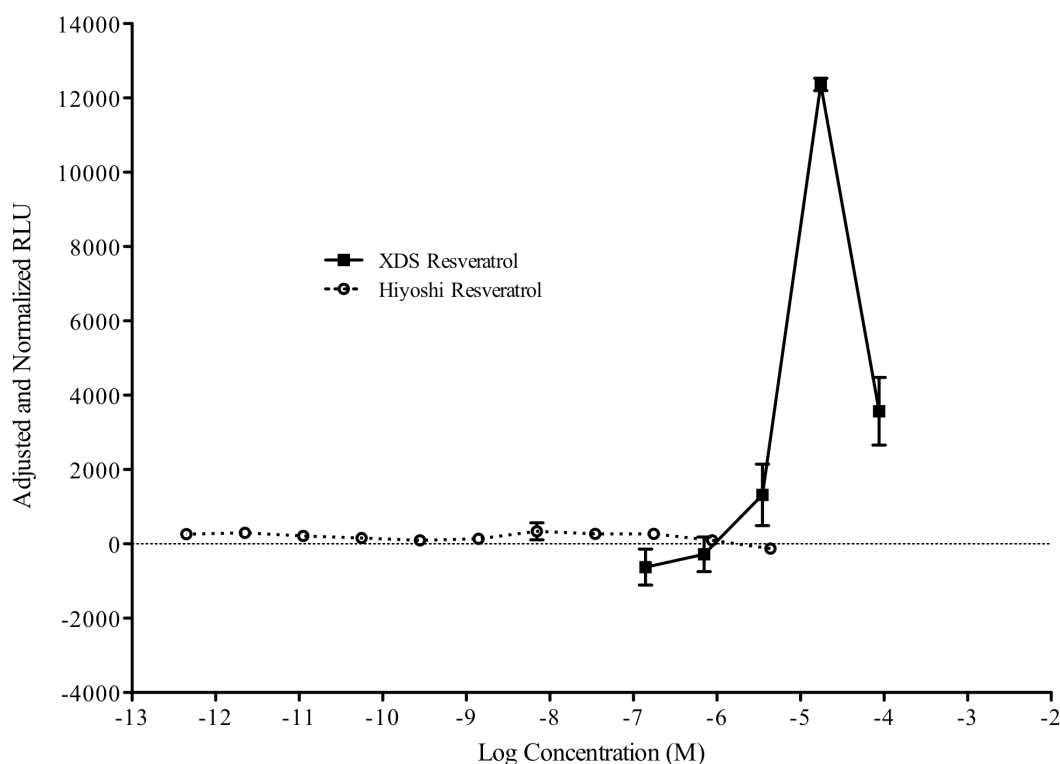
Figure 6-12 Fluoranthene Results at All Three Laboratories: Impact of Differences in Solubility on Comprehensive Test Results



Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; M = molar; RLU = relative light unit; XDS = Xenobiotic Detection Systems, Inc.

Each point represents the mean adjusted and normalized RLU value and SD from triplicate wells.

Figure 6-13 Resveratrol Results at XDS and Hiyoshi: Impact of Selecting the Incorrect Starting Concentration Based on Range Finder Results



Abbreviations: M = molar; RLU = relative light unit; XDS = Xenobiotic Detection Systems, Inc.

Each point represents the mean adjusted and normalized RLU value and SD from triplicate wells. Results for resveratrol at ECVAM were considered inadequate and are therefore not included here.

6.2.3 Interlaboratory Reproducibility of Phase 3 Antagonist Reference Substances

The classifications for each of the 41 substances that were tested once for antagonist activity at all three laboratories during Phase 3 were also used to evaluate the extent of interlaboratory agreement. Similar to the Phase 3 agonist test results, some of the substances tested in Phase 3 for antagonist activity produced results that were considered inadequate (i.e., substances failed to meet the decision criteria for either a positive or negative response as defined in **Section 2.7.2**). However, unlike the agonist test results, no substances tested for antagonist activity produced inadequate results in more than one laboratory. Therefore, all 41 Phase 3 substances tested for antagonist activity were included in the reproducibility assessment.

The three laboratories agreed on 93% (38/41) of the substances tested for antagonist activity. Most of these substances (85% [35/41]) were identified as negative for antagonist activity; three substances were positive for antagonist activity. There was discordance among the laboratories for the remaining three substances. One of these substances (diethylstilbestrol) was negative in 2/3 laboratories (XDS and ECVAM) but positive in one laboratory (Hiyoshi). The other two substances (clomiphene citrate and 17 α -estradiol) were discordant between the two laboratories that produced a definitive result (i.e., a negative result produced in one laboratory, a positive result in another laboratory, and an inadequate result in the third laboratory). It does not appear that any of these three discordant classifications can be explained by differences in solubility or interpretation of the range finder data.

If only those substances that produced a definitive result in all three laboratories are considered ($n = 36$), there was 100% agreement for 97% (35/36) of the substances tested. As mentioned previously, substances with inadequate data would be retested under the revised testing protocol, and conclusive results would therefore be expected for all test substances. Consequently, the high degree of intralaboratory reproducibility seen when all laboratories produce conclusive results is indicative of the level of performance expected using the revised protocol (**Annexes E and F**).

7.0 BG1Luc ER TA Data Quality

Good Laboratory Practice (GLP) guidelines are nationally and internationally recognized rules designed to ensure the quality and validity of laboratory data and records. To ensure the integrity, reliability, and accountability of a study, GLPs provide a standardized approach by which to report and archive laboratory data and records, and to prepare compliant test protocols (EPA 2006b, 2006a; FDA 2009; OECD 1998; Weinberg 2003). This section describes the extent to which the participating laboratories (XDS, ECVAM, and Hiyoshi) adhered to these guidelines during the validation study and the effect (if any) of any deviations in the quality of the data. This section also details how often each laboratory failed to generate data that met the plate acceptance criteria (see **Section 4.0**), necessitating repeat testing during the validation study.

7.1 Compliance with GLP Regulations

The BG1Luc ER TA validation study was conducted according to GLP guidelines at XDS and ECVAM, but not at Hiyoshi, which does not have a formal GLP program. However, prior to initiating the validation study, Hiyoshi provided a guidance document that outlined the quality control (QC) procedures that they would follow throughout the study. The guidance document is based on the OECD principles of GLP (see **Annex H2**). In addition, Hiyoshi follows the QC and quality assurance (QA) procedures included in the International Organization for Standardization (ISO) 9000 standards, which describe a series of internationally accepted good quality management practices that are applicable to laboratory testing (ISO 2000). However, ISO standards do not dictate the methods by which those requirements must be met. ISO 9001:2000, which was used by Hiyoshi, defines and describes requirements for the following standards:

- Quality Management System — requires written quality standards and a control system for all documents and records
- Management Responsibility — assigns the responsibility for all facets of the quality system, from creation to improvement, to the organization's senior management and requires a regular, documented review of the quality program
- Resource Management — requires that personnel be competent enough to provide quality work and that all facilities, equipment, supporting services, and training programs be sufficient to ensure quality product
- Product Realization — requires clear documentation on how design decisions are made, reviewed, validated, and controlled
- Measurement, Analysis, and Improvement — requires that all facets of the company be monitored, reviewed, and, when necessary, corrected

7.2 QA Audit Results

GLP compliance in each participating laboratory was determined by an independent QA review of various aspects of the study, including the following:

- Review of protocols and laboratory standard operating procedures (SOPs)
- Review of laboratory operations
- Review of data
- Review of the final report for each testing phase

All laboratory reports included QA statements that addressed whether the test methods and results accurately followed the test protocols and whether study reports accurately reflected the raw data produced during the study. The study project coordinator and assistant project coordinator also served as secondary QA reviewers for all data and information provided by study directors and/or study technical leads. QA review dates for each participating laboratory are provided in **Table 7-1**.

Table 7-1 Quality Assurance Review Dates

Laboratory	Phase	Review During Testing	Report Review
XDS	1	May–July 2007	March 2008
	2a	April 2008	November 2008
	2b	September 2008	November 2008
	3	October 2009	July 2010
	4	November 2009	July 2010
ECVAM	1	November 2007–January 2008	March 2008
	2a	October 2008	November 2008
	2b	NR	January 2010
	3	NR	January 2010
Hiyoshi	1	July–October 2007	February 2008
	2a	April 2008	November 2008
	2b	September 2010	February 2010
	3	September 2010	February 2010

Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; NR = not reviewed;
XDS = Xenobiotic Detection Systems, Inc.

The QA statements provided in final reports for all validation study phases completed at ECVAM and Hiyoshi (i.e., Phases 1, 2a, 2b, and 3) indicated that (1) the procedures used to conduct validation study testing followed the test method protocols and (2) study reports accurately reflected the raw data produced during the study. Phases 1, 2a, 3, and 4 at XDS also met these criteria. However, the XDS Phase 2b study report indicated that BG1Luc ER TA antagonist protocol procedures for assessing cell viability were not used in a consistent manner for five (apigenin, atrazine, *o,p'*-DDT, genistein, and resveratrol) of the eight antagonist substances tested. Therefore, testing results from these five Phase 2b substances were not used to evaluate antagonist activity. The validation study project coordinator reviewed cell viability assessment procedures with the XDS study director and quality assurance officer. Apigenin, atrazine, *o,p'*-DDT, genistein, and resveratrol were subsequently retested at XDS. These repeat testing results were then used to evaluate antagonist activity (see **Section 4.0, Table 4-13**).

7.3 Test Plate Failure Rates

As described in **Sections 2.7.1** and **2.7.2**, plate acceptance criteria were established based on results generated in reference standards and control wells. Failures due to results outside of the acceptable range could indicate poor-quality data. However, some of the plate failures may have been due more to overly stringent criteria that were established prior to testing of coded substances in Phase 2, as described in the following sections.

7.3.1 Phase 2a

Following Phase 2a of the validation study, NICEATM evaluated the failure rates of plates used during Phase 2a agonist and antagonist testing. The percentages of agonist and antagonist test plates that failed to meet acceptance criteria across the participating laboratories were 61% (33/54) and 38% (13/34), respectively:

- At XDS, 53% (8/15) of agonist plates and 43% (6/14) of antagonist plates did not meet acceptance criteria.
- At ECVAM, 80% (24/30) of agonist plates and 50% (7/14) of antagonist plates failed to meet acceptance criteria.
- At Hiyoshi, 11% (1/9) of agonist plates and 0% (0/6) of antagonist plates failed to meet acceptance criteria.

Based on these high failure rates, the plate acceptance criteria were reconsidered to determine if changes to these criteria could reduce the failure rates without compromising the ability of the test method to detect and quantify test substance agonist or antagonist activity. The test plate acceptance criteria that were considered for modification were (1) agonist E2 EC₅₀ and Met RLU control values and (2) antagonist Ral IC₅₀ and flavone control RLU values. Acceptance criteria based on the DMSO control RLU, agonist E2 reference standard fold induction, and antagonist Ral reference standard fold reduction values were not considered for modification because they are used to monitor background activity (i.e., vehicle control) and reference standard performance (i.e., positive control). The antagonist E2 control acceptance criterion was not considered for modification because it is required for determining test substance antagonist activity.

A comparison was made between qualitative (i.e., positive or negative classification) and quantitative (i.e., EC/IC₅₀ values) outcomes for test plates that met all acceptance criteria and those that failed to meet one or more criteria (see **Section 2.7** for Phase 2a acceptance criteria). The results of the qualitative evaluation of the relationship between agonist and antagonist test plate failure rates and acceptance criteria for these parameters are provided in **Tables 7-2** and **7-3**, respectively. The qualitative evaluation compared the overall ER TA activity classification of agonist and antagonist test substances for plates that passed and failed acceptance criteria. Results indicate that the ER TA activities (overall positive or negative classification) of substances tested on agonist plates that failed EC₅₀ and/or Met control acceptance criteria were equivalent to the ER TA activities for plates that passed acceptance criteria. Antagonist plates that failed IC₅₀ and/or flavone control acceptance criteria were equivalent to the ER TA activities for plates that passed acceptance criteria.

Table 7-2 Phase 2a Test Substance ER TA Agonist Activity for Plates That Passed or Failed Acceptance Criteria

Agonist Test Substance	Laboratory	Passed All Acceptance Criteria ^a	Failed E2 EC ₅₀ Only	Failed Met Only	Failed Both E2 EC ₅₀ and Met
Bisphenol A	XDS	POS (3/3)	POS (4/4)	NA	NA
	ECVAM	POS (3/3)	POS (7/7)	POS (3/3)	NA
	Hiyoshi	POS (3/3)	NA	POS (1/1)	NA
Bisphenol B	XDS	POS (3/3)	POS (4/4)	NA	NA
	ECVAM	POS (3/3)	POS (4/4)	NA	POS (2/2)
	Hiyoshi	POS (3/3)	NA	POS (1/1)	NA
Corticosterone	XDS	NEG (3/3)	NEG (4/4)	NA	NA
	ECVAM	POS (3/3)	POS (5/7)	POS (3/3)	NA
	Hiyoshi	NEG (4/4)	NA	NA	NA
Diethylstilbestrol	XDS	POS (3/3)	POS (4/4)	NA	NA
	ECVAM	POS (3/3)	POS (4/4)	NA	POS (2/2)
	Hiyoshi	POS (4/4)	NA	NA	NA

Abbreviations: E2 = 17 β -estradiol; EC₅₀ = half-maximal effective concentration; ECVAM = European Centre for the Validation of Alternative Methods; Met = methoxychlor; NA = not applicable; NEG = negative; POS = positive; XDS = Xenobiotic Detection Systems, Inc.

Agonist activity based on initial classification criteria as defined in **Section 2.7.1**.

^a Numbers in parentheses represent test results (POS or NEG) over the total number of test plates.

Table 7-3 Phase 2a Test Substance ER TA Antagonist Activity for Plates That Passed or Failed Acceptance Criteria

Antagonist Test Substance	Laboratory	Passed All Acceptance Criteria ^a	Failed Ral IC ₅₀ Only	Failed Flavone Control Only	Failed Both Ral IC ₅₀ and Flavone Control
Dibenzo[<i>a,h</i>]anthracene	XDS	POS (3/3)	POS (2/2)	NA	NA
	ECVAM	POS (3/3)	NA	NA	NA
	Hiyoshi	POS (3/3)	NA	NA	NA
<i>p</i> -n-Nonylphenol	XDS	NEG (3/3)	NEG (3/3)	NA	NA
	ECVAM	POS (3/3)	NA	NA	NA
	Hiyoshi	POS (3/3)	NA	NA	NA

Antagonist Test Substance	Laboratory	Passed All Acceptance Criteria ^a	Failed Ral IC ₅₀ Only	Failed Flavone Control Only	Failed Both Ral IC ₅₀ and Flavone Control
Progesterone	XDS	POS (3/3)	POS (2/3)	NA	NA
	ECVAM	POS (3/3)	NA	NA	NA
	Hiyoshi	POS (3/3)	NA	NA	NA
Tamoxifen	XDS	POS (3/3)	POS (3/3)	NA	NA
	ECVAM	POS (3/3)	NA	(1/2)	NA
	Hiyoshi	POS (3/3)	NA	NA	NA

Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; IC₅₀ = half-maximal inhibitory concentration; NA = not applicable; NEG = negative; POS = positive; Ral = raloxifene HCl; XDS = Xenobiotic Detection Systems, Inc.

Antagonist activity based on initial classification criteria as defined in **Section 2.7.2**.

^a Numbers in parentheses represent test results (POS or NEG) over the total number of test plates.

Table 7-4 provides the quantitative evaluation of the relationship between agonist and antagonist test plate failure rates and acceptance criteria. Agonist plates that passed all acceptance criteria are compared to those that failed the E2 EC₅₀ and Met RLU control value acceptance criteria. Antagonist plates that passed all acceptance criteria are compared to those that failed the Ral IC₅₀ and flavone control RLU value acceptance criteria. The quantitative evaluation compared EC₅₀ values that could be calculated for bisphenol A, bisphenol B, and diethylstilbestrol at XDS and ECVAM, and the IC₅₀ values that could be calculated for tamoxifen at XDS for plates that passed and failed acceptance criteria. Results indicate that agonist substance EC₅₀ values from plates that failed EC₅₀ and/or methoxychlor control acceptance criteria and tamoxifen IC₅₀ values from plates that failed IC₅₀ and/or flavone control acceptance criteria were not significantly different from plates that passed acceptance criteria ($p > 0.05$).

Table 7-4 Comparison of Phase 2a Test Substance EC₅₀/IC₅₀ Values for Plates That Passed or Failed Acceptance Criteria

Laboratory and Substance Evaluated	Agonist Plates That Passed All Acceptance Criteria			Agonist Plates That Did Not Pass E2 EC ₅₀ and/or Methoxychlor Acceptance Criteria			p Value ^a
	N	Mean EC ₅₀ Value	SD	N	Mean EC ₅₀ Value	SD	
XDS/BPA	3	8.8×10^{-2}	7.2×10^{-3}	4	9.9×10^{-2}	1.4×10^{-2}	0.40
ECVAM/BPA	3	1.9×10^{-1}	7.6×10^{-3}	10	1.6×10^{-1}	5.6×10^{-2}	0.16
XDS/BPB	3	3.9×10^{-2}	6.0×10^{-3}	4	4.3×10^{-2}	1.1×10^{-2}	0.63
ECVAM/BPB	3	4.2×10^{-2}	1.3×10^{-2}	4	7.5×10^{-2}	1.7×10^{-2}	0.06
XDS/DES	4	1.4×10^{-5}	5.0×10^{-6}	4	2.6×10^{-5}	1.1×10^{-5}	0.20

Laboratory and Substance Evaluated	Antagonist Plates That Passed All Acceptance Criteria			Antagonist Plates That Did Not Pass Ral/E2 IC ₅₀ and/or Flavone Acceptance Criteria			p Value ^a
	N	Mean IC ₅₀ Value	SD	N	Mean IC ₅₀ Value	SD	
XDS/TAM	4	1.5 x 10 ⁻¹	5.7 x 10 ⁻²	3	3.1 x 10 ⁻¹	8.8 x 10 ⁻²	0.11

Abbreviations: BPA = bisphenol A; BPB = bisphenol B; DES = diethylstilbestrol; E2 = 17β-estradiol; EC₅₀ = half-maximal effective concentration; ECVAM = European Centre for the Validation of Alternative Methods; IC₅₀ = half maximal inhibitory concentration; N = number of plates; Ral = raloxifene HCl; SD = standard deviation; TAM = tamoxifen; XDS = Xenobiotic Detection Systems, Inc.

Values are expressed in EC₅₀ values (μg/mL) except for TAM, which is expressed in IC₅₀ values (μg/mL).

^a p > 0.05 indicates that EC₅₀ or IC₅₀ values are not significantly different.

Based on this evaluation, it was determined that test plate acceptance criteria based on agonist E2 EC₅₀ and Met RLU control values could be eliminated without compromising the ability of the test method to detect and quantify test substance agonist or antagonist activity. The same was determined for antagonist Ral IC₅₀ and flavone control RLU values. The modified acceptance criteria for agonist and antagonist comprehensive testing are provided in **Sections 2.7.1** and **2.7.2**, respectively, and were used for all plates tested in the remainder of the validation study (i.e., Phases 2b, 3, and 4).

7.3.2 Phases 2b, 3, and 4 Failure Rates

The plate failure rates for the remaining phases of the study are provided in **Tables 7-5** and **7-6**. Results indicate that the modified acceptance criteria based on Phase 2a results significantly reduced the failure rates of agonist test plates in Phases 2b, 3, and 4 (≤ 27%) compared to the Phase 2a agonist test plate failure rate (61%). The failure rate of Phase 2b antagonist test plates (14%) was also significantly reduced compared to the Phase 2a antagonist test plate failure rate (38%). During Phases 3 and 4, the failure rates for antagonist test plates were only marginally decreased (36% and 35%, respectively).

Table 7-5 Test Plate Failure Rates for Agonists: Phases 2b–4

Phase	Laboratory	% of Plates That Failed Acceptance Criteria ^a
2b	XDS	0% (0/13)
	ECVAM	25% (4/16)
	Hiyoshi	19% (3/16)
	Total	16% (7/45)
3	XDS	26% (12/47)
	ECVAM	29% (10/35)
	Hiyoshi	0% (0/34)
	Total	19% (22/116)
4	XDS	27% (11/41)

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

^a Numbers in parentheses represent the number of test plates that failed acceptance criteria over the total number of plates tested.

Table 7-6 Test Plate Failure Rates for Antagonists: Phases 2b–4

Phase	Laboratory	% of Plates That Failed Acceptance Criteria
2b	XDS	0% (0/12)
	ECVAM	33% (6/18)
	Hiyoshi	0% (0/14)
	Total	14% (6/44)
3	XDS	47% (28/59)
	ECVAM	31% (11/36)
	Hiyoshi	13% (3/24)
	Total	36% (43/119)
4	XDS	35% (8/23)

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

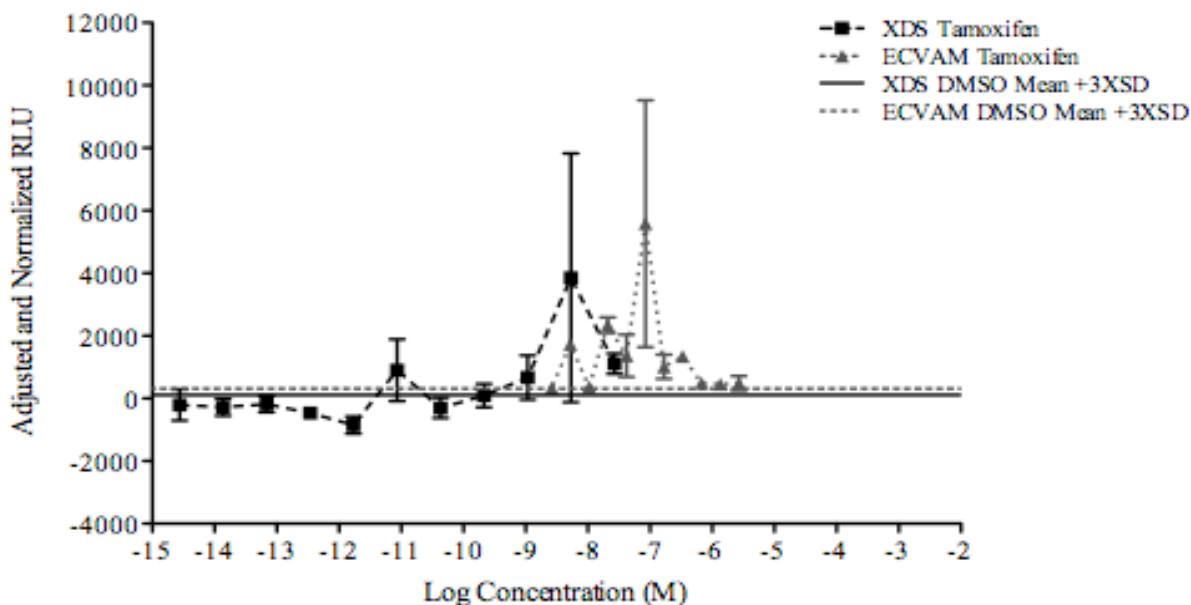
^a Numbers in parentheses represent the number of test plates that failed acceptance criteria over the total number of plates tested.

7.4 Inadequate Results

As described in **Section 2.0**, test substances were classified as positive, negative, or inadequate based on updated test method decision criteria. Inadequate data were identified as such based on those substances that failed to meet the decision criteria for either a positive or negative response as defined in **Sections 2.7.1** and **2.7.2**. The classification of data as “inadequate” is due to poor-quality data that could not be interpreted as valid because of major qualitative or quantitative limitations. Normally, substances with inadequate data would be retested, and conclusive results would therefore be expected for all test substances. However, because the updated classification system was developed after testing was complete, these substances were not retested.

As an example, tamoxifen test results at XDS and ECVAM failed to produce a clear concentration–response curve, and the resulting data had overlapping error bars due to one or more highly variable results (**Figure 7-1**).

Figure 7-1 Inadequate Test Results: Tamoxifen Tested at XDS and ECVAM



Abbreviations: DMSO = dimethyl sulfoxide; ECVAM= European Centre for the Validation of Alternative Methods;

M = molar; RLU = relative light unit; XDS = Xenobiotic Detection Systems, Inc.

Each point represents the mean adjusted and normalized RLU value and SD from triplicate wells.

While the actual test substance classifications based on BG1Luc ER TA results are presented in **Tables 4-13** and **4-14** (see **Section 4.0**), the frequency of inadequate data produced at each laboratory is summarized in **Table 7-7**. Inadequate test results in the agonist test method occurred from 3% (1/40) at Hiyoshi to 27% (11/41) at XDS. Antagonist testing produced far fewer inadequate results (3% to 5% of tests) but Hiyoshi again produced the fewest inadequate results.

Table 7-7 Summary of Test Results Classified as Inadequate

Phase	Laboratory	Agonist ^a	Antagonist ^a
Phase 2	XDS	0% (0/12)	0% (0/12)
	ECVAM	0% (0/12)	0% (0/12)
	Hiyoshi	0% (0/12)	0% (0/12)
Phase 3	XDS	27% (11/41)	5% (2/41)
	ECVAM	17% (7/41)	5% (2/41)
	Hiyoshi	3% (1/40)	3% (1/41)
Phase 4	XDS	16% (4/25)	4% (1/25)

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

^a Numbers in parentheses represent the number of inadequate results over the total number of substances tested.

7.5 Availability of Laboratory Notebooks or Other Records

All records are stored and archived by the participating laboratories and are available for inspection. NICEATM has all raw and reported data stored electronically, and the raw data for each test (in Microsoft Excel and GraphPad Prism files) are available upon request from NICEATM on compact disc(s). Long-term archival is available if deemed necessary. Requests can be made by mail, fax, or e-mail to Dr. William S. Stokes, NICEATM, NIEHS, P.O. Box 12233, MD EC-17, Research Triangle Park, NC, 27709, (phone) 919-541-2384, (fax) 919-541-0947, (e-mail) niceatm@niehs.nih.gov.

8.0 Other Scientific Reports

8.1 Summaries of Available Data from Studies Using the BG1Luc ER TA Test Method

This section reviews published studies that used BG1Luc4E2-based ER TA test methods to evaluate *in vitro* ER agonist or antagonist activity of a number of substances. Results for many of the substances described by Gordon et al. in 2003 and 2004 (see **Sections 8.1.3** and **8.1.4**) were also provided in the XDS submission (**Annex A**). Additionally, a separate study that compared the relative utility of qualitative and quantitative methods for determining BG-1 cell viability during the assay is described in Clark et al. (2007).

8.1.1 Rogers and Denison (2000)

Rogers and Denison (2000) describe the original development, optimization, and characterization of the BG-1 cell line, a stably transfected recombinant human ovarian cancer cell line. BG-1 cells were transfected using the pGudLuc7.0 plasmid, which contains a segment of the pGudLuc1.0 and the mouse mammary tumor viral promoter. The pGudLuc1.0 segment is hormone responsive but lacks glucocorticoid-responsive elements. The parent vector, pGudLuc7.0, was shown to be unresponsive to estrogen in BG-1 cells in the absence of EREs. After demonstration of estrogen-responsive luciferase activity in transiently transfected cells, a stably transfected, estrogen-responsive BG-1 clone was isolated and designated BG1Luc4E2. BG1Luc4E2 displayed constitutive activation of the luciferase gene under normal culture conditions, but this activity was greatly reduced when cells were grown in EFM. The estrogen-responsive induction of luciferase seen in BG1Luc4E2 cells that are grown in EFM is time and dose dependent. While maximal induction following exposure to 0.1 nM estradiol was seen at 20 hours, the minimum detection limit was between 0.1 and 1 pM estradiol. Cross-reactivity of BG1Luc4E2 cells with six other steroid hormones was also evaluated. Progesterone, testosterone, all-trans retinoic acid, and thyroid hormone did not induce luciferase activity, but dihydrotestosterone and dexamethasone produced slight induction (based on three independent experiments in which substances were considered positive for ER TA agonist activity when induction of luciferase was significantly different from control, at $p < 0.05$ as determined by a *t* test).

8.1.2 Jefferson et al. (2002)

This paper (Jefferson et al. 2002) describes a study that evaluated the ER TA activities of several phytoestrogens (biochanin A, coumestrol, daidzein, genistein, naringenin, taxifolin, zearalanol, and zearalenone) using a BG1Luc4E2-based test method. All substances except taxifolin tested positive for ER TA activity, with EC_{50} values ranging from 3.9×10^{-5} (zearalanol) to 1.2×10^0 $\mu\text{g/mL}$ (naringenin) as compared to 17β -estradiol (2.3×10^{-6} $\mu\text{g/mL}$) or diethylstilbestrol (4.9×10^{-6} $\mu\text{g/mL}$). The specific criteria used to determine negative ER TA response and the number of tests per substance were not provided. ER TA results were compared to uterotrophic bioassay results for the substances and showed agreement for all substances except daidzein and naringenin, which were weakly positive for ER TA activity (5.2×10^{-1} and 1.2×10^0 $\mu\text{g/mL}$, respectively) but negative when tested in the uterotrophic bioassay.

8.1.3 Gordon et al. (2003)

The 2003 International Dioxin Symposium (Boston, MA) presentation by Gordon et al. (2003) describes studies that evaluated the ER TA activities of 78 substances using a BG1Luc4E2-based test method. Of these substances, 29 had been previously tested in other ER TA assays that were not identified in the presented paper. The remaining 49 substances, which were classified by the presenter as environmental contaminants, had not been previously tested in ER TA assays. All substances were

tested independently at least three times, and ER TA activity was based on whether induction of luciferase was less or greater than three times the SD of the mean vehicle control value. Using these criteria, 61 substances were positive and 17 were negative for ER TA activity. (Note: A complete listing of results for individual substances was not provided. Graphical representations of concentration–response curves for 12 positive and 3 negative substances were provided as representative examples.) Results also indicated that the 29 substances previously tested were in agreement with the BG1Luc ER TA test method results, except for progesterone, which was negative in the BG1Luc ER TA test method but positive in other ER TA test methods.

8.1.4 Gordon et al. (2004)

The 2004 International Dioxin Symposium (Berlin, Germany) presentation by Gordon et al. (2004) describes studies that evaluated the ER TA activity of 13 commonly used organochlorine pesticides using a BG1Luc4E2-based test method. Each substance was tested independently at least three times, and ER TA activity was based on whether induction of luciferase was less or greater than three times the SD of the mean vehicle control value. Based on these criteria, 11 substances were positive and 2 were negative for ER TA activity. EC₅₀ values for those that tested positive for ER TA activity ranged from 1.3×10^{-6} (a-chlordane) to 1.2×10^{-5} M (2,4,5-trichlorophenoxyacetic acid) as compared to E2 (1.6×10^{-11}).

8.1.5 Gordon et al. (2005)

The 2005 International Dioxin Symposium (Toronto, Canada) presentation by Gordon et al. (2005) describes studies that evaluated the ER TA agonist activities of 10 commercially available sunscreens and eight substances commonly used as “non-active” sunscreen components (substances that are not used to protect against UV damage but rather as emulsifiers, emollients, lubricants, etc.) using a BG1Luc4E2 test method. The sunscreens and non-active sunscreen component substances were dissolved in methanol, serially diluted, and evaluated for ER TA agonist activity. Each substance was tested independently at least three times, and ER TA agonist activity was based on whether induction of luciferase was less or greater than three times the SD of the mean vehicle control value. This value was then translated into an E2 equivalent of 10 ng/g to control for differences in extraction recovery for individual substances (i.e., substances with E2 equivalents greater than 10 ng/g are considered positive for ER TA agonist activity). Nine of the 10 sunscreens tested positive for ER TA agonist activity, but only one of the eight non-active sunscreen component substances tested positive. The sunscreens that tested positive for ER TA activity had a range of 200 to 950 ng/g 17β-estradiol equivalents. The one non-active sunscreen component substance that tested positive for ER TA activity (a substance used for water resistance) had an E2 equivalent of 130 ng/g.

8.1.6 Clark et al. (2007)

The Clark et al. poster presentation (2007) from the 47th Annual Meeting of the Society of Toxicology (Charlotte, NC) describes a study that was conducted using the BG1Luc ER TA test method to determine if a qualitative method of assessing cell viability based on a visual observation was comparable to Promega Corporation’s CellTiter-Glo quantitative cell viability assay, which measures cell viability based on the generation of luminescence signal proportional to the amount of ATP in viable cells. The qualitative visual observation method is based on an assessment of cell density and morphology. The criteria for assessing and scoring cell viability are provided in **Table 8-1**.

Table 8-1 Visual Observation Scoring Table to Assess Cell Viability

Viability Score	Brief Description1
1	Normal Cell Morphology and Cell Density
2	Altered Cell Morphology and/or Small Gaps between Cells
3	Altered Cell Morphology and/or Large Gaps between Cells
4	Few (or no) Visible Cells

Comparison of the two cell viability assessment methods demonstrated that a score of 1 in the visual observation method corresponded to greater than 80% viability in the CellTiter-Glo assay. Visual observation scores of 2, 3, and 4 corresponded to 80–60%, 60–40%, and less than 40%, respectively, in the CellTiter-Glo assay. An assessment of cell viability is critical in determining whether reduction of ER TA activity is ER mediated or the result of cytotoxicity. The study showed that the visual observation method and the CellTiter-Glo assay are comparable for this assessment. Importantly, these results demonstrated that the simpler and more economical visual observation method can be used as effectively as the more complex and costly CellTiter-Glo, which requires testing on separate parallel plates.

9.0 Animal Welfare Considerations (Reduction, Refinement, and Replacement)

9.1 Reduction, Refinement, and Replacement Considerations

ICCVAM promotes the scientific validation and regulatory acceptance of new test methods that reduce, refine, or replace animal use where scientifically feasible. Reduction, refinement, and replacement are known as the “three Rs” of animal alternatives. These principles of humane treatment of laboratory animals are described as:

- Reducing animal use through improved science and experimental design
- Refining experimental procedures such that animal suffering is minimized
- Replacing animal models with non-animal procedures (e.g., *in vitro* technologies) where possible (Russell and Burch 1959)

Three *in vivo* methods are now commonly used by regulators to assess the estrogenic potential of substances: rat uterotrophic assay, rat pubertal female assay, and fish short-term reproduction assay. In addition, the “*in vitro*” rat uterine cytosol ER binding assay also requires the use of animals as a source of ER. Like the CERI-STTA, the BG1Luc ER TA test method will not directly replace any of these existing methods; however, it could be incorporated as part of a weight-of-evidence approach to reduce or eliminate the need for testing in these animal models. Currently, no *in vitro* test methods have been validated and accepted for use in the screening of both ER agonists and antagonists (ICCVAM 2002b). As discussed in **Section 1.0**, the EPA EDSP Tier 1 screening battery includes the CERI-STTA agonist test method, OPPTS 890.1300: Estrogen Receptor Transcriptional Activation (Human Cell Line [HeLa-9903]) (EPA 2009; OECD 2009)). The screening guideline also provides for the use of other scientifically valid methods. Therefore, the BG1Luc ER TA test method may be applicable for addressing the ER TA component of the EPA EDSP Tier 1 screening battery. Used in this context, the BG1Luc ER TA test method provides an opportunity to reduce animal use in endocrine disruptor testing by identifying substances that may enhance and/or inhibit the activation of the ER.

An evaluation of potential endocrine-active compounds (EACs) is required under European Commission Registration, Evaluation, Authorisation and Restriction of Chemicals [REACH] Regulation (EC) 1907/2006 (Bars et al. 2011; Bowman and Van Calster 2007; Harvey and Everett 2006; Løkke 2006; Marx-Stoelting et al. 2011). Validated *in vitro* methods may reduce animal use in this kind of testing program. “REACH-type” programs are also being adopted by Asian countries, and the availability of validated *in vitro* and *in silico* methods to screen/prioritize chemicals for these testing programs has potential to reduce animal use further. Following validation, the development of *in vitro* EAC assays into an OECD Test Guideline will broaden their potential for reducing animal use.

The BG1Luc ER TA method is being proposed as an independent part of a weight-of-evidence approach to prioritize potentially endocrine-active substances for further testing. Results from the BG1Luc ER TA test method were examined for concordance with published reports of ER binding. There was 97% (33/34) concordance between the BG1Luc ER TA results and ER binding data from the literature (see **Section 5.6**). In light of the excellent degree of agreement between ER binding and BG1Luc ER TA data (with no false negative results), it appears that evaluating results from BG1Luc ER TA agonist and antagonist testing may provide a viable alternative to conducting ER binding studies, which use animals as a source of ER. This cannot currently be accomplished with the only accepted ER TA method due to the inability of the CERI-STTA to assess ER antagonist activity.

Results from the BG1Luc ER TA test method were examined for concordance with published data from the uterotrophic assay (see **Section 5.7**). Based on a comparison with the *in vivo* uterotrophic assay classification, the 13 substances with data from the uterotrophic assay and conclusive test

results in the BG1Luc ER TA agonist test method produced overall concordance of 92% (12/13). All substances found positive in the uterotrophic assay were also positive in the BG1Luc ER TA method. The only discordant substance, butylbenzyl phthalate, was positive for ER agonist activity in the BG1Luc ER TA agonist test method and negative in the uterotrophic assay. These data indicate that the BG1Luc ER TA agonist test method has very good agreement with the *in vivo* results obtained with the uterotrophic assay, with no false negative results.

The development of a battery of *in vitro* and *in silico* methods that can replace animal testing for detecting chemicals that have the potential to interact with the endocrine system (i.e., EACs) is a biologically complex challenge. For example, a method for assessment of metabolites needs to be included with the *in vitro* assays, and assays for assessing the many modes of action of EACs on various tissues and species need to be developed and validated. The experience derived from validating and using the *in vitro* BG1Luc ER TA test method is expected to contribute to our knowledge and promote progress toward this goal. It should lead to the broader use of cell-based methods for EAC screening and could include the use of cells from other species.

9.2 Use of Animals in the BG1Luc ER TA Test Method

The BG1Luc ER TA test method utilizes cultured human ovary adenocarcinoma cells that endogenously express human ER and contain an estrogen-inducible gene expression system. Except for the fetal bovine sera used as part of the cell culture media, the test method does not require the use of animals.

10.0 Practical Considerations

Several issues are taken into account when assessing the practicality of using an alternative to an existing test method. In addition to performance evaluations, the following must be assessed:

- Laboratory equipment and supplies needed to conduct the alternative test method
- Level of personnel training required
- Labor costs
- Time required to complete the test method as compared to the existing test method

The time, personnel cost, and effort required to conduct the proposed test method must be considered reasonable when compared to those of the test method it is intended to replace. This section discusses the practical issues associated with using the BG1Luc ER TA test method for the determination of ER agonist and antagonist activity.

10.1 Transferability of the BG1Luc ER TA Test Method

Test method transferability addresses the ability of a method to be accurately and reliably performed by multiple laboratories (ICCVAM 2003b, 2003a), both those experienced in the particular type of procedure and laboratories with no prior experience. The transferability of the BG1Luc ER TA test method was demonstrated by the intra- and interlaboratory reproducibility studies in the validation study (see **Section 6.0**).

10.1.1 Facilities and Major Equipment

The facility requirements for conducting the BG1Luc ER TA test method include a standard laboratory setup for sterile cell culture procedures. The major equipment necessary is readily available and includes a laminar flow hood and a cell culture incubator. **Table 10-1** shows representative suppliers and estimated costs of this equipment.

Table 10-1 Example Suppliers and Costs of Major Equipment for the BG1Luc ER TA Test Method

Equipment	Example Supplier	Estimated Cost ^a
Laminar Flow Hood	Cole-Parmer	\$8,000–\$12,000
Cell Culture Incubator	Thomas Scientific	\$8,000–\$15,000

^a Estimated costs based on 2009 catalog prices

10.1.2 General Availability of Other Necessary Equipment and Supplies

The remaining equipment and supplies necessary to conduct the BG1Luc ER TA test method (e.g., microscopes, micropipettors, refrigerators/freezers, microtiter plates, cell culture supplies, sera, and reagents) are readily available in most cell culture laboratories or can be readily obtained from any of several scientific laboratory equipment and supply vendors.

10.1.3 BG1Luc4E2 Cell Line

The required BG1Luc4E2 cell line is available upon request from Dr. Michael S. Denison, Department of Environmental Toxicology, University of California, Davis.

10.2 BG1Luc ER TA Test Method Training Considerations

The level of training and expertise needed to conduct the BG1Luc ER TA test method should be similar to that needed for the HeLa-9903 ER TA test method, the only ER TA test method currently included on the EPA EDSP Tier 1 screening battery (see **Section 1.0**). Both methods require a moderate degree of technical capability and a high degree of skill in monitoring and maintaining appropriate cell growth conditions, troubleshooting cell culture problems, and analyzing and interpreting *in vitro* data. Accordingly, personnel should be trained in good cell culture practices, in the specialized culture procedures needed for this assay, and in safety and handling practices appropriate to the types of substances that may be tested in the laboratory (Hartung et al. 2002).

It is essential that all laboratory staff are trained to be aware of the need to minimize all sources of estrogenic contamination, which results in false positive outcomes.

10.3 Time and Cost Considerations

Most of the necessary equipment for conducting the BG1Luc ER TA test method is commonly found in laboratories that perform cell culture experiments. The one piece of nonstandard laboratory equipment is a microplate injecting luminometer (estimated cost is \$28,000), which is required for generating the RLU data used to establish a positive or negative result in the BG1Luc ER TA test method.

Supplies such as cell culture media, the reagents used to measure luciferase, and cell culture plasticware are available from numerous suppliers. An estimated cost for the BG1Luc ER TA test method, including relevant consumables (cell culture media, reagents, and supplies), is \$2000 per test substance for both agonist and antagonist testing (G. Clark, XDS, Inc., personal communication).

The BG1Luc ER TA test method takes approximately two days to perform (this includes a range finder test and at least one comprehensive test). The time estimate for the BG1Luc ER TA test method is similar to the two days necessary to conduct the CER1-STTA. The current cost of the CER1-STTA conducted at CER1 ranges from a minimum of \$1800 per test article based on at least 11 substances tested to a maximum of \$2500 per test article when one to five substances are tested (A. Ono, CER1, personal communication).

Commercially available *in vivo* test methods that are used to evaluate estrogenic activity are the uterotrophic and female pubertal assays, which take approximately 30 and 60 days to perform, respectively. The current approximate costs of commercially available uterotrophic and female pubertal assays are \$40,000 and \$140,000 per test substance, respectively (Willett and Sullivan 2010).

11.0 References

Food Quality Protection Act of 1996. 7 U.S.C. 136. Public Law 104-170.

Federal Food, Drug, and Cosmetic Act of 1938 (as amended). 21 U.S.C. 301 et seq. Public Law 75-717.

ICCVAM Authorization Act of 2000. 42 U.S.C. 285f-3. Public Law 106-545. Available: http://frwebgate.access.gpo.gov/cgi-bin/useftp.cgi?IPaddress=162.140.64.183&filename=publ545.pdf&directory=/disks/wais/data/106_cong_public_laws [accessed 17 September 2009].

Safe Drinking Water Act Amendments of 1996. 110 Stat 1613. Public Law 104-182.

Ankley G, Mihaich E, Stahl R, Tillitt D, Colborn T, McMaster S, et al. 1998. Overview of a workshop on screening methods for detecting potential (anti-) estrogenic/androgenic chemicals in wildlife. *Environ Toxicol Chem* 17(1): 68-87. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0031984814&partnerID=40&md5=384ec23dd1bcd9440752a9e625b52371>.

Baker VA. 2001. Endocrine disrupters - testing strategies to assess human hazard. *Toxicol In Vitro* 15(4-5): 413-419. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0034832175&partnerID=40&md5=71d6f5d9488cbcd4e130c3d06d9b5919>.

Baldwin WS, Curtis SW, Cauthen CA, Risinger JJ, Korach KS, Barrett JC. 1998. BG-1 ovarian cell line: an alternative model for examining estrogen-dependent growth in vitro. *In Vitro Cell Dev Bio Anim* 34(8): 649-654. Available: <http://www.jstor.org/stable/4294840>.

Ball LJ, Levy N, Zhao X, Griffin C, Tagliaferri M, Cohen I, et al. 2009. Cell type- and estrogen receptor-subtype specific regulation of selective estrogen receptor modulator regulatory elements. *Mol Cell Endocrinol* 299(2): 204-211. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-58649088924&partnerID=40&md5=2262a8459c9397be4fa5e61a2e0efd08>.

Bars R, Broeckaert F, Fegert I, Gross M, Hallmark N, Kedwards T, et al. 2011. Science based guidance for the assessment of endocrine disrupting properties of chemicals. *Regul Toxicol Pharmacol* 59(1): 37-46.

Bowers JL, Tyulmenkov VV, Jernigan SC, Klinge CM. 2000. Resveratrol acts as a mixed agonist/antagonist for estrogen receptors alpha and beta. *Endocrinology* 141(10): 3657-3667. Available: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11014220

Bowman D, Van Calster G. 2007. Reflecting on REACH: global implications of the European Union's chemicals regulation. *Nanotechnology Law and Business* 4(3): 375-383.

Brown M, Ning J, Ferreira JA, Bogener JL, Lubahn DB. 2009. Estrogen receptor- α and - β and aromatase knockout effects on lower limb muscle mass and contractile function in female mice. *Am J Physiol Endocrinol Metab* 296(4): E854-861. Available: <http://ajpendo.physiology.org/cgi/content/abstract/296/4/E854>.

Brown RP, Greer RD, Mihaich EM, Guiney PD. 2001. A critical review of the scientific literature on potential endocrine-mediated effects in fish and wildlife. *Ecotoxicol Environ Saf* 49(1): 17-25. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0034997186&partnerID=40&md5=a65fd57e9dc90c6d665a4bc7ee57a3ef>.

- Calabrese EJ, Baldwin LA. 2001. Hormesis: U-shaped dose responses and their centrality in toxicology. *Trends Pharmacol Sci* 22(6): 285-291. Available: <http://www.sciencedirect.com/science/article/pii/S0165614700017193>.
- Carey MF, Peterson CL, Stephen TS. 2009. *Transcriptional Regulation in Eukaryotes: Concepts, Strategies, and Techniques*. 2nd ed. Cold Spring Harbor, New York: Cold Spring Harbor Press.
- Chamness GC, Bannayan GA, Landry LA, Sheridan PJ, McGuire WL. 1979. Abnormal reproductive development in rats after neonatally administered antiestrogen (tamoxifen). *Biol Reprod* 21(5): 1087-1090. Available: <http://www.biolreprod.org/content/21/5/1087.abstract>.
- Chrzan BG, Bradford PG. 2007. Phytoestrogens activate estrogen receptor β 1 and estrogenic responses in human breast and bone cancer cell lines. *Mol Nutr Food Res* 51(2): 171-177.
- Clark GC, Gordon JD, Denison MS, Chu AC, Chu MD, Matherly C, et al. 2007. Comparison of Visual Inspection to CellTiter-Glo[®] in Evaluating Cytotoxicity in the LUMI-CELL[®] ER Bioassay. SOT 2007: 46th Annual Meeting & ToxExpo[™]; Charlotte, North Carolina; March 25-29, 2007.
- Collins-Burow BM, Burow ME, Duong BN, McLachlan JA. 2000. Estrogenic and antiestrogenic activities of flavonoid phytochemicals through estrogen receptor binding-dependent and -independent mechanisms. *Nutr Cancer* 38(2): 229-244. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0034434323&partnerID=40&md5=ea5f5fd8ffe20b0fa01d710297b1f7fc>.
- Combes RD. 2000. Endocrine disruptors: a critical review of *in vitro* and *in vivo* testing strategies for assessing their toxic hazard to humans. *Altern Lab Anim* 28(1): 81-118. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0033970280&partnerID=40&md5=1c2689d41c4691ddd8ae493430ead986>.
- EPA. 2006a. Good Laboratory Practice Standards. Federal Insecticide, Fungicide and Rodenticide Act. 40 CFR 160. Available: http://www.access.gpo.gov/nara/cfr/waisidx_06/40cfr160_06.html [accessed 1 November 2010].
- EPA. 2006b. Good Laboratory Practice Standards. Toxic Substances Control Act. 40 CFR 792. Available: http://www.access.gpo.gov/nara/cfr/waisidx_06/40cfr792_06.html [accessed 1 November 2010].
- EPA. 2009. Endocrine Disruptor Screening Program Test Guidelines: OPPTS 890.1300 - Estrogen Receptor Transcriptional Activation (Human Cell Line [HeLa-9903]). EPA 740-C-09-006. Washington, DC: U.S. Environmental Protection Agency. Available: http://www.epa.gov/ocspp/pubs/frs/publications/Test_Guidelines/series890.htm.
- Escande A, Pillon A, Servant N, Cravedi JP, Larrea F, Muhn P, et al. 2006. Evaluation of ligand selectivity using reporter cell lines stably expressing estrogen receptor alpha or beta. *Biochem Pharmacol* 71(10): 1459-1469. Available: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16554039.
- Evans WV, Diepenhorst EM. 1926. Further studies in luminescent grignard compounds. *J Am Chem Soc* 48(3): 715-723. Available: <http://dx.doi.org/10.1021/ja01414a026>.
- FDA. 2009. Good Laboratory Practice for Nonclinical Laboratory Studies. 21 CFR 58. Available: http://www.access.gpo.gov/nara/cfr/waisidx_08/21cfr58_08.html [accessed 1 November 2010].
- Foryst-Ludwig A, Kintscher U. 2010. Metabolic impact of estrogen signalling through ERalpha and ERbeta. *J Steroid Biochem Mol Biol* 122(1-3): 74-81. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-77956455483&partnerID=40&md5=dec90da705fb5c90a5f808382184a84a>.

- Fry DM, Toone CK. 1981. DDT-induced feminization of gull embryos. *Science* 213(4510): 922-924.
- Fujimoto N, Jinno N, Kitamura S. 2004. Activation of estrogen response element dependent transcription by thyroid hormone with increase in estrogen receptor levels in a rat pituitary cell line, GH3. *J Endocrinol* 181(1): 77-83. Available: <http://www.ncbi.nlm.nih.gov/pubmed>.
- Garrison PM, Tullis K, Aarts JM, Brouwer A, Giesy JP, Denison MS. 1996. Species-specific recombinant cell lines as bioassay systems for the detection of 2,3,7,8-tetrachlorodibenzo-p-dioxin-like chemicals. *Fundam Appl Toxicol* 30(2): 194-203. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-26844566935&partnerID=40&md5=17d651d7b8353c2f600f9ce359cd82e6>.
- Gelbke HP, Kayser M, Poole A. 2004. OECD test strategies and methods for endocrine disruptors. *Toxicology* 205(1-2): 17-25. Available: <http://www.sciencedirect.com/science/article/B6TCN-4D1V9VT-3/2/5990ec8b13e0670149872ec3a2e3b2f3>.
- Gordon JD, Chu AC, Chu MD, Denison MS, Clark GC. 2004. Detection of Estrogen Receptor Endocrine Disruptor Potency of Commonly Used Organochlorine Pesticides Using the LUMI-CELL ER Bioassay. *Dioxin 2004: 24th International Symposium on Halogenated Environmental Organic Pollutants and Persistent Organic Pollutants*; Berlin, Germany; September 6-10, 2004. *Organohalogen Compounds* 66: 169-174. Available: <http://www.dioxin20xx.org/pdfs/2004/04-151.pdf>.
- Gordon JD, Chu AC, Chu MD, Matherly C, Denison MS, Grune G, et al. 2005. Estrogenic Potency of Many Popular Sunscreens and It's "Non-Active" Components Detected Using the LUMI-CELL™ ER Bioassay. *Dioxin 2005: 25th International Symposium on Halogenated Environmental Organic Pollutants and Persistent Organic Pollutants*; Toronto, Canada; August 21-26, 2005. *Organohalogen Compounds* 67: 2685-2689.
- Gordon JD, Chu AC, Chu MD, Taylor CL, Denison MS, Clark GC. 2003. Validation of the LUMI-CELL™ ER Recombinant Bioassay for Rapid Evaluation of Chemicals for Potential Estrogenic Activity. *Dioxin 2003: 23rd International Symposium on Halogenated Environmental Organic Pollutants and Persistent Organic Pollutants*; Boston, Massachusetts; August 24-29, 2003. *Organohalogen Compounds* 65: 78-81.
- Greim HA. 2004. The endocrine and reproductive system: adverse effects of hormonally active substances? *Pediatrics* 113(4): 1070-1075. Available: <http://pediatrics.aappublications.org/cgi/content/abstract/113/4/S1/1070>.
- Guillette LJ, Gunderson MP. 2001. Alterations in development of reproductive and endocrine systems of wildlife populations exposed to endocrine-disrupting contaminants. *Reproduction* 122(6): 857-864. Available: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11732981.
- Hall JM, Couse JF, Korach KS. 2001. The multifaceted mechanisms of estradiol and estrogen receptor signaling. *J Biol Chem* 276(40): 36869-36872. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0035813112&partnerID=40&md5=4166927c2562aca04a18d642b19eba4e>.
- Harris HA. 2007. Estrogen receptor- β : recent lessons from *in vivo* studies. *Mol Endocrinol* 21(1): 1-13. Available: <http://mend.endojournals.org/cgi/content/abstract/21/1/1>.
- Hartung T, Balls M, Bardouille C, Blanck O, Coecke S, Gstraunthaler G, et al. 2002. Good cell culture practice: ECVAM good cell culture practice task force report 1. *Altern Lab Anim* 30(4): 407-414. Available: <http://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-0036657443&partnerID=40&rel=R6.5.0>

Harvey PW, Everett DJ. 2006. Regulation of endocrine-disrupting chemicals: critical overview and deficiencies in toxicology and risk assessment for human health. *Best Pract Res Clin Endocrinol Metab* 20(1): 145-165. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-33644642967&partnerID=40&md5=aedf3542fdcc74887499ae26e21264e8>.

Hass U, Dalgaard M, Jarfelt K, Kledal T. 2004. OECD Conceptual Framework for Testing and Assessment of Endocrine Disrupters as a basis for regulation of substances with endocrine disrupting properties. *TemaNord* 2004:555. Copenhagen, Denmark: Nordic Council of Ministers. Available: http://www.norden.org/en/publications/publications/2004-555/at_download/publicationfile.

Hayashi SI, Eguchi H, Tanimoto K, Yoshida T, Omoto Y, Inoue A, et al. 2003. The expression and function of estrogen receptor α and β in human breast cancer and its clinical application. *Endocr Relat Cancer* 10(2): 193-202. Available: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12790782.

Huet MC. 2000. OECD activity on endocrine disrupters test guidelines development. *Ecotoxicology* 9(1-2): 77-84. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0033830635&partnerID=40&md5=8b1aff81332436821770e790e012d732>.

ICCVAM. 2002a. Background Review Document. Current Status of Test Methods for Detecting Endocrine Disruptors: In Vitro Androgen Receptor Binding Assays. NIH Publication No. 03-4506. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/docs/endo_docs/final1002/arbndbrd/ARBd034506.pdf.

ICCVAM. 2002b. Background Review Document. Current Status of Test Methods for Detecting Endocrine Disruptors: In Vitro Androgen Receptor Transcriptional Activation Assays. NIH Publication No. 03-4507. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/docs/endo_docs/final1002/arta_brd/ARTA034507.pdf.

ICCVAM. 2002c. Background Review Document. Current Status of Test Methods for Detecting Endocrine Disruptors: In Vitro Estrogen Receptor Binding Assays. NIH Publication No. 03-4504. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/docs/endo_docs/final1002/erbndbrd/ERBd034504.pdf.

ICCVAM. 2002d. Background Review Document. Current Status of Test Methods for Detecting Endocrine Disruptors: In Vitro Estrogen Receptor Transcriptional Activation Assays. NIH Publication No. 03-4505. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/docs/endo_docs/final1002/erta_brd/ERTA034505.pdf.

ICCVAM. 2002e. Expert Panel Evaluation of the Validation Status of In Vitro Test Methods for Detecting Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays - Expert Panel Final Report. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/docs/endo_docs/expertpanfinalrpt/panelrpt1102.pdf.

ICCVAM. 2003a. ICCVAM Evaluation of In Vitro Test Methods for Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays. NIH Publication No. 03-4503. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/docs/endo_docs/edfinalrpt0503/edfinrpt.pdf.

ICCVAM. 2003b. 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. Available:
http://iccvam.niehs.nih.gov/SuppDocs/SubGuidelines/SD_subg034508.pdf [accessed 4 November 2010].
- ICCVAM. 2006. Finalized Addendum to ICCVAM Evaluation of In Vitro Test Methods for Detecting Potential Endocrine Disruptors. Research Triangle Park, NC:National Institute of Environmental Health Sciences. Available:
http://iccvam.niehs.nih.gov/docs/endo_docs/EDAddendFinal.pdf.
- ICCVAM. 2008. Biennial Progress Report: Interagency Coordinating Committee on the Validation of Alternative Methods. NIH Publication No. 08-6529. Research Triangle Park, NC:National Institute of Environmental Health Sciences. Available: <http://iccvam.niehs.nih.gov/docs/annrpt/BiennialRpt2007-508.pdf>.
- ICCVAM. 2011. Independent Scientific Peer Review Panel Report: Evaluation of the LUMI-CELL[®] ER (BG1Luc ER TA) Test Method. Research Triangle Park, NC:National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/docs/endo_docs/EDPRPrept2011.pdf.
- IPCS. 2002. Global Assessment of the State-of-the-Science of Endocrine Disruptors. WHO/PCS/EDC/02.2. Geneva:World Health Organization, International Programme on Chemical Safety. Available:
http://www.who.int/ipcs/publications/new_issues/endocrine_disruptors/en/index.html.
- ISO. 2000. ISO 9000: Quality Management Systems (2000). Geneva:International Organization for Standardization. Available: Available for purchase at <http://www.iso.org/iso/home.htm>.
- Jefferson WN, Padilla-Banks E, Clark G, Newbold RR. 2002. Assessing estrogenic activity of phytochemicals using transcriptional activation and immature mouse uterotrophic responses. *J Chromatogr B* 777(1-2): 179-189. Available: <http://www.sciencedirect.com/science/article/B6X0P-46DPF64-5/2/32646d99705a713de8cfc76d28dfb09f>.
- Jones LA, Hajek RA. 1995. Effects of estrogenic chemicals on development. *Environ Health Perspect* 103 Suppl(7): 63-67. Available:
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1518876/?tool=pubmed>.
- Judson RS, Houck KA, Kavlock RJ, Knudsen TB, Martin MT, Mortensen HM, et al. 2010. *In vitro* screening of environmental chemicals for targeted testing prioritization: the ToxCast project. *Environ Health Perspect* 118(4): 485-492. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-77951675220&partnerID=40&md5=deb959f6fe9b73fd3ff34c5a7adf6>.
- Kanno J, Onyon L, Peddada S, Ashby J, Jacob E, Owens JW. 2003a. The OECD program to validate the rat uterotrophic bioassay. Phase 2: Coded single-dose studies. *Environ Health Perspect* 111(12): 1550-1558. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0141451919&partnerID=40&md5=a08e0ad66564f5a9dbabda5caa81f0c1>.
- Kanno J, Onyon L, Peddada S, Ashby J, Jacob E, Owens JW. 2003b. The OECD program to validate the rat uterotrophic bioassay. Phase 2: Dose-response studies. *Environ Health Perspect* 111(12): 1530-1549. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0041698394&partnerID=40&md5=5d5ee3b9a8e1f1cde1bda9029d4dabed>.
- Kavlock R, Barr D, Boekelheide K, Breslin W, Breyse P, Chapin R, et al. 2006. NTP-CERHR expert panel update on the reproductive and developmental toxicity of di(2-ethylhexyl) phthalate. *Reprod Toxicol* 22(3): 291-399. Available: <http://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-33750945353&partnerID=40&rel=R6.0.0>

- Kavlock RJ. 1999. Overview of endocrine disruptor research activity in the United States. *Chemosphere* 39(8): 1227-1236. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0033213876&partnerID=40&md5=659f3fa3e6d5cdf072b58bef543aa90b>.
- Kelly BC, Ikonomou MG, Blair JD, Surridge B, Hoover D, Grace R, et al. 2009. Perfluoroalkyl contaminants in an Arctic marine food web: trophic magnification and wildlife exposure. *Environ Sci Technol* 43(11): 4037-4043.
- Komm BS, Kharode YP, Bodine PV, Harris HA, Miller CP, Lyttle CR. 2005. Bazedoxifene acetate: a selective estrogen receptor modulator with improved selectivity. *Endocrinology* 146(9): 3999-4008. Available: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15961563
- Kuiper GG, Lemmen JG, Carlsson B, Corton JC, Safe SH, van der Saag PT, et al. 1998. Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta. *Endocrinology* 139(10): 4252-4263. Available: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9751507.
- Latendresse JR, Bucci TJ, Olson G, Mellick P, Weis CC, Thorn B, et al. 2009. Genistein and ethinyl estradiol dietary exposure in multigenerational and chronic studies induce similar proliferative lesions in mammary gland of male Sprague-Dawley rats. *Reprod Toxicol* 28(3): 342-353. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-70349983562&partnerID=40&md5=b39922e6f718d5f9155d549203a50d6c>.
- Lee F, Hall CV, Ringold GM, Dobson DE, Luh J, Jacob PE. 1984. Functional analysis of the steroid hormone control region of mouse mammary tumor virus. *Nucleic Acids Res* 12(10): 4191-4206. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0021770607&partnerID=40&md5=aaeb01e494b326abfffe3ddb5503f6ec>.
- Løkke S. 2006. The precautionary principle and chemicals regulation: past achievements and future possibilities. *Environ Sci Pollut Res Int* 13(5): 342-349. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-33749631247&partnerID=40&md5=8d9c935217f87b3316d03128376b2a8b>.
- Marsaud V, Gougelet A, Maillard S, Renoir J-M. 2003. Various phosphorylation pathways, depending on agonist and antagonist binding to endogenous estrogen receptor α (ER α), differentially affect ER α extractability, proteasome-mediated stability, and transcriptional activity in human breast cancer cells. *Mol Endocrinol* 17(10): 2013-2027. Available: <http://mend.endojournals.org/cgi/content/abstract/17/10/2013>.
- Martin MT, Dix DJ, Judson RS, Kavlock RJ, Reif DM, Richard AM, et al. 2010. Impact of environmental chemicals on key transcription regulators and correlation to toxicity end points within EPA's ToxCast program. *Chem Res Toxicol* 23(3): 578-590. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-77949413202&partnerID=40&md5=670f4b2704799d4f7a5f37402398fa1c>.
- Marx-Stoelting P, Pfeil R, Solecki R, Ulbrich B, Grote K, Ritz V, et al. 2011. Assessment strategies and decision criteria for pesticides with endocrine disrupting properties relevant to humans. *Reprod Toxicol* 31(4): 574-584. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-79956359535&partnerID=40&md5=d47e92be57a1f461e430596fde944663>.
- Mohler ML, Narayanan R, Coss CC, Hu K, He Y, Wu Z, et al. 2010. Estrogen receptor beta selective nonsteroidal estrogens: seeking clinical indications. *Expert Opin Ther Pat* 20(4): 507-534. Available:

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=20302450.

Monje P, Boland R. 2001. Subcellular distribution of native estrogen receptor α and β isoforms in rabbit uterus and ovary. *J Cell Biochem* 82(3): 467-479. Available: <http://dx.doi.org/10.1002/jcb.1182>.

Morris JM, Van Wagenen G, McCann T, Jacob D. 1967. Compounds interfering with ovum implantation and development. II. Synthetic estrogens and antiestrogens. *Fertil Steril* 18(1): 18-34.

Newbold RR. 2008. Prenatal exposure to diethylstilbestrol (DES). *Fertil Steril* 89(2 Suppl): e55-e56. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-41749116162&partnerID=40&md5=46caa436699d70e4629b7bd4086586db>.

Newbold RR. 2010. Impact of environmental endocrine disrupting chemicals on the development of obesity. *Hormones (Athens)* 9(3): 206-217. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-77956297058&partnerID=40&md5=ef340b6dbcaec410aad7b3a355bb5af>.

Newbold RR, Jefferson WN, Padilla-Banks E. 2009. Prenatal exposure to bisphenol A at environmentally relevant doses adversely affects the murine female reproductive tract later in life. *Environ Health Perspect* 117(6): 879-885. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-67349143766&partnerID=40&md5=f2e0638ca0f78a102d88c7bdf9a67014>.

Newbold RR, Padilla-Banks E, Jefferson WN, Heindel JJ. 2008. Effects of endocrine disruptors on obesity. *Int J Androl* 31(2): 201-207. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-40349087016&partnerID=40&md5=0d64125e9de20062ab15759680f6bd5b>.

OECD. 1998. OECD Series on Good Laboratory Practice and Compliance Monitoring No. 1: OECD Principles on Good Laboratory Practice [as revised in 1997]. Paris:OECD Publishing. Available: http://www.oecd.org/document/63/0,3343,en_2649_34381_2346175_1_1_1_1,00.html [accessed 30 July 2010].

OECD. 2002a. OECD Conceptual Framework. Paris:Organisation for Economic Co-operation and Development. Available: http://www.oecd.org/document/58/0,3343,en_2649_34377_2348794_1_1_1_1,00.html.

OECD. 2002b. OECD Series on Testing and Assessment No. 21 - Detailed Review Paper: Appraisal of Test Methods for Sex Hormone Disrupting Chemicals. Paris:OECD Publishing. Available: [http://www.oecd.org/officialdocuments/displaydocument/?doclanguage=en&cote=env/jm/mono\(2002\)8](http://www.oecd.org/officialdocuments/displaydocument/?doclanguage=en&cote=env/jm/mono(2002)8).

OECD. 2005. Guidance Document No. 34: Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. Paris:OECD Publishing.

OECD. 2007. Test No. 440. Uterotrophic Bioassay in Rodents: A Short-Term Screening Test for Oestrogenic Properties [adopted 16 October 2007]. In: OECD Guidelines for the Testing of Chemicals, Section 4: Health Effects. Paris:OECD Publishing. Available: <http://dx.doi.org/10.1787/9789264067417-en>.

OECD. 2009. Test No. 455. Stably Transfected Human Estrogen Receptor- α Transcriptional Activation Assay for Detection of Estrogenic Agonist-Activity of Chemicals [adopted 7 September 2009]. In: OECD Guidelines for the Testing of Chemicals, Section 4: Health Effects. Paris:OECD Publishing. Available: <http://dx.doi.org/10.1787/9789264076372-en>.

Owens JW, Ashby J. 2002. Critical review and evaluation of the uterotrophic bioassay for the identification of possible estrogen agonists and antagonists: in support of the validation of the OECD

uterotrophic protocols for the laboratory rodent. *Crit Rev Toxicol* 32(6): 445-520. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0036445688&partnerID=40&md5=ec3847692aa5e2d3e33da92d4ea83f8b>.

Owens JW, Koeter HBWM. 2003. The OECD program to validate the rat uterotrophic bioassay: an overview. *Environ Health Perspect* 111(12): 1527-1529. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0141517411&partnerID=40&md5=9f565f88f848d3050927651d0dc47286>.

Park S-H, Kim K-Y, An B-S, Choi J-H, Jeung E-B, Leung PCK, et al. 2009. Cell growth of ovarian cancer cells is stimulated by xenoestrogens through an estrogen-dependent pathway, but their stimulation of cell growth appears not to be involved in the activation of the mitogen-activated protein kinases ERK-1 and p38. *J Reprod Dev* 55(1): 23-29. Available: http://www.jstage.jst.go.jp/article/jrd/55/1/55_23/_article.

Poland A, Wind M, Stokes WS, Allen D, Fitzpatrick S, Kulpa-Eddy J, et al. 2008. The NICEATM-ICCVAM Five-Year Plan: Creating a Path Forward to Reduce, Refine, and Replace Animal Testing. 47th Annual Meeting of the Society of Toxicology; Seattle, Washington.

Pujol P, Rey JM, Nirde P, Roger P, Gastaldi M, Laffargue F, et al. 1998. Differential expression of estrogen receptor-alpha and -beta messenger RNAs as a potential marker of ovarian carcinogenesis. *Cancer Res* 58(23): 5367-5373. Available: <http://cancerres.aacrjournals.org/content/58/23/5367.full.pdf+html>.

Rogers JM, Denison MS. 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 Mol Toxicol* 13(1): 67-82. Available: <http://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-0033936762&partnerID=40&rel=R6.0.0>

Rogers JM, Denison MS. 2002. Analysis of the antiestrogenic activity of 2,3,7,8-tetrachlorodibenzo-p-dioxin in human ovarian carcinoma BG-1 cells. *Mol Pharmacol* 61(6): 1393-1403. Available: <http://molpharm.aspetjournals.org/content/61/6/1393.abstract>.

Rozman KK, Bhatia J, Calafat AM, Chambers C, Culty M, Etzel RA, et al. 2006. NTP-CERHR expert panel report on the reproductive and developmental toxicity of genistein. *Birth Defects Res B Dev Reprod Toxicol* 77(6): 485-638. Available: <http://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-33845900198&partnerID=40&rel=R6.0.0>

Russell WMS, Burch RL. 1959. *The Principles of Humane Experimental Technique*. London: Methuen & Co. Ltd. [Reissued: 1992, Universities Federation for Animal Welfare, Herts England].

Segner H. 2005. Developmental, reproductive, and demographic alterations in aquatic wildlife: Establishing causality between exposure to endocrine-active compounds (EACs) and effects. *Acta Hydrochimica et Hydrobiologica* 33(1): 17-26. Available: <http://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-20444458366&partnerID=40&rel=R6.0.0>

Shang Y, Hu X, DiRenzo J, Lazar MA, Brown M. 2000. Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. *Cell* 103(6): 843-852. Available: <http://www.sciencedirect.com/science/article/B6WSN-4221G5G-4/2/298cf4bea7607ba7f0b0c3c9ee586f60>.

Skliris GP, Leygue E, Watson PH, Murphy LC. 2008. Estrogen receptor alpha negative breast cancer patients: estrogen receptor beta as a therapeutic target. *J Steroid Biochem Mol Biol* 109(1-2): 1-10.

Available: <http://www.sciencedirect.com/science/article/B6T8X-4R9JTS0-1/2/3e1baf36c8d6dea60467e3ee31e9764c>.

Soin T, Smagghe G. 2007. Endocrine disruption in aquatic insects: a review. *Ecotoxicology* 16(1): 83-93. Available: <http://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-33847120873&partnerID=40&rel=R6.0.0>

Sormo EG, Salmer MP, Jenssen BM, Hop H, Baek K, Kovacs KM, et al. 2006. Biomagnification of polybrominated diphenyl ether and hexabromocyclododecane flame retardants in the polar bear food chain in Svalbard, Norway. *Environ Toxicol Chem* 25(9): 2502-2511.

Stokes W, Wind M. 2009. Recent Progress and Future Directions at NICEATM-ICCVAM: Validation and Acceptance of Alternative Methods that Reduce, Replace, and Refine Animal Use [Abstract]. Seventh World Congress on Alternatives and Animal Use in the Life Sciences. Available: <http://iccvam.niehs.nih.gov/meetings/7thWC/NICEATMupdate.pdf>.

Takeyoshi M. 2006. Draft Report of Pre-validation and Inter-laboratory Validation For Stably Transfected Transcriptional Activation (TA) Assay to Detect Estrogenic Activity - The Human Estrogen Receptor Alpha Mediated Reporter Gene Assay Using hER-HeLa-9903 Cell Line. Japan:Chemicals Evaluation and Research Institute (CERI).

Tremblay A, Tremblay GB, Labrie F, Giguère V. 1999. Ligand-independent recruitment of SRC-1 to estrogen receptor beta through phosphorylation of activation function AF-1. *Mol Cell* 3(4): 513-519. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0033120030&partnerID=40&md5=9432bd8b91030e436024222c54c99daf>.

Tsai WT. 2006. Human health risk on environmental exposure to bisphenol-A: a review. *J Environ Sci Health C Environ Carcinog Ecotoxicol Rev* 24(2): 225-255. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-33751161665&partnerID=40&md5=2a496fbc34044774373c882975681e85>.

Tyler CR, Jobling S, Sumpter JP. 1998. Endocrine disruption in wildlife: a critical review of the evidence. *Crit Rev Toxicol* 28(4): 319-361. Available: <http://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-0031844641&partnerID=40&rel=R6.0.0>

vom Saal FS, Akingbemi BT, Belcher SM, Birnbaum LS, Crain DA, Eriksen M, et al. 2007. Chapel Hill bisphenol A expert panel consensus statement: Integration of mechanisms, effects in animals and potential to impact human health at current levels of exposure. *Reprod Toxicol* 24(2): 131-138. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-34548528825&partnerID=40&md5=8d71623fea54246f5741d53148a66350>.

Webb P, Lopez G, Uht R, Kushner P. 1995. Tamoxifen activation of the estrogen receptor/AP-1 pathway: potential origin for the cell-specific estrogen-like effects of antiestrogens. *Mol Endocrinol* 9(4): 443-456. Available: <http://mend.endojournals.org/cgi/content/abstract/9/4/443>.

Weinberg S. 2003. Good Laboratory Practice Regulations. 3rd ed. New York:Marcel Dekker, Inc.

Weiser MJ, Foradori CD, Handa RJ. 2008. Estrogen receptor beta in the brain: from form to function. *Brain Res Rev* 57(2): 309-320. Available: <http://www.sciencedirect.com/science/article/B6SYS-4P29K6S-1/2/35357a74f422df9ead733d717a2ea78d>.

Welshons WV, Wolf MF, Murphy CS, Jordan VC. 1988. Estrogenic activity of phenol red. *Mol Cell Endocrinol* 57(3): 169-178. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0023942664&partnerID=40&md5=012f0f6d671aadf4be7f68d346d222e4>.

- Whitten PL, Lewis C, Russell E, Naftolin F. 1995. Potential adverse effects of phytoestrogens. *J Nutr* 125(3 Suppl). Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0028800719&partnerID=40&md5=047ce1d791ee15a7915baed18c02b326>.
- Whitten PL, Naftolin F. 1992. Effects of a phytoestrogen diet on estrogen-dependent reproductive processes in immature female rats. *Steroids* 57(2): 56-61. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0026603747&partnerID=40&md5=f27209666e7916c55bf5d9e4bb432567>.
- Whitten PL, Naftolin F. 1998. Reproductive actions of phytoestrogens. *Baillieres Clin Endocrinol Metab* 12(4): 667-690. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0032467036&partnerID=40&md5=133a5c5a6ba81e1301228d1ea9d19264>.
- Whitten PL, Patisaul HB. 2001. Cross-species and interassay comparisons of phytoestrogen action. *Environ Health Perspect* 109(1 Suppl): 5-20. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0035051046&partnerID=40&md5=8b13acb186e7bc20fce201b60dc84059>.
- Whitten PL, Russell E, Naftolin F. 1992. Effects of a normal, human-concentration, phytoestrogen diet on rat uterine growth. *Steroids* 57(3): 98-106. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0026521082&partnerID=40&md5=5c0113fa9a9713b308157cbc6df9e785>.
- Willett C, Sullivan KM. 2010. Application of an Intelligent Testing Strategy to the US EPA Endocrine Disruptor Screening Program. 7th World Congress on Animal Alternatives; Berlin, Germany. Available: <http://www.mediapeta.com/peta/PDF/EDSP%20Feature/Non-Animal%20Methods/SOT09WillettSullivan.pdf>.
- Wong PSY, Matsumura F. 2006. Serum free BG-1 cell proliferation assay: a sensitive method for determining organochlorine pesticide estrogen receptor activation at the nanomolar range. *Toxicol In Vitro* 20(3): 382-394. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-33644761763&partnerID=40&md5=efa583ae9f54833796b4df2808f7c4ee>.
- Zhou R, Treeck O, Horn F, Ortmann O. 2005. Effects of prolonged tamoxifen treatment on receptor expression and apoptosis of ovarian cancer cells. *Gynecol Oncol* 96(3): 678-683. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-13944275572&partnerID=40&md5=324b47f2db6ea22e23b4db9a0ddd9834>.
- Zimniski SJ, Garola RE, Fendl K, Peterson CM. 1989. Endocrine characterization of a human ovarian carcinoma (BG-1) established in nude mice. *Steroids* 54(6): 593-606. Available: <http://www.sciencedirect.com/science/article/B6TC9-47F6YGM-V/2/1226a11f20506a2f03102a934a1f83e1>.

12.0 Glossary

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

Accuracy*: (a) The closeness of agreement between a test method result and an accepted reference value. (b) The proportion of correct outcomes of a test method. It is a measure of test method performance and is often used interchangeably with “concordance.”

Adenosine triphosphate (ATP): A nucleotide involved in energy metabolism and required for RNA synthesis. It occurs in all cells and is used to store energy in the form of high-energy phosphate bonds.

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

Androgen: A class of steroid hormones that includes testosterone and 5 α -dihydrotestosterone. These hormones are responsible for the development and maintenance of the male reproductive system.

Androgen receptor: The receptor to which androgens bind.

Antagonist: A substance that inhibits a response, e.g., transcription, when it binds to a specific receptor.

Assay*: An experimental system, often used interchangeably with “test” or “test method.”

BG-1: The BG1Luc4E2 cell line was derived from BG-1 immortalized adenocarcinoma cells that endogenously express both forms of the estrogen receptor (ER α and ER β) 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 promoter and the firefly luciferase gene.

Cell density: The degree of confluence of cells growing in a monolayer in a single well of a tissue culture plate.

Cell morphology: The shape and appearance of cells grown in a monolayer in a single well of a tissue culture plate. Cells that are dying often exhibit abnormal cellular morphology.

Charcoal/dextran treatment: Treatment of serum used in cell culture. Treatment with charcoal/dextran (often referred to as “stripping”) removes endogenous hormones and hormone-binding proteins.

Coded test substances: Substances labeled by code rather than name so that they can be tested and evaluated without knowledge of their identity or anticipation of test results. Coded test substances are used to avoid intentional or unintentional bias when evaluating laboratory or test method performance.

Coefficient of variation: A statistical representation of the precision of a test. It is expressed as a percentage and is calculated as follows:

$$\left(\frac{\text{standard deviation}}{\text{mean}} \right) \times 100$$

Comprehensive test: A test performed to determine an EC₅₀ or IC₅₀ value. Compared with the range finder test, the comprehensive test uses a smaller dilution factor for the concentrations tested.

The definitions in this Glossary are restricted to their use with respect to endocrine mechanisms and actions.

* Definition used by the Interagency Coordinating Committee on the Validation of Alternative Methods

Concordance*: The proportion of all substances tested that are correctly classified as positive or negative. It is a measure of test method performance, and it is often used interchangeably with “accuracy.”

Control: A substance with a known response selected for use during the research, development, protocol standardization, and validation of a proposed test method. Controls are used to evaluate the ongoing performance of a test method. All experimental controls must fall within established historical norms for an experiment to pass “acceptance criteria” and be considered valid.

Culture medium: An aqueous solution containing vitamins, minerals, and growth factors to support the growth of cells.

Cytotoxicity: The adverse effects resulting from interference with structures and/or processes essential for cell survival, proliferation, and/or function. For most substances, toxicity is a consequence of nonspecific alterations in “basal cell functions” (i.e., via mitochondria, plasma membrane integrity, etc.).

Definitive results: Data and calculations from an assay (excluding data from rejected plates or other inadequate data).

Dextran: A viscous or semiviscous polymer of glucose.

EC₅₀: The half-maximal effective concentration of an agonist test substance (concentration required to induce 50% of the maximum possible response).

Endocrine: Of or relating to the endocrine system, endocrine glands, and hormones.

Endocrine disruptor: A substance that interacts with the endocrine system to alter normal functioning. Endocrine disruptors may act directly by interfering with receptor binding or indirectly by altering hormone biosynthesis, transport, action, or metabolism.

Endocrine system: A system of glands throughout the body, the hormones they secrete, and the receptors that recognize and respond to the hormones.

Endpoint: The biological process, response, or effect assessed by a test method.

Essential test method components*: Structural, functional, and procedural elements of a validated test method that should be included in the protocol for a mechanistically and functionally similar proposed test method. These components include unique characteristics of the test method, critical procedural details, and quality control measures. Inclusion of essential test method components is necessary when the acceptability of a proposed test method is being evaluated based on performance standards derived from a mechanistically and functionally similar validated test method.

False negative*: An active substance incorrectly identified as negative by a test method.

False negative rate*: The proportion of all positive (active) substances falsely identified as negative. It is a measure of test method performance.

False positive*: An inactive substance incorrectly identified as positive by a test method.

False positive rate*: The proportion of all negative (inactive) substances falsely identified as positive. It is a measure of test method performance.

Fluorescence: The emission of visible or invisible radiation by certain substances as a result of incident radiation of a shorter wavelength, such as x-rays or ultraviolet light.

Good Laboratory Practice (GLP)*: Regulations promulgated by the U.S. Food and Drug Administration and the U.S. Environmental Protection Agency, and principles and procedures adopted by the Organisation for Economic Co-operation and Development and Japanese authorities.

GLP regulations cover record keeping, quality assurance, and laboratory practices for studies that will be the basis for data submissions to national regulatory agencies.

Hill function: A four-parameter logistic mathematical model relating the concentration of the test substance to the response (typically following a sigmoidal shape).

$$Y = \text{Bottom} + \frac{\text{Top} - \text{Bottom}}{1 + 10^{(\log \text{EC}_{50} - \log X) \text{HillSlope}}}$$

where Y = response (i.e., luciferase activity), X is the substance concentration producing the response, Bottom is the minimum response, Top is the maximum response, EC₅₀ is the substance concentration at the response midway between Top and Bottom, and HillSlope describes the slope of the curve.

IC₅₀: The half-maximal inhibitory concentration of an antagonist (concentration that causes 50% inhibition of the measured response).

Interlaboratory reproducibility*: A measure of whether different qualified laboratories, using the same protocol and test substances, can produce qualitatively and quantitatively similar results. Interlaboratory reproducibility is determined during the validation process and indicates the extent to which a test method can be transferred successfully among laboratories.

Intralaboratory repeatability*: The closeness of agreement between test results obtained within a single laboratory when the procedure is performed on the same substance under identical conditions within a given time period.

Intralaboratory reproducibility*: A measure of whether qualified people within the same laboratory can successfully replicate results using a specific test protocol at different times; the first stage of validation.

In vitro: Literally, in glass. Refers to assays that are carried out in an artificial system (e.g., in a test tube or Petri dish) and typically use single-cell organisms, cultured cells, cell-free extracts, or purified cellular components.

In vivo: In the living organism. Refers to assays performed in multicellular organisms.

Luciferase: An enzyme present in the cells of some bioluminescent organisms that catalyzes the oxidation of luciferin and ATP to produce luminescence.

Luminescence: The emission of radiation, especially of visible light caused by chemical or biochemical processes.

Luminometer: A device for measuring luminescence.

mRNA: Messenger ribonucleic acid (mRNA). The primary role of mRNA is to transport instructions related to the production of proteins essential to cell functioning from the genes to the rest of the cell.

Negative predictivity*: The proportion of correct negative responses among substances testing negative.

Peer review*: Objective review of data, a document, or proposal, and provision of recommendations, by an expert individual or group of individuals who have no conflict of interest with the outcome of the review.

Plasmid: A self-replicating circle of bacterial DNA. Plasmids can be artificially constructed and used as cloning vectors.

Positive predictivity*: The proportion of correct positive responses among substances testing positive.

Precipitate/precipitation: A solid substance, often in the form of crystals, separated from a solution, or the act of a solid substance separating from a solution.

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

Protocol*: The precise step-by-step description of a test, including the listing of all necessary reagents, criteria, and procedures for the valuation of the test data.

Protocol standardization: Selection of reference standards, controls, and performance standards for a protocol prior to initiation of validation efforts.

Q test: A simple statistical test to determine if a data point that appears to be different from the rest of the data points in a set may be discarded.

$$Q = \frac{\text{suspected outlier} - \text{closest value}}{\text{maximum value} - \text{minimum value}}$$

The resultant value, Q, is then compared to a table of critical values (Qc). If Q is larger than Qc, the data point is an outlier and can be discarded with 90% confidence. For example, in a data set with values of 100, 2655, and 241, the Q value is 0.95. For a set of three data points, Qc is 0.94. Q [0.95] is greater than Qc [0.94], so 2655 is an outlier and can be discarded.

Receptor: A protein or protein complex that binds to specific molecules to transport them elsewhere in the cell or to produce a chemical signal.

Receptor binding assay: An assay to measure the ability of a substance to bind to a hormone receptor protein, typically performed by measuring the ability of the substances to displace the bound natural hormone.

Reduction alternative*: A new or modified test method that reduces the number of animals required.

Reference standard: A reference substance used to demonstrate the adequacy of a test method. 17β-estradiol is the estrogenic reference standard, and raloxifene HCl is the anti-estrogenic reference standard for the BG1Luc ER TA test method.

Refinement alternative*: A new or modified test method that refines procedures to lessen or eliminate pain or distress in animals or enhance animal well-being.

Relative light unit (RLU): The unit used to characterize the endpoint of the BG1Luc ER TA test method, which is luminescence.

Relevance*: The extent to which a test method correctly predicts or measures the biological effect of interest in the species of interest. Relevance incorporates consideration of the “accuracy” or “concordance” of a test method.

Reliability*: A measure of the degree to which a test method can be performed reproducibly within and among laboratories over time. Reliability is assessed by calculating intra- and interlaboratory reproducibility and intralaboratory repeatability.

Reporter gene: A gene attached to a regulatory sequence of a gene of interest so that, when expression of the gene of interest is altered, activation of the reporter gene results in a quantifiable endpoint, such as luminescence.

Screen/screening test*: A rapid, simple test conducted for general classification of substances according to general categories of hazard. The results of a screen are generally used for preliminary decision making and to set priorities for more definitive tests.

Selection: Enrichment of stably transfected cells in tissue culture, usually by exposure to a substance that is noxious to nontransfected cells (e.g., exposure of cells to G418 kills cells that do not contain the G418 resistance vector).

Sensitivity*: The proportion of all positive substances that are classified correctly as positive in a test method. It is a measure of test method accuracy.

Specificity*: The proportion of all negative substances that are classified correctly as negative in a test method. It is a measure of test method accuracy.

Stable transfection: DNA encoding desirable genes is transfected into cultured cells in such a way that it is integrated into the cells' genome, resulting in the expression of those genes.

Standard operating procedures (SOPs)*: Formal written procedures that describe how specific laboratory operations are to be performed. These are required by Good Laboratory Practice (GLP) guidelines.

Tier 1 assay: An assay that is a component of the U.S. EPA Endocrine Disruptor Screening Program (EDSP) screening battery of tests. Tier 1 screening includes a battery of screening assays to identify substances with the potential to interact with the estrogen, androgen, or thyroid hormone systems.

Tier 2 assay: An assay that is a component of the U.S. EPA Endocrine Disruptor Screening Program (EDSP) testing battery. Tier 2 tests are longer in duration than Tier 1 tests and are intended to encompass a broad range of doses, life stages, and processes.

Transactivation: Induction of gene expression (often measured by a change in a chemical signal) in response to a transcription factor binding to DNA and activating adjacent proteins.

Transcription: Synthesis of RNA by RNA polymerases using a DNA template.

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.

Transfection: The process by which foreign DNA is introduced into a cell to change the cell's genotype.

Transferability*: The ability of a test method or procedure to be accurately and reliably performed in different, competent laboratories.

Transient transfection: DNA is transfected into cultured cells but is not permanently integrated into the cells genome and is retained for only two to three days.

Trypsin: An enzyme that cleaves proteins and can detach monolayer cells from a culture flask for resuspension.

Uterotrophic bioassay: An *in vivo* assay for estrogenic substances in which an increase in uterine weight compared with controls indicates positive estrogenic activity.

Validated method*: An accepted test method for which validation studies have been completed to determine its accuracy and reliability for a specific proposed use.

Validation*: The process by which the reliability and accuracy of a procedure are established for a specific purpose.

Vector: A small segment of DNA (frequently a plasmid or viral DNA) that is used to carry a foreign gene or DNA sequence into a cell.

Weight of evidence (process)*: The strengths and weaknesses of a collection of information used as the basis for a conclusion that may not be evident from the individual data.

Xenobiotic: A substance that is not produced by the organism of interest.

This page intentionally left blank

Annex A

Submission of XDS's LUMI-CELL™ ER High-Throughput System for Screening Estrogen-Like Chemicals for Review by ICCVAM (Received by NICEATM February 1, 2005)

The document is available electronically on the enclosed CD-ROM or at
<http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-BRDannex.htm>

This page intentionally left blank

Annex B

ICCVAM/NICEATM Evaluation – BG1Luc ER TA Submission

The document is available electronically on the enclosed CD-ROM or at
<http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-BRDannex.htm>

This page intentionally left blank

Annex C

NICEATM Report on the XDS BG1Luc ER TA Protocol Standardization Study Agonist and Antagonist Protocols

The document is available electronically on the enclosed CD-ROM or at
<http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-BRDannex.htm>

This page intentionally left blank

Annex D
Addendum to
ICCVAM Evaluation of *In Vitro* Test Methods for Detecting
Potential Endocrine Disruptors:
Estrogen Receptor and Androgen Receptor Binding and
Transcriptional Activation Assays
(NIH Publication No. 03-4503)

The document is available electronically on the enclosed CD-ROM or at
<http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-BRDannex.htm>

This page intentionally left blank

Annex E

ICCVAM/NICEATM BG1Luc ER TA – Agonist Protocol

The document is available electronically on the enclosed CD-ROM or at
<http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-BRDannex.htm>

This page intentionally left blank

Annex F

ICCVAM/NICEATM BG1Luc ER TA – Antagonist Protocol

The document is available electronically on the enclosed CD-ROM or at
<http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-BRDannex.htm>

This page intentionally left blank

Annex G

Materials Relating to Cell Viability

Annex G1
Cell Viability Manual

Annex G2
Quantitative versus Qualitative Assessment of Cell Viability

Annex G3
Viability Summaries

The documents are available electronically on the enclosed CD-ROM or at
<http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-BRDannex.htm>

This page intentionally left blank

Annex H

ICCVAM/NICEATM BG1Luc4E2 ER TA – Validation Study Design

Annex H1
Study Design and Work Plan

Annex H2
QA/QC Outline (Hiyoshi Corporation)

These documents are available electronically on the enclosed CD-ROM or at
<http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-BRDannex.htm>

This page intentionally left blank

Annex I

Substances Used for the Validation of the BG1Luc ER TA Test Method

Annex I1

ICCVAM-Recommended Substances – Structures, Chemical, and Product Class

Annex I2

Substances Used During BG1Luc ER TA Validation – Purity and Supplier Information

Annex I3

Substances Used During BG1Luc ER TA Validation – Test Substance Codes

These documents are available electronically on the enclosed CD-ROM or at
<http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-BRDannex.htm>

This page intentionally left blank

Annex J

ICCVAM/NICEATM BG1Luc ER TA – Prism Files: Prism Files for All Validation Study Phases

Graphs from all phases of the validation are available electronically on the enclosed CD-ROM or at

<http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-BRDannex.htm>

To view the graphs, use either GraphPad Prism® or the Prism® viewer. The Prism viewer is a free program for inspecting Prism files. The viewer opens any Prism file and allows viewing and printing of data, analyses, and results. The Prism viewer is free and will not expire.

The Prism® viewer can be installed from the enclosed CD-ROM
or from the GraphPad website at
<http://www.graphpad.com/prism/viewer.htm>

This page intentionally left blank

Annex K

Experimental Details

Annex K1

List of Test Substance Ship and Test Dates

Annex K2

Phase 1 Experiments Conducted to Establish Historical Databases

Annex K3

Phase 2a Experiments

Annex K4

Phase 2b Experiments

Annex K5

Phase 3 Experiments

Annex K6

Phase 4 Experiments

These documents are available electronically on the enclosed CD-ROM or at
<http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-BRDannex.htm>

This page intentionally left blank

Annex L

Within-Plate DMSO Control Values for BG1Luc ER TA Agonist and Antagonist Assays

The document is available electronically on the enclosed CD-ROM or at
<http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-BRDannex.htm>

This page intentionally left blank

Annex M

ICCVAM/NICEATM BG1Luc4E2 ER TA – Plate Redesign and Compilation of a Historical Database

The document is available electronically on the enclosed CD-ROM or at
<http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-BRDannex.htm>

This page intentionally left blank

Annex N

Literature Update Files

Annex N1
Binding Data from the Scientific Literature

Annex N2
ER TA Data from the Scientific Literature

Annex N3
Bibliography for the ER Binding and Transcriptional Activation Literature Update

These documents are available electronically on the enclosed CD-ROM or at
<http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-BRDannex.htm>

This page intentionally left blank

Appendix D

Independent Scientific Peer Review Panel Assessment

D1 Summary Minutes from the Independent Scientific Peer Review Panel Meeting on
 March 29-30, 2011 D-3

D2 Independent Scientific Peer Review Panel Report Evaluation of the LUMI-CELL[®] ER
 (BG1Luc ER TA) Test Method D-35

This page intentionally left blank

Appendix D1

Summary Minutes from the Independent Scientific Peer Review Panel Meeting on March 29-30, 2011

This page intentionally left blank

Summary Minutes

Independent Scientific Peer Review Panel Meeting

Evaluation of *In Vitro* Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening

William H. Natcher Conference Center

National Institutes of Health

Bethesda, MD

March 29 - 30, 2011

Peer Review Panel Members:

John G. Vandenberg, PhD (Peer Review Panel Chair)	Professor Emeritus, North Carolina State University, Department of Biology, Raleigh, NC
A. John Bailer, PhD	Distinguished Professor and Chair, Department of Statistics Miami University, Oxford, OH
Christopher Borgert, PhD	President & Principal Scientist, Applied Pharmacology and Toxicology, Inc., Gainesville, FL
Grantley Charles, PhD, DABT	Senior Scientist, Toxicology Drug Safety Evaluation, Allergan Inc. Irvine, CA
Daniel Desaulniers, PhD	Research Scientist, Health Canada, Hazard Identification Division, Ottawa, ON, Canada
J. Charles Eldridge, PhD	Professor of Physiology and Pharmacology, Wake Forest University Health Sciences, Winston-Salem, NC
Susan Jobling, PhD	Professorial Research Fellow, Institute for the Environment, Brunel University, Uxbridge, Middlesex, UK
William R. Kelce, PhD	Principal, Exponent®, Cary, NC
Hyung Kim, PhD	Associate Professor, College of Pharmacy, Pusan National University, Busan, Republic of Korea
Steven L. Levine, PhD	Science Fellow and Product Safety Manager, Monsanto Company, St. Louis, MO

Peer Review Panel Members:

Alberto Mantovani, DVM	Head of Food and Veterinary Toxicology Unit, Italian National Health Institute, (ISS - Istituto Superiore di Sanità, viale), Rome, Italy
Ellen Mihaich, PhD, DABT	President and Principal Scientist, Environmental and Regulatory Resources, LLC, Durham, NC
Hiroshi Ono, PhD	Scientific Advisor, Hatano Research Institute, Food and Drug Safety Center, Kanagawa, Japan
Sherry L. Ward, PhD, MBA	Consultant, BioTred Solutions, New Market, MD
Marc Weimer, PhD	Biostatistician, German Cancer Research Center, Department of Biostatistics, Heidelberg, Germany
James Wittliff, PhD, MD <i>hc</i> , FACB	Professor, Department of Biochemistry & Molecular Biology; Director, Hormone Receptor Laboratory; Director, Institute for Molecular Diversity & Drug Design, University of Louisville, Louisville, KY
James Yager, PhD	Professor and Director of Toxicology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD

ICCVAM and ICCVAM Endocrine Disruptor Working Group Members:

Surender Ahir, PhD	OSHA, Washington, D.C.
M. Cecilia Aguila, DVM	FDA, CVM, Rockville, MD
Don Bergfelt, PhD	EPA, Washington, D.C.
Michael Bolger, Ph.D, DABT	FDA, CFSAN, College Park, MD
Jeffrey Bray, PhD	FDA, CDER, Silver Spring, MD
Paul Brown, PhD	FDA, CDER, Silver Spring, MD
Kent Carlson, PhD	CPSC, Bethesda, MD
Warren Casey, PhD, DABT	NIEHS, NTP, Research Triangle Park, NC
Karen Davis-Bruno, PhD	FDA, CDER, Silver Spring, MD
Kenneth Delclos, PhD	FDA, NCTR, Jefferson, AR
Charles Eirkson, PhD	FDA, CVM, Rockville, MD
Ross Filice, MS	FDA, CDER, Silver Spring, MD
Suzanne Fitzpatrick, PhD DABT	FDA, OC, Silver Spring, MD
John R. “Jack” Fowle, III, PhD, DABT	EPA, OPP, Washington, D.C.

Kevin Gaido, PhD	FDA, CVM, Rockville, MD
David Hattan, PhD (<i>Chair</i>)	FDA, CFSAN, College Park, MD
Jerrold Heindel, PhD	NIEHS, OSTB, Research Triangle Park, NC
Huixiano Hong, PhD	FDA, NCTR, Jefferson, AR
Abigail (Abby) Jacobs, PhD	FDA, CDER, Silver Spring, MD
Devaraya Jagannath, PhD	FDA, CVM, Rockville, MD
Hajime Kojima, PhD	JaCVAM, NIHS, Tokyo, Japan
Annette McCarthy, PhD	FDA, CVM, Rockville, MD
Leslie McKinney, PhD	FDA, CDER, Silver Spring, MD
Catherine Richter, PhD	DOI, CERC, Columbia, MO
Jesudoss Rowland	EPA, Washington, D.C.
William Stokes, DVM, DACLAM	NIEHS, NTP, Research Triangle Park, NC
Julius Thigpen, PhD	NIEHS, Research Triangle Park, NC
Donald Tillitt, PhD	DOI, CERC, Columbia, MO
Thomas Umbreit, PhD	FDA, CDRH, Silver Spring, MD
Jon Wilkes, PhD	FDA, NCTR, Jefferson, AR

Invited Experts:

Joseph Haseman, PhD	Statistical Consultant, Raleigh, NC
---------------------	-------------------------------------

**European Centre for the Validation of Alternative Methods, ICCVAM EDWG
Liaison:**

Susanne Bremer, PhD	ECVAM, EC-Joint Research Centre, Ispra, Italy
Elise Grignard, PhD	ECVAM, EC-Joint Research Centre, Ispra, Italy

Public Attendees:

Public Attendee	Day		Affiliation
	1	2	
Patience Browne		✓	EPA, Washington, D.C.
Kent Carlson	✓		CPSC, Bethesda, MD
George Clark	✓	✓	Xenobiotic Detection Systems, Durham, NC
Abigail Jacobs	✓	✓	FDA/CDER, Silver Spring, MD
Susan Laessig	✓		EPA, Washington, D.C.
Raymond Lewis	✓	✓	Thermo Fisher Scientific, Pittsburgh, PA
Scott Lynn	✓	✓	EPA/OSCP, Washington, D.C.
Ruth McMillan	✓		Xenobiotic Detection Systems, Durham, NC
Thomas Umbreit	✓		FDA/CDRH, Silver Spring, MD
Kate Willett	✓	✓	People for the Ethical Treatment of Animals (PETA), Norfolk, VA
Cathy Yang	✓	✓	CertiChem, Inc., Austin, TX

NICEATM:

RADM William Stokes, DVM, DACLAM Director
 Warren Casey, PhD, DABT Deputy Director
 Debbie McCarley Special Assistant to the Director

Support Contract Staff— Integrated Laboratory Systems, Inc. (ILS):

David Allen, PhD Steven Morefield, MD
 Patricia Ceger, MS Michael Paris
 Jonathan Hamm, PhD Catherine Sprankle
 Linda Litchfield Linda Wilson

Abbreviations: CDER = Center for Drug Evaluation and Research; CDRH = Center for Devices and Radiological Health; CERC = Columbia Environmental Research Center; CFSAN = Center for Food Safety and Applied Nutrition; CPSC = U.S. Consumer Product Safety Commission; CVM = Center for Veterinary Medicine; DOI = U.S. Department of the Interior; EC = European Commission; ECVAM = European Centre for the Validation of Alternative Methods; EPA = U.S. Environmental Protection Agency; FDA = U.S. Food and Drug Administration; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; ILS = Integrated Laboratory Systems; JaCVAM = Japanese Center for the Validation of Alternative Methods; NICEATM = National Toxicology Program Interagency Center for Evaluation of Alternative Toxicological Methods; NIEHS = National Institute of Environmental Health Sciences; NIHS = National Institute of Health Sciences; NLM = National Library of Medicine; NTCR = National Center for Toxicological Research; NTP = National Toxicology Program; OC = Office of the Commissioner; OPP = Office of Pesticide Products; OSCP = Office of Science Coordination and Policy; OSHA = Occupational Safety and Health Administration; OSTB = Organ Systems Toxicology Branch; PETA = People for the Ethical Treatment of Animals.

TUESDAY, MARCH 29, 2011

Call to Order and Introductions

Dr. Vandenberg (Peer Review Panel Chair) called the meeting to order at 8:30 a.m. and introduced himself. He then asked all Peer Review Panel (Panel) members (those present and those in attendance via teleconference), Dr. Stokes, and Dr. Casey to introduce themselves and to state their name and affiliation for the record. Dr. Vandenberg stated that there would be ten public comment periods and he asked that those individuals interested in making a comment register at the registration desk. He requested that the public attendees hold questions and comments until the conclusion of the Panel's discussions. Dr. Vandenberg emphasized that the comments would be limited to seven minutes per individual per public comment session and requested that all comments should be brief and succinct. He deferred other introductions of those in attendance from the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), the ICCVAM Endocrine Disruptor Working Group (EDWG), Integrated Laboratory Systems, Inc (NICEATM staff), and members of the public until after the presentation of introductory remarks by Dr. Stokes.

Welcome and Opening Remarks from the NICEATM Director

Dr. Stokes of the National Institute of Environmental Health Sciences (NIEHS) and Director of NICEATM welcomed everyone to the Panel meeting being held on the main campus of the National Institutes of Health (NIH) in Bethesda, Maryland. He explained the NIH role as the nation's biomedical research agency and described the NIH budget and affiliated institutions. He stated that the NIH mission is to conduct science in the pursuit of fundamental knowledge about the nature and behavior of living systems, to apply that knowledge to extend healthy life, and to reduce the burdens of illness, injury, and disability.

Dr. Stokes stated that the NIH as well as the 15 agencies that are represented on ICCVAM greatly valued the expertise and input of the Panel on its review of the science behind these test methods. He said that ICCVAM is required by law to evaluate new methods and to make recommendations about their scientific validity, their usefulness, and limitations. The scientific peer review is incorporated as a critical and essential part of that evaluation process.

Dr. Stokes provided a brief overview of ICCVAM and NICEATM, and identified the 15 Federal agencies that comprise ICCVAM. He summarized the purpose and duties of ICCVAM (as described in the ICCVAM Authorization Act of 2000¹), noting that ICCVAM, as an interagency committee, does not carry out research and development or validation studies. Instead, ICCVAM, in conjunction with NICEATM, carries out critical scientific evaluations of the results of validation studies for proposed test methods to assess their usefulness and limitations for regulatory testing, and then makes formal recommendations to ICCVAM agencies. Dr. Stokes said that since the first recommendations were issued by ICCVAM in 1999, the committee and its members have reviewed and evaluated over 40 alternative safety-testing methods that have subsequently been endorsed by U.S. and/or international agencies. These methods, if they are used, can typically reduce, refine, and in some cases replace animal use for required regulatory testing. Most of these methods have been adopted as international test guidelines or incorporated into international guidance documents. NICEATM is a center of the National Toxicology Program (NTP) and is headquartered at NIEHS located in Research Triangle Park, North Carolina. It conducts and coordinates international validation studies and provides administrative and scientific support for ICCVAM. The NTP coordinates toxicology testing programs across the federal government.

¹ http://iccvam.niehs.nih.gov/docs/about_docs/PL106545.pdf

Dr. Stokes also stated that ICCVAM provides recommendations on how to advance new science and technology into standardized test methods by holding workshops. Experts are asked to provide advice on aspects that might improve or advance the methods that they review, so that the scientific community can move forward with methods that use few or no animals, can be conducted more efficiently, and can provide more predictive data.

Dr. Stokes defined basic validation as a determination of the usefulness and limitations of a test method for a specific purpose. Validation is more formally defined in ICCVAM documents as the process by which the reliability and relevance of a test method are established for a specific, defined purpose. Reliability is a measure of the extent to which a test method can be preformed reproducibly within and among laboratories over time, and relevance is defined as the extent to which a test method will correctly predict or measure the biological effect of interest. Adequate validation of a new test method is a prerequisite for consideration for use in regulatory decision-making by federal agencies. The law specifically states that agencies must determine the method to be valid for its intended purpose prior to requiring, recommending, or encouraging its use.

Dr. Stokes explained that the last step in evaluating the validity of a test method is independent scientific peer review. The proposed use of the test method must provide for equivalent or improved protection of human and/or animal health or the environment compared to the method that it is proposed to replace or be used in place of. Once validation studies are completed, then draft documents called background review documents (BRDs), as well as draft test method recommendations, are prepared by ICCVAM (through NICEATM). These are made available to the public and provided to the peer review panel. During the independent peer review an independent report is generated and made available to the public and to ICCVAM's advisory committee for comment. Finally, ICCVAM considers the Panel report, the comments from the public, and comments from the advisory committee and then develops a final test method evaluation report that is transmitted to federal agencies and, where appropriate, forwarded for international consideration.

Dr. Stokes concluded his opening remarks and thanked the Panel members for their commitment of expertise, time, and effort and acknowledged their important role in the ICCVAM test method evaluation process.

Overview of the ICCVAM Evaluation

Dr. Stokes provided an overview of the ICCVAM evaluation process for the validation study. He told the Panel that they would review the validation status of an *in vitro* endocrine disruptor assay (BG1Luc ER TA Test Method, hereafter known as BG1 method) to detect whether chemicals could interfere and interact with the estrogen receptor. He provided the audience with a brief discussion of how endocrine disruptor substances could interfere with the normal function of endogenous hormone signals, which can lead to abnormalities that have been shown in laboratory studies in terms of growth, development, and reproduction.

Process and Charge to the Panel

Dr. Stokes explained to the Panel that the BRD provided a comprehensive compilation of all the information and validation data supporting the validity of the BG1 method. The duties of the Panel included review of the BRD for its adequacy and completeness and then consideration of the draft test method recommendations and the extent that this documentation supported those recommendations.

Dr. Stokes reviewed the charge to the Panel:

- Review the draft BRD for completeness and to identify any errors or omissions
- Evaluate the information in the draft BRD to determine the extent to which each of the applicable criteria for validation and acceptance have been appropriately addressed
- Consider the draft test method recommendations and comment on the extent to which they are supported by the information and data in the BRD. Those recommendations

address the test method uses and limitations, recommended standardized protocols, recommended test method performance standards, and proposed future studies.

Dr. Stokes provided a brief timeline of the evaluation process. This process included publication of a public notice announcing a meeting and the availability of all the materials that have been provided to the Panel for review, implementation of Panel subcommittee meetings to determine initial draft positions, and reception of public comments for consideration. He explained that following the peer review meeting, the Panel would prepare and agree on a final report, and that Dr. Vandenberg would make a presentation at the June meeting of ICCVAM's Scientific Advisory Committee on Alternative Toxicological Methods (SACATM). ICCVAM and its Endocrine Disruptor Working Group (EDWG) will consider all this information, finalize its recommendations, and prepare the test method evaluation report for transmittal to federal agencies in the fall of 2011.

Dr. Stokes acknowledged the ICCVAM committee principal, alternate, and other representatives from the various agencies, as well as the participants on the EDWG from the various agencies. He also cited the international liaisons from ECVAM, the Japanese Center for the Validation of Alternative Methods (JaCVAM), and the NICEATM staff who organized the meeting under the leadership of Dr. Warren Casey.

Conflict of Interest Statement

Dr. Stokes reminded the Panel that they were meeting as an NIH special emphasis panel, under that charter, and he indicated that he would serve as the Designated Federal Official for the public meeting. He then read the conflict-of-interest statement and again asked members of the Panel to identify any potential conflicts for the record. Dr. Vandenberg asked the Panel members to declare any direct or indirect conflicts based on Dr. Stokes' statements and reminded the Panel that everyone had already signed a conflict-of-interest document.

Overview of Agenda

Dr. Vandenberg provided a general review of the agenda and outlined the process for reviewing each of the topics. Dr. Casey was to present the overview and background of the BG1 method and the validation study to the Panel on the first day of the meeting. Following that presentation, the Panel would break up into four separate groups with each group reviewing a specific area of the study. The leader of each group would then present the results of his or her group's discussion to the entire Panel when the whole Panel reconvened. All the panelists would have the opportunity to read the entire BRD, and those with specific areas of interest would have their comments discussed with the entire Panel. This would lead to a Panel discussion on the test methods, the validation of the data and the results, and the accuracy of the test method. The second day of the meeting would include Panel discussions on the test method reliability, the other studies that have come up in the interim, the animal welfare aspect, and other practical considerations. The Panel would also discuss the ICCVAM draft recommendations, the usefulness and limitations of the test methods, and potential future studies. Dr. Casey would present a summary of the BG1 method, and the Panel would discuss the test method performance standards. After each of these discussions and presentations, the public would have the opportunity to ask questions or make brief presentations.

Overview of the BG1Luc ER TA Test Method (LUMI-CELL® ER)

Historical Background

Dr. Casey provided the background for development of endocrine disruptor testing. He stated that the Federal Food, Drug, and Cosmetic Act sent a mandate to the U.S. Environmental Protection Agency (EPA) to develop a screening program, using appropriately validated methods or other scientifically relevant information, to determine whether certain substances may have an effect in humans that are

similar to an effect produced by naturally occurring estrogen. The Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) of the EPA recommended a two-tier screening program:

- Tier 1 – five *in vitro* and six *in vivo* tests, to identify substances with the potential to interact with the estrogen, androgen, or thyroid systems
- Tier 2 – a series of *in vivo* tests

In January 2004, Xenobiotic Detection Systems (XDS) nominated their LUMI-CELL® ER assay for a validation study to ICCVAM and in October of 2004, ICCVAM and SACATM considered the assay as a high priority method for validation. He described the prioritization of the *in vitro* assay based partly on the following factors:

- The method is faster and cheaper than any *in vivo* method
- A concentration-response curve is obtained from this method that is not always available in an *in vivo* study
- This method has an advantage over the currently existing *in vitro* method in that it has an agonist component and an antagonist component

Dr. Casey provided other prioritization criteria for *in vitro* test methods:

- They should be applicable to multiple agencies or programs
- They should be amenable to a high-throughput format as part of the Tox21 effort
- They should be applicable to the Organisation for Economic Co-operation and Development (OECD) conceptual framework, which is a weight-of-evidence approach to assess ED potential

He said that any assay that can be developed to help assess the endocrine disrupting potential of chemicals would add value to protecting human health. Dr. Casey noted that ICCVAM and NICEATM concluded that the test method is applicable to the criteria required by the agencies though the agencies are not obligated to accept.

Dr. Casey spoke on the potential of the BG1 method to reduce, refine, and replace animals. There is no direct replacement or refinement for the use of animals, but the BG1 method could be used as a substitute or an alternative test to the CERI STTA assay (Stably Transfected Human Estrogen Receptor- α Transcriptional Activation [STTA] assay validated by the Japanese company Chemicals Evaluation and Research Institute [CERI]). The BG1 method data showed that the assay has 100 percent specificity for detecting compounds that bind the estrogen receptors. Dr. Casey suggested that the BG1 method could possibly replace the direct uterine cytosol assay, which, although is an *in vitro* assay, does use animals as a source of ER. The BG1 method also has 100 percent sensitivity with the very small data set from the rat uterotrophic assay (i.e., any compound that tested positive in the uterotrophic assay was positive in BG1 method). Dr. Casey stated that the only validated ER TA method in use is the CERI STTA assay (an agonist only test), also known as OECD Test Guideline 455, which was directly adopted by the EPA for their EDSP program.

The BG1 method uses ovarian cancer cells that have both endogenously expressed ER-alpha and beta and cells that naturally have these receptors have the machinery in place to process signaling from those receptors. The BG1 method can identify antagonists, which creates the potential to identify a wider range of chemicals.

Validation Study Design

Dr. Casey said that a highly detailed standardized test method protocol was developed from October 2004 to October 2005 and ICCVAM recommended conducting an international multi-laboratory validation study. The study was organized by NICEATM, in conjunction with ECVAM and JaCVAM, and one laboratory was identified to perform the study in association each of the three respective centers. A four-phase validation study was initiated. Phase 1 generated historical data and

established that laboratories could adequately handle the reference standards and controls. Three replicate tests were run with four coded substances for agonism and four coded substances for antagonism during Phase 2a, the laboratory-qualification phase. Eight coded substances were tested for agonism and eight coded substances were tested for antagonism during Phase 2b, the qualification phase. Phase 3 included the testing of 41 coded substances one time each at all three laboratories. Only XDS participated in Phase 4 where 25 coded substances were tested. Additional retesting at XDS was completed in June 2010.

Background Review Document (BRD)

The BRD was drafted in 2010 (October through December) and reviewed by ICCVAM. The Federal Register (FR) notice for the peer review meeting was published in February 2011 and the BRD was made available to the public and the Panel.

Panel Discussion

Dr. Kelce initiated the Panel's discussion of Section 1 (Introduction) of the BRD for the BG1 method by noting that some typographical errors needed correction in the BRD and recommended that Phase 2 of the validation study should be known as the laboratory proficiency phase instead of the qualification phase. The Panel suggested including a more thorough discussion of the specific advantages of a transactivation assay relative to other *in vitro* ER assays such as providing advantages compared to binding assays since they examine estrogen agonist and antagonist activity, and detecting endocrine-active substances that potentially can act through non-receptor-mediated mechanisms. BG1 cells express both ER α and ER β and as such, possess the transcriptional machinery required for estrogen responsiveness.

BG1 cells have been stably transfected with four copies of the ERE and a luciferase reporter system. The Panel asked for clarification as to why there were four copies of the EREs and why the EREs are in their current orientation. Stably transfected cell lines have advantages over transiently transfected cell lines, including long-term utility without the need to transiently transfect cells each assay. The BG-1 test method has demonstrated responsiveness to estrogens and limited cross-reactivity with ligands of other steroid hormone receptors. Each of these important points should be emphasized in BRD.

Dr. Kelce asked whether ICCVAM's prioritization criteria and regulatory requirements were adequately discussed. The Panel was satisfied with the criteria discussion but agreed that the regulatory requirements need to be definitively detailed by their respective agencies worldwide. The specific regulatory use of data generated with this method has yet to be specifically defined (also applicable to the CERI STTA). Accordingly, it is essential that answers to the following questions be provided before making definitive conclusions regarding the usefulness and limitations of this assay:

- Is the BG1 method going to be added to the EPA Tier 1 battery?
- Will the BG1 method replace the STTA assay (if considered)?
- Will the BG1 method be used as a stand-alone screening assay for estrogen agonists and antagonists (i.e., replace the binding assay)?
- Will the BG1 method be developed into HTS screening assay?

The BRD should indicate that these issues are ill defined and that the agencies should provide input.

Dr. Borgert stated that ICCVAM criteria implied that all assays had to be validated for a specific purpose. However, the BG1 method has no defined use other than it is used in the same way as CERI STTA (which is also ill-defined as to the true use of the method). He said that one is trying to find substances that are ER agonists, and the definition of agonist is one that is positive in the CERI method, then circular reasoning is being used. He asked that more detail be provided to adequately address the validity of the method.

Dr. Borgert expressed concern as to how these assays would be used. He stated that the assays could not be adequately compared until the specific use of the test method is defined. He proposed the following questions:

- If a substance is positive for ER agonism, does that mean it has the potential to be estrogenic *in vivo*?
- If the assay cannot provide information on whether a substance will be active *in vivo*, what is the meaning?
- Do we believe that androgens can act at the ER and is that relevant?

Dr. Borgert said that his comments were not criticizing the effort, but suggested that the Panel was compromised by the lack of specificity by the intended use of the methods.

Dr. Yager stated that the assay includes a qualitative response (positive or negative) and a quantitative response. The positive or negative does not consider concentration response even if it is near 1 mM or 100 μ M, which would make it more or less irrelevant *in vivo*. He said that determination of an EC value misses what concentration is available to the cells. Cell culture conditions will affect absorption of compounds to different degrees, depending upon their structure and chemical properties and there is no proposal to determine what the free concentration is that is available to the cells. He stated that a second element is agonist versus antagonist and noted that the BRD states that *in vitro* transcription systems have the potential to detect antagonists. This is confusing and is another element of specificity. He asks if the compound is actually acting as an estrogen or whether the second element of specificity is antiestrogen.

Dr. Casey clarified that the CERI STTA assay has never been validated to detect antagonists, though in theory, it should be able to. He said that interpretation of these assays is more than just positive and negative. EC₅₀ values are collected but comparison of positive or negative to EC₅₀ values was not part of the validation study design. This study showed how the BG1 method related to the other transcriptional assay and to the binding assay and how EC₅₀ values generated in this assay correspond to published EC₅₀ values. The Panel agreed that more commentary is needed on the potency and dose-response since IC₅₀ and EC₅₀ data are collected according to the protocol.

Dr. Levine stated that the test system can also detect endocrine-active substances through non-receptor-mediated mechanisms and the EPA examines this in their matrix to detect potential activity. He accepted that the BG1 method could potentially detect non-receptor-mediated mechanisms and considered what types of diagnostic tools are available for such situations and how one would address that in the context of the screening program that many compounds are going through. Dr. Levine suggested that this issue and the accuracy and specificity aspects be discussed in the BRD.

Dr. Borgert said that EPA Test Guideline 890.1300 states that the aim of the transcription-activation assay is to evaluate the ability of a chemical to function as an ER ligand and activate an agonist response for screening and prioritization purposes but can also provide mechanistic information that can be used in a weight-of-evidence approach. He also said that if an estrogen agonist is being defined as one that is positive in this assay, then it is very much a circular reasoning process. The assay is validated for estrogen agonists that are defined by the fact that they produce a positive in this assay or in any transactivation assay and then the assay is deemed good because it identified the compounds that it identified. The real problem is there is specificity on what this means or how it will be used. Dr. Borgert requested that the document should state this conundrum. Dr. Kelce added that a positive result does not determine whether the result is ER-mediated or not. The only way to tell is to do a co-incubation with ICI. Another positive result will indicate that it is not receptor-mediated and this is the only way to know whether that satisfies that specific criterion mentioned in the CERI STTA assay. Dr. Borgert countered that if an estrogen agonist is defined as a compound that gives a positive response in a transcription-activation construct then he agrees with Dr. Kelce. However, he also said that if an estrogen agonist is actually a compound that is a ligand at estrogen receptors and

produces the cellular and physiological responses of an estrogen, then the two might not match exactly. This is observable with some of the compounds that produce a positive result in these assays, albeit at a potency so far removed from the natural ligand that they could never actually function as an estrogen in an organism.

Dr. Stokes stated that the test methods were screening assays that provide mechanistic or pathway-type information and that these types of assays will be prevalent in the future in the field of toxicology. He said that obtaining many nonspecific positive responses that are not ligand-mediated in a screening assay makes the usefulness of that assay questionable. He requested that the Panel provide insights on the likelihood that mechanisms other than a ligand interaction caused that positive response and how often this might occur. Dr. Borgert stated that the literature shows how the wrong ligand activates the receptor and can see that there are substances that legitimately are activating the receptor, but at concentrations that are not at all relevant.

The Panel agreed that the purpose and mechanistic basis of the BG1 method was adequately described with the caveats mentioned previously regarding vector construct and design. However, the Panel stated that the use of the proposed test method in an overall strategy of hazard or safety assessment is unclear. The BRD should indicate that this has yet to be defined. Additionally, the Panel asked that relevant regulatory agencies respond with answers to the previous questions to more clearly define this issue and suggested that the BRD should propose how these data should be used for safety assessment. Regulatory agencies have yet to define how individual assays within the Tier 1 battery will impact safety assessment.

Dr. Stokes stated that the draft recommendations provided to the Panel include proposed recommended uses and limitations for a specific use. The Panel can include information in the meeting report on demonstrated other uses or potential other uses that need further data, either from this test system or other test systems, to characterize that usefulness. Dr. Casey added that it would be very useful for the Panel to provide an assessment of the test method (e.g., is this method as good as the current test, should you get the same results, can a company use this to submit data to the EPA, is there enough data to replace the receptor binding assay). Additionally, recommendations on additional potential applications or suggestions for other validation studies would be welcome.

Public Comments

Dr. George Clark (Xenobiotic Detection Systems [XDS]) explained why the BG1 cells were used for the validation study. He stated that XDS wanted to develop a naturally responsive estrogen receptor in a human ovarian carcinoma along with four or five other different receptors to have a cell available that mimicked normal cells. He also said that the BG1 method has six logs of responsiveness and that you can obtain dose-response data and relative potency between the different systems. Dr. Clark also expressed that the EC₅₀ values were an important part of the validation because they could be used for dose setting and thereby reduce animal usage.

Dr. Kate Willett (PETA) spoke about the OECD performance-based test guideline concept of which the CERI STTA assay and, hopefully, the BG1 method will form the basis. This is a concept to expedite the validation of methods that are considered similar, i.e., they measure the same endpoint and use similar technology. In the review of the BG1 validation study, there will be the need to harmonize the two studies in terms of their specificity and sensitivity and output.

Additional Panel Discussion

Dr. Casey stated that he was on the OECD workgroup that was reviewing the performance-based test guideline concept and was aware of Dr. Willett's information.

The BG1Luc ER TA Test Method BRD Protocol

Dr. Casey presented an overview of the agonist and antagonist protocols that were developed for the validation study. He explained that the protocols were fine-tuned to be used by highly trained laboratory personnel because of the extreme level of sensitivity the test method is capable of achieving. Implementation of the protocols lends themselves to high variability. The agonist assay is a gain of function, meaning that when the ligand binds to the receptor, the receptor then binds to estrogen response elements and turns on the gene with an endpoint as a gain in signal. The antagonist assay is a loss of function where one starts with the cells in an induced state, increasing concentrations of test substance are added, and determination of whether there is a decrease in signal is the endpoint. The inherent problem is that cells that die produce a positive response and this is the number-one concern with this assay.

Both protocols have essential test method components: solvent controls, DMSO, reference standards (17-beta-estradiol for the agonist and raloxifene HCl for the antagonist) and two weak positive controls, which are several orders of magnitude in IC_{50} s or EC_{50} s from the reference standards. The maximum test substance concentration, unless otherwise limited by solubility or cytotoxicity is 1 mM for the agonist and 10 μ M for the antagonist. Seven concentrations at log10 intervals are tested in the range finder tests. EC_{50} or IC_{50} values are calculated using the four-Hill-parameter equation. Test acceptance criteria for the agonist assay include a minimum of threefold induction for the E2 standard. No data points are used where the visual determination of cell viability is less than 80%. The CellTiter-Glo ATP method for determining cytotoxicity shows that if you have less than 20 percent reduction in viability, then there is no effect on signal. The EPA steroidogenesis assay and the ISO9000 cytotoxicity test for medical devices allow visual observation of cytotoxicity.

Panel Discussion

Dr. Kelce led the discussion of Section 2 (Test Method Protocol) of the BRD. The Panel discussed the importance of a reliable cytotoxicity test. Agreement on the definition of cytotoxicity is difficult to achieve since reduction in cell numbers may be attributed to many factors, e.g., mitochondrial dysfunction, oxidative stress, and redox imbalance. The measurement of luciferase activity is problematic since the determination of cytotoxicity is subjective and only highly experienced laboratory personnel can make accurate determinations. The Panel suggested that a cross validation of the luciferase method and another cytotoxicity test method should be conducted.

The Panel noted typographical errors in the draft BRD, which are detailed in Appendix A and that although improvements to the protocol(s) during the course of the study were explained in the BRD, a better assessment of their impact on study results is needed to improve the document. The Panel agreed that the protocols appear complete and adequate in detail for a laboratory to conduct the study. However, some of the details in the protocols suggest requirements that should be more generic (i.e., less specific).

The Panel stated that the critical aspects of the test method protocol are adequately justified and described in the BRD. The use of visual assessment of cytotoxicity is subject to operator inconsistencies and may not always accurately reflect the viability of cells in culture. An accurate, objective cytotoxicity method is most critical for antagonist assays. Additional efforts should be undertaken to validate the utility of this approach for future use. The Panel suggested that a wider set of substances with known mechanisms of cytotoxicity should be tested and quantitative cytotoxicity methods are needed for developing new *in vitro* ER assays. The Panel agreed that the reference standards and controls proposed for the agonist and antagonist protocols are appropriate, but that future studies should consider including confirmation assays using a pure estrogen receptor antagonist (e.g., ICI 182,780). This would confirm ER binding behavior.

Test acceptance criteria were discussed and the Panel agreed that the plate acceptance criteria defined in the agonist and antagonist protocols appeared adequate. However the criteria for acceptance based on the DMSO controls could be too stringent and perhaps better criteria could be defined. The Panel also agreed that the proposed decision criteria for identifying a qualitative positive or negative response in the agonist and antagonist protocols were appropriate. However, the Panel suggested that potency and intrinsic activity relative to a known endogenous ligand are critical components in determining whether a substance is truly positive (or not) and this is not addressed in the current decision criteria. The Panel recommended inclusion of this component to improve the utility of the assay. Other recommendations by the Panel included 1) recording EC/IC₅₀ values (or other values, e.g., EC/IC₂₀) and associating them with the qualitative “+/-” call, and 2) implementing a relative potency approach by comparing test substances to known reference substances.

Public Comments

Dr. Clark stated that one can visually inspect a plate for cytotoxicity in less than 30 seconds and that other cytotoxicity endpoint assays take 15 to 20 minutes. He said that the test method is not a rapid, fast assay providing good data when using endpoints other than visual observations. He expressed that the test method is a screening assay and that this aspect is reflected in the protocol.

Additional Panel Discussion

Dr. Levine responded that the term “screening” applies to the endocrine disruption screening program, but people have to recognize that many of the screens are really tests and speed is not the most essential aspect of the test method. He said that assays currently conducted in the ED program take nearly a month or longer to complete for the end phase and they include extensive biochemical, histological, behavioral, and reproductive measurements. He believes the same level of rigor is needed in the *in vitro* assays as the *in vivo* assays and each assay is as important as the other in terms of data quality, data accuracy, and data reliability.

Validation Study Reference Substances

Dr. Casey presented background information on the validation study reference substances. He stated that NICEATM and ICCVAM performed an objective and thorough retrospective analysis of the data and updated the database (established in 2002) in 2010. The analysis used literature to classify reference substances that were well documented in peer-reviewed journals that consistently reported the call for the substances as one way or the other. In 2002, ICCVAM recommended a list of 78 reference substances to use to assess four different methods: estrogen and androgen binding and TA. Not all of the substances were intended to be estrogen receptor reference chemicals. There are certain compounds that should be used for reference classification of estrogen-active compounds.

Dr. Casey said that the definition of accuracy used for this study is the degree of closeness between a test method result and an anticipated reference. The literature citations used provided human ER TA assay results. If the call was positive in those assays in the literature, then the call was positive for the BG1 method. The intent of this study was to show that this assay is as sensitive as other ER TA assays that have been previously published. NICEATM reviewed all papers identified in the literature searches and thoroughly evaluated the quality of data presented.

Two criteria were based on the literature review (a minimum of two studies): a substance was labeled as a positive if it was reported as positive for activity in more than 50 percent of the studies; a substance was labeled as a negative if it was reported negative in all studies, and a minimum of two studies. Other substances were labeled presumed positive or presumed negative as a convenience for classification. The only substances we used in the study were the positive and the negative. Based on those criteria, 42 substances were identified for use in agonist accuracy (33 positive, 9 negative) and 25 substances for use in antagonist accuracy (3 positive, 22 negative). Dr. Casey suggested that one of the clear concerns regards the three reference antagonists, which is a reflection of the literature, and

may be a reflection of the real or not-real danger of these compounds. He said that it is difficult to find consistent data for antagonists and that reports presenting compounds as being positive are false positives because the controls were not run correctly. Almost all compounds become positive after testing substances at 10 μ M in these types of assays.

One of the criteria for reference substances is that there should be chemical diversification. The reference substances used in this study fall in many different categories and are a good representation of different chemical structures. All of the reference substances had well-referenced activity one way or another.

Panel Discussion

Dr. Levine addressed the notion that this study was a retrospective validation. He said that this is not entirely different, in many respects, from the CERi STTA validation that was a pre-validation and retrospectively evaluated as a validation. Dr. Casey replied that Dr. Levine's understanding of validation may be a matter of semantics, and that this study was not designed as a retrospective validation study but as a prospective validation study. There was a retrospective analysis of the results.

The Panel noted typographical errors in the draft BRD, which are detailed in Appendix A and also cited relevant information that should be included in this section that would improve the document. Specifically, the Panel asked that the BRD reflect how ECVAM ensured that the in-house ECVAM laboratory did not know the identity of the controlled substances. The Panel agreed that a broad range of different chemical classes and physicochemical properties were tested as part of this validation exercise. The list follows ICCVAM guidance and is largely applicable to chemicals and products that are screened to evaluate the potential for estrogenic activity. Several classes were represented by as few as one substance. It is questionable whether these classes are sufficiently represented as no conclusions on usefulness or limitations specific to these classes can be made. The Panel recommended testing more substances in both the agonist and antagonist protocols. Future testing should include compounds from under-represented classes and compounds that have surface-active properties. Also, identification of additional known negative compounds for agonist activity is necessary since less than 25% of the agonist substances used for the accuracy analysis are negative. It is difficult to investigate false positives in a new test system if the majority of test chemicals are positive.

The Panel concluded that the use of the majority classification among results to establish the consensus reference classification assigned to each reference substance is a reasonable strategy. The Panel stated that a 50% cut-off for definitive classification is not very strong. If the quality of the data used was additionally considered then it would strengthen the rankings. However, for assessment of a screening assay, where perfect performance could not be expected or required, this method is sufficient. The Panel noted that the BRD should provide better clarity on how data quality in the literature review was evaluated/considered and it would be useful to include a Klimsch code or similar approach to evaluating published literature data quality. The Panel also recommended consideration of additional sensitivity analyses that use something other than the majority classification ranking.

Public Comments

Dr. Willett stated that the validation study evaluates the reliability or reproducibility of the assay and the use of compounds with limited information is equal to testing an unknown with an unknown. She noted that *in vivo* validation assays are not dependent on this range of reference compounds. The default assumption is that the animal study is applicable to all chemicals, unless there is a reason for it not to be and the *in vitro* method is the opposite. Perhaps there are physical properties of the assay

itself that may limit its applicability. The disparity between the two approaches is stark and may be prejudiced.

Validation Study Data and Results

Dr. Allen provided information on the validation study data and results. He stated that all records are publicly available and all of the original data were submitted to NICEATM as Microsoft Excel® and GraphPad Prism® files. They are readily available electronically for anyone. The study consisted of four phases, ultimately testing all 78 substances in at least one laboratory and testing 53 substances in three laboratories (52 in the Japanese laboratory due to Phenobarbital being unattainable). He discussed the DMSO control results and the E2 reference standard data. Fold induction was calculated on agonist test plates by dividing the highest average RLU value from the E2 reference standard by the average DMSO control RLU. Differences in solubility testing only impacted three substances in the data set, where androstenedione, 2-sec-butylphenol, and fluoranthene had some discordant results among the laboratories because of the starting concentration that was chosen for comprehensive testing. Cell viability assessment was not evaluated when assessing discordance among laboratories.

Dr. Allen stated that Phases 2 and 3 included agonist testing for 53 coded test substances at XDS and ECVAM (31 and 33 positive results, respectively; XDS had 10 negatives and 12 inadequates; ECVAM had 13 negatives 7 inadequates). Phase 4 included 25 additional substances to round out the list of 78 tested at XDS (7 positives, 14 negatives, and 4 inadequates). Antagonist testing resulted in very few positives, many negatives, two inadequates at XDS and ECVAM, and one at Hiyoshi. Among the 25 additional substances there are 4 positives, 20 negatives, and 1 inadequates.

Dr. Allen outlined the data quality aspects of the study. Both XDS and ECVAM conducted their studies according to Good Laboratory Practices (GLPs). The Japanese laboratory (Hiyoshi) used procedures based on principles of GLP that are outlined in the ISO 9000 standards. Laboratory reports contained the QA statements and any findings associated with those laboratory audits, and the NICEATM project coordinators also served as a secondary QA of data that were imported into the graphical software.

Plate acceptance criteria were established based on results generated in reference standards and control wells. Protocol standardization in the first phase of the validation study, as well as Phase 2, where historical data were generated, included acceptance criteria based on the historical databases. Dr. Allen said that high plate failures, 61 percent for agonists and 38 percent for antagonists, occurred in the first part of Phase 2 and some were an indicator of poor-quality data. The plate acceptance criteria were reconsidered, particularly the agonist E2 EC₅₀ and the methoxychlor RLU control value, in addition to the antagonist raloxifene IC₅₀ and the flavone control RLU values. The qualitative and quantitative outcomes for test plates that met all acceptance criteria versus those that failed to meet one or more of the acceptance criteria were considered. Analyses showed that there were test plates whether they met criteria or not, where the answer was the same, i.e., E2 EC₅₀ and the methoxychlor RLUs, in addition to, for the antagonist assay, the raloxifene IC₅₀ and the flavone control RLUs. These were dropped as acceptance criteria and the changes were incorporated in the subsequent phases of the validation study. The impact of changing those acceptance criteria showed a reduction in percent failures for Phase 2a (61 percent failures to fewer than 30 percent failures). Similarly, Phase 2b plate failure rates for the antagonist assay were reduced by half. Phases 3 and 4 showed marginally decreased failure rates.

Substances were classified as positive, negative, or inadequate. Inadequate data were identified based on those substances that did not meet decision criteria for either a positive or a negative response. This classification was always due to poor-quality data that could not be interpreted because of major qualitative and quantitative limitations. Normally these substances would be retested, but this exercise was actually retrospective.

Panel Discussion (Data and Results)

Dr. Levine suggested bridging the CERI STTA and BG1 methods under a common protocol and to verify the performance of the assay and the performance standards that are in the BRD, perhaps slightly changed based on recommendations of the panel.

The Panel noted typographical errors, suggested edits, and identified points needing clarification in Section 4 of the BRD, which are detailed in Appendix A. The Panel agreed that the data for studies used to evaluate the accuracy and reliability of the BG1 method appeared to be provided. The interlaboratory studies of the BG1 method conducted by NICEATM, JaCVAM, and ECVAM have been included in the accuracy and reliability assessments provided in the BRD. The Panel noted that the NICEATM-ICCVAM question asking if all known relevant data for all studies used to evaluate the accuracy and reliability of the BG1 method had been provided to the Panel is misleading. The Panel questioned if sufficient data from other screens that add to the support of a particular finding have been included, i.e., has there been enough consideration of results from ER binding, uterotrophic, etc. to address accuracy.

The Panel suggested that potency evaluation and relative comparisons of potency to reference substances would be useful and that suggested statistical tests of trend be included in the evaluation of a positive call. The Panel recommended that the test results should include: characterization of activity in addition to a positive/negative call, identification of the scale of measurement needed to compare tests, and evaluation of descriptive endpoints vs. continuous. Users of the test method should consider diagnostic testing as an example of comparing results from multiple tests. The Panel agreed that evaluation of the data resulting from the BG1 method requires further attention because analyses described often involve some transformation of the response relative to control responses and the variability in the control responses appears to be ignored in these constructions. The assumption of no downturn in the dose-response models implies that some preprocessing of the data points occurs to remove values that violate this pattern.

The Panel recommended that the criteria for an acceptable concentration response should be developed to only allow data that is sufficient for concentration response modeling. The current criteria of three points with non-overlapping standard deviations (SDs) are not sufficient. The Panel recommended additional discussion of the 4-parameter model to justify its use. The Panel stated that interpretation of the results should not rely solely on statistics but also on scientific judgment and should incorporate consideration of the nature and shape of the dose-response relationship and, if needed, the reproducibility of the response in independent experiments. Criteria should be established for acceptability of data to estimate potency values. The estimated values for EC₅₀ and IC₅₀ are presented as point estimates without any error being associated with them. There is uncertainty associated with these estimates, and it should be reported (e.g., confidence interval).

The Panel agreed that there was adequate documentation showing that coded substances were tested and experiments were conducted without knowledge of the identity of the substances being tested. This was critical for the evaluation of reliability and accuracy.

Public Comments

Dr. Willett stated that similar issues were discussed when interpreting the results from the CERI STTA validation study. The committee implemented a relative potency index that related the EC₅₀ of a test chemical to the EC₅₀ of 17-beta-estradiol. Instead of getting just a positive/negative result, there was a definite spread that could be divided into classes.

Panel Discussion (Data Quality)

The Panel noted typographical errors in Section 7 (Data Quality) of the BRD, which are detailed in Appendix A. The Panel suggested that the BRD should include: availability of audit results, statistical

evaluations, and methods for calculations; a time period for requiring this data be stored; recommendation of a secondary location for data backups.

The Panel agreed that the extent of adherence to national and international GLP guidelines for all submitted *in vitro* and *in vivo* test data and the use of coded substances and coded testing was adequately presented and that any deviations and alterations to original protocols including use of coded substances were relatively well described. However, formal training records of laboratory personnel were not available to confirm compliance with GLPs and Good Cell Culture Practices (GCCPs).

The Panel suggested that the BRD should include a description of the quality of the reference data, i.e., data from CERI STTA, data used to generate the ICCVAM reference classifications, data from the uterotrophic assay, and data from the rat ER binding assay. Consideration should be given to assigning greater weight to certain reference data, and a thorough description of the weighting methodology should be provided.

The Panel recommended that the responsible personnel for the quality assurance (QA) aspect of the study should be identified in the BRD. It was not clear to the Panel whether all appropriate QA steps have been performed. All audits should be documented in an appendix that is part of the BRD. The Panel asked that the BRD include additional explanation about the aspect of the updated classification system being developed after testing was complete and substances not being retested.

Test Method Accuracy

Dr. Casey provided background information on test method accuracy and the interpretation of results. A retrospective analysis was performed due to the way the decision criteria were originally laid out. The criteria were intended to estimate an effective lowest observed concentration based on a dose-response curve. The curve reaches a point that is statistically different than control, which equals a difference in response. The criterion for that threshold was three times the standard deviation of the DMSO control (three times above the DMSO baseline). Theoretically, the signal is above background. The characteristics of this type of classification system are that any value above that threshold constitutes a positive response. Any value that is statistically different than the DMSO control was a positive. If all the values were below that line, the test substance was a negative. There are some curves that cannot be classified because they are really close and those are called equivocal. Many of the curves should probably not have passed quality control, but the protocol did not allow the laboratories to exclude those. There was no quality metric around the data themselves. All the quality metrics were set around the DMSO controls, with the assumption that once the DMSO controls were adequate that should imply good-quality data for the rest of the plate.

Dr. Casey explained how the new decision criteria were determined. The two issues that were driving this issue were the high background and the variability with the data. Historical data were reviewed and the commonalities for clearly positive test results were evaluated. That evaluation showed that a general S-shaped curve with three non-overlapping data points on the slope was the common aspect. This was applied to all the other data, even the data that have high variability and if the substances were not positive by these criteria or negative by the other criteria, then were inadequate and the call could not be made. The criteria that were determined for what a positive should be in this assay is a dose response: positive slope, a peak or a plateau, and three data points with non-overlapping standard deviations. The other criterion, an amplitude of 2000 (20 percent of the E2 reference), is associated with the amount of noise at the baseline. Anything below that is negative response.

Panel Discussion

The Panel recommended that the primary comparison for the BG1 method should be the accepted reference method (CERI STTA) and this comparative analysis should be in Section 5.0 of the BRD. The Panel agreed that the ICCVAM reference consensus classification is an excellent additional

reference method for assessing the accuracy of the BG1 method and is essential due to deficiencies in the CER1 data.

The Panel stated that there is no definitive way to assess the performance/accuracy of a screening assay. Therefore, this kind of novel comparison is useful, and the comparative results are good enough. The endorsement of a screening assay cannot be based on strict correlation to any one other method, but should be judged on a weight-of-evidence approach that includes all performance comparisons, the repeatability results, and scientific judgment regarding the biological relevance of the test system.

The Panel provided additional commentary on accuracy. BG1 method accuracy should be discussed in terms of how the BG1 method performance statistics compare to previously endorsed *in vitro* test methods. Demonstration of “agreement” between the two data sets (e.g., comparison of EC₅₀ and IC₅₀ values with reference data using regression) should use more appropriate methods than regression (e.g. Bland-Altman Plots and Limits-of-Agreement). Accuracy analyses should be based on results from individual labs along with a consensus classification determined based upon a “majority approach” using the three testing laboratories.

The Panel stated that a validation based upon unequivocal classifications might result in overly optimistic assessments of test method performance. The accuracy of the method in the validation should use unequivocal reference classifications from multiple test systems that have passed specific data quality measures, as was performed. This increases confidence in the results by eliminating questionable responses, positive or negative. The “accuracy” assessment for a screening assay can only be approximated, especially one without sufficient comparative data from a similar assay. Substances that result in equivocal reference classifications may provide additional insights in to aspects of the test method. Comparing the new test method with other methods is reasonable and indices of accuracy should be calculated only with data that meet certain quality control measures. If the data included in the ICCVAM classification do not meet these criteria they should not be used. Reference method data quality is an uncertainty for all of the reference methods used in the BRD because reference data quality is not provided in Section 7.0. Therefore, this criterion should not be imposed unfairly on these additional analyses. Use of data from other methods for comparison could be questioned since they may not be truly measuring the same biological impact. Accordingly, emphasis should be placed on describing the different purpose of the various tests, their advantages and disadvantages (the document should include a single section to provide these detailed explanations).

The Panel stated that assessment of biological impact is an admirable goal, but mechanisms are not fully understood. The indices of accuracy for assay performance have been used in all previous validation study analyses, and the BG1 method should not be subjected to different criteria. For example, the concordance of the BG1 method with the rat ER binding assay exceeds that for other assays endorsed as scientifically valid by validation authorities.

The potential for the assay to identify a greater portion of false positives (agonist testing) or false negatives (antagonist testing) than indicated in the results should be mentioned (due to low number of negative test substances). This is not seen as any reason for not endorsing the method for use in a test battery or other weight-of-evidence approach, but is important for users to understand. The contingency tables used to generate the summary statistics should be included in the document.

The Panel recognized that the original test criteria were inadequate and revisions had to be made throughout the study. Although this should not be a precedent going forward (optimal decision criteria should be selected *a priori*), the Panel acknowledged that changes made were performed in an appropriate manner.

The Panel agreed that the BG1 method is effective for generating data, in an amount and of a scientific value that is at least equivalent to the data generated from existing tests (i.e., U.S. EPA Endocrine Disruptor Screening Program Test Guideline OPPTS 890.1300: Estrogen Receptor Transcriptional Activation [CERI STTA]). The Panel concluded that the BG1 method protocol is an improvement over the CERI protocol given the extent to which the BG1 method protocol was standardized and optimized. Additionally, the Panel stated that ER binding assays indicate high concordance and therefore suggest that the outcomes of the stably transfected ER TA assay can provide reliable information about the biological effect of chemicals mediated by ER-ligand interactions. However, there appear to be sufficient data to consider replacing the rat uterine cytosol ER binding assay with the BG1 method. The Panel recommended that NICEATM and ICCVAM identify and use other available rat uterine cytosol ER binding data for these comparisons (i.e., substances beyond the 78 tested in the BG1 validation study). As an additional activity, the Panel recommended an evaluation of recombinant (human and animal) ER binding assays as a replacement for the rat uterine cytosol ER binding assay note.

Public Comments

Dr. Lynn congratulated the Panel for the expertise and diversity represented by the members. He stated that the rationale and discussion about the qualifications and the guidelines was well done.

Dr. Laessig addressed the Panel's discussion on whether or not to recommend replacing one assay with this one. This assay in particular is very oriented toward looking at human health because of the way the system has been designed. But the EPA screening battery also needs to take into account ecological effects. Some assays may be redundant but could provide useful information both human and ecological effects.

Additional Panel Discussion

The Panel continued general discussion of the various Panel-generated answers to the questions provided by NICEATM and ICCVAM concerning the information in the BRD. The Panel approved all of the answers subject to general wordsmithing.

Adjournment

Dr. Vandenberg adjourned the Panel for the day at 5:25 p.m., to reconvene at 8:30 a.m. on Wednesday, March 30, 2011.

WEDNESDAY, MARCH 30, 2011

Call to Order and Introductions

Dr. Vandenberg called the meeting to order at 8:30 a.m. and asked Panel members and other participants to introduce themselves. Dr. Stokes again read the conflict of interest statement and reminded the Panel that each person had signed the appropriate form.

Test Method Reliability

Dr. Allen provided a synopsis of the reliability section (Section 6) of the BRD. He defined reliability as a measure of the degree to which a test method can be performed and is calculated by intra- and inter-laboratory reproducibility and inter-laboratory repeatability. Reliability is based on the reproducibility of the reference standards and controls of the reference substances themselves within and between laboratories.

Dr. Allen discussed the intra-laboratory reproducibility of reference standards and controls, i.e., the agonist DMSO control RLU values and the antagonist DMSO control and E2 control RLU reference values. He explained that 12 substances were tested three times at all the three labs in Phase 2 and that there is 100 percent agreement within each laboratory for each of the three repeat tests for both agonists and antagonists. Similarly for inter-laboratory reproducibility, there was data from Phases 2 and 3 for substances that were tested in all three laboratories. Only those substances that produced a definitive result in at least two of the three laboratories were used for the reproducibility analyses. In agonist testing, two-thirds of the laboratories got the same answer with regard to positive-negative calls. There was 100 percent agreement across all labs for antagonist testing. For Phase 3, 83 percent of the substances had agreement across the laboratories in the agonist phase. For antagonist testing, most of the substances tested produced agreement among the laboratories, with only three instances showing discordance.

Panel Discussion

Dr. Mihaich initiated the discussion on test method reliability (Section 6 of the BRD) and presented the Panel with the questions proposed by NICEATM. The Panel concluded that not enough data were presented to ensure that a thorough analysis of intra- and inter-laboratory reproducibility had been adequately evaluated. Many of the figures in the BRD do not adequately present the variability associated with the test method and in fact may confuse variability with outliers and these sources of variability should be explored using appropriate methods, e.g., comparison of CVs and formal statistical analyses for evaluating variance components. The Panel stated that reproducibility had been addressed on three data levels: raw data (e.g., DMSO control measurements), derived endpoints (e.g., EC₅₀), and prediction of estrogenic activity. The analysis is descriptive but no statistical testing was performed which avoided sample size, power issues, and definition of equivalence margins. The Panel agreed that while this approach is acceptable, there are some issues that require additional discussion:

- Criteria used for determining what is an appropriate way to characterize intra- and/or inter-laboratory variability could include comparisons to established test methods that have been considered acceptable.
- Variability of EC₅₀ estimation needs to be summarized in Section 6 of the BRD. Summarizing EC₅₀ values from different compounds using plots, summary statistics, or agreement measures would provide further insight into the overall variability and reproducibility of the assay, e.g., compare the BG1 method reproducibility to reproducibility of similar test methods (e.g., CER1 STTA).
- CER1 STTA analysis of variance components provided components that could be useful for a direct comparison. This includes calculating the mean and standard deviation of

logEC₅₀ to complement the current analysis and to allow for a more direct comparison to the analyses of other assays that also included these calculations (e.g. CERI STTA).

- Sensitivity of analyses to assumptions that are made should be investigated.
- Assessment of reproducibility might be too optimistic by excluding substances from reproducibility analysis for which definite results were not determined.
- Rationale for selection of the substances used to evaluate intra- and inter-laboratory reproducibility should be more clearly described in the BRD.

Public Comments

Dr. Clark stated that there are variants with unknowns that you are trying to assay, and it is unknown exactly how the biological mechanism works. Members of the Panel discussed further variants such as DMSO controls, reference standards, and EC₅₀ values.

Additional Panel Discussion

Dr. Stokes added that ICCVAM has not published strict criteria that define what is acceptable reproducibility between and within a laboratory. There is inherent variability in test systems and variability due to differences in laboratory operations. The goal is to determine whether the variability is reduced sufficiently to get the same type of response back even with that background variation. ICCVAM does not define what is acceptable and not acceptable but wants to know how variability impacts the outcome of the assay.

Other Studies

Dr. Casey stated that Section 8 of the BRD provided summaries of other publications that related to the BG1 method. He said that numerous chemicals have been tested in this cell system, not necessarily with this particular protocol, in addition to the 78 chemicals that have been tested in this study. Although there are no published reports that this assay is being considered for validation, the test method is being evaluated in the Tox 21 effort (the high throughput testing 1536-well plate format) at NIEHS. Researchers are obtaining a threefold induction, and feel that the test method is highly transferable from 96-well plate format to the high-density plate format.

Panel Discussion

Dr. Borgert initiated the discussion on the BRD's presentation of other studies and presented the conclusions of the Panel. The Panel unanimously agreed that all the relevant data identified in published studies that employ this test method have been adequately considered. The Panel suggested obtaining QSAR-based predictions for ER binding from the literature on some of the validation study chemicals to evaluate a comparison to the BG1 and rat ER binding results.

Public Comments

There were no public comments.

Animal Welfare

Panel Discussion

Dr. Borgert provided the leadership on the discussion of animal welfare. The Panel concluded that further discussion was required to determine the extent to which the test method will reduce, refine, or replace animal use. In order to fully understand how this method will impact the 3Rs, there needs to be a better overview of the EPA EDSP Tier I screening battery and the proposed context into which this test method will fit in terms of the overall testing scheme. Additional discussion of the following topics would provide additional ways that a validated BG1 method could contribute to reducing animal use:

- Determine whether the BG1 method will be performed prior to- or simultaneously with- the other *in vitro* and *in vivo* assays
- Define the implications for possible Tier II testing, which will likely incorporate significant animal use

The Panel stated that comments provided in the EPA's 2011 budget state their intent to phase out the EDSP and to include high-throughput ED-detecting assays in the ToxCast screening battery, which should lead to reductions in animal use for regulatory testing in the long term. The BG1 method will probably contribute to the development of this screening battery. The Panel also determined that a major problem with *in vitro* ED testing and the EDSP screening battery is the lack of an *in vitro* method to evaluate metabolism. The inclusion of a metabolism component as part of any *in vitro/in silico* ED test battery will provide a more biologically relevant assessment of ED activity. Since there are many *in vitro* metabolism methods being used in drug development and being used for *in vitro* testing of other toxicity endpoints (e.g., genotoxicity), the importance of including metabolism as part of all future studies on *in vitro* ED assays should be mentioned in Section 9.0 of the BRD.

The Panel was concerned that implications of BG1 method validation in the EU and Japan were not mentioned in the BRD. ED chemicals are substances of very high concern (SVHC) in the REACH testing program, and validated *in vitro* methods have the potential to reduce animal use in these kinds of large testing programs. REACH-type programs are also being adopted by Asian countries, so the availability of validated *in vitro* and *in silico* methods to screen/prioritize chemicals for these testing programs has the potential to reduce animal use. Additionally, following validation, the development of *in vitro* ED assays into an OECD TG will broaden their potential for reducing animal use. The Panel agreed that concordance of the BG1 method with the rat cytosol ER binding assay of 97% and based on 33/34 test substances suggests the BG1 method can "replace" the rat ER binding assay within the EDSP Tier 1 battery (and similar test batteries). The Panel believes that the excellent concordance of the BG1 method exceeds the "performance" of other methods that have been endorsed as scientifically valid. Additionally, an assessment based on 34 test substances could provide sufficient confidence (power analysis could be conducted to confirm).

The Panel stated that the excellent concordance of the BG1 method with the rat uterotrophic assay (92%, 12/13) indicates that the BG1 method is an excellent candidate assay for replacing the uterotrophic assay, and thereby reducing animal use. The small data set, however, is not sufficient to recommend endorsement of the BG1 method as a replacement at this time. Therefore, the Panel could recommend the BG1 method as a high priority for additional studies. Retrospective analyses may be sufficient. If necessary, a prospective study could be conducted to further compare these methods, preferably by identifying additional materials already evaluated in the uterotrophic assay. *In vitro* metabolism must be included as part of the prospective component of this study. Also, since the BG1 method was already subjected to an extensive interlaboratory study, consideration for an abbreviated assessment (e.g., 1-2 laboratories) should be considered to reduce time and costs.

The Panel provided the following text as a possible concluding paragraph for Section 9 of the BRD:

The development of a battery of *in vitro* and *in silico* methods that can totally replace animal testing for detecting chemicals that have the potential to cause an adverse effect interact with the endocrine system is a biologically complex problem. For example, a method for the assessment of metabolized test substances need to be included with the *in vitro* assays, and assays for assessing the many modes of action of EDs on various tissues and species need to be developed and validated. The experience derived from validating and using the *in vitro* BG1 method is expected to contribute to our knowledge and promote progress toward this goal. It should lead to the broader use of cell-based methods for ED screening, and could include the use of cells from other species.

Public Comments

There were no public comments.

Practical Considerations

Dr. Casey stated that some of the practical considerations of the test method are 1) it should be considered as part of a weight of evidence approach even though it does not replace animal use, 2) it is a possibility that the test method may replace the rat uterine cytosol assay, and 3) it is extremely important to have well trained personnel that operate under a GLP environment.

Panel Discussion

Dr. Borgert led the Panel discussion on practical considerations and provided a list of practical considerations that should be addressed:

- Provide better clarification regarding the availability of the BG-1 cell line which has not been placed in a public repository; efforts should be undertaken to do so or to otherwise ensure the continued availability of the cell line.
- Employ a less subjective cytotoxicity assessment than the current visual method to improve transferability and implementation of the method by laboratories.
- Associate costs of equipment and supplies with the date they were acquired.

The Panel agreed that the level of training and expertise required to conduct the test method are reasonable for its wide use and that the protocol should recommend the need for adherence to GCCPs and GLPs in order to minimize variability.

Public Comments

Dr. Clark stated that one of the reasons that XDS developed the test method was to get dose response data that could be used to set doses in animal studies and reduce the number of animals used in range finding tests.

Dr. Willett stated that it might be appropriate for the Panel to make recommendations for improving the use of this BG1 method regardless of how it compares to the rat uterotrophic, assay, for example. One of those recommendations might be to explore the use of different metabolism systems with the BG1 method. This has implications more broadly for *in vitro* assays and might be included in recommendations for future studies. Additionally, considerations for measuring the actual concentration could be part of the recommendations. She also suggested that general recommendations that the Panel could make regarding changes in future validation exercises that might facilitate the validation of these types of assays. She acknowledged this kind of validation exercise cannot be physically be done for all of approximate 300 ToxCast assays but any recommendations in terms of improving the evaluation process would be very helpful.

Additional Panel Discussion

Panel members discussed the use of the acronym EDs (i.e., endocrine disruptors) and asked whether a new acronym should be used (e.g., EAC, endocrine active chemical).

Dr. Jacobs said that the use of the term ED and endocrine disruption for assays, which only assess possible interaction with the system, has caused the FDA problems. They should be called potentially interactive substances. She does not believe that endocrine disruptors and an adverse *in vivo* effect can be defined in an *in vitro* assay.

The BG1Luc ER TA Test Method Draft Recommendations on Usefulness and Limitations

Dr. Casey presented a restatement of what was contained in the BRD and listed the ICCVAM recommendations.

- ICCVAM proposes that the BG1 ER TA test method can be used as a screening test to identify substances with *in vitro* agonist activity. This use is based on an evaluation of available validation data and corresponding accuracy and reliability. ICCVAM concludes that the accuracy of this assay is at least equivalent to the current ER TA included in the regulatory testing guidance.
- ICCVAM proposes that the BG1 ER TA test method can be used as a screening test to identify substances with *in vitro* antagonist activity. This use is based on an evaluation of available validation data and corresponding accuracy and reliability.
- ICCVAM recommends that the protocols in the BRD that the Panel has reviewed should be used in any further testing with this cell system.
- There are some of the limitations that have been identified: test substances must be soluble in DMSO, cannot react adversely with DMSO or cell culture media, and cannot have endogenous luminescence or naturally inhibit luciferase activity.
- ICCVAM states that the concordance of this assay is similar with that of the rat uterine cytosol assay and the BG1 ER TA test method has potential replace the rat uterine cytosol assay.

Panel Discussion

The Panel agreed that based on the overall test method performance, as presented in the BRD, the BG1 method is capable of identifying substances as potential *in vitro* ER agonists and antagonists, provided the acceptance criteria described in the recommended test method protocol are met. In addition, accessibility to the cell line still needs to be resolved, which clearly can impact the future use of this assay, including as a screening test in a contract laboratory setting.

The Panel suggested that additional analyses could be performed that would strengthen the understanding of how well this screen performs compared to the current CER1 STTA assay and within the battery as a whole. It is important that all efforts have been made to make use of the data that are available before additional laboratory work is conducted. The Panel agreed that the assay can be used as a screening test to identify substances with *in vitro* estrogen agonist properties and can be used in place of CER1 STTA for regulatory testing but recommended additional reliability analyses (which may lead to revising draft ICCVAM recommendations pending their outcome).

The Panel also agreed that the assay could be used as a screening test to identify substances with *in vitro* estrogen antagonist activity although error rates may not be precisely estimated in the antagonist assay nor necessarily representative of the population of chemicals that may be tested.

Public Comments

There were no public comments.

The BG1Luc ER TA Test Method Draft Recommendations on Future Studies

Panel Discussion

Dr. Casey presented the Panel's recommendations for future studies. He stated that the Panel agreed that the available data supported the ICCVAM draft recommendations for the BG1 method in terms of the proposed future studies. In addition to the proposed studies, the Panel recommended that additional future studies should include:

- Efforts to validate the utility of the current cytotoxicity evaluation

- Searches for fully-defined media to replace FBS
- Attempts to account for compound metabolism/activation
- Conversion of the BG1 method (+/- metabolism) into a HTS assay format for use in ToxCast and other screening programs

The Panel stated that the concordance of the BG1 method with the rat cytosol ER binding assay suggests that the BG1 method and binding assays produce similar results. Regulatory agency clarification is needed to determine if both assays are necessary.

Public Comments

There were no public comments.

Additional Panel Discussion

Additional discussion included comments on who would actually perform future studies. Dr. Stokes stated that any recommendations relevant to advancing the usefulness and limitations of this assay are appreciated. Who carries out that work depends on who has the resources, as well as if it is a federal agency, where they have the statutory authority to do such work. Research agencies in the federal government can carry on R&D type work, but that does not preclude a non-government organization from performing such work as well. He said that the BG1 cell line is currently available through a commercial entity. Performance standards were developed for evaluating commercially available methods or those with some intellectual property rights associated with them. For agencies to endorse a method that is copyrighted, trademarked, or sold by a commercial firm, the basis by which that is considered acceptable must be stated so that anyone else can create a similar model and know what criteria it has to meet to also be considered acceptable. It is important to understand that perspective on the availability of the test. This method has been published in the open literature as to how the cell line was created, so that others could do that if they wanted to.

The BG1Luc ER TA Test Method Performance Standards

Dr. Casey stated that the performance standards are used to evaluate the accuracy and reliability of other proposed test methods, sometimes referred to as me-too methods. If they can meet these criteria, then they are considered functionally equivalent to the BG1 assay. The performance standards include the essential test method components, all critical aspects of the assay, a list of reference substances to be tested, and a defined set of accuracy statistics or of metrics to use for accuracy and reliability to insure that another method is the same as the current method.

The essential test method components include: a human ovarian cell line, stably transfected with luciferase; an appropriate solvent; test substance concentration up to one millimolar for agonist testing and up to ten micromolar for antagonist testing; a minimum of seven concentrations at log ten intervals; an evaluation of cytotoxicity; the use of reference standards, weak and positive controls, and solvent control.

He said that the interpretation of results is not a statistical approach but an empirical approach. If a positive curve is observed, then the EC₅₀ should be calculated if possible. The criteria for the antagonist tests include testing a ten-micromolar limit concentration and performing an IC₅₀ calculation if possible.

These are 34 proposed reference substances and all should be tested as coded materials. Performance of the BG1 method requires significant training to become proficient. In order to demonstrate equivalence, test as many chemicals as possible. All laboratory personnel should treat every chemical like it is E2 to avoid contamination issues and high variability. Discordant results with a new test compared to the BG1 method should be discussed in terms of the ability of the test methods to detect a similar range.

Panel Discussion (Performance Standards)

The Panel assessed the adequacy of the performance standards adequate for assessing the accuracy and reliability of test method protocols that are based on similar scientific principles and whether those methods measure the same biological effect as the BG1 method. The following concepts were proposed by the Panel and should be addressed in Section 1 of the BRD.

- The ‘intended purpose’ for the assay in the context of the screening battery has not been directly defined. The lack of clear purpose for the assay has profound implications for the setting of performance standards. Without a clear purpose, only provisional performance standards can be set. The detailed technical comments of the Panel should be understood in the context of this provisional nature. In turn, issues related to cytotoxicity concerns and assay performance, including the number of known negatives evaluated, are intimately related to the setting of performance standards and must be understood in the context of the provisional nature of the validation effort.
- A statistically significant difference between control and treated is an inadequate delimiter of agonist response. This is readily apparent from the chemicals deemed to produce positive responses. For example, both the 890.1300 assay and the BG1 assay are said to produce a positive response with methyltestosterone. Does this mean that methyltestosterone is to be considered an estrogen agonist, a ‘potential’ estrogen agonist, or something else? Or, is some distinction being made between “... the ability of a chemical to function as an ER α ligand and activate an agonist response, ...” and “... the ability of a chemical to function as an ER α ligand and estrogen agonist.” In other words, has the validation effort made a distinction between a chemical ‘activating an agonist response’ versus actually functioning as an agonist? If so, what is the qualitative and quantitative relationship between a chemical that activates an agonist response and one that is actually an agonist? These questions need clarification before any meaningful validation can ensue.
- For example, the 20% cutoff for cytotoxicity might be a good standard if the purpose of the assay were to identify chemicals with potential to activate an agonist response, and an agonist response in this assay had been defined as intrinsic activity and potency at the ER α at least 25% and 5%, respectively, relative to 17 β -estradiol (percentages not to be taken as a recommendations, but merely for the sake of interjecting sufficient detail to make the point clear). With that type of clear, detailed definition of “agonist response,” it would then be possible to determine whether the performance criteria for the assay were adequate for assessing the accuracy and reliability of the test method.
- Intrinsic activity is the ability of a drug-receptor complex to produce a maximum functional response. Intrinsic activity is sometimes used interchangeably with efficacy; however, intrinsic activity refers to a cellular response whereas efficacy is more often used in the context of a clinical response. Assuming equivalent pharmacokinetic parameters and affinity, a drug with greater intrinsic activity would be expected to have greater efficacy. Affinity is a measure of how tightly a drug binds to a particular receptor, and is often defined by the dissociation constant. Potency is the intensity of effect produced per unit of drug, and is a function of intrinsic activity and affinity.
- Neither the BG1 nor 890.1300 assay measures a complete agonist response, defined as a chemical that binds to an estrogen receptor in a cell and triggers an estrogenic response by that cell. Instead, these assays measure the first two steps of an agonist response at the alpha subtype of the estrogen receptor (ER α) via an artificial construct that couples binding and activation of the receptor complex to a reporter construct, for the BG1 assay, the enzyme luciferase. Activation of luciferase is not a normal physiological response of estrogens; hence, this step is not considered a component of estrogen agonist activity. Because neither the BG1 or the 890.1300 assay measures a bona fide agonist action,

validation would require defining both qualitatively and quantitatively the relationship between the stated purpose of the method and the activity actually measured by the assay. This relationship has not been characterized, quantitatively or qualitatively, for either the BG1 or the 890.1300 series assay.

Panel Discussion (Essential Test Method Components)

The Panel determined that the justification for some of the essential test method components is not clear and recommended clarification of the following points be addressed in the BRD:

- The most critical point for the type of cell line used is that it should include the appropriate “machinery.” The specific tissue source, cell type, and species may not be critical.
- Explain why the maximal concentrations were suggested.
- Explain why log10 spacing of seven concentrations is needed.
- Address the evaluation of cytotoxicity and how modeling of the possibility of such a response could also be considered.
- Explain why 20% of maximal response is the cut-point for significant response.
- Address how a test of negative trend might be preferred to the suggestion of “three points with non-overlapping error bars.”
- Define what the error bars correspond to (e.g., standard deviations, standard errors, half-width of confidence intervals). Use of non-overlapping confidence intervals is a conservative way to declare difference between parameter and better guidance is needed on this.

Panel Discussion (Reference Substances)

The Panel stated that in general, the criteria used to select the performance standards reference substances are adequate. However, one could question the appropriateness of defining reference substances as positive based upon >50% of ER TA studies indicating a positive response. Given that the quality of reference data has not been detailed the Panel cannot definitively determine appropriateness of the reference substances.

The Panel agreed that the list of reference substances upon which to evaluate the performance of functionally and mechanistically similar test methods would be considered adequate if there were more negatives and proportionally fewer positives in the list for agonist testing. The list has reasonable overall diversity and reflects the extensive effort to obtain relevant information. There are enough substances to lend sufficient robustness to an assessment. There is a good range of estrogenic activity over several orders of magnitude, as well as a few confounders to assess the robustness of the assay and methodology. There may be an opportunity to revisit the list of reference substances and make modifications based on experience gained in the assay. Including discordant chemicals on the reference substance list is important because they are critical for truly characterizing the limitations of the assay. The potent estrogens on the reference list should not be missed and there could be some tolerance for discordance for the weaker acting reference substances. Any discordant results should be discussed in terms of the ability of the test method to detect a similar range of potencies and intrinsic activity and chemical/product classes.

Panel Discussion (General Comments)

The Panel determined that the number of repeat experiments to evaluate intra- and interlaboratory reproducibility appeared to be a good starting point for evaluation of the test method. The intralaboratory assessment (based on at least three tests with 12 chemicals) and the interlaboratory reproducibility assessment through one trial (three laboratories) allowed for ‘real world’ evaluation of the methodology on the validation reference set of materials. The Panel recommended that the BG1 method should use a range of accuracy (or perhaps the lowest %) and include a metric of potency and

intrinsic activity in the accuracy evaluation (a measure of uncertainty needs to be included). There is also a need to establish tolerance to the vehicle (ideally at least 1%). The test report should include potency and sensitivity analyses based on EC₅₀ values (agonist) or IC₅₀ values (antagonist).

Adjournment

After the discussion, Dr. Hayes adjourned the Panel for the day at 5:31 p.m.

North Carolina State University is a land-grant university and a constituent institution of The University of North Carolina

**College of Agriculture and Life Sciences
Department of Biology**

NC STATE UNIVERSITY

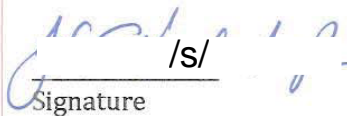
127 David Clark Labs
Campus Box 7617
Raleigh, NC 27695-7617

919.515.2741
919.515.5327 (fax no.)

Warren Casey
NIEHS
P.O. Box 12233
Mail Stop: K2-16
Research Triangle Park, NC 27709

Dear Dr. Casey-
The Meeting Summary Minutes, Independent Scientific Peer Review Panel Meeting
Evaluation of In Vitro Estrogen Receptor Transcriptional Activation Test Method for
Endocrine Disruptor Chemical Screening accurately summarizes the Peer Review
Panel meeting of March 29 - 30, 2011, in Bethesda, MD.

Sincerely,

 /s/ _____
Signature Printed Name

John G. Vandenberg

Sept. 28, 2011

Date

This page intentionally left blank

Appendix D2

Independent Scientific Peer Review Panel Report: Evaluation of the LUMI-CELL[®] ER (BG1Luc ER TA) Test Method

This page intentionally left blank

**Independent Scientific Peer Review Panel Report
Evaluation of the LUMI-CELL[®] ER (BG1Luc ER TA)
Test Method**

May 2011

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

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

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

National Toxicology Program
P.O. Box 12233
Research Triangle Park, NC 27709

Independent Peer Review Panel Report

This document is available electronically at
http://iccvam.niehs.nih.gov/docs/endo_docs/EDPRPrept2011.pdf

When referencing this document, please cite as follows:

Interagency Coordinating Committee on the Validation of Alternative Methods. 2011. Independent Scientific Peer Review Panel Report: Evaluation of the LUMI-CELL[®] ER BG1Luc ER TA Test Method. Research Triangle Park, NC:National Institute of Environmental Health Sciences.

The findings and conclusions of this report are those of the Independent Scientific Peer Review Panel and should not be construed as representing the official views of ICCVAM or its member agencies.

Independent Peer Review Panel Report

Table of Contents

List of Abbreviations and Acronyms	iv
Members of the Independent Scientific Peer Review Panel	vi
Preface	vii
Executive Summary	ix
Overview	
Use of the BG1Luc ER TA Test Method to Identify Substances as Potential <i>In Vitro</i> Estrogen Receptor Agonists or Antagonists	1
I. Review of the BG1Luc ER TA Test Method BRD for Errors and Omissions	2
II. Evaluation of the Validation Status of the BG1Luc ER TA Test Method	3
1. Introduction and Rationale for the Test Method	3
2. Test Method Protocol	4
3. Substances Used for the Validation Study	5
4. Data and Results	6
5. Test Method Accuracy	8
6. Test Method Reliability (Intra- and Interlaboratory Reproducibility)	10
7. Data Quality	10
8. Consideration of All Available Data and Relevant Information	11
9. Animal Welfare Considerations	11
10. Practical Considerations	13
III. Draft ICCVAM Test Method Recommendations on the BG1Luc ER TA Test Method to Identify Substances with Estrogen Agonist and/or Antagonist Activity	14
1. Test Method Usefulness and Limitations	14
2. Test Method Protocols	14
3. Future Studies	14
4. Performance Standards	15
5. Additional Comments on Essential Test Method Components	17
6. Additional Comments on Test Method Performance	17
IV. References	18
Appendix A	
Errors and Omissions	21
Appendix B	
Peer Review Panel Member Biosketches	33

Independent Peer Review Panel Report

List of Abbreviations and Acronyms

BG1Luc ER TA	LUMI-CELL [®] BG-1Luc4E2 ER TA test method
BPA	Bisphenol A
BRD	Background review document
CERI	Chemicals Evaluation and Research Institute (Japan)
CV	Coefficient of variation
DMSO	Dimethyl sulfoxide
EAC	Endocrine-active chemical
EC ₅₀	Half-maximal effective concentration
ECVAM	European Centre for the Validation of Alternative Methods
ED	Endocrine disruptor
EDSP	EPA Endocrine Disruptor Screening Program
EDSTAC	EPA Endocrine Disruptor Screening and Advisory Committee
EDTA	Endocrine Disruptor Testing and Assessment (OECD)
EPA	U.S. Environmental Protection Agency
ER	Estrogen receptor
ERE	Estrogen-responsive element
FBS	Fetal bovine serum
FIFRA	U.S. Federal Insecticide, Fungicide, and Rodenticide Act
GCCP	Good Cell Culture Practices
GLP	Good Laboratory Practices
IC ₅₀	Concentration of the test substance that inhibits the reference estrogen response by 50%
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
ISO	International Organization for Standardization
JaCVAM	Japanese Center for the Validation of Alternative Methods
KoCVAM	Korean Center for the Validation of Alternative Methods
NICEATM	U.S. National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods
NIEHS	U.S. National Institute of Environmental Health Sciences
NIH	U.S. National Institutes of Health
OECD	Organisation for Economic Co-operation and Development
OPPTS	Office of Prevention, Pesticides and Toxic Substances
QA	Quality assurance
QSAR	Quantitative structure-activity relationship
REACH	Registration, Evaluation and Authorisation of Chemicals
RLU	Relative light unit
RUC	Rat uterine cytosol
SACATM	Scientific Advisory Committee on Alternative Toxicological Methods

Independent Peer Review Panel Report

SD	Standard deviation
SOP	Standard operating procedure
STTA	Stably transfected human estrogen receptor- α transcriptional activation
TA	Transcriptional activation
TG	Test Guideline (OECD)

Independent Peer Review Panel Report

Members of the Independent Scientific Peer Review Panel

John G. Vandenberg, PhD (Panel Chair), Professor Emeritus, Department of Biology, North Carolina State University, Raleigh, NC

A. John Bailer, PhD, Distinguished Professor and Chair, Department of Statistics, Miami University, Oxford, OH

Christopher Borgert, PhD, President and Principal Scientist, Applied Pharmacology and Toxicology, Inc., Gainesville, FL

Grantley Charles, PhD, DABT, Senior Scientist, Toxicology, Drug Safety Evaluation, Allergan Inc., Irvine, CA

Daniel Desaulniers, PhD, Research Scientist, Hazard Identification Division, Health Canada, Ottawa, ON, Canada

J. Charles Eldridge, PhD, Professor of Physiology and Pharmacology, Wake Forest University Health Sciences, Winston-Salem, NC

William R. Kelce, PhD, Principal, Exponent[®], Cary, NC

Hyung Kim, PhD, Associate Professor, College of Pharmacy, Pusan National University, Busan, Republic of Korea

Steven L. Levine, PhD, Science Fellow and Product Safety Manager, Monsanto Company, St. Louis, MO

Alberto Mantovani, DVM, Head of Food and Veterinary Toxicology Unit, Italian National Health Institute (ISS - Istituto Superiore di Sanità, viale), Rome, Italy

Ellen Mihaich, PhD, DABT, President and Principal Scientist, Environmental and Regulatory Resources, LLC, Durham, NC

Hiroshi Ono, PhD, Scientific Advisor, Food and Drug Safety Center, Hatano Research Institute, Hadano, Japan

Sherry L. Ward, PhD, MBA, Consultant, BioTred Solutions, New Market, MD

Marc Weimer, PhD, Biostatistician, Department of Biostatistics, German Cancer Research Center, Heidelberg, Germany

James Wittliff, PhD, MD hc, FACB, Professor, Department of Biochemistry and Molecular Biology, Director, Hormone Receptor Laboratory, Director, Institute for Molecular Diversity and Drug Design, University of Louisville, Louisville, KY

James Yager, PhD, Professor and Director of Toxicology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD

Independent Peer Review Panel Report

Preface

The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) convened an international independent scientific peer review panel (Panel) meeting on March 29-30, 2011, at the Natcher Conference Center in Bethesda, Maryland. The Panel, which included 16 expert scientists from six countries, evaluated the LUMI-CELL[®] estrogen receptor (ER) transcriptional activation (TA) test method (BG1Luc ER TA), an *in vitro* TA assay used to identify chemicals that can interact with human ERs.

During the March 2011 public meeting, the Panel discussed the test method, listened to public comments, and developed conclusions and recommendations for ICCVAM. The Panel focused on the following areas: (1) review of the ICCVAM draft background review document (BRD) for completeness and identification of errors or omissions of existing relevant data or information that should be included, (2) evaluation of the information in the draft BRD to determine the extent to which each of the applicable ICCVAM criteria for validation and acceptance of toxicological test methods had been appropriately addressed, and (3) consideration of the ICCVAM draft test method recommendations and commentary on the extent to which they are supported by the information provided in the draft BRD for the following:

- Proposed test method uses and limitations
- Proposed recommended standardized protocols
- Proposed test method performance standards
- Proposed future studies

This report details the Panel's independent conclusions and recommendations. ICCVAM will consider this report and all relevant public comments as it develops final test method recommendations. The ICCVAM final test method recommendations will be forwarded to U.S. Federal agencies for their consideration in accordance with the ICCVAM Authorization Act of 2000 (42 U.S.C. 285f-3).

The Panel gratefully acknowledges the efforts of NICEATM staff for an outstanding effort in coordinating the logistics of the Panel meeting and in the preparation of materials for its review. Finally, as Panel Chair, I want to thank each Panel member for her or his thoughtful and objective review of this test method.

John G. Vandenberg, PhD
Chair, BG1Luc ER TA Test Method Peer Review Panel
May 2011

Independent Peer Review Panel Report

This page intentionally left blank

Independent Peer Review Panel Report

Executive Summary

This report describes the conclusions and recommendations of an international independent scientific peer review panel (Panel). The Panel was charged by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) with evaluating the validation status of the BG1Luc estrogen receptor (ER) transcriptional activation (TA) test method according to established Federal and international criteria (ICCVAM 1997). The Panel also commented on ICCVAM draft recommendations regarding the usefulness and limitations of the test method and proposed performance standards.

The Panel considered the results of an international interlaboratory validation study that included laboratories in the United States, Italy, and Japan. Based on their evaluation of these data, the Panel agreed with ICCVAM's draft test method recommendation that the BG1Luc ER TA test method can be used to identify substances with *in vitro* estrogenic and anti-estrogenic activity. Based on results of concordance analyses for a limited number of substances, the Panel further concluded that the BG1Luc ER TA test method could be considered as a replacement for other *in vitro* assays that may provide substantially similar information, specifically the Chemicals Evaluation and Research Institute stably transfected transactivation assay (CERI STTA) and the rat uterine cytosol (RUC) ER binding assay. The Panel noted that additional analysis could further support this recommendation, particularly regarding the RUC ER binding assays.

The Panel endorsed the draft ICCVAM-recommended test method protocols and noted several advantages provided by this assay over the currently accepted test method for this endpoint, including the robust test method protocol, the validated testing range, and the ability to detect substances with *in vitro* anti-estrogenic activity. However, the Panel also noted that careful analysis of cytotoxicity is critical for correctly interpreting results. The Panel expressed a preference for using quantitative approaches for such a measurement. The Panel recommends that a potency endpoint, such as the half-maximal effective and/or inhibitory concentration (EC/IC₅₀), be included in each study report and that the uncertainty associated with these estimates should also be reported. The Panel considered the descriptive approach for evaluating test method reliability acceptable but suggested additional statistical analyses that could be performed to better characterize and understand variability.

The Panel agreed with the draft ICCVAM-recommended future studies and suggested additional studies that should be conducted to expand the usefulness of the BG1Luc ER TA test method. The Panel recommended additional evaluations of the utility of the current categorical assessment of cytotoxicity and advocated for the implementation of a quantitative method for its replacement. The Panel also recommended studies to add *in vitro* metabolism (compound activation or inactivation) to the test method. This addition could expand the utility of this and other ER TA test methods. The Panel recommended that additional efforts focus on expanding the reference substance list, and subsequently the BG1Luc ER TA test results, with additional negative agonist and positive antagonist test substances.

Finally, the Panel concurred that the draft ICCVAM performance standards could be used to evaluate the validation status of test methods that are functionally and mechanistically similar to the BG1Luc ER TA test method. The Panel considered the list of performance standards reference substances to be adequate. The Panel noted that ideally more negatives should be included but recognized that data on such substances are not currently available. When evaluating test method accuracy, the Panel strongly supported quantification of relative agonist and antagonist activity in addition to the dichotomous call of positive or negative. In addition, the Panel concluded that the potent estrogens on the reference list should not be misclassified, but there could be some tolerance for discordance for the weakly active reference substances. Discordant results need to be discussed in terms of the ability of the test method to detect a similar range of potencies and intrinsic activities compared to current

Independent Peer Review Panel Report

validated test methods. Discordant results for particular chemicals or product classes also need to be discussed.

Independent Peer Review Panel Report

Overview**Use of the BG1Luc ER TA Test Method to Identify Substances as Potential
In Vitro Estrogen Receptor Agonists or Antagonists**

The overall question that the international independent scientific peer review panel (Panel) considered is whether the validation status of the BG1Luc estrogen receptor (ER) transcriptional activation (TA) test method has been adequately characterized for its intended purpose and whether it is sufficiently accurate and reliable to be used to identify substances with estrogen agonist and/or antagonist activity.

The Panel discussed the intended use of this assay and the potential for its inclusion in a regulatory testing battery. Panel members agreed that the BG1Luc ER TA test method and the Chemicals Evaluation and Research Institute stably transfected transactivation assay (CERI STTA) are similarly capable of assessing *in vitro* estrogen receptor (ER) agonist activity. In addition, the BG1Luc ER TA test method is capable of detecting *in vitro* estrogen antagonists. Clarification of the intended use of these assays in regulatory decision making, particularly in the context of the U.S. EPA's Endocrine Disruptor Screening Program (EDSP), would enable a better understanding of the relative merits of the various screening assays for their intended purpose.

In the absence of clear regulatory guidance, the Panel recommends that the BG1Luc ER TA test method be endorsed as a scientifically valid method for assessing the *in vitro* estrogen agonist and antagonist activity of compounds within a test battery or tiered testing scheme. The Panel recommends that the BG1Luc ER TA method be considered as a replacement for other *in vitro* assays that, in combination, may provide substantially similar information, specifically the CERI STTA assay. This is supported by the following findings:

- The concordance of the BG1Luc ER TA test method with the CERI STTA assay suggests that the BG1Luc ER TA test method and the CERI assay produce similar results.
- The thoroughness and transparency of the BG1Luc ER TA method validation process compare favorably with other *in vitro* assays.
- The detailed BG1Luc ER TA agonist and antagonist protocols permit ease of use.
- The detailed and publically available BG1Luc ER TA data permits thorough evaluation of the performance of the method.
- The endogenous expression of both ER α and ER β in BG1Luc4E2 cells allows *in vitro* activity through both receptors to be assessed in the BG1Luc ER TA test method. Endogenous expression of the receptor and its related endogenous cellular machinery may be an advantage over receptors that are stably transformed into an immortal cell line and constitutively expressed at high levels.

Independent Peer Review Panel Report

I. Review of the BG1Luc ER TA Test Method BRD for Errors and Omissions

The Panel noted typographical errors in the draft background review document (BRD), which are detailed in Appendix A of this report.

The Panel also cited relevant information that should be included that would improve the BRD. These include:

- In Figure 1-1 of the BRD, Phase 2 should be changed to Laboratory Proficiency Phase because the laboratory qualification should already have been demonstrated.
- Ongoing improvements were made to the protocol(s) during the course of these validation studies. A better assessment of their impact on the final study results is needed.
- The Panel has recommended additional reliability analyses; therefore, Section 6.0 of the BRD and Lines 43-49 and 76-82 of the draft recommendations of the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) may need to be revised pending their outcome.
- Data quality sections for the reference test method data were not included in the draft BRD. The Panel recommends they be added for CERI STTA, ICCVAM literature-based reference data, and RUC ER binding assay.
- Sections II and III of this peer review panel report detail additional information gaps cited by the Panel.

II. Evaluation of the Validation Status of the BG1Luc ER TA Test Method

1. Introduction and Rationale for the Test Method

a. Is the historical background provided for the BG1Luc ER TA test method and the rationale for its development adequate?

While the historical background provided in the BRD is adequate, it would be improved by including a more thorough discussion of the specific advantages of this transactivation assay relative to other *in vitro* ER assays (see Rogers and Denison 2000). Accordingly, each of the following important points should be emphasized in the Background section of the BRD:

- Transactivation assays provide advantages compared to ER binding assays in that ER transcriptional activation (TA) assays have the potential to assess both *in vitro* estrogen agonist and antagonist activity. They can also detect endocrine-active substances that elicit effects on estrogen-regulated pathways through non-receptor-mediated mechanisms.
- BG-1 cells endogenously express both ER α and ER β and, consequently, possess the full transcriptional machinery required for estrogen responsiveness.
- Stably transfected cell lines have advantages over transiently transfected cell lines, including long-term utility without the need to transiently transfect cells for each assay.
- The BG1Luc ER TA test method has demonstrated *in vitro* responsiveness to estrogens and limited cross-reactivity with ligands of other steroid hormone receptors.
- A discussion should be added to detail the four copies of the estrogen-responsive element (ERE) and their orientation, which mimics that in the fish vitellogenin promoter and thereby indicates a native system.

b. Are the current regulatory testing requirements and ICCVAM prioritization criteria adequately discussed and up to date?

ICCVAM's prioritization criteria are adequately discussed. If possible, the BRD should compare the phased Organisation for Economic Co-operation and Development (OECD) conceptual framework approach with the EPA Tier 1 battery approach, as discussed at the March 2011 Panel meeting. If possible, as new/other worldwide regulatory agencies develop endocrine disruptor (ED) testing strategies, this information should be updated via addenda.

c. Are the purpose and mechanistic basis of the BG1Luc ER TA test method adequately described?

The purpose of this test method as a screen for *in vitro* estrogen receptor agonist and antagonist activity has been adequately stated. The mechanistic basis of the test method also is adequately described.

d. Is the description of the use of the proposed test method in an overall strategy of hazard or safety assessment adequate?

The use of the proposed test method in the overall strategy of hazard or safety assessment of endocrine-active chemicals (EACs) is unclear. There has been no clear regulatory guidance on how the method will be used in the EPA Endocrine Disruptor Screening Program (EDSP).

Because the BG1Luc ER TA method has been validated to assess compounds for *in vitro* estrogen agonist and antagonist activity, the BG1Luc ER TA method should be considered as a replacement for the CERi STTA (which detects only agonist activity) and the RUC ER binding assays.

Independent Peer Review Panel Report

If, in the future, a receptor binding assay is still deemed necessary, the Panel recommends an evaluation of recombinant (human and animal) ER binding assays as replacements for the RUC ER binding assay.

The BRD should also propose how the BG1Luc ER TA method could be used in screening, prioritization, and safety assessments for other regulatory testing programs (see comment above regarding the OECD conceptual framework).

2. Test Method Protocol

a. Are the protocols complete and adequate in detail for a laboratory to conduct the study, including (1) a description of the material and equipment needed to conduct the test and (2) a description of what is measured and how the data are used to identify positive and negative results?

The protocols in Annexes E and F of the BRD appear complete and adequate in detail for a laboratory to conduct the study. Although the protocols state that alternative vendors for technical and cell culture equipment and supplies may be used, this should be more clearly stated in the BRD body.

The approaches to evaluate data were reviewed and the Panel's recommendations on quantification of activity are presented below in the response to Section II Question 4b.

b. Overall, are critical aspects of the test method protocol, as outlined in the ICCVAM Submission Guidelines, adequately justified and described in the BRD?

Critical aspects of the test method protocol are adequately justified and described in the BRD. However, the subjective nature of the visual assessment method for determining cytotoxicity raised concerns about its application with regard to future usage of the test method protocols.

These concerns are:

- Visual assessment of cytotoxicity is subject to operator inconsistencies and may not always accurately reflect the viability of cells in culture. This is particularly relevant as the use of the method expands to other laboratories.
- The current cytotoxicity evaluation is in line with currently accepted practices (ISO 2009). The International Organization for Standardization (ISO) document compared the use of the subjective visual assessment of cytotoxicity with the objective MTT assay with regard to medical devices. This type of comparison has not but should ideally be completed in the future for chemicals evaluated in the BG1Luc ER TA.
- An accurate, objective cytotoxicity method is most critical for the antagonist assay because it measures a loss of function mediated through the ER. Therefore, an ER-mediated decrease in activity must be delineated from that resulting from cytotoxicity alone.

For future test method use, the Panel recommends:

- Testing of a wider set of substances with known mechanisms of cytotoxicity to further validate the qualitative cytotoxicity method
- Use of quantitative cytotoxicity methods when developing new *in vitro* ER assays, which could also allow for normalization of relative light unit (RLU) responsiveness

Independent Peer Review Panel Report

i. Do you agree with reference standards and controls proposed for the agonist and antagonist protocols? Are there other reference standards and/or controls that you would consider more appropriate?

The reference standards and controls proposed for the agonist and antagonist protocols are appropriate. Future studies should consider including confirmation assays using a pure ER antagonist (e.g., ICI 182,780), which would confirm ER binding behavior.

ii. Do you agree with the plate acceptance criteria as defined in the agonist and antagonist protocols? Are there additional criteria that should be routinely used?

The plate acceptance criteria defined in the agonist and antagonist protocols appear adequate. However, the criteria for acceptance based on the dimethyl sulfoxide (DMSO) controls could be too stringent if, for example, luminometer replacement is necessary. In such a situation, baseline luminometer values might change. The expectation is that the laboratory would perform adequate repeat baseline and positive and negative controls to assure reproducibility but would not repeat an entire validation study. A suggestion was made to use an acceptable range in RLUs induced by a specific concentration of 17 β -estradiol. Such criteria would be independent of equipment and of the choice of vehicle; however, this approach would require validation.

iii. Do you agree with the proposed decision criteria for identifying a positive or negative response in the agonist and antagonist protocols?

The proposed decision criteria for identifying a qualitative positive or negative response in the agonist and antagonist protocols are acceptable (see response to Section II Question 4b).

However, potency and intrinsic activity relative to a known endogenous ligand are critical components in determining whether a substance is truly positive, and this is not addressed in the current decision criteria. This assessment could include a relative potency approach by comparing to known reference substances. Including this component would result in dramatic improvement in the utility of this assay.

The Panel recommends that the half-maximal effective and/or inhibitory concentrations (EC/IC₅₀s) (or potency endpoints such as EC/IC₂₀ or the concentration associated for a particular fold induction) be included in each study report and discussed in the Conclusions section in association with the qualitative dichotomous positive or negative calls. Uncertainty associated with these estimates should also be reported (e.g., confidence intervals).

3. Substances Used for the Validation Study

a. Do you consider the database for the BG1Luc ER TA test method representative of a sufficient range of chemical classes and physicochemical properties and that it would be applicable to any of the types of chemicals and products that are typically tested for estrogenic activity? If not, what are the relevant chemical classes/properties (other than those that are identified as limitations in the previous ICCVAM BRD [ICCVAM 2003, 2006]) that should be tested with caution, or not evaluated using this test method? What chemicals or products should be evaluated to fill this data gap?

The chemicals tested in this validation exercise represent a broad range of different chemical classes and physicochemical properties and represent a census of available information. The list follows ICCVAM guidance (ICCVAM 2003, 2006) and is largely applicable to chemicals and products that would be screened to evaluate potential *in vitro* estrogenic activity. Testing an adequate range of activities and a structurally diverse group of chemicals is important. In this validation study, estimated EC₅₀s ranged seven orders of magnitude and varied in terms of their coverage of chemical

Independent Peer Review Panel Report

classes. However, several chemical classes in Table 3-7 of the BRD are represented by as few as one substance; therefore, in such cases, no conclusions on usefulness or limitations specific to these classes can be made.

As new chemicals are tested in both agonist and antagonist protocols (e.g., data generated from ongoing EDSP Tier 1 testing and/or new ER TA protocols), the Panel recommends including a subset of the newly tested chemicals in any future validation studies. In particular, inclusion of data from chemicals in under-represented classes, compounds with known surface-active properties that could perturb the cell system, and compounds that are known to be negative for agonist activity are necessary. This last recommendation concerning negative compounds is particularly important because less than 25% of the agonist substances used for the accuracy analysis are negative. While the Panel acknowledges that great effort went into identifying substances that would fit into this category, it is difficult to investigate false positives in a new test system if the majority of test chemicals are positive.

b. Do you agree with the methodology used to establish the consensus reference classification that was assigned to each reference substance?

The use of the majority classification criteria among study results (i.e., >50%) to establish the consensus reference classification for each reference substance is a reasonable strategy. However, the criteria used to evaluate the quality of the data obtained from the literature to determine the reference classification needs to be described in the BRD. A ranking method such as Klimisch criteria (Klimisch et al. 1997), which focuses primarily on the reliability of the data, would provide clarity on the relative quality of the reference data and strengthen the classification. In addition, a sensitivity analysis could be performed to assess how the results of the validation study vary based on the method of reference classification. Alternatively, the concordance assessment could be repeated with a different classification criterion, such as declaring a chemical positive if at least one laboratory declared a positive result.

4. Data and Results

a. Have all known data for all studies used to evaluate the accuracy and reliability of the BG1Luc ER TA test method been provided?

Data for all studies used to evaluate the accuracy and reliability of the BG1Luc ER TA test method appear to have been provided. The interlaboratory studies of the BG1Luc ER TA test method conducted by NICEATM, JaCVAM, and ECVAM have been included in the accuracy and reliability assessments provided in the BRD. The Panel is unaware of additional studies available for comparison to the BG1Luc ER TA test method for accuracy and reliability.

b. Are the statistical and nonstatistical approaches to evaluate the data resulting from studies conducted with the BG1Luc ER TA test method appropriate? What other approaches could have been used?

Approaches for assessing test method accuracy

The Panel strongly supports the quantification of activity as a complement to the dichotomous call of a positive or negative response. Interpretation of the results should not rely solely on statistics but also on scientific judgment and should incorporate consideration of the nature and shape of the dose-response relationship and, if needed, the reproducibility of the response in independent experiments.

Independent Peer Review Panel Report

At a very general level, the Panel recommends that the following questions be explicitly addressed in the BRD:

1. What is the role of statistical modeling in the quantification of activity in this assay?
2. What aspects of the dose-response relationship are of greatest interest? For example, potency evaluation may be of interest.
3. Would characterization of a chemical's estrogen receptor activity be most relevant and best utilized for comparison to that of a reference compound's estrogenic response?
4. Are there statistical criteria that can be used to support the decision of a positive call?
 - Should statistical tests of trend be included in the evaluation of a positive call?

ICCVAM-recommended criteria (ICCVAM 2003) incorporating appropriate statistical methods and sound scientific judgment for classifying a substance as an ER agonist or antagonist are essential for ensuring the credibility of the results. However, the evaluation of data from studies conducted with the BG1Luc ER TA test method would benefit from specific attention to the following detailed comments.

The criteria for making a positive call

As defined in the BRD, the criteria for a positive call are based on comparisons to experimental results from the current set of test chemicals and sound scientific judgment. While the Panel views this as an acceptable strategy, a statistical test for trend is a reasonable approach for making a positive call (Bretz and Hothorn 2003).

The criteria for estimating potency

Potency estimation requires an adequate concentration–response model (Ritz 2010). If a particular model has been selected for biological or statistical reasons, then some processing of data may be required. For example, a Hill model was fit to the RLU responses in the current BRD. This model assumes a positive or neutral slope in the dose-response relationship. Points that deviated from a positive/neutral slope were excluded from the analysis. The rules for such preprocessing of data must be clearly stated.

The estimated values for EC₅₀ and IC₅₀ presented in the BRD were point estimates without any error associated with them. Uncertainty associated with these estimates should be reported (e.g., confidence intervals).

Potency estimates can be broader than just EC₅₀ and IC₅₀ parameters. For example, the concentration associated with a specific fold induction may be a relevant potency endpoint.

Additional considerations

The BRD should explicitly define all data transformations and normalizations. For example, the procedure to calculate the adjusted RLUs should be clearly described. The variability in the DMSO control responses appears to be ignored for the background correction of the adjusted RLUs. An alternative is to include the control response to help estimate the baseline in the concentration response model.

Approaches for assessing reliability

The appropriateness of methods used for reliability assessment is discussed in Section 6.0 (below).

- c. **For each set of data, is the discussion of whether coded substances were tested and whether experiments were conducted without knowledge of the identity of the substances being tested adequately documented?**

The discussion of whether coded substances were tested and whether experiments were conducted without knowledge of the identity of the substances being tested is adequately documented.

Section 3.9 of the BRD describes procurement, coding, and distribution. Section 4.4.2 of the BRD

Independent Peer Review Panel Report

states that substances in Phases 2, 3, and 4, which were critical for the evaluation of reliability and accuracy, were coded.

The only additional information that should be provided is a confirmation of the coding for the controlled substances purchased regionally. Specifically, BRD Section 3.9 states that ECVAM and JaCVAM procured “controlled substances” from their regional suppliers. This raises the question of how they ensured that the laboratories did not know the identity of the tested substances—this should be clarified in the BRD.

5. Test Method Accuracy

a. The current accuracy analysis is based primarily on overall concordance with the ICCVAM reference consensus classification, which is based on results from *in vitro* ER TA test methods. Are these data adequate for assessing the accuracy of the test method?

The ICCVAM reference consensus classification was the primary reference method used for assessing BG1Luc ER TA test method accuracy. Some additional measures of agreement were performed in comparison to the CERi STTA, RUC ER binding assay, and the uterotrophic assay. As described in Section 5.5 of the BRD, median estimated EC₅₀ and IC₅₀ values were compared with reference data using regression. If analysis of “agreement” between the two data sources is a major concern, statistical approaches are available that might provide more insight into the strength of agreement between assay results (e.g., Bland-Altman Plots and Limits-of-Agreement [Assessing Agreement 2007; Bland and Altman 1986, 1999]).

The endorsement of a screening assay cannot be based on its strict agreement with any one other method but should be judged on a weight-of-evidence approach that includes all performance comparisons, the repeatability results, and scientific judgment regarding the biological relevance of the test system.

b. Do you agree that accuracy of the BG1Luc ER TA test method should be based only on those substances for which an unequivocal reference classification can be assigned?

The accuracy of the BG1Luc ER TA test method should be assessed using unequivocal reference classifications from other human ER TA tests. This increases confidence in the results by eliminating questionable responses. One caution is that validation study results based upon unequivocal classifications may result in overly optimistic assessments of test method performance. Substances that result in equivocal reference classifications may provide additional insights into aspects of the test method.

c. Other concordance analyses included in the BRD are based on direct comparisons to *in vivo* (uterotrophic) or *in vitro* (CERi STTA, ER binding assay) test methods. How much emphasis should be placed on these comparisons relative to the comparisons to the ICCVAM classification?

Comparing the new test method with other methods assessing the same mode of action is reasonable. Indices of agreement should be calculated only with data that meets certain quality control measures. Comparison of data from other methods could be confounded if alternative methods are not measuring an endpoint based upon the same biological mechanism. Accordingly, emphasis should be placed on describing the different purpose of the various tests, as well as their advantages and disadvantages.

As noted above, data quality is an uncertainty for all of the reference methods used in the BG1Luc ER TA BRD.

Independent Peer Review Panel Report

- d. Based on your responses to Questions 5 a-c, has the relevance (e.g., accuracy/concordance, sensitivity, specificity, false positive, and false negative rates) of the BG1Luc ER TA test method been adequately evaluated? If not, what other analyses should be performed?**

The interpretation of the estimated error rates should be put into context. The first observation is that this set of test chemicals necessarily represented a limited set of substances. As with all previous validation studies of this type, the performance of this assay with new classes of chemicals will be difficult to predict. This is not seen as any reason for not endorsing the method for use in a test battery or other weight-of-evidence approach, but it is an important issue for users to understand.

The second observation is that confidence intervals for the calculated performance statistics, including the false positive and false negative error estimates, should be provided. Additionally, the contingency tables used to generate the summary statistics should be included in the document, including tables showing results from the individual laboratories.

The criteria for declaring a chemical as positive changed after Phase 3 was completed and were retrospectively applied to all of the previous data. The Panel recognizes that the original criteria were inadequate. This should not be a precedent going forward because optimal decision criteria should be selected *a priori*. However, for this study, the Panel affirms that the changes made were performed in an appropriate manner.

- e. Do you agree that this test method is effective for generating data in an amount and of a scientific value that is at least equivalent to the data generated from existing tests (i.e., U.S. EPA Endocrine Disruptor Screening Program Test Guideline OPPTS 890.1300: Estrogen Receptor Transcriptional Activation [EPA 2009])?**

The BG1Luc ER TA test method is effective for generating data that are functionally equivalent to the data generated from existing tests (i.e., U.S. EPA Endocrine Disruptor Screening Program Test Guideline OPPTS 890.1300: Estrogen Receptor Transcriptional Activation [EPA 2009]).

In fact, the BG1Luc ER TA protocol could be considered an improvement over the CER1 protocol given the extent to which the BG1Luc ER TA protocol was standardized and optimized. The Panel recommends that the CER1 protocol be revised to meet the standards established in the BG1Luc ER TA protocol.

- f. Based on comparison of results from the ER TA with results from RUC ER binding assays, are there sufficient data to conclude that positive and/or negative results from BG1Luc ER TA can adequately identify ER binding potential compared to the RUC? If not, what other analyses and/or additional information would be necessary to draw such a conclusion?**

RUC ER binding assays indicate agreement with the BG1Luc ER TA test method results and suggest that the outcomes of the stably transfected ER TA test method can provide insight regarding the biological effect of chemicals mediated by ER–ligand interactions. There appears to be sufficient data to consider replacing the rat uterine cytosol ER binding assay with the BG1Luc ER TA test method, especially if ER confirmation assays are completed as part of the BG1Luc ER TA test method procedure (see Section II.2.b.i). Therefore, the Panel recommends that NICEATM and ICCVAM identify and use other available ER binding data for these comparisons (i.e., substances beyond the 78 tested in the BG1Luc ER TA validation study). As an additional activity, the Panel recommends an evaluation of recombinant (human and animal) ER binding as a replacement for the RUC ER binding assays.

Independent Peer Review Panel Report

6. Test Method Reliability (Intra- and Interlaboratory Reproducibility)

a. Has the intra- and interlaboratory reproducibility of the BG1Luc ER TA test method been adequately evaluated? If not, what other analyses should be performed? Are any limitations apparent based on this intralaboratory reproducibility assessment?

The Panel combined the responses to both intra- and interlaboratory reproducibility evaluations because they are similar.

In general, insufficient data are presented to ensure that a thorough analysis of intra- and interlaboratory reproducibility has been performed. While reproducibility has been addressed on three data levels (raw data [e.g., DMSO control measurements], derived endpoints [e.g., EC₅₀], and prediction of estrogenic activity), the analysis is descriptive. No formal inference has been carried out. The descriptive approach is acceptable; however, some additional analyses could be performed to better understand the components of assay response variability.

Many of the figures in the BRD do not adequately present the variability associated with the test method and in fact may confuse variability with outliers. For example, it is better to display a scatter plot of data instead of means and standard deviations when only three data points are present at each concentration. Sources of variability could be explored using appropriate methods such as descriptively comparing coefficients of variation (CVs) or using more formal statistical methods to estimate variance components.

The potency discussion in the BRD did not capture the variability in these quantities. This could be captured by calculating the mean and standard deviation of the log EC₅₀ for each chemical in the BG1Luc ER TA validation study. These quantities could then be compared to the analysis of other established test methods considered acceptable that also include these calculations (e.g., CERI STTA).

The rationale for selection of substances used to evaluate intra- and interlaboratory reproducibility should be more clearly described in the BRD. It should be noted that excluding substances from the analysis of reproducibility for which definite results were not determined may produce an assessment of reproducibility that is too optimistic.

7. Data Quality

a. Is the extent of adherence to national and international Good Laboratory Practice (GLP) guidelines for all submitted test data and the use of coded substances and coded testing adequately presented?

Validation guidance states that “Ideally, all data supporting the validity of a test method should be obtained and reported in accordance with Good Laboratory Practice (GLP) guidelines. Aspects of data collection not performed according to GLPs must be fully described, along with their potential impact” (ICCVAM 1997). Two out of the three laboratories in this study followed GLPs, and the third laboratory was ISO 9000 certified. Therefore, the extent of adherence to national and international GLP guidelines is adequately presented. The use of coded substances and coded testing in the validation studies is adequately described. Deviations and alterations to original protocols including use of coded substances were relatively well described, such as acquisition of controlled agents (e.g., androgens, barbiturates, etc.) for testing.

However, while training of laboratory personnel was reportedly conducted, details that would be included in formal training records were not available to confirm compliance with GLPs and Good Cell Culture Practices (GCCP).

Independent Peer Review Panel Report

An essential subsection is missing from BRD Section 7.0 that should include a description of the quality of the reference data. Several reference methods were used to assess BG1Luc ER TA test method accuracy and concordance. Therefore, the “quality” of the reference data for the following methods should be summarized in the BRD: CER1 STTA, data used to generate ICCVAM reference classifications, uterotrophic assay, and RUC ER binding assay. If this information is incomplete or unavailable, it should be explained in the BRD. As discussed in Section 3 of the BRD, regarding the ICCVAM reference classifications, consideration could be given to assigning greater weight to certain reference data, and a thorough description of the weighting methodology should be provided.

Section 7.2 of the BRD states that an independent quality assurance (QA) review was conducted at each of the participating laboratories. This should have included a process audit of the assay at each laboratory (at the time of testing) to ensure that standard operating procedures (SOPs) were followed. Any data transfers made during calculations and statistical analysis also need to be audited to ensure that errors are caught and corrected. The responsible QA personnel in each lab should also be identified in this section. Based on the information provided in the BRD, it is not clear if all of these steps were performed. All audits should be documented in an appendix to the BRD. Ideally, an independent audit of all data from all laboratories should be conducted upon the conclusion of an interlaboratory validation study.

The following statement is included in Section 7.4 of the BRD: “since the updated classification system was developed after testing was complete, these substances were not retested.” This seems problematic from a validation study perspective. However, as previously discussed in the response to Question 5d, the Panel affirmed that the changes made were performed in an appropriate manner and recommended that the explanation be added to the BRD.

Section 7.5 of the BRD should include a statement about the availability of audit results, statistical evaluations, and methods for calculations. Additionally, the time requirements and location(s) for study data retention/storage should be described, as well as the secondary location for storage of backup copies of study data.

8. Consideration of All Available Data and Relevant Information

- a. **Based on available information contained in the draft BG1Luc ER TA BRD, have all the relevant data identified in published or unpublished studies that employ this test method been adequately considered? Are there other comparative test method data that were not considered in the draft BRD but are available for consideration? If yes, please explain how to obtain such data.**

To the extent the Panel can determine, all the relevant data identified in published studies that employ this test method have been adequately considered. The Panel is not aware of any existing unpublished studies.

Quantitative structure-activity relationship (QSAR)-based predictions for ER binding on some of the substances tested in this study are likely available in the literature, and a comparison to BG1Luc ER TA and RUC ER binding results would be of value.

9. Animal Welfare Considerations

- a. **Is the extent to which the BG1Luc ER TA test methods will reduce, refine, or replace animal use adequately characterized and discussed? If not, then what should be added?**

The extent to which the BG1Luc ER TA test methods will reduce, refine, or replace animal use requires further discussion. In order to fully understand how this method will impact the 3Rs, the

Independent Peer Review Panel Report

context into which this test method would fit into the overall EPA EDSP Tier I screening battery is needed. For example, if the BG1Luc ER TA test method is just added to the battery of individual screening assays, the BRD should simply state that it would not reduce animal use at this time. However, additional consideration of strategic testing schemes provides possible approaches by which a validated BG1Luc ER TA test method could contribute to reducing animal use. For example, these could include:

- Defining whether the BG1Luc ER TA test method will be performed prior to or simultaneously with the other *in vitro* and *in vivo* assays
- Defining the implications for possible Tier II testing, which will likely incorporate significant animal use

Comments provided in the EPA’s 2011 budget (EPA 2011) state the EPA’s intent to phase out the EDSP and to include high-throughput ED-detecting assays in the ToxCast screening battery, which could lead to reductions in animal use for regulatory testing over the long term. The BG1Luc ER TA test method could contribute to the development of this screening battery.

A major limitation of *in vitro* EAC testing and the EDSP screening battery is the lack of an *in vitro* method to evaluate metabolism. The inclusion of a metabolism component as part of any *in vitro/in silico* EAC test battery will provide a more biologically relevant assessment of EAC activity. Because there are many *in vitro* metabolism methods being used in drug development and being used for *in vitro* testing of other toxicity endpoints (e.g., genotoxicity), the importance of including metabolism in all future assays for *in vitro* EAC activity needs to be discussed in the BRD.

Implications of a validated BG1Luc ER TA test method for use in the European Union and Japan were not discussed in the BRD. An evaluation of potential EACs is required under the EU chemical regulation REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals) (Bars et al. 2011; European Commission 2006; Kovarich et al. 2011; Marx-Stoelting et al. 2011), and validated *in vitro* methods may reduce animal use in this kind of testing program. “REACH-type” programs are also being adopted by Asian countries, so the availability of validated *in vitro* and *in silico* methods to screen/prioritize chemicals for these testing programs has further potential to reduce animal use. Following validation, the development of *in vitro* EAC assays into an OECD Test Guideline will broaden their potential for reducing animal use.

The concordance of the BG1Luc ER TA test method with the RUC ER binding assay of 97% (33/34 substances) suggests the BG1 assay could “replace” the rat ER binding assay within the EDSP Tier 1 battery (and similar test batteries). This excellent concordance of the BG1Luc ER TA test method exceeds the “performance” of other methods that have been endorsed as scientifically valid. If necessary, additional retrospective data analyses could be conducted. This topic was extensively discussed by the Panel, and additional Panel recommendations are provided in the Overview and Sections II.1.d and II.5.f of this report.

Based on the concordance of the BG1Luc ER TA test method with the rat uterotrophic assay (92%, 12/13), the BG1Luc ER TA test method is a potential candidate for reducing the use of the uterotrophic assay, which in turn would reduce animal use. The small data set, however, is not sufficient to recommend endorsement of BG1Luc ER TA as a replacement at this time. Therefore, the BG1Luc ER TA should be considered as a high priority for additional studies. A retrospective analysis may be sufficient. However, if necessary, a prospective study could be conducted to further compare these methods, preferably by identifying additional materials already evaluated in the uterotrophic assay. *In vitro* metabolism is an essential component of any prospective study that compares the BG1Luc ER TA to an *in vivo* assay. Also, since the BG1Luc ER TA was already subjected to an extensive interlaboratory study, consideration for an abbreviated assessment (e.g., 1 or 2 laboratories) could be considered to reduce time and costs.

Independent Peer Review Panel Report

Comparisons of the BG1Luc ER TA test method to the uterotrophic bioassay (or any other *in vivo* test results) would benefit from including analytical methods to measure the extent of free chemical exposure to target organs *in vivo* or to the cells in culture (Gülden and Seibert 2003).

The following text is suggested as a concluding paragraph for BRD Section 9:

The development of a battery of *in vitro* and *in silico* methods that can replace animal testing for detecting chemicals that have the potential to interact with the endocrine system (i.e., endocrine active chemicals [EAC]) is a biologically complex problem. For example, a method for assessment of metabolites needs to be included with the *in vitro* assays, and assays for assessing the many modes of action of EACs on various tissues and species need to be developed and validated. The experience derived from validating and using the *in vitro* BG1Luc ER TA test method is expected to contribute to our knowledge and promote progress toward this goal. It should lead to the broader use of cell-based methods for EAC screening, and could include the use of cells from other species.

10. Practical Considerations

a. Are the practical considerations associated with the BG1Luc ER TA test method adequately described and are there any points that would seem to preclude its transferability and implementation by other laboratories?

There are a few practical considerations that should be addressed:

- There needs to be better clarification regarding the availability of the BG1Luc4E2 cell line. This cell line has not been placed in a public repository and therefore additional efforts should be undertaken to do so or to otherwise ensure the continued availability of the cell line. It is the understanding of the Panel that the BG1Luc4E2 cell line is available upon request for a reasonable fee. Written confirmation of this needs to be obtained, and indicated as such in the BRD.
- The costs of equipment and supplies are provided, but these costs should be associated with the date they were acquired (i.e., it is not clear if the costs provided are current). This information should be updated in the BRD. Additionally, it should be clarified that the vendors used during this update are suggestions only and that equipment and supplies are not restricted to specific vendors.

b. Are the apparent level of training and expertise required to conduct the BG1Luc ER TA test method reasonable for its wide use?

The level of training and expertise required to conduct the BG1Luc ER TA test method are reasonable for its wide use. In order to minimize variability, the protocol should place greater emphasis on the need for adherence to Good Cell Culture Practices, Good Laboratory Practices, training of personnel and contamination control. Best practices for data analysis, presentation and data illustration (graphs, tables, etc.) should also be used.

Independent Peer Review Panel Report

III. Draft ICCVAM Test Method Recommendations on the BG1Luc ER TA Test Method to Identify Substances With Estrogen Agonist and/or Antagonist Activity

1. Test Method Usefulness and Limitations

a. Do you agree that the available data and test method performance (accuracy and reliability) support the ICCVAM draft recommendations for the BG1Luc ER TA test method in terms of the proposed test method usefulness and limitations?

Based on overall test method performance as presented in the BRD, the BG1Luc ER TA test method is capable of identifying substances as potential *in vitro* ER agonists or antagonists, as is recommended by ICCVAM, provided the acceptance criteria described in the recommended test method protocol are met.

However, careful analysis of cytotoxicity is critical for correctly interpreting results, particularly with regard to the antagonist protocol. While it appears the qualitative approach can be used successfully for some chemicals, as indicated in the responses to Question II.4, a quantitative approach is recommended as the primary method for evaluating cytotoxicity. In addition, accessibility to the cell line still needs to be resolved, which clearly can impact the future use of this assay, including its use as a screening test in a contract laboratory setting.

The Panel also noted that limitations due to interlaboratory variability may not be fully explained by the data analysis provided in the draft BRD. In addition, test method accuracy statistics (concordance, false positives and negatives, etc.) need to be updated to include the confidence intervals.

If ICCVAM concurs with the Panel's recommendations, then the recommendation that BG1Luc ER TA be considered for replacing the RUC ER binding assay in test batteries and tiered testing schemes such as EDSP Tier 1 battery needs to be moved to this section of the ICCVAM recommendations.

2. Test Method Protocols

a. Do you agree that the available data support the ICCVAM draft recommendations for the BG1Luc ER TA test method procedure in terms of the proposed test method standardized protocols? If not, what recommendations would you make?

The available data support the ICCVAM draft recommendations for the BG1Luc ER TA test method procedure in terms of the proposed test method standardized protocols with the caveat that the visual assessment of cytotoxicity needs to be better standardized and validated for chemicals or replaced by a quantitative method (see related comments on the test method protocol in Section 2b and below relative to future studies). Additionally, appropriate quality control measures should be further emphasized in the protocols to ensure reproducibility (i.e., GLPs, GCCPs, training, and contamination control).

3. Future Studies

Do you agree that the available data support the ICCVAM draft recommendations for the BG1Luc ER TA test method in terms of the proposed future studies? If not, then what recommendations would you make? Please explain your answer.

The available data support the ICCVAM draft recommendations for the BG1Luc ER TA test method in terms of the proposed future studies. Before additional work is recommended, it is important that all efforts have been made to make use of information that is available.

Independent Peer Review Panel Report

In addition to those proposed studies, recommended future studies include the following:

- Efforts to assess the utility of the current visual assessment of cytotoxicity evaluation for chemicals, or a quantitative method, should be assessed for its replacement.
- Future studies to account for compound metabolism/activation could expand the utility of this and other ER TA methods.
- A search for fully defined media to replace fetal bovine serum is recommended.
- A study to assess protocol components that might impact intra- and/or interlab variability in BG1Luc ER TA test results is recommended.
- An effort to expand the reference substance list, and subsequently the BG1Luc ER TA test results, with additional negative agonist and positive antagonist test substances is recommended. Pazos et al. (2010) is a possible source for additional reference substances.

4. Performance Standards

ICCVAM has developed draft performance standards consisting of essential test method components, a minimum list of reference substances, and expected accuracy and reliability values. These are proposed for evaluating the acceptability of proposed test methods that are mechanistically and functionally similar to the BG1Luc ER TA test method. **The overall question for the Panel is do you consider these performance standards adequate for assessing the accuracy and reliability of test method protocols that are based on similar scientific principles and that measure the same biological effect as the BG1Luc ER TA test method?**

a. Do you agree with the selection and prioritization criteria used to select the performance standards reference substances?

In general the criteria used to select the performance standards reference substances are adequate. As previously noted, the use of the majority classification criteria (i.e., greater than 50%) among results to establish the consensus reference determination that was assigned to each reference substance is a reasonable strategy. However, the criteria used to evaluate the quality of the data obtained from the literature to determine the reference classification could impact the appropriateness of such a strategy. The use of a ranking method such as using Klimisch criteria (Klimisch et al. 1997), which focuses essentially on the reliability of the data, would strengthen the resulting activity determination. In addition, some type of sensitivity analysis could be performed that would be more robust than a simple majority classification.

b. Do you consider the number of substances included in the list of reference substances to be an adequate number upon which to evaluate the performance of functionally and mechanistically similar test methods? If not, how many reference substances should be tested?

Based upon the currently available data, the list of reference substances upon which to evaluate the performance of functionally and mechanistically similar test methods is adequate. Ideally, more negatives should be included, but the Panel recognizes that data on such substances are not currently available. See additional discussion below.

Independent Peer Review Panel Report

- c. Do you consider the types of substances included in the reference substance list, with regard to relative estrogenic activity and physicochemical characteristics to be representative of the overall diversity of substances that are likely to be tested for potential estrogenic activity?**

The reference substance list has reasonable overall diversity and reflects the extensive effort to obtain all available relevant information. The list allows for a wide range of estrogenic activity over several orders of magnitude, as well as a few confounders to assess the robustness of the assay and its methodology (e.g., TPA, sodium azide, ammonium perchlorate, cycloheximide, and actinomycin D). However, there may be an opportunity to revisit the list of reference substances and make modifications based on experience gained in the assay subsequent to future testing. For example, more negatives and proportionally fewer positives should be included in the list for agonist testing when possible.

- d. Are there other types of information relevant to estrogen agonist or antagonist activity that should be considered in order to demonstrate an adequately diverse reference list? If yes, please explain what additional information should be included.**

As noted above, the reference list is adequately diverse, but it would benefit from the inclusion of additional negative substances once they become available.

- e. “Discordant chemicals” are also included in the reference list as substances that could be studied to evaluate if the proposed modifications might provide improved performance relative to the BG1Luc ER TA test method. Please comment on the appropriateness of including these specific substances in the reference list. Should more and/or different substances be included? If so, how many more and what are they? Should testing these substances be required?**

Discordant chemicals should be included on the reference substance list. These discordant chemicals, which include weakly active or non-active chemicals, transcriptional inhibitors, and general cytotoxicants, are critical for truly characterizing the limitations of the assay. While more compounds of a similar type might have some added value, there are enough to lend sufficient robustness to an assessment. In order to place any “discordant” results into proper context, some metric of potency and intrinsic activity should be included in the accuracy evaluation.

- f. Are there any substances on the proposed reference list for which a discordant result would be considered unacceptable and would therefore signal that a proposed test method is not considered scientifically valid, regardless of its overall performance?**

The potent estrogens on the reference list should not be missed. There could be some tolerance for discordance for the weakly active reference substances. Therefore, any “discordant” results should be discussed in terms of the ability of the test method to detect a similar range of potencies and intrinsic activity, as well as the chemical/product class. This could also facilitate the interpretation and utility of the data in a possible “weight of evidence” assessment of comparative assays.

- g. Do you consider the number of repeat experiments to be adequate to evaluate intra- and interlaboratory reproducibility?**

The number of repeat experiments to evaluate intra- and interlaboratory reproducibility appears to be a good start. The intralaboratory and interlaboratory reproducibility assessments allow for a reasonable evaluation of the methodology on the validation reference set of materials.

However, as discussed in Section II.6 of this report, while reproducibility has been addressed on three data levels (raw data [e.g., DMSO control measurements], derived endpoints [e.g., EC₅₀], and

Independent Peer Review Panel Report

prediction of estrogenic activity), the analysis is descriptive. No formal inference has been carried out. The descriptive approach is acceptable; however, there are some additional analyses that could be performed to better understand these components of assay response variability.

5. Additional Comments on Essential Test Method Components

The justification for some of the essential test method components is not clear. For example:

- The type of cell line does not have to be specified in the performance standards. It is critical to ensure that any cell line used incorporates the appropriate “endogenous cellular machinery,” as stated in the BRD, but the specific tissue source, type, and species may not be critical.
- Evaluation of cytotoxicity is necessary. The Panel recommends a quantitative approach be used when developing new ER TA methods, particularly with regard to an antagonist protocol, where false positives can result from cytotoxicity.

6. Additional Comments on Test Method Performance

The criteria for accuracy included in the performance standards should be based on meeting or exceeding the lowest accuracy that was noted among the participating laboratories. As noted previously, the Panel strongly supports the quantification of activity as a complement to the dichotomous call of a positive or negative response. Interpretation of the results should not rely solely on statistics but also on scientific judgment and should incorporate consideration of the nature and shape of the dose-response relationship and, if needed, the reproducibility of the response in independent experiments. A test of trend might be preferred to the BRD’s suggestion of “three points with non-overlapping error bars.”

- The BRD should define what statistic corresponds to the error bars included in the criteria for a positive response (i.e., standard deviations, standard errors, half-width of confidence intervals). Additionally, the idea of nonoverlapping confidence intervals is a conservative way to declare difference between parameters and is not clear guidance.
- A metric of potency and intrinsic activity should be included in the accuracy evaluation; appropriate measures of uncertainty should also be included. Therefore, the reporting requirements should include the metric of potency that is used.
- Sensitivity analyses based on EC₅₀ (agonist) or IC₅₀ (antagonist) should also be considered.

Independent Peer Review Panel Report

IV. References

- Assessing Agreement (Special Issue). 2007. J Biopharm Stat 17(4).
- Bars R, Broeckaert F, Fegert I, Gross M, Hallmark N, Kedwards T, et al. 2011. Science based guidance for the assessment of endocrine disrupting properties of chemicals. Regul Toxicol Pharmacol 59(1): 37-46.
- Bland JM, Altman DG. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1: 307-310.
- Bland JM, Altman DG. 1999. Measuring agreement in method comparison studies. Stat Methods Med Res 8: 135-160.
- Bretz F, Hothorn LA. 2003. Statistical analysis of monotone or non-monotone dose-response data from in vitro toxicological assays. Altern Lab Anim 31(Suppl) 1: 81-96.
- EPA. 2009. Endocrine Disruptor Screening Program Test Guidelines. OPPTS 890:1300: Estrogen Receptor Transcriptional Activation (Human Cell Line (HeLa-9903)). EPA 740-C-09-006. Washington, DC:Office of Prevention, Pesticides and Toxic Substances, U.S. Environmental Protection Agency.
- EPA. 2011. FY 2012 EPA Budget in Brief. EPA-190-S-11-001. Washington, DC:U.S. Environmental Protection Agency, 60-62.
- European Commission. 2006. Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. OJ L 396.
- Gülden M, Seibert H. 2003. In vitro-in vivo extrapolation: estimation of human serum concentrations of chemicals equivalent to cytotoxic concentrations in vitro. Toxicology 189: 211-222.
- 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. NIH Publication No. 97-3981. Research Triangle Park, North Carolina:National Institute of Environmental Health Sciences.
- ICCVAM. 2003. ICCVAM Evaluation of *In Vitro* Test Methods for Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays. NIH Publication No. 03-4503. Research Triangle Park, North Carolina:National Institute of Environmental Health Sciences.
- ICCVAM. 2006. Addendum to ICCVAM Evaluation of *In Vitro* Test Methods for Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays. NIH Publication No. 03-4503. Research Triangle Park, North Carolina:National Institute of Environmental Health Sciences.
- ISO. 2009. ISO 10993, Biological evaluation of medical devices—Part 5:Tests for *in vitro* cytotoxicity. Geneva:International Organization for Standardization.
- Klimisch HJ, Andreae E, Tillmann U. 1997. A systematic approach for evaluating the quality of experimental and ecotoxicological data. Regul Toxicol Pharmacol 25: 1-5.
- Kovarich S, Papa E, Gramatica P. 2011. QSAR classification models for the prediction of endocrine disrupting activity of brominated flame retardants. J Hazard Mater 190(1-3): 106-112.

Independent Peer Review Panel Report

Marx-Stoelting P, Pfeil R, Solecki R, Ulbrich B, Grote K, Ritz V, et al. 2011. Assessment strategies and decision criteria for pesticides with endocrine disrupting properties relevant to humans. *Reprod Toxicol* 31(4): 574-584.

Pazos P, Pellizzer C, Stummann TC, Hareng L, Bremer S. 2010. The test chemical selection procedure of the European Centre for the Validation of Alternative Methods for the EU Project ReProTect. *Reprod Toxicol* 30(1): 161-199.

Ritz C. 2010. Toward a unified approach to dose-response modeling in ecotoxicology. *Environ Toxicol Chem* 29(1): 220-229.

Rogers JM, Denison MS. 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 Mol Toxicol* 13: 67-82.

Independent Peer Review Panel Report

This page intentionally left blank

Independent Peer Review Panel Report

Appendix A
Errors and Omissions

Independent Peer Review Panel Report

The Panel has identified several errors that should be corrected, as well as omissions of existing relevant data or information that should be included. These are detailed below.

General Format

In general, table titles are not explicit enough and should better describe the content of the tables.

Throughout the text the use of the wording “normal” should be discouraged:

- “... the ~~normal~~ function ...”.
- “... interfere with ~~normal~~ estrogen activity. ...”

Preface

Page xvii

Spell out ER, AR and TA at first use; add citations for BRDs (3rd line up from bottom)

Page xvii (last paragraph)

Add “identification of substances with *in vitro* ER agonist or antagonist activity, as predictive of activity *in vivo*”.

Page xviii - top line

Add citation

Page xviii - last paragraph statement

“...for identification of substances with *in vitro* ER agonist or antagonist activity”

Suggest deleting “*in vitro*” from this and other similar statements in the BRD as this conflicts with your own definition of the scientific basis for the BG1Luc ER TA Luc ER TA on page 1-11, which says the *in vitro* ER TA assays are to identify... that might interfere with normal estrogen activity *in vivo*. Including “*in vitro*” also conflicts with the original described goals of the EPA for this kind of assay to be validated “to determine whether certain substances may have an effect in humans....” - see page 1-2.

Page xix - 6th line from top

“these proposed...” should be “the proposed....”

Page xix

The second charge to the panel (3rd line from top) is “assessing the extent that established validation and acceptance criteria have been adequately addressed.” To do this NICEATM/ICCVAM selectively provided the peer panel with a list of questions that effectively cover some of the criteria but not others. The second question in this set of “questions to the panel” should list the validation and acceptance criteria, so that the panel members can easily review them and respond effectively to this second charge identified for the panel. The remaining NICEATM questions can then follow.

Executive Summary

The Executive Summary should be a stand-alone document and provide a clear explanation of how the assay is used to identify a substance as positive or negative as well as how substances were classified as positives or negatives. This information will improve the reader’s understanding of results.

Page xxi

Spell out ER, AR and TA at first use; add citations for BRDs (first paragraph)

This appendix documents errors and omissions identified by the Panel, as communicated to NICEATM. Editing of Panel comments by NICEATM has been limited to only that necessary to ensure clarity.

Independent Peer Review Panel Report

Page xxii (end of paragraph just before protocol section)

The ES does not make it clear to a reader, as a stand-alone document, the role of the peer review and the Panel and that this is a draft BRD. Therefore, add statement something like this to the end of the paragraph: “This draft version of the BRD will be reviewed by an international independent scientific peer review panel (Panel) that will meet on March 29-30, 2011. The Panel’s comments will be considered by ICCVAM before ICCVAM recommendations on the BG1Luc ER TA test method are finalized.”

Page xxii

Protocol section should briefly explain how the assay is used to identify a substance as positive or negative, and why the qualitative endpoint is being used for the proposed validated assay.

Page xxii-xxiii - Substances used

Add citation for identification of 78 substances; add statement explaining how substances were classified as +/- for ER TA activity; the numbers don’t add up - make it easier for reader to understand the numbers of positives and negatives identified and used in this study.

Page xxiii - Accuracy section

More clearly define the reference method - “preponderance of published data” - of what kind of data? Average response from 3 labs should not be used. In real testing one lab will be used, therefore each lab should be assessed independently for correlation to reference method(s). First time CERI STTA is mentioned it should briefly be explained as the only currently accepted *in vitro* method.... The numbers in this section don’t add up without further explanation - check their correctness and simplify this for the reader.

Page xxvi (last paragraph)

Add “antagonist” to this statement: “....identify substances with estrogen agonist and antagonist activity.”

Page xxvi (last paragraph)

Add statement something like this to the end of the page: These ICCVAM draft recommendations will be reviewed by the scientific peer review panel at the March 2011 meeting, and the Panel’s comments will be considered by ICCVAM before they are finalized.

1.0 Introduction

The Introduction should provide the advantages of the BG1Luc ER TA test method. Included in this information, this section should include some of the background behind the preference for a TA assay over a binding assay (there are parts to this here and there) but might consider pulling some from the Rogers paper:

Binding assays and caveats leading to why TA assays are better than binding assays (agonist and antagonist activity can be assessed, done at physiological temperatures etc).

Transient transfection assay vs. development of a stable cell line and why a stable cell is better.

For this ER TA cell line the properties that make it best suitable for this use (little metabolic activity, appropriate transcription machinery for hormone responsive cells, no other steroid hormone receptors are expressed to confound the assay, etc.)

Page 1-2: 3rd line

Delete “*in vitro*” for same reason previously discussed

Line 33-35 in the Public Health Perspective section (1.2)

Says “EDs are widespread in our environment...”, though most of the substances have not yet been judged to be endocrine disrupting, since the definitive test methods have not been established.

Diethylstilbestrol is obviously an endocrine disruptor by its clinical and epidemiological evidence,

Independent Peer Review Panel Report

and bisphenol-A may deserve to be an ED by many supporting experimental results, but other chemicals are in purgatory so far. Thus, “EDs are widespread ...” should be “Potential EACs are widespread...”

Page 1-2: Historical Background

Lists 3 Acts as requiring EPA to develop EAC test methods, and then the next statement refers to only one Act. Please clarify and correct.

Pages 1-2 and 1-13

Remove the Fenner-Crisp and Fisher reference, which appears to be only an abstract and not needed to define EACs.

The BG1Luc ER TA Luc ER TA should be defined as an appropriate screen (section 1.5.2) to be integrated in an accurate, comprehensive and cost-effective *in vitro* testing battery (section 1.5.5)

Section 1.4 describes that the interlaboratory validation study was conducted by three laboratories in three different regions. Why they were restricted to three, and do three suffice the validation? I am not making an objection but want to confirm the background opinion.

Page 1-5

Delete “academia” from 4th paragraph unless you can identify academic labs that conduct high throughput testing of chemicals.

Page 1-6: section 1.5.2

Move this statement “An appropriate screen such as....” to the end of the section, and revise as follows: Depending on how it is used, an appropriate screen such as BG1Luc ER TA Luc ER TA has the potential to limit human and ecological.....

Page 1-7: section 1.5.3

Not sure last statement is correct - how would BG1Luc ER TA assay eliminate need for testing in these [Tier 1?] animal models?

Page 1-3

A section titled “Regulatory rationale for BG1Luc ER TA” should be added to the Introduction section. The EPA EDSP is the motivation for this entire decade-long validation process and relevant information from the EPA EDSP should be summarized in Section 1.0, including a summary of the regulatory use of the current and the proposed EDSP test batteries. Especially relevant are: a) the recent policy document regarding how to use Tier I data - this should be summarized for easy understanding, b) the 11 Tier 1 assays were [pre?]validated by the EPA/OECD, and c) when is Tier 2 testing required? The validation of the CER1 ER TA reference method should be described and references to relevant documents provided. Current and future OECD TG activities could be described in this section, as well as EU and Japan regulatory implications.

Section 1.7 describes more precisely the interlaboratory validation study. It is not clear how the laboratory qualification was done. It seems to have been done in the Phase 2 depicted in Fig. 1-1, but it is better to be stated also in the text.

Page 1-8: section 1.5.5

Seems to exaggerate savings; BG1Luc ER TA as part of a battery of *in vivo* and *in vitro* tests would not provide cost and time savings; its potential use as a screening assay compared to an animal test might....If BG1Luc ER TA leads to development and validation of a more complete battery of *in vitro* methods or tiered testing scheme then the potential to reduce animal use would be more likely.

Page 1-9: section 1.7

The statement about KoCVAM should be removed. Other table entries are not described in this degree of detail in the text so it seems out of place.

Independent Peer Review Panel Report

2.0 Protocol Components

Discuss the sequence of the vector itself and that it has been closely examined for the presence of cryptic activation sites which could confound the data (I assume this was done as with computers it is quite easy to assess – not like the old days). The use of 4 ERE's with varying orientations in the construct design should be discussed.

A section on reference method protocol(s) is missing from the BRD.

Protocol section should clearly explain when over the course of the study changes were made to the protocol and test method evaluation criteria, and why these changes would not affect study results that are being evaluated to assess the scientific validity of the assay. Validation study guidance states that: "Prevalidation is the process by which testing laboratories are selected and demonstrate competence in performing the testing procedures, and during which the test protocols are standardized. It is important that this be established in advance of formal validation procedures" (http://iccvam.niehs.nih.gov/docs/about_docs/validate.pdf, page 33).

Section 2.3: Preparation and use of cells

The BRD protocol section does not adequately describe the cell culture procedures (split ratios, number of cells seeded per well, storage conditions, passage numbers acceptable for validated assays, etc.)

Section 2.2.1

Explain how the cell line was characterized including relevant citations, and state what components are critical to the development of a "replacement" cell line (cell line characterization criteria). This information is needed for development of the performance standards.

Section 2.2.1

Should identify source(s) for the cell line.

Table 2-4

Add column showing corresponding % cell viability.

3.0 Substances Used

L7: complete the dotted line.

L8: "... EA TA ..." should be ER TA.

L10: "... EA TA ..." should be ER TA.

L13 and 14

Consistency between both titles "... to assess ..."

Also correct the title accordingly in the text.

L27

Citation order 2002 a to d.

Page 3-9, line 54-55: add citation.

Footnote of table 3-2

L96-97: ER = ~~endocrine~~ estrogen (?) receptor

L100, 101, and 102: (2010) Could this citation be incomplete?

L147: "... EA TA ..." should be ER TA.

L156: ~~... and vice versa ...~~ delete.

L176: "... EA TA ..." should be ER TA.

Independent Peer Review Panel Report

L190-191: "... in conjunction with CERI ..."

L194-199: The number of substances listed by the bullets is 13, not 15 (?).

Table 3-3, 3-4, 3-5, 3-6

Add a footnote to explain the column "ICCVAM Consensus Classification". For example: estrogenic activity based on a literature review for effects on ER binding, ER-TA based on CERI, and uterotrophic response.

Consider the following to reduce the number of tables: The only new information in Table 3-5 and 3-6 is the Study Phase. Table 3-5 and 3-6 can be deleted and replaced by a single column inserted in Table 3-3 and 3-4. There is enough space to insert this column in Table 3-3 and 3-4. For example, 17 β -estradiol is listed as agonist in Phase 1 and 3, therefore in this new column the number "1,3" can be added beside 17 β -estradiol.

A footnote can be added to the table to explain the implication of the Study Phase.

In general the Table titles poorly describe the content of the tables.

L221: "Substances Used to Assess ~~for~~ ...". To be consistent with title 3.6.

L232-250

If needed, this is an area where the length of the document can be reduced by presenting Section 3.7 and 3.8 into an annex. This new annex can be cited with the description of Table 3-1. The information presented in section 3.7 and 3.8 is redundant with the information already presented in Table 3-1. The text in both sections (3.7 and 3.8), as well as the "incomplete" titles of Table 3-7 and 3-8, failed at justifying the repetition of the same data but in a different format. It is understood that the idea is to demonstrate that the validation process involved a wide distribution of chemicals based on chemical and product classes, but Table 3-1 was sufficient to achieve this. In addition, the unavoidable fact that the same chemical can be assigned to more than one category is reducing the importance of these tables.

L262: methyl-testosterone

In section 3.9 it says that ECVAM and JaCVAM procured "controlled substances" from their regional suppliers. This raises the question how ECVAM ensured that the identity of the tested substances was not known to the (in-house?) ECVAM laboratory. This should be clarified in the BRD.

4.0 Test Method Data and Results

L20-21: Delete, not useful.

L76-77: "Hiyoshi reported the lowest values ...". This is not obvious, is this useful?

L103: "... (~~i.e., the methoxychlor control must be positive~~)."

L147

Not clear, rewrite adjustment calculation. What is meant by "the maximum Ral response to 10,000...". Perhaps an example could be provided.

L160-161: Not clear. Perhaps an example could be provided.

L200 and L204: Flouranthene should be Fluoranthene.

Table 4-11. "Agonist Range Finder Starting Concentrations ~~Based on Solubility in the culture medium~~".

Table 4-12 "Antagonist Range Finder Starting Concentrations ~~Based on Solubility in the culture medium~~".

Independent Peer Review Panel Report

L224: "... comprehensive testing. This ~~which~~ was ..."

L231: The reference to Table 4-11 is not appropriate, it should be 4-13.

L235: The reference to Table 4-12 is not appropriate, it should be 4-14.

L250-251: The table numbers are wrong and should be 4-13, 4-14, and 4-15.

L266: this footnote should refer to IC50 not EC50.

Section 4.4.1

According to table 4-14 substance 17-alpha-estradiol was not "negative" but "inadequate" at Hiyoshi.

Section 4.4.1

According to table 4-14 substance clomiphene citrate was not "negative" but "inadequate" at XDS.

Table 4-14: Coumestrol at ECVAM is N(0/2). Why is it a "negative" if no runs support this result?

In table 4-14

ECVAM classification of p-n-nonylphenol is N (2/3), but in 6.1.7 it says that there were no "inadequate" data generated at any lab during analysis of antagonist reference substances and that there was 100% agreement within each lab for each of the three repeat tests. But does N(2/3) not mean +-- or I--?

Table references in section 4.4.1

5.0 Accuracy

Page 5-2

Would be good to start this section out with the definition of accuracy from validation guidance document and a citation to the guidelines you used (ILS said OECD).

Page 5-2, lines 32-34

Pos/Neg calls were based on average results from 3 labs does not appear valid way to evaluate the data. When an organization submits a sample for testing, it will be submitted to one lab. They need to know what confidence they can have in getting the correct result from one lab (this issue overlaps with inter-lab variability assessment). A justification could be included here explaining that the BG1Luc ER TA is being evaluated as a screening assay rather than a replacement assay, and that there is usually no exact way to assess performance of an assay that is part of a battery or tiered scheme. Additionally, NICEATM/ILS developed the novel approach of developing weight-of-evidence results from the literature, etc.

Pages 5-7 and 5-8

Replace Tables 5-4 and 5-6 with tables presented at meeting showing accuracy calculations for individual labs and combined.

Table 5-2

Discuss why negative substances are likely to be I or NT. Why are there so few negative substances tested and how will this effect the results?

Table 5-7: alpha symbol for estradiol and ethinyl estradiol.

Page 5-7, line 100: after "antagonist testing" add (Table 5-3)

L130: "... listed in Table ~~5-5~~ 5-7."

L136: "... listed in Table ~~5-5~~ 5-7."

Independent Peer Review Panel Report

L151

The following addition can be considered: “Although EC50 values can differ by a log between methods (even two logs for norethynodrel), this relatively high correlation ...”

Table 5-10

The origin of the EC50 values can be added as a footnote (median XDS, ECVAM, Hiyoshi (?)).

L203: Owens and Koester 2003 (same modification required in the list of references).

In appendix B it says “The model calculates the best fit for the Top, Bottom, HillSlope, and EC50 parameters. See Section 11.6.5 for more details.” Where is section 11.6.5?

Classification criteria are given in section 2.7.1.5. Please explain what you mean by “error bars”. Standard deviation? Standard error?

Table 5-2

Inconsistencies with table 4-13, e.g. 17-alpha-Estradiol POS(2/2) at Hiyoshi in table 5-2, but P(1/1) in table 4-13.

Table 5-3

Inconsistencies with table 4-14, e.g. 4-Hydroxytamoxifen POS(1/1) at XDS in table 5-3, but P(3/3) in table 4-14.

Table 5-7: explain how BG1Luc ER TA and CER1 results were determined.

Tables 5-8 and 5-9

Errors - Table 5-7 lists 27 total test substances (not 26 like used in the tables) and 5 negatives (not 4 like used in these tables). If correct, the calculations shown need to be revised and the numbers reported on lines 117-119 corrected.

Section 5.7

Mention and cite results of additional paper comparing Lumi-Cell assay with uterotrophic assay (Jefferson, Padilla-Banks et al. 2002) that was mentioned in the XDS submission (Annex A).

There have been sufficient toxicity test method validation studies resulting in endorsed alternative methods so that the conclusions for BG1Luc ER TA accuracy should be discussed in terms of how the BG1Luc ER TA performance statistics compare to previously endorsed *in vitro* test methods.

6.0 Test Method Reliability

Section 6.0, lines 19-20

Cite validation guidance document rather than secondary source for definition of reliability.

Section 6.1.3

Add table showing mean, range, and CVs for agonist test substances in the 3 labs (or refer to tables in Section 4.0). These are the real intra-lab variability data to be evaluated in this section of the BRD.

Section 6.1.7

Add table showing mean, range, and CVs for antagonist test substances in the 3 labs (or refer to tables in section 4.0). These are the real intra-lab variability data to be evaluated in this section of the BRD.

Section 6.2.2 (agonist phase 3)

Says that there was discordance among the laboratories for six substances: dicofol, fluoranthene, butylphenol, androstenedione, clomiphene citrat, resveratrol. What about dexamethasone? For this substance definitive results are available in table 4-13 for ECVAM and Hiyoshi, but results are discordant. This would change table 6-9 as well.

Footnotes of table 6-3

Says “each of three replicate tests”. But in some cases more than three replicates are available, e.g.

Independent Peer Review Panel Report

Corticosterone at Hiyoshi (Table 4-13). This means that for different compounds certainty about reproducibility is different. Same issue in table 6-7.

Section 6.1.5, Table 6-5: add units.

Section 6.1.6 line 185: “As indicated in **Table 6-3...**” Is this the right table?

Section 6.2.1 line 236-237: Are you referring to the correct table? There are no such results in Table 4-11 and 4-12 (?).

Section 6.2.1, Table 6-8: footnote “d” is wrong. ~~The substance was classified as negative in the third laboratory.~~

Section 6.2.2, line 269: “(... defined in Section ~~2.7.1.4~~.” should it be 2.7.1.5?

Table 6-9: footnote f an h: corrections needed.

Line 358-359

“... under the revised testing protocol, ...”. After the word protocol, insert within parenthesis the section where the revised protocol is described.

Section 6.1.7

Says “Although the classifications for some of the test substances differed among the laboratories”, but according to table 4-14 and section 6.2.1 “Among the substances tested for antagonist activity, there was 100% agreement among the three laboratories for all 12 substances.”

Section 6.1.7

Says “... within each laboratory for each of the three repeat tests”, but according to table 4-14 some substances were tested 4 times.

Table 6-2

Plates 2 and 3 of ECVAM have identical means and SDs and the means of plates 4 and 5 are clearly different. In contrast, in figure 6-2 means of plates 2 and 3 are different and means of plates 4 and 5 are similar. Also inconsistencies in other tables? E.g. table L-3, plates 6, 7,10 of Hiyoshi (compare with figure 6-3).

Table 6-2

Mean E2 EC50 of ECVAM is 1.1e-11. Are the horizontal lines in figure 6-4 means? If so, In figure 6-4 the respective ECVAM mean seems to be smaller than 1.0e-11. In Table 6-6 mean Flavone Control Value for XDS is 3774, but in figure 6-10 it looks like the mean is smaller. Are means in the figures calculated without the values excluded to minimize scale distortion?

Footnotes of table 6-9

There are some inconsistencies (e.g. footnote h +--- , ...) that should be corrected.

Section 6.2.1, line 237

Reference to tables 4-13 and 4-14?

(p. 6-3, Figure 6-1) - What is the meaning of “plate” here? Should we expect that these values would be the same?

(p. 6-5, Figure 6-3) – All figures have a tremendous amount of wasted space (above 10000 in this figure). If the figures are rescaled, then the variability in points will be clearer. Figure 6-6 could be much improved.

(p. 6-6, Table 6-2) – SD of EC50 estimates is essentially a SE estimate. Are each of the constituent EC50 values estimated with the same precision?

(p. 6-7, Figure 6-5) – Footnotes describe points omitted to “minimize scale distortion” – this omission removes points that are potentially dramatic outliers.

Independent Peer Review Panel Report

There have been sufficient toxicity test method validation studies that the conclusions of intra and inter-lab variability could be discussed in terms of how the BG1Luc ER TA reliability compares to previously endorsed methods - CVs, % correct classifications, other analysis.

Additional suggestions for criteria to evaluate intra and intra lab variability were provided in Panel comments for Section 6.0.

7.0 Data Quality

Sub-section on Reference Methods Data Quality is missing.

Section 7.5

Add statement on the availability of audit results, statistical evaluations, and methods for calculations; descriptions for the specific length of time that study data will be stored (and where), and where data backups will be stored (a secondary location).

Line 122-124: This sentence does not clearly explain Table 7-4.

Line 145: "... provided in Section 2.7.1.3 ...". You probably meant section 2.7.1.4.

Line 169: "... defined in Section 2.12.3." This section does not exist.

Line 182-183:

"... presented in Tables 4-11 and 4-12...". Wrong tables, you probably meant 4-13 and 4-14 (?).

8.0 Other Scientific Reports

No edits.

9.0 Animal Welfare

The organization and wording of this entire section can be improved. A main consideration is that the content of this section does not appear to contradict the results or statements made elsewhere in the BRD.

Additional content is also suggested for this section in the Panel comments for Section 9.0.

Test Method Recommendations

L29: Perhaps the reader can be referred to the glossary for the definition of accuracy and reliability.

L32

Should "definitive results" be in quote, and added to the glossary to indicate that it is the data excluding rejected plates and inadequate data.

Lines 56-57 - numbers possibly incorrect.

L59-68 and L83-89

Perhaps these paragraphs can be used to create a separate section numbered 1.1.3 to discuss the limitations of the assay. Note that L61-64 list new experiments (mixtures, volatiles) to be presented with the others in section 1.3. Also, L62-64 (volatiles...) are redundant with L125-127.

L104 and L114

Verify both expressions on both lines, perhaps L114 should have been: "...100% specificity (no false positive)". (?)

L119

"... to ~~replace~~ reduce the need for the uterotrophic bioassay." What would be the alternative test if a substance cannot be tested in the BG1Luc ER TA assay?

Independent Peer Review Panel Report

Section 1.1.1: Lines 43-49 on test method reliability

May need to be revised following Panel discussions and revision of the draft BRD.

Lines 56-57 - numbers possibly incorrect.

Section 1.1.2

Lines 76-82 on test method reliability may need to be revised following Panel discussions and revision of the draft BRD.

Add Section 1.1.3

Add recommendation that BG1Luc ER TA be endorsed for replacing rat ER binding assay in test batteries and tiered testing schemes such as EDSP Tier 1 battery.

Performance Standards

Performance Standards section (Tab 6, p. xii, lines 272-307)

What is the justification for some of the essential model components? For example, why were these maximal concentrations suggested? Why log10 spacing of 7 concentrations? Evaluation of cytotoxicity is reasonable although modeling of the possibility of such a response could also be considered. Why is 20% of maximal response the cut-point for significant response?

ER TA antagonist testing (Tab 6, p. 6, lines 476-493) – concentration curve defined by a baseline followed by a negative slope – isn't it sufficient to require a negative slope? (will always have a baseline). A test of negative trend might be preferred to the suggestion of "three points with non-overlapping error bars." Error bars corresponding to what? Standard deviations? Standard Errors? Half-width of confidence intervals? The idea of non-overlapping confidence intervals is an incredibly conservative way to declare difference between parameters. This is not clear guidance.

Would Figure 1 (line 470) reflect a real study result? The guideline of log10 spacing of 7 concentrations wouldn't span the -13 to -4 log(conc) range depicted here. Add revised Figures 1 and 2.

Appendices

(Appendix B – Agonist Protocol, p. B-6, line 135)

A four parameter model is suggested here that captures a pattern of growth between two horizontal asymptotes. A couple of concerns: 1) won't "Bottom" = 0 and "Top" = 10000 when you do relative scaling of responses? If so, then you are not working with 2 parameters and not 4; 2) data are known to have a downturn at high doses – this is not consistent with the assumed model. How do you routinely address this? Are you dropping the highest concentration responses in this case? Could you model downturn at the higher concentration levels?

(Appendix B – Figures 12-2 through 12-7)

Better graphics to not have so much wasted white space in the graphs.

(Appendix B – p. B-39, Figure 13-1)

The layout of the test plate is unclear to me. Can more explanation of rows A-H and columns 1-12 be provided?

Independent Peer Review Panel Report

This page intentionally left blank

Independent Peer Review Panel Report

Appendix B

Peer Review Panel Member Biosketches

Independent Peer Review Panel Report

Panel Member Biosketches

A. John Bailer, PhD

Dr. Bailer received a PhD in 1986 from the University of North Carolina at Chapel Hill and is currently the Distinguished Professor and Chair, Department of Statistics at Miami University. He is a member of the American Statistical Association, International Biometric Society, International Statistical Institute (elected 2004), International Association of Statistical Education, Society for Risk Analysis, Society of Environmental Toxicology and Chemistry, and Sigma Xi. Dr. Bailer has served as a reviewer for numerous statistical and toxicological journals including *Biometrics*, *Risk Analysis*, *Environmental Toxicology and Chemistry*, *American Journal of Industrial Medicine*, *American Journal of Public Health*, *Occupational and Environmental Medicine*, *Environmental and Ecological Statistics*, *Environmetrics*, *Environmental and Molecular Mutagenesis*, *Environmental Health Perspectives*, *Fundamental and Applied Toxicology*, *Biological and Environmental Statistics*, *Journal of the American Statistical Association*, and *Journal of Pharmacokinetics and Biopharmaceutics*, *Cancer Research*. He has served as Secretary/Treasurer, Vice-President, and President of the Cincinnati Chapter of the American Statistical Association, and is a member of the International Statistical Institute Council (2009–2013) and is Council liaison, Committee on Statistics in the Life Sciences. Dr. Bailer has authored or coauthored 121 peer-reviewed journal articles, 4 books, and 34 book chapters and technical reports.

Christopher J. Borgert, PhD

Dr. Borgert received a PhD in 1991 from the University of Florida College of Medicine, Department of Pharmacology and Therapeutics. He completed a postdoctoral fellowship at the Center for Environmental and Human Toxicology at the University of Florida Research and Technology Park. He is currently the President and Principal Scientist at Applied Pharmacology and Toxicology, Inc. Dr. Borgert has served on numerous panels and committees focused on toxicological assessment of mixtures and human health impacts and was a member of the OECD Peer Review Panel for Validation of the Uterotrophic Assay, and the EPA Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC), the EDSTAC Screening & Testing Workgroup, and was Co-Chair of the EDSTAC Communication & Outreach Workgroup. Dr. Borgert has served as a reviewer for *Chemosphere*, *Environmental Research*, *Environmental Toxicology and Chemistry*, *Human and Ecological Risk Assessment*, *International Journal of Toxicology*, *International Workshop on Quantitative Structure-Activity Relationships (QSARs) in Environmental Sciences*, *Journal of Agricultural and Food Chemistry*, *Nonlinearity in Biology, Toxicology, and Medicine*, *Regulatory Toxicology and Pharmacology*, *Toxicological Sciences*, and *Toxicology and Applied Pharmacology*. He is currently a member of *Toxicology Forum*, *Society of Environmental Toxicology and Chemistry*, *Southeast Chapter, Society of Environmental Toxicology and Chemistry*, and the *Society of Toxicology*, *International Society of Regulatory Toxicology & Pharmacology*. Dr. Borgert has authored or coauthored 20 peer-reviewed journal articles, 5 book chapters, and has given 35 invited presentations.

Grantley D. Charles, PhD, DABT

Dr. Charles received his PhD from the University of Florida, Gainesville, in Pharmacology and Toxicology and completed a postdoctoral fellowship at the Dow Chemical Company. While at Dow, he received a Dow Chemical special recognition award for research conducted on endocrine disruptors. Dr. Charles is currently a Senior Scientist at Allergan, Inc. He was a member of the expert panel for the ICCVAM/NICEATM review of the Validation Status of *In Vitro* Test Methods for Detecting Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays. He also served as a member of the Scientific Advisory Committee on Alternative Test Methods (SACATM) for ICCVAM/NICEATM, and the American Chemistry Council (formerly Chemical Manufacturer's Association) ad hoc *In Vitro* Endocrine Disruptor subcommittee. He is a reviewer for *Toxicological Sciences*, *Journal of Toxicology and Environmental*

Independent Peer Review Panel Report

Health, Toxicology In Vitro, Journal of Pharmacy & Pharmacology, Reproductive Toxicology and Birth Defects Research, Expert Opinion in Pharmacotherapy, Expert Opinion in Drug Metabolism and Toxicology. He is currently a member of the Society of Toxicology and the Southern California Chapter of the Society of Toxicology and a Diplomate of the American Board of Toxicology. Dr. Charles has authored or coauthored 16 peer-reviewed journal articles and given 4 invited presentations.

Daniel Desaulniers, MSc, PhD

Dr. Desaulniers received a Master of Sciences degree in Veterinary Anatomy and Physiology and a PhD in Biomedical Sciences from Montreal University, St-Hyacinthe, Quebec. He is currently a Research Scientist, Health Canada, Healthy Environments and Consumer Safety Branch. His research has included using *in vitro* MCF-7 cell proliferation and the uterotrophic bioassay to evaluate the estrogenic activity of chemicals. He participated in the Mammary Gland Evaluation and Risk Assessment Workshop, was a reviewer for the OECD Test Guideline on the Uterotrophic Bioassay in Rodents, and was a member of the National Advisory Council for Environmental Policy and Technology (NACEPT) Endocrine Disruptor Methods Validation Subcommittee (EDMVS). He has served as a reviewer for 11 scientific journals, including Environmental Health Perspectives, International Journal of Toxicology, Teratology, Toxicology and Industrial Health, and Toxicological Sciences. Dr. Desaulniers has authored or coauthored 21 peer-reviewed journal articles, 5 government reports, and over 90 posters. He has given 16 invited presentations.

John Charles Eldridge, PhD

Dr. Eldridge received his PhD in Endocrinology from the Medical College of Georgia, Augusta and completed post-doctoral work as an Attaché de Recherche, Institut Nationale de la Santé et de la Recherche Médicale in Bordeaux, France. He is currently a Professor at the Wake Forest University School of Medicine in the Department of Physiology and Pharmacology. His interests include the endocrinology of reproduction, steroid hormone biochemistry and receptor mechanisms, and hormone and drug assay methodology. His research has focused on endocrine toxicology of xenobiotic chemicals, especially triazine herbicides, and hormones of stress and endocrine aspects of aging. Dr. Eldridge is a program reviewer for the National Institute on Aging, the National Center for Research Resources at NIH, the Health and Environmental Effects Research Laboratory (EPA), and is a member (ad hoc) of the FIFRA Science Advisory Panel (EPA). He is also a consultant (ad hoc) for reproductive and endocrine toxicology for Novartis Animal Health, Inc. and was a member of the EPA Endocrine Disruptors Methods Validation Committee. Dr. Eldridge holds or has held membership in numerous professional societies, including the American Society for Reproductive Medicine, the Endocrine Society (emeritus), Society for the Study of Reproduction (emeritus), the American Society for Andrology, and the Society for Neuroscience (emeritus). He is reviewer for a number of scientific journals including the Journal of Pharmacology and Experimental Therapeutics, Environmental Health Perspectives, the Journal of Toxicology and Environmental Health, Neuroendocrinology, Biology of Reproduction, and the Endocrine Journal. Dr. Eldridge has authored or coauthored 53 peer-reviewed journal articles, 16 books or monographs, 16 book chapters, and 69 abstracts for national or international meetings. He has given 24 invited lectures or seminars.

William R. Kelce, MS, PhD, ATS

Dr. Kelce received his PhD in Physiology/Toxicology from the University of Missouri–Columbia and completed a postdoctoral fellowship in Toxicology at The Johns Hopkins University. He is currently vice-president of preclinical development and clinical operations at POZEN Corporation and is also an Adjunct Associate Professor at the University of North Carolina School of Medicine, Department of Pediatrics and the Laboratories for Reproductive Biology. Dr. Kelce previously served as Director of Developmental and Reproductive Toxicology at Pfizer Global Research and Development. He received the Young Andrologist Award from the American Society of Andrology and the U.S. EPA Gold Medal Award (Scientist of the Year). As a postdoctoral fellow, he received First Place, Young

Independent Peer Review Panel Report

Investigator Award—Developmental and Reproductive Toxicology Section from the Society of Toxicology. He is a member of the International Life Sciences Institute Developmental and Reproductive Toxicology Section. He has served on numerous panels and committees focusing on endocrine disruptors, including the EPA Endocrine Disruptor Methods Validation Advisory Committee, the EPA Endocrine Disruptor Standardization and Validation Task Force, where he served as Chair of the *In Vitro* Screening Work Group, and as a member of the ICCVAM Endocrine Disruptor Peer Review Panel. He is a member of the Drug Information Association, Society of Toxicology (Editorial Board Member, Toxicology and Applied Pharmacology and member of the Board of Publications for Toxicological Sciences), the Endocrine Society, American Society of Andrology (Editorial Board Member, Journal of Andrology), American College of Toxicology, Teratology Society (Editorial Board Member, Birth Defects Research: Developmental and Reproductive Toxicology), Midwest Teratology Association (Steering Committee Member), and the Academy of Toxicological Scientists. Dr. Kelce has authored or coauthored 46 peer-reviewed journal articles and 13 invited reviews and book chapters. He has given 44 invited presentations.

Hyung Sik Kim, PhD

Dr. Kim received his PhD in Pharmacy from College of Pharmacy, Sung Kyun Kwan University. He is currently an Associate Professor in the Laboratory of Molecular Toxicology, College of Pharmacy, at Pusan National University, in the Republic of Korea, where he conducts research in the development of biomarkers and mechanisms of action of environmental xenobiotics at molecular and cellular levels. Dr. Kim is a member of Society of Toxicology, the American Association for Cancer Research, the Society of Study for Reproduction, the Korean Cancer Association, the Korean Society of Environmental Toxicology, the Korean Society of Applied Pharmacology, the Korean Society of Toxicology (Editorial Board), the Pharmaceutical Society of Korea, and the Korean Society of Food Hygiene Safety. Dr. Kim has authored or coauthored 95 peer-reviewed journal articles and 5 books.

Steven Levine, PhD

Dr. Levine received his PhD in Zoology from Miami University (Ohio) and was an NIEHS Postdoctoral Fellow in Toxicology at Pennsylvania State University. He is currently the Product Safety Manager – Ecotoxicology & Risk Assessment; Quantitative Bioassay Team Lead and Science Fellow in the Regulatory Sciences Department at Monsanto Company in St. Louis, MO. Dr. Levine's primary research interests include development of functional assays with insecticidal traits in biotechnology-derived products, aquatic and terrestrial toxicology, molecular mechanisms of steroidogenesis, and approaches for probabilistic risk assessment. He was a BIAC representative to the OECD Working Group of National Coordinators for Test Guidelines and Endocrine Disruption Testing Advisory Committee, chaired Crop Life America's (CLA) Endocrine Disruptor Group and served on the CLA and American Chemistry Council's Ecological Risk Assessment Working Group. He also served on the EPA Endocrine Disruption Methods Validation Advisory Committee. Dr. Levine was the President of the regional Society of Environmental Toxicology and Chemistry (SETAC) and is a member of SETAC and the Society of Toxicology. Dr. Levine has authored or coauthored 21 articles in peer-reviewed journals, 33 abstracts/presentations at national and international meetings, nine of which were invited. He has given eight invited lectures at universities and other nonmeeting venues.

Alberto Mantovani, DVM, MSc

Dr. Mantovani received a DVM from the University of Bologna and a MSc in Veterinary Public Health from the University of Edinburgh. He is currently the head of the Food and Veterinary Toxicology Unit within the Department of Food Safety and Veterinary Public Health of the Italian National Health Institute (ISS). The main research topics of the unit are endocrine disruptors and trace elements. Dr. Mantovani chaired the Endocrine Disruptors Technical Working Group within the SCALE Project to set the scientific bases of the "Environment and Health Action Plan" of the European Commission, chaired the Italian pilot project on endocrine disruptors, and participated in

Independent Peer Review Panel Report

the EU Working Group “Specialized Experts in the Field of Mutagenic, Carcinogenic and Teratogenic Substances.” He was the Italian Expert on toxicology at the Safety of Residues Working Party of the Committee for Veterinary Medicinal Products–European Agency for Evaluation of Medicinal Products; and currently participates on the OECD Working Group on Endocrine Disrupters Testing and Assessment. Dr. Mantovani has authored or coauthored 38 peer-reviewed journal articles since 2001.

Ellen Mihaich, PhD, DABT

Dr. Mihaich received her PhD from the Integrated Toxicology Program in the School of Forestry and Environmental Studies at Duke University. She is a Diplomate of the American Board of Toxicology and is currently the owner and principal scientist of Environmental and Regulatory Resources, LLC, specializing in environmental toxicology, risk assessment, and regulatory services. She is also an Adjunct Professor at Duke University. Dr. Mihaich was formerly Manager of Environmental Toxicology at Rhodia, Inc., and a Senior Environmental Toxicologist at Rhône-Poulenc. She served as a domestic and international consultant on environmental toxicology at both companies. Her responsibilities included management of environmental toxicology-based issues and environmental toxicology programs for chemical enterprises, monitoring and evaluating general and specialty laboratory and field environmental toxicology studies, plant site and chemical-use risk assessment, and interaction with international industry groups and regulatory agencies. Dr. Mihaich served as President of the Society of Environmental Toxicology and Chemistry (SETAC) and was on the Board of Directors for SETAC North America and the World Council. She is currently on the Scientific Advisory Board of the Strategic Environmental Research and Development Program (SERDP) of the Department of Defense. Dr. Mihaich previously served on the EPA Ecological Processes and Effects Committee, Deriving Aquatic Life Criteria for Emerging Contaminants, and ICCVAM peer panels on endocrine disruptors (2002) and the Frog Embryo Teratogenesis Assay (2000). Dr. Mihaich has authored or coauthored more than 25 publications in peer-reviewed journals or government publications and over 40 abstracts at national or international meetings.

Hiroshi Ono, MD, PhD

Dr. Ono received his MD from the Tohoku University School of Medicine and his PhD in Medical Research from Tohoku University. He is currently a scientific adviser to Hatano Research Institute, Food and Drug Safety Center. Dr. Ono has served as a visiting lecturer of General Medical Science at Azabu University School of Environmental Health Science, and toxicology at Yamanashi Medical College and at Tohoku University School of Medicine. He also served as a longtime delegate to the Meeting of National Coordinators of Test Guidelines Programme of Organisation for Economic Cooperation and Development. He is a member of the Japanese Society of Pharmacology, Japanese Society of Toxicology, Japanese Society for Alternatives to Animal Experiments, Japanese Society for Endocrine Disruptors Research, the Society of Toxicology, and EUROTOX-Association of European Toxicologists and Toxicological Societies. He received the Tanabe Prize for the notable publication of the year from the Japanese Society of Toxicology. Dr. Ono has 23 original articles and 49 reviews published in Japanese and 95 original articles published in English.

John G. Vandenberg, PhD

Dr. Vandenberg received a Masters degree in Zoology from Ohio University and a PhD in Zoology from Pennsylvania State University. He is currently a Professor Emeritus in the Department of Biology at North Carolina State University. In recent years, Dr. Vandenberg has focused his research on the effects of endocrine disruptors on development and later reproductive performance in rodents. In 2002 he received the Holladay Medal, NC State University’s highest faculty award. He is a founding board member and former chair of the North Carolina Association for Biomedical Research (NCABR) and served as a fellow and former president of the Animal Behavior Society. He was a member of the Institute for Laboratory Animal Research Council of the National Academies of Science and was on the committee that wrote the *Guide for the Care and Use of Animals* (both 1986

Independent Peer Review Panel Report

and 1996). Dr. Vandenberg has also served on several additional committees at the National Academies of Science, including as Chair of Committee on Animal Biotechnology: Science Based Concerns. He was named a National Associate of the National Academies of Science in 2003. He has been a member of several scientific review committees at the NSF, EPA and the NIH. Most recently, he served on the National Toxicology Program's Expert Panel on Human Risks of Exposure to Bisphenol A (BPA). Dr. Vandenberg is a present or past member of several scientific societies and editorial boards. He has edited 2 books, published over 100 chapters and review papers, and given many invited scientific presentations.

Sherry Ward, PhD, MBA

Dr. Ward received her PhD in Biochemistry from Michigan State University, an MBA from the University of Maryland University College (UMUC), and an executive MS in Technology Management from UMUC. She is currently a consultant with BioTred Solutions in New Market, Maryland. Dr. Ward has expertise in *in vitro* toxicology, scientific writing and project management, grant proposal review, and grant writing. She also has experience in biotechnology market research, commercialization, and strategy development. Dr. Ward is a contributing editor to AltTox.org. She is an adjunct faculty member at UMUC in Biotechnology & Project Management. She has animal welfare experience and has served since 2006 on the board of the International Foundation for Ethical Research. As a Staff Scientist at the Gillette Company, she developed, characterized, and drafted patent applications for the first human conjunctival epithelial cell lines and gained experience in bioassay development and validation. Dr. Ward has served on numerous scientific panels and committees and was a panel member and presenter at the ICCVAM symposia on mechanisms of ocular injury and recovery and minimizing pain and distress in ocular toxicity testing held at the NIH in May 2005. She has been actively involved with trade organizations and served on the European Cosmetic, Toiletry and Perfumery Association Eye Irritation Task Force and the ILSI-HESI Alternatives to Animals Task Force. Dr. Ward's experience in models of eye irritation and mechanisms of injury is reflected in 19 publications in peer-reviewed journals, 4 unpublished validation or prevalidation documents related to ICCVAM activities, 17 presentations, 28 abstracts, and a patent. She is a member of the Hopkins Medical and Surgical Association and the Washington Academy of Sciences.

Marc Weimer, PhD

Dr. Weimer received a PhD in Neurophysiology from the University of Hohenheim, Germany, and an MS in Methods and Models from FernUniversität in Hagen, Germany. He joined the Department of Biostatistics, German Cancer Research Center (DKFZ), in Heidelberg in 2006 as a biostatistician. Dr. Weimer's primary areas of work are toxicogenomics and development and validation of alternative methods to animal experiments. As a statistical consultant, he has been involved in national and international research projects aimed at reducing, refining, and replacing animal testing in toxicology. His main interests include dose-response modeling, agreement statistics, and toxicogenomics. Funded by ECVAM, he has been responsible for the statistical evaluation of the quality of *in vitro* assays developed within ReProTect, a project of the European Union advancing alternative methods in reproductive toxicity. Dr. Weimer has authored or coauthored 19 peer-reviewed journal articles.

James Wittliff, PhD, MD *hc*, FACB

Dr. Wittliff received his PhD in Molecular Biology from the University of Texas at Austin and completed postdoctoral studies in the Biology Division of Oak Ridge National Laboratory in Tennessee. He is currently a Professor of Biochemistry and Molecular Biology in the Graham Brown Cancer Center in the School of Medicine at the University of Louisville with additional appointments as Research Professor of Surgery and Director of the Institute for Molecular Diversity & Drug Design (IMD). He is also the Director of the Hormone Receptor Laboratory at the University and has held numerous professorships at universities in Europe, Asia, and Africa. Dr. Wittliff's research interests include mechanisms and applications of steroid and peptide hormone action in disease, biochemical

Independent Peer Review Panel Report

techniques and concepts for detection and treatment of cancer, and laser capture microdissection and its use in proteomics and genomics. He was among the first to prove that the appearance of estrogen receptors in breast cancer predicted a patient's response to hormone therapy. Dr. Wittliff has researched the biological properties and cellular roles of estrogen and progesterone receptors in human cancers and the actions of estrogen mimics acting as endocrine disruptor compounds (EDC). Formerly at NEN/DuPont, Dr. Wittliff developed the original FDA-approved kits for assessing receptors in biopsies, celebrated as a major contribution to laboratory medicine. He was the Principal Investigator or the Investigator for a number of funded studies on genomic approaches to disease, including a Genomic Approach for Assessing Clinical Outcome of Breast Cancer using Cells Isolated by Laser Capture Microdissection. Dr. Wittliff has served on numerous panels and committees, including the ICCVAM Endocrine Disruptor Peer Review Panel (2002). He is a member of several professional societies including the Endocrine Society, the American Association for Cancer Research, and the American Association for the Advancement of Science. Dr. Wittliff has authored or coauthored over 250 peer-reviewed publications and holds patents on methods and apparatus for measurement of the effect of test compounds on signal transduction at the receptor level, quantitative immunohistochemistry, breast cancer signatures, and gene expression profiles.

James Yager, Jr., PhD

Dr. Yager received his PhD from the University of Connecticut, Storrs Campus, in Cell and Developmental Biology and conducted postdoctoral studies at the McArdle Laboratory for Cancer Research at the University of Wisconsin. He is currently a Professor in Preventive Medicine and Toxicology in the Department of Environmental Health Sciences at the Johns Hopkins Bloomberg School of Public Health with a joint appointment in the Department of Oncology. He has administrative responsibility as the Senior Associate Dean for Academic Affairs. He was formerly a Professor of Anatomy and an Adjunct Professor in the Biochemistry Program at Dartmouth College. Dr. Yager has served as the Program Director and Principal Investigator for the Training Program in Environmental Health Sciences, Director of the Division of Toxicological Sciences, and Director of the Molecular Toxicology Program of the NIEHS-supported Center in Urban Environmental Health. His research interests include mechanisms of promotion of hepatocarcinogenesis by estrogenic xenobiotics, mechanisms of estrogen-induced oxidative DNA damage in liver and human breast epithelium, and the role of genetic susceptibility in human cancer through polymorphisms in biotransformation enzymes involved in estrogen oxidative metabolism. Dr. Yager serves on various committees and task forces for several professional societies, including the American Association of Cancer Research (AACR), the American Society for Investigative Pathology (ASIP, FASEB), and the Society of Toxicology. Dr. Yager serves or has served on various advisory boards and chartered review panels, including the EPA Endocrine Disruptor Methods Validation Subcommittee and the ICCVAM Scientific Review Panel to evaluate the validation status of in vitro estrogen and androgen receptor binding and transcriptional activation assays (2002). He is a peer reviewer for numerous journals, including Biochemical Pharmacology, Cancer Research, Chemical–Biological Interactions, Molecular Carcinogenesis, Science, and the Journal of the National Cancer Institute. Dr. Yager is on the editorial board for the Journal of Environmental Pathology, Toxicology and Oncology; In Vitro-Cell & Developmental Biology; Toxicology Sciences; and Chemical Research in Toxicology. He has authored or coauthored 84 peer-reviewed journal articles, 15 book chapters, 66 abstracts or presentations at national and international meetings; and he and has given over 50 invited presentations.

This page intentionally left blank

Appendix E

Federal Register Notices and Public Comments

E1	69 FR 21564 - <i>In Vitro</i> Endocrine Disruptor Test Methods: Request for Comments and Nominations (No public comments received).....	E-3
E2	71 FR 13597 - Notice of Availability of a Revised List of Recommended Reference Substances for Validation of <i>In Vitro</i> Estrogen and Androgen Receptor Binding and Transcriptional Activation Assays: Request for Comments and Submission of <i>In Vivo</i> and <i>In Vitro</i> Data (No public comments received).....	E-7
E3	74 FR 62317 - Evaluation of <i>In Vitro</i> Estrogen Receptor Transcriptional Activation and <i>In Vitro</i> Cell Proliferation Assays for Endocrine Disruptor Chemical Screening: Request for Nominations for an Independent Expert Peer Review Panel and Submission of Relevant <i>In Vitro</i> and <i>In Vivo</i> Data	E-11
E4	76 FR 4113 - Announcement of an Independent Scientific Peer Review Panel Meeting on an <i>In Vitro</i> Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening; Availability of Draft Background Review Document (BRD); Request for Comments	E-23
E5	76 FR 23323 - Meeting of the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM).....	E-37
E6	76 FR 28781 - Independent Scientific Peer Review Panel Report: Evaluation of the Validation Status of an <i>In Vitro</i> Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening: Notice of Availability and Request for Public Comments (No public comments received)	E-55
E7	Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) Comments: SACATM Meeting on June 16-18, 2011	E-59

This page intentionally left blank

Appendix E1

69 FR 21564

***In Vitro* Endocrine Disruptor Test Methods: Request for Comments and Nominations**

(No public comments received)

This page intentionally left blank

DEPARTMENT OF HEALTH AND HUMAN SERVICES**Public Health Service****National Toxicology Program (NTP), National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH), NTP Interagency Center for the Evaluation of Alternative Test Methods (NICEATM); In Vitro Endocrine Disruptor Test Methods: Request for Comments and Nominations**

SUMMARY: The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) have identified *in vitro* endocrine disruptor screening methods as a priority for validation. ICCVAM has published guidelines for development of *in vitro* endocrine-disruptor estrogen and androgen receptor binding and transcriptional activation assays. In these guidelines, ICCVAM recommends that priority be given to assays that (1) do not require the use of animal tissue as the receptor source, but rather use recombinant-derived proteins and (2) do not use radioactive materials. On behalf of the ICCVAM, NICEATM invites the nomination for validation studies of *in vitro* test methods that meet these recommendations and for which there are standardized test method protocols, pre-validation data, and proposed validation study designs. At this time, ICCVAM has received nominations for two *in vitro* endocrine-disruptor screening methods purported to meet these recommendations. Information on the nominated methods is posted on the ICCVAM/NICEATM Web site (<http://iccvam.niehs.nih.gov>) or available from

NICEATM (contact information provided below). ICCVAM will consider nominations and comments received in response to this notice and develop recommended priorities for proposed evaluation and validation studies of endocrine disruptor screening methods.

Request for Comments and Nomination of In Vitro Endocrine Disruptor Test Methods

Comments and nominations submitted in response to this notice should be sent by mail, fax, or e-mail to NICEATM (Dr. William S. Stokes, Director, NICEATM, NIEHS, 79 T. W. Alexander Drive, P.O. Box 12233, MD EC-17, Research Triangle Park, NC 27709, (phone) 919-541-2384, (fax) 919-541-0947, (e-mail) iccvam@niehs.nih.gov) by June 7, 2004, in order to ensure their consideration by the ICCVAM.

SUPPLEMENTARY INFORMATION: In May 2003, ICCVAM published a report entitled, "ICCVAM Evaluation of *In Vitro* Test Methods for Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays" (NIH Publication No. 03-4503; available: <http://iccvam.niehs.nih.gov/methods/endocrine.htm>). During its evaluation of *in vitro* endocrine disruptor screening assays, ICCVAM recommended that preference be given to development of assays that (1) do not require the use of animal tissue as the receptor source, but rather use recombinant-derived proteins and (2) do not use radioactive materials. ICCVAM also recommended minimum procedural standards that should be incorporated in standardized test method protocols and minimum lists of chemicals that should be used for validation studies. ICCVAM subsequently received nominations of two methods for validation studies. The first nomination is for a biosensor system that can assess estrogen receptor binding and transcriptional activation. The second nomination is for a stably transfected recombinant cell-based transcriptional method. The methods meet the ICCVAM's recommendations for studies that do not require the use of animals as a receptor source or use radioactive materials. Both methods detect receptor agonist and antagonist activity.

ICCVAM reviewed the two nominations described above and

unanimously approved the following draft recommendation: "Evaluation studies for *in vitro* receptor binding and transcriptional activation test methods that do not require the use of animals should receive a high priority for support. Prior to the initiation of such studies, the proposed validation studies should be evaluated for adherence to relevant recommendations in the report: "ICCVAM Evaluation of *In Vitro* Test Methods for Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays" (NIH Publication No. 03-4503) by the ICCVAM Endocrine Disruptor Working Group (EDWG) and NICEATM."

ICCVAM subsequently presented these nominations and its recommendation to the SACATM at its March 10-11, 2004 meeting. SACATM concurred with ICCVAM that endocrine disrupting screening assays should be a priority.

Background Information on ICCVAM and NICEATM

ICCVAM is an interagency committee composed of representatives from 15 Federal regulatory and research agencies that use, generate, or disseminate toxicological information. ICCVAM promotes the development, validation, regulatory acceptance, and national and international harmonization of toxicological test methods that more accurately assess the safety or hazards of chemicals and products and test methods that refine, reduce and replace animal use. The ICCVAM Authorization Act of 2000 (Pub. L. 106-545, available at <http://iccvam.niehs.nih.gov/about/PL106545.htm>) established ICCVAM as a permanent interagency committee of the NIEHS under the NICEATM. NICEATM administers the ICCVAM and provides scientific support for ICCVAM and ICCVAM-related activities. NICEATM and ICCVAM work collaboratively to evaluate new and improved test methods applicable to the needs of Federal agencies. Additional information about ICCVAM and NICEATM can be found at the following Web site: <http://iccvam.niehs.nih.gov>.

Dated: April 9, 2004.

Samuel H. Wilson,
Deputy Director, National Institute of Environmental Health Sciences.

[FR Doc. 04-8980 Filed 4-20-04; 8:45 am]

BILLING CODE 4140-01-P

This page intentionally left blank

Appendix E2

71 FR 13597

Notice of Availability of a Revised List of Recommended Reference Substances for Validation of *In Vitro* Estrogen and Androgen Receptor Binding and Transcriptional Activation Assays: Request for Comments and Submission of *In Vivo* and *In Vitro* Data

(No public comments received)

This page intentionally left blank

Toxicological Methods (NICEATM) announces the availability of an addendum to the report entitled, “Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) Evaluation of *In Vitro* Test Methods for Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays” [NIH Publication 03–4503]. The addendum describes the rationale for proposed revisions to the original list of recommended reference substances for validation of *in vitro* estrogen receptor (ER) and androgen receptor (AR) binding and transcriptional activation (TA) assays. The original list was made publicly available in June 2003 (**Federal Register**, Vol. 68, No. 106, pp. 33171–33172, June 3, 2003). NICEATM requests public comments on the substances proposed as substitutes for six of the 78 substances in the original list. Data are also requested from *in vitro* and *in vivo* studies evaluating the estrogenic and androgenic activity of the 78 substances in the revised list of reference substances.

DATES: Comments and data submissions should be received by May 1, 2006.

ADDRESSES: Correspondence should be sent by mail, fax, or e-mail to Dr. William S. Stokes, NICEATM Director, NIEHS, P. O. Box 12233, MD EC–17, Research Triangle Park, NC, 27709, (phone) 919–541–2384, (fax) 919–541–0947, (e-mail) niceatm@niehs.nih.gov.

SUPPLEMENTARY INFORMATION:

Background

In April 2000, the Environmental Protection Agency (EPA) asked ICCVAM to evaluate the validation status of *in vitro* ER and AR binding and TA assays that were proposed as possible components of the EPA Endocrine Disruptor Screening Program Tier 1 screening battery. ICCVAM agreed to evaluate these test methods based on their potential interagency applicability and public health significance. NICEATM, which administers and provides scientific support for ICCVAM, subsequently compiled available data and information on *in vitro* ER and AR binding and TA assays in four draft Background Review Documents (BRDs) (available at <http://iccvam.niehs.nih.gov/methods/endocrine.htm>).

In collaboration with the ICCVAM Endocrine Disruptor Working Group, NICEATM organized an independent scientific evaluation of the validation status of the four types of *in vitro* endocrine disruptor screening test

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Toxicology Program (NTP), NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM); Notice of Availability of a Revised List of Recommended Reference Substances for Validation of *In Vitro* Estrogen and Androgen Receptor Binding and Transcriptional Activation Assays: Request for Comments and Submission of *In Vivo* and *In Vitro* Data

AGENCY: National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH).

ACTION: Request for Comments and Submission of Data.

SUMMARY: The National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative

methods on May 20–21, 2002, in Research Triangle Park, NC (**Federal Register**, Vol. 66, No. 57, pp. 16278–16279, March 23, 2001 and **Federal Register**, Vol. 66, No. 67, pp. 16415–16416, April 5, 2002) (available at <http://iccvam.niehs.nih.gov/methods/endocrine.htm>).

The final BRDs and the ICCVAM Test Method Evaluation Report, which includes the expert panel report, public comments, and other relevant documents, were published in May 2003 and announced in a **Federal Register** notice (Vol. 68, No. 106, pp. 33171–33172, June 3, 2003) (available at <http://iccvam.niehs.nih.gov/methods/endocrine.htm>).

NICEATM recently reviewed the commercial availability and cost for the 78 substances recommended by ICCVAM for use in *in vitro* ER and AR binding and TA validation studies. A minimum of 44 substances are recommended for AR binding and TA assays, while a minimum of 53 substances are recommended for ER binding and TA assays. This review indicated that three substances [anastrozole, CGS 18320B, fadrozole] are not commercially available, one substance has restricted commercial availability [ICI 182,780] and six others [actinomycin D, hydroxyflutamide, 4-hydroxytamoxifen, methyltrienolone, 12-O-tetradecanoylphorbol-13-acetate, zearalenone] have costs that are considered excessive. ICCVAM proposes replacing the four substances that are not commercially available or have restricted availability with ones having similar ER and AR activity profiles [4-hydroxyandrostenedione, chrysin, dicofol, raloxifene HCl]. Suitable replacements (19-nortestosterone and resveratrol) were identified for methyltrienolone and zearalenone, respectively, for two of the expensive substances. NICEATM would also prefer to replace four of the highly priced substances [actinomycin D, hydroxyflutamide, 4-hydroxytamoxifen, 12-O-tetradecanoylphorbol-13-acetate], but has been unable to identify suitable replacements because of their unique activity profiles and/or chemical/physical properties. The revised list of 78 substances and a discussion about the proposed revisions are included and discussed in the “Addendum to the 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> see “Test Method Evaluations”) or by contacting NICEATM (see **ADDRESSES** above.) ICCVAM will finalize this list

after considering any public comments received and forward it to U.S. Federal agencies for their information and consideration.

Request for Comments and Request for Data

NICEATM requests public comments on the four substances (listed above) proposed as replacements for substances on the list that are not readily commercially available. NICEATM also requests public comments on the proposed replacements for the two expensive substances for which replacements have been identified, and suggestions for replacements for the four expensive substances that remain on the recommended list.

In order to update the reference substance database, NICEATM request data from completed *in vitro* studies using or evaluating ER and AR binding and/or TA assays, and information about ongoing or planned studies using or evaluating these test methods. NICEATM also requests the submission of data from animal studies that have evaluated the endocrine activity of chemicals using, for example, the uterotrophic, Hershberger, intact male, or male/female pubertal assays. NICEATM is especially interested in receiving additional data or information on any of the 78 substances included in the reference list. NICEATM previously requested data from completed studies using or evaluating ER and AR binding and/or TA assays, and information about ongoing or planned *in vitro* or *in vivo* studies using or evaluating these test methods (**Federal Register**, Vol. 66, No. 57, pp. 16278–16279, March 23, 2001). Submitted data will be used to update and supplement the existing NICEATM database; the current database can be accessed in the ICCVAM Test Method Evaluation Report [NIH Publication No. 03–4503] and the four final BRDs on ER and AR binding and TA assays [NIH Publication No. 03–4504, 03–4505, 03–4506, and 03–4507] (available at <http://iccvam.niehs.nih.gov/methods/endocrine.htm>).

When submitting chemical and protocol information/test data, please reference this **Federal Register** notice and provide appropriate contact information (name, affiliation, mailing address, phone, fax, e-mail, and sponsoring organization, as applicable).

NICEATM prefers data to be submitted as copies of pages from study notebooks and/or study reports, if available. Raw data and analyses available in electronic format may also be submitted. If data are published in the peer-reviewed literature, citations

should be provided. Each submission for a chemical should preferably include the following information, as appropriate:

- Common and trade name
- Chemical Abstracts Service Registry Number (CASRN)
- Chemical class
- Product class
- Commercial source
- *In vitro* test protocol used
- *In vitro* test results
- *In vivo* test protocol used
- *In vivo* test results
- The extent to which the study complied with national or international Good Laboratory Practice (GLP) guidelines
- Date and testing organization

Background Information on ICCVAM and NICEATM

ICCVAM is an interagency committee composed of representatives from 15 Federal regulatory and research agencies that use or generate toxicological information. ICCVAM conducts technical evaluations of new, revised, and alternative methods with regulatory applicability and promotes the scientific validation and regulatory acceptance of toxicological test methods that more accurately assess the safety and hazards of chemicals and products and that refine, reduce, or replace animal use. The ICCVAM Authorization Act of 2000 (Pub. L. 106–545) establishes ICCVAM as a permanent interagency committee of the NIEHS under the NICEATM. NICEATM administers the ICCVAM and provides scientific and operational support for ICCVAM-related activities. NICEATM and ICCVAM work collaboratively to evaluate new and improved test methods applicable to the needs of Federal agencies. Additional information about ICCVAM and NICEATM can be found at the following Web site: <http://www.iccvam.niehs.nih.gov>.

Dated: March 7, 2006.

Samuel H. Wilson,

Deputy Director, National Institute of Environmental Health Sciences and National Toxicology Program.

[FR Doc. E6–3763 Filed 3–15–06; 8:45 am]

BILLING CODE 4140–01–P

Appendix E3

74 FR 62317

Evaluation of *In Vitro* Estrogen Receptor Transcriptional Activation and *In Vitro* Cell Proliferation Assays for Endocrine Disruptor Chemical Screening: Request for Nominations for an Independent Expert Peer Review Panel and Submission of Relevant *In Vitro* and *In Vivo* Data

Public Comments Received in Response to 74 FR 62317

- Nancy Flourney (University of Missouri)E-16
- Keith Houck, PhD (National Center for Computational Toxicology, U.S. EPA)E-17
- Robert C. Renner (Water Research Foundation)E-18
- Keith Houck, PhD (National Center for Computational Toxicology, U.S. EPA)E-20
- M Pilar Vinardell (Universitat de Barcelona).....E-21
- Joanne Zurlo, PhD (The National Academies).....E-22

This page intentionally left blank

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Toxicology Program (NTP); NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM); Evaluation of In Vitro Estrogen Receptor Transcriptional Activation and In Vitro Cell Proliferation Assays for Endocrine Disruptor Chemical Screening: Request for Nominations for an Independent Expert Peer Review Panel and Submission of Relevant In Vitro and In Vivo Data

AGENCY: National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH).

ACTION: Request nominations for an independent expert panel and submission of relevant data.

SUMMARY: NICEATM, in collaboration with the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), is planning to convene an independent scientific peer review panel (hereafter, Panel) to assess the validation status of an *in vitro* stably-transfected estrogen

receptor (ER) transcriptional activation (TA) Assay (LUMI-CELL® ER assay) and an *in vitro* cell proliferation assay (CertiChem MCF-7 Cell Proliferation assay) for their usefulness and limitations in determining whether and to what extent chemicals can interact with estrogen receptors *in vitro*.

Validated assays that can detect the interaction of chemicals with specific hormone receptors including the ER are included in the U.S. Environmental Protection Agency (EPA) Endocrine Disruptor Screening Program (EDSP) (<http://www.epa.gov/endo/pubs/assayvalidation/status.htm>). The two assays that will undergo peer review are currently undergoing validation studies to determine their usefulness and limitations for the EDSP. Any other existing data from these two assays are requested to ensure that all available relevant data are considered by the Panel. Data from other existing *in vitro* and *in vivo* assays for the 78 reference substances used for the validation studies (available at http://iccvam.niehs.nih.gov/docs/endo_docs/EDAddendFinal.pdf) are requested for use in characterizing the expected *in vitro* and *in vivo* activity of these 78 reference substances. At this time NICEATM requests:

- Nominations of expert scientists for consideration as potential Panel members.

- Submission of existing data from the LUMI-CELL® ER and the CertiChem MCF-7 Cell Proliferation assays.

- Submission of data from *in vivo* or other *in vitro* assessments for the 78 reference substances recommended by ICCVAM for the validation of *in vitro* ER and AR binding and TA test methods (available at http://iccvam.niehs.nih.gov/docs/endo_docs/EDAddendFinal.pdf).

DATES: Submit nominations and data by January 11, 2010. Data submitted after this date will be considered in the evaluation, where feasible.

ADDRESSES: Submit nominations and data electronically by e-mail to niceatm@niehs.nih.gov, or via the NICEATM-ICCVAM Web site at http://iccvam.niehs.nih.gov/contact/FR_pubcomment.htm. Nominations and data may also be sent by mail or fax to Dr. William S. Stokes, Director, NICEATM, NIEHS, P.O. Box 12233, Mail Stop: K2-16, Research Triangle Park, NC 27709, (telephone) 919-541-2384, (fax) 919-541-0947, (e-mail). Courier address: NIEHS, NICEATM, 530 Davis Drive, Room 2034, Morrisville, NC 27560.

FOR FURTHER INFORMATION CONTACT: Dr. William S. Stokes, (telephone) 919-541-2384, (fax) 919-541-0947 and (e-mail) niceatm@niehs.nih.gov.

SUPPLEMENTARY INFORMATION:

Background

In April 2000, the EPA requested that ICCVAM evaluate the validation status of *in vitro* ER and AR binding and TA assays for potential use in the proposed EPA EDSP. ICCVAM and NICEATM compiled available relevant data for 137 existing assays and compiled data were submitted to an independent expert panel for review. This panel concluded that there were no adequately validated *in vitro* ER- or AR-based test methods (the panel's report is available on the NICEATM-ICCVAM Web site at http://iccvam.niehs.nih.gov/methods/endocrine/end_EPrt.htm). Based on these conclusions and recommendations, along with comments from the public, ICCVAM recommended minimum procedural standards and a list of 78 reference substances that should be used to standardize and validate *in vitro* ER and AR binding and TA test method protocols. These recommendations were made publicly available in the report: *ICCVAM Evaluation of the In Vitro Methods for Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays* (available at: http://iccvam.niehs.nih.gov/methods/endocrine/end_TMER.htm). The list of 78 reference substances was subsequently modified because of cost and availability considerations and published in a separate Addendum (available at: http://iccvam.niehs.nih.gov/docs/endo_docs/EDAddendFinal.pdf).

Two *in vitro* assays to detect ER agonists and antagonists were subsequently nominated to ICCVAM for validation studies in response to an ICCVAM request (69 FR 21564): The LUMI-CELL® ER assay developed by Xenobiotic Detection Systems, Inc. (XDS) and the CertiChem MCF-7 Cell Proliferation assay developed by CertiChem, Inc. (CertiChem). Based on preliminary results provided for these test methods and comments from the public and the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM; 69 FR 21564 and 71 FR 60748, respectively), ICCVAM and its Endocrine Disruptor Working Group recommended a high priority for validation studies for the LUMI-CELL® ER and CertiChem MCF-7 Cell Proliferation assays.

An international interlaboratory validation study of the LUMI-CELL® ER assay is currently nearing completion. The study includes three laboratories sponsored by NICEATM, the European Centre for the Validation of Alternative Methods, and the Japanese Center for the Validation of Alternative Methods. An intralaboratory validation study of the MCF-7 Cell Proliferation assay has been completed by CertiChem in conjunction with NICEATM, and an interlaboratory study is planned.

NICEATM will prepare draft background review documents (BRDs) following completion of the validation studies that will provide comprehensive summaries of available data, analyses of test method accuracy and reliability, and related information characterizing the current validation status of each of the assays. The draft BRDs will form the basis for draft ICCVAM test method recommendations on usefulness and limitations, standardized test method protocols, future studies, and performance standards that will subsequently be provided to the Panel and made available to the public. The Panel will meet in public session to review the validation status of the LUMI-CELL® ER, MCF-7 Cell Proliferation assays, and any of the other assays for which there are adequate data available. The Panel will comment on the extent to which the BRD supports draft ICCVAM test method recommendations. The Panel may also consider the results for other assays with incomplete validation databases to determine their current validation status and to identify data gaps that need to be addressed in order to further characterize their usefulness and limitations for the EDSP. Meeting information, including dates, locations, and public availability of the BRDs will be announced in future **Federal Register** notices and will also be posted on the ICCVAM/NICEATM Web site (http://iccvam.niehs.nih.gov/methods/endocrine/end_eval.htm).

Request for Nominations of Scientific Experts

NICEATM requests nominations of scientists with relevant knowledge and experience to serve on the Panel. Areas of relevant expertise include, but are not limited to, biostatistics, cellular biology, endocrinology, molecular genetics, regulatory toxicology, reproductive toxicology, and test method validation. Each nomination should include the nominee's name, affiliation, contact information (*i.e.*, mailing address, email address, telephone, and fax numbers), *curriculum vitae*, and a brief summary

of relevant experience and qualifications.

Request for Data

NICEATM invites the submission of relevant *in vitro* and *in vivo* data and information for reference substances on the list of 78 substances recommended by ICCVAM for standardizing and validating *in vitro* ER and AR binding and TA test methods (available at http://iccvam.niehs.nih.gov/docs/endo_docs/EDAddendFinal.pdf) or other substances for which data exists from the two *in vitro* test methods described in this notice. Relevant *in vivo* data may include, but are not limited to: Multi-generational reproductive and developmental toxicity studies, uterotrophic bioassays, and short term assays assessing changes in phenotypic parameters such as anogenital distance, time of vaginal opening, nipple retention, and preputial separation delays in males.

Although data can be accepted at any time, data received by January 11, 2010 will ensure consideration during the ICCVAM evaluation process. Relevant data received after this date will be considered during the ICCVAM evaluation process where feasible. All information submitted in response to this notice will be made publicly available and may be incorporated into future NICEATM and ICCVAM reports and publications as appropriate.

When submitting data, please reference this **Federal Register** notice and provide appropriate contact information (name, affiliation, mailing address, phone, fax, e-mail, and sponsoring organization, as applicable).

NICEATM prefers that data be submitted as copies of pages from study notebooks and/or study reports, if available. Laboratory data and analyses available in electronic format may also be submitted. Each submission for a substance should preferably include the following information, as appropriate:

- Common and trade name
- Chemical Abstracts Service Registry Number (CASRN)
- Commercial source
- *In vivo* or *in vitro* test protocol used
- Individual animal or *in vitro* responses at each observation time (*i.e.*, laboratory data)
- The extent to which the data were collected in accordance with national/international Good Laboratory Practice guidelines
- Date and testing organization
- Physical and chemical properties (*e.g.*, molecular weight, pH, water solubility, *etc.*)

**Background Information on ICCVAM,
NICEATM, and SACATM**

ICCVAM is an interagency committee composed of representatives from 15 Federal regulatory and research agencies that use or generate toxicological information. ICCVAM conducts technical evaluations of new, revised, and alternative methods with regulatory applicability and promotes the scientific validation and regulatory acceptance of toxicological test methods that more accurately assess the safety and hazards of chemicals and products and that refine, reduce, and replace animal use. The ICCVAM Authorization Act of 2000 (42 U.S.C. 2851–3, available at http://iccvam.niehs.nih.gov/docs/about_docs/PL106545.pdf) established ICCVAM as a permanent interagency committee of the NIEHS under NICEATM. NICEATM administers ICCVAM and provides scientific and operational support for ICCVAM-related activities. NICEATM and ICCVAM work collaboratively to evaluate new and improved test methods applicable to the needs of Federal agencies. Additional information about ICCVAM and NICEATM is available on the NICEATM–ICCVAM Web site: <http://iccvam.niehs.nih.gov>.

SACATM was established January 9, 2002 and is composed of scientists from the public and private sectors (67 FR 11358). SACATM provides advice to the Director of the NIEHS, ICCVAM, and NICEATM regarding the statutorily mandated duties of ICCVAM and activities of NICEATM. Additional information about SACATM, including the charter, roster, and records of past meetings, can be found at <http://ntp.niehs.nih.gov/go/167>.

Dated: November 16, 2009.

John R. Bucher,

Associate Director, National Toxicology Program.

[FR Doc. E9–28278 Filed 11–25–09; 8:45 am]

BILLING CODE 4140–01–P

From: Flournoy, Nancy
Date: Mon, 30 Nov 2009 14:48:21 -0500
To: "iccvam-all@list.niehs.nih.gov"
Subject: Vol 51, Issue 1 expert recommendations

Endocrine receptor experts

Shanna Swan
Fred vomSaal
Wade Welshons
Nancy Flournoy

Nancy Flournoy
Professor and Chair
Department of Statistics
University of Missouri
Columbia, MO 65211

Subject: scientific review panel nomination
Date: Wednesday, December 2, 2009 7:34 PM
From: Houck.Keith
To: NIEHS NICEATM

Hello,

Please see attached nomination and CV.

Thank you,

~~~~~  
Keith Houck, Ph.D.  
National Center for Computational Toxicology  
Office of Research & Development  
U.S. Environmental Protection Agency  
Research Triangle Park, NC 27711  
~~~~~

Subject: FW: Nominations?

Date: Thursday, December 3, 2009 4:11 PM

From: Rob Renner

To: Frank Deal

Cc: Catherine Sprankle

Dear Mr. Frank Deal,

On behalf of the Water Research Foundation, I would like to nominate Dr. Shane Snyder as a member of the expert panel on estrogen receptor mediated bioassays for endocrine disruptor screening. Dr. Snyder served on the EDMVS and EDMVAC committees in the past and was also nominated for these panels by the Water Research Foundation. Our foundation has funded many projects related to endocrine disruptors in water, including projects using in vitro bioassays for estrogenic compound screening. Dr. Snyder has been the principal investigator on many of our projects related to EDCs and is a member of our expert panel for a strategic initiative on EDCs and pharmaceuticals. I believe that Dr. Snyder is ideally suited for this important panel and we are certain he has the scientific merit and perspective on water industry needs that would be invaluable on this expert panel.

Name: Shane Snyder

Mailing Address: Harvard School of Public Health

Department of Environmental Health

Exposure, Epidemiology and Risk Program

[REDACTED]

Boston, MA 02215

Email Address: [REDACTED]

Phone: [REDACTED]

Thank you for your consideration of Dr Snyder for your expert panel. Dr. Snyder's CV and biosketch are provided as attached files, but please feel free to contact me if you require additional information.

Warmest regards,

Rob

Robert C. Renner
Executive Director
Water Research Foundation (formerly AwwaRF)
303.347.6150
Fax 303.730.0851
rrenner@WaterResearchFoundation.org

Water Research Foundation

This email has been scanned by the MessageLabs Email Security System.
For more information please visit <http://www.messagelabs.com/email>

NICEATM Scientific Review Panel Request for Nominations (74 FR 62317)

02 December 2009

Name: Keith A. Houck
Affiliation: U.S. EPA
Address: [REDACTED]
Research Triangle Park, NC 27711
email: [REDACTED]
Telephone: [REDACTED]
Fax: [REDACTED]

I would like to nominate myself to the NICEATM independent scientific review panel to assess the validation status of two in vitro assays for endocrine disruptor screening (74 FR 62317). I believe I have a very appropriate background and expertise to serve on this panel. I have more than 15 years of relevant experience in industry and government positions. This experience includes developing, validating, operations, reviewing and interpreting in vitro assays. For several of these years, I directed a group of 10 scientists with a focus on developing assays for nuclear receptors, including the endocrine receptors, and screening large libraries of compounds to identify novel ligands. I have highlighted significant achievements in this area below. Please also see my attached C.V.

- Developed/validated over 100 in vitro assays for use in high-throughput screening or follow-up screening in support of drug discovery efforts at Eli Lilly & Co.
- Developed/validated several medium- to high-throughput screens for toxicity screening in support of lead characterization efforts (Ames II, PXR induction, CYP1A1 induction, cytotoxicity).
- Co-founded and chaired Assay Protocol Approval Committee that reviewed validation status for all high-throughput screens and secondary assays run at Eli Lilly & Co. for many years. Also was leader in defining guidelines used for the validation.
- Led Transcriptional Regulation Group at Eli Lilly & Co. that developed, validated and screened large chemical libraries for over 30 human nuclear receptor targets using both biochemical and cellular assays.
- Led screening effort in Tox21 collaboration (EPA/NTP/NCGC) against a panel of 12 nuclear receptors including key endocrine receptors.
- Participated in two recent workshops focused on endocrine disruptor screening.
- Served on several NIH Roadmap on Assay Development for High Throughput Molecular Screening Grant Review Panels
- Lead scientist in EPA's ToxCast program that makes extensive use of a wide variety of in vitro screening assays to support a predictive toxicology effort.

Subject: Evaluation of Endocrine Disruptor Screening Assays: Expert Panel
Nominations

Date: Wednesday, December 2, 2009 9:56 AM

From: Pilar Vinardell

To: NIEHS NICEATM

Dear Dr William S. Stokes

I am pleased to send you my /curriculum vitae/ and a brief summary of relevant experience, in answer to the request for Nominations of Scientific Experts to serve as member of an independent peer review panel on *Evaluation of In Vitro Estrogen Receptor Transcriptional Activation and In Vitro Cell Proliferation Assays for Endocrine Disruptor Chemical Screening* (Federal Register Notice, Vol. 74, No. 227, November 27, 2009).

Best regards,

--

Dra M Pilar Vinardell
Directora Dep. Fisiologia
Facultat de Farmàcia
Universitat de Barcelona

Subject: nomination for committee

Date: Tuesday, December 1, 2009 3:35 PM

From: Zurlo, Joanne

To: NIEHS NICEATM

Conversation: nomination for committee

I would like to nominate Dr. James D. Yager from The Johns Hopkins Bloomberg School of Public Health for the expert panel on Evaluation of Endocrine Disruptor Screening Assays. In 2002, he served on the Scientific Review Panel to evaluate the validation Status of in vitro estrogen and androgen receptor binding and transcriptional assays.

Joanne Zurlo

Joanne Zurlo PhD

Director, Institute for Laboratory Animal Research

The National Academies

500 Fifth Street, NW

Washington, DC 20001

Internet: www.nationalacademies.org/ilar

Appendix E4

76 FR 4113

Announcement of an Independent Scientific Peer Review Panel Meeting on an *In Vitro* Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening; Availability of Draft Background Review Document (BRD); Request for Comments

Public Comments Received in Response to 76 FR 4113

- Catherine Willett, PhD (People for the Ethical Treatment of Animals) and Patricia L. Bishop, MS (People for the Ethical Treatment of Animals).....E-28

This page intentionally left blank

SUMMARY: NICEATM, in collaboration with the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), announces a public meeting of an independent scientific peer review panel (Panel) to evaluate the validation status of LUMI-CELL® ER (BG1Luc ER TA), an *in vitro* transcriptional activation (TA) assay used to identify chemicals that can interact with human estrogen receptors (ERs). Validated assays that can detect the interaction of chemicals with specific hormone receptors, including ERs, have been accepted and included in the U.S. Environmental Protection Agency (EPA) Endocrine Disruptor Screening Program (EDSP) (<http://www.epa.gov/endo/pubs/assayvalidation/status.htm>). Consequently, the BG1Luc ER TA may be applicable for addressing the ER TA component of the EPA EDSP Tier 1 screening battery.

At this meeting, the Panel will review the draft BRD for the BG1Luc ER TA and evaluate the extent to which established validation and acceptance criteria have been appropriately addressed. The Panel also will be asked to comment on the extent to which the information included in the BRD supports ICCVAM's draft test method recommendations.

NICEATM invites public comments on the draft BRD and draft ICCVAM test method recommendations. These documents are available on the NICEATM-ICCVAM Web site at: <http://iccvam.niehs.nih.gov/methods/endocrine/PeerPanel11.htm>.

DATES: The meeting will be held on March 29–30, 2011, from 8:30 a.m. to 5 p.m. each day. In order to facilitate planning for this meeting, persons wishing to attend are asked to register by March 15, 2011, via the NICEATM-ICCVAM Web site (<http://iccvam.niehs.nih.gov/contact/reg-form-EDpanel.htm>). Comments should be sent by March 10, 2011.

ADDRESSES: The meeting will be held at the National Institutes of Health (NIH), William H. Natcher Conference Center, 45 Center Drive, Bethesda, MD 20892. Persons needing special assistance in order to attend, such as sign language interpretation or other reasonable accommodation, should contact 301–402–8180 (voice) or 301–435–1908 TTY (text telephone) at least seven business days before the event.

FOR FURTHER INFORMATION CONTACT: Dr. Warren Casey, Deputy Director, NICEATM, NIEHS, P.O. Box 12233, Mail Stop: K2–16, Research Triangle Park, NC 27709, (telephone) 919–541–2384, (fax) 919–541–0947, (e-mail)

DEPARTMENT OF HEALTH AND HUMAN SERVICES

Independent Scientific Peer Review Panel Meeting on an *In Vitro* Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening; National Toxicology Program (NTP); NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM); Announcement of an Independent Scientific Peer Review Panel Meeting on an *In Vitro* Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening; Availability of Draft Background Review Document (BRD); Request for Comments

AGENCY: National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH), HHS.

ACTION: Meeting announcement and request for comments.

niceatm@niehs.nih.gov. Courier address: NICEATM, NIEHS, 530 Davis Drive, Room 2035, Morrisville, NC 27560.

SUPPLEMENTARY INFORMATION:

Background

In January 2004, Xenobiotics Detection Systems, Inc. (XDS, Durham, NC) nominated their LUMI-CELL® TA (BG1Luc ER TA) Test Method for an interlaboratory validation study to be coordinated by NICEATM. This method uses BG-1 cells, a human ovarian carcinoma cell line that was stably transfected with an estrogen-responsive luciferase reporter gene, to measure whether and to what extent a substance induces or inhibits TA activity via ER mediated pathways. Included in the nomination package were test results from XDS for 56 of the 78 ICCVAM Reference Substances for agonist activity and 16 of the 78 ICCVAM Reference Substances for antagonist activity. These studies were funded primarily by a Small Business Innovation Research (SBIR) grant (SBIR43ES010533-01) from the NIEHS.

In accordance with the ICCVAM nomination process, NICEATM conducted a pre-screen evaluation of the nomination package to determine the extent to which it addressed the ICCVAM prioritization criteria and adherence to the ICCVAM recommendations for the standardization and validation of *in vitro* endocrine disruptor test methods. Based on this evaluation, ICCVAM recommended a high priority for validation studies for the BG1Luc ER TA test method. The NIEHS subsequently agreed to support the validation study in light of its participation as one of the three NTP agencies, whose mission includes the development and validation of improved testing methods.

The international interlaboratory validation study of the BG1Luc ER TA test method has been completed. The study included three laboratories sponsored by NICEATM, the European Centre for the Validation of Alternative Methods, and the Japanese Center for the Validation of Alternative Methods.

NICEATM and ICCVAM have prepared a draft BRD that provides comprehensive summaries of data, analyses of test method accuracy and reliability, and related information characterizing the current validation status of the test method. The draft BRD forms the basis for ICCVAM test method recommendations on usefulness and limitations, standardized test method protocols, future studies, and performance standards.

Peer Review Panel Meeting

This meeting will take place March 29–30, 2011, at the National Institutes of Health (NIH) William H. Natcher Conference Center, 45 Center Drive, Bethesda, MD 20892. It will begin at 8:30 a.m. and is scheduled to conclude each day at approximately 5 p.m. The meeting is open to the public at no charge, with attendance limited only by the space available. The Panel will consider the draft ICCVAM BRD, recommendations, and performance standards for the test method and evaluate the extent to which the draft ICCVAM test method recommendations are supported by the information provided in the draft BRD.

Additional information about the meeting, including a roster of the Panel members and the draft agenda, will be posted on the NICEATM–ICCVAM Web site at <http://iccvam.niehs.nih.gov/methods/endocrine/PeerPanel11.htm> two weeks before the meeting. This information will also be available after that date by contacting NICEATM (see **FOR FURTHER INFORMATION CONTACT**).

Attendance and Registration

In order to facilitate planning for this meeting, persons wishing to attend are asked to register by March 15, 2011, via the NICEATM–ICCVAM Web site at <http://iccvam.niehs.nih.gov/contact/register-EDpanel.htm>.

Availability of the Documents

The draft BRD and draft ICCVAM test method recommendations will be posted no later than February 1, 2011 on the NICEATM–ICCVAM Web site (<http://iccvam.niehs.nih.gov/methods/endocrine/PeerPanel11.htm>) or may be obtained by contacting NICEATM (see **FOR FURTHER INFORMATION CONTACT**).

Request for Public Comments

NICEATM invites the submission of written comments on the draft BRD, draft ICCVAM test method recommendations, and draft performance standards by March 10, 2011. NICEATM prefers that comments be submitted electronically via the NICEATM–ICCVAM Web site (http://iccvam.niehs.nih.gov/contact/FR_pubcomment.htm) or via e-mail to niceatm@niehs.nih.gov. Written comments may also be sent by mail, fax, or e-mail to Dr. Casey (see **FOR FURTHER INFORMATION CONTACT**). When submitting written comments, please refer to this **Federal Register** notice and include appropriate contact information (name, affiliation, mailing address, phone, fax, e-mail, and sponsoring organization, if applicable). NICEATM will post all comments on the NICEATM–ICCVAM

Web site (<http://iccvam.niehs.nih.gov>) identified by the individual's name and affiliation or sponsoring organization (if applicable). NICEATM will provide these comments to the Panel and ICCVAM agency representatives and make them available to the public at the meeting.

Opportunity will be provided for members of the public to present oral comments at designated times during the peer review. Up to seven minutes will be allotted per speaker. If you wish to present oral statements at the meeting (one speaker per organization), contact NICEATM (see **FOR FURTHER INFORMATION CONTACT**) by March 2, 2011. Please provide a written copy of your comments with contact information (name, affiliation, mailing address, phone, fax, e-mail, and sponsoring organization, if applicable) when registering to make oral comments. If it is not possible to provide a copy of your statement in advance, please bring 40 copies to the meeting for distribution to the Panel and to supplement the record. Written statements can supplement and expand the oral presentation. Please provide NICEATM with copies of any supplementary written statement using the guidelines outlined above.

Summary minutes and the Panel's final report will be available following the meeting on the NICEATM–ICCVAM Web site (<http://iccvam.niehs.nih.gov>). ICCVAM will consider the Panel's conclusions and recommendations and any public comments received in finalizing their test method recommendations for the test method.

Background Information on ICCVAM and NICEATM

ICCVAM is an interagency committee composed of representatives from 15 Federal regulatory and research agencies that use or generate toxicological and safety testing information. ICCVAM conducts technical evaluations of new, revised, and alternative methods with regulatory applicability and promotes the scientific validation and regulatory acceptance of toxicological and safety-testing methods that more accurately assess the safety and hazards of chemicals and products and that reduce, refine (decrease or eliminate pain and distress), and replace animal use. The ICCVAM Authorization Act of 2000 (42 U.S.C. 285l-3, available at http://iccvam.niehs.nih.gov/docs/about_docs/PL106545.pdf) established ICCVAM as a permanent interagency committee of the NIEHS under NICEATM. NICEATM administers ICCVAM and provides scientific and operational support for ICCVAM-related activities. NICEATM and ICCVAM work collaboratively to

evaluate new and improved test methods applicable to the needs of Federal agencies. Additional information about ICCVAM and NICEATM is available on the NICEATM–ICCVAM Web site at <http://iccvam.niehs.nih.gov>.

Dated: January 13, 2011.

John R. Bucher,

Associate Director, National Toxicology Program.

[FR Doc. 2011–1329 Filed 1–21–11; 8:45 am]

BILLING CODE 4140–01–P

March 10, 2011

Dr Warren Casey
Deputy Director
NICEATM
National Institute of Environmental Health Sciences
PO Box 12233, K2-16
Research Triangle Park, NC 27709

And via e-mail to: niceatm@niehs.nih.gov

Re: **76 FR 4113; January 24, 2011; Independent Scientific Peer Review Panel Meeting on an *In Vitro* Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening; National Toxicology Program (NTP); NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM); Request for Comments.**



HEADQUARTERS
501 FRONT ST.
NORFOLK, VA 23510
757-622-PETA
757-622-0457 (FAX)

Dear Dr Casey:

People for the Ethical Treatment of Animals (PETA) is the world's largest animal rights organization, with over 2 million members and supporters. We appreciate the opportunity to comment regarding the draft Background Review Document (BRD) and draft ICCVAM test method recommendations (TMR).

General Comments

We are pleased that ICCVAM is recommending the BG1Luc ER TA test method as a screening assay to identify substances with estrogen agonist and antagonist activity. We support the finding that this assay can be applied to a wide range of substances and can be routinely used to prioritize substances for further testing. We also appreciate the thoroughness of the BRD and the development of Performance Standards for the BG1Luc ER TA assay. We support the conclusion that the BG1Luc ER TA assay is equivalent to the OPPTS 890.1300/CERI STTA method; however, since the CERI STTA validation report has been published,¹ it would be useful to include a quantitative comparison and to compare chemicals used to assess accuracy as compared to ER binding and uterotrophic assays. While the BRD is thorough, it contains a large amount of repetitive information, which, if removed, could significantly shorten the document.

Additionally, we support any recommendations that could lead to reduction, refinement or replacement of animal testing. These include the recommendations that: 1) the BG1Luc ER TA

¹ OECD. 2006. Draft Report of Pre-validation and Inter-laboratory Validation For Stably Transfected Transcriptional Activation (TA) Assay to Detect Estrogenic Activity - The Human Estrogen Receptor Alpha Mediated Reporter Gene Assay Using hER-HeLa-9903 Cell Line. Available at: <http://www.oecd.org/dataoecd/7/27/37504278.pdf> (accessed 6 March 2011).

test be considered for quantitative, rather than just qualitative, assessment of estrogen agonist and antagonist activity,² 2) the BG1Luc ER TA test be incorporated as part of a weight of evidence approach to reduce or eliminate the need for testing in animal models such as the female rat pubertal, rat uterotrophic and fish short-term reproductive assays.³

While we are pleased with the results of the study, we have major concerns with the length of time it took to validate this test, which, when nominated in January 2004, already had a considerable amount of relevant data associated with it. We are disappointed that this review took seven years and was not completed in time for inclusion in the initial phase(s) of the Environment Protection Agency's Endocrine Disruptor Screening Program. The prolonged review has also affected development of a Performance Based Test Guideline (PBTG) for estrogen receptor transcriptional activation assays.

The delay of this validation process was likely exacerbated by the inclusion of 78 reference chemicals, many of which are not well characterized, in the validation process. New methods should be validated with reference chemicals whose activities are extremely well characterized. Following validation, other chemicals with suspected activity or limited data can then be characterized – it is inappropriate to mix the two. In addition, the ICCVAM list of 78 chemicals were described as chemicals “that should be used to standardize and validate *in vitro* ER and AR binding and TA test methods.”⁴ Since the BG1Luc test is concerned with ER TA agonist and antagonist effects only, we have to question why all 78 would be included.

Specific Comments

1.0 Draft ICCVAM Recommendations: the BG1Luc ER TA test method

Lines 36 – 37: The characterization of L-thyroxine as a “false negative” is misleading considering that this chemical is not well characterized (see Table 1). In fact, it later states in Section 5.2.1 that this substance was classified by ICCVAM as positive based on two reports of positive agonist activity and one report of no agonist activity, hardly a definitive set of evidence.

Lines 76 – 78: There seems to be something missing from this sentence.

Lines 99 – 110. Although assessment of both agonist and antagonist activity is an advantage of the BG1Luc ER TA test method over the CER1 STTA method, transcriptional activation assays support but do not definitively prove receptor mediation. Binding studies are performed to confirm a receptor binding mechanism of action, and therefore cannot be replaced by a transcriptional activation assay. A more appropriate recommendation would be to validate an ER binding assay that uses a human recombinant ER. The CER1 STTA method is currently being validated for antagonist activity.⁵

² NICEATM Draft ED BRD: BG1Luc ER TA Test Method – Section 5.0, p. 5-11.

³ NICEATM Draft ED BRD: BG1Luc ER TA Test Method – Section 9.0, p. 9-2.

⁴ NICEATM Draft ED BRD: BG1Luc ER TA Test Method – Section 1.0 Introduction, p. 1-3.

⁵ Workplan for the Test Guidelines Programme. 2010. Organization for Economic Coordination and Development (OECD) (www.oecd.org/dataoecd/54/29/46034089.pdf) (accessed 6 March 2010).

2.5.1 Solubility Testing

Solubility in 100% DMSO is not reflective of the solubility upon dilution in the culture medium – many compounds can be completely soluble in DMSO yet form precipitate when diluted in aqueous solution – this adjustment could lead to serious miscalculations of solubility in many cases.

3.0 Substances Used to Evaluate Test Method Accuracy

Table 3 – 2: The 78 reference substances, chosen based on “a preponderance of evidence found in a review of the scientific literature” includes several substances with very little information. The substances listed in Table 3-2 should be graded with respect to the confidence of a positive or negative determination based on the quantity and quality of available data as we have illustrated below in Table 1. In Table 1, substances with substantial, definitive data are not shaded, substances with a moderate amount of information are lightly shaded, and substances with little information are darkly shaded. Substances with low confidence (e.g. those darkly shaded) should be deleted from the reference list, and should not have been used in validation studies.

ICCVAM, in considering which substances to use to assess the accuracy of the agonist and antagonist activity, selected “only those substances that could be definitively classified as POS or NEG.” Many (but not all) of the substances with little supporting data were not tested by ECVAM or Hiyoshi as indicated in **Table 2 for agonists and Table 3 for antagonists.**

Table 1. Copy of Table 3–2: Substances graded by amount of substantiating information.

ICCVAM Reference Substance	CASRN	ER TA Agonist Activity	ER TA Antagonist Activity	ER Binding Activity	CER ER TA Activity	Uterotrophic Activity
12 – O – tetradecanoylphorbol-13-acetate	16561-29-8	PN (nt)	PN (nt)	PN (nt)	nt	nt
17-β estradiol	50-28-2	POS (226/226)	PP (1/1)	POS (160/160)	POS	nt
17-α estradiol	57-91-0	POS (10/10)	PP (1/1)	POS (15/15)	POS	POS (nt/+)
17-α ethinyl estradiol	57-63-6	POS (21/21)	NEG (0/9)	POS (32/32)	POS	POS (+/+)
17β-trenbolone	10161-33-8	PP (1/1)	PN (nt)	PN (nt)	POS	nt
19-nortestosterone	434-22-0	POS (3/3)	PP (1/1)	PP (1/7)	nt	nt
2-sec-butylphenol	89-72-5	PN (0/1)	PN (nt)	POS (2/2)	NEG	nt
2,4,5-trichlorophenoxyacetic acid	93-76-5	PP (1/3)	PP (1/2)	PP (1/3)	nt	nt
4-androstenedione	63-05-8	PP (1/1)	PN (0/1)	PP (1/5)	NEG	nt
4-cumylphenol	599-64-4	POS (4/4)	PN (nt)	POS (3/3)	POS	nt
4-hydroxy androstenedione	566-48-3	PP (1/2)	PN (nt)	PP (nt)	nt	nt
4-hydroxytamoxifen	68047-06-3	PP (17/56)	POS (27/27)	POS (36/36)	nt	nt
4-tert-octylphenol	140-66-9	POS (20/23)	PN (nt)	POS (20/20)	POS	POS (nt/+)
5α-dihydrotestosterone	521-18-6	POS	NEG (0/3)	POS (17/18)	nt	POS (nt/+)

ICCVAM Reference Substance	CASRN	ER TA Agonist Activity (15/17)	ER TA Antagonist Activity	ER Binding Activity	CER ER TA Activity	Uterotrophic Activity
Actinomycin D	50-76-0	PN (nt)	PN (nt)	PN (nt)	nt	nt
Ammonium perchlorate	7790-98-9	PN (nt)	PN (nt)	PN (nt)	nt	nt
Apigenin	520-36-5	POS (25/25)	NEG (0/11)	POS	POS	nt
Apomorphine	58-00-4	PN (nt)	PN (nt)	PN (nt)	nt	nt
Atrazine	1912-24-9	NEG (0/29)	PN (0/1)	PP (2/19)	NEG	nt
Bicalutamide	90357-06-5	NEG (0/5)	PN (nt)	PN (nt)	nt	nt
Bisphenol A	80-05-7	POS (64/64)	NEG (0/12)	POS (46/47)	POS	POS (+/+)
Bisphenol B	77-40-7	POS (5/5)	PN (0/1)	POS (2/2)	POS	POS (nt/+)
Butylbenzyl phthalate	85-68-7	POS (11/13)	NEG (0/3)	POS (10/19)	POS	NEG (-/-)
Chrysin	480-40-0	POS (6/9)	NEG (0/4)	PP (2/10)	nt	nt
Clomiphene citrate	50-41-9	POS (3/4)	PP (1/1)	POS (8/8)	POS	nt
Corticosterone	50-22-6	NEG (0/5)	PP (1/3)	NEG (0/6)	NEG	nt
Coumestrol	479-13-0	POS (29/29)	NEG (0/8)	POS (38/38)	POS	nt
Cycloheximide	66-81-9	PN (nt)	PP (nt)	PN (nt)	nt	nt
Cyproterone acetate	427-51-0	PP (1/6)	PN (0/1)	PP (1/2)	nt	nt
Daidzein	486-66-8	POS (38/38)	NEG (0/6)	POS (32/35)	POS	POS (nt/+)
Dexamethasone	50-02-2	PP (2/6)	PP (1/1)	PP (1/4)	nt	nt
Di-n-butyl phthalate	84-74-2	PP (5/10)	NEG (0/3)	POS (7/13)	nt	NEG (-/-)
Dibenzo[a,h] anthracene	53-70-3	PP (1/2)	PP (nt)	PN (0/1)	nt	nt
Dicofol	115-32-2	POS (4/6)	NEG (0/2)	POS (2/2)	nt	nt
Diethylhexyl phthalate	117-81-7	PP (4/9)	NEG (0/3)	PP (4/8)	NEG	NEG (nt/-)
Diethylstilbestrol	56-53-1	POS (41/41)	NEG (0/2)	POS (52/52)	POS	nt
Estrone	53-16-7	POS (25/27)	PP (1/2)	POS (29/29)	POS	POS (nt/+)
Ethyl paraben	120-47-8	POS (5/5)	PN (nt)	POS (4/5)	POS	nt
Fenarimol	60168-88-9	POS (5/6)	PN (0/1)	POS (2/2)	nt	nt
Finasteride	98319-26-7	PN (nt)	PN (0/1)	PN (0/1)	nt	nt
Flavone	525-82-6	PP (2/5)	PP (1/1)	PP (3/13)	nt	nt
Fluoranthene	206-44-0	PN (nt)	PN (nt)	PN (0/1)	nt	nt
Fluoxymestrone	76-43-7	PN (nt)	PN (nt)	PN (0/1)	nt	nt
Flutamide	13311-84-7	NEG (0/5)	PN (0/1)	NEG (0/2)	nt	nt
Genistein	446-72-0	POS (99/101)	NEG (0/13)	POS (64/64)	POS	POS (+/+)
Haloperidol	52-86-8	PN (0/1)	PN (nt)	PN (0/1)	nt	nt
Hydroxyflutamide	52806-53-8	NEG (0/2)	PN (nt)	PP (1/4)	nt	nt
Kaempferol	520-18-3	POS (22/22)	NEG (0/9)	POS (19/19)	POS	nt
Kepone	143-50-0	POS (13/17)	NEG (0/2)	POS (14/15)	POS	nt
Ketoconazole	65277-42-1	PN (0/1)	PN (nt)	PN (0/1)	NEG	nt
L-thyroxine	51-48-9	POS (2/3)	PN (nt)	POS (2/2)	nt	nt
Linuron	330-55-2	NEG (0/7)	PN (nt)	POS (2/3)	NEG	nt

ICCVAM Reference Substance	CASRN	ER TA Agonist Activity	ER TA Antagonist Activity	ER Binding Activity	CER ER TA Activity	Uterotrophic Activity
Medroxyprogesterone acetate	71-58-9	PP (1/2)	PN (0/1)	POS (2/2)	NEG	nt
meso-hexestrol	84-16-2	POS (3/3)	PN (nt)	POS (11/11)	nt	nt
Methyl testosterone	58-18-4	POS (4/5)	PP (1/2)	POS (2/3)	POS	nt
Mifepristone	84371-65-3	PP (3/6)	NEG (0/3)	POS (4/6)	NEG	nt
Morin	480-16-0	PP (1/1)	PN (nt)	POS (3/3)	POS	nt
Nilutamide	63612-50-0	PN (nt)	PN (nt)	PN (nt)	nt	nt
Norethynodrel	68-23-5	POS (4/4)	NEG (2/2)	POS (7/7)	POS	na
o,p'-DDT	789-02-6	POS (24/25)	NEG (0/3)	POS (20/22)	nt	POS (+/nt)
Oxazepam	604-75-1	PN (nt)	PN (nt)	PN (nt)	nt	nt
p-n-nonylphenol	104-40-5	POS (9/9)	NEG (0/2)	POS (21/21)	NEG	IC (+/-)
p,p'-DDE	72-55-9	POS (5/7)	NEG (2/2)	PP (5/15)	nt	nt
p,p'-methoxychlor	72-43-5	POS (23/26)	PP (1/5)	POS (16/26)	POS	IC (+/-)
Phenobarbital	50-06-6	NEG (0/2)	PN (nt)	PN (0/1)	nt	nt
Phenolphthalin	81-90-3	PN (0/1)	PN (nt)	POS (2/2)	NEG	nt
Pimozide	2062-78-4	PN (nt)	PN (nt)	PN (nt)	nt	nt
Procymidone	32809-16-8	NEG (0/4)	PN (nt)	PP (2/5)	nt	nt
Progesterone	57-83-0	PP (3/15)	NEG (0/2)	PP (2/20)	NEG	nt
Propylthiouracil	51-52-5	PN (nt)	PN (nt)	PN (nt)	nt	nt
Raloxifene HCl	82640-04-8	PP (7/31)	POS (13/13)	POS (16/16)	NEG	nt
Reserpine	50-55-5	PN (0/1)	PN (nt)	PN (0/1)	NEG	nt
Resveratrol	501-36-0	POS (24/37)	NEG (0/16)	POS (9/12)	nt	nt
Sodium azide	26628-22-8	PN (0/1)	PN (nt)	PN (nt)	nt	nt
Spirolactone	52-01-7	NEG (0/3)	PN (nt)	PN (0/1)	NEG	nt
Tamoxifen	10540-29-1	POS (15/22)	POS (20/22)	POS (46/46)	POS	nt
Testosterone	58-22-0	PP (4/9)	PN (0/1)	PP (5/12)	POS	nt
Vinclozolin	50471-44-8	PP (6/13)	PN (0/1)	POS (3/5)	POS	nt

4.2.9 Weak Agonist Positive Control: Flavone

It is not clear why flavone was chosen as the weak antagonist positive control as there is scant data to support such a conclusion. The extremely high CV's noted indicate that estrogen antagonist activity of flavones is variable and is a poor candidate for a control substance.

4.3 Solubility Test Results

It does not appear that differences among the labs in range finder starting concentrations were ever fully explained. Initially in Phases 1 and 2, this was attributed to problems associated with log scale dilutions in the 1% DMSO medium. Protocols were modified after Phase 2 to use test substance solubility in 100% DMSO as the starting concentration for range finder testing. However, differences persisted in Phase 3 (Tables 4-11 and 4-12) and all three labs rarely had the same starting concentration for each substance tested.

4.4.2 BG1Luc ER TA Agonist and Antagonist Data

The table numbers in lines 250-251 should be 4-12, 4-14 and 4-15, not 4-11, 4-12, and 4-13.

5.0 Accuracy of the BG1Luc ER TA

5.1 Substances Used for Accuracy Analysis

Table 5-2: Most of the substances with little supporting data were not tested by either ECVAM or Hiyoshi (Tables 2 and 3 below) and it is not clear why they are included in the analysis. If a substance is not tested in two out of three laboratories during the validation, a consensus determination cannot be established.

The discordance in lab results for atrazine, corticosterone, and dicofol (Table 5-2) was never fully explained in the report. Atrazine and corticosterone are well-substantiated negative agonists, yet ECVAM reported a positive response for these. The discordance with dicofol (two positives, one negative) may be illustrative of the moderate amount of substantiating evidence for this substance.

Table 2. Copy of Table 5-2: Agonist substances with little or moderate substantiating data indicated.

Agonist		Classification							
Substance	CASRN	ICCVAM	Lumi Cell	XDS		ECVAM		Hiyoshi	
17 α -Estradiol	57-91-0	POS	POS	POS	(1/1)	POS	(3/3)	POS	(2/2)
17 α -Ethinyl Estradiol	57-63-6	POS	POS	POS	(3/3)	POS	(3/3)	POS	(3/3)
17 β -Estradiol	50-28-2	POS	POS	POS	(1/1)	POS	(1/1)	POS	(1/1)
19-Nortestosterone	434-22-0	POS	POS	POS	(1/1)	NT		NT	
4-Cumylphenol	599-64-4	POS	POS	POS	(1/1)	POS	(1/1)	POS	(1/1)
4-tert-Octylphenol	140-66-9	POS	POS	I	(1/1)	POS	(1/1)	POS	(2/2)
5 α -Dihydrotestosterone	521-18-6	POS	I	I	(1/1)	I	(1/1)	POS	(1/1)
Apigenin	520-36-5	POS	POS	POS	(1/1)	POS	(1/1)	POS	(1/1)
Atrazine	1912-24-9	NEG	NEG	NEG	(3/3)	POS	(3/3)	NEG	(3/3)
Bicalutamide	90357-06-5	NEG	NEG	NEG	(1/1)	NT		NT	
Bisphenol A	80-05-7	POS	POS	POS	(3/3)	POS	(3/3)	POS	(3/3)
Bisphenol B	77-40-7	POS	POS	POS	(3/3)	POS	(3/3)	POS	(3/3)
Butylbenzyl phthalate	85-68-7	POS	POS	POS	(3/3)	POS	(3/3)	POS	(3/3)
Chrysin	480-40-0	POS	POS	POS	(2/2)	NT		NT	
Clomiphene citrate	50-41-9	POS	I	I	(1/1)	NEG	(1/1)	POS	(1/1)
Corticosterone	50-22-6	NEG	NEG	NEG	(3/3)	POS	(3/3)	NEG	(3/3)
Coumestrol	479-13-0	POS	POS	POS	(1/1)	POS	(1/1)	POS	(1/1)
Daidzein	486-66-8	POS	POS	POS	(1/1)	POS	(1/1)	POS	(1/1)
Dicofol	115-32-2	POS	POS	POS	(1/1)	NEG	(1/1)	POS	(1/1)
Diethylstilbestrol	56-53-1	POS	POS	POS	(3/3)	POS	(3/3)	POS	(3/3)
Estrone	53-16-7	POS	POS	POS	(1/1)	POS	(1/1)	POS	(1/1)

Agonist		Classification							
Substance	CASRN	ICCVAM	Lumi Cell	XDS		ECVAM		Hiyoshi	
Ethyl paraben	120-47-8	POS	POS	I	-1	POS	(1/1)	POS	(1/1)
Fenarimol	60168-88-9	POS	POS	POS	(1/1)	NT		NT	
Flutamide	13311-84-7	NEG	I	I	-1	NT		NT	
Genistein	446-72-0	POS	POS	POS	(3/3)	POS	(3/3)	POS	(4/4)
Hydroxy Flutamide	52806-53-8	NEG	NEG	NEG	(1/1)	NEG	(1/1)	NEG	(1/1)
Kaempferol	520-18-3	POS	POS	POS	(1/1)	POS	(1/1)	POS	(1/1)
Kepone	143-50-0	POS	POS	POS	(1/1)	POS	(1/1)	POS	(1/1)
L-Thyroxine	51-48-9	POS	NEG	NEG	(1/1)	NT		NT	
Linuron	330-55-2	NEG	NEG	NEG	(1/1)	NT		NT	
meso-Hexestrol	84-16-2	POS	POS	POS	(1/1)	POS	(1/1)	POS	(1/1)
Methyl testosterone	58-18-4	POS	POS	POS	(3/3)	POS	(1/1)	POS	(2/2)
Norethynodrel	68-23-5	POS	POS	POS	(2/2)	POS	(1/1)	POS	(2/2)
o,p'-DDT	789-02-6	POS	POS	POS	(3/3)	POS	(3/3)	POS	(3/3)
p-n-Nonylphenol	104-40-5	POS	POS	POS	(3/3)	POS	(3/3)	POS	(2/3)
p,p'-Methoxychlor	72-43-5	POS	POS	POS	(1/1)	POS	(1/1)	POS	(2/2)
p,p'-DDE	72-55-9	POS	I	I	(1/1)	I	(1/1)	NEG	(1/1)
Phenobarbital	50-06-6	NEG	NEG	NEG	(1/1)	NEG	(1/1)	NT	
Procymidone	32809-16-8	NEG	I	I	(1/1)	NT		NT	
Resveratrol	501-36-0	POS	I	POS	(1/1)	I	(1/1)	NEG	(1/3)
Spironolactone	52-01-7	NEG	NEG	NEG	(1/1)	NT		NT	
Tamoxifen	10540-29-1	POS	I	I	(1/1)	I	(1/1)	POS	(1/1)

Table 3. Copy of Table 5-3: Antagonist substances with little or moderate substantiating data indicated.

Antagonist		Classification							
Substance	CASRN	ICCVAM	Lumi Cell	XDS		ECVAM		Hiyoshi	
17 α -Estradiol	57-91-0	NEG	NEG	NEG	(1/1)	NEG	(1/1)	NEG	(1/1)
4-Hydroxytamoxifen	68047-06-3	POS	POS	POS	(1/1)	I	(2/2)	POS	(1/1)
5 α -Dihydrotestosterone	521-18-6	NEG	NEG	NEG	(1/1)	NEG	(1/1)	NEG	(1/1)
Apigenin	520-36-5	NEG	NEG	NEG	(3/3)	NEG	(3/3)	NEG	(4/4)
Bisphenol A	80-05-7	NEG	NEG	NEG	(1/1)	NEG	(1/1)	NEG	(1/1)
Butylbenzyl phthalate	85-68-7	NEG	NEG	NEG	(3/3)	NEG	(3/3)	NEG	(4/4)
Chrysin	480-40-0	NEG	NEG	NEG	(1/1)	NT		NT	
Coumestrol	479-13-0	NEG	NEG	NEG	(1/1)	NEG	(1/1)	NEG	(1/1)
Daidzein	486-66-8	NEG	NEG	NEG	(1/1)	NEG	(1/1)	NEG	(1/1)
Di-n-butyl phthalate	84-74-2	NEG	NEG	NEG	(1/1)	NEG	(1/1)	NEG	(1/1)
Dicofol	115-32-2	NEG	NEG	NEG	(1/1)	NEG	(1/1)	NEG	(1/1)
Diethylhexyl	117-81-7	NEG	NEG	NEG	(1/1)	NEG	(1/1)	NEG	(1/1)

phthalate									
Diethylstilbestrol	56-53-1	NEG	NEG	NEG	(1/1)	NEG	(1/1)	POS	(1/1)
Genistein	446-72-0	NEG	NEG	NEG	(3/3)	NEG	(3/3)	NEG	(3/3)
Kaempferol	520-18-3	NEG	NEG	NEG	(1/1)	NEG	(1/1)	NEG	(1/1)
Kepone	143-50-0	NEG	NEG	NEG	(1/1)	NEG	(1/1)	NEG	(1/1)
Mifepristone	84371-65-3	NEG	NEG	NEG	(1/1)	NT		NT	
Norethynodrel	68-23-5	NEG	NEG	NEG	(1/1)	NEG	(1/1)	NEG	(1/1)
o,p'-DDT	789-02-6	NEG	NEG	NEG	(3/3)	NEG	(3/3)	NEG	(3/3)
p-n-Nonylphenol	104-40-5	NEG	NEG	NEG	(3/3)	NEG	(3/3)	NEG	(3/3)
p,p'-DDE	72-55-9	NEG	NEG	NEG	(1/1)	NEG	(1/1)	NEG	(1/1)
Progesterone	57-83-0	NEG	NEG	NEG	(3/3)	NEG	(3/3)	NEG	(3/3)
Raloxifene HCl	82640-04-8	POS	POS	POS	(1/1)	POS	(1/1)	POS	(1/1)
Resveratrol	501-36-0	NEG	NEG	NEG	(3/3)	NEG	(3/3)	NEG	(3/3)
Tamoxifen	10540-29-1	POS	POS	POS	(3/3)	POS	(3/3)	POS	(3/3)

5.4 Comparison of BG1Luc ER TA Results with CER1 STTA (OPPTS 890.1300)

This qualitative comparison is helpful for determining the relative utility of the two assays; however, it would be more informative to include a quantitative comparison as well, as we have done in **Table 4** below. During the OECD validation of the CER1 STTA assay, it was decided that a useful exercise would be to use the ER STTA assay as a proof-of-concept for a Performance-Based Test Guideline (PBTG). The objective is to use two validated assays, in this case the CER1 STTA assay and now the BG1Luc ER TA assay (agonist version) to create a set of performance standards that can be used to evaluate and expedite validation of subsequent similar assays. To compare assays that may generate different types of data and utilize different decision criteria, it is useful to present data as a Relative Potency Index (RPI) in addition to EC₅₀. The RPI is the EC₅₀ of the positive control divided by the EC₅₀ of the chemical multiplied by 100. We suggest that the RPI be added to Table 5-7.

In addition, several chemicals that were tested in the validation of the CER1 STTA method are missing from the comparison in Table 5-7, including clomiphene citrate, methoxychlor and tamoxifen.

6.0 Test Method Reliability

6.1.6 Antagonist E2 Control Values

Line 185: Table 6-3 should be Table 6-6.

9.0 Animal Welfare Considerations

Lines 32 – 35: Contain a direct repeat of lines 17 – 19 and should be deleted.

Based on a 97% concordance (33/34) of findings from the BG1Luc ER TA assay and the ER rat cytosol binding assay it is suggested that the former could serve as a replacement for the latter. Following the same logic, a 92% concordance (12/13, with no false negatives) should argue for

that the BG1Luc ER TA assay could replace the uterotrophic assay. In fact, in the interest of reducing animal use, a strong recommendation should be made to investigate the use of *in vitro* metabolizing systems with the BG1Luc ER TA assay so that the ER TA could definitively replace the uterotrophic assay.

Lines 52 – 54: Contain a direct repeat of lines 20 -22 and should be deleted.

10.1.3 BG1LUC 4E2 Cell line

If the line is available only from a private academic lab, will supply and quality control (e.g. passage number) be an issue?

10.3 Time and Cost Considerations

Lines 67 and 75: Authors' names are misspelled: should be Willett and Sullivan.

In conclusion, we find the BG1Luc test to be an accurate method for both qualitatively and quantitatively assessing the ER TA agonist and antagonist potential of a wide range of substances. We urge you to further enhance the utility of this method by pursuing many of the recommendations in the report as well as our recommendations, such as incorporating the use of *in vitro* metabolizing systems. We also ask that you reconsider your list of 78 substances when validating future ER/AR binding and TA tests and only use chemicals that have been definitively evaluated for their effects. Finally, in light of the need for new tests that can reduce, refine or replace animals in testing, we suggest a thorough examination of the validation process used in this study to determine if there are ways to make future studies more streamlined and time-efficient while still meeting the needs of public health and welfare.

Sincerely,

/s/

Catherine Willett, PhD
Science Policy Advisor
Regulatory Testing Division
Tel: 617-522-3487

/s/

Patricia L. Bishop, M.S.
Research Associate
Regulatory Testing Division
Tel: 757-390-0564

Appendix E5

76 FR 23323

**Meeting of the Scientific Advisory Committee on Alternative Toxicological Methods
(SACATM)**

Public Comments Received in Response to 76 FR 23323

- Nancy Beck, PhD (Physicians Committee for Responsible Medicine) and Samantha Dozier, PhD (People for the Ethical Treatment of Animals).....E-41
- Matthew Stoner, PhD (CertiChem, Inc.)E-52

This page intentionally left blank

DEPARTMENT OF HEALTH AND HUMAN SERVICES

Meeting of the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM)

AGENCY: National Toxicology Program (NTP), National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH).

ACTION: Meeting announcement and request for comments.

SUMMARY: Pursuant to section 10(a) of the Federal Advisory Committee Act, as amended (5 U.S.C. Appendix 2), notice is hereby given of a meeting of SACATM on June 16–17, 2011, at the Hilton Arlington Hotel, 950 North Stafford Street, Arlington, VA 22203. The meeting is open to the public with attendance limited only by the space available. The meeting will be videocast through a link at (<http://www.niehs.nih.gov/news/video/live>). SACATM advises the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), the NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), and the Director of the NIEHS and NTP regarding statutorily mandated duties of ICCVAM and activities of NICEATM. **DATES:** The SACATM meeting will be held on June 16 and 17, 2011. The meeting is scheduled from 8:30 a.m. Eastern Daylight Time to 5:30 p.m. on June 16 and 8:30 a.m. until adjournment on June 17. All individuals who plan to attend are encouraged to register online at the NTP Web site (<http://ntp.niehs.nih.gov/go/32822>) by June 9, 2011. In order to facilitate planning, persons wishing to make an oral presentation are asked to notify Dr. Lori White, NTP Designated Federal Officer, via online registration, phone, or email by June 9, 2011 (see **ADDRESSES** below). Written comments should also be received by June 9, 2011, to enable review by SACATM and NIEHS/NTP staff before the meeting.

ADDRESSES: The SACATM meeting will be held at the Hilton Arlington Hotel, 950 North Stafford Street, Arlington, VA 22203. Public comments and other correspondence should be directed to Dr. Lori White (NTP Office of Liaison,

Policy and Review, NIEHS, P.O. Box 12233, MD K2–03, Research Triangle Park, NC 27709; telephone: 919–541–9834 or e-mail: whitelord@niehs.nih.gov). Courier address: NIEHS, 530 Davis Drive, Room 2136, Morrisville, NC 27560. Persons needing interpreting services in order to attend should contact 301–402–8180 (voice) or 301–435–1908 (TTY). Requests should be made at least 7 days in advance of the meeting.

SUPPLEMENTARY INFORMATION:

Preliminary Agenda Topics and Availability of Meeting Materials

Preliminary agenda topics include:

- NICEATM–ICCVAM Update
- Regulatory Acceptance of ICCVAM-Recommended Alternative Test Methods
- Report on Peer Review Panel Meeting: Evaluation of an *In Vitro* Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening
- Federal Agency Research, Development, Translation, and Validation Activities Relevant to the NICEATM–ICCVAM Five-Year Plan
- Nominations to ICCVAM: Botulinum *In Vitro* Assays, *In Vitro* Pyrogen Assay Validation
- Outcome/Recommendations from the ICCVAM Workshop Series on Best Practices for Regulatory Safety Testing
- Outcomes/Recommendations from the International Workshop on Alternative Methods to Reduce, Refine, and Replace the Use of Animals in Vaccine Potency Testing: State of the Science and Future Directions
- Update from the Korean Center for the Validation of Alternative Methods
- Update from Health Canada
- Update from the Japanese Center for the Validation of Alternative Methods
- Update from the European Centre for the Validation of Alternative Methods

A copy of the preliminary agenda, committee roster, and additional information, when available, will be posted on the NTP Web site (<http://ntp.niehs.nih.gov/go/32822>) or available upon request (see **ADDRESSES** above). Following the SACATM meeting, summary minutes will be prepared and available on the NTP Web site or upon request.

Request for Comments

Both written and oral public input on the agenda topics is invited. Written comments received in response to this notice will be posted on the NTP Web site. Persons submitting written comments should include their name, affiliation (if applicable), and

sponsoring organization (if any) with the document. Time is allotted during the meeting for presentation of oral comments and each organization is allowed one time slot per public comment period. At least 7 minutes will be allotted for each speaker, and if time permits, may be extended up to 10 minutes at the discretion of the chair. Registration for oral comments will also be available on-site, although time allowed for presentation by on-site registrants may be less than for pre-registered speakers and will be determined by the number of persons who register at the meeting. In addition to in-person oral comments at the meeting, public comments can be presented by teleconference line. There will be 50 lines for this call; availability will be on a first-come, first-served basis. The available lines will be open from 8 a.m. until 5 p.m. on June 16 and 8:30 a.m. to adjournment on June 17, although public comments will be received only during the formal public comment periods, which will be indicated on the preliminary agenda.

The access number for the teleconference line will be provided to registrants by email prior to the meeting.

Persons registering to make oral comments are asked to do so through the online registration form (<http://ntp.niehs.nih.gov/go/32822>) and to send a copy of their statement to Dr. White (see **ADDRESSES** above) by June 9, 2011, to enable review by SACATM, NICEATM–ICCVAM, and NIEHS/NTP staff prior to the meeting. Written statements can supplement and may expand the oral presentation. If registering on-site and reading from written text, please bring 40 copies of the statement for distribution and to supplement the record.

Background Information on ICCVAM, NICEATM, and SACATM

ICCVAM is an interagency committee composed of representatives from 15 Federal regulatory and research agencies that require, use, generate, or disseminate toxicological and safety testing information. ICCVAM conducts technical evaluations of new, revised, and alternative safety testing methods with regulatory applicability and promotes the scientific validation and regulatory acceptance of toxicological and safety testing methods that more accurately assess the safety and hazards of chemicals and products and that reduce, refine (decrease or eliminate pain and distress), or replace animal use. The ICCVAM Authorization Act of 2000 (42 U.S.C. 285f–3) established ICCVAM as a permanent interagency committee of the NIEHS under

23324

Federal Register / Vol. 76, No. 80 / Tuesday, April 26, 2011 / Notices

NICEATM. NICEATM administers ICCVAM, provides scientific and operational support for ICCVAM-related activities, and conducts independent validation studies to assess the usefulness and limitations of new, revised, and alternative test methods and strategies. NICEATM and ICCVAM work collaboratively to evaluate new and improved test methods and strategies applicable to the needs of U.S. Federal agencies. NICEATM and ICCVAM welcome the public nomination of new, revised, and alternative test methods and strategies for validation studies and technical evaluations. Additional information about ICCVAM and NICEATM can be found on the NICEATM-ICCVAM Web site (<http://iccvam.niehs.nih.gov>).

SACATM was established in response to the ICCVAM Authorization Act [Section 285I-3(d)] and is composed of scientists from the public and private sectors. SACATM advises ICCVAM, NICEATM, and the Director of the NIEHS and NTP regarding statutorily mandated duties of ICCVAM and activities of NICEATM. SACATM

provides advice on priorities and activities related to the development, validation, scientific review, regulatory acceptance, implementation, and national and international harmonization of new, revised, and alternative toxicological test methods. Additional information about SACATM, including the charter, roster, and records of past meetings, can be found at <http://ntp.niehs.nih.gov/go/167>.

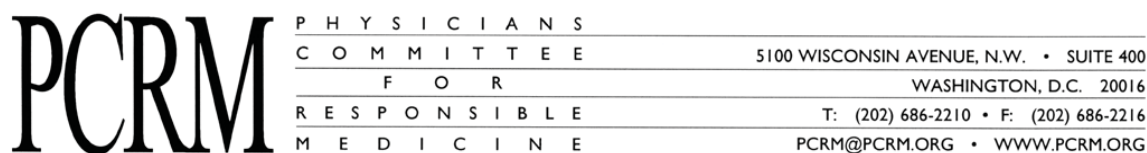
Dated: April 18, 2011.

John R. Bucher,

Associate Director, National Toxicology Program.

[FR Doc. 2011-10020 Filed 4-25-11; 8:45 am]

BILLING CODE 4140-01-P



June 9, 2011

Dr. Lori White
NTP Office of Liaison, Policy and Review
National Institute of Environmental Health Sciences
PO Box 12233, MD K2-03
Research Triangle Park, NC 27709

Dear Dr. White:

The following comments are submitted on behalf of the more than two million members and supporters of the Physicians Committee for Responsible Medicine (PCRM) and People for the Ethical Treatment of Animals (PETA) in response to the nominations of *in vitro* botulinum and *in vitro* pyrogen assays to ICCVAM for validation. (April 26, 2011; Federal Register 76[80]:23323). Our organizations are committed to replacing and reducing animal use with the best available *in vitro* and computational science.

MAT Pyrogenicity Assay

We support an expanded domain of applicability for the nominated MAT, but we have suggestions for how to proceed with validation that differ from those in the nomination documented submitted. A summary of the major recommendations are listed below, followed by more detailed comments.

Major Recommendations

1. Coordinate and collect data from smaller, product-specific validation studies from individual companies that still use the RPT rather than require that a large validation study be conducted.
2. Expand efforts to gather information on pyrogenicity testing from manufacturers of parenterals, biologics, and devices as well as the FDA
3. Perform a current literature review on the immune response to pyrogens.

Background

BioTest has nominated its Monocyte Activation Test (MAT), an IL-1 β enzyme-linked immunosorbent assay (ELISA) using cryopreserved blood, for broader validation. This assay was one of the five *in vitro* pyrogenicity assays approved by ICCVAM in 2008 for

the detection of pyrogenic contamination. These five methods were approved for the detection of Gram-negative endotoxin in human parenteral drugs, subject to product specific validation. BioTest has proposed a validation study to expand the domain of applicability to include non-endotoxin pyrogens as well as products other than parenterals, including devices and biologics, with the intent of fully replacing the rabbit pyrogen test (RPT). Although there were abundant data on the ability of the five methods to detect both endotoxin and non-endotoxin pyrogen in a variety of product classes, ICCVAM found the data inadequate for validation purposes, and limited the validation to endotoxin in parenterals.

Due to the perceived limitations of the previous validation studies, pyrogencity testing of biologics and devices and testing for non-endotoxin pyrogens still requires animal-based testing. The two other primary methods for pyrogen testing are the Rabbit Pyrogen Test (RPT) and the Limulus Amebocyte Lysate (LAL). Although the LAL is an *in vitro* test, it is not a non-animal test, as it relies on the blood of horseshoe crabs. The LAL is only capable of detecting endotoxin, which means the RPT is used when non-endotoxin pyrogens are a concern. The LAL has other technical limitations, such as incompatibility with certain types of drugs and biologics that may also lead to the use of rabbits for detection of pyrogenic contamination.

While we appreciate BioTest's intent to expand the use of the MAT in order to replace the RPT, we are concerned about the rabbit use proposed for the validation study. BioTest has suggested a validation study that includes the RPT and LAL along with the MAT. Inclusion of these assays in parallel is an attempt to address the ICCVAM recommendations for future studies enumerated in the 2008 Test Method Evaluation Report (TMER), section 2.3. BioTest also proposes to include endotoxin and non-endotoxin standards (lipotechoic acid and crude preparations from Gram-positive bacteria), a pro-inflammatory substance, parenteral pharmaceuticals, biologics, and devices. However, we question the need for parallel LAL and RPT testing given the inability of the LAL to detect non-endotoxin pyrogens and the abundance of existing LAL and RPT reference data available for comparison and extrapolation.

Product Specific Validation Versus a Large-scale Validation Study

If all the reference standards and classes of products proposed are tested in rabbits, this study could lead to significant animal use. The number of animals who would be consumed by parallel testing is one of the reasons that RPT studies were not conducted as part of the original validation study performed by the European Center for the Validation of Alternative Methods (ECVAM). Another reason cited by ECVAM is the fact that it is common practice for a manufacturer to validate pyrogen tests for every given product. Rather than conducting a massive and animal-intensive validation study, ECVAM opted for a smaller study to demonstrate the general applicability and validity of the methods for regulatory purposes, leaving validation of the assays for additional pyrogens and product classes up to manufacturers. This sensible approach should be applied here to prevent the duplicative use of rabbits in an ICCVAM validation, which would then be followed by a product specific validation anyway. ICCVAM and BioSentinel should take advantage of RPTs currently taking place for regulatory purposes and facilitate product

specific validation of the MAT. Collection of this data could, over time, fulfill data needs for validation of the MAT.

ICCVAM's 2008 Background Review Document (BRD) also acknowledges that product-specific validation of these methods is ultimately part of regulatory practice in the U.S. and in the European Union. Again, this leads us to question whether a large scale validation study including non-endotoxin pyrogens and products beyond parenteral drugs is necessary, particularly since there are numerous studies demonstrating proof of concept for expanded use. There is little doubt that the MAT method nominated, as well as the other 4 versions of the MAT, will work for all pyrogens in a wide array of products (this point is further elaborated later in these comments).

We urge ICCVAM to consider facilitating the expanded use of the MATs via coordination of smaller product-specific validation studies rather than causing additional harm to many animals. ICCVAM could collaborate with FDA to encourage parenteral, biologics, and device manufacturers that still use the RPT to conduct and share the results of their own MAT validation studies. A coordinated effort could build a body of data that would supplant the perceived need for a large prospective study. For example, in 2009, Pfizer and the Center for Biologics Evaluation and Research at FDA published a paper examining the use of certain cell lines for the detection of non-endotoxin pyrogens in a biological product¹. This paper illustrates the type of collaboration and data-sharing that would be useful to advance pyrogenicity testing. One step in this direction is NICEATM's request for data on non-endotoxin pyrogens that appeared in the May 23 issue of the Federal Register. Perhaps this can be used as a starting point to establish collaborations with companies that submit data.

Enhanced Information Gathering

When determining how to proceed, ICCVAM must gather additional information to inform the approach. For example, it would be very useful to know when and how often the RPT is still performed and whether it is used because the LAL will not suffice due to technical limitations or because of a need to address non-endotoxin contamination. If the need to detect non-endotoxin pyrogen is rarely the reason that the RPT is performed, then the need for a validation study of this magnitude is unclear. Other information useful for informing this process includes an understanding of the following: which of the five MATs is most used; whether there is reluctance to use the whole blood-based assays; what product-specific validation studies have been conducted and submitted to FDA.

Scientific Support for the Expanded Use of MATs

There is already very compelling scientific evidence to support the expanded use of MATs without further testing. We question how much more evidence is necessary for ICCVAM to consider the science valid. At the time of the 2008 validation of the MATs for detection of endotoxin in parenteral drugs, there were data available from numerous studies assessing the ability of the MAT's to detect both endotoxin and non-endotoxin pyrogens in a variety of parenterals and some biologics. There were also data available

¹ Huang L, et al. 2009. Use of Toll-Like Receptor Assays to Detect and Identify Microbial Contaminants in Biological Products. *Journal of Clinical Microbiology*. 47(11):3427-3434

directly comparing the MAT with parallel RPTs involving data from hundreds of rabbits. Beyond this evidence, the biological and mechanistic basis of the MATs is well understood. There is little question that the MAT is capable of detecting both endotoxin and non-endotoxin pyrogens^{2, 3}.

The basis of the MATs is the detection of particular cytokines, either IL-1 β or IL-6. These cytokines are known to mediate fever in humans. In fact, they are an integral part of the molecular pathways leading to fever, regardless of the source of pyrogen. Over the last 10 years or so, tremendous progress has been made in understanding how the immune system recognizes and responds to pathogens and pyrogens, which are really just fragments of pathogens^{4, 5}. In the parlance of modern immunology, pyrogens are referred to as pathogen-associated molecular patterns, or PAMPs. Pyrogens/PAMPs are recognized and bound by a few different Toll-like receptors (TLRs), which sit at the top of convergent molecular pathways leading to fever. All of these pathways involve IL-1 β and IL-6 production as part of the signaling necessary to induce inflammation and fever to deal with microbial threats (Table 1 and Figure 1). Because this anti-microbial response is critical for survival, we have evolved a limited number of conserved signal transduction pathways to mediate this response. By measuring the induction of IL-1 β and IL-6, we can be certain that we are detecting all pyrogenic contamination.

ICCVAM seems to be fixated on the distinction between endotoxin and non-endotoxin pyrogens, but this is an arbitrary distinction in the case of the MATs. These assays are capable of detecting all pyrogens through TLR signaling cascades, leading to induction of IL-1 β and IL-6. The limitation of the LAL, which can only detect endotoxin, may be responsible for the apprehension that is preventing broader application of the MATs. ICCVAM's concerns are misplaced, since MATs are based on a completely different, highly conserved, universal mammalian mechanism--unlike the LAL.

It is illogical to make a dramatic distinction between endotoxin and non-endotoxin pyrogens, while making no such distinction between all the different types of non-endotoxin pyrogens. Endotoxin is one molecular entity, lipopolysaccharide (LPS) from the cell wall of Gram-negative bacteria, but non-endotoxin pyrogens are not one type of substance. Non-endotoxin pyrogens include every other pyrogen aside from LPS, including surface proteins, secreted proteins/toxins, lipoproteins, glycoprotein, lipotechoic acid, peptidoglycan, and nucleic acids. These PAMPs/non-endotoxin pyrogens come from fungal, viral, parasitic, and bacterial (both Gram-positive and Gram-negative) sources. MATs measuring IL-1 β or IL-6 have been shown to detect PAMPs/pyrogens regardless of their composition or origin, which addresses the concern

² Schindler S, et al. 2009. Development, Validation, and Applications of the Monocyte Activation Test for Pyrogens Based on Human Whole Blood. *ALTEX*. 26(4):265-277.

³ Banerjee, S and Mohanan, P.V. 2011. Inflammatory Response to Pyrogens Determined by a Novel ELISA Method Using Human Whole Blood. *Journal of Immunological Methods*. *In Press*.

⁴ Abdul-Sater, A.A. et al. 2009. Inflammasomes Bridge Signaling Between Pathogen Identification and the Immune Response. *Drugs Today*. 45(Suppl. B): 105-112.

⁵ Mogensen, T.H. 2009. Pathogen Recognition and Inflammatory Signaling in Innate Immune Defenses. *Clinical Microbiology Reviews*. 22(2):240-273.

about differentiating between endotoxin and non-endotoxin pyrogens, as well as between the different types of non-endotoxin pyrogens.

Although there are likely to be differences in the timing and levels of cytokine induction between different pyrogens (even from different sources of LPS), these differences can be accounted for by comparing the cytokine profiles arising from LPS and a variety of other PAMPs (i.e. non-endotoxin pyrogens) to ensure the timing of the assay is optimized to detect peak cytokine induction from a broad range of pyrogens. Once optimization is ensured, it is simply a matter of comparing the level of IL-1 β or IL-6 induced by a test substance to the levels induced by a pyrogen standard over a concentration range spanning the fever threshold. Stated another way, the concentration of LPS known to cause fever in rabbits (or humans) can be tested in the MAT and the levels of cytokines produced in response can be used to establish the threshold for gauging pyrogenicity. Bridging between studies in this manner should be sufficient to replace the parallel rabbit testing suggested.

Useful information on the timing of IL-1 β and/or IL-6 induction arising from non-endotoxin pyrogenic contamination may already be available. IL-1 β and/or IL-6 has been well documented using MATs for numerous gram positive species (*Alcyclobacillus acidocaldarius*, *Staphylococcus aureus*, *Corynebacterium diphtheriae*, *Clostridium tetani*, *Bacillus subtilis*, *Bacillus stearothermophilus*, *Micrococcus luteus*, *Lactobacillus plantarum*, group B streptococcus, and *Streptococcus pneumoniae*), mycobacteria, and numerous fungal species (*Candida albicans*, *Aspergillus niger*, *Aspergillus versicolor*, *Alternaria alternate*, *Cladosporium cladosporoides*, and *Penicilium crustosum*). Induction of IL-1 β and/or IL-6 have also been demonstrated for specific non-endotoxin pyrogens/ PAMPs including lipotechoic acid, diphtheria toxin, tetanus toxin, spores of Gram positives and fungi, as well as peptidoglycan and other constituents of both Gram-positive and Gram-negative cell walls. Furthermore, many of these analyses were performed using biologics and devices as well as parenterals, so there is a strong precedent for the detection of non-endotoxin in products other than parenteral drugs.

Need for an Updated Literature Review

Before proceeding with further validation of the nominated MAT, we strongly recommend conducting an updated review of the literature. We are surprised and disappointed that a more exhaustive review of current findings was not submitted in support of the nomination. We hope that the 2008 BRD and TMER are not relied upon for information on the state of the science. The TMER contains outdated references and reaches some misinformed conclusions, for example "The development of tests based on the production of such cytokines [IL-1 β , IL-6, and TNF] from human white blood cells or cell lines appears to correlate well with the induction of fever in both the RPT and humans. However, the RPT detects a whole organ/body fever response; whereas, the proposed test methods detect only cytokine secretion. Evidence to suggest that detection of IL-1 β or IL-6 is necessarily an indication of a febrile reaction is lacking."⁶ Numerous

⁶ Hoffman, H.M. and Brydges, S.D. 2011. Genetic and Molecular Basis of Inflammasome-mediated Disease. *Journal of Biological Chemistry*. 286(13):10889-10896.

studies and reviews have been published in the last several years to address these concerns and significantly inform the validation process.

We hope that, in accordance with ICCVAM's Congressional mandate, the increased use of MATs can be achieved quickly and simply. It is troubling that approval for an expanded domain of applicability is still an issue, given the extensive scientific evidence to support detection of non-endotoxin pyrogens in varied products. The MATs are affordable and practical, based on straightforward, well-established, and simple ELISA "technology" that is widely used throughout even the most modestly equipped labs. There appear to be few downsides to the MATs, which detect a broader array of pyrogens than the LAL and are more sensitive than the RPT.

Botulinum Neurotoxin (BoNT) Activity Assays

Clostridium botulinum toxin testing applications range from food safety needs to vaccine potency-type tests, and also include field tests for outbreaks in lakes and in animals such as waterfowl, horses, cows, domestic poultry, fish, and fish-eating birds. Botulinum toxin is produced by *Clostridium botulinum*, a gram positive, spore-producing, anaerobic bacterium that is capable of producing up to seven different serotypes of botulinum toxin. It is a potent neurotoxin that: (1) can contaminate food sources, (2) can be used in pharmaceutical applications, and (3) is also thought to be a concern related to military defense applications. Types A, B, E, and F are inherited chromosomally, while types C and D are transmitted to *C. botulinum* via bacteriophages, and type G is exchanged on plasmids. Serotype A is used most often in pharmaceutical preparations.

United States regulatory and governmental agencies such as the Food & Drug Agency's (FDA) Center for Food Safety and Applied Nutrition (CFSAN), FDA's Center for Biologics Evaluation & Research (CBER) and The US Geological Survey's National Wildlife Health Center (USGS NWHC) currently rely on the mouse bioassay for detecting *C. botulinum*-based toxins. It has long been proposed that a user-friendly, rigorous, non-animal-based replacement to the mouse bioassay (MBA) is needed due to ethical concerns as described below.

The MBA for *C. botulinum* toxin detection requires a dilution series of the toxin to be injected intraperitoneally into multiple mice. The resulting poisoning causes an often slow and painful death, ultimately culminating in respiratory paralysis. The MBA requires at least 48 mice for testing in food safety and approximately 100 mice per potency test for a single batch of Botox. The MBA accounts for the use of an estimated 74,000 mice in a single Botox manufacture's lab in one year and for 600,000 mice used worldwide per year by pharmaceutical companies, alone.

In addition to animal welfare concerns, the MBA also has scientific and practical shortcomings. One study in California found that only 68% of tests run on the serum from 73 patients with wound botulism resulted in positive MBA tests⁷, illustrating a

⁷ Wheeler C, Inami G, Mohle-Boetani J, Vugia D. Sensitivity of mouse bioassay in clinical wound botulism. Clin Infect Dis. 2009;48:1669-73.

significant false negative rate by the MBA. The MBA is incapable of distinguishing BoNT serotype unless neutralization tests with each antisera are carried out in parallel. Additionally, the MBA takes a week to complete, requires trained staff, ample cage space within animal facilities, licensing for the use of the mice, and considerable costs for the week-long test and the staff to carry it out.

In order to move past reliance on the MBA, a desirable test method candidate for monitoring and quantifying *C. botulinum* toxins must be: (1) indicative of active toxin(s), (2) inexpensive, (3) rapid, and (4) sensitive, and have the capability to be used in field monitoring. ICCVAM received a nomination to evaluate three botulinum neurotoxin detection methods developed by BioSentinel, Inc., (Madison, WI) and has requested public comment on each of the three methods.

BoTest™

BoTest™ botulinum neurotoxin (BoNT) detection kits are capable of detecting *C. botulinum* A, B, D, E, F and G proteolytic activity. BoTest™ is intended to be used with purified samples and for drug discovery. Depending on the serotype being tested, BoTest™ offers mouse-level or near mouse-level sensitivity. In contrast to the MBA, the detection capabilities by BoTest™ are in real time, the output signal wavelength is tunable, and the system requires only a small amount of training to use.

A pharmaceutical company has validated BoTest™ assays for the quantification of drug formulations and products and is in the process of performing comparability studies between BoTest™ and the MBA. This valuable data should be used by ICCVAM as part of the requirement for method validation.

Currently, specific BoTest™ substrates are available for six of the seven serotypes of botulinum toxins. BoTest™ A/E uses a SNAP-25-based reporter, while BoTest™ B/D/F/G assay uses synaptobrevin specific for those subtypes. The BoTest substrates are comprised of much larger fragments from the substrate proteins than competing commercial assays and therefore the BoNT has much higher affinity for these substrates than those used in other commercially prepared reporter systems, although the data to support this was not provided in the supporting documents. Because these assays quantify the endopeptidase activity of the BoNT using reporters linked to modified endogenous BoNT targets, the BoTest™ assays are biologically relevant.

BoTest™ takes advantage of Forster Resonance Energy Transfer (FRET) donor-acceptor pair fluorescence. BoNT cleavage of the substrate leads to decrease of visible yellow fluorescent protein (YFP) emission with a concomitant increase in cyan fluorescent protein (CFP) emission. The reactions are measurable in real-time and emissions can be quantified and enzymatic activity can therefore be determined.

Cost comparison by BioSentinel of BoTest™ with the MBA shows BoTest™ is much less expensive, at a cost \$875.00 per test (including labor) compared with an estimated \$6000.00 for the MBA, which would also result in the suffering and death of up to 300

mice. BoTest™ can be completed in as little as two hours while the mouse test commonly takes a week or more.

BoTest™ is biologically relevant and has sensitivity in the femtomolar and picomolar range, making this method unique among commercially available assays and also a strong contender to completely replace animal testing as part of a suite of MBA replacement assays.

BoTest™ Matrix

BoTest™ Matrix kits were specifically designed to be used in complex matrices (blood, serum, water, pharmaceutical products, and food). The Matrix test kits can detect botulinum A and E serotypes and use magnetic beads to capture and concentrate BoNT-containing matrices.

The reporter substrates are the same as those used for BoTest™ and also use FRET detection of reporter cleavage in real time. Sensitivity rivals that of the MBA (femtomolar and picomolar range), but is higher throughput. Costs for the BoTest Matrix™ kit is estimated \$1930.00 at a commercial laboratory while the MBA would cost \$6000.00. The BoTest Matrix™ kits are readily transferable between labs and require little training and, like the BoTest™, take as little as two hours to complete compared to a week for the MBA.

BoCell™

The BoCell™ A assay uses an engineered cell line that responds to intracellular BoNT proteolytic activity by use of a stably transfected reporter and is specific for serotype A. The BoCell™ assay is intended to be used as a direct replacement for the mouse bioassay. The substrate is a SNAP-25-based sequence fused to two fluorescent reporter proteins. BoNT cleavage activity is detected by either a loss FRET pairing or by destruction of the C-terminal fluorophore.

BoCell™ can be carried out in any lab capable of basic tissue culture and requires minimal training, which is in contrast to the training and certification required for the MBA.

BoCell™ offers a high throughput method of BoNT detection without use of any animals and with minimal equipment. At this time, the BoCell™ is not as sensitive as the MBA (2 – 3 orders of magnitude less sensitive), but for many applications, this level of sensitivity is acceptable. BioSentinel is recommending BoCell™ in combination with BoTest™ or BoTest Matrix™ assays for applications that require increased sensitivity. Depending on the application, the combination of two or three of the assays can meet the needed specificity and sensitivity.

Cost estimates are not completed, but running costs are thought to be in line with maintaining a typical adherent cell line. BoCell™ can be completed in 24 to 96 hours, depending on the application while the MBA typically takes a week to complete.

As a complete suite of BoNT-detecting and quantifying assays, the BoTest™, BoTest Matrix™, and BoCell™ assays appear to be quite promising and deserving of ICCVAM-sponsored validation. These assays are capable of saving hundreds of thousands of mice from painful deaths while protecting food supplies, monitoring susceptible wildlife, and quantifying Botox batch potency.

ICCVAM must follow its Congressional mandate and implement the validation of this cost and life-saving suite of assays with all due speed, thus eliminating the use of mice for *C. botulinum* toxin detection and quantification. In planning this validation study, we request that ICCVAM use the available data showing BoTest™'s capabilities with respect to quantifying drug formulations and related products, as well as any data that the company has gathered regarding comparability between BoTest™ to the MBA, and apply these data sets towards validation efforts.

Thank you for your attention to these comments on validation of Botulinum neurotoxin activity assays and expanded validation of a MAT pyrogencity assay. We can be reached for questions at the contact information below.

Sincerely,

[Redacted]

Nancy Beck, Ph.D.
Science and Policy Adviser
Physicians Committee for Responsible Medicine
nbeck@pcrm.org
202-527-7345

[Redacted]

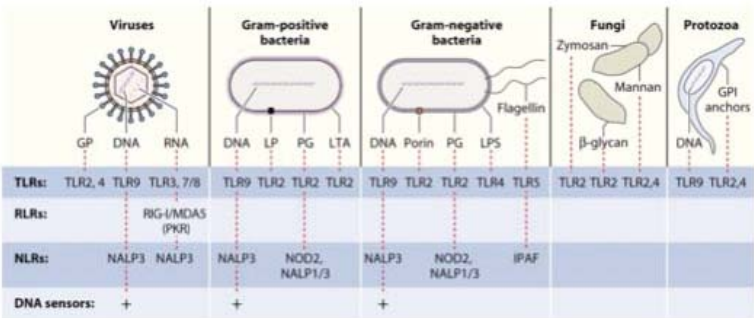
Samantha Dozier, Ph.D.
Policy Advisor, Medical Testing Issues
Regulatory Testing Division
People for the Ethical Treatment of Animals

Table 1: Recognition of Pyrogens by TLRs

Receptor	Cellular localization	Microbial component(s)	Origin(s)
TLRs			
TLR1/TLR2	Cell surface	Triacyl lipopeptides	Bacteria
TLR2/TLR6	Cell surface	Diacyl lipopeptides	Mycoplasma
		Lipoteichoic acid	Gram-positive bacteria
TLR2	Cell surface	Lipoproteins	Various pathogens
		Peptidoglycan	Gram-positive and -negative bacteria
		Lipoarabinomannan	Mycobacteria
		Porins	Neisseria
		Envelope glycoproteins	Viruses (e.g., measles virus, HSV, cytomegalovirus)
		GP1-mucin	Protozoa
		Phospholipomannan	Candida
		Zymosan	Fungi
		β -Glycan	Fungi
TLR3	Cell surface/endosomes	dsRNA	Viruses
TLR4	Cell surface	LPS	Gram-negative bacteria
		Envelope glycoproteins	Viruses (e.g., RSV)
		Glycoinositolphospholipids	Protozoa
		Mannan	Candida
		HSP70	Host
TLR5	Cell surface	Flagellin	Flagellated bacteria
TLR7/8	Endosome	ssRNA	RNA viruses
TLR9	Endosome	CpG DNA	Viruses, bacteria, protozoa

Mogensen, T. H.. 2009. Clin. Microbiol. Rev. 22(2):240-273

FIG. 1. Recognition of PAMPs from different classes of microbial pathogens



Mogensen, T. H.. 2009. Clin. Microbiol. Rev. 22(2):240-273



CertiChem, Inc.
11212 Metric Blvd. Suite 500
Austin, TX 78758

Dr. Lori White
NTP Office of Liaison, Policy and Review
NIEHS, P.O. Box 12233, MD K2-03
Research Triangle Park, NC 27709

9 June, 2011

Dear Dr. White:

After reviewing the BG1Luc ER TA Agonist and Antagonist Protocols, the scientists at CertiChem, Inc. believe this cell line has excellent utility. We fully support and appreciate the tremendous effort put forth by the ICCVAM/NICEATM team led by Dr. Stokes and Dr. Casey.

While this is a strong assay, we have recommendations that should allow the BG1Luc ER TA assay to be more versatile. We propose changing various aspects of the protocol to allow efficient implementation by a liquid handling system. While these changes should not be significant alterations to the protocol, they could greatly increase throughput. Specifically, some liquid handling systems used to automate (roboticize) assays are incapable of vertical serial dilutions in the Y-axis, so a plate design that can accommodate serial dilutions in the X-axis is preferred. Additionally, the option to use ICI 182,780, instead of Raloxifene HCl in some assays would offer further improvement. Requested changes for both the agonist/antagonist protocols are in four general areas:

- 1) Allow flexible plate template design and orientation
- 2) Allow increased number of replicates of weak positive (Methoxychlor/Tamoxifen) & negative controls
- 3) Allow increased number of concentrations for strong positive controls (estradiol/raloxifene): recommend to use 8 concentrations in Range-finder experiments and 12 concentrations in comprehensive experiments
- 4) - Allow serial dilution for Range-finder and Comprehensive experiments to be performed in 2% dimethylsulfoxide (DMSO) in estrogen-free medium (EFM) (for example, allow serial dilutions in 2% DMSO EFM to be overlaid on 100uL cell seeding volume to achieve final 200uL volume and 1% DMSO)

These recommendations are given more specifically in the tables below.

**Table entries refer to version 25 January 2011 DRAFT BG1Luc ER TA Performance Standards:
Appendix B (Agonist Protocol) by line numbers**

Document Section	Line Number	Current Protocol	Robotized Protocol	Recommended Modification
6.1	144-147	7-point serial dilution for Range Finder	8-point RF and serial dilutions that run horizontally, rather than vertically on the 96-well plate	Allow the use of a bigger range of testing concentration to help identify the right concentrations for strong ER agonists
8.2	248-257	Estrogen-Free DMEM Medium	Estrogen-free RPMI medium	Allow the use of any phenol red-free medium, especially since cells are maintained already in RPMI medium
8.2	248-257	4.5% charcoal/dextran treated FBS	Combination of charcoal/dextran treated FBS and calf serum	Allow the replacement or supplementation of medium with varying percentage of FBS and/or calf serum to decrease the background
9.2	489-501	200,000 cells/mL for seeding 40,000 cells per well in 200 uL	Seed 40,000 cells/well in 100 uL	Allow 100 uL seeding volume so that treatments may be overlaid without firstly removing the seeding medium
10.1	510-525	Solubility testing in 4 mL conical tubes	Solubility testing in various sized microcentrifuge tubes and glass vials	Allow the use of any appropriately sized microcentrifuge or glass vessel to be used for solubility testing
10.2	535-645	Preparations of controls and test substances and serial dilutions (for Range Finder and Comprehensive assays) are performed in glass tubes and with very specific volumes.	Serial dilutions are performed in 96-well Deep Well Plates that are amenable to roboticized procedure.	Allow serial dilutions in any type of vessel and with smaller or larger volumes, so long as DMSO is used and specified dilution is achieved.
11.4	741-746	Injecting luminometer is used	Luciferase assay reagent is dispensed by liquid handling system and not by luminometer injectors	Allow dispensing of luciferase assay reagent by liquid handling system to speed up the plate measurement.
12	912-917	7-point serial dilution for Range Finder	8-point RF and serial dilutions that run horizontally, rather than vertically on the 96-well plate	Allow the use of a bigger range of testing concentration to help identify the right concentrations for strong ER agonists



CertiChem, Inc.
11212 Metric Blvd. Suite 500
Austin, TX 78758

**Table entries refer to version 25 January 2011 DRAFT BG1Luc ER TA Performance Standards:
Appendix C (Antagonist Protocol) by line numbers**

Document Section	Line Number	Current Protocol	Robotized Protocol	Recommended Modification
	5.2 96-107	Fixed 17beta-estradiol (E2) testing concentration	E2 concentration is based on historical comprehensive testing for E2 and approximates E2 EC80.	Allow modification of E2 concentration, as-needed, since cellular responses may drift over time and it is necessary to achieve an appropriate level of E2 activity.
	5.2 96-107	Flavone is used (as is tamoxifen)	Use dibenzanthracene (DBA) -	Allow the use of tamoxifen as already allowed in changes made to the BG1 assay
	6.1 143-147	7-point serial dilution for Range Finder	8-point RF and serial dilutions that run horizontally, - rather than vertically on the 96-well plate -	Allow the use of a bigger range of testing concentration to help identify the right concentrations for strong ER antagonists
	8.2 244-254	Estrogen-Free DMEM Medium	Estrogen-free RPMI medium -	Allow the use of RPMI phenol red-free medium, especially since cells are maintained already in RPMI medium
	8.2 244-254	4.5% charcoal/dextran treated FBS	Combination of charcoal/dextran treated FBS and calf serum	Allow the replacement or supplementation of medium with varying percentage of FBS and/or calf serum to decrease the background luciferase activity
	9.2 472-479	200,000 cells/mL for seeding 40,000 cells per well in 200 uL	Seed 40,000 cells/well in 100 uL	Allow 100 uL seeding volume so that treatments may be overlaid without firstly removing the seeding medium
	10.1 492-505	Solubility testing in 4 mL conical tubes	Solubility testing in various sized microcentrifuge tubes and glass vials	Allow the use of any appropriately sized microcentrifuge or glass vessel to be used for solubility testing
	11.0 to 12.4 508-636	Preparations of controls and test substances and serial dilutions (for Range Finder and Comprehensive assays) are performed in glass tubes and with very specific volumes.	Serial dilutions are performed in 96-well Deep Well Plates that are amenable to robotized procedure.	Allow serial dilutions in 96-well Deep Well Plates and with smaller or larger volumes, so long as DMSO is used and specified dilution is achieved.
	14	7-point serial dilution for Range Finder	8-point RF and serial dilutions that run horizontally, rather than vertically on the 96-well plate	Allow the use of a bigger range of testing concentration to help identify the right concentrations for strong ER antagonists

Again, we appreciate all the work that the ICCVAM/NICEATM team is doing and we are excited about the - opportunity to provide our input in this review process. Please do not hesitate to ask for any clarification or - assistance. -

Sincerely, -

Matthew Stoner, PhD -
Senior Research Scientist, CertiChem, Inc. -
Matthew.Stoner@CertiChem.com -
512.339.0550

This page intentionally left blank

Appendix E6

76 FR 28781

**Independent Scientific Peer Review Panel Report: Evaluation of the Validation Status
of an *In Vitro* Estrogen Receptor Transcriptional Activation Test Method for Endocrine
Disruptor Chemical Screening: Notice of Availability and Request for Public
Comments**

(No public comments received)

This page intentionally left blank

DEPARTMENT OF HEALTH AND HUMAN SERVICES**Independent Scientific Peer Review Panel Report: Evaluation of the Validation Status of an In Vitro Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening: Notice of Availability and Request for Public Comments**

AGENCY: Division of the National Toxicology Program (DNTP), National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH).

ACTION: Notice of availability and request for comments.

SUMMARY: The NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), on behalf of the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), convened an independent international scientific peer review panel (hereafter, Panel) on March 29–30, 2011, to evaluate the validation status of the LUMI-CELL® (BG1Luc ER TA) test method, an *in vitro* transcriptional activation (TA) assay used to identify chemicals that can interact with human estrogen receptors (ERs). The Panel report is now available on the NICEATM–ICCVAM Web site at: http://iccvam.niehs.nih.gov/docs/endo_docs/EDPRPrept2011.pdf or by contacting NICEATM (see **ADDRESSES**). The report contains (1) the Panel's evaluation of the validation status of the test method and (2) the Panel's comments on the draft ICCVAM test method recommendations. NICEATM invites public comment on the Panel report.

DATES: Written comments on the Panel report should be received by July 5, 2011.

ADDRESSES: NICEATM prefers that comments be submitted electronically by e-mail to niceatm@niehs.nih.gov. Comments can also be submitted via the NICEATM–ICCVAM Web site at http://iccvam.niehs.nih.gov/contact/FR_pubcomment.htm. Written comments can be sent by mail or fax to Dr. Warren Casey, Deputy Director, NICEATM, NIEHS, P.O. Box 12233, Mail Stop: K2–16, Research Triangle Park, NC 27709; (fax) 919–541–0947. Courier address: NIEHS, NICEATM, 530 Davis Drive, Room 2035, Durham, NC 27713.

FOR FURTHER INFORMATION CONTACT: Dr. Warren Casey: (telephone) 919–316–4729, (fax) 919–541–0947, (e-mail) niceatm@niehs.nih.gov.

SUPPLEMENTARY INFORMATION:**Background**

In January 2011, NICEATM announced the convening of an independent scientific peer review panel to review and comment on the draft background review document (BRD) summarizing available data, reliability and accuracy of the BG1Luc ER TA test method, the draft recommendations, as well as the availability of the draft documents for public comment (76 FR 4113). The Panel met in public session on March 29–30, 2011, at the Natcher Conference Center in Bethesda, MD. The Panel reviewed the draft ICCVAM BRD for completeness, errors, and omissions of any existing relevant data or information. The Panel also evaluated the information in the draft documents to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM, 2003a) had been appropriately addressed. The Panel then considered the ICCVAM draft recommendations and commented on the extent that the recommendations were supported by the information provided in the draft BRD.

In January 2004, Xenobiotic Detection Systems, Inc. (XDS, Durham, NC) nominated their LUMI-CELL® BG1Luc ER TA test method for an interlaboratory validation study. This method uses BG–1 cells, a human ovarian carcinoma cell line that is stably transfected with an estrogen-responsive luciferase reporter gene to measure whether and to what extent a substance induces or inhibits TA activity via ER mediated pathways (Denison and Heath-Pagliuso, 1998). Included in the nomination package were test results from XDS for 56 of the 78 ICCVAM reference substances for agonist activity and 16 of the 78 ICCVAM reference substances for antagonist activity. These studies were funded primarily by an NIEHS Small Business Innovation Research (SBIR) grant (SBIR43ES010533–01).

In accordance with the ICCVAM nomination process, NICEATM conducted a preliminary evaluation of the nomination package to determine the extent to which it addressed the ICCVAM prioritization criteria and adherence to the ICCVAM recommendations for the standardization and validation of *in vitro* endocrine disruptor test methods (ICCVAM, 2003b). ICCVAM and the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) recommended that the BG1Luc ER TA test method should be

considered a high priority for interlaboratory studies based upon the lack of adequately validated test methods and the regulatory and public health need for such test methods. Based on this evaluation, ICCVAM recommended that:

- The BG1Luc ER TA test method should be considered a high priority for interlaboratory validation studies as an *in vitro* test method for the detection of test substances with ER agonist and antagonist activity.
- Validation studies should include coordination and collaboration with the European Centre for the Validation of Alternative Methods (ECVAM) and the Japanese Center for the Validation of Alternative Methods (JaCVAM) and include one laboratory in each of the three respective geographic regions (United States, Europe, and Japan).
- In preparation for the interlaboratory validation study, XDS should conduct protocol standardization studies with an emphasis on filling data gaps in the antagonist protocol for the BG1Luc ER TA.

The NIEHS subsequently agreed to support the validation study in light of its role as one of the three NTP agencies, whose mission includes the development and validation of improved testing methods. Based on the results of this study, ICCVAM is now reviewing the validation status of this test method for identification of substances with *in vitro* ER agonist or antagonist activity. NICEATM and the ICCVAM Interagency Endocrine Disruptors Working Group prepared a draft BRD that provides a comprehensive description and the data from the validation study used to assess the accuracy and reliability of the BG1Luc ER TA test method. ICCVAM also developed draft recommendations for its use.

Availability of the Peer Panel Report

The Panel's conclusions and recommendations are detailed in the *Independent Scientific Peer Review Panel Report: Evaluation of the Validation Status of the BG1Luc4E2 ER TA (LUMICELL), an In Vitro Transcriptional Activation Assay Used to Identify Chemicals That Can Interact with Human Estrogen Receptors* which is available along with the draft documents reviewed by the Panel and the draft ICCVAM test method recommendations at <http://iccvam.niehs.nih.gov/methods/endocrine/PeerPanel11.htm>.

Request for Public Comments

NICEATM invites the submission of written comments on the Panel report. When submitting written comments, please refer to this **Federal Register** notice and include appropriate contact information (name, affiliation, mailing address, phone, fax, e-mail, and sponsoring organization, if applicable). All comments received will be made publicly available via the NICEATM–ICCVAM Web site at <http://iccvam.niehs.nih.gov/methods/endocrine/PeerPanel11.htm>. ICCVAM will consider the Panel report along with public comments and comments made by SACATM at their June 16–17, 2011 meeting (67 FR 23323) when finalizing test method recommendations. Final ICCVAM recommendations will be published in an ICCVAM test method evaluation report, which will be forwarded to relevant Federal agencies for their consideration. The evaluation report will also be available to the public on the NICEATM–ICCVAM Web site at <http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-TMER.htm> and by request from NICEATM (see **ADDRESSES** above).

Background Information on ICCVAM, NICEATM, and SACATM

ICCVAM is an interagency committee composed of representatives from 15 Federal regulatory and research agencies that require, use, generate, or disseminate toxicological and safety testing information. ICCVAM conducts technical evaluations of new, revised, and alternative safety testing methods with regulatory applicability and promotes the scientific validation and regulatory acceptance of toxicological and safety testing test methods that more accurately assess the safety and hazards of chemicals and products and that refine (decrease or eliminate pain and distress), reduce, and replace animal use. The ICCVAM Authorization Act of 2000 (42 U.S.C. 2851–3) established ICCVAM as a permanent interagency committee of the NIEHS under NICEATM. NICEATM administers ICCVAM and provides scientific and operational support for ICCVAM-related activities and conducts independent validation studies to assess the usefulness and limitations of new, revised, and alternative test methods and strategies. NICEATM and ICCVAM welcome the public nomination of new, revised, and alternative test methods and strategies applicable to the needs of U.S. Federal agencies. Additional information about ICCVAM and NICEATM can be found on the

NICEATM–ICCVAM Web site (<http://iccvam.niehs.nih.gov>).

SACATM was established in response to the ICCVAM Authorization Act [Section 2851-3(d)] and is composed of scientists from the public and private sectors. SACATM advises ICCVAM, NICEATM, and the Director of the NIEHS and NTP regarding statutorily mandated duties of ICCVAM and activities of NICEATM. SACATM provides advice on priorities and activities related to the development, validation, scientific review, regulatory acceptance, implementation, and national and international harmonization of new, revised, and alternative toxicological test methods. Additional information about SACATM, including the charter, roster, and records of past meetings, can be found at <http://ntp.niehs.nih.gov/go/167>.

References

- Denison MS, Heath-Pagliuso S. 1998. The Ah receptor: A regulator of the biochemical and toxicological actions of structurally diverse chemicals. *Bull Environ Contam Toxicol* 61(5): 557–568.
- ICCVAM. 2003a. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03–4508. Research Triangle Park, NC.
- ICCVAM. 2003b. ICCVAM Evaluation of In Vitro Test Methods For Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays.

Dated: May 11, 2011.

John R. Bucher,
Associate Director, National Toxicology Program.

[FR Doc. 2011–12264 Filed 5–17–11; 8:45 am]

BILLING CODE 4140–01–P

Appendix E7

Scientific Advisory Committee on Alternative Toxicological Methods (SACATM)

Comments

SACATM Meeting on June 16 - 17, 2011

The following is excerpted from the minutes of the SACATM meeting convened on June 16-17, 2011. The full meeting minutes are available online at <http://ntp.niehs.nih.gov/go/8202>

This page intentionally left blank

Summary Minutes from the June 16-17, 2011 SACATM Meeting
Arlington Hilton, Arlington, VA

VI. Report on Peer Review Panel Meeting: Evaluation of an *In Vitro* Estrogen Receptor (ER) Transcriptional Activation (TA) Test Method for Endocrine Disruptor Chemical (EDC) Screening

Introduction and Overview of Proposed Methods and Applications: The BG1Luc ER TA (LUMI-CELL®) Test Method to Identify Substances with Estrogen Agonist and/or Antagonist Activity

NICEATM Deputy Director Dr. Warren Casey briefed SACATM on the proposed endocrine disruptor test method. The EPA has been mandated to develop a screening program to detect EDCs so it asked ICCVAM to evaluate existing validated *in vitro* EDC screening tests. ICCVAM found none, leading ICCVAM and SACATM to make validation of such a test a high priority. In response, there was a nomination from Xenobiotic Detection Systems (XDS) for its LUMI-CELL® assay, a luciferase reporter assay that detects estrogen-binding activity. The assay is based in human ovarian carcinoma (BG-1) cells, with endogenous ER-alpha and ER-beta. The test provides a concentration-response, and so can assess both potency and efficacy. There are nearly identical protocols for both agonists and antagonists. The agonist assay involves gain of function, while the antagonist test measures loss of function, both based upon luciferase levels.

Dr. Casey provided a timeline for the project, beginning in January 2004 with the nomination of the assay by XDS, through the public peer review meeting in Bethesda in March 2011. He reviewed the definition of validation and ICCVAM's validation criteria, as well as the four phases of the international validation study, which was sponsored by NICEATM-ICCVAM, JaCVAM, and ECVAM.

When the testing was completed, accuracy and reproducibility were assessed. The agonist test method was 97% accurate, had 96% sensitivity and 100% specificity. The antagonist method was 100% accurate, with 100% sensitivity and 100% specificity. The agonist method showed 100% intra-laboratory reproducibility of the substances tested independently three times. Inter-

Summary Minutes from the June 16-17, 2011 SACATM Meeting
Arlington Hilton, Arlington, VA

laboratory reproducibility was 81%. For the antagonist methods, intra-laboratory reproducibility was 100%, while inter-laboratory reproducibility was 89%. In a comparison of the BG1Luc ER TA with the ER binding assay, there was 97% concordance. Compared with the Chemical Evaluation and Research Institute (CERI) Stably Transfected Human Estrogen Receptor Transcriptional Activation (STTA) assay, overall there was 86% concordance using 26 reference substances. Based on the validation program, ICCVAM recommended the use of the BG1Luc ER TA as a screening test to identify substances with estrogen agonist and antagonist activity, with the highest test substance concentration limited to 10 μ M for the antagonist assay. ICCVAM also developed and released performance standards for the assays.

ICCVAM conducted a peer review panel meeting March 29-30, 2011, to consider the recommendations, performance standards, and background data. The panel consisted of 16 scientists from 6 countries. Following the SACATM meeting, the Endocrine Disruptor Working Group will consider SACATM comments and the panel report and finalize ICCVAM's test method evaluation report. Ultimately, in fall 2011, the ICCVAM recommendations will be forwarded to Federal agencies, and a draft test guideline will be forwarded to OECD.

Summary of the Independent Scientific Peer Review Panel Evaluation of the Validation Status of the LUMI-CELL ER® (BG1Luc ER TA) Test Method

Dr. John Vandenberg of North Carolina State University (retired), who chaired the Peer Review Panel ("the Panel"), briefed SACATM on the meeting.

He reviewed ICCVAM's charges to the Panel and its recommendations. The Panel agreed with ICCVAM that the BG1Luc ER TA could be used as a screening tool to identify substances with *in vitro* estrogen agonist and antagonist activity. It considered the test method protocol to be complete and adequate in detail, and agreed with ICCVAM about the needs for future studies. The Panel also suggested that such future studies could address metabolic activation, that the reference substance list and associated database could be expanded with additional negative agonist and positive antagonist substances as they are identified, and that efforts could be made to identify a quantitative cytotoxic method. It also concurred with the draft ICCVAM performance standards and some modifications to expand applicability of the performance standard.

Public Comments

Dr. Niemi called for public comments and noted written comments had been submitted from CertiChem, Inc.

Dr. Catherine Willett, Associate Director of Regulatory Testing for People for the Ethical Treatment of Animals (PETA), reported that PETA lauded the Panel and supported the recommendations, both the main finding recommending the test method and the other recommendations. She congratulated the Panel on its review, saying it was "an incredibly thorough, well-done, well-reviewed validation study." She listed several panel recommendations that PETA supported: (1) designation of the assay as an alternative for the CERI STTA assay and the rat uterine cytosol assay, (2) development and validation of ER

Summary Minutes from the June 16-17, 2011 SACATM Meeting Arlington Hilton, Arlington, VA

binding assays using recombinant receptors for both humans and other animals, (3) development and use of a metabolism component, (4) inclusion of potency evaluations to quantify activity, (5) evaluation of the quality of the data used to classify the original ICCVAM reference substances (6) discussion of the use of assay, and (7) discussion of the animal reduction potential.

She conveyed several additional PETA recommendations: (1) revise the chemical list to follow up on the evaluation and updating of the chemical reference list, and adding the new information to a publicly searchable database; (2) ensure that the best characterized chemicals are used for future assay evaluations; (3) identify new reference chemicals in underrepresented chemical classes; (4) consider the use of the assay to reduce animal testing, such as its use in addition to screening and prioritization, revising the structure of the Endocrine Disruptor Screening Program (EDSP) Tier 1 assessment by performing *in vitro* assays prior to animal testing, and the adoption of a weight-of-evidence approach that could be used to further reduce or eliminate estrogen receptor-related animal tests; and (5) evaluate the data quantitatively using a Relative Potency Index relative to a standard reference chemical, to allow quantitative comparison to the CERI STTA and to other assays

She noted the study had taken 7 years to complete, and so was not included in Phase I of the EDSP. She said a more efficient process is needed in light of the large number of new assays emerging. She recommended the Panel note issues that contributed to the length of the review in its report, and include recommendations for avoiding those issues in future reviews.

Dr. Niemi recognized Dr. Fowle, who was at that point prepared to respond to Dr. Hansen's request regarding data on adoption of alternative test methods.

Dr. Fowle said data had last been collected August 26, 2010, regarding 12 assays, which were grouped from the larger assay population: LLNA: 241, Corrositex: 0, Up and Down Assay: 1,139, EpiSkin/EpiDerm: 2, BCOP: 14, ICE: 0, *In Vitro* pyrogen tests: 0, Cytosensor: 0, EpiOcular: 3, LumiCell: 0, CertiChem: 0, Total: 1,399.

He mentioned that those figures may make it appear that EPA and others are not committed to reducing, refining and replacing animal use, and asked that he be allowed to comment at some point about some of the things EPA is doing to achieve the 3Rs. Dr. Niemi asked Dr. Fowle to hold those comments for later in the meeting.

SACATM Discussion

Dr. Corcoran, lead discussant, said the EDC method evaluation seemed to be "a tour de force," and commended the work of the Panel. He said he would like more information about the quality of data issue that had been commented upon in the Panel's report, specifically the criteria involving ranking and sensitivity analysis, or tests for trends in terms of the criteria for evaluating positive and negative compounds. He asked Dr. Vandenberg to comment on whether the Panel was proposing a higher and new standard for all assays of this nature. Dr. Vandenberg said it would be presumptuous for the Panel to do so, in terms of attempting to direct what other panels might do. On the other hand, he said, it would be fine for other panels

Summary Minutes from the June 16-17, 2011 SACATM Meeting
Arlington Hilton, Arlington, VA

to adopt the standards described by this one. Dr. Corcoran asked for clarification on the Panel's conclusion that there were insufficient data to term the evaluation a "thorough" analysis, although it was termed as "adequate." This was, he said, due to the use of a descriptive versus a formal, inferential assessment of the data. Dr. Vandenberg said it was hard for the Panel to consider the analysis to be thorough, since there would always be things that had not been thought of. Thus, their description of the analysis was adequate. Statistically, he said the analysis of the data was considered to be adequate, with no fault found. Dr. Casey added it was always difficult to get statisticians to agree on anything, so some of the comments pointed to ways things could have been done differently, statistically, particularly EC₅₀ calculations. Dr. Corcoran said he had been hoping to hear that ICCVAM was moving toward a new standard for quality of data.

Dr. Corcoran added he would like to have seen more information in the document on the implications of the assay for use in Europe and Japan. Dr. Vandenberg said that was not specifically discussed as it related to the background document, but it did come up during the discussion, and there were foreign representatives present who brought some of those issues. Dr. Corcoran said he would like to have seen validation conducted in one set of known agonists and antagonists, and then movement into a second set of yet-untested agonists and antagonists, thus incorporating a two-step process. He recognized it had already been a 7-year, \$3 million process, but nonetheless objected to validation based on only one set of compounds. Dr. Casey said every positive and every negative they could find had been tested, but the chemical space was very small for well-referenced compounds; just 38 compounds fit the criteria. Dr. Corcoran maintained since the protocol was changed over the course of the 7 years, having two sets of data would have helped, even if it involved splitting up the known compounds. Despite his comments, Dr. Corcoran said the review was "a very impressive body of work."

Dr. Elmore, lead discussant, agreed with the previous comments, as well as the conclusions and recommendations contained in the report. He felt, however, the BG1 cell line needs to be better characterized. He recommended the cell line be placed in a repository to ensure access and availability in the future.

Dr. Meyer, lead discussant, was also impressed with the work of the Panel, calling it "very comprehensive and very clear." She strongly supported the idea that cytotoxic changes be quantified. She noted that although validation normally means the replacement of an *in vivo* method with an *in vitro* method, in this case, an *in vitro* method is to be replaced by another *in vitro* method. She questioned the priority of whether ICCVAM should be funding such an effort, given the large number of animals still being used in other areas. Dr. Meyer noted the introduction of the non-radioactive LLNAs would actually replace animal use, but that the EDC assay is a screening method, and that she was uncomfortable with expending too many resources on such an approach. She wondered whether the current method could not be further developed to work on antagonists. She also asked about harmonization for *in vitro* methods. She mentioned it would be helpful to have a formula in the document on how the fold-reduction was calculated and commented on a lack of clarity for expressing the performance standard.

Summary Minutes from the June 16-17, 2011 SACATM Meeting
Arlington Hilton, Arlington, VA

Regarding priority, Dr. Stokes said the developer nominated the method for validation studies in 2005, and at that time it was given a very high priority by SACATM. He said in the case of a positive, that such information could be used along with other mechanistic data to move forward with characterizing whether or not the compound is in fact an *in vivo* endocrine disruptor. Dr. Stokes said regarding the comparison with the current method that has been adopted by the EPA and is in their guidelines, this was done because the adoption had occurred after the validation study was initiated. He said, "But we didn't even know about that method in 2005, that it even existed, but it was moving along and a couple of years later, yes, we did find out that it was going through validation as well. This study was nominated and a validation study was initiated before there was any knowledge of the other method."

Dr. Fowle said in terms of maximizing the utility of tests, clearly things have evolved, and some of the earlier screens that were developed for validation occurred a number of years ago. He said it's really important, if these screens get used, that they get linked very closely in terms of working with the regulatory agencies and the users who'll be using them, to make sure these assays will be used, and will be used for purposes which will help advance the mission. He said Dr. Meyer raised some very good points in terms of the resources available. ICCVAM focuses on validation of alternative methods to animal tests, and he thinks it's very important to focus on replacements for animal tests. EPA's policy and approach for using the EDSP is such that it probably will not be using this assay. He said he thought it just sort of underlines the importance of having very close communications at the beginning, middle, and end. He alluded to the history of EPA discussions with Drs. Stokes, Bucher, and Birnbaum as they tried to build on the lessons learned to try to do a better job in the future. He suggested having a retreat or similar meeting to look at the good things ICCVAM has done, see what might be improved, and figure out how to move forward. Dr. Birnbaum agreed with Dr. Fowle, but reminded everyone that the purpose of some *in vitro* tests is to answer a very specific question. She said this test determines whether a substance is an agonist or antagonist for ER α and ER β , but there are other ways that chemicals can be endocrine disruptors, e.g., of the estrogen signaling system, and this test is not identifying them.

Dr. Wilson, lead discussant, concurred with previous comments, as well as the need for a follow-up meeting with ICCVAM to focus on trying to determine an overview of the various assays currently in use. He noted to run an assay is as much an art as a science, and that it should be moved more toward the science. So a focused discussion with experts to understand the limitations of the current assays and see whether any stand out would be helpful to further the state of the science. For the EDC assay, he agreed with Dr. Elmore regarding better characterization of the cell line. He cautioned that use of the phrase "endocrine disruptor" carries an obvious stigma, and suggested a careful definition of what is or is not an endocrine disruptor be put into the background information of the document.

Dr. Casey noted the figures regarding accuracy, reliability, and reproducibility had been approached in a thoughtful manner, with choices having been made among potential approaches. Dr. Meyer suggested revising the specific section she had earlier referred to as problematic. Regarding usability of the assay in high throughput screening, Dr. Casey said it was currently being evaluated at NCGC, and that it works well in a 384-well format, but it may

Summary Minutes from the June 16-17, 2011 SACATM Meeting
Arlington Hilton, Arlington, VA

not provide adequate signal-to-noise in a 1536-well format. He continued by noting the cells for the assay are co-owned by XDS and Dr. Michael Denison at the University of California, Davis, and that Dr. Denison was reluctant to put the cells into a repository because he wishes to maintain control of them. He does make them freely available to academic and government labs through a formal licensing process. Dr. Vandenberg said the Panel had discussed the issue of cell line availability at length, and did all they could to ensure access to the cell line. Dr. Toth expressed concern about drift in the cell line over time, asking whether there are quality control measures to ensure such drift would not take place. Dr. Casey said positive and negative controls are run with each test, but that currently there is not a way to track the genetic stability of the line. Dr. Stokes said all of the *in vitro* assays use acceptance criteria for the positive controls, so there must be a response within that acceptance range. Thus, if the cells have changed and the response has been decreased to below that threshold for an acceptable positive control response, or if it exceeds the upper limit of it, it would not be a good run and it would indicate that perhaps the cells had changed, become contaminated, or were the wrong cells.

Regarding the history of the assay, Dr. Stokes noted the EDSP was mandated by laws in 1996. The LUMI-CELL ER[®] was developed in response to a Small Business Innovative Research (SBIR) topic issued by NIEHS in the late 1990s in response to considerable interest at the time. The SBIR grant to develop the EDC method was supported by NIEHS and NIH grant funds. Dr. Birnbaum added that since NIH supported the development of cell lines, they should be fully available. Relevant to agencies' involvement, Dr. Stokes said there is an Endocrine Disruptor Working Group that includes representatives from all of the ICCVAM agencies. Dr. Stokes said the working group had EPA representatives on it who were kept abreast of the study design, chemical selection, and protocols, which were all run by that group before this testing went forward. He clarified that all of the agencies in ICCVAM had the opportunity for input into this validation study. New members have been integrated into ICCVAM and SACATM, and the work of the previous members may have been forgotten. He said NICEATM-ICCVAM is trying to make sure as much information as possible is reflected in the final evaluation reports that go out to the agencies and to the public.

Appendix F

Relevant Endocrine Disruptor Regulations and Testing Guidelines

F1	Table of Relevant Endocrine Disruptor Test Regulations	F-3
F2	EPA Endocrine Disruptor Screening Program Test Guidelines OPPTS 890.1300: Estrogen Receptor Transcriptional Activation (Human Cell Line (HeLa-9903)) (October 2009).....	F-7
F3	OECD Test Guideline 455: Stably Transfected Human Estrogen Receptor- α Transcriptional Activation Assay for Detection of Estrogenic Agonist-Activity of Chemicals (Adopted September 2009).....	F-33
F4	Weight of Evidence Guidance: Evaluating Results of EDSP Tier 1 Screening to Identify Candidate Chemicals for Tier 2 Testing (Draft for Public Comment)	F-51

This page intentionally left blank

Appendix F1

Table of Relevant Endocrine Disruptor Test Regulations

Note to the Reader:

Regulations may be updated in the future. It is recommended that users review the most current version of all regulations identified.

Electronic versions of United States Code (U.S.C.) can be obtained at:
<http://www.gpoaccess.gov/uscode/index.html>

Electronic versions of the Code of Federal Regulations (CFR) can be obtained at:
<http://www.gpoaccess.gov/cfr/index.html>

This page intentionally left blank

Endocrine Disruption Testing: Relevant US Federal Laws, Regulations, Guidelines, and Recommendations				
Agency, Center, or Office	Regulated Products	Statutory Safety Requirements	Regulations	Guidelines and Recommendations
FDA/CDER	Pharmaceuticals	Federal Food, Drug, and Cosmetic Act (U.S.C. Title 21, Chapter 9) Public Health Service Act (U.S.C. Title 42, Chapter 6A)	None Specific to Endocrine Disruption	No Specific Guidelines, Guidances, or Recommendations
EPA/OPPTS	Chemicals as defined by Section 5 of the Act Pesticides Deleterious substances added to food	Toxic Substances Control Act (U.S.C. Title 15, Chapter 53) Federal Insecticide, Fungicide, and Rodenticide Act (U.S.C. Title 7, Chapter 6) Federal Food, Drug, and Cosmetic Act (U.S.C. Title 21, 346a)	None Specific to Endocrine Disruption	OPPTS 890.1300 (2009)
CPSC	Consumer Products	Federal Hazardous Substances Act (U.S.C. Title 15, Chapters 1261- 1278)	None Specific to Endocrine Disruption	No Specific Guidelines, Guidances, or Recommendations
OSHA	Chemicals	Occupational Safety and Health Act of 1970 (U.S.C. Title 29, Chapter 15)	None Specific to Endocrine Disruption	No Specific Guidelines, Guidances, or Recommendations

Relevant Endocrine Disruption Regulations and Guidelines Europe		
Agency, Center, or Office	Regulated Products	Regulations and Directives
EU	Substances and Mixtures	No Specific Regulations or Directives
	Pesticides	No Specific Regulations or Directives
	Chemicals	No Specific Regulations or Directives
Relevant Endocrine Disruption Regulations and Guidelines International		
Organizations	Regulated Products	Guidelines, Guidance, and Recommendations
GHS	Chemicals	No Specific Guidelines, Guidances, or Recommendations
ISO	Medical Devices	No Specific Guidelines, Guidances, or Recommendations
OECD	Chemicals	OECD Test Guideline 455 (2009)
ICH	Pharmaceuticals	No Specific Guidelines, Guidances, or Recommendations

Abbreviations: CDER = Center for Drug Evaluation and Research; CFR = Code of Federal Regulations; CPSC = U.S. Consumer Product Safety Commission; EC = European Community; EEC = European Economic Community; EPA = U.S. Environmental Protection Agency; EU = European Union; FDA = U.S. Food and Drug Administration; GHS = Globally Harmonized System of Classification and Labelling of Chemicals; ICH = International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use; ISO = International Organization for Standardization; NA = not available; OECD = Organisation for Economic Co-operation and Development; OPPTS = Office of Prevention, Pesticides and Toxic Substances; OSHA = U.S. Occupational Safety and Health Administration; US = United States; U.S.C. = United States Code

Appendix F2

EPA Endocrine Disruptor Screening Program Test Guidelines OPPTS 890.1300: Estrogen Receptor Transcriptional Activation (Human Cell Line (HeLa-9903)) (October 2009)

This page intentionally left blank



United States
Environmental Protection
Agency

Prevention, Pesticides
and Toxic Substances
(7101)

EPA 740-C-09-006
October 2009

Endocrine Disruptor Screening Program Test Guidelines

OPPTS 890.1300:
Estrogen Receptor
Transcriptional Activation
(Human Cell Line (HeLa-
9903))

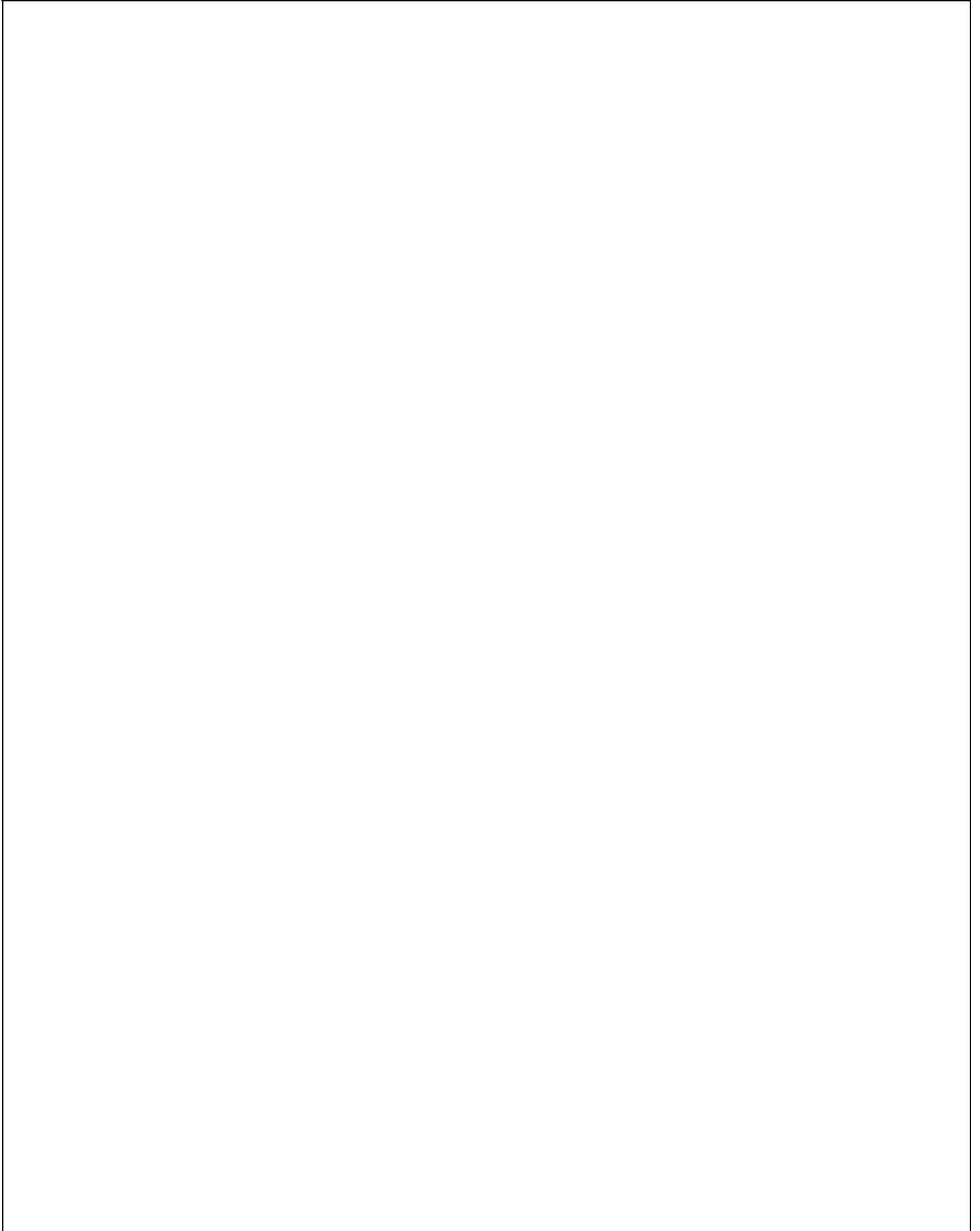


NOTICE

This guideline is one of a series of test guidelines established by the Office of Prevention, Pesticides and Toxic Substances (OPPTS), United States Environmental Protection Agency for use in testing pesticides and chemical substances to develop data for submission to the Agency under the Toxic Substances Control Act (TSCA) (15 U.S.C. 2601, *et seq.*), the Federal Insecticide, Fungicide and Rodenticide Act (FIFRA) (7 U.S.C. 136, *et seq.*), and section 408 of the Federal Food, Drug and Cosmetic (FFDCA) (21 U.S.C. 346a).

The OPPTS test guidelines serve as a compendium of accepted scientific methodologies and protocols that are intended to provide data to inform regulatory decisions under TSCA, FIFRA, and/or FFDCA. This document provides guidance for conducting the test, and is also used by EPA, the public, and the companies that are subject to data submission requirements under TSCA, FIFRA and/or the FFDCA. As a guidance document, these guidelines are not binding on either EPA or any outside parties, and the EPA may depart from the guidelines where circumstances warrant and without prior notice. The procedures contained in this guideline are strongly recommended for generating the data that are the subject of the guideline, but EPA recognizes that departures may be appropriate in specific situations. You may propose alternatives to the recommendations described in these guidelines, and the Agency will assess them for appropriateness on a case-by-case basis.

For additional information about OPPTS harmonized test guidelines and to access the guidelines electronically, please go to <http://www.epa.gov/oppts> and select "Test Methods & Guidelines" on the left side navigation menu. You may also access the guidelines in <http://www.regulations.gov> grouped by Series under Docket ID #s: EPA-HQ-OPPT-2009-0150 through EPA-HQ-OPPT-2009-0159, and EPA-HQ-OPPT-2009-0576.



OPPTS 890.1300: Estrogen Receptor Transcriptional Activation (Human Cell Line (HeLa-9903))

(a) Scope.

- (1) **Applicability.** This guideline is intended to meet testing requirements of the Toxic Substances Control Act (TSCA) (15 U.S.C. 2601, *et seq.*), the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) (7 U.S.C. 136, *et seq.*), and the Federal Food, Drug, and Cosmetic Act (FFDCA) (21 U.S.C. 346a).
- (2) **Background.** The Endocrine Disruptor Screening Program (EDSP) reflects a two-tiered approach to implement the statutory testing requirements of FFDCA section 408(p) (21 U.S.C. 346a). In general, EPA intends to use the data collected under the EDSP, along with other information, to determine if a pesticide chemical, or other substances, may pose a risk to human health or the environment due to disruption of the endocrine system.

This test guideline is intended to be used in conjunction with other guidelines in the OPPTS 890 series that make up the full screening battery under the EDSP to identify substances that have the potential to interact with the estrogen, androgen, or thyroid hormone (Tier 1 “screening”). The determination will be made on a weight-of-evidence basis taking into account data from the Tier 1 assays and other scientifically relevant information available. The fact that a substance may interact with a hormone system, however, does not mean that when the substance is used, it will cause adverse effects in humans or ecological systems.

Chemicals that go through Tier 1 screening and are found to have the potential to interact with the estrogen, androgen, or thyroid hormone systems will proceed to the next stage of the EDSP where EPA will determine which, if any, of the Tier 2 tests are necessary based on the available data. Tier 2 testing is designed to identify any adverse endocrine-related effects caused by the substance, and establish a quantitative relationship between the dose and that endocrine effect.

- (3) **Source.** OPPTS developed this guideline through a process of harmonization with the testing guidance and requirements published by the Organization for Economic Cooperation and Development (OECD) (Ref. 16).
- (b) Purpose.** *In vitro* transcriptional activation (TA) assays are based upon the production of a reporter gene product induced by a chemical, following binding of the chemical to a specific receptor and subsequent downstream transcriptional activation. TA assays using activation of reporter genes are screening assays

that have long been used to evaluate the specific gene expression regulated by specific nuclear receptors, such as the estrogen receptors (ERs) (**Refs. 3, 4, 5 & 6**). They have been proposed for the detection of estrogenic transactivation regulated by the ER (**Refs. 7, 8, & 9**). The nuclear ERs exist as at least two subtypes, termed α and β , encoded by distinct genes and with different tissue distribution, relative ligand binding affinities and biological functions. Nuclear ER α mediates the classic estrogenic response, therefore models currently being developed to measure ER activation mainly relate to ER α . The aim of this TA assay is to evaluate the ability of a chemical to function as an ER α ligand and activate an agonist response, for screening and prioritization purposes but can also provide mechanistic information that can be used in a weight of evidence approach. The definitions and abbreviations used in this Test Guidelines are described in Appendix 1.

- (c) **Initial Considerations and Limitations.** Estrogen agonists act as ligands for ERs, and may activate the transcription of estrogen responsive genes. This interaction may have the potential to trigger adverse health effects by disrupting estrogen-regulated systems. This Test Guideline describes an assay that evaluates TA mediated by the hER α . This process is considered to be one of the key mechanisms of possible endocrine disruption related health hazards, although there are also other important endocrine disruption mechanisms. These include:

- ☐ Actions mediated via other nuclear receptors linked to the endocrine system and interactions with steroidogenic enzymes
- ☐ Metabolic activation or deactivation of hormones
- ☐ Distribution of hormones to target tissues
- ☐ Clearance of hormones from the body

This Test Guideline exclusively addresses TA of an estrogen-regulated reporter gene by agonist binding to the hER α , and therefore it should not be directly extrapolated to the complex *in vivo* situation of estrogen regulation of cellular processes. Furthermore, this Test Guideline does not address antagonist interaction with the hER α and subsequent effect on transcription.

This test method is specifically designed to detect hER α -mediated TA by measuring chemiluminescence as the endpoint. However, non-receptor-mediated luminescence signals have been reported at phytoestrogen concentrations higher than 1 μ M due to the over-activation of the luciferase reporter gene (**Refs. 10 & 11**). While the dose response curve indicates that true activation of the ER system occurs at lower concentrations, luciferase expression obtained at high concentrations of phytoestrogens or similar compounds suspected of producing phytoestrogen-like over-activation of the luciferase reporter gene needs to be examined carefully in stably transfected ER TA assay systems (Appendix 2).

- (d) **Principle of the Test.** The TA assay using a reporter gene technique is an *in vitro* tool that provides mechanistic data. The assay is used to signal binding of the estrogen receptor with a ligand. Following ligand binding, the receptor-ligand complex translocates to the nucleus where it binds specific DNA response elements and transactivates a firefly 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. Luciferase activity can be evaluated quickly and inexpensively with a number of commercially available test kits.

The test system provided in this guideline utilizes the hER α -HeLa-9903 cell line, which is derived from a human cervical tumor, with two stably inserted constructs:

- ☐ The hER α expression construct (encoding the full-length human receptor).
- ☐ A firefly luciferase reporter construct bearing five tandem repeats of a vitellogenin Estrogen-Responsive Element (ERE) driven by a mouse metallothionein (MT) promoter TATA element.

The mouse MT TATA gene construct has been shown to have the best performance, and so is commonly used. Consequently this hER α -HeLa-9903 cell line can measure the ability of a test chemical to induce hER α -mediated transactivation of luciferase gene expression.

Data interpretation for this assay is based upon whether or not the maximum response level induced by a test chemical equals or exceeds an agonist response equal to 10% of that induced by a maximally inducing (1 nM) concentration of the positive control (PC) 17 β estradiol (E2) (*i.e.*, the PC10). Data analysis and interpretation are discussed in greater detail in section (f)(1) through (f)(3).

- (e) **Procedure.**

- (1) **Cell Line.** Use the stably transfected hER α -HeLa-9903 cell line for the assay. The cell line can be obtained from the Japanese Collection of Research Bioresources (JCRB) Cell Bank¹.

Use only cells characterized as mycoplasma-free in testing. RT PCR (Real Time Polymerase Chain Reaction) is the method of choice for a sensitive detection of mycoplasma infection (**Refs. 12, 13 & 14**).

¹ JCRB Cell Bank : National Institute of Biomedical Innovation, 7-6-8 Asagi Saito, Ibaraki-shi, Osaka 567-0085, Japan Fax: +81-72-641-9812

- (2) **Stability of the Cell Line.** To monitor the stability of the cell line, use E2, 17 α -estradiol, 17 α -methyltestosterone, and corticosterone as the reference chemicals, and include a complete concentration response curve in the test concentration range provided in Table 1 at least once each time the assay is performed. Comparative results to those provided in Table 1 are recommended.
- (3) **Cell Culture and Plating Conditions.** Maintain cells in Eagle's Minimum Essential Medium (EMEM) without phenol red, supplemented with 60 mg/L of antibiotic Kanamycine and 10% dextran-coated-charcoal-treated fetal bovine serum (DCC-FBS), in a CO₂ incubator (5% CO₂) at 37 \pm 1°C. Upon reaching 75-90% confluency, cells can be subcultured at 10 mL of 0.4 \times 10⁵ – 1 \times 10⁵ cells/mL for 100 mm cell culture dish. Suspend cells with 10% FBS-EMEM (which is the same as EMEM with DCC-FBS) and then plate into wells of a microplate at a density of 1 \times 10⁴ cells/100 μ L/well. Next, pre-incubate the cells in a 5% CO₂ incubator at 37 \pm 1°C for 3 hours before the chemical exposure. Use plastic-ware free of estrogenic activity.

To maintain the integrity of the response, grow the cells for more than one passage from the frozen stock in the conditioned media and do not culture them for more than 40 passages. For the hER α -HeLa-9903 cell line, this will be less than three months.

The DCC-FBS can be prepared as described in Appendix 3, or obtained from commercial sources.

- (4) **Acceptability Criteria.**
- (i) **Positive and Negative Reference Chemicals.** Prior to and during the study, verify the responsiveness of the test system using the appropriate concentrations of a strong estrogen: E2, a weak estrogen (17 α -estradiol), a very weak agonist (17 α -methyltestosterone) and a negative compound (corticosterone). Acceptable range values derived from the validation study are given in Table 1 (**Ref. 2**). Include these 4 concurrent reference chemicals with each experiment. It is recommended that the results fall within the given limits. If this is not the case, it is suggested that the reason for the failure to meet the acceptability criteria be determined (*e.g.*, cell handling, and serum and antibiotics for quality and concentration) and the assay repeated. Obtaining values within the recommended range will help ensure minimum variability of EC₅₀, PC₅₀ and PC₁₀ values. Consistent use of materials for cell culturing is also essential. The four concurrent reference chemicals, which are included in each experiment (conducted under the same conditions including the materials, passage level of cells and technicians), ensure the sensitivity of the

assay when the PC₁₀s of the three positive reference chemicals, and the PC₅₀s and EC₅₀s (when they can be calculated (see Table 1)) fall within the recommended ranges.

Table 1. Acceptable Range Values of the 4 Reference Chemicals for the STTA Assay (means \pm 2 standard deviations).

Name	logPC ₅₀	logPC ₁₀	logEC ₅₀	Hill slope	Test range
17 β -Estradiol (E2) CAS No: 50-28-2	-11.4 ~ -10.1	<-11	-11.3 ~ -10.1	0.7 ~ 1.5	10 ⁻¹⁴ ~ 10 ⁻⁸ M
17 α -Estradiol CAS No: 57-91-0	-9.6 ~ -8.1	-10.7 ~ -9.3	-9.6 ~ -8.4	0.9 ~ 2.0	10 ⁻¹² ~ 10 ⁻⁶ M
Corticosterone CAS No: 50-22-6	—	—	—	—	10 ⁻¹⁰ ~ 10 ⁻⁴ M
17 α -Methyltestosterone CAS No: 58-18-4	-6.0 ~ -5.1	-8.0 ~ -6.2	—	—	10 ⁻¹¹ ~ 10 ⁻⁵ M

- (ii) **Positive and Vehicle Controls.** Test the positive control (PC) (1 nM of E2) at least in triplicate in each plate. Test the vehicle that is used to dissolve a test chemical as a vehicle control (VC) at least in triplicate in each plate. If the PC uses a different vehicle than the test chemical, include another vehicle control at least in triplicate on the same plate with the PC in addition to this original vehicle control.
- (iii) **Fold-induction.** The target mean luciferase activity of the PC (1 nM E2) is at least 4-fold that of the mean vehicle control on each plate. This criterion is established based on the reliability of the endpoint values from the validation study (historically between four- and 30-fold).

With respect to the quality control of the assay, the target fold-induction corresponding to the PC₁₀ value of the concurrent PC (1 nM E2) is to be greater than 1+2SD (standard deviations) of the fold-induction value (=1) of the concurrent VC. For prioritization purposes, the PC₁₀ value can be useful to simplify the data analysis required compared to a statistical analysis. Although a statistical analysis provides information on significance, such an analysis is not a quantitative parameter with respect to concentration-based potential, and so is less useful for prioritization purposes.

- (5) **Chemicals to Demonstrate Laboratory Proficiency.** Prior to testing unknown chemicals in the STTA assay, confirm the responsiveness of the test system by each laboratory, at least once for each newly prepared batch of cell stocks taken from the frozen stock by independent testing of the 10 proficiency chemicals listed in Table 2. Perform this at least in

duplicate, on different days, and compare the results to Table 2. Please justify any deviations.

Table 2. List of Proficiency Chemicals.

Compound	CAS No.	Class ²	Test concentration range	Note
Diethylstilbestrol (DES)	56-53-1	Positive	10 ⁻¹⁴ - 10 ⁻⁸ M	
17 α -Ethinyl estradiol (EE)	57-63-6	Positive	10 ⁻¹⁴ - 10 ⁻⁸ M	
Hexestrol	84-16-2	Positive	10 ⁻¹³ - 10 ⁻⁷ M	
Genistein	446-72-0	Positive	10 ⁻¹² - 10 ⁻⁵ M	Cytotoxic at (0.01) ⁴ , 0.1 and 1 mM
Estrone	53-16-7	Positive	10 ⁻¹² - 10 ⁻⁶ M	
Butyl paraben	94-26-8	Positive	10 ⁻¹¹ - 10 ⁻⁴ M	Cytotoxic at (0.1) ⁴ and 1 mM
1,3,5-Tris(4hydroxyphenyl)benzene ¹	15797-52-1	Positive	10 ⁻¹² - 10 ⁻⁵ M	Cytotoxic at 100 μ M. PCmax approx 15% of PC Binds to hER α and has ER antagonist activity
Dibutyl phthalate (DBP)	84-74-2	Negative ³	10 ⁻¹¹ - 10 ⁻⁴ M	Cytotoxic at 1 mM
Atrazine	1912-24-9	Negative	10 ⁻¹¹ - 10 ⁻⁴ M	Cytotoxic ⁴ at 1 mM
Corticosterone	50-22-6	Negative	10 ⁻¹⁰ - 10 ⁻⁴ M	If not cytotoxic at 1 mM, then that is to be the highest tested concentration

¹Compound selected to challenge solubility and cytotoxicity.

²See Table 5 for definitions of positive and negative.

³Negative for ER α mediated transcriptional activation but may not be negative for non-ER β mediated transcriptional activation. Thus a positive result in this assay with DBP would indicate that the system is detecting other than pure ER α mediated activity and is therefore unacceptable.

⁴Cytotoxicity is close to 80%.

- (6) **Vehicle.** Use dimethyl sulfoxide (DMSO), or appropriate solvent, at the same concentration used for the different positive and negative controls and the test chemicals as the concurrent vehicle control. Dissolve each test substance in a solvent that solubilizes that test substance and is miscible with the cell medium. Water, ethanol (95% to 100% purity) and DMSO are suitable vehicles. If DMSO is used, do not exceed 0.1% (v/v). For any vehicle, demonstrate that the maximum volume used is not cytotoxic and does not interfere with assay performance.
- (7) **Preparation of Test Chemicals.** Generally, dissolve the test chemicals in DMSO or other suitable solvent, and serially dilute with the same solvent at a common ratio of 1:10 in order to prepare solutions for dilution with media.
- (8) **Solubility and Cytotoxicity: Considerations for Range Finding.** Conduct a preliminary test to determine the appropriate concentration range of chemical to be tested, and to ascertain whether the test chemical

may have any solubility and cytotoxicity problems. Initially, chemicals are tested up to the maximum concentration of 1 µl/ml, 1 mg/ml, or 1 mM, whichever is the lowest. Based on the extent of cytotoxicity or lack of solubility observed in the preliminary test, perform the first definite run for the test chemical at log serial dilutions starting at the maximum acceptable concentration (e.g., 1 mM, 100 µM, 10 µM, etc.). Note the presence of cloudiness, precipitate or cytotoxicity. Adjust concentrations in the second, and if necessary third run as appropriate to better characterize the concentration-response curve and to avoid concentrations which are found to be insoluble or to induce excessive cytotoxicity.

For ER agonists, the presence of increasing levels of cytotoxicity can significantly alter or eliminate the typical sigmoidal response and are a consideration when interpreting the data. Use cytotoxicity testing methods that can provide information regarding 80% cell viability, utilizing an appropriate assay based upon laboratory experience.

Should the results of the cytotoxicity test show that the concentration of the test substance has reduced the cell number by 20% or more, this concentration is regarded as cytotoxic, and concentrations at or above the cytotoxic concentration should be excluded from the evaluation.

- (9) **Chemical Exposure and Assay Plate Organization.** The procedure for chemical dilutions (Steps-1 and 2) and exposure to cells (Step-3) can be conducted as follows:

Step 1: Dilute each test chemical by serial dilution in DMSO, or appropriate solvent, and add to the wells of a microtitre plate to achieve final serial concentrations as determined by the preliminary range finding test (typically in a series of, for example 1 mM, 100 µM, 10 µM, 1 µM, 100 nM, 10 nM, 1 nM, 100 pM, and 10 pM (10^{-3} - 10^{-11} M)) for triplicate testing.

Step 2: Chemical dilution: First dilute 1.5 µL of the test chemical in the solvent to a concentration of 500 µL of media.

Step 3: Chemical exposure of the cells: Add 50 µL of dilution with media (prepared in Step-2) to an assay well containing 10^4 cells/100 µL/well.

The recommended final volume of media required for each well is 150 µL.

Test samples and reference chemicals can be assigned as shown in Table 3.

Table 3. Example of Plate Concentration Assignment of the Reference Chemicals in the Assay Plate.

Row	17 α -Methyltestosterone			Corticosterone			17 α -Estradiol			E2		
	1	2	3	4	5	6	7	8	9	10	11	12
A	conc 1 (10 μ M)	→	→	100 μ M	→	→	1 μ M	→	→	10 nM	→	→
B	conc 2 (1 μ M)	→	→	10 μ M	→	→	100 nM	→	→	1 nM	→	→
C	conc 3 (100 nM)	→	→	1 μ M	→	→	10 nM	→	→	100 pM	→	→
D	conc 4 (10 nM)	→	→	100 nM	→	→	1 nM	→	→	10 pM	→	→
E	conc 5 (1 nM)	→	→	10 nM	→	→	100 pM	→	→	1 pM	→	→
F	conc 6 (100 pM)	→	→	1 nM	→	→	10 pM	→	→	0.1 pM	→	→
G	conc 7 (10 pM)	→	→	100 pM	→	→	1 pM	→	→	0.01 pM	→	→
H	VC	→	→	→	→	→	PC	→	→	→	→	→

Plate controls = VC: Vehicle control (DMSO); PC: Positive control (1 nM E2)

Test the reference chemicals (E2, 17 α -Estradiol, 17 α -methyl testosterone and corticosterone) in every run (Table 3). Include 1) PC wells treated with 1 nM of E2 that can produce maximum induction of E2, and 2) VC wells treated with DMSO (or appropriate solvent) alone, in each test assay plate (Table 4). If cells from different sources (e.g., different passage number, different lot, etc.) are used in the same experiment, the test the reference chemicals with each cell source.

Table 4. Example of Plate Concentration Assignment of Test and Plate Control Chemicals in the Assay Plate.

Row	Test Chemical 1			Test Chemical 2			Test Chemical 3			Test Chemical 4		
	1	2	3	4	5	6	7	8	9	10	11	12
A	conc 1 (10 μ M)	→	→	1 mM	→	→	1 μ M	→	→	10 nM	→	→
B	conc 2 (1 μ M)	→	→	100 μ M	→	→	100 nM	→	→	1 nM	→	→
C	conc 3 (100 nM)	→	→	10 μ M	→	→	10 nM	→	→	100 pM	→	→
D	conc 4 (10 nM)	→	→	1 μ M	→	→	1 nM	→	→	10 pM	→	→
E	conc 5 (1 nM)	→	→	100 nM	→	→	100 pM	→	→	1 pM	→	→
F	conc 6 (100 pM)	→	→	10 nM	→	→	10 pM	→	→	0.1 pM	→	→
G	conc 7 (10 pM)	→	→	1 nM	→	→	1 pM	→	→	0.01 pM	→	→
H	VC	→	→	→	→	→	PC	→	→	→	→	→

Confirm the lack of edge effects, as appropriate, and if edge effects are suspected, alter the plate layout to avoid such effects. For example, a plate layout excluding the edge wells can be employed.

After adding the chemicals, incubate the assay plates in a 5% CO₂ incubator at 37 \pm 1°C for 20-24 hours to induce the reporter gene products.

Special considerations will need to be applied to those compounds that are highly volatile. In such cases, nearby control wells may generate false positives, and it is important they be considered in light of expected and historical control values. In the few cases where volatility may be of concern, the use of “plate sealers” may help to effectively isolate individual wells during testing, and is therefore recommended in such cases.

Conduct repeat definitive tests for the same chemical on different days, to ensure independence.

- (10) **Luciferase assay.** A commercial luciferase assay reagent [e.g., Steady-Glo® Luciferase Assay System (Promega, E2510, or equivalents)] or a standard luciferase assay system (e.g., Promega, E1500, or equivalents) can be used for the assay, as long as the results match the acceptability criteria as defined in this assay. Select the assay reagents based on the sensitivity of the luminometer to be used. When using the standard luciferase assay system, use the Cell Culture Lysis Reagent (e.g., Promega, E1531, or equivalents) before adding the substrate. Follow the manufacturer’s instructions when using the luciferase reagent.
- (f) **Analysis of Data.** To obtain the relative transcriptional activity to PC (1 nM of E2), the luminescence signals from the same plate can be analyzed according to the following steps (other equivalent mathematical processes are also acceptable):
 - Step 1: Calculate mean value for the VC.
 - Step 2: Subtract the mean value of the VC from each well value to normalize the data.
 - Step 3: Calculate the mean for the normalized PC.
 - Step 4: Divide the normalized value of each well in the plate by the mean value of the normalized PC (PC=100%). The final value of each well is the relative transcriptional activity for that well compared to the PC response.
 - Step 5: Calculate the mean value of the relative transcriptional activity for each concentration group of the test chemical. There are two dimensions to the response: the averaged transcriptional activity (response) and the concentration at which the response occurs (see following section).
- (1) **Considerations for Induction of EC₅₀, PC₅₀ and PC₁₀.** The full concentration response curve is required for the calculation of the EC₅₀, but this may not always be achievable or practical due to limitations of the test concentration range (for example due to cytotoxicity or solubility problems). However, as the EC₅₀ and maximum induction level (corresponding to the top value of the Hill-equation) are informative parameters, report these parameters where possible. For the calculation

of EC₅₀ and maximum induction level, use appropriate statistical software (e.g., Graphpad Prism statistical software).

If the Hill's logistic equation is applicable to the concentration response data, calculate the EC₅₀ by the following equation (**Ref 15**):

$$Y = \text{Bottom} + (\text{Top} - \text{Bottom}) / (1 + 10^{\exp((\log \text{EC}_{50} - X) \times \text{Hillslope}))}$$

Where:

X is the logarithm of concentration; and,

Y is the response and Y starts at the Bottom and goes to the Top in a sigmoid curve.

Bottom is fixed at zero in the Hill's logistic equation.

For each test chemical, provide the following:

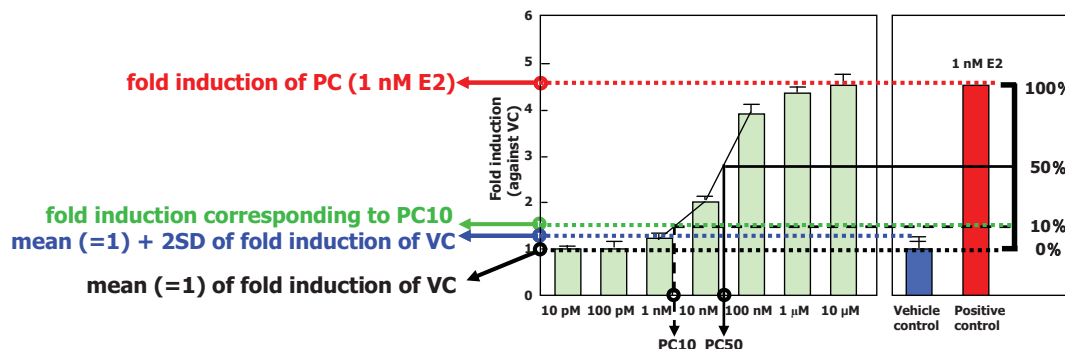
- ☐ The RPC_{Max} which is the maximum level of response induced by a test chemical, expressed as a percentage of the response induced by 1 nM E2 on the same plate, as well as the PC_{Max} (concentration associated with the RPC_{Max}).
- ☐ For positive chemicals, the concentrations that induce the PC₁₀ and, if appropriate, the PC₅₀.

The PC_x value can be calculated by interpolating between 2 points on the X-Y coordinate, one immediately above and one immediately below a PC_x value. Where the data points lying immediately above and below the PC_x value have the coordinates (a,b) and (c,d) respectively, then the PC_x value may be calculated using the following equation:

$$\log[\text{PC}_x] = \log[c] + (x - d) / (d - b)$$

Descriptions of PC values are provided in Figure 1 below.

Figure 1. Example of How to Derive PC-values. The PC (Positive control; 1 nM of E2) is included on each assay plate.



(2) **Performance Standards.**

To be acceptable, the following are considered:

- The mean luciferase activity of the positive controls (1 nM E2) is at least 4-fold that of the mean vehicle control on each plate.
- The fold induction corresponding to the PC₁₀ value of the concurrent PC (1 nM E2) is greater than 1+2SD of the fold induction value (=1) of the VC (vehicle control).
- The results of 4 reference chemicals are within the acceptable range (Table 1).
- Be reproducible.

- (3) **Data Interpretation Criteria.** Base the results on two (or three) independent runs. If two runs give comparable and therefore reproducible results, it is not necessary to conduct a third run. Data interpretation criteria are shown in Table 5. Positive results will be characterized by both the magnitude of the effect and the concentration at which the effect occurs. Expressing results as a concentration at which a 50% (PC₅₀) or 10% (PC₁₀) of positive control values are reached accomplishes both of these goals. However, a test chemical is determined to be positive, if the maximum response induced by the test chemical (RPC_{Max}) is equal to or exceeds 10% of the response of the positive control in at least two of two or two of three runs, while a test chemical is considered negative if the RPC_{Max} fails to achieve at least 10% of the response of the positive control in two of two or two of three runs.

Table 5. Positive and Negative Decision Criteria.

Positive	If the RPC _{Max} is obtained that is equal to or exceeds 10% of the response of the positive control in at least two of two or two of three runs.
Negative	If the RPC _{Max} fails to achieve at least 10% of the response of the positive control in two of two or two of three runs.

The EPA intends to provide a calculation spreadsheet with the posting of this guideline on the Agency's Web site (**Ref. 17**) that may be utilized to determine PC₁₀, PC₅₀ and PC_{Max}.

Obtaining PC₁₀ or PC₅₀ values at least twice is sufficient, unless the resulting base-line for data in the same concentration range shows variability with an unacceptably high coefficient of variation (CV; %). In such a case, the data may not be considered reliable and it is recommended that the source of the high variability be identified. The target CV of the raw data triplicates (i.e. luminescence intensity data) of the data points that are used for the calculation of PC₁₀ is less than 20%.

Meeting the acceptability criteria indicates the assay system is operating properly, but it does not ensure that any particular run will produce accurate data. Duplicating the results of the first run is the best insurance that accurate data were produced (see above).

Where more information is required in addition to the screening and prioritization purposes of this TG for positive test compounds, particularly for PC₁₀-PC₄₉ chemicals, as well as chemicals suspected to over stimulate luciferase, it can be confirmed that the observed luciferase-activity is solely an ER α -specific response, using an ER α antagonist (see Appendix 3).

(g) **Test Report.** Include the following information in the test report:

- ☐ **Test substance:**
 - Identification information (e.g., molecular weight, lot, supplier, expiration date) and CAS Number, if known
 - Physical nature and purity
 - Physicochemical properties relevant to the conduct of the study
 - Stability of the test substance
- ☐ **Solvent/Vehicle:**
 - Characterization (nature, supplier and lot)
 - Justification for choice of solvent/vehicle
 - Solubility and stability of the test substance in solvent/vehicle, if known
- ☐ **Cells:**
 - Type and source of cells
 - Number of cell passages
 - Methods for maintenance of cell cultures
- ☐ **Test conditions:**
 - Report cytotoxicity data (and justifications for the method of choice) and solubility limitations, as well as:

- Composition of media, CO₂ concentration
- Concentration of test chemical
- Volume of vehicle and test substance added
- Incubation temperature and humidity
- Duration of treatment
- Cell density during treatment
- Positive and negative reference chemicals
- Duration of treatment period
- Luciferase assay reagents (Product name, supplier and lot)
- Acceptability and data interpretation criteria.

☐ **Reliability check:**

- Fold inductions for each assay plate
- Actual logEC₅₀, logPC₅₀, logPC₁₀ and Hill slope values for concurrent reference chemicals

☐ **Results:**

- Raw and normalized data of luminescent signals
- Concentration-response relationship, where possible
- RPC_{Max}, P_{Max}, PC₅₀ and/or PC₁₀ values, as appropriate
- EC₅₀ values, if appropriate
- Statistical analyses, if any, together with a measure of error (e.g., SEM, SD, CV or 95% CI) and a description of how these values were obtained.

☐ **Discussion of the results.**

☐ **Conclusion.**

(h) **References.**

1. CERI (2006). Draft validation report of TA assay using HeLa-hER-9903 to detect estrogenic activity. Available at: http://www.oecd.org/document/62/0,3343,en_2649_34377_2348606_1_1_1_1,00.html.
2. Takeyoshi, M., Yamasaki, K., Sawaki, M., Nakai, M., Noda, S. and Takatsuki, M. (2002). The efficacy of endocrine disruptor screening tests in detecting anti-estrogenic effects downstream of receptor-ligand interactions. *Toxicol. Lett*, 126, 91-98.
3. Jefferson, W.N., Padilla-Banks, E., Clark, G. and Newbold R. (2002). Assessing estrogenic activity of phytochemicals using transcriptional activation and immature mouse uterotrophic responses. *J. Chromat. B.*, 777, 179-189.
4. Sonneveld, E., Riteco, J.A., Jansen, H.J., Pieterse, B., Brouwer, A., Schoonen, W.G. and van der Burg, B. (2006). Comparison of *in vitro* and *in vivo* screening

- models for androgenic and estrogenic activities. *Toxicol. Sci.*, 89, 173-187.
5. Escande, A., Pillon, A., Servant, N., Cravedi, J.P., Larrea, F., Muhn, P., Nicolas, J.C., Cavailles, V. and Balaguer, P. (2006). Evaluation of ligand selectivity using reporter cell lines stably expressing estrogen receptor alpha or beta. *Biochem. Pharmacol.*, 71, 1459-1469.
 6. Gray, L.E. Jr. (1998). Tiered screening and testing strategy for xenoestrogens and antiandrogens. *Toxicol. Lett.*, 102-103, 677-680.
 7. EDSTAC (1998). Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) Final Report. Available at: <http://www.epa.gov/scipoly/oscpendo/pubs/edspoverview/finalrpt.htm>.
 8. 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/methods/endocrine.htm#fineval>.
 9. Escande, A., Pillon, A., Servant, N., Cravedi, J.P., Larrea, F., Muhn, P., Nicolas, J.C., Cavailles, V. and Balaguer, P. (2006). Evaluation of ligand selectivity using reporter cell lines stably expressing estrogen receptor alpha or beta. *Biochem. Pharmacol.*, 71, 1459-1469.
 10. Kuiper, G.G., Lemmen, J.G., Carlsson, B., Corton, J.C., Safe, S.H., van der Saag, P.T., van der Burg, B. and Gustafsson, J.A. (1998). Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta. *Endocrinol.*, 139, 4252-4263.
 11. Spaepen M, Angulo AF, Marynen P, Cassiman JJ. (1992). Detection of bacterial and mycoplasma contamination in cell cultures by polymerase chain reaction. *FEMS Microbiol Lett.* 78(1), 89-94.
 12. Kobayashi H, Yamamoto K, Eguchi M, Kubo M, Nakagami S, Wakisaka S, Kaizuka M, Ishii H (1995). Rapid detection of mycoplasma contamination in cell cultures by enzymatic detection of polymerase chain reaction (PCR) products. *J Vet Med Sci.* 57(4), 769-71.
 13. Dussurget O, Roulland-Dussoix D (1994). Rapid, Sensitive PCR-based Detection of Mycoplasmas in Simulated Samples of Animal Sera. *Appl Environ Microbiol.* 60(3), 953-9.
 14. De Lean A, Munson, P.J. Rodbard D (1978). Simultaneous Analysis of Families of Sigmoidal Curves: Application to Bioassay, Radioligand Assay, and Physiological Dose-response Curves. *Am J Physiol.* 235, E97-E102.
 15. OECD (2005). Guidance Document on The Validation and International

Acceptance of New or Updated Test Methods for Hazard Assessment. OECD Series on Testing and Assessment. No. 34. Paris, France.

16. EPA. OPPTS Harmonized Test Guidelines. Available on-line at: <http://www.epa.gov/oppts> (select "Test Methods & Guidelines" on the left side navigation menu). You may also access the guidelines in <http://www.regulations.gov> grouped by Series under Docket ID #s: EPA-HQ-OPPT-2009-0150 through EPA-HQ-OPPT-2009-0159, and EPA-HQ-OPPT-2009-0576.

Appendix 1

Definitions and Abbreviations

Agonist: A substance that binds to a specific [receptor](#) and triggers a response in the [cell](#). It mimics the action of an [endogenous ligand](#) that binds to the same receptor.

Antagonist: A type of receptor [ligand](#) or chemical that does not provoke a biological response itself upon binding to a [receptor](#), but blocks or dampens [agonist](#)-mediated responses.

Anti-estrogenic activity, the capability of a chemical to suppress the action of 17 β -estradiol mediated through estrogen receptors.

CV: Coefficient of variation

Cytotoxicity: the harmful effects to cell structure or function ultimately causing cell death and can be a result of 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.

DCC-FBS: Dextran-coated charcoal treated fetal bovine serum.

DMSO: Dimethyl sulfoxide

E2: 17 β -estradiol

EC₅₀ value, the concentration of agonist that provokes a response halfway between the baseline (Bottom) and maximum response (Top).

EE: 17 α -ethynyl estradiol

ER; Estrogen receptor

ERE: Estrogen Response Element

Estrogenic activity, the capability of a chemical to mimic 17 β -estradiol in its ability to bind to and activate estrogen receptors. hER α mediated specific estrogenic activity can be detected in this Test Guideline.

FBS: Fetal bovine serum

hER α : Human estrogen receptor alpha

MT: Metallothionein

OHT: 4-Hydroxytamoxifen

PC: Positive control

PC₁₀: the concentration of a test chemical at which the response in an agonist assay is 10% of the response induced by positive control (E2 at 1nM) in each plate

PC₅₀: the concentration of a test chemical at which the response in an agonist assay is 50% of the response induced by positive control (E2 at 1nM) in each plate

PC_{Max}: the concentration of a test chemical inducing the RPCMax

RPC_{Max}: maximum level of response induced by a test chemical, expressed as a percentage of the response induced by 1 nM E2 on the same plate

RT PCR: Real Time polymerase chain reaction

SD: Standard deviation

STTA: Stably Transfected Transcriptional Activation Assay.

TA: Transcriptional activation

Validation, a process based on scientifically sound principles by which the reliability and relevance of a particular test, approach, method, or process are established for a specific purpose. Reliability is defined as the extent of reproducibility of results from a test within and among laboratories over time, when performed using the same standardized protocol. The relevance of a test method describes the relationship between the test and the effect in the target species and whether the test method is meaningful and useful for a defined purpose, with the limitations identified. In brief, it is the extent to which the test method correctly measures or predicts the (biological) effect of interest, as appropriate (16).

VC: The vehicle that is used to dissolve test and control chemicals is tested solely as vehicle without dissolved chemical.

Appendix 2

False Positives: Assessment of Non-receptor Mediated Luminescence Signals

1. **False Positives**

False positives might be generated by non-ER-mediated activation of the luciferase gene, or direct activation of the gene product or unrelated fluorescence. Such effects are indicated by an incomplete or unusual dose-response curve. If such effects are suspected, examine the effect of an ER antagonist (*e.g.*, 4-hydroxytamoxifen (OHT) at non-toxic concentration) on the response. The pure antagonist ICI 128780 may not be suitable for this purpose as a sufficient concentration of ICI 128780 may decrease the vehicle control value, and this will affect the data analysis.

To ensure validity of this approach, the following needs to be tested in the same plate:

- Agonistic activity of the unknown chemical with / without 10 μ M of OHT
- Vehicle Control (VC)(in triplicate)
- OHT (in triplicate)
- 1 nM of E2 (in triplicate) as agonist Positive Control (PC)
- 1 nM of E2 + OHT (in triplicate)

3. **Data Interpretation Criteria**

Note: *Treat all wells with the same concentration of the vehicle.*

- If the agonistic activity of the unknown chemical is NOT affected by the treatment with ER antagonist, it is classified as “Negative”.
- If the agonistic activity of the unknown chemical is completely inhibited, apply the decision criteria.
- If the agonistic activity at the lowest concentration is equal to, or is exceeding, PC10 response the unknown chemical is inhibited equal to or exceeding PC10 response. The difference in the responses between the non-treated and treated wells with the ER antagonist is calculated and considered as the true response to be used for the calculation of the appropriate parameters to enable a classification decision to be made.

4. **Data Analysis**

- Check the performance standard.
- Check the CV between wells treated under the same conditions.
- Calculate the mean of the VC.

- Subtract the mean of VC from each well value **not** treated with OHT.
- Calculate the mean of OHT.
- Subtract the mean of the VC from each well value treated with OHT.
- Calculate the mean of the PC.
- Calculate the relative transcriptional activity of all other wells relative to the PC.

Appendix 3

Preparation of Serum treated with Dextran Coated Charcoal (DCC)

The treatment of serum with dextran-coated charcoal (DCC) is a general method for removal of estrogenic compounds from serum that is added to cell medium, in order to exclude the biased response associated with residual estrogens in serum. 500 mL of fetal bovine serum (FBS) can be treated by this procedure.

Components

The following materials and equipment will be needed:

Materials

- Activated charcoal
- Dextran
- Magnesium chloride hexahydrate ($\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$)
- Sucrose
- 1 M HEPES buffer solution (pH 7.4)
- Ultrapure water produced from a filter system

Equipment

- Autoclaved glass container (size should be adjusted as appropriate)
- General Laboratory Centrifuge (that can set temperature at 4°C.)

Procedure

The following procedure is adjusted for the use of 50 mL centrifuge tubes:

[Day-1] Prepare dextran-coated charcoal suspension with 1 litre of ultrapure water containing 1.5 mM of MgCl_2 , 0.25 M sucrose, 2.5 g of charcoal, 0.25 g dextran and 5 mM of HEPES and stir it at 4°C, overnight.

[Day-2] Dispense the suspension in 50 mL centrifuge tubes and centrifuge at 10000 rpm at 4°C for 10 minutes. Remove the supernatant and store half of the charcoal sediment at 4°C for the use on Day-3. Suspend the other half of the charcoal with FBS that has been gently thawed to avoid precipitation, and heat-inactivated at 56°C for 30 minutes, then transfer into an autoclaved glass container such as an Erlenmeyer flask. Stir this suspension gently at 4°C, overnight.

[Day-3] Dispense the suspension with FBS into centrifuge tubes for centrifugation at 10000 rpm at 4°C for 10 minutes. Collect FBS and transfer into the new charcoal sediment prepared and stored on Day-2. Suspend the charcoal sediment and stir this suspension gently in an autoclaved glass container at 4°C, overnight.

[Day-4] Dispense the suspension for centrifugation at 10000 rpm at 4°C for 10 minutes and sterilize the supernatant by filtration through 0.2 µm sterile filter. This DCC treated FBS should be stored at -20°C and can be used for up a year.

This page intentionally left blank

Appendix F3

**OECD Test Guideline 455: Stably Transfected Human Estrogen Receptor- α
Transcriptional Activation Assay for Detection of Estrogenic Agonist-Activity of
Chemicals
(Adopted September 2009)**

This page intentionally left blank

OECD GUIDELINE FOR THE TESTING OF CHEMICALS

Stably Transfected Human Estrogen Receptor- α Transcriptional Activation Assay for Detection of Estrogenic Agonist-Activity of Chemicals

INTRODUCTION

1. The OECD initiated a high-priority activity in 1998 to revise existing, and to develop new, Test Guidelines for the screening and testing of potential endocrine disrupting chemicals. The OECD conceptual framework for testing and assessment of potential endocrine disrupting chemicals comprises five levels, each level corresponding to a different level of biological complexity (1). The Transcriptional Activation (TA) assay described in this Test Guideline is a level 2 “*in vitro* assay, providing mechanistic information”. The validation study of the Stably Transfected Transactivation Assay (STTA) by the Japanese Chemicals Evaluation and Research Institute (CERI) using the hER α -HeLa-9903 cell line to detect estrogenic agonist activity mediated through human estrogen receptor alpha (hER α) demonstrated the relevance and reliability of the assay for its intended purpose (2).

2. *In vitro* TA assays are based upon the production of a reporter gene product induced by a chemical, following binding of the chemical to a specific receptor and subsequent downstream transcriptional activation. TA assays using activation of reporter genes are screening assays that have long been used to evaluate the specific gene expression regulated by specific nuclear receptors, such as the estrogen receptors (ERs) (3) (4) (5) (6). They have been proposed for the detection of estrogenic transactivation regulated by the ER (7) (8) (9). The nuclear ERs exist as at least two subtypes, termed α and β , encoded by distinct genes and with different tissue distribution, relative ligand binding affinities and biological functions. Nuclear ER α mediates the classic estrogenic response, therefore models currently being developed to measure ER activation mainly relate to ER α . The aim of this TA assay is to evaluate the ability of a chemical to function as an ER α ligand and activate an agonist response, for screening and prioritisation purposes but can also provide mechanistic information that can be used in a weight of evidence approach.

3. Definitions and abbreviations used in this Test Guideline are described in Annex 1.

INITIAL CONSIDERATIONS AND LIMITATIONS

4. Estrogen agonists act as ligands for ERs, and may activate the transcription of estrogen responsive genes. This interaction may have the potential to trigger adverse health effects by disrupting estrogen-regulated systems. This Test Guideline describes an assay that evaluates TA mediated by the hER α . This process is considered to be one of the key mechanisms of possible endocrine disruption related health hazards, although there are also other important endocrine disruption mechanisms. These include (i) actions mediated via other nuclear receptors linked to the endocrine system and interactions with steroidogenic enzymes, (ii) metabolic activation or deactivation of hormones, (iii) distribution of hormones to target tissues, and (iv) clearance of hormones from the body. This Test Guideline exclusively addresses TA of an estrogen-regulated reporter gene by agonist binding to the hER α , and therefore it should not be directly extrapolated to the complex *in vivo* situation of estrogen regulation of cellular processes. Furthermore, this Test

© OECD, (2009)

You are free to use this material for personal, non-commercial purposes without seeking prior consent from the OECD, provided the source is duly mentioned. Any commercial use of this material is subject to written permission from the OECD.

Guideline does not address antagonist interaction with the hER α and subsequent effect on transcription.

5. This test method is specifically designed to detect hER α -mediated TA by measuring chemiluminescence as the endpoint. However, non-receptor-mediated luminescence signals have been reported at phytoestrogen concentrations higher than 1 μ M due to the over-activation of the luciferase reporter gene (10) (11). While the dose-response curve indicates that true activation of the ER system occurs at lower concentrations, luciferase expression obtained at high concentrations of phytoestrogens or similar compounds suspected of producing phytoestrogen-like over-activation of the luciferase reporter gene needs to be examined carefully in stably transfected ER TA assay systems (Annex 2).

6. It is recognized that this assay using the hER α -HeLa-9903 cell line is only one of several ER transcriptional activation assays currently being developed and validated. It is, therefore the intention that a generic performance based Test Guideline will replace this Test Guideline as soon as such guideline is developed and approved.

PRINCIPLE OF THE TEST

7. The TA assay using a reporter gene technique is an *in vitro* tool that provides mechanistic data. The assay is used to signal binding of the estrogen receptor with a ligand. Following ligand binding, the receptor-ligand complex translocates to the nucleus where it binds specific DNA response elements and transactivates a firefly 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. Luciferase activity can be evaluated quickly and inexpensively with a number of commercially available test kits.

8. The test system provided in this Test Guideline utilises the hER α -HeLa-9903 cell line, which is derived from a human cervical tumor, with two stably inserted constructs: (i) the hER α expression construct (encoding the full-length human receptor), and (ii) a firefly luciferase reporter construct bearing five tandem repeats of a vitellogenin Estrogen-Responsive Element (ERE) driven by a mouse metallothionein (MT) promoter TATA element. The mouse MT TATA gene construct has been shown to have the best performance, and so is commonly used. Consequently this hER α -HeLa-9903 cell line can measure the ability of a test chemical to induce hER α -mediated transactivation of luciferase gene expression.

9. Data interpretation for this assay is based upon whether or not the maximum response level induced by a test chemical equals or exceeds an agonist response equal to 10% of that induced by a maximally inducing (1 nM) concentration of the positive control (PC) 17 β estradiol (E2) (*i.e.* the PC10). Data analysis and interpretation are discussed in greater detail in paragraphs 34- 45.

PROCEDURE

Cell Lines

10. The stably transfected hER α -HeLa-9903 cell line should be used for the assay. The cell line can be obtained from the Japanese Collection of Research Bioresources (JCRB) Cell Bank¹, upon signing a Material Transfer Agreement (MTA).

¹ JCRB Cell Bank : National Institute of Biomedical Innovation, 7-6-8 Asagi Saito, Ibaraki-shi, Osaka 567-0085, Japan Fax: +81-72-641-9812

OECD/OCDE

455

11. Only cells characterised as mycoplasma-free should be used in testing. RT PCR (Real Time Polymerase Chain Reaction) is the method of choice for a sensitive detection of mycoplasma infection (12) (13) (14).

Stability of the cell line

12. To monitor the stability of the cell line, E2, 17 α -estradiol, 17 α -methyltestosterone, and corticosterone should be used as the reference chemicals and a complete concentration-response curve in the test concentration range provided in Table 1 should be measured at least once each time the assay is performed, and the results should be in agreement with the results provided in Table 1.

Cell Culture and Plating Conditions

13. Cells should be maintained in Eagle's Minimum Essential Medium (EMEM) without phenol red, supplemented with 60 mg/L of antibiotic Kanamycine and 10% dextran-coated-charcoal-treated fetal bovine serum (DCC-FBS), in a CO₂ incubator (5% CO₂) at 37 \pm 1°C. Upon reaching 75-90% confluency, cells can be subcultured at 10 mL of 0.4 x 10⁵ – 1 x 10⁵ cells/mL for 100 mm cell culture dish. Cells should be suspended with 10% FBS-EMEM (which is the same as EMEM with DCC-FBS) and then plated into wells of a microplate at a density of 1 x 10⁴ cells/100 μ L/well. Next, the cells should be pre-incubated in a 5% CO₂ incubator at 37 \pm 1°C for 3 hours before the chemical exposure. The plastic-ware should be free of estrogenic activity.

14. To maintain the integrity of the response, the cells should be grown for more than one passage from the frozen stock in the conditioned media and should not be cultured for more than 40 passages. For the hER α -HeLa-9903 cell line, this will be less than three months.

15. The DCC-FBS can be prepared as described in Annex 3, or obtained from commercial sources.

Acceptability Criteria

Positive and Negative Reference Chemicals

16. Prior to and during the study, the responsiveness of the test system should be verified using the appropriate concentrations of a strong estrogen: E2, a weak estrogen (17 α -estradiol), a very weak agonist (17 α -methyltestosterone) and a negative compound (corticosterone). Acceptable range values derived from the validation study are given in Table 1 (2). These 4 concurrent reference chemicals should be included with each experiment and the results should fall within the given acceptable limits. If this is not the case, the cause for the failure to meet the acceptability criteria should be determined (e.g. cell handling, and serum and antibiotics for quality and concentration) and the assay repeated. Once the acceptability criteria have been achieved, to ensure minimum variability of EC50, PC50 and PC10 values, consistent use of materials for cell culturing is essential. The four concurrent reference chemicals, which should be included in each experiment (conducted under the same conditions including the materials, passage level of cells and technicians), can ensure the sensitivity of the assay because the PC10s of the three positive reference chemicals should fall within the acceptable range, as should the PC50s and EC50s where they can be calculated (see Table 1).

455

OECD/OCDE

Table 1. Acceptable range values of the 4 reference chemicals for the STTA assay (means \pm 2 standard deviations)(SD).

Name	logPC50	logPC10	logEC50	Hill slope	Test range
17 β -Estradiol (E2) CAS No: 50-28-2	-11.4 ~ -10.1	<-11	-11.3 ~ -10.1	0.7 ~ 1.5	10 ⁻¹⁴ ~ 10 ⁻⁸ M
17 α -Estradiol CAS No: 57-91-0	-9.6 ~ -8.1	-10.7 ~ -9.3	-9.6 ~ -8.4	0.9 ~ 2.0	10 ⁻¹² ~ 10 ⁻⁶ M
Corticosterone CAS No: 50-22-6	—	—	—	—	10 ⁻¹⁰ ~ 10 ⁻⁴ M
17 α -Methyltestosterone CAS No: 58-18-4	-6.0 ~ -5.1	-8.0 ~ -6.2	—	—	10 ⁻¹¹ ~ 10 ⁻⁵ M

Positive and Vehicle Controls

17. The positive control (PC) (1 nM of E2) should be tested at least in triplicate in each plate. The vehicle that is used to dissolve a test chemical should be tested as a vehicle control (VC) at least in triplicate in each plate. In addition to this VC, if the PC uses a different vehicle than the test chemical, another VC should be tested at least in triplicate on the same plate with the PC.

Fold-induction

18. The mean luciferase activity of the PC (1 nM E2) should be at least 4-fold that of the mean VC on each plate. This criterion is established based on the reliability of the endpoint values from the validation study (historically between four- and 30-fold).

19. With respect to the quality control of the assay, the fold-induction corresponding to the PC10 value of the concurrent PC (1 nM E2) should be greater than 1+2SD of the fold-induction value (=1) of the concurrent VC. For prioritisation purposes, the PC10 value can be useful to simplify the data analysis required compared to a statistical analysis. Although a statistical analysis provides information on significance, such an analysis is not a quantitative parameter with respect to concentration-based potential, and so is less useful for prioritisation purposes.

Chemicals to Demonstrate Laboratory Proficiency

20. Prior to testing unknown chemicals in the STTA assay, the responsiveness of the test system should be confirmed by each laboratory, at least once for each newly prepared batch of cell stocks taken from the frozen stock by independent testing of the 10 proficiency chemicals listed in Table 2. This should be done at least in duplicate, on different days, and the results should be comparable to Table 2 and any deviations should be justified.

Table 2. List of Proficiency Chemicals

Compound	CAS No.	Class ²	Test concentration range	Note
Diethylstilbestrol (DES)	56-53-1	Positive	10^{-14} – 10^{-8} M	
17 α -Ethinyl estradiol (EE)	57-63-6	Positive	10^{-14} – 10^{-8} M	
Hexestrol	84-16-2	Positive	10^{-13} – 10^{-7} M	
Genistein	446-72-0	Positive	10^{-12} – 10^{-5} M	Cytotoxic at (0.01) ⁴ , 0.1 and 1 mM
Estrone	53-16-7	Positive	10^{-12} – 10^{-6} M	
Butyl paraben	94-26-8	Positive	10^{-11} – 10^{-4} M	Cytotoxic at (0.1) ⁴ and 1 mM
1,3,5-Tris(4hydroxyphenyl)benzene ¹	15797-52-1	Positive	10^{-12} – 10^{-5} M	Cytotoxic at 100 μ M. PCmax approximately 15% of PC Binds to hER α and has ER antagonist activity
Dibutyl phthalate (DBP)	84-74-2	Negative ³	10^{-11} – 10^{-4} M	Cytotoxic at 1 mM
Atrazine	1912-24-9	Negative	10^{-11} – 10^{-4} M	Cytotoxic ⁴ at 1 mM
Corticosterone	50-22-6	Negative	10^{-10} – 10^{-4} M	If not cytotoxic at 1 mM, then that should be the highest tested concentration

¹ Compound selected to challenge solubility and cytotoxicity.

² See Table 5 for definitions of positive and negative.

³ Negative for ER α mediated TA but may not be negative for non-ER α mediated TA. Thus a positive result in this assay with DBP would indicate that the system is detecting other than pure ER α mediated activity and is therefore unacceptable.

⁴ Cytotoxicity is close to 80%.

Vehicle

21. Dimethyl sulfoxide (DMSO), or appropriate solvent, at the same concentration used for the different positive and negative controls and the test chemicals should be used as the concurrent VC. Test substances should be dissolved in a solvent that solubilizes that test substance and is miscible with the cell medium. Water, ethanol (95% to 100% purity) and DMSO are suitable vehicles. If DMSO is used, the level should not exceed 0.1% (v/v). For any vehicle, it should be demonstrated that the maximum volume used is not cytotoxic and does not interfere with assay performance.

Preparation of Test Chemicals

22. Generally, the test chemicals should be dissolved in DMSO or other suitable solvent, and serially diluted with the same solvent at a common ratio of 1:10 in order to prepare solutions for dilution with media.

Solubility and Cytotoxicity: Considerations for Range Finding.

23. A preliminary test should be carried out to determine the appropriate concentration range of chemical to be tested, and to ascertain whether the test chemical may have any solubility and cytotoxicity problems. Initially, chemicals are tested up to the maximum concentration of 1 µl/ml, 1 mg/ml, or 1 mM, whichever is the lowest. Based on the extent of cytotoxicity or lack of solubility observed in the preliminary test, the first definite run should test the chemical at log-serial dilutions starting at the maximum acceptable concentration (*e.g.* 1 mM, 100 µM, 10 µM, etc.) and the presence of cloudiness or precipitate or cytotoxicity noted. Concentrations in the second, and if necessary third run should be adjusted as appropriate to better characterise the concentration-response curve and to avoid concentrations which are found to be insoluble or to induce excessive cytotoxicity.

24. For ER agonists, the presence of increasing levels of cytotoxicity can significantly alter or eliminate the typical sigmoidal response and should be considered when interpreting the data. Cytotoxicity testing methods that can provide information regarding 80% cell viability should be used, utilising an appropriate assay based upon laboratory experience.

25. Should the results of the cytotoxicity test show that the concentration of the test substance has reduced the cell number by 20% or more, this concentration is regarded as cytotoxic, and the concentrations at or above the cytotoxic concentration should be excluded from the evaluation.

Chemical Exposure and Assay Plate Organisation

26. The procedure for chemical dilutions (Steps-1 and 2) and exposure to cells (Step-3) can be conducted as follows:

Step-1: Each test chemical should be serially diluted in DMSO, or appropriate solvent, and added to the wells of a microtitre plate to achieve final serial concentrations as determined by the preliminary range finding test (typically in a series of, for example 1 mM, 100 µM, 10 µM, 1 µM, 100 nM, 10 nM, 1 nM, 100 pM, and 10 pM (10^{-3} - 10^{-11} M)) for triplicate testing.

Step-2: Chemical dilution: First dilute 1.5 µL of the test chemical in the solvent to a concentration of 500 µL of media.

Step-3: Chemical exposure of the cells: Add 50 µL of dilution with media (prepared in Step-2) to an assay well containing 10^4 cells/100 µL/well.

The recommended final volume of media required for each well is 150 µL.

Test samples and reference chemicals can be assigned as shown in Table 3.

Table 3.: Example of plate concentration assignment of the reference chemicals in the assay plate

Row	17 α -Methyltestosterone			Corticosterone			17 α -Estradiol			E2		
	1	2	3	4	5	6	7	8	9	10	11	12
A	conc 1 (10 µM)	→	→	100 µM	→	→	1 µM	→	→	10 nM	→	→
B	conc 2 (1 µM)	→	→	10 µM	→	→	100 nM	→	→	1 nM	→	→
C	conc 3 (100 nM)	→	→	1 µM	→	→	10 nM	→	→	100 pM	→	→
D	conc 4 (10 nM)	→	→	100 nM	→	→	1 nM	→	→	10 pM	→	→
E	conc 5 (1 nM)	→	→	10 nM	→	→	100 pM	→	→	1 pM	→	→
F	conc 6 (100 pM)	→	→	1 nM	→	→	10 pM	→	→	0.1 pM	→	→
G	conc 7 (10 pM)	→	→	100 pM	→	→	1 pM	→	→	0.01 pM	→	→
H	VC	→	→	→	→	→	PC	→	→	→	→	→

OECD/OCDE

455

Plate controls = VC: Vehicle control (DMSO); PC: Positive control (1 nM E2)

27. The reference chemicals (E2, 17 α -Estradiol, 17 α -methyl testosterone and corticosterone) should be tested in every run (Table 3). PC wells treated with 1 nM of E2 that can produce maximum induction of E2 and VC wells treated with DMSO (or appropriate solvent) alone should be included in each test assay plate (Table 4). If cells from different sources (*e.g.* different passage number, different lot, etc.) are used in the same experiment, the reference chemicals should be tested for each cell source.

Table 4.: Example of plate concentration assignment of test and plate control chemicals in the assay plate

Row	Test Chemical 1			Test Chemical 2			Test Chemical 3			Test Chemical 4		
	1	2	3	4	5	6	7	8	9	10	11	12
A	conc 1 (10 μ M)	→	→	1 mM	→	→	1 μ M	→	→	10 nM	→	→
B	conc 2 (1 μ M)	→	→	100 μ M	→	→	100 nM	→	→	1 nM	→	→
C	conc 3 (100 nM)	→	→	10 μ M	→	→	10 nM	→	→	100 pM	→	→
D	conc 4 (10 nM)	→	→	1 μ M	→	→	1 nM	→	→	10 pM	→	→
E	conc 5 (1 nM)	→	→	100 nM	→	→	100 pM	→	→	1 pM	→	→
F	conc 6 (100 pM)	→	→	10 nM	→	→	10 pM	→	→	0.1 pM	→	→
G	conc 7 (10 pM)	→	→	1 nM	→	→	1 pM	→	→	0.01 pM	→	→
H	VC	→	→	→	→	→	PC	→	→	→	→	→

28. The lack of edge effects should be confirmed, as appropriate, and if edge effects are suspected, the plate layout should be altered to avoid such effects. For example, a plate layout excluding the edge wells can be employed.

29. After adding the chemicals, the assay plates should be incubated in a 5% CO₂ incubator at 37 \pm 1°C for 20-24 hours to induce the reporter gene products.

30. Special considerations will need to be applied to those compounds that are highly volatile. In such cases, nearby control wells may generate false positives, and this should be considered in light of expected and historical control values. In the few cases where volatility may be of concern, the use of “plate sealers” may help to effectively isolate individual wells during testing, and is therefore recommended in such cases.

31. Repeat definitive tests for the same chemical should be conducted on different days, to ensure independence.

Luciferase assay

32. A commercial luciferase assay reagent [*e.g.* Steady-Glo® Luciferase Assay System (Promega, E2510, or equivalents)] or a standard luciferase assay system (Promega, E1500, or equivalents) can be used for the assay, as long as the acceptability criteria is met. The assay reagents should be selected based on the sensitivity of the luminometer to be used. When using the standard luciferase assay system, Cell Culture Lysis Reagent (Promega, E1531, or equivalents) should be used before adding the substrate. The luciferase reagent should be applied following the manufacturers' instructions.

455

OECD/OCDE

ANALYSIS OF DATA

33. To obtain the relative transcriptional activity to PC (1 nM of E2), the luminescence signals from the same plate can be analysed according to the following steps (other equivalent mathematical processes are also acceptable):

Step 1. Calculate mean value for the VC.

Step 2. Subtract the mean value of the VC from each well value to normalise the data.

Step 3. Calculate the mean for the normalised PC.

Step 4. Divide the normalised value of each well in the plate by the mean value of the normalised PC (PC=100%).

The final value of each well is the relative transcriptional activity for that well compared to the PC response.

Step 5. Calculate the mean value of the relative transcriptional activity for each concentration group of the test chemical. There are two dimensions to the response: the averaged transcriptional activity (response) and the concentration at which the response occurs (see following section).

EC50, PC50 and PC10 induction considerations

34. The full concentration-response curve is required for the calculation of the EC50, but this may not always be achievable or practical due to limitations of the test concentration range (for example due to cytotoxicity or solubility problems). However, as the EC50 and maximum induction level (corresponding to the top value of the Hill-equation) are informative parameters, these parameters should be reported where possible. For the calculation of EC50 and maximum induction level, appropriate statistical software should be used (*e.g.* Graphpad Prism statistical software).

35. If the Hill's logistic equation is applicable to the concentration response data, the EC50 should be calculated by the following equation (15):

$$Y = \text{Bottom} + (\text{Top} - \text{Bottom}) / (1 + 10^{\exp((\log \text{EC50} - X) \times \text{Hill slope}))}$$

Where:

X is the logarithm of concentration; and,

Y is the response and Y starts at the Bottom and goes to the Top in a sigmoid curve.

Bottom is fixed at zero in the Hill's logistic equation.

36. For each test chemical, the following should be provided:

(i) The RPCMax which is the maximum level of response induced by a test chemical, expressed as a percentage of the response induced by 1 nM E2 on the same plate, as well as the PCMax (concentration associated with the RPCMax); and

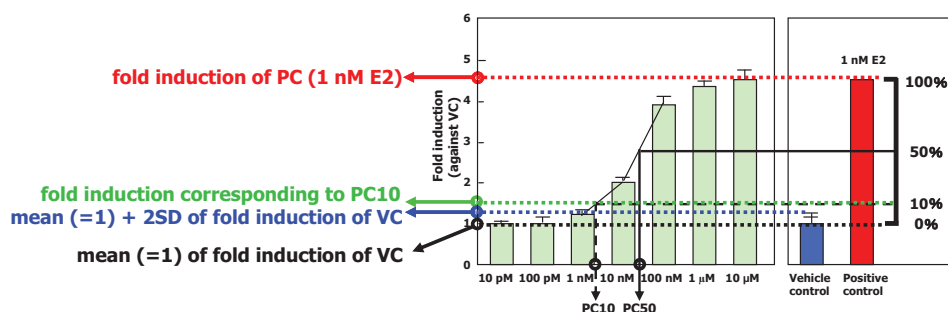
(ii) For positive chemicals, the concentrations that induce the PC10 and, if appropriate, the PC50.

37. The PCx value can be calculated by interpolating between 2 points on the X-Y coordinate, one immediately above and one immediately below a PCx value. Where the data points lying immediately above and below the PCx value have the coordinates (a,b) and (c,d) respectively, then the PCx value may be calculated using the following equation:

$$\log[\text{PCx}] = \log[c] + (x-d)/(d-b)$$

38. Descriptions of PC values are provided in Figure 1 below.

Figure 1: Example of how to derive PC-values. The PC (1 nM of E2) is included on each assay plate



39. The results should be based on two (or three) independent runs. If two runs give comparable and therefore reproducible results, it is not necessary to conduct a third run. To be acceptable, the results should:

- Meet the performance standard requirements:
 - The mean luciferase activity of the PC (1 nM E2) should be at least 4-fold that of the mean VC on each plate
 - The fold induction corresponding to the PC10 value of the concurrent PC (1 nM E2) should be greater than 1+2SD of the fold induction value (=1) of the VC.
 - The results of 4 reference chemicals should be within the acceptable range (Table 1).
- Be reproducible.

Data Interpretation Criteria

Table 5. : Positive and negative decision criteria

Positive	If the RPCMax is obtained that is equal to or exceeds 10% of the response of the positive control in at least two of two or two of three runs.
Negative	If the RPCMax fails to achieve at least 10% of the response of the positive control in two of two or two of three runs.

40. Data interpretation criteria are shown in Table 5. Positive results will be characterised by both the magnitude of the effect and the concentration at which the effect occurs. Expressing results as a concentration at which a 50% (PC50) or 10% (PC10) of PC values are reached accomplishes both of these goals. However, a test chemical is determined to be positive, if the maximum response induced by the test chemical (RPCMax) is equal to or exceeds 10% of the response of the PC in at least two of two or two of three runs, while a test chemical is considered negative if the RPCMax fails to achieve at least 10% of the response of the positive control in two of two or two of three runs.

455

OECD/OCDE

41. The calculations of PC10, PC50 and PCMax can be made by using a spreadsheet available with the Test Guideline on the OECD public website².

42. It should be sufficient to obtain PC10 or PC50 values at least twice. However, should the resulting base-line for data in the same concentration range show variability with an unacceptably high coefficient of variation (CV; %) the data may not be considered reliable and the source of the high variability should be identified. The CV of the raw data triplicates (*i.e.* luminescence intensity data) of the data points that are used for the calculation of PC10 should be less than 20%.

43. Meeting the acceptability criteria indicates the assay system is operating properly, but it does not ensure that any particular run will produce accurate data. Duplicating the results of the first run is the best insurance that accurate data were produced, see paragraphs 41 and 42.

44. Where more information is required in addition to the screening and prioritisation purposes of this TG for positive test compounds, particularly for PC10-PC49 chemicals, as well as chemicals suspected to over-stimulate luciferase, it can be confirmed that the observed luciferase-activity is solely an ER α -specific response, using an ER α antagonist (see Annex 3).

TEST REPORT

45. The test report should include the following information:

Test substance:

- identification data and CAS Number, if known;
- physical nature and purity;
- physicochemical properties relevant to the conduct of the study;
- stability of the test substance.

Solvent/Vehicle:

- characterisation (nature, supplier and lot);
- justification for choice of solvent/vehicle;
- solubility and stability of the test substance in solvent/vehicle, if known.

Cells:

- type and source of cells;
- number of cell passages;
- methods for maintenance of cell cultures.

Test conditions:

cytotoxicity data (and justifications for the method of choice) and solubility limitations should be reported, as well as:

² [<http://www.oecd.org/env/testguidelines>]

OECD/OCDE

455

- composition of media, CO₂ concentration;
- concentration of test chemical;
- volume of vehicle and test substance added;
- incubation temperature and humidity;
- duration of treatment;
- cell density during treatment;
- positive and negative reference chemicals;
- duration of treatment period;
- Luciferase assay reagents (Product name, supplier and lot);
- acceptability and data interpretation criteria.

Reliability check:

- Fold inductions for each assay plate.
- Actual logEC₅₀, logPC₅₀, logPC₁₀ and Hill slope values for concurrent reference chemicals.

Results:

- Raw and normalised data of luminescent signals;
- Concentration-response relationship, where possible;
- RPCMax, PCMax, PC₅₀ and/or PC₁₀ values, as appropriate;
- EC₅₀ values, if appropriate;
- Statistical analyses, if any, together with a measure of error (*e.g.* SD, CV or 95% confidence interval) and a description of how these values were obtained.

Discussion of the results***Conclusion***

LITERATURE

1. OECD (2007), OECD Conceptual Framework for the Testing and Assessment of Endocrine Disrupting Chemicals, in annex 2 of: OECD Guideline for the Testing of Chemicals No. 440, the Uterotrophic Bioassay in Rodents: A Short-term Screening Test for Oestrogenic Properties, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]
2. CERI (2006), Draft validation report of TA assay using HeLa-hER-9903 to detect estrogenic activity. [Available at: http://www.oecd.org/document/62/0,3343,en_2649_34377_2348606_1_1_1_1,00.html]
3. Takeyoshi, M., Yamasaki, K., Sawaki, M., Nakai, M., Noda, S. and Takatsuki, M. (2002), The efficacy of endocrine disruptor screening tests in detecting anti-estrogenic effects downstream of receptor-ligand interactions. *Toxicol. Lett.* 126, 91-98.
4. Jefferson, W.N., Padilla-Banks, E., Clark, G. and Newbold R. (2002), Assessing estrogenic activity of phytochemicals using transcriptional activation and immature mouse uterotrophic responses. *J. Chromat. B.*, 777, 179-189.
5. Sonneveld, E., Riteco, J.A., Jansen, H.J., Pieterse, B., Brouwer, A., Schoonen, W.G. and van der Burg, B. (2006), Comparison of *in vitro* and *in vivo* screening models for androgenic and estrogenic activities. *Toxicol. Sci.*, 89, 173-187.
6. Escande, A., Pillon, A., Servant, N., Cravedi, J.P., Larrea, F., Muhn, P., Nicolas, J.C., Cavailles, V. and Balaguer, P. (2006), Evaluation of ligand selectivity using reporter cell lines stably expressing estrogen receptor alpha or beta. *Biochem. Pharmacol.*, 71, 1459-1469.
7. Gray, L.E. Jr. (1998), Tiered screening and testing strategy for xenoestrogens and antiandrogens. *Toxicol. Lett.*, 102-103, 677-680.
8. EDSTAC (1998), Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) Final report. Available at: [<http://www.epa.gov/scipoly/oscpendo/pubs/edspoverview/finalrpt.htm>]
9. 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/methods/endocrine.htm#fineval>]
10. Escande, A., Pillon, A., Servant, N., Cravedi, J.P., Larrea, F., Muhn, P., Nicolas, J.C., Cavailles, V. and Balaguer, P. (2006), Evaluation of ligand selectivity using reporter cell lines stably expressing estrogen receptor alpha or beta. *Biochem. Pharmacol.*, 71, 1459-1469.
11. Kuiper, G.G., Lemmen, J.G., Carlsson, B., Corton, J.C., Safe, S.H., van der Saag, P.T., van der Burg, B. and Gustafsson, J.A. (1998), Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta. *Endocrinol.*, 139, 4252-4263.
12. Spaepen, M., Angulo, A.F., Marynen, P. and Cassiman, J.J. (1992), Detection of bacterial and mycoplasma contamination in cell cultures by polymerase chain reaction. *FEMS Microbiol. Lett.* 78(1), 89-94.
13. Kobayashi, H., Yamamoto, K., Eguchi, M., Kubo, M., Nakagami, S., Wakisaka, S., Kaizuka, M. and Ishii H (1995), Rapid detection of mycoplasma contamination in cell cultures by enzymatic detection of polymerase chain reaction (PCR) products. *J. Vet. Med. Sci.* 57(4), 769-771.
14. Dussurget, O. and Roulland-Dussoix D. (1994), Rapid, sensitive PCR-based detection of mycoplasmas in simulated samples of animal sera. *Appl. Environ. Microbiol.* 60(3), 953-9.
15. De Lean, A., Munson, P.J. and Rodbard D. (1978), Simultaneous analysis of families of sigmoidal curves: application to bioassay, radioligand assay, and physiological dose-response curves. *Am. J. Physiol.* 235, E97-E102.
16. OECD (2005), Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. OECD Series on Testing and Assessment No. 34, OECD, Paris. Available at: [<http://www.oecd.org/env/testguidelines>]

OECD/OCDE

455

ANNEX 1

Definitions and abbreviations

Agonist: A substance that binds to a specific receptor and triggers a response in the cell. It mimics the action of an endogenous ligand binds to the same receptor.

Antagonist: A type of receptor ligand or chemical that does not provoke a biological response itself upon binding to a receptor, but blocks or dampens agonist-mediated responses.

Anti-estrogenic activity, the capability of a chemical to suppress the action of 17 β -estradiol mediated through estrogen receptors.

CV: Coefficient of variation

Cytotoxicity: the harmful effects to cell structure or function ultimately causing cell death and can be a result of 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.

DCC-FBS: Dextran-coated charcoal treated fetal bovine serum.

DMSO: Dimethyl sulfoxide

E2: 17 β -estradiol

EC50 value, the concentration of agonist that provokes a response halfway between the baseline (Bottom) and maximum response (Top).

EE: 17 α -ethynyl estradiol

ER; Estrogen receptor

ERE: Estrogen Response Element

Estrogenic activity, the capability of a chemical to mimic 17 β -estradiol in its ability to bind to and activate estrogen receptors. hER α mediated specific estrogenic activity can be detected in this Test Guideline.

FBS: Fetal bovine serum

hER α : Human estrogen receptor alpha

MT: Metallothionein

OHT: 4-Hydroxytamoxifen

PC: Positive control (1 nM of E2)

PC10: the concentration of a test chemical at which the response in an agonist assay is 10% of the response induced by positive control (E2 at 1nM) in each plate

PC50: the concentration of a test chemical at which the response in an agonist assay is 50% of the response induced by positive control (E2 at 1nM) in each plate

PCMax: the concentration of a test chemical inducing the RPCMax

455

OECD/OCDE

RPCMax: maximum level of response induced by a test chemical, expressed as a percentage of the response induced by 1 nM E2 on the same plate

RT PCR: Real Time polymerase chain reaction

SD: Standard deviation

STTA: Stably Transfected Transcriptional Activation Assay.

TA: Transcriptional activation

Validation, a process based on scientifically sound principles by which the reliability and relevance of a particular test, approach, method, or process are established for a specific purpose. Reliability is defined as the extent of reproducibility of results from a test within and among laboratories over time, when performed using the same standardised protocol. The relevance of a test method describes the relationship between the test and the effect in the target species and whether the test method is meaningful and useful for a defined purpose, with the limitations identified. In brief, it is the extent to which the test method correctly measures or predicts the (biological) effect of interest, as appropriate (16).

VC (Vehicle control): The vehicle that is used to dissolve test and control chemicals is tested solely as vehicle without dissolved chemical.

OECD/OCDE

455

ANNEX 2

False positives: Assessment of non-receptor mediated luminescence signals

1. False positives might be generated by non-ER-mediated activation of the luciferase gene, or direct activation of the gene product or unrelated fluorescence. Such effects are indicated by an incomplete or unusual dose-response curve. If such effects are suspected, the effect of an ER antagonist (*e.g.* 4-hydroxytamoxifen (OHT) at non-toxic concentration) on the response should be examined. The pure antagonist ICI 128780 may not be suitable for this purpose as a sufficient concentration of ICI 128780 may decrease the VC value, and this will affect the data analysis.

2. To ensure validity of this approach, the following needs to be tested in the same plate:

- Agonistic activity of the unknown chemical with / without 10 µM of OHT
- VC (in triplicate)
- OHT (in triplicate)
- 1 nM of E2 (in triplicate) as agonist PC
- 1 nM of E2 + OHT (in triplicate)

3. ***Data interpretation criteria***

Note: All wells should be treated with the same concentration of the vehicle.

- If the agonistic activity of the unknown chemical is NOT affected by the treatment with ER antagonist, it is classified as “Negative”.
- If the agonistic activity of the unknown chemical is completely inhibited, apply the decision criteria.
- If the agonistic activity at the lowest concentration is equal to, or is exceeding, PC10 response the unknown chemical is inhibited equal to or exceeding PC10 response. The difference in the responses between the non-treated and treated wells with the ER antagonist is calculated and this difference should be considered as the true response and should be used for the calculation of the appropriate parameters to enable a classification decision to be made.

4. ***Data analysis***

Check the performance standard.

Check the CV between wells treated under the same conditions.

1. Calculate the mean of the VC
2. Subtract the mean of VC from each well value **not** treated with OHT
3. Calculate the mean of OHT
4. Subtract the mean of the VC from each well value treated with OHT
5. Calculate the mean of the PC
6. Calculate the relative transcriptional activity of all other wells relative to the PC.

455

OECD/OCDE

ANNEX 3

Preparation of Serum treated with Dextran Coated Charcoal (DCC)

1. The treatment of serum with dextran-coated charcoal (DCC) is a general method for removal of estrogenic compounds from serum that is added to cell medium, in order to exclude the biased response associated with residual estrogens in serum. 500 mL of fetal bovine serum (FBS) can be treated by this procedure.

Components

2. The following materials and equipment will be required:

Materials

Activated charcoal
Dextran
Magnesium chloride hexahydrate ($\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$)
Sucrose
1 M HEPES buffer solution (pH 7.4)
Ultrapure water produced from a filter system

Equipment

Autoclaved glass container (size should be adjusted as appropriate)
General Laboratory Centrifuge (that can set temperature at 4°C)

Procedure

3. The following procedure is adjusted for the use of 50 mL centrifuge tubes:

[Day-1] Prepare dextran- coated charcoal suspension with 1 L of ultrapure water containing 1.5 mM of MgCl_2 , 0.25 M sucrose, 2.5 g of charcoal, 0.25 g dextran and 5 mM of HEPES and stir it at 4°C, overnight.

[Day-2] Dispense the suspension in 50 mL centrifuge tubes and centrifuge at 10000 rpm at 4°C for 10 minutes. Remove the supernatant and store half of the charcoal sediment at 4°C for the use on Day-3. Suspend the other half of the charcoal with FBS that has been gently thawed to avoid precipitation, and heat-inactivated at 56°C for 30 minutes, then transfer into an autoclaved glass container such as an Erlenmeyer flask. Stir this suspension gently at 4°C, overnight.

[Day-3] Dispense the suspension with FBS into centrifuge tubes for centrifugation at 10000 rpm at 4°C for 10 minutes. Collect FBS and transfer into the new charcoal sediment prepared and stored on Day-2. Suspend the charcoal sediment and stir this suspension gently in an autoclaved glass container at 4°C, overnight.

[Day-4] Dispense the suspension for centrifugation at 10000 rpm at 4°C for 10 minutes and sterilise the supernatant by filtration through 0.2 µm sterile filter. This DCC treated FBS should be stored at -20°C and can be used for up a year.

Appendix F4

Weight of Evidence Guidance: Evaluating Results of EDSP Tier 1 Screening to Identify Candidate Chemicals for Tier 2 Testing (Draft for Public Comment)

This page intentionally left blank

**WEIGHT OF EVIDENCE GUIDANCE: EVALUATING RESULTS OF EDSP TIER 1
SCREENING TO IDENTIFY CANDIDATE CHEMICALS FOR TIER 2 TESTING**

DRAFT FOR PUBLIC COMMENT

1. PURPOSE AND SCOPE OF DOCUMENT

The purpose of this document is to set forth some of the general principles, criteria, and considerations EPA generally believes to be relevant using a weight-of-evidence (WoE) approach to evaluate data submitted as part of EPA's two-tiered paradigm for screening and testing chemicals for endocrine activity under the Endocrine Disruptor Screening Program (EDSP). This paper was developed by EPA to provide guidance to EPA staff and managers who will be reviewing data submitted in response Orders for Tier 1 screening that began October 29, 2009 under the Endocrine Disruptor Screening Program (EDSP). Additionally, outside parties submitting data may be interested to know how the results from Tier 1 screening are being evaluated. This paper provides general guidance and is not binding on either EPA or any outside parties. The use of language such as "will," "is," "may," "can," or "should" in this paper does not connote any requirement for either EPA or any outside parties. As such, EPA may depart from the guidance where circumstances warrant and without prior notice.

A WoE evaluation is a process where potentially relevant studies are judged in a professional manner for quality. Thereafter, a summary statement is developed indicating the potential effects of the compound, the mode of action (MOA), and other relevant information. It is not a process that simply involves tallying the number of positive and negative results within and among studies. Critical assessment of an entire body of available data is taken into account for consistency, coherence, and biological plausibility (e.g., see USEPA, 2002 & 2005). Principles articulated in this document are equally applicable to a WoE evaluation of data from individual assays with multiple endpoints, as well as across the suite of assays in the Tier 1 screening battery. In addition, these principles would be generally relevant to the review of other scientifically relevant information (OSRI) submitted in response to test orders that request OSRI to be considered in lieu of designated screening assays in the Tier 1 battery. Most of the principles presented in this document are not unique to chemicals with potential endocrine activity but are commonly used for WoE evaluations conducted by EPA (USEPA 1991; 1992; 1996; 2002; 2005). The criteria discussed in this document are based on EPA's experience in developing and applying risk assessment guidelines involving cancer, reproductive and developmental toxicity, and ecological toxicity. Important considerations include the use of expert judgment formed through the scientific process, current understanding of endocrine mechanisms of toxicity, and knowledge of other fields of toxicology (e.g., developmental, reproductive, neurological and immunological toxicology, and toxicokinetics). This document provides a transparent scientific approach for broadly evaluating Tier 1 screening data to determine if additional Tier 2 testing is necessary.

This document also is expected to comply with the provision in the Office of Management and Budget Terms of Clearance for the Information Collection Request for the first list of chemicals to be screened under the EDSP and direction in the House Appropriations Committee for the Interior and Environment FY 2010 report (HR 2996, [http://thomas.loc.gov/cgi-bin/cpquery/R?cp111:FLD010:@1\(hr180\)\)](http://thomas.loc.gov/cgi-bin/cpquery/R?cp111:FLD010:@1(hr180))) that directed EPA to:

“develop and publish criteria for evaluating the results of Tier 1 screening and determining whether a chemical should undergo Tier 2 analysis within one year of enactment. The process should allow for public input.”

2. HISTORICAL BACKGROUND

This section provides an abbreviated overview of EPA's EDSP. A more detailed history of the program can be found at its website (<http://www.epa.gov/endo>) and in other documents or websites referenced herein.

In 1996, Congress amended section 408 of the Federal Food, Drug, and Cosmetic Act (FFDCA) to require EPA to:

“develop a screening program, using appropriate validated test systems and other scientifically relevant information, to determine whether certain substances may have an effect in humans that is similar to an effect produced by a naturally occurring estrogen, or other such endocrine effect as the Administrator may designate” [21 U.S.C. 346a(p)]. (<http://www.epa.gov/pesticides/regulating/laws/fqpa/>)

Fundamental to the EDSP is a two-tiered approach involving a battery of Tier 1 screening assays and individual Tier 2 tests designed to identify and characterize chemicals with the potential to interact with the estrogen, androgen, and/or thyroid (E, A and/or T) hormonal systems. Tier 1 consists of a battery of complementary *in vitro* and short term *in vivo* assays designed to maximize sensitivity for detecting interactions with E, A, and/or T. Interactions with E, A, and/or T are evaluated using a range of key endpoints involving the MOA [e.g., receptor binding and gene transcription, steroidogenesis, hypothalamic-pituitary-gonadal (HPG) and –thyroid (HPT) axes] across gender and various taxa (e.g., rodents, amphibians, and fish) as indicated in Table 1. The diversity in endocrine endpoints within and among the Tier 1 screening assays is expected to provide corroborating information and support a WoE evaluation to yield a decision as to whether or not the chemical under evaluation requires additional testing in Tier 2.

Tier 2 testing consists of a group of individual *in vivo* tests designed to include males and females with an intact hypothalamic-pituitary-gonadal axis, multiple pathways of exposure and life-stages, and various taxa to further identify and characterize chemical-induced interactions with E, A, and/or T for risk assessment. Although the endocrine system is included, Tier 2 tests are designed to quantify dose-response relationships in a larger context of toxicity and potential adversity that may involve other biological

systems (e.g., neurological, immunological, hepatic, renal, and cardiovascular) to be used for risk assessment. While the Tier 2 mammalian two-generation reproductive toxicity test is considered valid, other Tier 2 tests are at various stages in the validation process (Table 2).

Table 1: EDSP Tier 1 battery of screening assays and complementary modes of action*.

Screening Assays	*Modes of Action							
	Receptor Binding				Steroidogenesis		HPG Axis	HPT Axis
	E	Anti-E	A	Anti-A	E	A		
<i>In vitro</i>								
ER Binding	■	■						
ER α Transcriptional Activation	■							
AR Binding			■	■				
Steroidogenesis H295R					■	■		
Aromatase Recombinant					■			
<i>In vivo</i>								
Uterotrophic	■							
Hershberger			■	■				
Pubertal Male			■	■		■	■	■
Pubertal Female	■	■			■		■	■
Fish Short-term Reproduction (male & female)	■	■	■	■	■	■	■	
Amphibian Metamorphosis								■

*A mode of action is defined as a sequence of key events and processes, starting with interaction of an agent with a cell, proceeding through operational and anatomical changes, and resulting in an adverse outcome (USEPA, 2005). These assays encompass certain key events within a mode of action (e.g., receptor binding) as well as certain pathways (e.g., steroidogenesis) through which a chemical can interact with the E, A, or T hormonal systems.

Table 2. EDSP Tier 2 tests.

Mammalian two-generation reproductive toxicity test
*Avian two-generation toxicity test - Japanese quail two-generation toxicity test
*Amphibian growth and development test - Larval Amphibian Growth and Development Assay
*Fish multigeneration test - Medaka Multigeneration Test
*Invertebrate two-generation test - Mysid two-generation test

*Proposed Tier 2 test currently at various stages of the validation process.

Most of the proposed Tier 1 screening assays completed the validation process in 2008 and another in 2009. Subsequent to review [Federal Register Notice of January 24, 2008 (73 FR 4216)] by the Federal Insecticide Fungicide and Rodenticide Act Scientific Advisory Panel (FIFRA SAP) and, based on a final SAP report (SAP, 2008), EPA accepted the current EDSP Tier 1 battery (Table 1). Availability of test guidelines for each of the 11 screening assays in Tier 1 was published in a Federal Register Notice October 21, 2009 (74 FR 54416).

3. TESTING METHODS AND ENDPOINTS

3.1. EDSP Tier 1 screening assays

The basis for the endpoints in each of the *in vitro* and *in vivo* Tier 1 screening assays has been described (<http://www.epa.gov/scipoly/oscpendo/pubs/assayvalidation/tier1battery.htm>) in Integrated Summary Reports associated with the validation process (see “peer review of individual assays” at website), summarized in a Technical Review Document for SAP review of the battery (see “external review” at website) and listed in EPA test guidelines for each of the assays (see “Test guidelines for Series 890” at website). The endpoints associated with each assay were determined through the validation process to be sensitive and specific to detect interaction with the E, A, or T hormonal systems.

For chemicals having estrogen- or androgen-like activity, the *in vitro* receptor binding assays provide information on the potential binding characteristics of a compound. Comparatively, the *in vitro* estrogen receptor transcriptional activation or steroidogenic assays are cell-based and provide mechanistic as well as some functional information on induction/inhibition of gene transcription or degree of steroid hormone production, respectively.

In vivo assays integrate effects in a whole organism and provide apical as well as mechanistic information from one or multiple endpoints within an assay. Agonistic or antagonistic E-, A-, or T-dependent changes can be detected in association with reproductive development (e.g., vaginal opening and preputial separation), organ weights (e.g., ovaries, uterus, testes, prostate, and thyroid), and corresponding

histology of target organs. Mechanistic information such as thyroid hormones measured in the pubertal and amphibian assays and vitellogenin in the fish assay can be correlated with apical information within the same assays. The use of gonadectomized rats in the uterotrophic and Hershberger assays provides mechanistic information on specific estrogen- (*i.e.*, uterus) or androgen- (*i.e.*, prostate, seminal vesicles, levator ani-bulbocavernosus muscle, Cowper's glands, and glans penis) dependent target organs. As with the *in vitro* assays, the results of the uterotrophic or Hershberger assays are limited to E and A receptor function. On the other hand, results involving intact animals (*i.e.*, female and male pubertal, amphibian, and fish assays) generally provide more systems information, since more sites of action are involved along the HPG and HPT axes that target some of the same E-, A-, or T-dependent endpoints.

3.2. Other Scientifically Relevant Information (OSRI)

EPA's approach to the submission and use of OSRI as part of the EDSP has previously been described by EPA and is available in the Federal Docket Management System (EPA-HQ-OPPT-2007-1080-0032). In general, OSRI consists of data from assays that satisfy the same function as EDSP Tier 1 assays or may include data that are indicative of a potential consequence or adverse effect resulting from a chemical-induced change in the E, A, and/or T hormonal systems. Hence, interference of endocrine function may come from effects measured in standard toxicity test guidelines or other comparable toxicity studies. Typically, OSRI-types of studies are not designed to provide definitive information on the modes or mechanisms of toxicity, but are generally focused on measured adverse effects (*e.g.*, ability to become pregnant, duration of gestation, signs of difficult or prolonged parturition, sex ratio, or feminization or masculinization of offspring, number of pups, stillbirths, gross pathology, and histopathology of the vagina, uterus, ovaries, testis, epididymis, seminal vesicles, prostate, and thyroid) representing permanent changes with organizational or functional consequences. These studies may also encompass a range of life-stages (*e.g.*, two-generation reproductive study), treatment durations and doses, and provide information generated by relevant routes of exposure.

4. WEIGHT-OF-EVIDENCE APPROACH

In evaluating whether additional testing is warranted in Tier 2, EPA anticipates that the following key questions would typically be considered as part of EPA's WoE approach:

- Do existing data provide adequate evidence to conclude whether there is a potential for the chemical to interfere with the normal function of the E, A, and/or T hormonal systems?
- If the data indicate a potential to interact with those specific endocrine systems, which hormone system is impacted (E, A, and/or T)?

Determination of whether the evidence suggests the substance is or is not a candidate for Tier 2 testing is based on evaluation of all relevant data, including any Tier 1 results. This WoE analysis is conducted on a case-by-case basis by assessing all of the individual lines of evidence (Section 4.1) and performing an integrated analysis of the data (Section 4.2).

4.1. Analysis of individual studies

In any WoE analysis, a full evaluation of each relevant study is conducted and documented. In general, the evaluation of individual studies includes characterization of the following:

- Nature of the effect(s) seen in the study(ies) (e.g., were the effects seen in *in vitro* and/or *in vivo* assays; were the effects persistent or transient changes; were the effects molecular/biochemical changes or adverse outcomes);
- specificity and sensitivity of the effect(s);
- dose- and time- dependent changes, if available;
- potency of responses and magnitude or severity of changes; and
- consistency and relationship of the different effects seen within a study.

Both statistical and biological significance of the observed effects are relevant in evaluating study results. In general, the results of relevant studies are assumed to be indicative of interactions with the endocrine system unless data are available that demonstrate otherwise (e.g., evidence that the effect is not the consequence of an interaction with the endocrine system but a consequence of excessive toxicity to a non-endocrine system). To aid in determining the level of confidence in a study, the strengths of the study as well as any attendant limitations and uncertainties shall be considered.

4.2. Integrated analysis of data

Weight-of-evidence assessments must be conducted on a case-by-case basis. EPA's WoE analyses will generally include consideration of the information as follows:

- Quality of data and the extent to which effects can be replicated within a laboratory and across different laboratories;
- strengths and limitations of *in vitro* and *in vivo* results;
- number and type of effects induced and potency, magnitude, and severity of effects;
- consistency, pattern, range, and interrelationships of effects observed across studies, species, strains, and sexes;
- conditions under which effects occur (e.g., dose, route, duration); and
- understanding of MOA and biological plausibility of responses.

These considerations are part of evaluating the evidence as a whole and determining whether or not a chemical has the potential to interact with the endocrine

system via E, A, and/or T hormonal pathways. In examining the balance of positive and negative results, the relative sensitivity and specificity of the measured endpoints would also be considered. Tier 1 *in vitro* screens can provide some insight into MOA. In general, however, Tier 1 *in vivo* results would carry greater weight than *in vitro* results because *in vitro* assays inherently lack physiological conditions associated with whole tissues or organs and, therefore, have nil or very limited ability to represent metabolic processes and pathways leading to endocrine disruption. The relationship between endpoints and their impact on normal endocrine function would also typically be significant factors in this determination. Concordant effects found in multiple interrelated endpoints generally imply a compromise in endocrine function, in contrast to isolated or discordant effects. Totality of the evidence is evaluated to determine whether such effects can potentially occur across taxa. Generally, consistent positive or negative effects across studies and taxa increase confidence in the determination of whether or not a chemical has the potential to interact with the endocrine system. Additionally, if marginal or weak relationships exist with regard to dose, severity, magnitude, and/or incidence, consideration of other available information may also be appropriate in determining whether further testing in Tier 2 is warranted. This could include consideration of other critical effects (non-endocrine), dose response, what is understood about the underlying basis (*i.e.*, toxicity MOA) of these critical effects (*i.e.*, non-endocrine or an endocrine MOA not covered by Tier 1) and their human relevance, and potential for exposure. Other supportive evidence may also be used in the WoE evaluation including pertinent data on related chemicals, metabolism or toxicokinetics, and results of computational models.

4.3. Summary and conclusions of WoE approach

A summary of the WoE analysis is expected to transparently state and explain conclusions. It should explain the selection of certain studies or effects as the key basis for conclusions. In general, this characterization should be limited to the most significant and relevant data, conclusions, and uncertainties.

Summary statements for a WoE analysis should generally address key elements as follows:

- Each E, A, and T pathway, including species, gender, and life stages;
- uncertainties and the extent these uncertainties impact the conclusions;
- relative weight placed on studies and effects (*i.e.*, points at which choices are made of critical effects or studies and why); and
- inconsistent or conflicting results.

If Tier 2 testing is indicated, *i.e.*, effects are seen that are mediated through the E, A, and/or T hormonal systems, to the extent permitted by the available data, the summary should address any potential impact of the results of Tier 2 tests to risk assessment for that chemical and provide rationale for any conclusions. This may include, to the extent supported by the available data, conclusions regarding the species in which additional

testing is warranted and the likelihood that the effect may occur at a lower dose than effects seen in existing studies.

REFERENCES

EDSTAC (1998) Endocrine Disruptor Screening and Testing Advisory Committee, Final Report, Volume I-II.

<http://www.epa.gov/scipoly/oscpendo/pubs/edspoverview/finalrpt.htm>.

SAP (2008) A set of scientific issues being considered by the Environmental Protection Agency regarding: The Endocrine Disruptor Screening Program (EDSP) Proposed Tier 1 Screening Battery. SAP Minutes No. 2008-03.

<http://www.epa.gov/scipoly/sap/meetings/2008/march/minutes2008-03-25.pdf>.

USEPA (U.S. Environmental Protection Agency) (1991) Guidelines for developmental toxicity risk assessment. Federal Register 56(234):63798-63826. Available from:

<http://cfpub.epa.gov/ncea/raf/recordisplay.cfm?deid=23162>.

USEPA (U.S. Environmental Protection Agency) (1996) Guidelines for reproductive toxicity risk assessment. Federal Register 61(212):56274-56322. Available from:

<http://cfpub.epa.gov/ncea/raf/recordisplay.cfm?deid=2838>.

USEPA (U.S. Environmental Protection Agency) (1998) Guidelines for neurotoxicity risk assessment. Federal Register 63(93):26926-26954. Available from:

<http://cfpub.epa.gov/ncea/raf/recordisplay.cfm?deid=12479>.

USEPA (U.S. Environmental Protection Agency) (2002) A Review of Review of the Reference Dose and Reference Concentration Processes. Risk Assessment Forum, Washington EPA/630/P-02/002F December 2002 Final Report

<http://www.epa.gov/raf/publications/pdfs/rfd-final.pdf>

USEPA (U.S. Environmental Protection Agency) (2005) Guidelines for carcinogen risk assessment Risk Assessment Forum, Washington, DC. EPA/630/P-03/001F. Available from: [http://www.epa.gov/raf/publications/pdfs/CANCER_GUIDELINES_FINAL_3-25-](http://www.epa.gov/raf/publications/pdfs/CANCER_GUIDELINES_FINAL_3-25-05.PDF)

[05.PDF](http://www.epa.gov/)<http://www.epa.gov/>

This page intentionally left blank

National Toxicology Program
P.O. Box 12233
Research Triangle Park, NC 27709



U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
National Institutes of Health



NIEHS
National Institute of
Environmental Health Sciences



NTP
National Toxicology Program
U.S. Department of Health and Human Services

