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 *In Vitro* ER Binding Test Methods C-67

3.3.3 Substances Used to Evaluate BG1Luc ER TA Concordance with the *In Vivo* 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 *In Vitro* 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 *In Vivo* Uterotrophic Assay Data C-123

Table 5-15 Concordance of BG1Luc ER TA Agonist Classification and *In Vivo* 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	17β-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 Disrupters 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	*Federal Register*
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)
<u>NICEATM</u>
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

<u>ECVAM</u>
Elise Grignard, PhD
Susanne Bremer, PhD

<u>JaCVAM</u>
Hajime Kojima, PhD – Director, JaCVAM
Atsushi Ono, PhD – National Institute of Health Sciences

<u>KoCVAM</u>
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. 285*l*-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 Vandenbergh 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 Sprankle; 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, 17β-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 [EC_{50}] and half-maximal inhibitory concentration [IC_{50}] values) are usually obtained for positive

[1] The CERI-STTA assay (OECD 2009) uses the hERα-HeLa-9903 human cervical cancer cell line to detect estrogenic agonist activity mediated through human estrogen receptor alpha (hERα).

results. EC_{50} and IC_{50} 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 EC_{50} 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 IC_{50} 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 CERI-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 Disrupters 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 ERα and a reporter gene have been stably transfected. HeLa-9903 cells do not express endogenous ERα or ERβ. 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 ERα and ERβ (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 ERβ 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 ERβ-specific substances, it is known that a number of substance types (e.g., isoflavones) are ERβ selective (Chrzan and Bradford 2007; Escande et al. 2006; Kuiper et al. 1998; Mohler et al. 2010), with more potent responses through ERβ than through ERα (Chrzan and Bradford 2007; Kuiper et al. 1998). The BG1Luc ER TA test method, using cells that express both ERα and ERβ, allows for the potential detection of a wider range of substances than test methods that use cells expressing only the ERα 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:

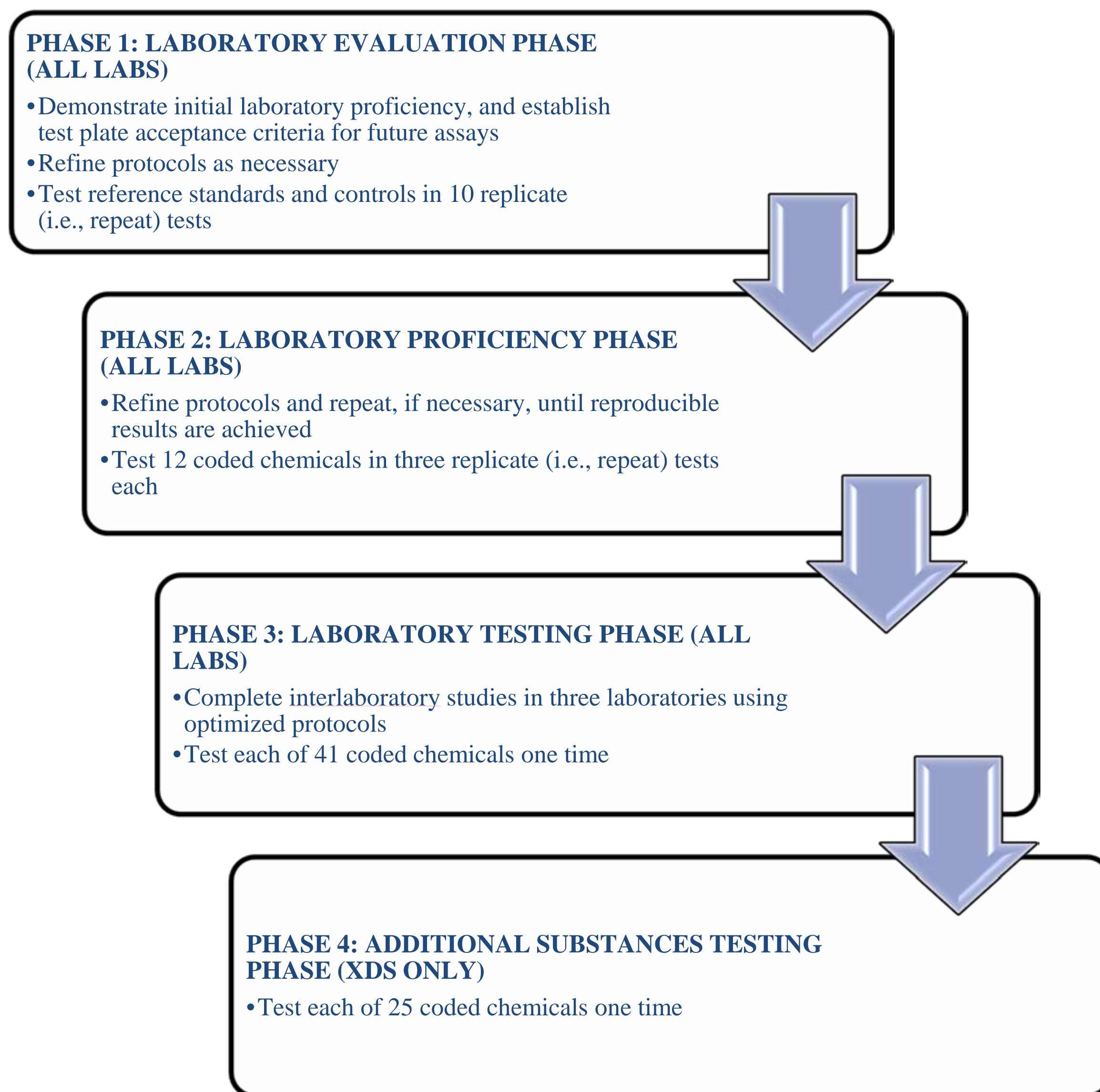
- Ensure that the study adhered to the principles stated in OECD Guidance Document Number 34 for prospective validation studies (OECD 2005)
- Develop a Statement of Work for the laboratories
- Determine timelines and deliverables
- Arrange for purchasing, coding, and distributing test substances to the 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 ERα and ERβ. 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 µM 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 CO_2-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 (EC_{50}) or the half-maximal inhibitory concentration (IC_{50}). Separate protocols are used to identify substances that possess ER agonist or antagonist activity, although the two protocols share most major components.

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 EC_{50} 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% (IC_{50}) 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.
- EC_{50} or IC_{50} 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 ERα and ERβ (Wong and Matsumura 2006), although ERα 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, pipetters, repeat pipetters, pipette tips
- Sterile, disposable tissue culture plasticware (e.g., 25-cm^2 and 75-cm^2 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% CO_2/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-cm^2 tissue culture flasks, and incubated at 37°C ± 1°C, 90% ± 5% humidity, and 5% ± 1% CO_2/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, 17β-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 EC_{50} 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 = 17β-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 EC_{50} two to three orders of magnitude higher than E2 ($EC_{50} = 3 \text{ x } 10^{-12}$ 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 EC_{50} 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 ($EC_{50} = 6$ µM), approximately six orders of magnitude higher than E2 ($EC_{50} = 3 \text{ x } 10^{-12}$ 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 IC_{50} 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 IC_{50}.

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 IC_{50} at least three orders of magnitude higher than the reference antagonist (2003a), Ral (IC_{50}= 2.24 x 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 IC_{50} = 4.3 x 10^{-7} M, which was consistent with the single literature reference for this compound (reported IC_{50} = ~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 μM. In addition, NICEATM–ICCVAM could not identify in the literature any substances classified as positive for ER antagonism with an IC_{50} > 10 μM. Therefore, the SMT established 10 μM as the upper limit of utility for determining antagonist activity in the BG1Luc ER TA test method. Because the 10 μM would preclude the use of flavone as a weak antagonist control (IC_{50} = 15 μM), 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 μM 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 μM.

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 μM in the antagonist assay consistently produce false positive responses in this loss-of-function assay**.** Consequently, the SMT established 10 μM 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, EC_{50} values from serial dilutions of BPA derived from replicates using outside wells were compared to EC_{50} values derived from replicates using inside wells. The comparisons showed no significant differences between EC_{50} 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 EC_{50} 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 EC_{50} 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 ± 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 EC_{50} 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 IC_{50} 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 IC_{50} 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 IC_{50} value must be within 2.5 times the SD of the historical mean IC_{50} 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-*O*-Tetradecanoylphorbol-13-acetate	16561-29-8	Hydrocarbon (Cyclic)	Laboratory Chemical	>99.5	LC Laboratories
17α-Estradiol	57-91-0	Steroid	Pharmaceutical, Veterinary Agent	99.5	Sigma-Aldrich Corporation
17α-Ethinyl estradiol	57-63-6	Steroid	Pharmaceutical, Veterinary Agent	≥98.0	Sigma-Aldrich Corporation
17ß-Estradiol	50-28-2	Steroid	Pharmaceutical, Veterinary Agent	98.0	Sigma-Aldrich Corporation
17ß-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-*sec*-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-*tert*-Octylphenol	140-66-9	Phenol	Chemical Intermediate, Pharmaceutical Intermediate	99.3	Chem Service, Inc.
5α-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-*n*-butyl phthalate	84-74-2	Ester, Phthalic Acid	Cosmetic Ingredient, Industrial Chemical, Plasticizer	≥98.0	City Chemical LLC
Dibenzo[*a*,*h*]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.
o,p'-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
p-n-Nonylphenol	104-40-5	Phenol	Chemical Intermediate	99.6	Alfa Aesar GmbH
p,p'- Methoxychlor	72-43-5	Hydrocarbon (Halogenated)	Pesticide, Veterinary Agent	99.1	Chem Service, Inc.
p,p'-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.
Phenolpthalin	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 Science℠ (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., IC_{50}, EC_{50})

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-*O*-Tetradecanoylphorbol-13-acetate	16561-29-8	PN (NT)	PN (NT)	PN (NT)	NT	NT
17α Estradiol	57-91-0	POS (10/10)	PN (0/1)	POS (15/15)	POS	POS (NT/+)
17α Ethinyl estradiol	57-63-6	POS (21/21)	NEG (0/9)	POS (32/32)	POS	POS (+/+)
17ß Estradiol	50-28-2	POS (226/226)	PN (0/1)	POS (160/160)	POS	NT
17ß-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-*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)	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-*tert*-Octylphenol	140-66-9	POS (20/23)	PN (NT)	POS (20/20)	POS	POS (NT/+)
5α-Dihydrotestosterone	521-18-6	POS (15/17)	NEG (0/3)	POS (17/18)	NT	POS (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]
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)	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-*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)	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

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]
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
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
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
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 EC_{50} 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}
17ß-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}

Substance	CASRN	ICCVAM Consensus Classification[a]	Mean EC_{50} (M)[b]
Diethylstilbestrol	56-53-1	POS	1.29×10^{-7}
Estrone	53-16-7	POS	8.33×10^{-8}
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
meso-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}
o,p'-DDT	789-02-6	POS	1.67×10^{-4}
p-n-Nonylphenol	104-40-5	POS	1.59×10^{-6}
p,p'-Methoxychlor	72-43-5	POS	1.56×10^{-4}
p,p'-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); EC_{50} = 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 CERI, and uterotrophic response.

[b] Mean EC_{50} 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]
17α–Ethinyl estradiol	57-63-6	NEG	NA
4-Hydroxytamoxifen	68047-06-3	POS	1.93×10^{-8}
5α-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-*n*-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
o,p'-DDT	789-02-6	NEG	NA
p-n-Nonylphenol	104-40-5	NEG	NA
p,p'-DDE	72-55-9	NEG	NA
Progesterone	57-83-0	NEG	NA
Raloxifene HCl	82640-04-8	POS	6.23×10^{-8}
Resveratrol	501-36-0	NEG	NA
Tamoxifen	10540-29-1	POS	1.26×10^{-6}

Abbreviations: CASRN = CAS Registry Number (American Chemical Society); IC_{50} = half-maximal inhibitory concentration; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; M = molar; NA = not applicable; NEG = negative; POS = positive.

[a] Estrogenic activity based on a literature review for effects on ER binding, ER TA based on CERI, and uterotrophic response.

[b] Mean IC_{50} 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 hERα- 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 hERα-HeLa-9903 cell line is derived from a human cervical tumor with two stably inserted constructs: (1) the hERα 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 inbibition, 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	17ß-Estradiol	50-28-2	Steroid	Pharmaceutical, Veterinary Agent	POS
1	*p,p*'-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	17α-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	*o,p*'-DDT	789-02-6	Hydrocarbon (Halogenated)	Pesticide	POS
2b	*p*-n-Nonylphenol	104-40-5	Phenol	Chemical Intermediate	POS
2b	Vinclozolin	50471-44-8	Heterocyclic Compound	Fungicide	PP
3	12-*O*-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	17α-Estradiol	57-91-0	Steroid	Pharmaceutical, Veterinary Agent	POS
3	17ß-Estradiol	50-28-2	Steroid	Pharmaceutical, Veterinary Agent	POS
3	2-*sec*-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-tert-Octylphenol	140-66-9	Phenol	Chemical Intermediate, Pharmaceutical Intermediate	POS
3	5α-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-*n*-butyl phthalate	84-74-2	Ester, Phthalic Acid	Cosmetic Ingredient, Industrial Chemical, Plasticizer	PP
3	Dibenzo[*a*,*h*]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	*meso*-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	*p,p*'-Methoxychlor	72-43-5	Hydrocarbon (Halogenated)	Pesticide, Veterinary Agent	POS
3	*p,p*'-DDE	72-55-9	Hydrocarbon (Halogenated)	Pesticide Intermediate	POS
3	Phenobarbital	50-06-6	Heterocyclic Compound, Pyrimidine	Pharmaceutical, Veterinary Agent	NEG
3	Phenolpthalin	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	17ß-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	Fluoxymestrone	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 CERI, 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	17ß-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[*a*,*h*]anthracene	53-70-3	Polycyclic Compound	Laboratory Chemical, Natural Product	PP
2a	*p*-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	*o,p'*-DDT	789-02-6	Hydrocarbon (Halogenated)	Pesticide	NEG
2b	Resveratrol	501-36-0	Hydrocarbon (Cyclic)	Natural Product	NEG
3	12-*O*-Tetradecanoylphorbol-13-acetate	16561-29-8	Hydrocarbon (Cyclic)	Laboratory Chemical	PN
3	17α-Estradiol	57-91-0	Steroid	Pharmaceutical, Veterinary Agent	PN
3	17α-Ethinyl estradiol	57-63-6	Steroid	Pharmaceutical, Veterinary Agent	NEG
3	17ß-Estradiol	50-28-2	Steroid	Pharmaceutical, Veterinary Agent	PN
3	2-*sec*-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-*tert*-Octylphenol	140-66-9	Phenol	Chemical Intermediate, Pharmaceutical Intermediate	PN
3	5α-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-*n*-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	*meso*-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	*p,p*'-DDE	72-55-9	Hydrocarbon (Halogenated)	Pesticide Intermediate	NEG
3	*p,p*'- 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	Phenolpthalin	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	17ß-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	Spironolactone	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, 5α-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, 5α-dihydrotestosterone, or testosterone. Therefore, ECVAM obtained the required EU regulatory permissions for the importation of 4-androstenedione, 5α-dihydrotestosterone, and testosterone, which were subsequently procured by the NTPSI and exported to the ECVAM laboratory accordingly. JaCVAM was able to procure 4-androstenedione, 5α-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 EC_{50}, Ral reference standard IC_{50}, 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 EC_{50} 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 EC_{50} 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 EC_{50} 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: EC_{50} = 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 ≤ 17%). Variability remained low throughout subsequent phases of the validation study (CV ≤ 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 IC_{50} 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 $\geq$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 IC_{50} 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 IC_{50} 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; IC_{50} = 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 = 17β-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-androstendione 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*-bultyphenol 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∝-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}
o,p' -DDT	2b	354.5	100	2.82×10^{-4}	100	2.82×10^{-4}	10	2.82×10^{-5}
p-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-*O* -Tetradecanoylphorbol-13-acetate	3	616.8	1000	1.62×10^{-3}	100	1.62×10^{-4}	10	1.62×10^{-5}
17∝-Estradiol	3	272.4	1000	3.67×10^{-3}	1000	3.67×10^{-3}	10	3.67×10^{-5}
17ß-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}
5∝-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}
17ß-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[*a,h*]anthracene	2a	278.4	10	3.59×10^{-5}	10	3.59×10^{-5}	10	3.59×10^{-5}
p -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}
o,p'-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-*O*-Tetradecanoylphorbol-13-acetate	3	616.8	1000	1.62×10^{-3}	1000	1.62×10^{-3}	10	1.62×10^{-5}
17∝-Estradiol	3	272.4	1000	3.67×10^{-3}	100	3.67×10^{-4}	10	3.67×10^{-5}
17∝-Ethinyl estradiol	3	296.4	100	3.37×10^{-4}	10	3.37×10^{-5}	10	3.37×10^{-5}
17ß-Estradiol	3	272.4	1000	3.67×10^{-3}	100	3.67×10^{-4}	10	3.67×10^{-5}
2-*sec*-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-*tert*-Octylphenol	3	206.3	1000	4.85×10^{-3}	100	4.85×10^{-4}	10	4.85×10^{-5}
5∝-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-*n*-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}
meso-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}
p,p'-DDE	3	318.0	1000	3.14×10^{-3}	100	3.14×10^{-4}	10	3.14×10^{-5}
p,p'-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}
17ß-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
Fluoxymestrone	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-∝ 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)
p-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[*a*,*h*]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-*n*-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)
p,p'-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∝-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-∝ 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-ß 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)
meso-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)
p,p'- 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-*tert*-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; EC_{50} = 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 EC_{50} 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[*a*,*h*]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)
p-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)
o,*p*' -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]	Classifi-cation[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-∝ 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-*n*-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]
p,p'-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∝-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∝ Estradiol	3	XDS	4.26×10^{-6}	-	-	1/2	P (1/1)
		ECVAM	-	-	-	0/1	N (1/1)
		Hiyoshi	-	-	-	0/1	I (1/1)
17ß 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)
meso-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]	Classifi-cation[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)
p,*p*'- 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-*tert*-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); IC_{50} = 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 IC_{50} 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
17ß-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
Fluoxymestrone	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: EC_{50} = half-maximal effective concentration; I = inadequate (positive or negative classification could not be determined because of poor-quality data); IC_{50} = half-maximal inhibitory concentration; M = molar; N = negative; NC = not calculated; P = positive.

[a] EC_{50} 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] IC_{50} 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

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

a = positive in 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

Concordance = ([a+d]/[a+b+c+d])

Sensitivity = (a/[a+c])

Specificity = (d/[b+d])

False positive rate = (b/[b+d])

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∝-Dihydrotestosterone
- Flutamide
- Procymidone
- Resveratol
- 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∝-Estradiol	57-91-0	POS	POS	POS (1/1)	POS (3/3)	POS (2/2)
17∝-Ethinyl estradiol	57-63-6	POS	POS	POS (3/3)	POS (3/3)	POS (3/3)
17ß-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∝-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
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*'-DDE	72-55-9	POS	I	I (1/1)	I (1/1)	NEG (1/1)
p,*p*'-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
17α–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)
5α-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)
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)

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 ERα agonist activity of test substances is currently the only ER TA test method accepted by regulatory agencies. This test system utilizes the hERα-HeLa-9903 cell line, which is derived from a human cervical tumor, with two stably inserted constructs: the hERα 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]
17α-Estradiol	57-91-0	POS	POS	POS
17α-Ethinyl estradiol	57-63-6	POS	POS	POS
17ß-Estradiol	50-28-2	POS	POS	POS
4-Cumylphenol	599-64-4	POS	POS	POS
4-*tert*-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
p-n-Nonylphenol	104-40-5	POS	POS	NEG
p,p'-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

		CERI-STTA Classification		
		Positive	**Negative**	**Total**
ICCVAM Consensus Classification	**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 EC_{50} and IC_{50} 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., EC_{50} and IC_{50} values) are usually obtained for positive results. EC_{50} and IC_{50} 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 EC_{50} and **Table 5-11** for IC_{50} comparisons. Regression analyses of these data are presented in **Figures 5-1** and **5-2**, respectively.

Based on EC_{50} values obtained for 26 substances, the correlation coefficient between the log EC_{50} 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 EC_{50} 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 IC_{50} values obtained for three substances, the correlation coefficient between the log IC_{50} 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 EC_{50} Values for Substances Used to Generate EC_{50} Linear Regression

Substance	BG1Luc ER TA Median EC_{50} (M)	ICCVAM Reference Data Median EC_{50} (M)
17α-Estradiol	3.02×10^{-10}	5.20×10^{-09}
17α-Ethinyl estradiol	7.09×10^{-12}	5.20×10^{-11}
17β-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-*tert*-Octylphenol	2.08×10^{-08}	1.00×10^{-07}

Substance	BG1Luc ER TA Median EC_{50} (M)	ICCVAM Reference Data Median EC_{50} (M)
5α-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}
meso-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}
o,p'-DDT	4.22×10^{-07}	1.69×10^{-06}
p-n-Nonylphenol	2.50×10^{-06}	3.60×10^{-07}
p,p'-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}

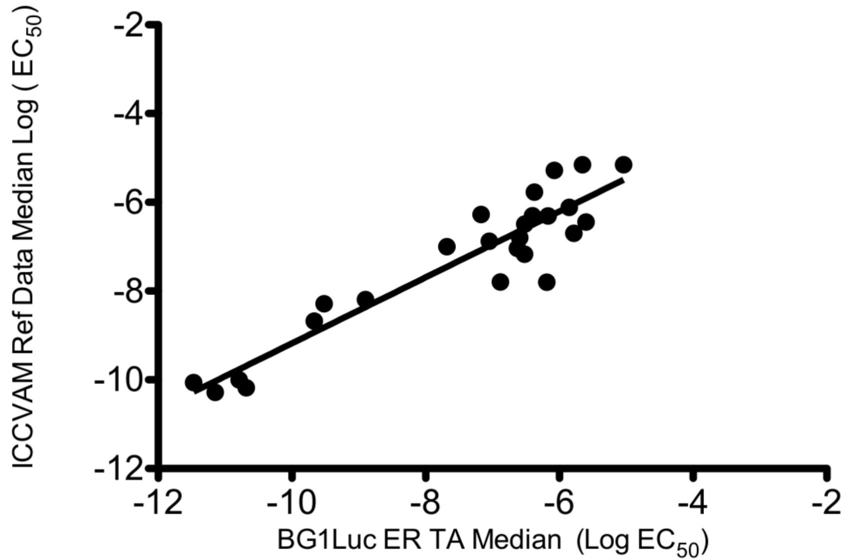
Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; EC_{50} = half-maximal effective concentration; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; M = molar.

Table 5-11 Median IC_{50} Values for Substances Used to Generate IC_{50} 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; IC_{50} = half-maximal inhibitory concentration; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; M = molar.

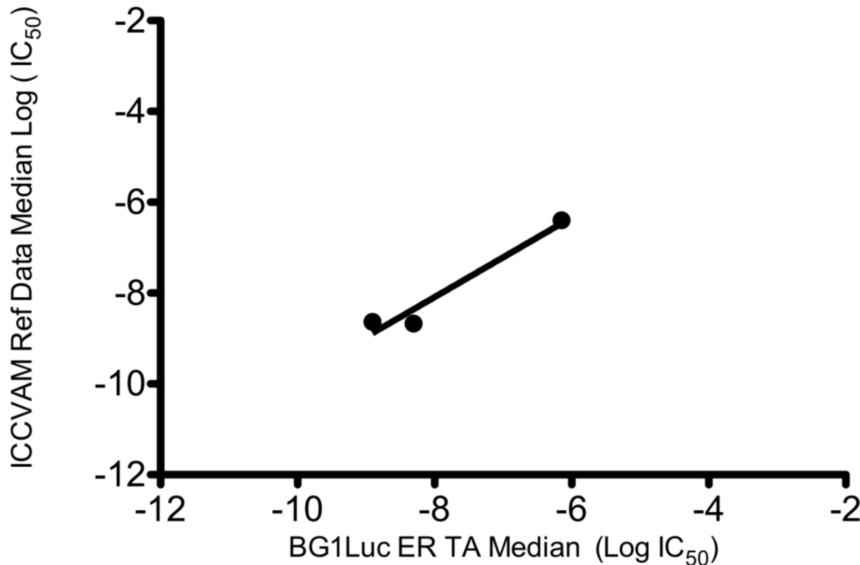
Figure 5-1 Relationship of EC_{50} Values Obtained in the BG1Luc ER TA Test Method and EC_{50} Values from ICCVAM Reference Data



Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; EC_{50} = half-maximal effective concentration; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods.

Each point in this figure represents a median EC_{50} value obtained in the BG1Luc ER TA test method compared with the median ICCVAM EC_{50} value (from the literature reference data updated in 2010, discussed in **Section 3** and provided in **Annex N**).

Figure 5- 2 Relationship of IC_{50} Values Obtained in the BG1Luc ER TA Test Method and IC_{50} Values from ICCVAM Reference Data



Abbreviations: BG1Luc ER TA = LUMI-CELL BG1Luc4E2 ER TA test method; IC_{50} = half-maximal inhibitory concentration; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods.

Each point in this figure represents a median IC_{50} value obtained in the BG1Luc ER TA test method compared with the median ICCVAM IC_{50} 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 x 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)
17α-Estradiol	57-91-0	POS	I	POS	POS
17α-Ethinyl estradiol	57-63-6	POS	NEG	POS	POS
17ß-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
meso-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
o,p'-DDT	789-02-6	POS	NEG	POS	POS
p-n-Nonylphenol	104-40-5	POS	NEG	POS	POS
p,p'-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]
17α Estradiol	57-91-0	POS	POS	NT	POS
17α Ethinyl estradiol	57-63-6	POS	POS	POS	POS
4-*tert*-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
o,p'-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**
***In Vivo* 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.)

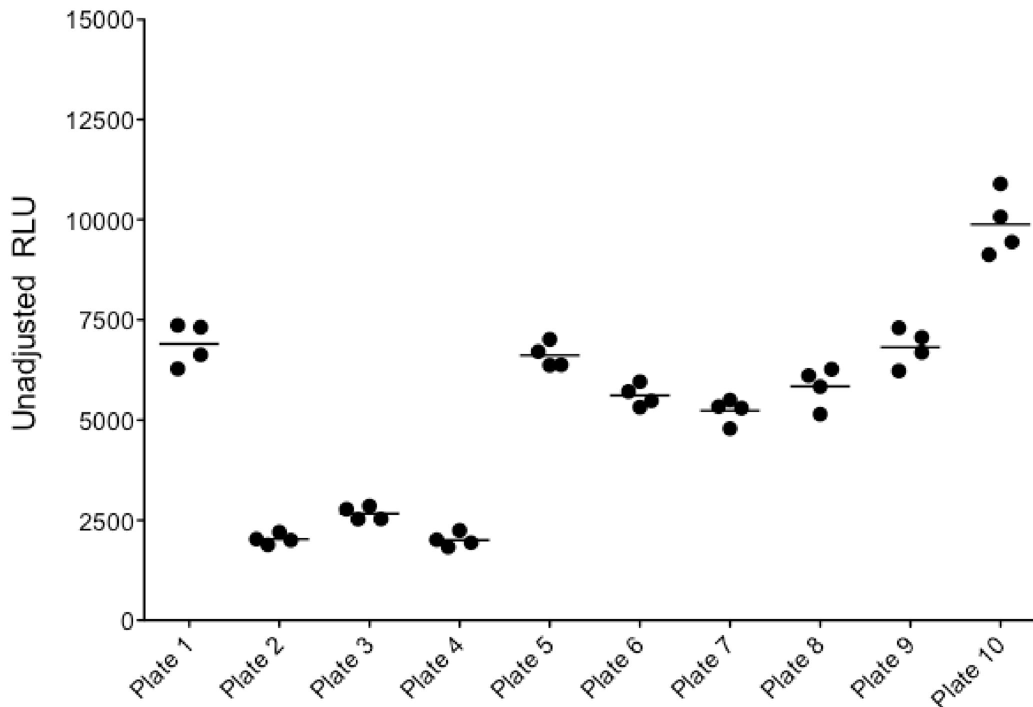
Table 6-1 Agonist Within-Plate DMSO Control Data

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.

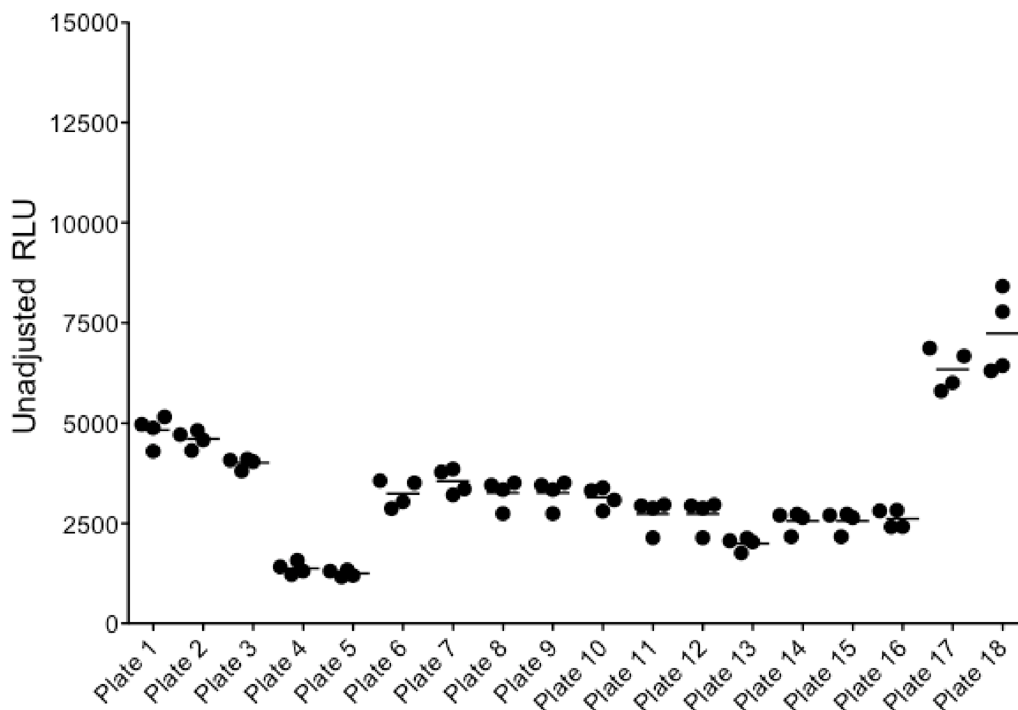
Figures 6-1 through **6-3** show the within-plate agonist DMSO control RLU values for Phase 1 of the validation study as examples of the low variability for this parameter. As discussed above, within-plate CVs were low throughout the validation study.

Figure 6-1 Agonist DMSO Control Within-Plate RLU Values During Phase 1 at XDS



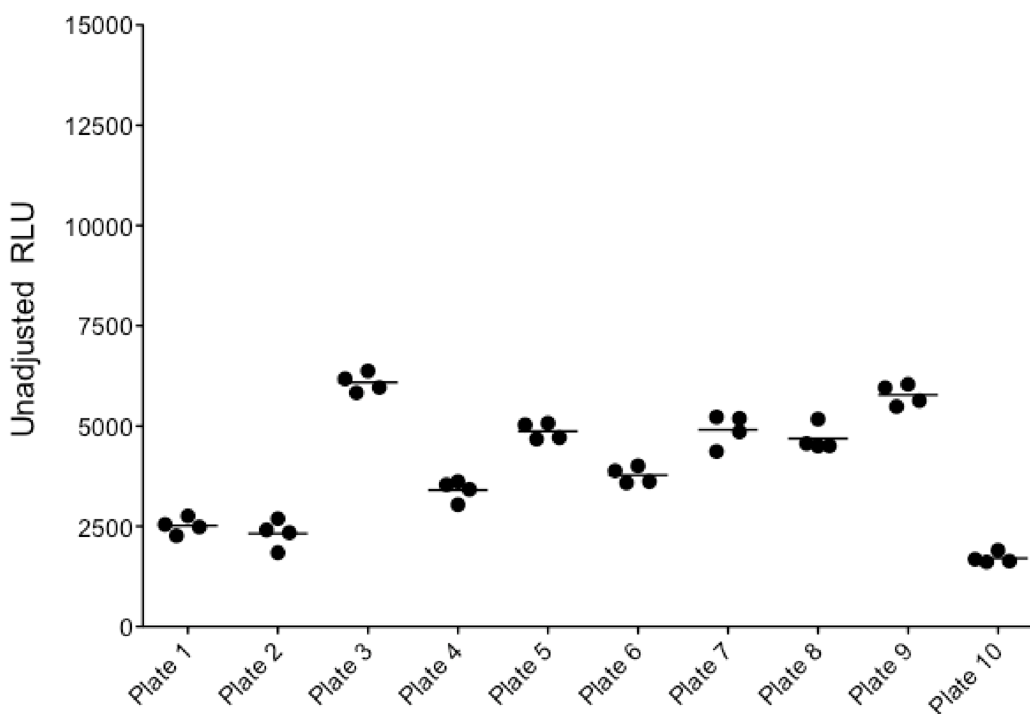
Abbreviations: DMSO = dimethyl sulfoxide; RLU = relative light unit.
Each point represents the non-normalized DMSO value for a single well in a 96-well plate.
Within-plate DMSO variance at XDS during Phase 1 was fairly low, with coefficients of variation ranging from 5% to 9%.

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 EC_{50} and Methoxychlor Control

Although E2 reference standard EC_{50} 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 EC_{50} and Methoxychlor Control Values

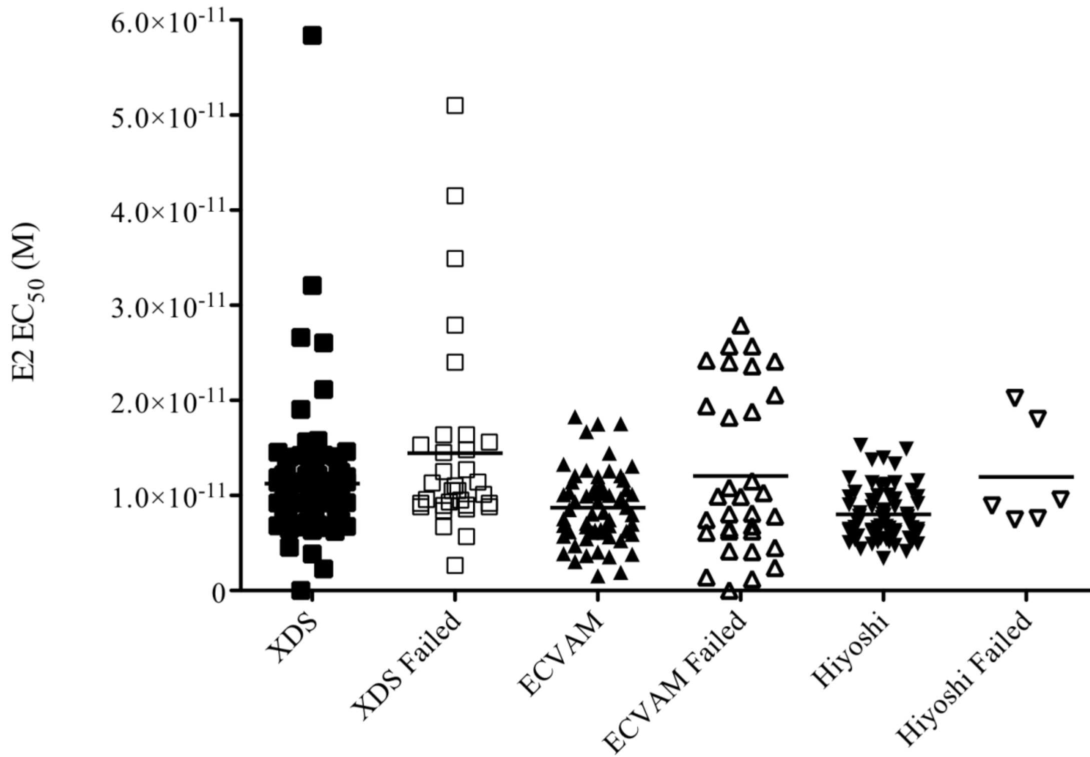
Laboratory	Mean	SD	N
E2 Reference Standard EC_{50} (M)			
XDS	1.1×10^{-11}	6.7×10^{-12}	93
ECVAM	1.1×10^{-11}	1.9×10^{-11}	60
Hiyoshi	8.0×10^{-12}	2.8×10^{-12}	65
Methoxychlor (RLU)			
XDS	6075	1283	93
ECVAM	6246	1609	60
Hiyoshi	8029	1233	65

Abbreviations: EC_{50} = 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 EC_{50} 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 EC_{50} 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 EC_{50} at XDS, all outlier values among the parameters evaluated were associated with these failed plates.

Figure 6-4 Agonist E2 Reference Standard EC_{50} Values



Abbreviations: EC_{50} = half-maximal effective concentration; ECVAM = European Centre for the Validation of Alternative Methods; M = molar; XDS = Xenobiotic Detection Systems, Inc.

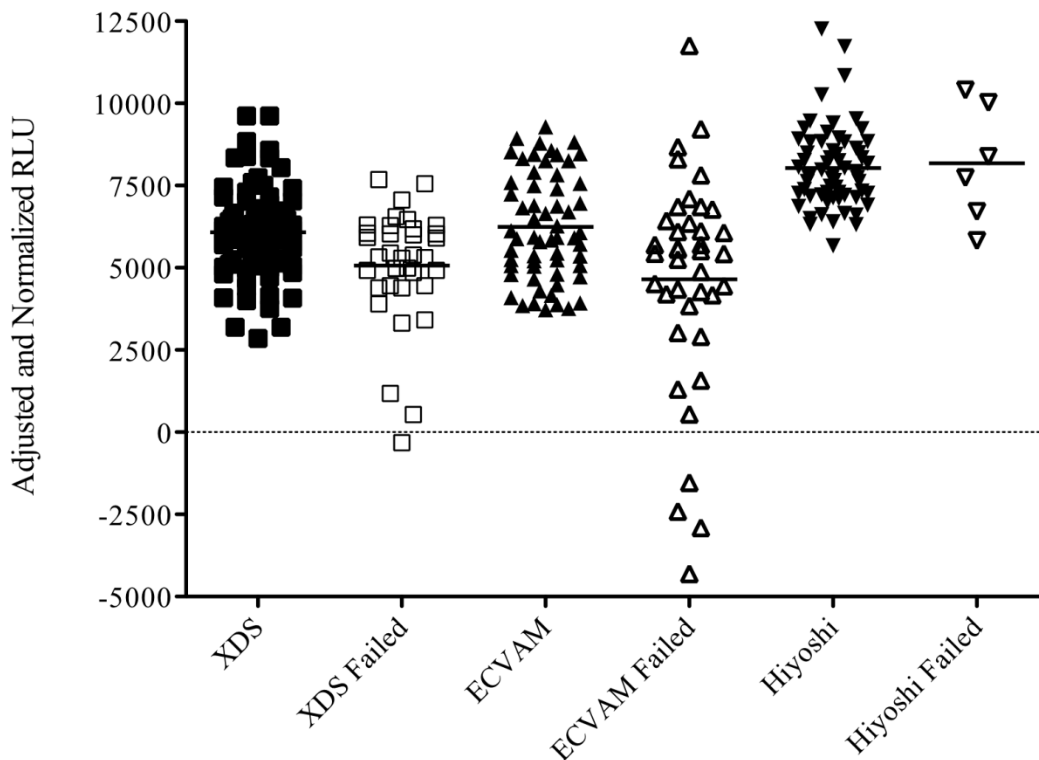
Each point represents a single plate.

An EC_{50} value (1.18×10^{-9} M) from one experiment that failed acceptance criteria at XDS was excluded from the graph to minimize scale distortion.

EC_{50} 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 EC_{50} 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.)

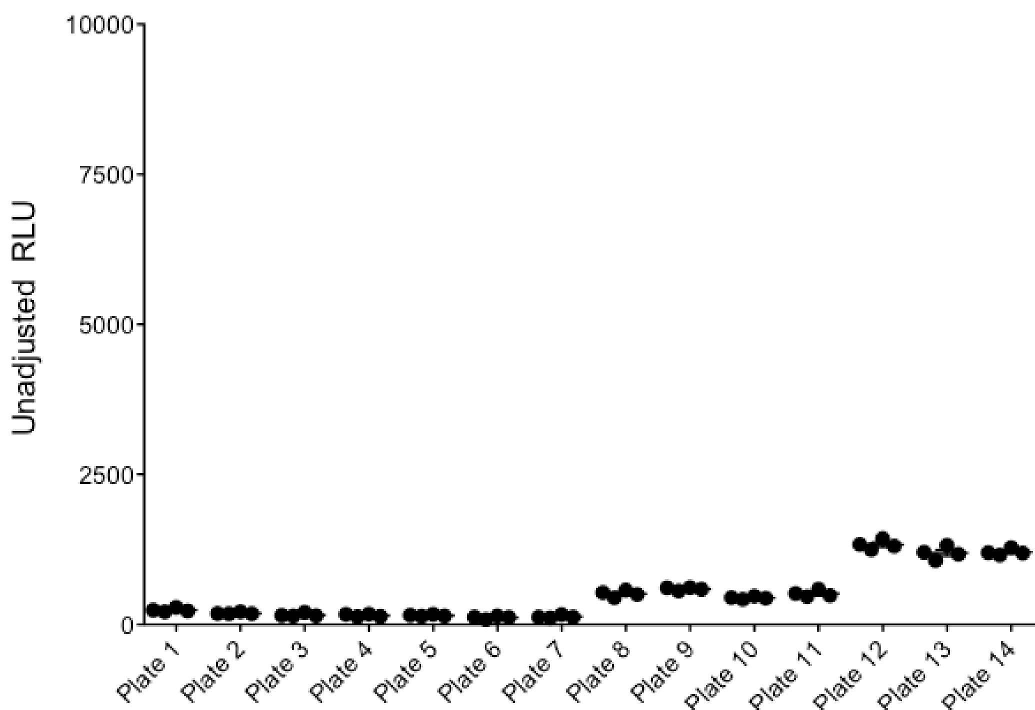
Table 6-4 Antagonist DMSO Control Values

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.

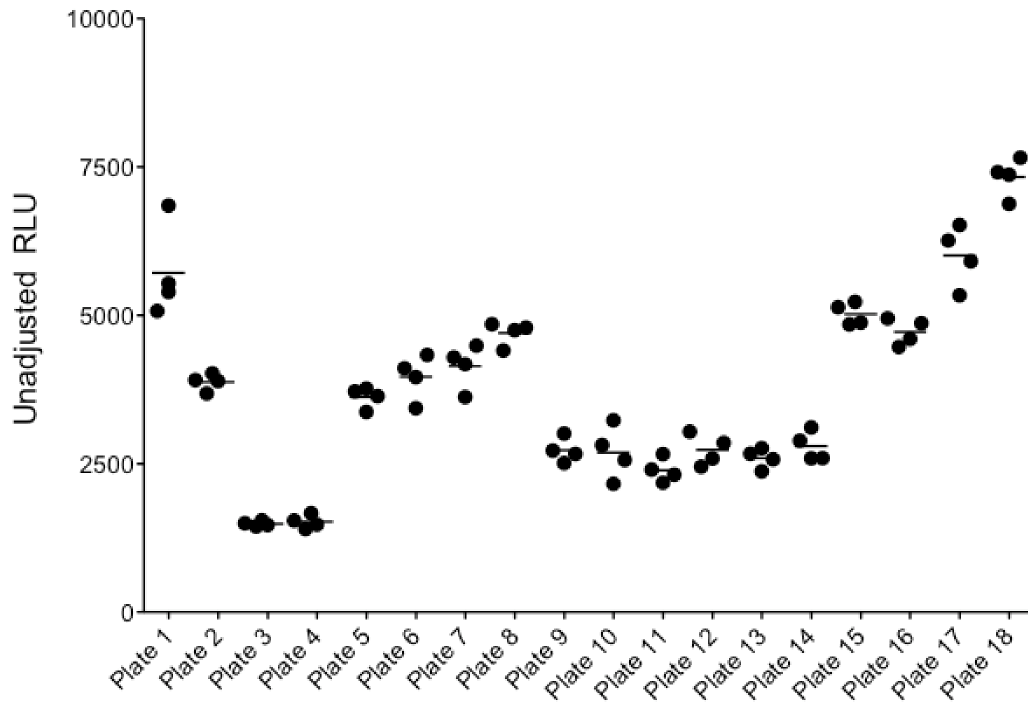
Figures 6-6 through **6-8** show the within-plate agonist DMSO control RLU values for Phase 1 of the validation study as examples of the low variability for this parameter. As discussed above, within-plate CVs were low throughout the validation study.

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

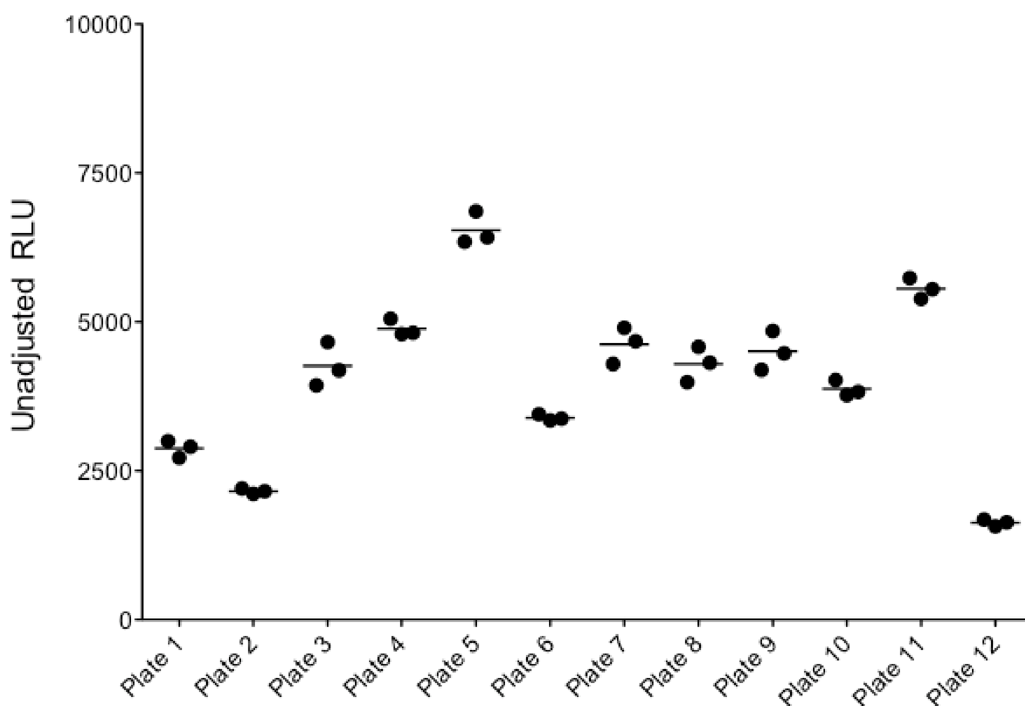


Abbreviations: DMSO = dimethyl sulfoxide; ECVAM = European Centre for the Validation of Alternative Methods; RLU = relative light unit.

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 IC_{50} and Flavone Control Values

Although Ral reference standard IC_{50} 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 IC_{50} and Flavone Control Values

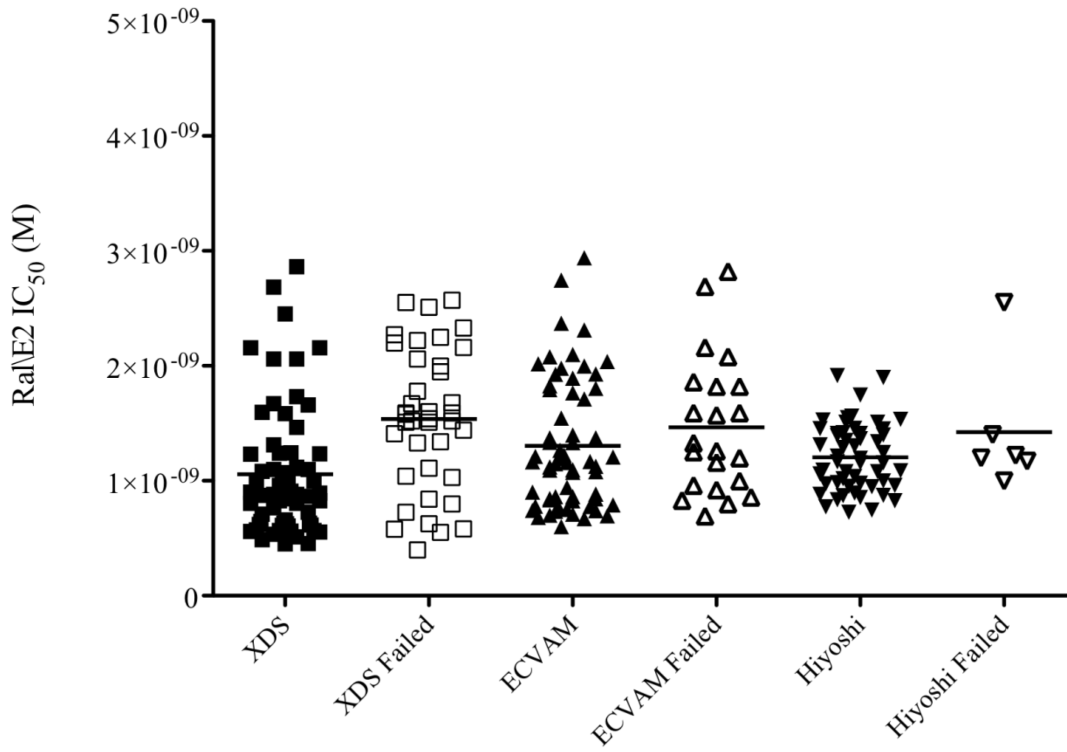
Laboratory	Mean	SD	N
Raloxifene Reference Standard IC_{50} (M)			
XDS	1.1×10^{-9}	5.6×10^{-10}	79
ECVAM	1.3×10^{-9}	5.6×10^{-10}	62
Hiyoshi	1.2×10^{-9}	2.9×10^{-10}	53
Flavone (RLU)			
XDS	3774	1366	79
ECVAM	599	468	62
Hiyoshi	873	772	53

Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; IC_{50} = 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 IC_{50} 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 IC_{50}, 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.

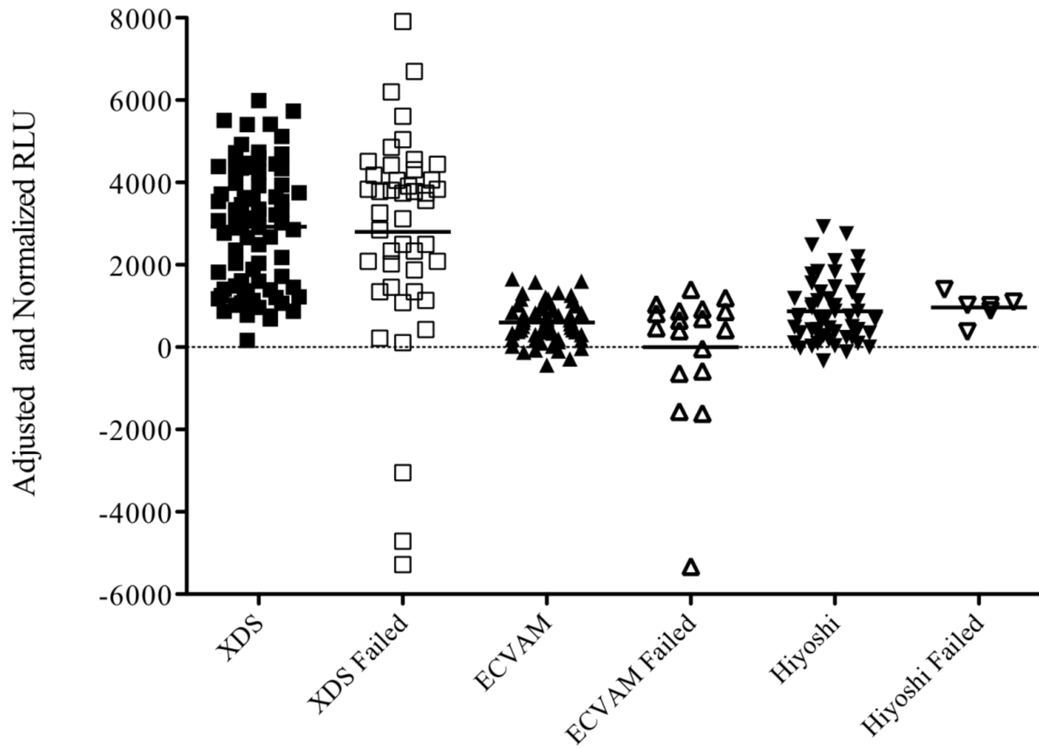
Figure 6-9 Antagonist Raloxifene Reference Standard IC_{50} Values



Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; IC_{50} = 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



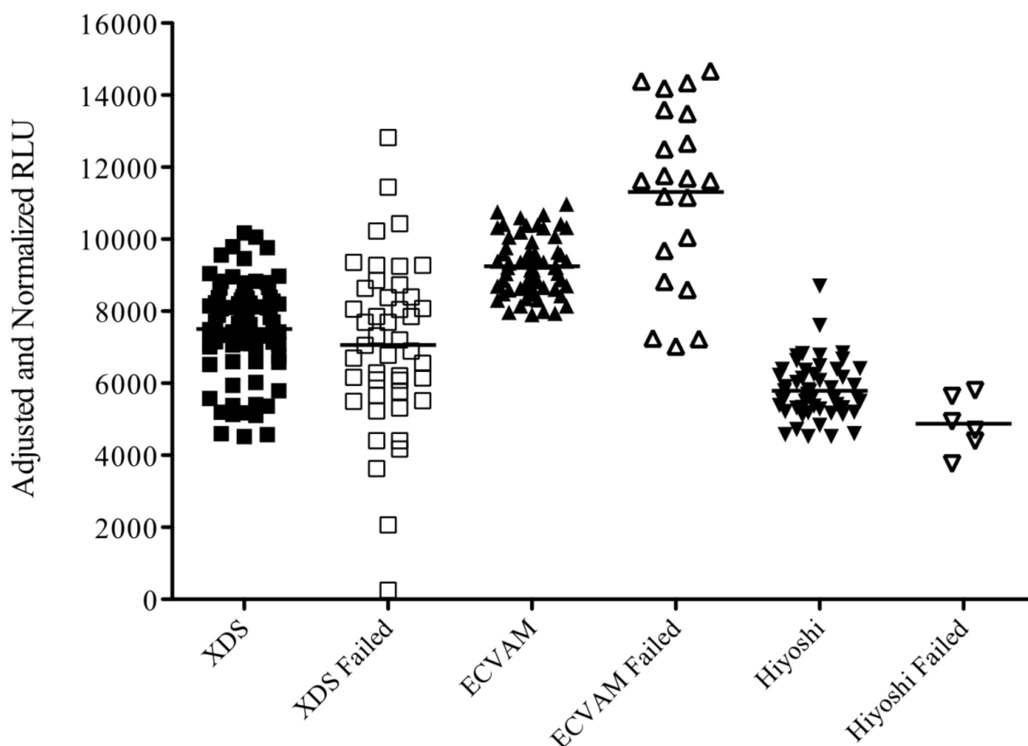
Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; RLU = relative light unit; XDS = Xenobiotic Detection Systems, Inc.

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.

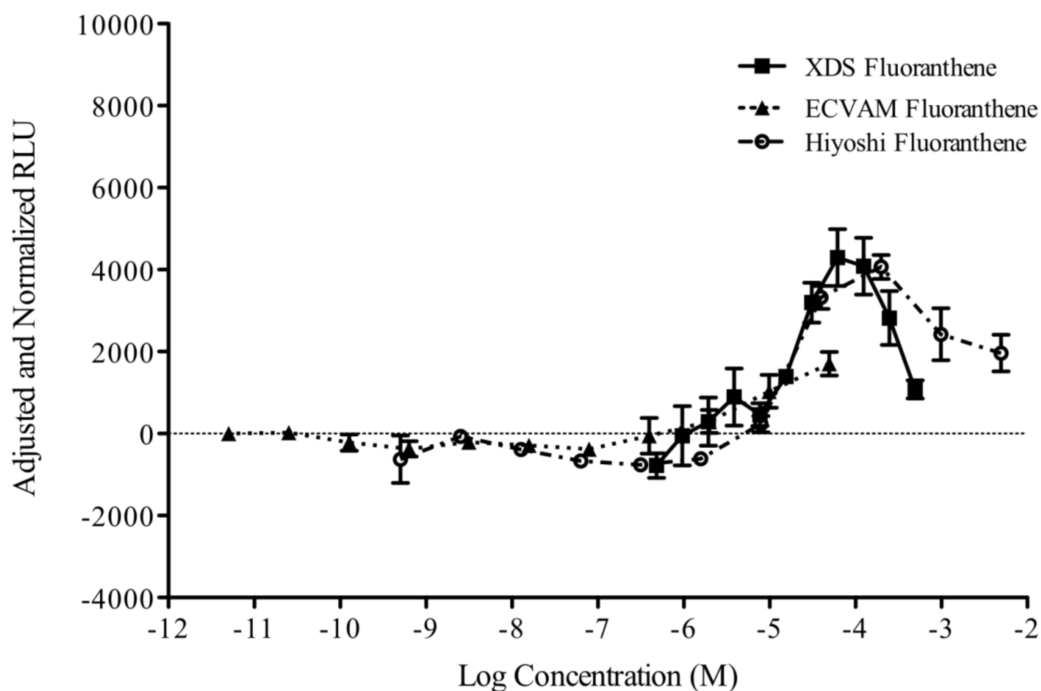
++I indicates that the substance was classified as positive at two of three laboratories but had inadequate data in the third.

--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**).

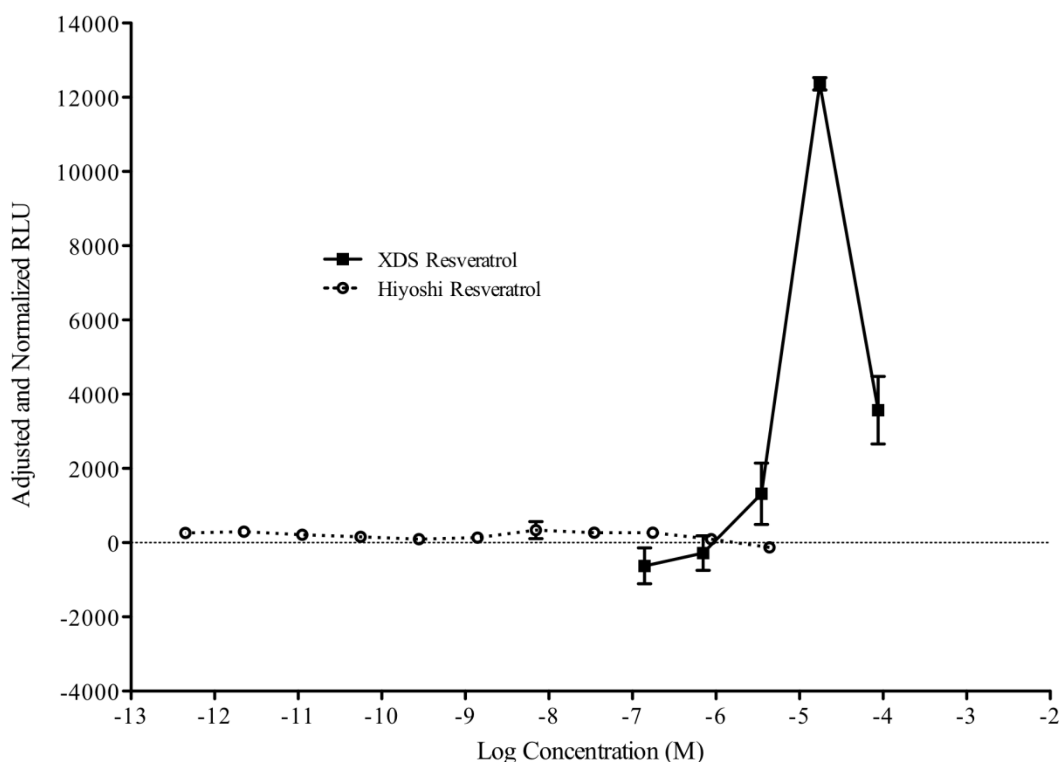
Figure 6-12 Fluoranthene Results at All Three Laboratories: Impact of Differences in Solubility on Comprehensive Test Results



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 17α-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 EC_{50} and Met RLU control values and (2) antagonist Ral IC_{50} 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/IC_{50} 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 EC_{50} and/or Met control acceptance criteria were equivalent to the ER TA activities for plates that passed acceptance criteria. Antagonist plates that failed IC_{50} 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 = 17β-estradiol; EC_{50} = 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[*a,h*]anthracene	XDS	POS (3/3)	POS (2/2)	NA	NA
	ECVAM	POS (3/3)	NA	NA	NA
	Hiyoshi	POS (3/3)	NA	NA	NA
p-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; IC_{50} = 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 EC_{50} and Met RLU control value acceptance criteria. Antagonist plates that passed all acceptance criteria are compared to those that failed the Ral IC_{50} and flavone control RLU value acceptance criteria**.** The quantitative evaluation compared EC_{50} values that could be calculated for bisphenol A, bisphenol B, and diethylstilbestrol at XDS and ECVAM, and the IC_{50} values that could be calculated for tamoxifen at XDS for plates that passed and failed acceptance criteria. Results indicate that agonist substance EC_{50} values from plates that failed EC_{50} and/or methoxychlor control acceptance criteria and tamoxifen IC_{50} values from plates that failed IC_{50} 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 EC_{50}/IC_{50} 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	1.5×10^{-1}	5.7×10^{-2}	3	3.1×10^{-1}	8.8×10^{-2}	0.11

Abbreviations: BPA = bisphenol A; BPB = bisphenol B; DES = diethylstilbestrol; E2 = 17ß-estradiol; EC_{50} = half–maximal effective concentration; ECVAM = European Centre for the Validation of Alternative Methods; IC_{50} = 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 EC_{50} values (µg/mL) except for TAM, which is expressed in IC_{50} values (µg/mL).

[a] $p > 0.05$ indicates that EC_{50} or IC_{50} values are not significantly different.

Based on this evaluation, it was determined that test plate acceptance criteria based on agonist E2 EC_{50} 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 IC_{50} 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.

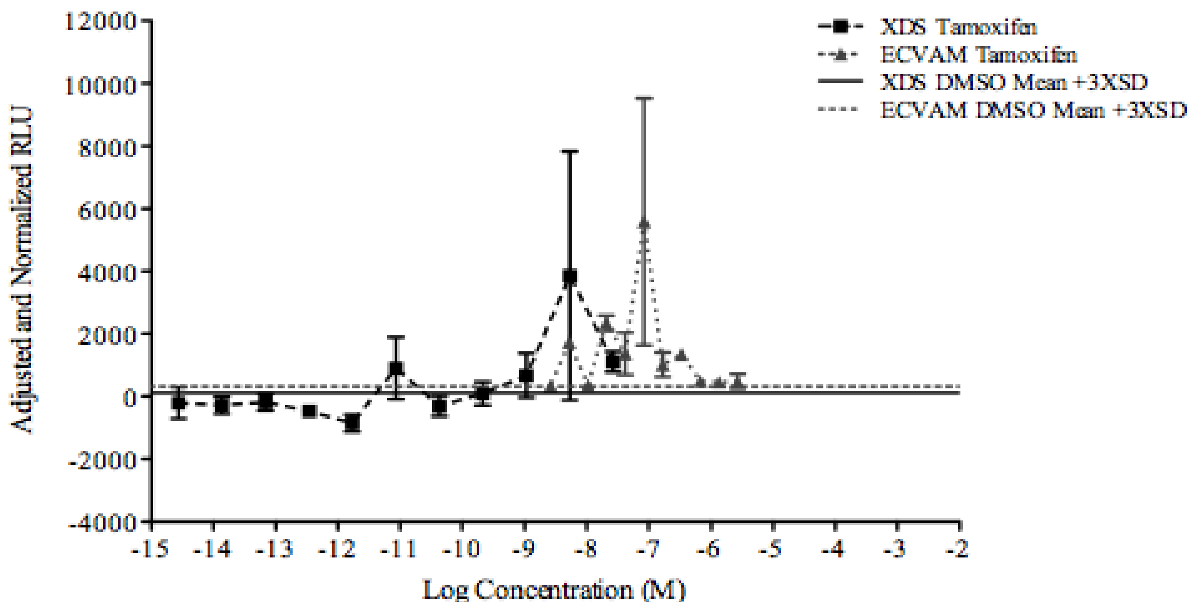
[a] Numbers in parentheses represent the number of test plates that failed acceptance criteria over the total number of plates tested.

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} µg/mL (naringenin) as compared to 17β-estradiol (2.3×10^{-6} µg/mL) or diethylstilbestrol (4.9×10^{-6} µg/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} µg/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. EC_{50} 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 **r**epresentative 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 CERI-STTA. The current cost of the CERI-STTA conducted at CERI 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, CERI, 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. 285*l*-3. Public Law 106-545. Available: http://frwebgate.access.gpo.gov/cgi-bin/useftp.cgi?IPaddress=162.140.64.183&filename=publ545.pdf&directory=/diska/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 JI, 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 Habor 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 β1 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=deb959f6fe9b73fdfb3ffd34c5a7adf6.

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, Breysse 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 (ERα), differentially affect ERα 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=ef340b6dbcaaec410aad7b3a355bb5af.

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.

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 Vitr 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, Giguére 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 5α-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 (ERα and ERβ) and have been stably transfected with the plasmid pGudLuc7.ERE. This plasmid contains four copies of a synthetic oligonucleotide containing the estrogen response element upstream of the mouse mammary tumor viral 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 EC_{50} or IC_{50} 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.

EC_{50}: 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).

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

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, EC_{50} is the substance concentration at the response midway between Top and Bottom, and HillSlope describes the slope of the curve.

IC_{50}: 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