ENTRIX

Report

Multi-Laboratory Validation of the H295R Steroidogenesis Assay to Identify Modulators of Testosterone and Estradiol Production

Prepared by:

Markus Hecker, Ph.D.

ENTRIX, Inc.

Saskatoon, SK

Canada

John P. Giesy, Ph.D.

Toxicology Centre

University of Saskatchewan

Saskatoon, SK

Canada

Gary Timm

U.S. Environmental Protection Agency

Offices of Science Coordination & Policy

1200 Pennsylvania Avenue, NW

Washington, DC

USA

Prepared for:

U.S. Environmental Protection Agency

February 2008

TABLE OF CONTENTS

TABLE OF TABLES

TABLE OF FIGURES

1 ABBREVIATIONS

AOE	City University of Hong Kong
DIFVR	Danish Institute for Food and Veterinary Research
EC	Effective Concentration
GER	University of Heidelberg
IHCP	Institute for Health and Consumer Protection
KFDA	Korea Food and Drug Administration
LOEC	Lowest Observable Effect Concentration
MSU	Michigan State University
NOEC	No Observable Effect Concentration
OECD	Organization for Economic Co-operation and Development
VMG NA	Validation Management Group for Non-animal Testing for Endocrine Disruptors
REACH	Registration, Evaluation and Authorization of Chemicals
RTP	US Environmental Protection Agency
T	Testosterone
E2	17β-estradiol

2 KEY PERSONNEL

Project Officer:

Gary E. Timm
Office of Science Coordination and Policy (7201M)
U.S. Environmental Protection Agency
Washington, DC 20460, USA
Tel: +1-202-564-8474
FAX: +1-202-564-8483
Email: Timm.Gary@epamail.epa.gov

_____ _____

Signature *Date*

Principal Investigators:

Prof. John P. Giesy, Ph.D.
University of Saskatchewan
Department of Veterinary Biomedical Sciences
Toxicology Centre
44 Campus Drive
Saskatoon, SK S7N 5B3
Canada

TEL: +1 (306) 966-2096
FAX: +1 (306) 966-4796
Email: jgiesy@aol.com

_____ _____

Signature *Date*

Dr. Markus Hecker, Ph.D.
University of Saskatchewan
Toxicology Centre
44 Campus Drive
Saskatoon, SK S7N 5B3
Canada

TEL: +1 (306) 966-5233
FAX: +1 (306) 966-4796
Email: mhecker@ENTRIX.com

_____ _____

Signature *Date*

Study Manager/Coordinator:

Dr. Markus Hecker, Ph.D.
ENTRIX. Inc.
Toxicology Centre
44 Campus Drive
Saskatoon, SK S7N 5B3
Canada

TEL: +1 (306) 966-5233
FAX: +1 (306) 966-4796
Email: mhecker@ENTRIX.com

_____ _____

Signature *Date*

3 INTRODUCTION

Recently, studies have indicated potential links between the exposure to natural and human-made substances in the environment and adverse effects on the endocrine and reproductive systems of mammals, birds, reptiles, amphibians, and fish (EPA, 1997; Kavlock et al. 1996). In response to emerging concerns that these substances may alter the function of endocrine systems and result in adverse effects to human health, the U.S. Congress included a provision in the Food Quality Protection Act of 1996 adding section 408 to the Federal Food Drug and Cosmetic Act. This section of the FFDCA requires EPA to:

> *… develop a screening program, using appropriate validated test systems and other scientifically relevant information, to determine whether certain substances may have an effect in humans that is similar to an effect produced by a naturally occurring estrogen, or other such endocrine effect as the Administrator may designate [21 U.S.C. 346 (p)].*

Subsequent to passage of the Act, EPA formed the Endocrine Disruptor screening and Testing Advisory Committee (EDSTAC), a committee of scientists and stakeholders that EPA charged to provide it with recommendations on how to implement its Endocrine Disruptor Screening Program (EDSP). Upon recommendations from EDSTAC, the EDSP was expanded using the Administrator's discretionary authority to include the androgen and thyroid hormone systems and wildlife effects. EPA accepted the EDSTAC's recommendations for a two-tier screening program (EPA 1998).

One of the assays recommended by EDSTAC as a Tier 1 screen was an in vitro a rodent minced-testis assay screen to detect chemicals with the potential to disrupt steroid hormone production (EDSTAC, 1998). The objective of the steroidogenic screen assay is to detect any substance that would disrupt estrogen and/or androgen gonadal steroid hormone production. In this way, the assay will complement the other Tier 1 assays and provide the necessary breadth and depth to detect substances that could be classified as endocrine disruptors. The steroidogenic assay is intended to identify xenobiotics that have as their target site(s) the endogenous components that comprise the intracellular biochemical pathway beginning with the sequence of reactions occurring after the receptor, up through and including the production of the terminal steroid hormones, i.e. testosterone (males) and estradiol/estrone (females). The steroidogenic assay is not intended to identify substances

that affect steroidogenesis due to effects on the hypothalamus, pituitary gland, and storage or release of gonadal steroid hormones. The most promising assay for use as a screen, which will meet the objectives as described above, will be a relatively fast, inexpensive, technically simple assay that identifies substances that alter gonadal steroid hormone production due to direct effects on the enzymes or other endogenous components of the steroidogenic pathway found in the testis and ovary.

Despite its long history of use, the rodent sliced testes assay had not been optimized. EPA conducted a series of studies to optimize the assay and evaluate its suitability to serve its function in the battery. Preliminary inter-laboratory studies showed high levels of variability within and between laboratories (Battelle, 2005); however, the seemingly insurmountable problem of assessing cytotoxicity specific to Leydig cells lead EPA's advisory committee to recommend that EPA abandon further work on this assay (EDMVAC, 2005). As a consequence, there was a need for a less variable and more reliable in vitro test systems as alternatives to the sliced testes assay. Based on studies conducted by ENTRIX for the EPA (Giesy et al. 2002; Hilscherova et al. 2004) and other reports as noted in section 3.2, the human H295R adreno-carcinoma cell line has been shown to possess all of the enzymes of the steroidogenic pathway and thus could serve as an *in vitro* model of steroidogenesis..

Development and standardization of the assay as a screen for steroidogenesis was carried in a multi –step process under EPA contracts with the ENTRIX Corporation. After initial assay development, US EPA presented a progress report on the development of the H295R assay to a committee of the Organization for Economic Cooperation and Development and invited member countries to join the US in its further standardization and validation. This invitation was accepted by Japan and Denmark. More recently laboratories in Germany, Hong Kong, and Korea have joined the validation effort. The initial product of this effort will be a validated protocol that can be used in the initial testing phase of the EDSP which is scheduled to begin in August 2008. The final product of the validation effort will be an OECD Test Guideline for assessing the potential of chemicals to affect steroid hormone synthesis.

This document reports on the validation of a cell-based screening assay using the H295R cell line to identify chemicals that act to alter steroidogenic process in humans and wildlife as part of the EDSP and OECD validation programs.

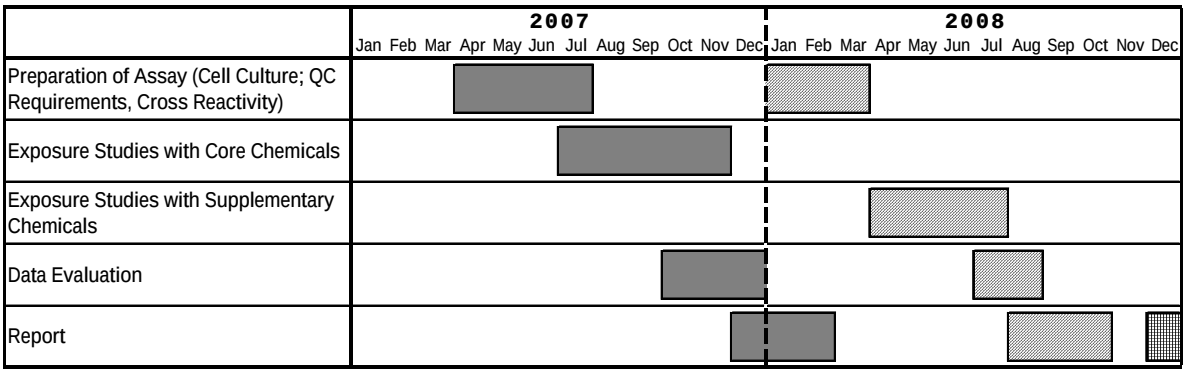
Validation:

Validation is a scientific process designed to characterize the operational characteristics and limitations of a test method and to demonstrate its reliability and relevance for a particular purpose. OECD Guidance Document 34 provides the principles of test validation and practical guidance for validation that are followed by OECD. These principles were set forth in the report from a workshop on validation in Solna (OECD 1996) and are consistent with the approaches used in Europe by the European Center for Validation of Alternative Methods (ECVAM 1995) and the U.S. Interagency Coordinating Committee on Validation of Alternative Methods (ICCVAM 1997).

3.1 Time Lines

Because the H295R Steroidogenesis Assay is being validated both as part of US-EPA's EDSP and the OECD Test Guidelines Program, the proposed validation efforts are being conducted under to two different timelines (Table 3.1). The overall duration of the validation studies was set at 19 months (April 2007 through October 2008). Due to time restraints regarding the EDSP — battery selection for Tier 1 must be completed in early 2008 to meet the August 2008 deadline for beginning testing — a peer review to meet the needs of the US EDSP was to be held after completion of studies on the 12 core chemicals. The validation studies for OECD purposes were then to be completed by including the extended set of 18 additional chemicals. Target date for completion of the laboratory work and data evaluation is August 31, 2008. A draft report will be submitted to OECD by October, 2008. A summary presentation shall be given to OECD at the 2008 VMG NA meeting.

Table 3.1: *Timeline for the H295R Steroidogenesis Assay validation studies to be submitted to US-EPA's EDSP and OECD for peer review. Gray shaded bars: EDSP & OECD; Diagonal pattern bars: OECD; Square pattern bar: Report to OECD at VMG NA meeting*



3.2 H295R Cell Line

The human H295R adreno-carcinoma cell line has been shown to be a useful *in vitro* model for steroidogenic pathways and processes (Hecker et al. 2006; Hilscherova et al. 2004; Sanderson et al. 2002). The H295R cell line expresses genes that encode for all the key enzymes for steroidogenesis (Gazdar et al. 1990; Rainey et al. 1993) (Figure 1). This is a unique property because *in vivo* expression of these genes is tissue and developmental stage-specific with typically no one tissue or one developmental stage expressing all of the genes involved in steroidogenesis. H295R cells have physiological characteristics of zonally undifferentiated human fetal adrenal cells (Gazdar et al. 1990). The cells represent a unique *in vitro* system in that they have the ability to produce the steroid hormones found in the adult adrenal cortex and the gonads, allowing testing for effects on both corticosteroid synthesis and the production of sex steroid hormones such as androgens and estrogens. There are several additional advantages to the use of the H295R cell line over other systems currently being evaluated as Tier I assays. One advantage the H295R *in vitro* assay has over tissue-based assays is that it permits the direct assessment of the potential impact of a chemical on cell viability/cytotoxicity. This is an important feature as it allows for the discrimination between effects that are due to cytotoxicity or due to the direct interaction of chemicals with steroidogenic pathways, which is not possible in tissue explants systems that consist of multiple cell types of varying sensitivities and functionalities. In addition, the NCI-H295R cells are commercially available from the

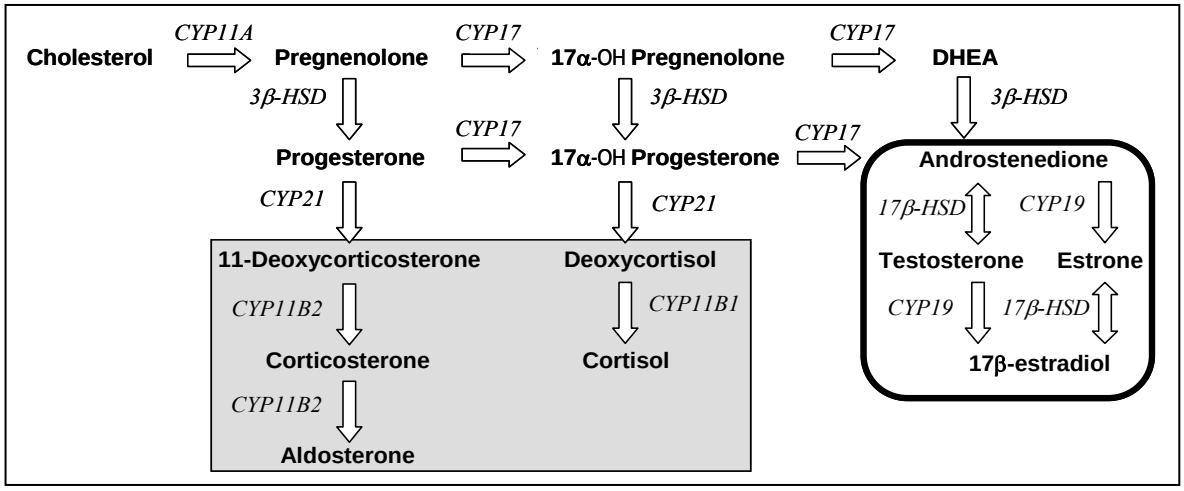


Figure 3.1: *Steroidogenic pathway in H295R cells. Enzymes are in italics, hormones are bolded and arrows indicate the direction of synthesis. Gray background indicates corticosteroid pathways/products. Sex steroid pathways/products are circled. CYP = cytochrome P450; HSD = hydroxysteroid hydrogenase.*

American Type Culture Collections (ATCC CRL-2128; ATCC, Manassas, VA, USA). Thus, these cells are available to everybody and no costly permissions are required as is the case for many other cell systems. The H295R has the advantage over other tests such as the minced testis assay that it allows for the detection of both increases and decreases in the production of both T and E2, and thus permits to assess chemicals with the potential to induce or inhibit this enzyme. Furthermore, H295R cells contain the complete suite of steroidogenic enzymes for the production of corticosteroids and sex steroids, and thus, enable the research of any target site within the steroidogenic pathway downstream of cholesterol in addition to those investigated in this study. Finally, the use of the immortalized H295R cells addresses the need for *in vitro* test systems as alternatives to tissue explant assays which require the use of live animals.

3.3 The H295R Steroidogenesis Assay

Based on the promising results obtained during initial studies researching the potential of the H295R cells to detect effects of chemicals on steroidogenesis including the production of testosterone, estradiol, and progestins (Hecker et al. 2006), a standardized H295R Steroidogenesis Assay protocol was developed (Figure 3.2; APPENDIX I). In brief, the assay is performed under standard cell culture conditions in 24-well culture plates. After an acclimation period of 24 h, cells are exposed for 48 h to multiple concentrations of the test chemical in triplicate. In parallel, a plate with known inhibitors and inducers of hormone production is run as a quality control (QC). At the end of the exposure period, the medium is removed from each well and hormones are extracted using ethyl ether (note: some hormone detection assays may not need extraction; in these cases the medium can directly be used in the assay). Cell viability in each well is analyzed immediately after removal of medium. Concentrations of hormones in medium can be measured using a variety of methods including the use commercially available hormone detection kits and/or instrumental techniques (LC-MS), again making the assay accessible to most laboratories.

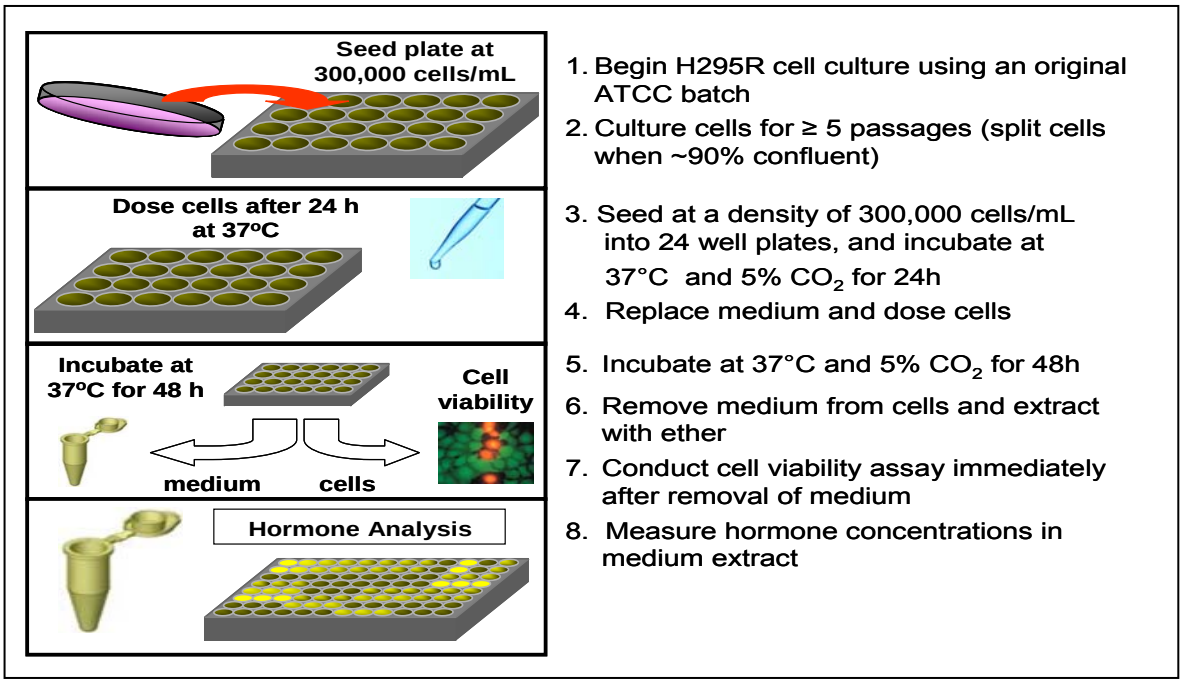


Figure 3.2: *H295R Steroidogenesis Assay to measure effects of chemicals on production of testosterone (T) and estradiol (E2).*

3.4 Pre-validation Studies Overview

An initial inter-laboratory pre-validation study has been conducted to evaluate the H295R Steroidogenesis protocol described in Section 3.3 using a limited set of three model chemicals at five independent laboratories. These studies demonstrated that the test protocol is very promising as a reproducible, transferable, sensitive, economic, and precise method to test for chemical effects on the production of T and E2 (Hecker et al. 2007). The three model compounds used in this study included forskolin, prochloraz, and fadrozole that had known modes of interaction with steroidogenic pathway. Comparison of changes in hormone production by H295R cells treated with these chemicals revealed a high degree of reproducibility of the tested protocol among five independent laboratories (Figure 3). However, the results also indicated that H295R cells appear to maintain some variability concerning their hormone producing capacities that was related to cell passage, which is likely to be due to the undifferentiated characteristics of the cells (Gazdar et al. 1990). Interestingly, both the direction and extent of the changes in hormone production with cell passage were predictable and reproducible among different laboratories (Hecker et al. 2006b). Furthermore, it was demonstrated that, despite the differences in absolute production of hormones that occurred as a function of cell passage, the relative response of

the H295R cells exposed to chemicals remained constant among the different laboratories (Hecker et al. 2007a; Figure 3.3). In addition, different data evaluation techniques that normalize the data to correct for differences due to cell passage have shown promise in terms of accounting for these variations in hormone production. For instance, expression of responses, both as changes relative to the controls and as percent of maximum efficacy, represent promising approaches. The development of appropriate data evaluation approaches is part of the validation study described in this document.

Based on the findings of these pre-validation studies the protocols to be used in this validation study were revised such that quality criteria for the exposure to known inducers and inhibitors of T and E2 production to serve as positive controls were defined, and that the cell culture protocols were adapted to reflect optimum hormone production patterns of the cells (APPENDIX I).

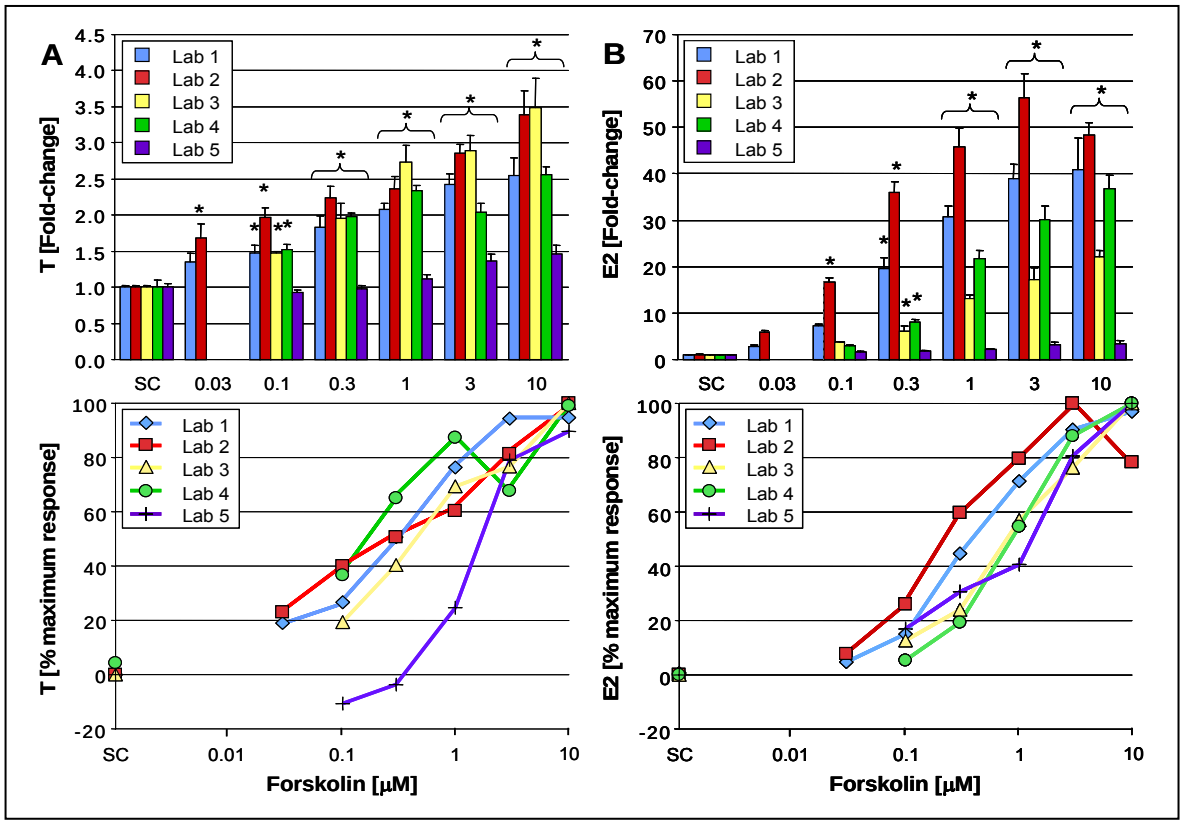


Figure 3.3: *Changes in T (A) and E2 (B) production by H295R cells measured by five (5) independent laboratories (Lab 1 – Lab 5). Data are expressed as relative changes compared to the SC (upper graphs) and percent of the maximum hormone concentration measured across all doses (maximum induction = 100%; lower graphs) observed after exposure to forskolin for 48 h. Data represents the mean of three independent exposure experiments. Error bars = standard error of the mean (SEM). * = $p < 0.05$.*

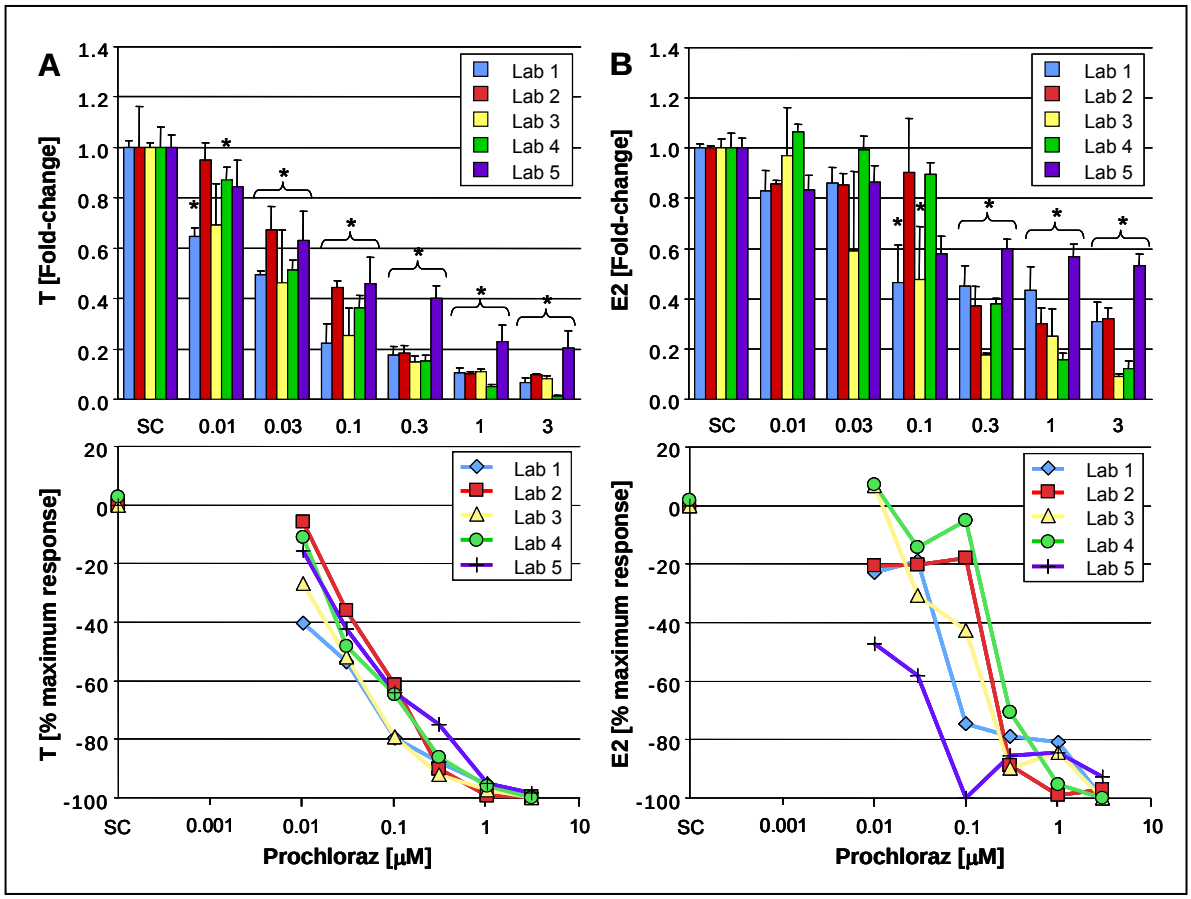


Figure 3.4: *Changes in T (A) and E2 (B) production by H295R cells measured by five (5) independent laboratories (Lab 1 – Lab 5). Data are expressed as relative changes compared to the SC (upper graph) and percent of the least hormone concentration measured across all doses (maximum suppression = -100%; lower graph) observed after exposure to prochloraz for 48 h. Data represents the mean of three independent exposure experiments. Error bars = standard error of the mean (SEM). * = $p < 0.05$.*

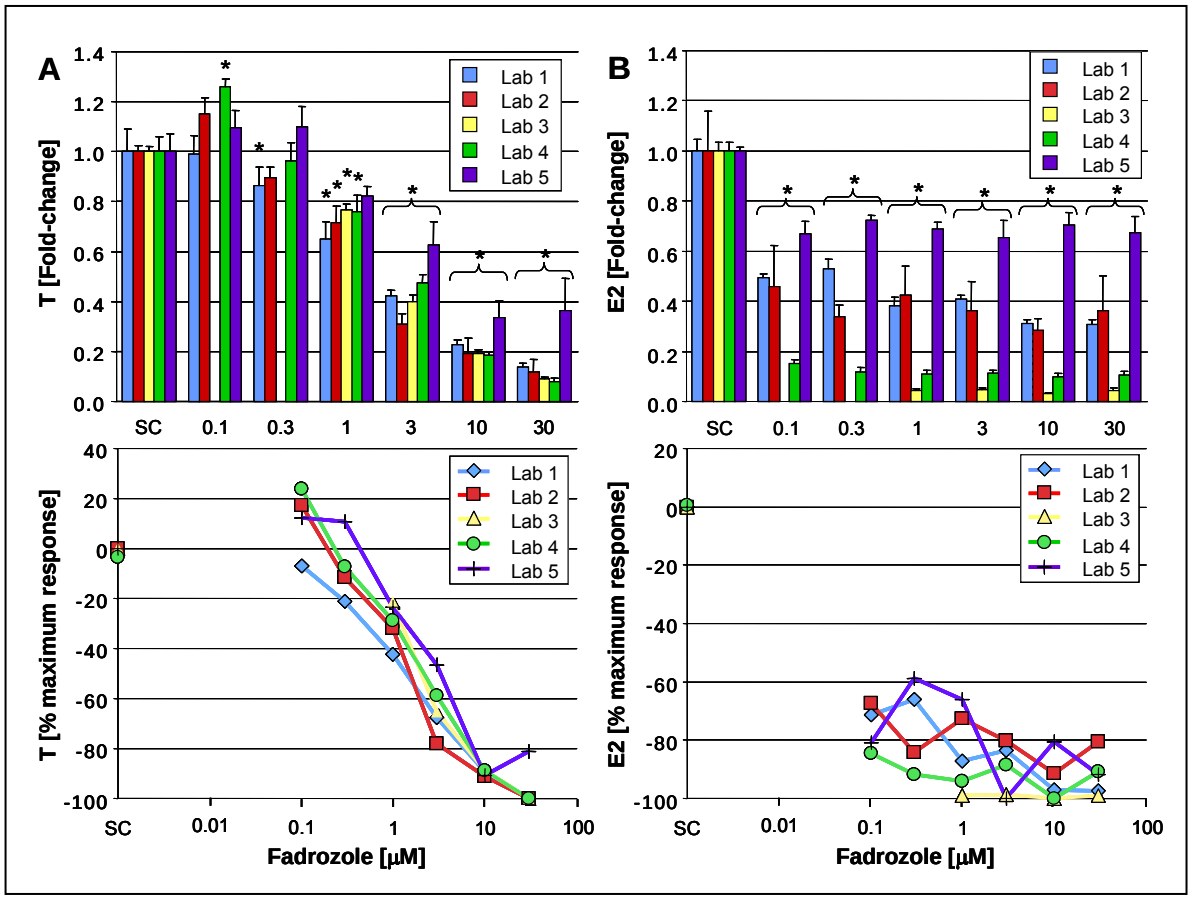


Figure 3.5: *Changes in T (A) and E2 (B) production by H295R cells measured by five (5) independent laboratories (Lab 1 – Lab 5). Data are expressed as relative changes compared to the SC (upper graph) and percent of the least hormone concentration measured across all doses (maximum suppression = -100%; lower graph) observed after exposure to fadrozole for 48 h. Data represents the mean of three independent exposure experiments. Error bars = standard error of the mean (SEM). * = $p < 0.05$.*

4 OBJECTIVES OF INTER-LABORATORY VALIDATION STUDY

The inter-laboratory comparison and validation study was conducted to evaluate the performance of the optimized H295R cell culture and exposure protocols (APPENDICES I & II) in different laboratories.

The specific objectives of this study are as follows:

1. Test the revised H295R Steroidogenesis Assay protocol at seven independent laboratories;
2. Compare the specific changes in the production of T and E2 in response to 12 model compounds with different mode of interactions with steroidogenic pathways (weak, medium and strong inhibitors and inducers; negative chemicals);
3. Validate the quality control criteria to be met when conducting the H295R Steroidogenesis Assay;
4. Assess the transferability, reproducibility, sensitivity and applicability of the assay;
5. Identify and validate appropriate data evaluation procedures.
6. Apply the developed protocols and evaluation procedures to 14 supplemental chemicals with largely unknown types of interaction with the production of T and E2 (lead laboratory only).

Ongoing studies including six of the seven laboratories that participated in the here presented validation studies will apply the developed protocols and data evaluation procedures to assess the potential of 18 supplemental chemicals with largely unknown types of interaction with the production of sex steroid hormones

5 SELECTION OF PARTICIPATING LABORATORIES

A total number of seven laboratories with different levels of experience in conducting the H295R Steroidogenesis Assay were selected to participate in the validation of this assay. Different proficiencies regarding the utilization of the assay are essential for an objective evaluation of the appropriateness of the test protocols and their transferability. As a consequence, the proficiency levels of the laboratories that are part of the validation group range from extensive to no previous experience. Extensive experience is represented by the lead laboratory (U of S/MSU), through which the assay was originally developed, and by three core laboratories (RTP, DIFVR, and CERI) that have been involved in the early pre-validation and optimization studies since 2005. A second group, consisting of two laboratories (GER and AOE), was chosen based on their participation in the most recent series of pre-validation studies that were conducted in 2006 and are representative of laboratories with limited experience with the assay. Finally, one laboratory (KFDA) was included that has never conducted the H295R Steroidogenesis Assay protocol. All laboratories that were part of the international validation team and their contact information are listed below (Chapter 5.1).

5.1 Participating Laboratories

University of Saskatchewan (U of S; *Lead Laboratory)*
Toxicology Centre
Contact:
Dr. Markus Hecker, Ph.D.
44 Campus Drive
Saskatoon, SK S7N 5B3, Canada

TEL: +1 (306) 966-5233
FAX: +1 (306) 966-4796
Email: mhecker@ENTRIX.com

US Environmental Protection Agency (RTP)
Endocrinology Laboratory
Research Triangle Park
Contact:
Dr. Ralph Cooper, Ph.D.
Endocrinology Branch, MD-72
TRD, NHEERL, US EPA, USA
TEL: +1-919-541-4084
FAX: +1-919-541-5138
Email: cooper.Ralph@epamail.epa.gov

Chemicals Evaluation and Research Institute (CERI)
Chemicals Assessment Center
Contact:
Dr. Yumi Akahori, Ph.D.
1600, Shimo-Takano, Sugito-machi,
Kitakatsuhika-gun, Saitama, Japan
TEL: +81-3-5804-6136
FAX: +81-3-5804-6149
Email: akahori-yumi@ceri.jp

Department of Toxicology and Risk Assessment (DIFVR)
National Food Institute
Technical University of Denmark
Contact:
Christine Nellemann, Ph.D.
Mørkhøj Bygade 19
DK-2860 Søborg
Denmark

TEL: +45-72-34-7549
Email: CNE@food.dtu.dk

University of Heidelberg (GER)
Heidelberger Institut für Zoologie
Contact:
Dr. Henner Hollert, Ph.D.
Im Neuenheimer Feld 230
69120 Heidelberg, Germany

TEL: ++ 49 6221 54-5650
FAX: ++ 49 6221 54-6162
Email: Henner.Hollert@urz.uni-heidelberg.de

City University of Hong Kong (AOE)
Department of Biology and Chemistry
Contact:
Dr. Margaret Murphy, Ph.D.
83 Tat Chee Avenue
Hong Kong, SAR, China

TEL: +852-2788-7368
Email: mbmurphy@cityu.edu.hk

National Institute of Toxicological Research (KFDA)

Korea Food and Drug Administration (KFDA)
Contact:
Soon Young HAN, Ph.D.
Endocrine Toxicology Team
194 Tongil-ro, Eunpyung-gu
Seoul 122-704, Korea

TEL: +82-2-380-1877~9
FAX: +82-2-380-1879
Email: soonyoungh@kfda.go.kr

6 SELECTION OF CHEMICALS

To evaluate the validity of the H295R Steroidogenesis Assay as a screen for potential effects of suspected endocrine disrupting chemicals on the production of T and E2, a total of 30 chemicals were identified. These chemicals were selected based on their known or suspected endocrine activity, or a lack thereof, and include various inhibitors and inducers of different strengths as well as positive and negative controls. Prior to initiation of the validation studies with the international laboratory team, all chemicals were to be pre-analyzed by the lead laboratory using the H295 Steroidogenesis Assay. It should be noted that for the purpose of the validation studies in context with the EDSP only the 12 core chemicals were tested by all laboratories. The remaining 18 chemicals (with few exceptions) were only tested by the lead laboratory at the time this report was written. However, for completeness the 18 supplementary chemicals are discussed in the subsequent sections but no data is provided for these in this report. The status of testing of the chemical groups described in sections 6.1 and 6.2 for each laboratory as of the time this report was written is provided (Table 6.1).

***Table 6.1**: Number of chemicals tested for interaction with the production of testosterone (T) and estradiol (E2) at each laboratory as of the time this report was written. The number refer to the code number assigned to each chemical at the initiation of studies, and do not reflect the order in which compounds are listed in tables 6.1 and 6.2. This was done because all chemicals were to be tested blind to avoid any potential bias. Note: at the time this report was written not all laboratories had tested all chemicals yet.*

	Core chemicals [code #]		**Supplementary chemicals [code #]**	
Lab #	***T***	***E2***	***T***	***E2***
1	1-12	1-12	13-15; 17-19; 21-28	13-15; 17-19; 21-28
2	1-12	1-12	---	---
3	1-12	1-12	---	---
4	1-8	1-8	---	---
5	5-8	---	---	---
6	1-12	1-12	---	---
7	---	---	---	---

To reduce the work load for individual laboratories, each group tested a total of 18 chemicals. These were comprised of the so called “core group” of 12 compounds that are tested in parallel by all laboratories. In addition to these 12 compounds, each laboratory

conducted assays of a different set of six of the 18 supplementary chemicals in the assay. That is, the 18 chemicals were divided into three sub-groups of six chemicals and each chemical sub-group was to be tested by two laboratories (total number of laboratories = 6, so that with every two laboratories testing 6 different compounds all 18 remaining compounds were analyzed).

Where possible the test set of chemicals was harmonized with those used in other steroidogenesis assays currently under development or in validation (e.g. REACH program).

6.1 Core chemicals

Out of the 30 chemicals, a core set of 12 compounds was established that were analyzed by all laboratories participating in the validation studies (Table 6.2). Selection of these core chemicals was made in accordance with accepted guidelines for the validation of screening type test systems (OECD 2005, Hartung 2004). All compounds included in this core set were previously reported to exhibit a specific type of direct or indirect interaction with the production of T and/or E2 as measured either by means of the H295R Steroidogenesis Assay or by other steroidogenic tests including tissue explant assays and/or *in vivo* studies. In addition, this core chemical test set included a number of negative chemicals that were not expected to elicit any effect on the endpoints measured here at non-cytotoxic concentrations. The inclusion of such negative chemicals is of importance because it allows evaluation of the specificity of a test system with regard to the endpoints of interest. The suite of positive chemicals utilized in this core chemical set was chosen to reflect different types and strengths of interactions with the production of the hormones analyzed here.

***Table 6.2**: Core chemicals and their hypothesized mode of action selected for the H295R Steroidogenesis Assay validation studies. Due to the nature of the validation studies, conduct of experiments using coded chemicals, here we do not distinguish between core and supplementary chemicals. Chemicals are sorted in alphabetical order.*

Name	CAS #	Mode of action	Product class	Effect type
Aminoglutethimide	125-84-8	Inhibits CYP19 aromatase and other cytochrome P450 enzymes	Pharmaceutical (phased out)	Medium to weak inhibitor of T and E2 production.
Atrazine	1912-24-9	Aromatase inducer in vitro	Herbicide	Weak inducer of E2 production.
Benomyl	17804-35-2	Aromatase inducer in vitro	Fungicide	Weak inhibitor of T production; Weak inducer or negative for E2 production. Has been shown to induce aromatase activity in human ovarian tumor cells (KGN).

Ethane dimethane sulfonate (EDS)	4672-49-5	Cytotoxic		No effect expected at non-cytotoxic concentrations.
Forskolin	66575-29-9	Cyclic-AMP second messenger system	Pharmaceutical	Strong inducer of T and E2 production.
Human chorionic gonadotropin (hcG)	9002-61-3	Binds to GtH receptor	Peptide hormone	No effect on T and E2 production in H295R cells.
Letrozole	112809-51-5	Specifically inhibits catalytic aromatase activity.	Pharmacytical	Strong inhibitor of E2 production. Weak inhibitor of T production.
Molinate	2212-67-1	Anti-cholinesterase/ neurotoxicant. Note: In vitro, molinate is a poor inhibitor of esterase activity, whereas molinate sulfoxide, a major metabolite of molinate in rats, and molinate sulfone were shown to be potent inhibitors of esterase activity, suggesting that metabolic activation of molinate is required in vivo.	Pesticide	Weak inducer of E2 and negative/weak inhibitor of T production.
Nonoxynol-9	26027-38-3	Unknown	Excipients, Pharmaceutical aid [surfactant], Pharmaceutical aid [wetting and or solubilizing agent], Spermaticide	Unknown.
Paraben (Butyl paraben)	94-26-8	ER binder	Preservative in food, cosmetics, toiletries, pharmaceutical.	Weak inducer of E2, and weak inhibitor of T production.
Prochloraz	67747-09-5	General inhibitor of microsomal cytochrome P450 mixed function oxidases.	Fungicide	Strong inhibitor of T and E2 production.
Trilostane	13647-35-3	3B-HSD competitive inhibitor	Pharmaceutical, used in treatment of Cushings disease	Strong inducer of T and E2 production.

6.2 Supplementary chemicals

In addition to the 12 core chemicals described in the previous section, 18 compounds were selected for additional testing (Table 6.3). To reduce the burden on each laboratory, the 18

chemicals were divided into 3-subsets of 6 chemicals each and each subset will be tested by two laboratories.

Selection of these additional 18 chemicals was made based on the range of putative effects, as well as general toxic properties and technical feasibility (e.g. availability of the compound, ownership rights, etc.). The types of effects were categorized as strong, medium, and weak inducers and inhibitors of production of testosterone, estradiol, or both hormones as well as negative compounds. All decisions were discussed and made in agreement with the US-EPA and the OECD advisory group.

Table 6.3: *Supplemental chemicals and their hypothesized mode of action selected for the H295R Steroidogenesis Assay validation studies. Due to the nature of the validation studies, conduct of experiments using coded chemicals, here we do not distinguish between core and supplementary chemicals. Chemicals are sorted in alphabetical order.*

Name	CAS #	Mode of action	Product class	Effect type
2,4-Dinotrophenol	51-28-5	Cell toxicant: phosphorylation uncoupler	Industrial chemical	No known endocrine function other than cell toxicity and altered bioenergetics.
Bisphenol A	80-05-7	Cyclic-AMP second messenger system; proported ER binder	Monomer in polycarbonate plastics	Unknown. Some evidence that alters Progesterone in vitro, but mechanism may or may not be c-AMP second messenger system. For all steroidogenesis assay, will need to be specific for endpoint of assay. Tested positive for ER binding in vitro and in uterrotrophic assay.
Bromocriptine	25614-03-3	Dopamine receptor agonist/ affects production of pituitary hormones	Pharmaceutical: treatment of pituitary tumors, Parkinson's, anti-aging	Shown to suppress prolactin secretion while increasing growth hormone levels.
Cyanoketone	4248-66-2	3beta-HSD inhibitor		Unknown. An androstenolone-nitrile compound with steroidogenesis-blocking activity
Danazol	17230-88-5	3HSD; P450c17 (17 hydrolase/C17-20 lyase); 17KSR	Agricultural Chemical, Antineoplastic agents, Contraceptives, postcoital, synthetic, Drug / Therapeutic Agent	Unknown

Di (2-ethylhexyl) phthalate (DEHP)	117-81-7	Inhibits FSH-stimulated c*AMP* accumulation. Effects have been demonstrated at the level of P450scc and aromatase. Note: Compound that has been hypothesized to be active is the metabolite MEHP, not DEHP.	Polyvinyl additive	Metabolite monoethylhexyl phthalate (MEHP) has been shown to suppress aromatase and estradiol production in female rat primary granulosa cells. Parent compound is not considered active.
Fenarimol	60168-88-9	Aromatase inhibition	Fungicide	Shown to inhibit aromatase (CYP19) in vitro, evidence from in vivo studies not as unequivocal
Finasteride	98319-26-7	5-a reductase inhibitor	Pharmaceutical, therapeutic agent for prostrate cancer, hirsutism, and alopecia	Unknown
Flutamide	13311-84-7	P450c17 (17 hydrolase/C17-20lyase)	Pharmaceutical	Unknown
Genistein	446-72-0	Anti-oxidant, topoisomerase inhibitor/ tyrosine kinase inhibitor	Pharmaceutical	Weak inducer of E2 and weak inhibitor of T production. Weak estrogen receptor agonist
Glyphosate (Roundup)	1071-83-6		Herbicide	Unknown. Has not shown to conclusively affect reproduction in laboratory in vivo studies.
Ketoconazole	65277-42-1	Inhibiting the microsomal cytochrome P450 mixed function oxidases. This drug inhibits 17 alpha-hydroxylase, C17-20 lyase, and the cholesterol-side-chain cleavage enzyme	Fungicide	Strong inhibitor of T production; Medium inhibitor of E2 production; Induces progesterone production.
Piperonyl butoxide	51-03-6	Cytochrome P450 inhibitor	Pesticide synergist	Unknown. This compound is used to inhibit several P450s involved in metabolism but not necessarily steroidogenesis.
Prometon	1610-18-0	Photosynthetic inhibitor	Wide-spectrum herbicide	Weak inducer of E2 production; Negative for T.
RU-486/mifepristone	84371-65-3	Negative for ER very weakly positive for AR at high conc., blocking the progesterone receptor, incr. levels of EST.	Pharmaceutical	Unknown
Spironolactone	52-01-7	Antiandrogen action through inhibition of 17α hydroxlase;Glucocorticoid & PXR-ligand	Pharmaceutical	Unknown
Taxol (Paclitaxel)	33069-62-4	P450 SCC	Anti cancer drug	Unknown

Vinclozolin	50471-44-8	Metabolized to Mi and M2, which are strong AR antagonists	Fungicide	Weak inducer of and moderate inhibitor of T production.

7 TEST PROCEDURE

7.1 QA criteria

Laboratories were required to demonstrate competence in performing all of the procedures that are part of the H295R Steroidogenesis Assay prior to testing chemicals. The quality controls that were part of the actual conduct of the assay to allow for the evaluation of the assay performance during each experiment also served as the benchmarks for determining laboratory competence prior to the initiation of chemical testing.

7.1.1 Performance of H295R Cells under Standard Culture Conditions

The qualifying experiment required growing cells for 5 passages, seeding and exposing them as described for the QC plates in the H295R exposure protocol, and measuring E2 and T in the cell media using pre-validated hormone detection methods. The threshold concentrations for basal production of testosterone (T) and estradiol (E2) relative to the minimum detectable level (MDL) of the assay method employed at each laboratory is given for each test chemical (Table 7.1). Since production of E2 in passage 5 cells may not be sufficient to detect decreases greater than 1.5-times the response after exposure to an inhibitor (Hecker et al. 2006b), during the qualifying experiments it was only expected that the laboratory showed conformance with the performance criteria for E2 induction after exposure to the stimulator forskolin (Table 7.1).

Table 7.1: *Performance criteria to be met by each laboratory during the qualifying experiments. Criteria were selected based on the previous pre-validation studies investigating the effects of forskolin and prochloraz on the production of T and E2 by H295R cells in five independent laboratories (Hecker et al. 2007b). Numbers represent average (replicate wells of a given dose) change in hormone production relative to the solvent control (SC = 1) in a given plate.*

	Testosterone	***Estradiol***
Basal Production	*≥ 2.5-times MDL*	*≥ 2.5-times MDL*
Induction (10uM forskolin)	*≥ 2-times SC*	*≥ 15-times SC*
Inhibition (3uM prochloraz)	*≤ 0.5-times SC*	*n/a*

7.1.2 Performance of H295R Cells during Exposure Experiments

A necessary step in the validation of an assay is to define the quality control (QC) criteria. A QC plate was designed to include wells for control (i.e., blank), non-treated wells, solvent control wells, wells with a known inducer (forskolin) and wells with a known inhibitor (prochloraz). The first step in the QC validation studies was to measure the alterations in hormone production in cells exposed to forskolin and prochloraz in the QC plates. In addition, selected wells are dedicated to the evaluation of maximum cytotoxicity and the viability assay is performed on the entire plate. (Figure 7.1) The target performance criteria to be achieved for the QC plate are described below and were selected based on the data obtained for the exposure to these chemicals at five laboratories during the pre-validation studies (Hecker et al. 2007b):

1. Blank: basal hormone production of T and E2 had to be at least 2.5-times the MDL of the hormone detection assay used.
2. Solvent control: basal hormone production of T and E2 had to be at 2.5-times the MDL of the hormone detection assay used.
3. Inducer: *Forskolin (10 μM):* ≥ 10-times induction of E2 production, and ≥ 2-times induction of T production after 48 h.
4. Inhibitor: *Prochloraz (3 μM):* ≥ 50% reduction of E2 and T production compared to the SCs after 48 h.
5. Cytotoxicity positive wells are treated for 15 - 30 min with 70% methanol after harvesting the media for hormone assessment

	1	2	3	4	5	6
A	Blank [a]	Blank [a]	Blank [a]	Blank [a] + MeOH [b]	Blank [a] + MeOH [b]	Blank [a] + MeOH [b]
B	DMSO 1uL	DMSO 1uL	DMSO 1uL	DMSO 1uL + MeOH [b]	DMSO 1uL + MeOH [b]	DMSO 1uL + MeOH [b]
C	FOR 1uM	FOR 1uM	FOR 1uM	PRO 0.3uM	PRO 0.3 M	PRO 0.3uM
D	FOR 10uM	FOR 10uM	FOR 10uM	PRO 3uM	PRO 3uM	PRO 3uM

[a] Blank wells receive medium only.

[b] Methanol (MeOH) was added **after** the exposure is terminated and the medium is removed from these wells.

Figure 7.1*: Plate layout for QC-plate to be analyzed together with each chemical exposure experiment. PRO = prochloraz; FOR = forskolin; MeOH = methanol.*

Other quality criteria include an evaluation of the variation between replicate wells, replicate experiments, linearity, and sensitivity of hormone detection systems, variability between replicate hormone measures of the same sample. One criterion for the acceptance of data generated during an experiment was the validity of the SC in terms of its position within the standard curve range. Allowable location of the SC was within the 75% range below the upper part (maximum optical density [OD] or similar response measured by hormone detection system) and 75% above the lower part (minimum OD or similar response measured by hormone detection system) of the linear range of standard curve for inducers, and inhibitors, respectively (Figure 7.2). Dilutions of medium (extracts) in the hormone detection assay were to be selected accordingly.

Table 7.2: *Acceptable ranges and/or variation (%) for H295R assay QC parameters.*

	Comparison Between	***T***	***E2***
Basal hormone production in blanks and SCs	*Fold-greater than MDL*	*≥ 2.5-fold*	*≥ 2.5-fold*
Exposure Experiments - Within Plate CV for SCs (Replicate Wells)	*Absolute Concentrations*	*≤ 30%*	*≤ 30%*
Exposure Experiments - Between Plate CV for SCs (Replicate Experiments)	*Fold-Change*	*≤ 30%*	*≤ 30%*
Hormone Detection System – Sensitivity	*Detectable fold-decrease relative to SC*	*≥ 2.5-fold*	*≥ 2.5-fold*
Hormone Detection System – Replicate Measure CV for SCs	*Absolute Concentrations*	*≤ 25%*	*≤ 25%*
Medium Extraction – Recovery of Internal ^{3}H Standard (If Applicable)	*CPM*	*≥ 65% Nominal*[a]	

[a] *Note: Hormone concentration should be normalized for % recovery in each well.*

The QC plate was used in all assays, and the ability to meet these criteria was intended to be used as the qualifier to accept data generated during an experiment. In those instances where the data did not meet these criteria, the experiment typically was to be repeated. A second level of QC was required for the acceptance of the test chemical data generated during the validation studies. These QC criteria were to be defined on the basis of the variation observed during the validation experiments, and were to include but were not limited to measures of effective concentrations (EC5, EC10 and EC50) and fold-changes relative to the SCs (see section 10).

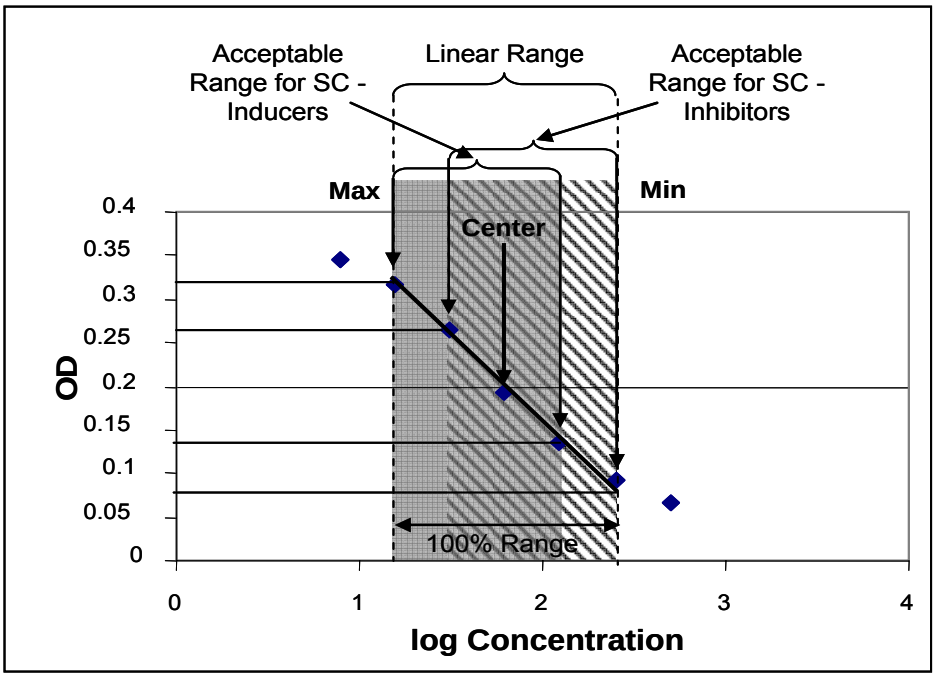


Figure 7.2: *Example of hormone detection system standard curve indicating acceptable range of hormone concentration of solvent control (SC) sample (not corrected for dilution in assay). Max = upper limit of linear range; Min = lower limit of linear range. Grey shaded area = 75% range below the maximum OD of the linear part of the standard curve. Diagonally striped area = 75% range above the minimum OD of the linear part of the standard curve.*

7.2 Testing of chemicals

Validation of the H295R Steroidogenesis Assay was conducted in two main phases, each focusing on a certain suite of chemicals. In the first phase the core set of 12 chemicals was to be analyzed by all laboratories. In the second phase, the remaining 18 compounds were to be tested in three sub-sets of six chemicals, each of which will be analyzed by two laboratories.

Prior to initiation of the actual exposure experiments each chemical had to be tested for potential interference with the hormone detection system utilized. This was of particular relevance for antibody based assays such as ELISAs and RIAs because it has been previously shown that some chemicals can interfere to a certain extent with these tests (Puddefoot et al. 2002; Villeneuve, personal communication; Hecker, personal communication). The results from a simple test can not be used to correct for possible interference with the hormone detection system, nor can it be used to eliminate the compound of interest from being evaluated in the H295R assay, it will provide critical information regarding the validity of the data generated.

7.2.1 Core chemicals

After satisfactorily demonstrating that the laboratory could meet the performance criteria listed in Table 7.1, the laboratory was allowed to proceed with evaluation of the 12 core chemicals. In this phase each chemical was to be tested in two-three independent experiments as a measure of intra-laboratory variation. Testing of core chemicals represents the key aspect of these studies and the results are critical in confirming the basic utility of the assay. As pointed out in Section 6, the majority of the core chemicals have been well characterized in terms of their interaction with the production of T and E2 or other endocrine processes. It was decided in conjunction with the US-EPA that completion of this stage of the validation studies and the result of the supplementary chemicals in the lead lab will suffice for the submission of the assay for peer-review and validation through the EDSP and would allow EPA to include the H295R assay in the Tier 1 battery for the first group of chemicals for which screening will begin in August 2008.

7.2.2 Supplementary chemicals

For validation purposes through OECD, an additional 18 chemicals were tested using the H295R Steroidogenesis Assay following the procedures given above. These studies were be initiated after phase I (testing of core chemicals) has been completed. The aim of this second phase of supplementary chemical testing was to evaluate the broader application of the assay to less well characterized compounds and chemicals of unknown or uncertain modes of action.

7.3 Data interpretation

One purpose of this validation study was to come to a decision regarding the interpretation and evaluation of data obtained with the H295R Steroidogenesis Assay. It was the aim to identify and harmonize appropriate statistical procedures to be used with the assay. Defining data evaluation criteria is essential in the validation and later implementation of the assay and will be a key component of the standardized study protocols that will come out of these efforts, and that are to be submitted for review through OECD member countries. To achieve these goals the data achieved during the experiments involving the 12 core chemicals have been subjected to a number of different data evaluation procedures, and from which the most appropriate approach was to be selected for the final study protocols.

In previous discussions with both the US-EPA and OECD, it was decided that results obtained with the H295R Steroidogenesis Assay would be evaluated using a semi-quantitative approach that will group data as negative, weak, medium or strong inducers or inhibitors of steroidogenesis. Decisions regarding the evaluation approach utilized in this validation study were decided together with the laboratories involved in the OECD validation of the assay as well as with the OECD advisory group. Prior to being subject to the analysis procedures discussed below, data were to be categorized into effectors, negatives, and uncertain chemicals. This categorization was conducted considering a combination of basic comparative statistical procedures, magnitude of change, and dose-response profiling. A summary of these parameters is provided (Table 7.3), and a detailed discussion and selection of criteria and approaches based on the data generated during this validation effort is given in section 10.

Table 7.3: *Data categorization parameters for the pre-categorization of results obtained with the H295R Steroidogenesis Assay prior to statistical assessment.*

Parameter	Criterion
Type-I error	Response has to be statistically significant different from the SC: $\alpha < 0.05$
Pattern	Data has to follow dose response type profile at non-cytotoxic doses, or doses that do not interfere with the hormone detection assay (note: response can be bi-phase such as increase at lower and decrease at higher doses but changes randomly observed at only a few concentration within the dose range are to excluded).
Interference with hormone detection assay	In the case of strong interference of the chemical of interest with the hormone detection system utilized this data is to be omitted. In the case of weak to moderate interference, results may be corrected for the % interference.
Cell viability	Only non-cytotoxic concentration (> 80% cell viability) are to be included

In addition, where permitted by the data, dose-response curves were to be defined that would allow both the determination of effective concentrations (ECs) of a chemical to alter a hormone concentration and its relative potency. It was assumed that the study protocols would only permit the description of dose-response curves for medium and strong inhibitors/inducers. This was because the dose-range for chemicals to be tested in the validation studies covered a wide range of concentrations due to the relatively wide spacing of concentrations; however, this results in lower resolution. Weak inducers that only show effects at the greatest doses, therefore, would be required to be re-analyzed using a different dosing regime to allow for the description of full dose-response curves. The OECD validation management group decided that such retesting was not necessary given the primary purpose of this assay: a screen to identify potential inducers and/or inhibitors of T and E2 production. The proposed approach allowed the grouping of chemicals based on their potency to alter the production of T and E2, and provided additional toxicological information for medium and strong effectors.

A key component of the data evaluation procedure was the integration of cell viability data from each assay to identify acceptable dose-ranges that can be used in evaluating the potency of a test chemical. Given the relatively large difference among doses (10-fold) there is a certain degree of uncertainty regarding the true cytotoxicity threshold of a test compound. Therefore, care must be taken when strong cytotoxicity is observed for a chemical. In this study all doses that exhibited cytotoxicity greater 20% were omitted for further evaluation.

8 DATA PROCESSING AND STATISTICS

All data are expressed as mean +/- standard deviation (SD). To evaluate the relative increase/decrease in chemically altered hormone production, results were normalized to the mean solvent control (SC) value for each assay (i.e., each 24-well plate of cells used to test a given chemical), and results are expressed as changes relative to the SC in each exposure plate (Equation 1). Furthermore, the average response in each well was divided by the relative cell viability measured in the same well to normalize for possible differences due to variations in the number of live cells (note: as stated in 7.3; Data interpretation, all data obtained for wells with > 20% cytotoxicity was omitted).

Fold-change = Concentration per well / average concentration of SCs in same plate (1)

Prior to conducting statistical analyses, the assumption of normality and variance homogeneity was evaluated. Normality was evaluated using standard probability plots or any other appropriate statistical method (e.g. Shapiro-Wilk's test). If the data were not normal, transformation of the data to approximate a normal distribution was attempted. If the data were normally distributed or approximate normal distribution, differences between chemical treatments and solvent controls (SCs) were analyzed using the parametric Dunnett's test. If data were not normally distributed, an appropriate non-parametric test was be used (e.g. Kruskal Wallis, Steel's Many-one rank test). Effective concentrations (EC_{50}s) were calculated using a probit model. Differences were considered significant at $p < 0.05$.

8.1 Intra-laboratory statistics

The participating laboratories provided all data to the lead laboratory for statistical analysis. Data was reported both in form of mean responses +/- 1 SD for each well measured and as raw data. Mean response data were entered by each laboratory into the data sheet templates provided for this purpose by the lead laboratory. Raw data was submitted to the lead laboratory in a format of choice by the participating laboratories, and included the following information:

1. Standard and calibration curves for all analytical assays conducted
2. Each replicate measure in form of the original data provided by the instrument utilized for a specific analysis (e.g. as OD, fluorescence units, CPM, etc.)
3. Hormone extraction recovery data (if applicable)

Each laboratory was responsible for conducting assay-related statistical analysis (e.g. variation between replicate measures of the same well) in compliance with the QC criteria required for this assay. In those cases where significant deviations from target criteria were observed (e.g. increased variation between replicate measures of the same sample, significant differences between standard curves of the same assay), the deviations were to be reported to the lead laboratory. The lead laboratory then decided in conjunction with the participating laboratory and the OECD advisory group regarding the acceptability of these data.

8.2 Inter-laboratory statistics

Conduct of inter-laboratory data evaluation and statistical analyses was the responsibility of the lead laboratory. Data was subjected to the statistical procedures described above (Chapter 9.0). Statistical analysis procedures and results were submitted to the US-EPA and the OECD advisory group for independent review. All data and analysis results are provided in form of summarized data and raw data including all analysis procedures and steps the data was subjected to (APPENDICES III – IX).

9 RESULTS

9.1 QA criteria

9.1.1 Performance of H295R Cells under Standard Culture Conditions

9.1.1.1 Basal hormone Production (Blanks)

Threshold concentrations for basal production of testosterone (T) and estradiol (E2) varied significantly among laboratories (Table 9.1). The greatest and the least T concentrations were observed at Labs 1and 5 with 6575 and 626 pg/mL, respectively. For E2, concentrations ranged between 123 and 12 pg/mL at Labs 1 and 3, respectively. With the exception of T at Lab 6, the coefficients of variation (CVs) at the different laboratories were between 20 and 40% for T and between 9 and 40% for E2. Among laboratory CVs were 57 and 62% for T and E2, respectively. All laboratories met the performance criterion for the basal hormone production, which was defined as a 2.5-fold greater concentration than the MDL for each hormone (Table 7.1).

Table 9.1: *Comparison of basal production of testosterone (T) and estradiol (E2) as measured in the blanks of the QC-plates among laboratories. SD = Standard deviation; CV = Coefficient of variation (%); Max = Maximum hormone concentration observed during all experiments; Min = Minimum hormone concentration observed during all experiments; Lab = Laboratory.* [a] *Only data from two QC-plates available.* [b] *No E2 data available from this laboratory.*

	Lab 1		Lab 2		Lab 3		Lab 4		Lab 4		Lab 6	
	T	*E2*	*T*	*E2*	*T*	*E2*	*T*	*E2*	*T*[a]	*E2*[b]	*T*	*E2*
Mean	5694	131	1718	50	4616	20	2057	54	1384	---	2881	92
SD	1246	11	576	20	953	6	825	19	1071	---	574	12
CV	22	9	34	40	21	30	40	36	77	---	20	13
Max	6575	139	2473	73	6175	27	3106	77	2141	---	3342	104
Min	4813	123	870	14	3344	11	1292	31	626	---	2239	80

9.1.1.2 Initial QC-Plate Experiment Responses

Changes in hormone production relative to the solvent controls as determined in the QC-plate experiment conducted to evaluate the performance of the H295R prior to initiation of the exposure experiments showed a comparable trend among laboratories with exception of E2 production after exposure to 0.3 μM prochloraz at Labs 3 and 4 (Figure 9.1). However, these laboratories achieved similar responses compared to all other groups during later QC-plate experiments (Figures 9.3 and 9.4), and thus, all subsequent data from these laboratories was accepted if not stated otherwise. While there was some variability among laboratories for the production of both hormones (Table 9.2), this variability did not exceed 89% for the forskolin exposures, and a CV of up to 75% was observed for the T data in the prochloraz exposures. Within laboratory variation was substantially less with CVs not exceeding 30% for either the SCs or the forskolin treatment groups (Table 9.3). The greater CVs observed after exposure to prochloraz are due to the fact that hormone concentrations were close to the method detection limits of the utilized hormone quantification assays. Regardless of the observed inter-laboratory variation, all laboratories could demonstrate conformance with the data performance criteria outlined in chapter 7.1.2 with the exception of T measured by Labs 2 & 3 after exposure to forskolin. Here only 1.5- and 1.6-fold instead of the desired 2-fold increases were observed. However, considering that the trend for T after exposure of cells to forskolin was similar to that reported by the other groups, and because in the subsequent experiments greater than 2-fold changes were observed at these lab (see chapter 9.1.3), it was decided to accept data from these laboratories. Nevertheless, care should be taken when evaluating T data of laboratories for potential inducers when less than 2-fold changes are observed in the respective QC-plate of an experiment.

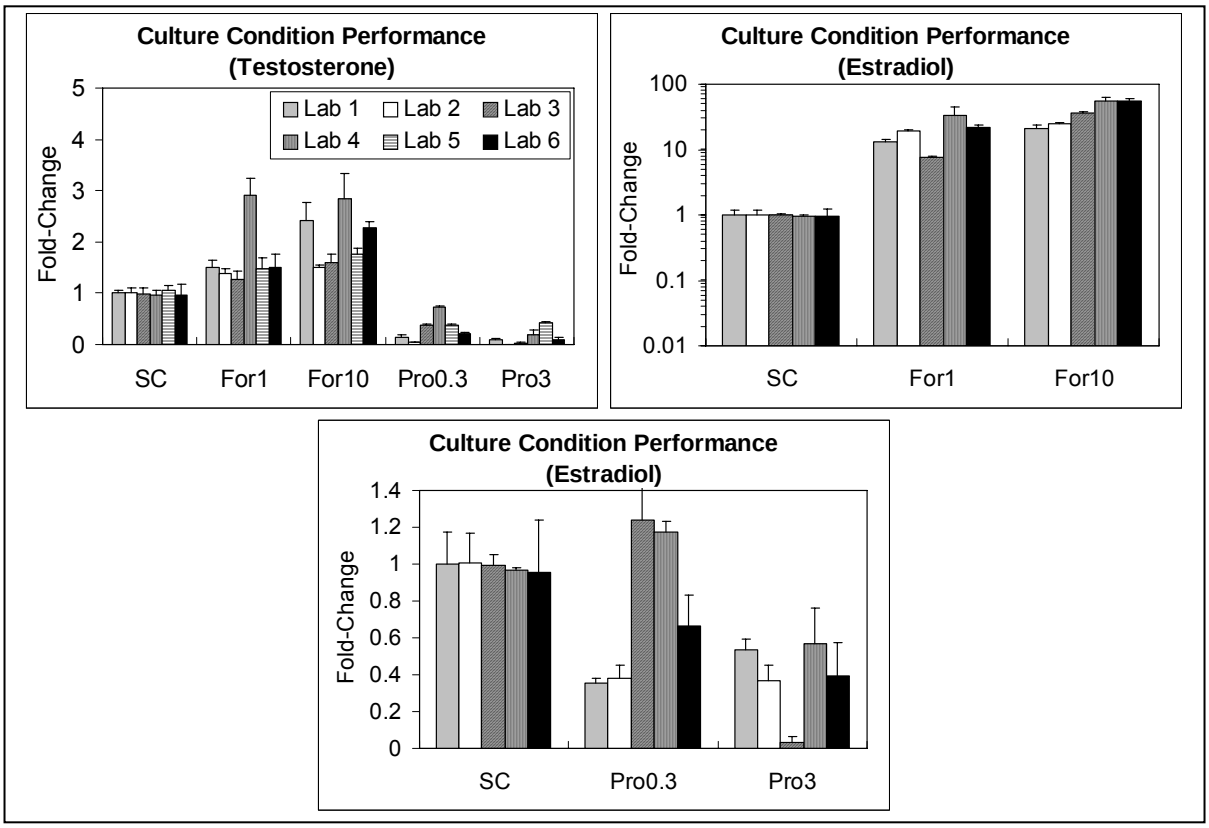


Figure 9.1: *Comparison of changes in the concentrations of testosterone (T) and estradiol (E2) relative to the solvent controls (SC=1) in the QC-plate experiment conducted to evaluate the performance of the H295R prior to initiation of the exposure experiments. For 1 = 1 μM Forskolin; For 10 = 10 μM Forskolin; Pro 0.3 = 0.3 μM Prochloraz; Pro 3 = 3 μM Prochloraz. Error bars = 1x standard deviation. Bars represent means of three replicate well (Lab 4 only 2 replicate wells). No data for E2 available from Lab5.*

Table 9.2: *Among laboratory coefficients of variation (CVs) for changes in concentrations of testosterone (T) and estradiol (E2). Calculations were based on data expressed relative to the solvent controls (SC=1) from the QC-plate experiment conducted to evaluate the performance of the H295R prior to initiation of the exposure experiments. For 1 = 1 μM Forskolin; For 10 = 10 μM Forskolin; Pro 0.3 = 0.3 μM Prochloraz; Pro 3 = 3 μM Prochloraz.*

	Among Lab CV	
	T	**E2**
For 1	13	76
For 10	65	89
Pro 0.3	54	36
Pro 3	75	26

Table 9.3: *Within laboratory coefficients of variation (CVs) for changes in concentrations of testosterone (T) and estradiol (E2). Calculations were based on absolute data from the QC-plate experiment conducted to evaluate the performance of the H295R prior to initiation of the exposure experiments. For 1 = 1 μM Forskolin; For 10 = 10 μM Forskolin; Pro 0.3 = 0.3 μM Prochloraz; Pro 3 = 3 μM Prochloraz.*[a] *No E2 data available from this laboratory.*

	Testosterone					
	Lab1	***Lab2***	***Lab3***	***Lab4***	***Lab5***	***Lab6***
SC	6	10	10	8	16	24
For 1	9	7	12	11	4	18
For 10	14	3	10	17	2	6
Pro 0.3	29	30	8	0	8	11
Pro 3	41	64	50	52	3	41
	Estradiol					
	Lab1	***Lab2***	***Lab3***	***Lab4***	***Lab5***[a]	***Lab6***
SC	18	16	5	1	---	30
For 1	7	6	8	34	---	13
For 10	12	5	6	11	---	9
Pro 0.3	8	18	42	5	---	25
Pro 3	11	21	89	33	---	47

9.1.2 Performance of H295R Cells during Exposure Experiments

As in the standard culture performance experiments, with a few exceptions, all laboratories complied with the key quality performance parameters such as fold induction or percent inhibition of hormone production after exposure to forskolin and prochloraz, respectively and observable fold-change (Figures 9.2, 9.3 & 9.4). However, in some cases basal hormone production was such that no decreases in E2, or T (one laboratory only) production could be measured (Table 9.4). Another issue that was observed in rare occasions was an increase in variation among replicate wells such that the data could not be used for further evaluations. This, however, occurred only at one laboratory during a single experiment (Lab 4; Chemicals 5-8; Experiment 1), where the average CV of the SCs was 48%, which is almost 20% greater than the QC criterion of 30% for this parameter. None of the results obtained during these experiments were used for the data evaluation as described in Chapter 10.2 (Table 9.4). However, it should be noted that the instances where issues with sensitivity and/or variability occurred, these events were rare and did not impact the overall validity and usability of data produced during these studies. Overall, only 4 or 10% of all experiments for T and E2 were excluded due to these issues. The

reason for the slightly greater percentage of non-usable data for E2 is probably due to the fact that basal E2 production was relatively low compared to T (Hecker et al. 2006b).

Relative changes in the production of T and E2 after exposure to forskolin and prochloraz in the QC plates was comparable both within and among laboratories (Figures 9.3, 9.4 & 9.5), indicating that the H295R Steroidogenesis Assay functioned properly at all laboratories. Coefficients of variation for relative changes measured after exposure to forskolin and prochloraz were between 2 and 21% and between 3 and 45%, respectively, for T, and between 4 and 88% and between 9 and 61%, respectively, for E2. Overall, the QC-plates were a good indicator of the performance of the H295R Steroidogenesis Assay during a specific experiment.

Table 9.4: *Experiments and chemicals for which basal hormone concentration of testosterone (T) and estradiol (E2) was such that no decreases and/or increases after chemical exposure could be observed. Number indicate number of laboratories in which these issues occurred. Exp = Number of Repeat Experiment in which incidence occurred*

	Testosterone			Estradiol		
Chemical	***Exp 1***	***Exp 2***	***Exp 3***	***Exp 1***	***Exp 2***	***Exp 3***
Atrazine		1				
Aminoglutethimide		1	1		1	1
Benomyl				1		
EDS				1		
Forskolin						1
HCG				1		
Letrozole				1		
Molinate				1		
Nonoxynol-9						
Paraben				1		
Prochloraz			1		1	2
Trilostane						

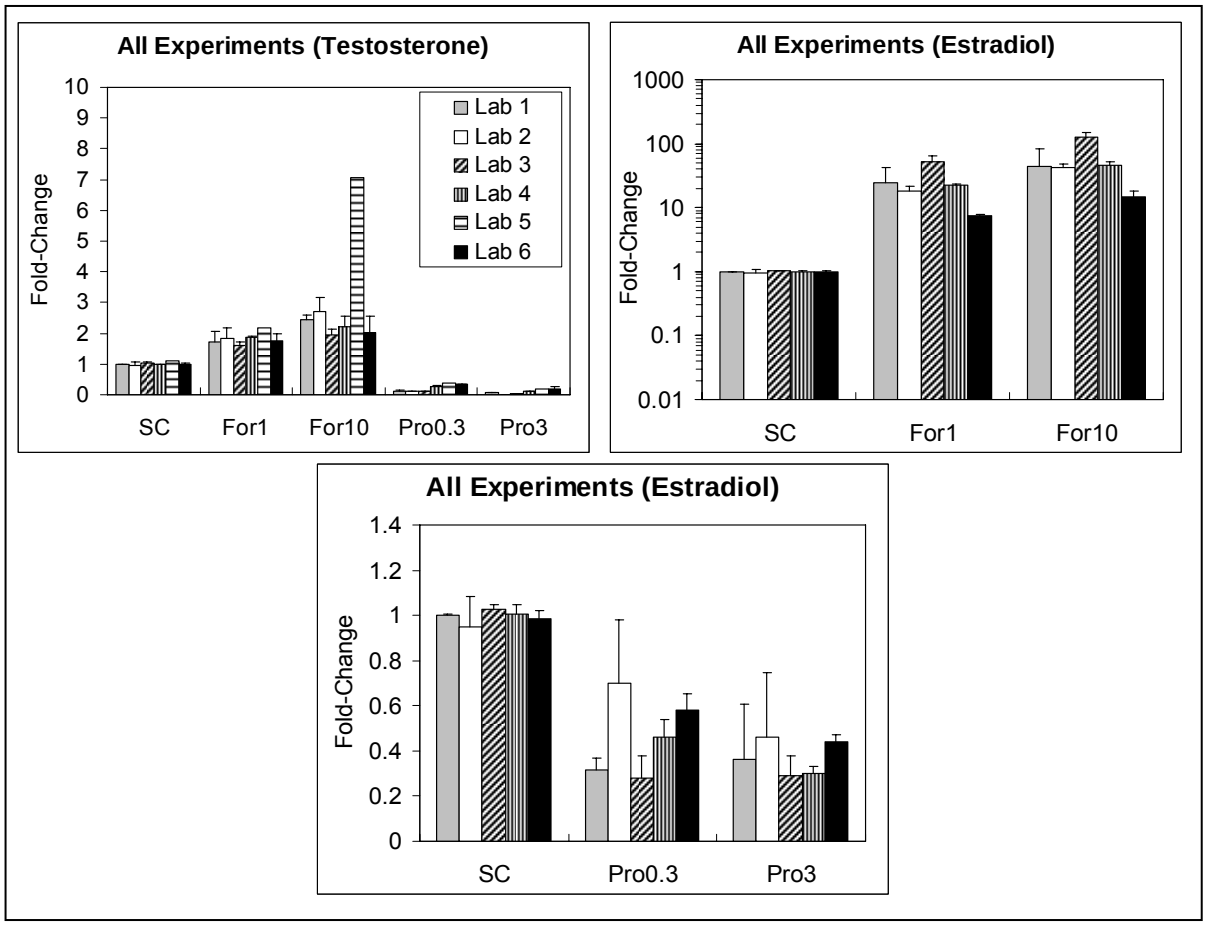


Figure 9.2: *Comparison of changes in the concentrations of testosterone (T) and estradiol (E2) relative to the solvent controls (SC=1) in the QC-plates among laboratories (Lab). For 1 = 1 µM Forskolin; For 10 = 10 µM Forskolin; Pro 0.3 = 0.3 µM Prochloraz; Pro 3 = 3 µM Prochloraz. Error bars = 1x standard deviation. Bars represent means of three replicate well (Lab 4 only 2 replicate wells). No data for E2 available from Lab5.*

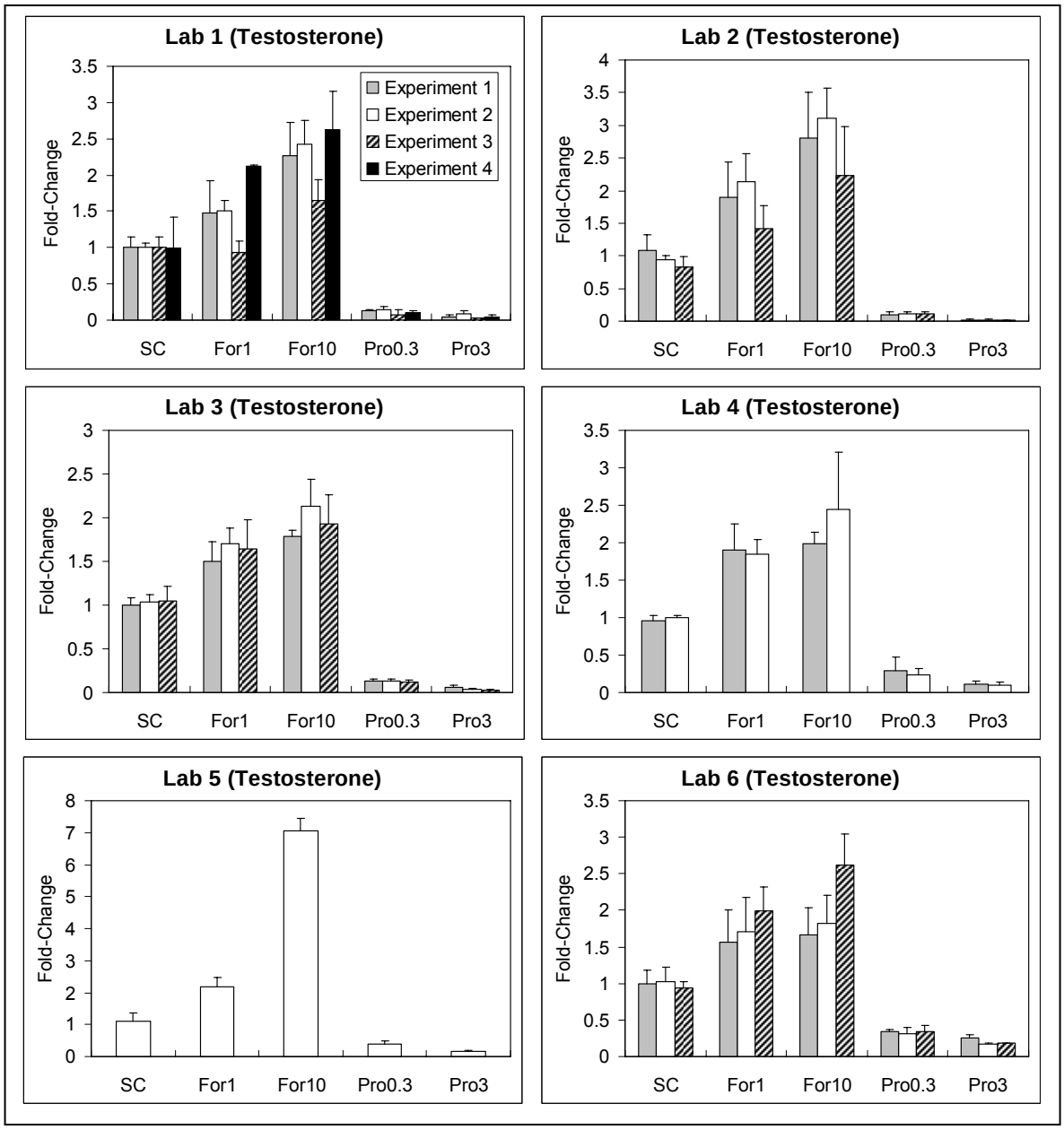


Figure 9.3: *Comparison of changes in the concentrations of testosterone (T) relative to the solvent controls (SC=1) in the QC-plat s run parallel to the exposure experiments at each laboratory (Lab). Each group represents a batch of chemicals for which a parallel QC-plate was run. For 1 = 1 μM Forskolin; For 10 = 10 μM Forskolin; Pro 0.3 = 0.3 μM Prochloraz; Pro 3 = 3 μM Prochloraz. Error bars = 1x standard deviation. Bars represent means of three replicate well (Lab 4 only 2 replicate wells).*

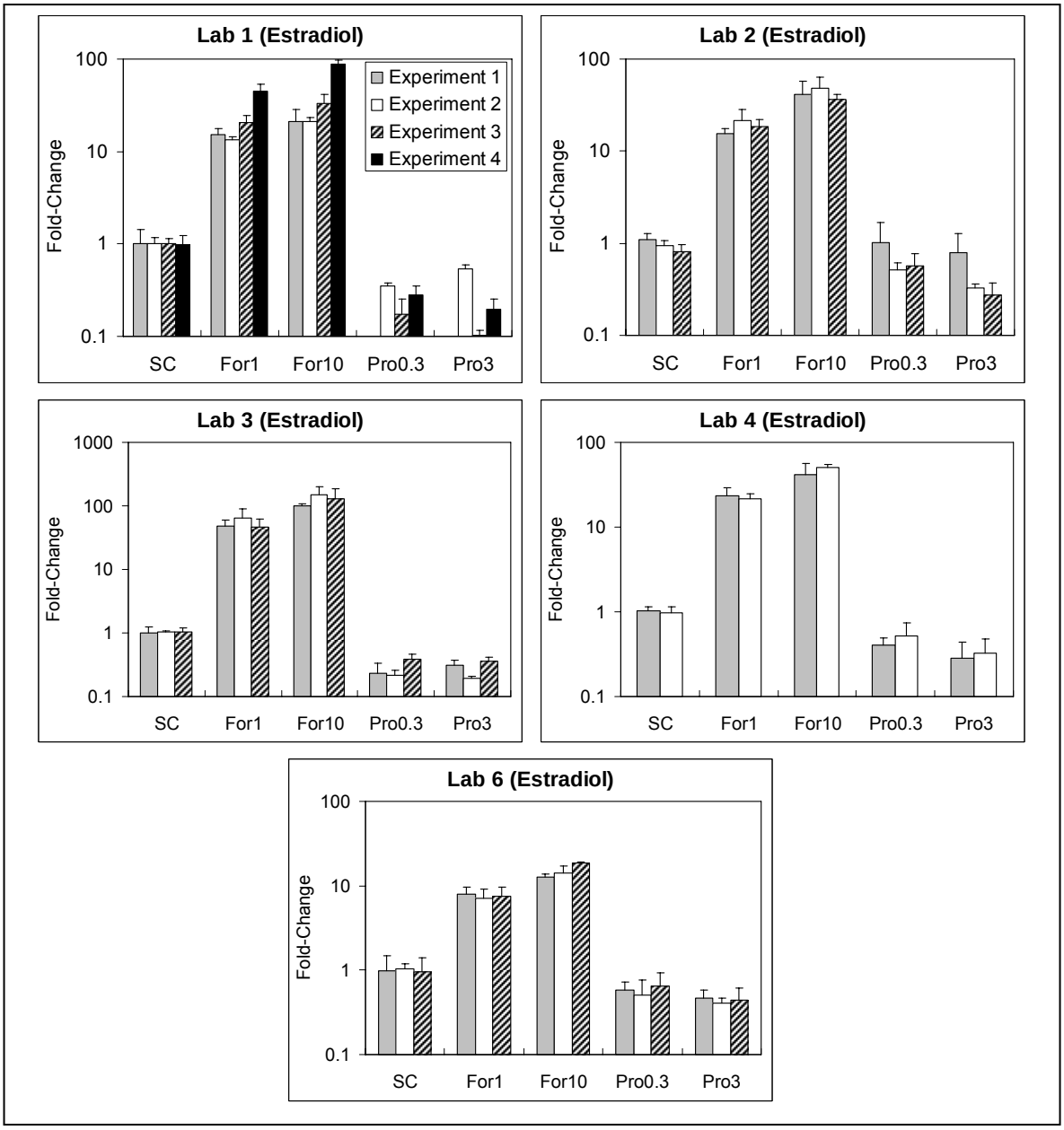


Figure 9.4: *Comparison of changes in the concentrations of estradiol (E2) relative to the solvent controls (SC=1) in the QC-plates run parallel to the exposure experiments at each laboratory (Lab). Each group represents a batch of chemicals for which a parallel QC-plate was run. For 1 = 1 μM Forskolin; For 10 = 10 μM Forskolin; Pro 0.3 = 0.3 μM Prochloraz; Pro 3 = 3 μM Prochloraz. Error bars = 1x standard deviation. Bars represent means of three replicate well (Lab 4 only 2 replicate wells). No data for E2 available from Lab5.*

9.2 Testing of chemicals

At the time this report was prepared, data from five laboratories were available for analysis. These data sets included the complete set of 12 core chemicals for four laboratories, as well as 8 out of the 12 core chemicals for the fourth laboratory. T data as measured by LC-MS were also available from a sixth group for letrozole, paraben, molinate and EDS. In addition, as discussed in Section 9.1.3, 4% of the T experiment and 10% of the E2 experiment were excluded due to issues related to basal hormone production (Table 9.4). Four of the five laboratories (1, 2, 3 and 4) had participated in the pre-validation studies, and therefore, were considered experienced with regard to the assay. The last laboratory (Lab 6), however, never conducted the H295R Steroidogenesis Assay prior to this validation study, and thus, was considered to provide information regarding the true transferability of the assay protocol. With very few exceptions (see subsequent sections) the performance of this laboratory was comparable to that of the other groups. Therefore, the data set presented here from the four laboratories can be assumed to be representative of the performance of the H295R Steroidogenesis Assay.

9.2.1 Cell Viability

Out of the 12 chemicals tested only four compounds were found to be cytotoxic (Figures 9.5 & 9.6). Benomyl, paraben, and prochloraz were cytotoxic at only the greatest dose tested (100 μM) with the exception of benomyl when measured at Labs 1 and 6 where cytotoxicity occurred either at the two greatest concentrations or where no cytotoxicity was observed, respectively. For nonoxyenol-9, effects on cell viability were observed at 10 μM for 2 out 4 laboratories that evaluated this compound. As a result, cytotoxic concentrations for these chemicals were excluded from further data analysis. An increase in cell viability greater than that observed in the solvent controls was observed at 4 out of 5 laboratories in the forskolin experiments (Fig. 9.5). Maximum inductions in cell viability relative to the controls observed for this chemical were 126 and 136% at the greatest dose at Labs 1 and 2, respectively, and 137% at Lab 6 at 10μM. In addition to the forskolin exposures, a greater than 20% increase in cell viability was observed for Lab 2 experiments with nonoxynol-9 (10 μM), trilostane (100 μM) and prochloraz (10 and 100 μM). This trend was opposite to that reported by the other groups. It is assumed that there might have been a technical problem with the cell viability assay because the hormone concentration data obtained for these doses behaved in a manner that was similar to those reported by the other groups (see Chapter 9.2.2). To verify this hypothesis, the cell viability experiments for prochloraz, benomyl, and trilostane were re-run by this laboratory. The data obtained during this

second set of experiments was in accordance with those obtained by the other groups (Figure 9.6). Finally, Lab 6 observed an increase in cell viability with molinate for doses greater or equal to 1µM, a trend that was not observed by the other groups. It is unclear what the reason for this increase in cell viability was.

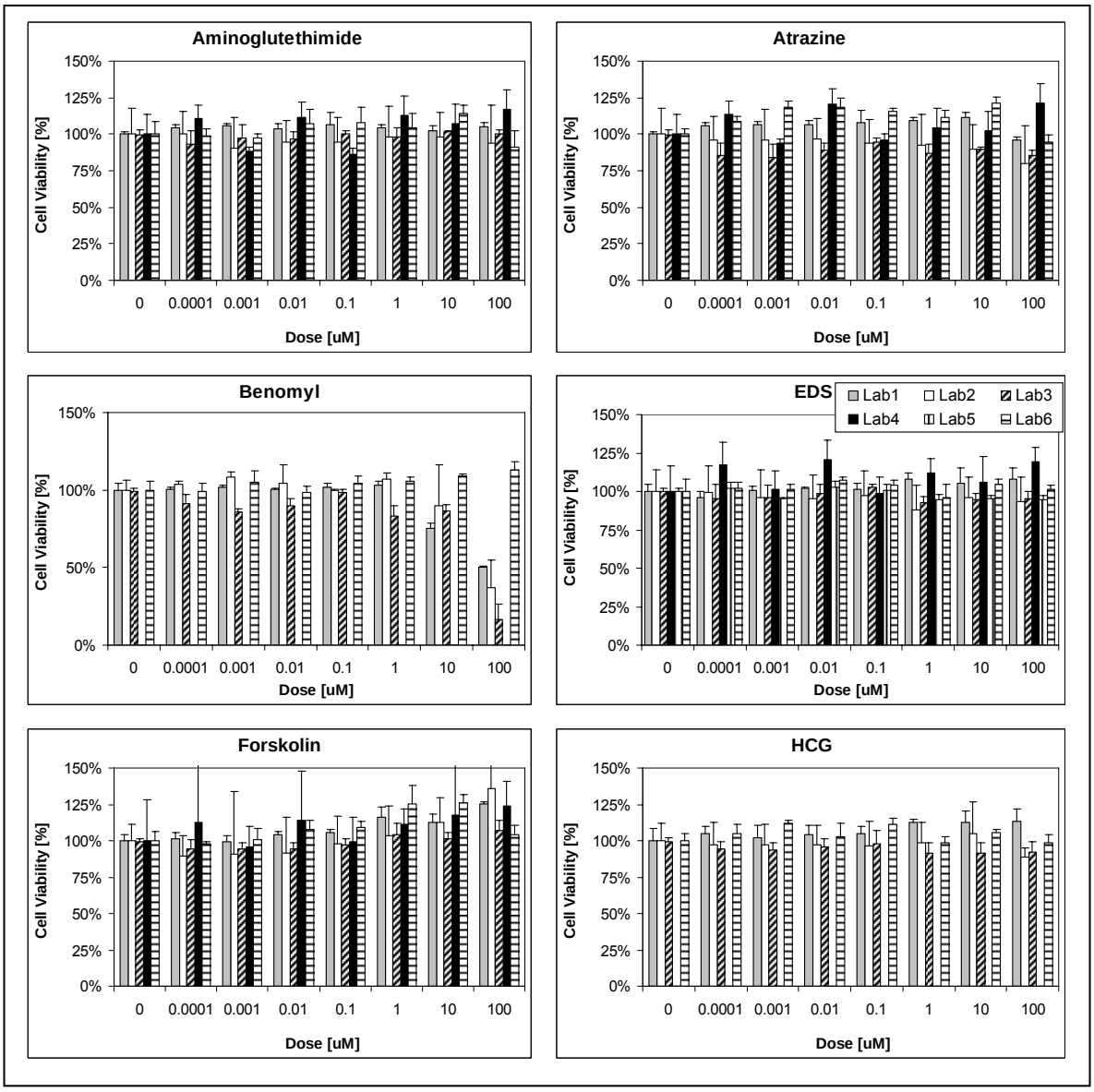


Figure 9.5: *Comparison of cell viability among laboratories (Lab) after exposure to Aminoglutethimide, Atrazine, Benomyl, EDS, Forskolin and HCG. Cell viability is expressed relative to the solvent controls (SC = 100%) in each plate. Error bars = 1 x SD. Lab4: No data available for HCG and benomyl. Lab6: No data available for aminogluthetimide, atrazine, forskolin and HCG.*

Given the variations (slight inductions or reductions in the number of viable cells), it was decided to normalize all data from wells with cell viabilities greater 80% for cell viability by dividing the hormone response by the relative viability (SC=1) in each well. All data from wells with cell viability of equal to or less than 80% was not considered for further evaluation due to potential interference through cytotoxicity.

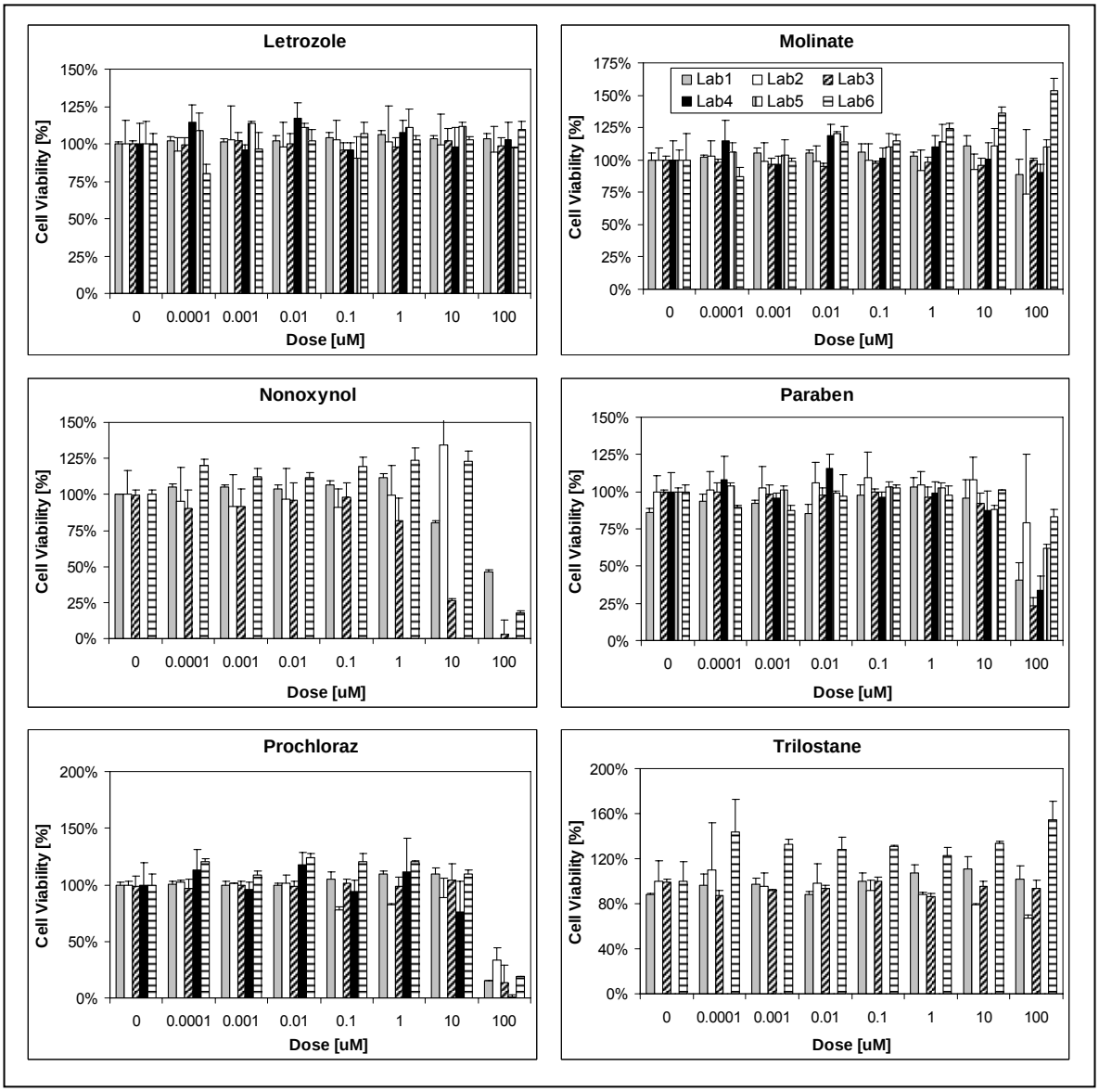


Figure 9.6: *Comparison of cell viability among laboratories (Lab) after exposure to Letrozole, Molinate, Nonoxynol-9, Paraben, Prochloraz and Trilostane. Cell viability is expressed relative to the solvent controls (SC = 100%) in each plate. Error bars = 1 x SD. Lab4: No data available for nonoxynol-9 and trilostane. Lab6: No data available for nonoxynol-9, prochloraz and trilostane.*

9.2.2 Relative changes after Exposure to Core Chemicals

9.2.2.1 Testosterone

There were marked differences in the response of T production after exposure of H295R cells to the 12 core chemicals (Figures 9.7 – 9.9). With few exceptions, the effects observed were comparable among laboratories and could be grouped in three different types of effects: inducers (Figure 9.7), inhibitors (Figure 9.8) and negative reference chemicals (Figure 9.9). Among the inducers, exposure to trilostane resulted in greatest fold changes (>10-fold induction) in T concentration when compared to solvent controls. The least fold-changes were observed for the atrazine exposures where induction of T production all were less than 1.5-fold with the exception of Labs 2 and 6, at which maximum inductions were 2.4- and 1.5-fold, respectively. Exposure to prochloraz resulted in greater15-fold reductions of T production at the greatest dose tested (100 µM) at all laboratories with the exception of Lab 4 where up to 4.5-fold reductions were observed. Exposure to the other inhibitors resulted in less than 4-fold changes in T production.

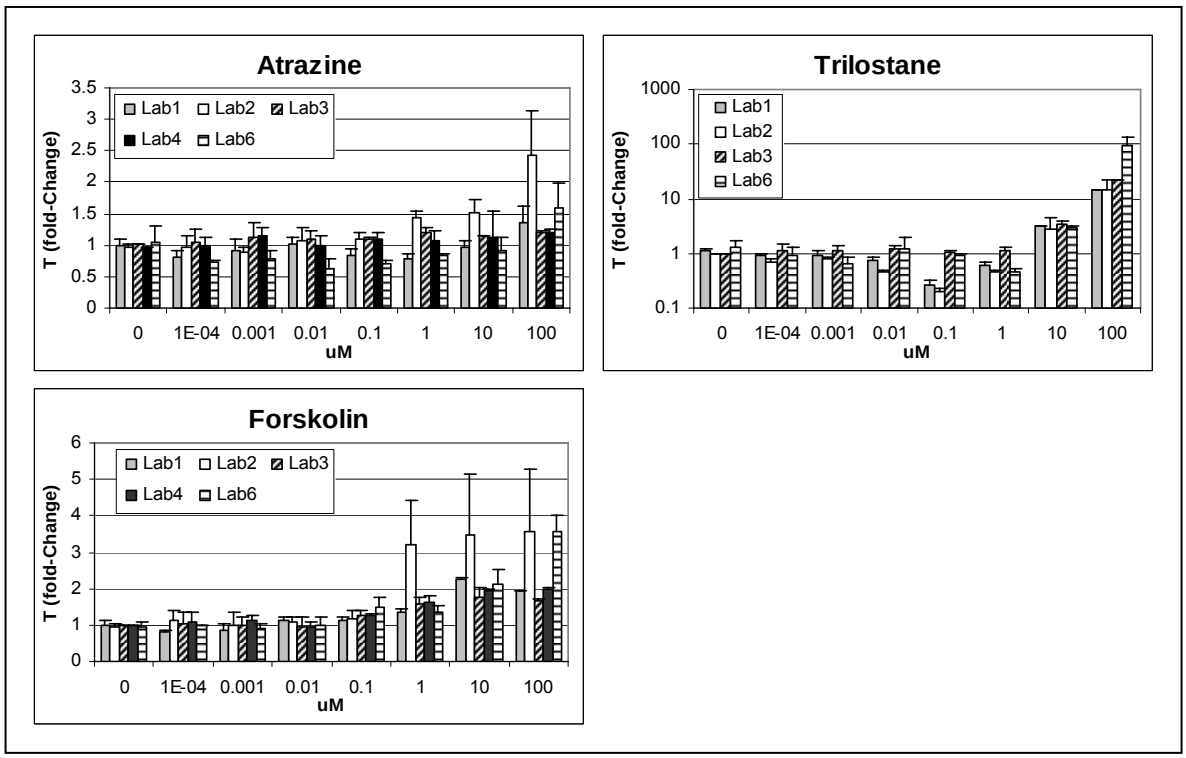


Figure 9.7: *Comparison of changes in the concentrations of testosterone (T) relative to the solvent controls (SC=1) after exposure to forskolin, trilostane, atrazine and paraben. Error bars = 1x standard deviation. Bars represent means of three replicate experiments (exceptions: Lab1 – only one replicate experiment was conducted; Lab3 – only two replicate experiments were conducted). Lab4: No data available for trilostane. Lab6: No data available for atrazine, forskolin and trilostane.*

When chemicals exhibited a less than 1.5-fold change in T production they were categorized as negatives (Fig. 9.9). Some of these negative chemicals could have been categorized as inhibitors in individual cases (Molinate: Lab 4; Benomyl: Lab1). However, even in situations where an inhibition was observed at an individual laboratory, this change was always less than 2-fold, and typically did not follow a dose-dependent trend. In the case of nonoxynol-9, a decrease in T concentrations at non-cytotoxic concentrations at two of four laboratories for which data was available was observed. Inhibitions were 29 and 47% relative to the SCs for Labs 1 (1 μM) and 2 (10 μM), respectively. However, it should be noted that at Lab 2 10μM nonoxynol-9 cause an average increase in cell viability of 38%, and thus, the observed reduction in T production may be an artifact due to the correction for cell viability, especially as no such increase was observed at any of the other groups.

Variation between laboratories did not exceed 2-fold for a given dose with the exception of trilostane (see Chapter 9.2.2.3 for discussion). Among group CVs for inducers were always less than 1.5-fold (no trilostane).

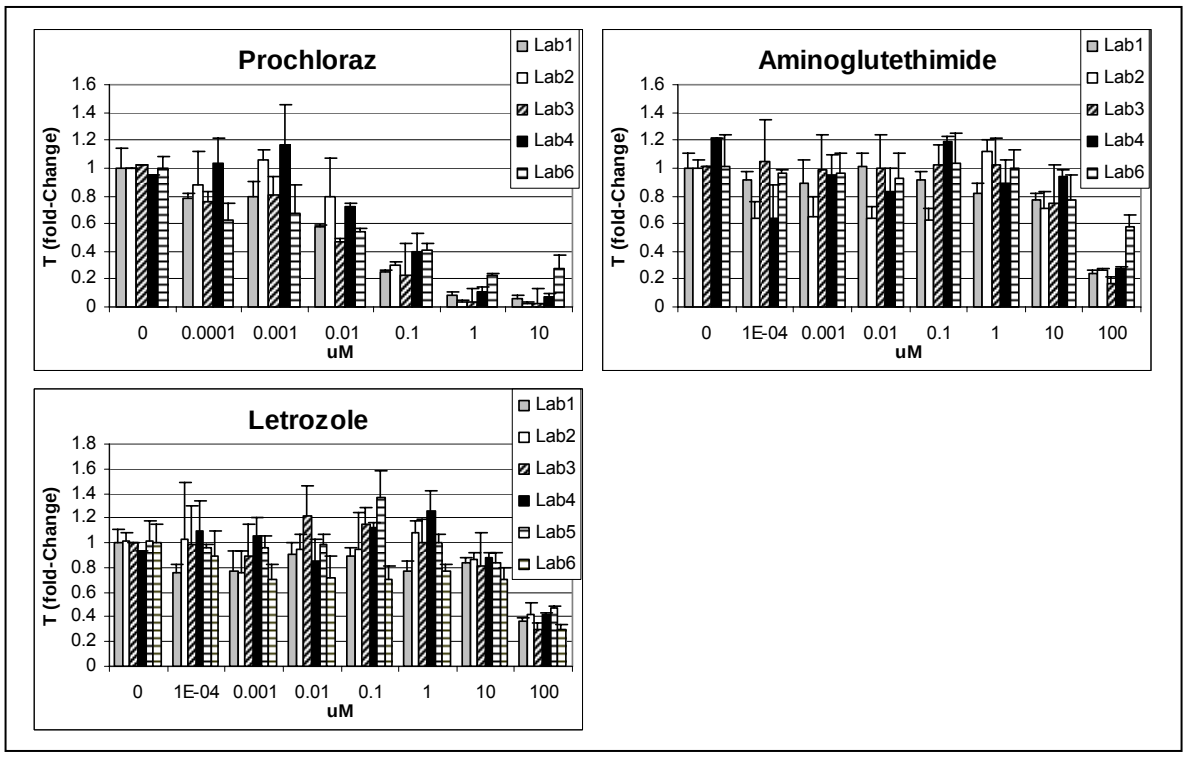


Figure 9.8: *Comparison of changes in the concentrations of testosterone (T) relative to the solvent controls (SC=1) after exposure to prochloraz, aminoglutethimide, and letrozole. Error bars = 1x standard deviation. Bars represent means of three replicate experiments (exceptions: Lab1 – only one replicate experiment was conducted; Lab3 – only two replicate experiments were conducted). Lab6: No data available for prochloraz and aminogluthetimide.*

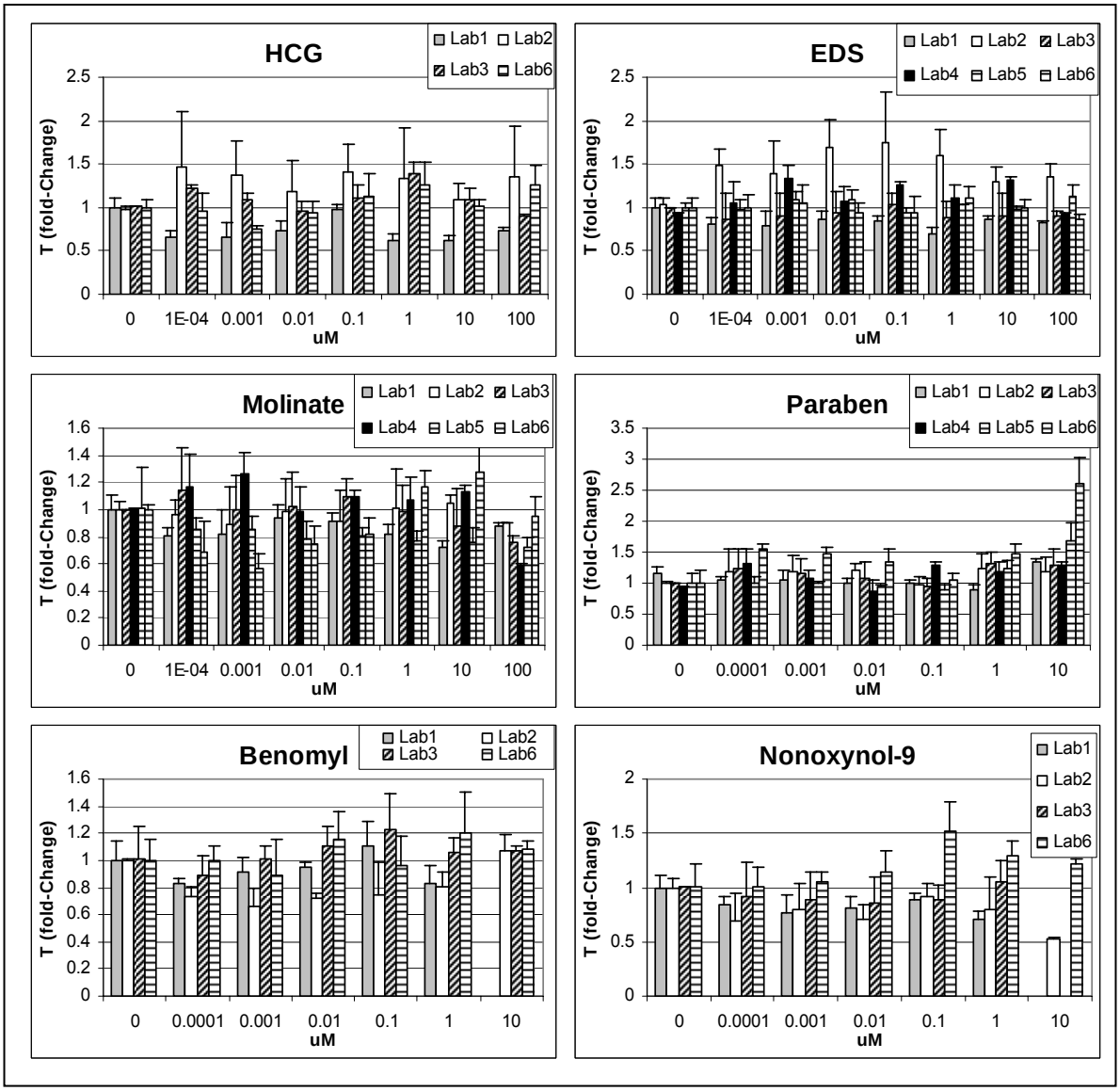


Figure 9.9: *Comparison of changes in the concentrations of testosterone (T) relative to the solvent controls (SC=1) after exposure to HCG, EDS, molinat, benomyl and nonoxynol-9. Error bars = 1x standard deviation. Bars represent means of three replicate experiments (exceptions: Lab1 – only one replicate experiment was conducted; Lab3 – only two replicate experiments were conducted). Lab4: No data available for HCG, benomyl and nonoxynol-9. Lab6: No data available for HCG, benomyl and nonoxynol-9.*

9.2.2.2 Estradiol

Significant differences in the response of E2 production was observed for H295R cells exposed to 12 core chemicals (Figures 9.10 – 9.12). The direction of effects for each chemical was comparable among laboratories with the exception of the Lab2 trilostane data (see Chapter 9.2.3 for discussion). Overall, the types of effect were slightly different than those observed for T with the majority of the chemicals acting as inducers of E2 production (Figure 9.10). Three chemicals inhibited E2 concentrations (letrozole, prochloraz and aminoglutethimide; Figure 9.11) while HCG, EDS, benomyl and nonoxynol-9 (Figure 9.12) did not elicit any clear (> 1.5-fold) effects at non-cytotoxic doses.

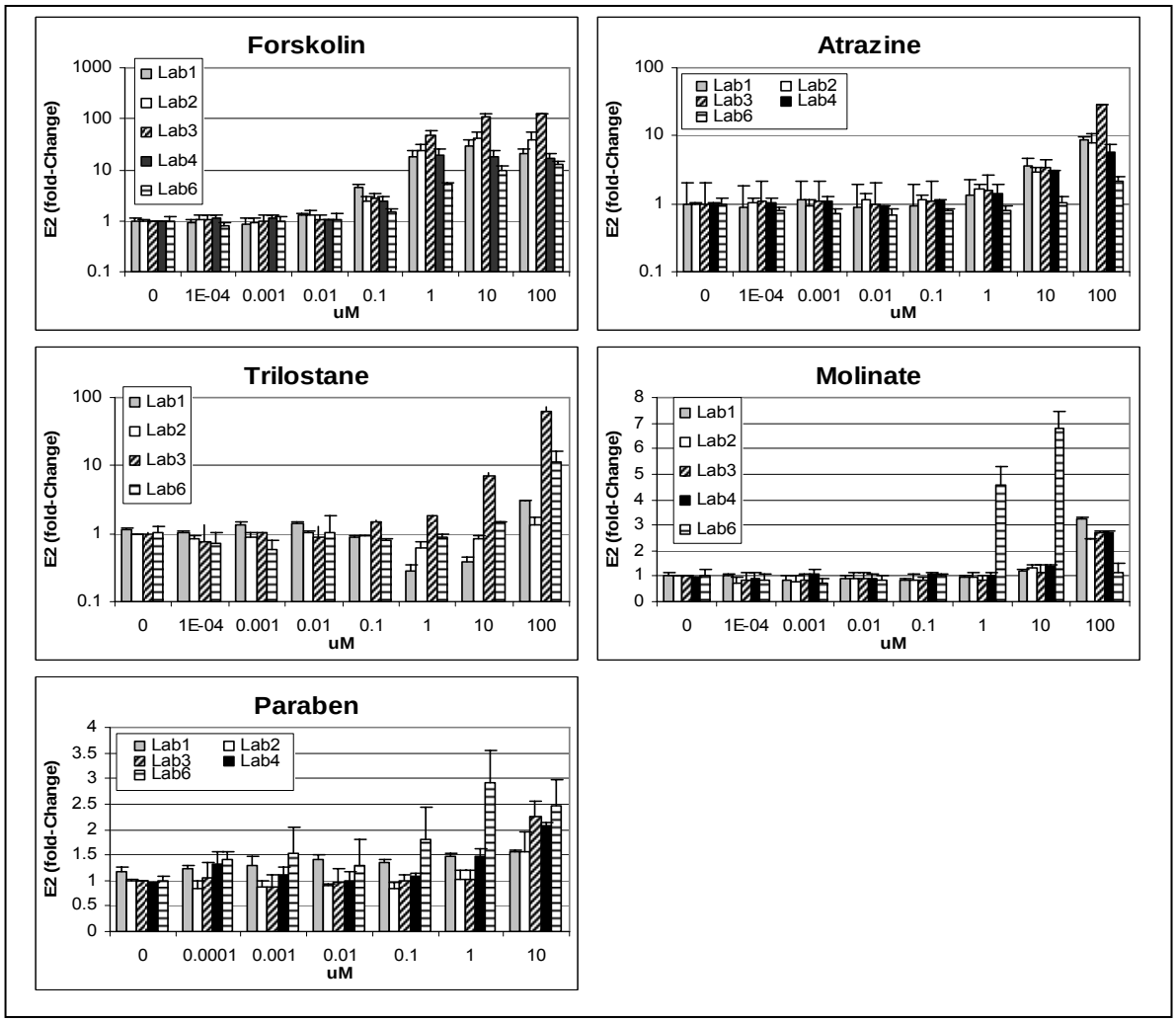


Figure 9.10: *Comparison of changes in the concentrations of estradiol (E2) relative to the solvent controls (SC=1) after exposure to forskolin, atrazine, trilostane,molinate,and paraben. Error bars = 1x standard deviation. Bars represent means of three replicate experiments (exceptions: Lab1 – only one replicate experiment was conducted; Lab3 – only two replicate experiments were conducted). Lab4: No E2 data available for trilostane. Lab6: No data available.*

The strength of the response to the exposure with chemicals that increased production of E2 ranged between 20-fold or greater (Forskolin) to <3-fold (Paraben). The dose at which effects occurred was not related to the magnitude of the response. While forskolin resulted in increases in E2 production at doses greater or equal to 0.1 μM exposure to other inducers typically did not reveal effects at doses less than 1 μM. Exposure to letrozole and prochloraz resulted in marked reductions of E2 at doses greater 0.001 and 0.1 μM, respectively. Exposure to aminoglutethimide, in contrast, only caused a clear reduction in E2 concentrations at the greatest dose tested.

Variation between laboratories did not exceed 2-fold for a given dose with the exception of trilostane (see Chapter 9.2.2.3 for discussion).

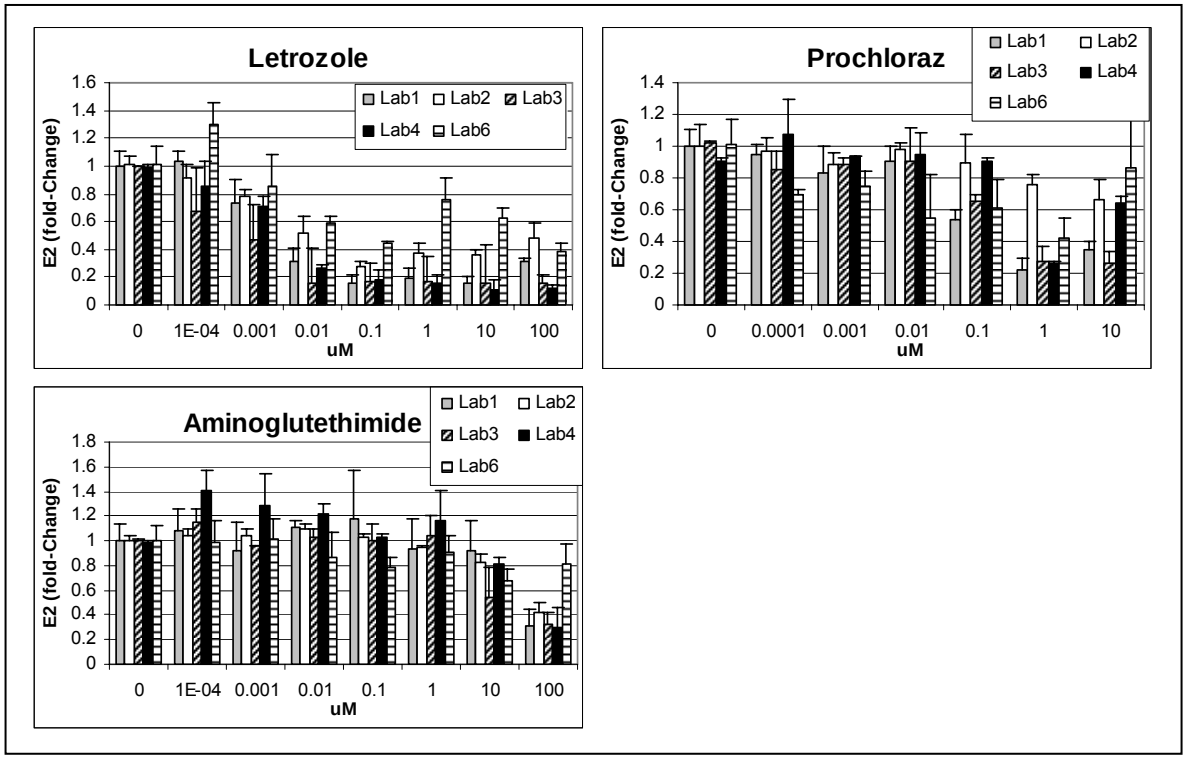


Figure 9.11: *Comparison of changes in the concentrations of estradiol (E2) relative to the solvent controls (SC=1) after exposure to letrozole, prochloraz and aminoglutethimide, letrozole. Error bars = 1x standard deviation. Bars represent means of three replicate experiments (exceptions: Lab1 – only one replicate experiment was conducted; Lab3 – only two replicate experiments were conducted). Lab6: No E2 data available.*

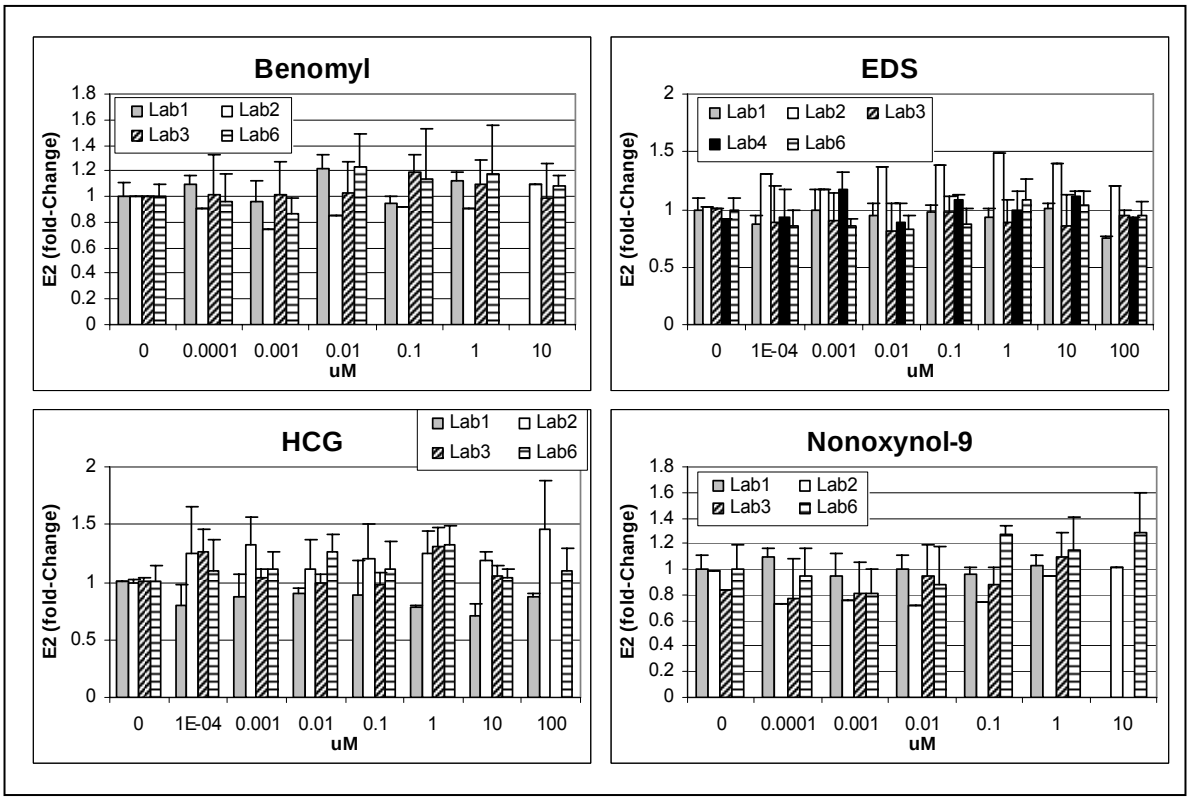


Figure 9.12: *Comparison of changes in the concentrations of estradiol (E2) relative to the solvent controls (SC=1) after exposure to HCG, EDS, and benomyl. Error bars = 1x standard deviation. Bars represent means of three replicate experiments (exceptions: Lab1 – only one replicate experiment was conducted; Lab3 – only two replicate experiments were conducted). Lab4: No data available for benomyl, , HCG and nonoxynol-9. Lab6: No E2 data available.*

9.2.3 Confounding Factors – Interference with Hormone Detection Assays

The analysis of cross-reactivity of the greatest doses of each chemical with the antibodies of the immunoassays utilized at the majority of the laboratories revealed marked interaction of some few test compounds (Table 9.5). Interestingly, all laboratories with the exception of Lab2 showed a very great interaction of the estradiol immunoassay with trilostane. Similarly, cross-reactivity of this chemical was also reported for the T antibodies, albeit less pronounced. Only the greatest chemical dose was analyzed for cross-reactivity with the antibodies, and thus, no adjustment of the dose-response curves could be performed. However, an initial attempt to correct for the interaction with the antibodies at this greatest dose (greatest three doses for Lab1) indicated that while the induction of E2 after exposure to trilostane is likely to be solely due to this cross-reactivity while the induction of T could

not be explained by this factor (Figure 9.12). Similar interaction of trilostane with hormone detection systems have been also observed by other authors (Puddefoot et al. 2002; Villeneuve, personal communication). Furthermore, nonoxynol-9, paraben, and prochloraz also interacted with the E2 immunoassays. However, since the cross-reactivity for prochloraz, paraben, and nonoxynol-9 was either low at the greatest doses tested or these doses were excluded due to marked cytotoxicity, this factor had no effect on the interpretation of the results. However, further analyses are required to address possible uncertainties resulting from the interference of a test chemical with the hormone detection system utilized.

Table 9.3: *Interference (pg hormone/mL @ the greatest dose tested) of the 12 core chemicals with the immunoassays used for the determination of testosterone (T) and estradiol (E2). The numbers indicate the concentration (pg/mL) of T or E2 measured by the respective hormone immunoassay in supplemented stock medium with 100µM of the chemical of interest added. Note: This medium has not been in contact with the cells.*

	Interference with Hormone Detection Assay							
	Testosterone (pg/ml @ 100µM)				***Estradiol (pg/ml @ 100µM)***			
	Lab 1*	**Lab 2***	**Lab 3***	**Lab 4***	**Lab 1***	**Lab 2***	**Lab 3+**	**Lab 4***
Aminoglutethimide								24
Atrazine				59				
Benomyl				n/d				n/d
EDS								22
Forskolin								
HCG				n/d				n/d
Letrozole	33		237	34				12
Molinate								11
Nonoxynol	137			n/d	**394**	11		n/d
Paraben	n/d			20	n/d	**15**	**24**	**20**
Prochloraz	94	15	169	57			**20**	11
Trilostane	**3719**	**6554**	**44303**	n/d	**802**	9	**1112**	n/d

* only chemicals that were @ >9pg/mL interference were considered
+ only chemicals that were @ >5pg/mL interference were considered
n/d = not determined

Bold numbers	Result has or is likely to have impact on final result
Shaded bold numbers	Not of relevance due to parallel cytotoxicty that resulted in omission of data point

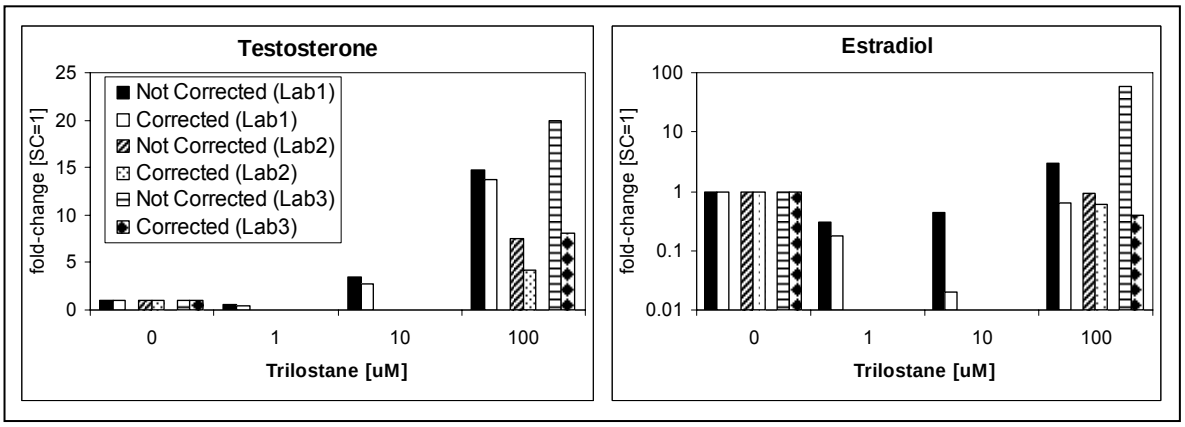


Figure 9.13: *Changes in the concentrations of testosterone (T) and estradiol (E2) relative to the solvent controls (SC=1) after exposure to trilostane with and without adjustment of final hormone concentration for interference with hormone detection system.*

10 DATA EVALUATION

In the previous sections, despite some minor remaining uncertainties, it was demonstrated that the H295R Steroidogenesis Assay is a sensitive, reproducible, transferable, and specific test for the determination of chemical effects on the production of T and E2. One of the key questions that needs to be addressed during the development and validation of an *in vitro* assay for screening purposes is the format in which the data are presented and the type of analyses that will be needed for a meaningful interpretation of these data and classification of chemicals of concern. In this section we discuss several different approaches for the presentation and evaluation of the data obtained during this phase of the validation studies then select an approach for the assessment of data obtained with the H295R Steroidogenesis Assay. In a concluding section, we then will subject an extended data set including 14 additional chemicals that has been produced by the lead laboratory to the selected data evaluation procedures to verify the selected approach with a number of compounds with unknown modes of interaction with the assay.

Based on results obtained during the pre-validation studies (Hecker et al. 2006b & 2007b) it was decided that data should be expressed as fold-change relative to the solvent control due to variation in basal hormone production that was a function of cell passage and/or freeze/thaw cycle (for a detailed discussion please refer to the above listed reports). For this validation study, several different data evaluation approaches were utilized including the definition of induction and inhibition thresholds, calculation of effective concentrations (EC50) where possible, and expression of data relative to a certain

concentration of the model inducers and inhibitors analyzed on the QC-plates run in parallel to each exposure experiment (percent of control [PC] approach). This approach has been successfully applied in previous validation efforts of other *in vitro* tests such as hormone receptor transcriptional assays (Takeyoshi, 2007). It is based on the concept of the potency of a test chemical relative to a model inducer or inhibitor that was analyzed in the QC-plate run in parallel with each experiment. In this study we used to PC endpoints, PC50 and PCmax:

- PC50 - the concentration of a test chemical at which the measured induction or inhibition is 50% of the activity induced by a model inducer (forskolin) or model inhibitor (prochloraz), respectively, analyzed in a parallel QC-plate.
- PCmax - the concentration of a test chemical at which the measured response is the maximum induction or inhibition compared to the activity caused by a model inducer (forskolin) or model inhibitor (prochloraz), respectively, analyzed in a parallel QC-plate. Both the concentration and % of maximum activity of the positive control should be given.

10.1 Fold-Change Evaluation

Data expressed as fold-change relative to the SC was subjected to an ANOVA followed by the Dunnett's test to identify significant changes from the SC. Furthermore, changes were classified into four categories to allow for separation of chemicals by the magnitude of effect the evoked (Table 10.1).

Table 10.1: *Classification system for the effects of the 12 core chemicals on testosterone (T) and estradiol (E2) production by H295R cells. Note: Only statistically significant effects are to be considered.*

Category	Inducer	Change	Inhibitor	Change
Weak	⇧	<2-fold	⇩	<2-fold
Medium	⇧⇧	2- to <4-fold	⇩⇩	2- to <4-fold
Strong	⇧⇧⇧	4- to <20-fold	⇩⇩⇩	4- to <10-fold
Very Strong	⇧⇧⇧⇧	≥20-fold	⇩⇩⇩⇩	≥10-fold

A comparison of lowest observed effect concentrations (LOECs), which was defined as the least concentration at which a statistically significant change occurs, revealed comparable dose-ranges among laboratories with only a few exceptions (Table 10.2). These exceptions included the responses in T production in the molinate and nonoxynol-9

experiments conducted at Labs 4 and 1, respectively. For molinate, Lab 4 reported a significant and very strong decrease in T production in one of three experiments. Results obtained during all other repeat experiments were comparable to the responses observed for this chemical at the other laboratories. For nonoxynol-9 at Lab 1, a weak decrease in T concentrations was observed (29 +/- 14% decrease compared to SCs) that was not reported by the other groups.

The magnitude of change in T and E2 production for experiments conducted with the core chemicals was similar among laboratories. In 75 and 67% of all experiments results did not deviate by more than one category among laboratories for T and E2, respectively. Some of this variability, however, originated from Lab 6, which conducted the assay the first time and also tested each chemical only in one experiment. While most of the data obtained by this group still showed the same trend as that obtained by the other laboratories (none of the results showed an opposite trend), it can be concluded that multiple repeat experiments (preferably three) are desired to be able to address some of this variation. Most of the remaining variability was associated with Lab 4. This group only conducted two repeat measurements (duplicate wells) per treatment and experiment, and thus, the power to detect significant changes was less than for the other laboratories.

It should also be noted that the data obtained from the trilostane experiments has to be assessed with care due to the uncertainties resulting from the marked cross-reactivity that this chemical exhibited with most of the hormone detection assays utilized by the laboratories.

In summary, the categorical system (Table 10.1) in combination with the LOEC data was capable of clearly identifying inducers and inhibitors of different strengths for both T and E2 production in the H295R steroidogenesis assay. One uncertainty with this approach that still needs additional evaluation is for chemicals that only weakly affect hormone production; it is not always possible to distinguish between weak effectors and negative chemicals. In the subsequent sections we attempted to further characterize the results obtained during these validation studies to improve the evaluation approach by utilizing regression models, such as probit models, to determine effective concentrations and potency curves relative to the model inducer and inhibitor forskolin and prochloraz, respectively.

***Table 10.2**: Lowest observed effect concentrations (LOECs; measured by Dunnett's test) and strength and direction of change (arrows; see Table 10.1 for explanation) for testosterone (T) and estradiol (E2) after exposure to the twelve core chemicals. Ranges refer to maximum values measured in repeat-experiments. nd – not detectable; --- chemical not analyzed. Gray shaded cells – uncertainty due to cross-reactivity.*

	Fold-Change (Testosterone)											
	Lab 1[a]		Lab 2		Lab 3		Lab 4		Lab 5[a]		Lab 6[a]	
	LOEC	*Max Change*	*LOEC*	*Max Change*	*LOEC*	*Max Change*	*LOEC*	*Max Change*	*LOEC*	*Max Change*	*LOEC*	*Max Change*
Aminoglutethimide	100	⇩⇩⇩	10 - 100	⇩⇩	10-100	⇩⇩⇩	1-100	⇩⇩	---	---	100	⇩
Atrazine	100	⇧	10 - 100	⇧⇧	100-nd	⇧	Nd		---	---	100	⇧
Benomyl	nd		nd		Nd		---	---	---	---	nd	
EDS	nd		nd		Nd		Nd		nd		nd	
Forskolin	10	⇧⇧	1	⇧⇧	0.1 - 1	⇧	1	⇧⇧	---	---	10	⇧⇧
HCG	nd		nd		Nd		---	---	---	---	nd	
Letrozole	100	⇩⇩	100	⇩⇩	100[a]	⇩⇩	100	⇩⇩	100	⇩	100	⇩⇩
Molinate	nd		nd		100 - nd	⇩	10 – nd[b]	(⇧⇧⇧⇧)	nd		nd	
Nonylhenol	1	⇩	10-nd	⇩	Nd		---	---	---	---	nd	
Paraben	10	⇧	1 - nd	⇧	1 - nd	⇧	1-nd	⇧	10	⇧	10	⇧⇧
Prochloraz	0.0001	⇩⇩⇩⇩	0.1-1	⇩⇩⇩⇩	0.0001 - 0.01	⇩⇩⇩⇩	0.001 - 0.1	⇩⇩⇩	---	---	0.0001	⇩⇩
Trilostane	1	⇧⇧⇧	0.01-10	⇧⇧⇧⇧	10-100	⇧⇧⇧⇧	---	---	---	---	100	⇧⇧⇧⇧

	Fold-Change (Estradiol)											
	Lab 1[a]		Lab 2		Lab 3		Lab 4		Lab 5		Lab 6[a]	
	LOEC	*Max Change*	*LOEC*	*Max Change*	*LOEC*	*Max Change*	*LOEC*	*Max Change*	*LOEC*	*Max Change*	*LOEC*	*Max Change*
Aminoglutethimide	100	⇩⇩	10[a]	⇩⇩	10	⇩⇩	Nd	⇩⇩	---	---	nd	
Atrazine	10	⇧⇧⇧	1 - 10	⇧⇧	10	⇧⇧⇧⇧	10-100	⇧⇧⇧	---	---	100	⇧⇧
Benomyl	nd		nd		nd[a]		---	---	---	---	100	⇧
EDS	nd		nd		Nd		Nd		---	---	nd	
Forskolin	1	⇧⇧⇧⇧	1	⇧⇧⇧⇧	1	⇧⇧⇧⇧	1	⇧⇧⇧	---	---	1	⇧⇧⇧
HCG	nd		100- nd	⇧	nd[a]		---	---	---	---	nd	
Letrozole	0.001	⇩⇩⇩	0.0001-0.001	⇩⇩	0.0001-0.001	⇩⇩⇩	0.01	⇩⇩⇩⇩	---	---	(0.001)[c]	⇩⇩
Molinate	100	⇧⇧	10 - nd	⇧⇧	100	⇧⇧	100	⇧⇧	---	---	nd[d]	(⇧⇧⇧)
Nonoxynol-9	nd		nd		Nd		---	---	---	---	nd	
Paraben	1	⇧	10-nd	⇧⇧	10	⇧⇧	1-10	⇧⇧	---	---	1	⇧⇧
Prochloraz	0.1	⇩⇩⇩	10[a]	⇩⇩	0.1	⇩⇩⇩	1	⇩⇩	---	---	1	⇩⇩
Trilostane	10	⇧⇧	100-nd	⇧	100	⇧⇧⇧⇧	---	---	---	---	100	⇧⇧⇧

[a] only one experiment was conducted or considered for data evaluation; [b] Effect only observed in one of three experiments; [c] Effect occurred at all doses except for 0.1 μM;
[d] No effect at greatest dose but at lesser doses

10.2 Application of Regression Type Models – Effective Concentrations (ECs)

An alternative approach to further characterize the results obtained during the validation studies was to use regression models to calculate EC50s with probit-transformed data. The EC50 is defined as the chemical concentration that provokes a response halfway between the solvent control baseline and the maximum response for the chemical. For the 12 core chemicals tested in this study, only two chemicals had sufficient T data for conducting a probit analysis while for E2, only 4 chemicals had sufficient data for this type of analysis. For chemicals that had sufficient data, the analyses revealed comparable EC50s for most of these chemicals at the different laboratories. The exceptions to this were for Lab 4 atrazine and paraben T data, Lab 3 benomyl T and E2 and paraben T data, Lab 6 benomyl and forskolin T data, Lab 2 molinate, nonoxynol-9 and paraben T data, and Lab 1 nonoxynol-9 T data (Table 10.3). Inconsistencies in the determination whether a chemical could be categorized as an effector (either an inhibitor or an inducer) vs a negative occurred in 18 and 4% of all cases for T and E2, respectively, and the maximum difference between EC50s among all laboratories for a given chemical was less than one order of magnitude with exception of the Lab 6 T data in the forskolin exposure experiment. However, such inconsistencies only occurred for weak effectors and none of the strong interacting compounds or the negatives EDS and HCG showed such variation among laboratories.

One major issue with using a regression model to calculate EC50 was the lack of sufficient data to fully describe dose-response profiles for many of the chemicals tested. While the data from most of the strong effectors was sufficient to derive a full dose-response curve necessary for conducting a probit regression analysis, medium and weak inducers/inhibitors of hormone production did not produce data that described maximum induction or inhibition values of hormone production such that regression model could be used. Therefore, in these cases Table 10.3 reports that the EC50 is a value greater than the no effect concentration (NOEC). Furthermore, for some chemicals calculation of effective concentrations was complicated by the occurrence of decreases in cell viability at doses near the upper testing range used in this study, and by the relatively large spacing (order of magnitude) of doses. Given these uncertainties and limitations, it can be concluded that the utilization of this regression approach, using the current H295R Steroidogenesis Assay protocol is of limited value in categorizing chemicals.

However, with the exception of chemicals that limit the dose range due to cytotoxicity, the H295R assay protocol could be refined to allow for the determination of a

more detailed and complete dose-response curves that would allow the calculation of effective concentrations (ECx). Thus, the current H295R assay protocol could be used as an initial screen to identify chemicals of potential concern and then a second experiment could be conducted with chemical that tested positive, but with a dosing-regime more closely grouped around the concentration for which a response was observed.

Table 10.3: *EC50s(μM) calculated for the changes in testosterone (T) and estradiol (E2) production after exposure to the twelve core chemicals using a probit model. --- chemical not analyzed. Greater than values indicate insufficient data to derive a complete dose response curve, as a result the data is present as > NOEC. Gray shaded cells – uncertainty due to cross-reactivity.*

	Testosterone					
	Lab 1	**Lab 2**	**Lab 3**	**Lab 4**	**Lab 5**	**Lab 6**
Aminoglutethimide	>10	>10	>10	>10	---	>10
Atrazine	>10	>10	>10		---	>10
Benomyl	>10	>10		---	---	
EDS						
Forskolin	>1	1.3	0.30	0.65	---	45
HCG				---	---	
Letrozole	>10	>10	>10	>10	>10	>10
Molinate	>10		>10	>10	>10	>0.1
Nonoxynol-9	>1	>10		---	---	
Paraben	>10				3.7	2.93
Prochloraz	0.0099	0.068	0.0075	0.028	---	0.0026
Trilostane	>10	>10	>10	---	---	>10
	Estradiol					
	Lab 1	**Lab 2**	**Lab 3**	**Lab 4**	**Lab 5**	**Lab 6**
Aminoglutethimide	>10	>10	>10	>10	---	>1
Atrazine	>10	>10	>10	>10	---	>10
Benomyl	>10	>10		---	---	>10
EDS					---	
Forskolin	0.65	1.94	1.75	1.62	---	3.33
HCG				---	---	
Letrozole	0.0022	0.0037	0.00052	0.0025	---	>0.001
Molinate	>10	>10	>10	>10	---	>0.1
Nonoxynol-9				---	---	
Paraben	>1	>10	>10	2.7	---	>0.1
Prochloraz	0.068	0.27	0.040	>0.1	---	>0.01
Trilostane	>10	>10	>10	---	---	>10

10.3 Application of the Percent of Control (PC) Concept

To address some of the issues encountered with the probit regression analysis method, an alternative approach was used. In this approach, hormone production data from chemically exposed cells were normalized to the hormone concentrations associated with either the prochloraz (inhibitor) or forskolin (inducer) that were analyzed in the QC-plate that was run in parallel with each experiment. The normalization was done in accordance with the nature of the observed response, that is a chemical exposure that resulted in a reduction in hormone production was normalized to prochloraz while chemical the elevated hormone production were normalized to forskolin. Normalization of inducers was made to the maximum forskolin dose for both T and E2 while for inhibitors the lesser prochloraz dose (0.3μM) was used for E2 due to the potential for interference at the prochloraz. The selection of the prochloraz normalization dose seems justified given that this dose elicited maximum or close to maximum responses at most laboratories. Based on this normalization procedure, it was possible to derive normalized responses for most chemicals that were equal to or greater than 50% of the response observed for forskolin (inducers) and prochloraz (inhibitors), i.e., the PC50. For chemicals where a response did not reach 50% of the response of the appropriate model chemical, the dose at which the maximum response occurred (PCmax) was reported. Using this evaluation approach chemicals were grouped based on their relative potency compared to the model compounds (Tables 10.4). Compounds were classified as "negatives" when no dose-dependent and statistically significant differences compared to the SCs as determined in the above presented fold-change approach occurred. The categories in which the different chemicals could be grouped were similar to those determined using the fold-change approach with few exceptions (Section 10.1). Thus, when the data are evaluated as percent effect of the response caused by either forskolin or prochloraz, the results appear to objectively reflect the true response of chemical on steroidogenesis. The advantage of this approach is that it integrates both the dose at which an effect occurs as well as the magnitude of the change evoked by a chemical.

One uncertainty that still needs to be addressed pertains to the classification of T inducers. For this case, the model inducer forskolin only caused a moderate induction, which is likely the cause for the observed variation in the classification of chemicals as inducers (Tables 10.4). While the utilization of forskolin allowed the categorization of chemicals based on the strength of the response they elicited, we are further evaluating additional chemicals regarding their potential to induce T production by H295R cells, and will consider replacing forskolin with such a compound in the QC plate that represents a

strong inducer of T (desired maximum induction ≥ 10-fold). One candidate that is currently explored as a replacement chemical for forskolin is the growth promoter trenbolone. Initial results indicate that trenbolone is a strong inducer of both T and E2.

While the PC approach provides valuable information regarding the classification of a chemical as an inducer or inhibitor of T and/or E2 production, it should only be used in combination with the fold-change evaluation procedures described in chapter 10.1.

Table 10.4: *Chemical concentrations (μM) required to elicit 50% of the response caused by forskolin (inducers) and prochloraz (inhibitors). Increasing intensities of yellow/red indicate increasing inhibiting potency of a chemical; Increasing intensities of blue indicate increasing inducing potency of a chemical. Turquoise represents negative testing compound. Where a PC50 was not achieved, the PCmax is indicated. Numbers in bracket indicate the maximum percentage of induction or inhibition observed at the maximum dose tested. Negative = no dose-dependent and significant effect relative to the SCs was observed.*

	T					
	Lab1a	**Lab2**	**Lab3**	**Lab4**	**Lab5**	**Lab6**
Prochloraz	0.11	0.040	0.006	0.033	---	0.036
Aminoglutethimide	100 (24%)	43	25	46	---	54
Letrozole	43	47	31	59	54	26
Nonoxynol-9				---	---	
EDS						
HCG				---	---	
Benomyl				---	---	
Paraben	0.047				100 (7.7%)	0.52
Atrazine	100 (4%)	100 (33%)	100 (3.6%)	100 (32%)	---	33
Molinate	100 (31%)	100 (25%)	100 (26%)	77		0.33
Forskolin	17	0.41	0.56	0.22	---	0.031
Trilostane	0.014	0.0068	1.6	---	---	4.3
	E2					
	Lab1a	**Lab2**	**Lab3**	**Lab4**	**Lab5**	**Lab6**
Letrozole	0.00015	0.0020	0.00015	0.0077	---	0.0044
Prochloraz	0.052	0.040	0.202	0.20	---	0.0046
Aminoglutethimide	29	11	7	66	---	28
EDS					---	
HCG				---	---	
Nonoxynol-9				---	---	
Benomyl				---	---	
Molinate	100 (10%)				---	
Paraben	10 (12%)				---	10 (13%)
Atrazine	100 (36%)	100 (10%)	100 (18%)	100 (28%)	---	100 (11%)
Trilostane	1.028	0.3417	100 (42%)	---	---	85.0
Forskolin	0.3	1.62	2.22	4.46	---	2.244

[a] Maximum dose tested was 10 μM. Next greater dose exhibited significant cytotoxicity.

Color Codes:

Inducers		**Inhibitors**		**Negative**
	≤ 1 μM		≤ 0.01 μM	
	≤10 μM		≤ 1 μM	
	≤100 μM		≤ 10 μM	**Assay Interference**
	Max dose (>20%)		≤ 100 μM	
	Max dose (<20%)		Max dose (> 20%)	

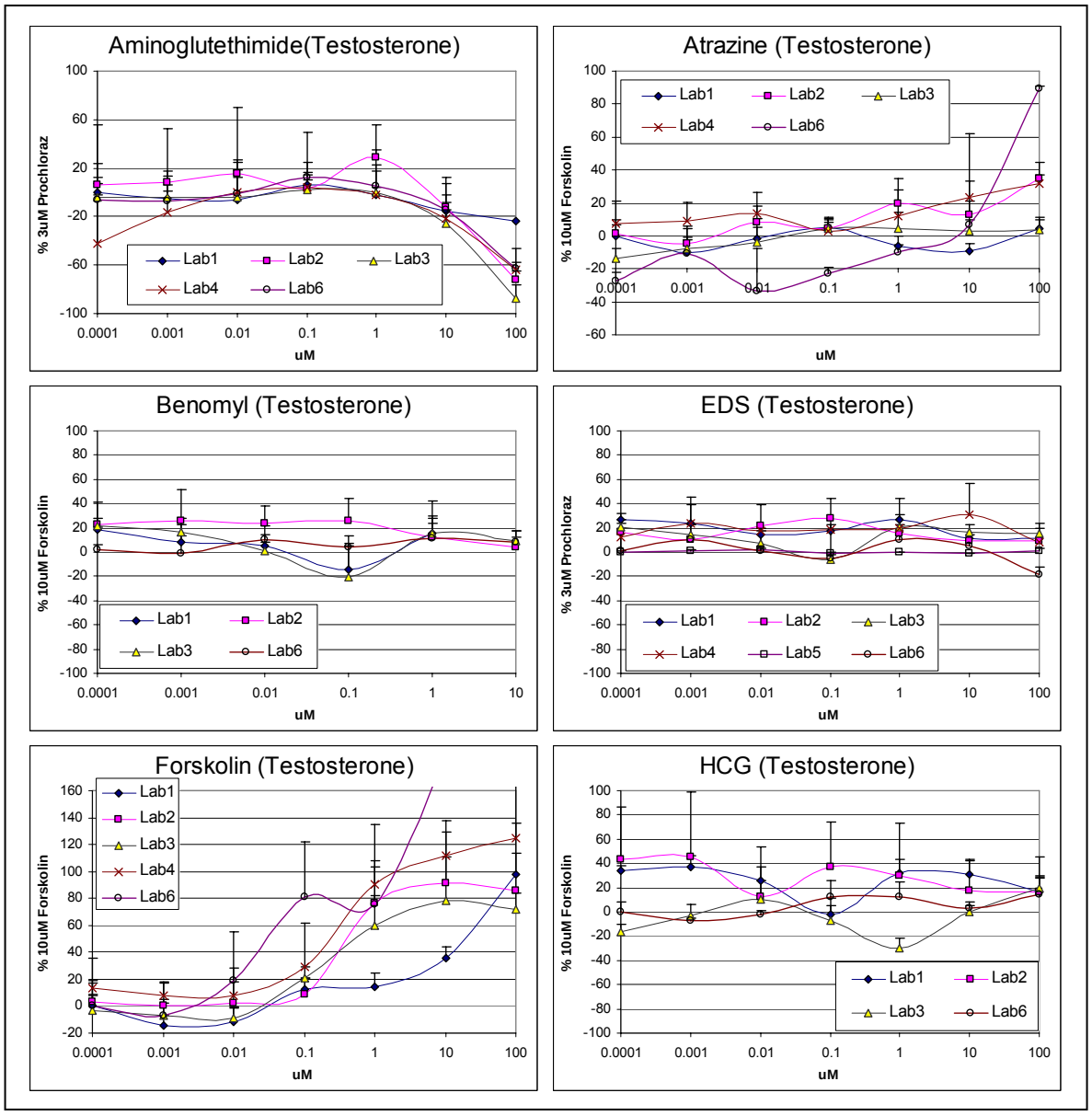


Figure 10.1: *Comparison of changes in the production of testosterone (T) relative to the induction caused by 10 µM forskolin (100%; inducers) and to the inhibition caused by 3 µM prochloraz (-100%; inhibitors) after exposure to aminoglutethimide, atrazine, benomyl, EDS, forskolin and HCG . Data points represent means of three replicate experiments (exceptions: Lab1 – only one replicate experiment was conducted; Lab3 – only two replicate experiments were conducted). Error bars = 1x standard deviation.*

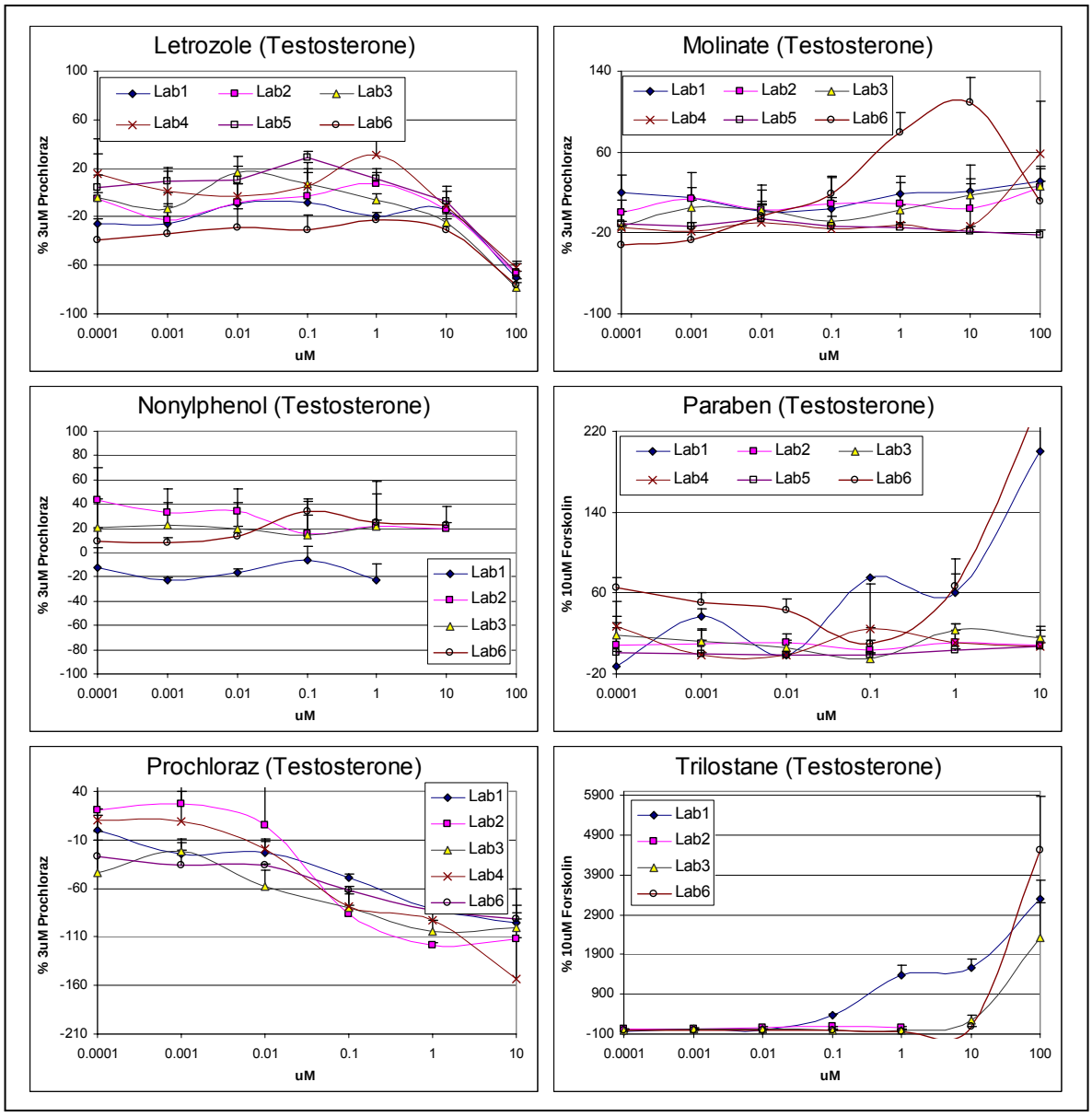


Figure 10.2: *Comparison of changes in the production of testosterone (T) relative to the induction caused by 10 µM forskolin (100%; inducers) and to the inhibition caused by 3 µM prochloraz (-100%; inhibitors) after exposure to letrozole, molinate, nonoxynol-9, paraben, prochloraz and trilostane . Data points represent means of three replicate experiments (exceptions: Lab1 – only one replicate experiment was conducted; Lab3 – only two replicate experiments were conducted). Error bars = 1x standard deviation.*

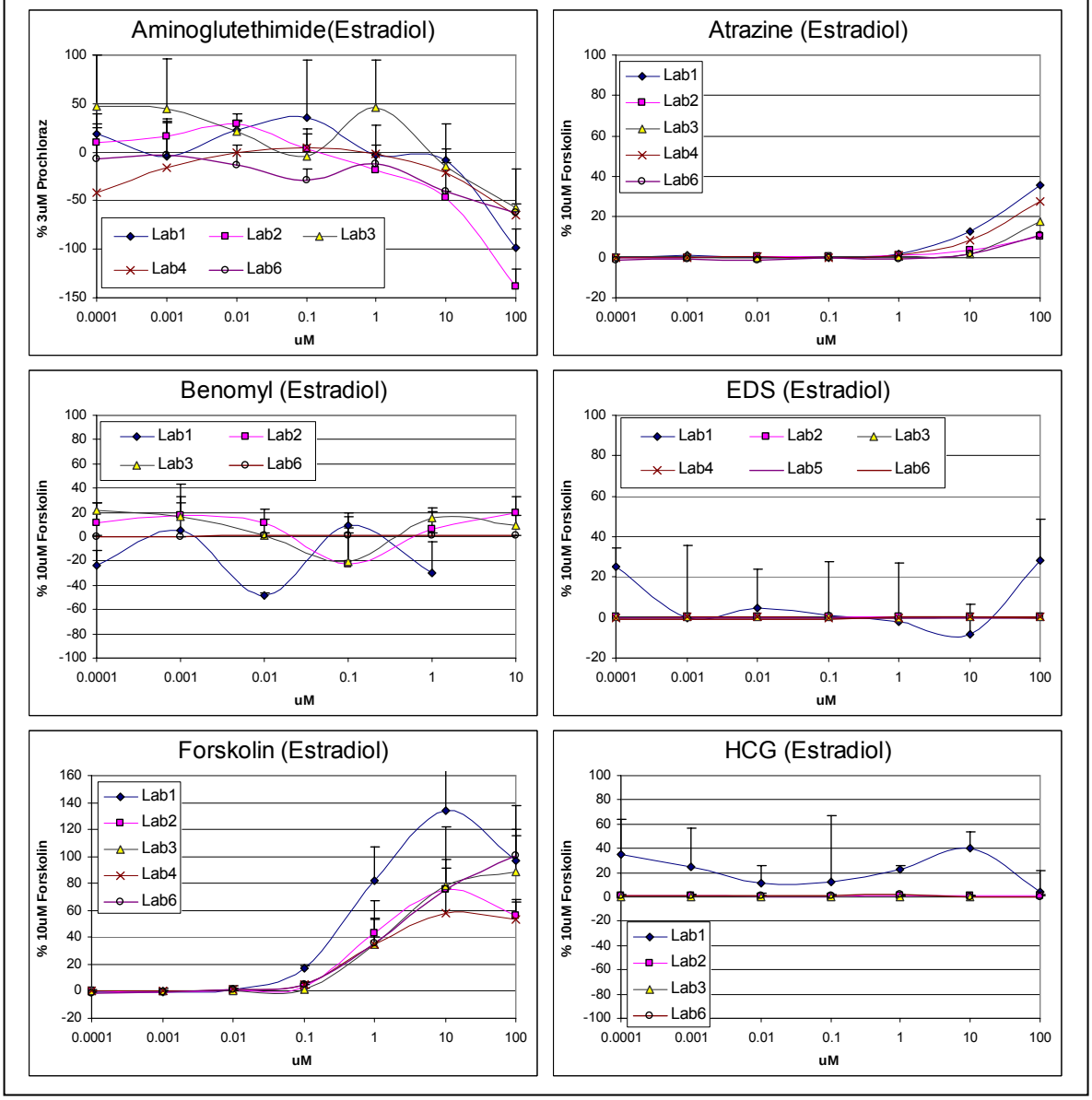


Figure 10.3: *Comparison of changes in the production of estradiol (E2) relative to the induction caused by 10 µM forskolin (100%; inducers) and to the inhibition caused by 0.3 µM prochloraz (-100%; inhibitors) after exposure to aminoglutethimide, atrazine, benomyl, EDS, forskolin and HCG . Data points represent means of three replicate experiments (exceptions: Lab1 – only one replicate experiment was conducted; Lab3 – only two replicate experiments were conducted). Error bars = 1x standard deviation.*

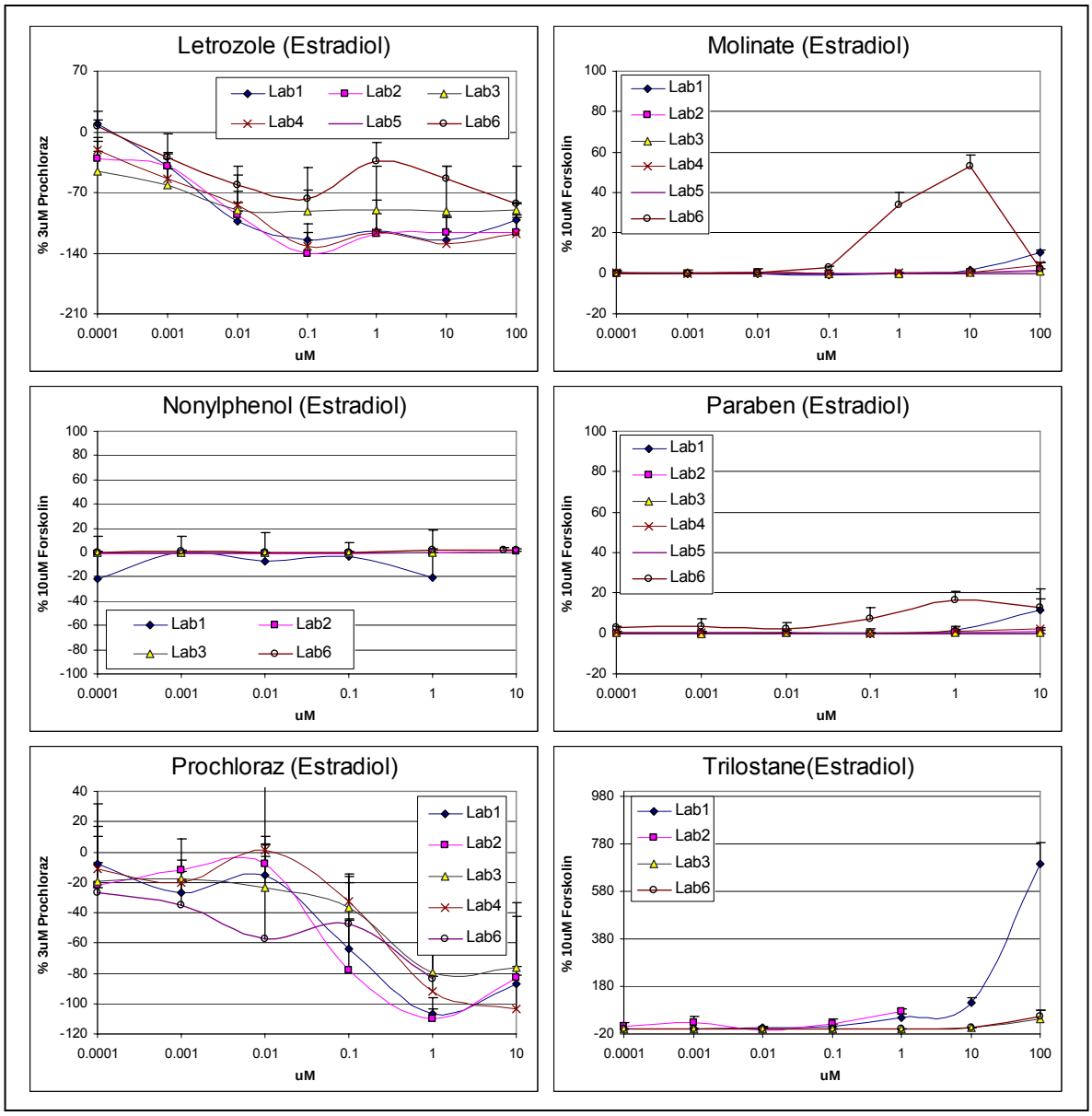


Figure 10.4: *Comparison of changes in the production of testosterone (E2) relative to the induction caused by 10 µM forskolin (100%; inducers) and to the inhibition caused by 0.3 µM prochloraz (-100%; inhibitors) after exposure to letrozole, molinate, nonoxynol-9, paraben, prochloraz and trilostane . Data points represent means of three replicate experiments (exceptions: Lab1 – only one replicate experiment was conducted; Lab3 – only two replicate experiments were conducted). Error bars = 1x standard deviation.*

11 APPLICATION OF THE H295R STEROIDOGENESIS ASSAY TO A SELECTION OF SUPLEMENTARY CHEMICALS

The above described and discussed approaches and protocols have been applied to evaluate the potential of 14 of the in chapter 6.2 listed supplementary chemicals. The data presented for these 14 chemicals was derived from a series of experiments (no replicate experiments per compound were conducted) conducted by the lead laboratory. As for the experiments with the 12 core chemicals, data obtained from wells with less or equal to 80% cell viability was excluded from further analysis. While over 60% of all chemicals showed reductions in cell viability of greater or equal to 20%, no single compound revealed cytotoxicity at doses less than 100µM (Figure 11.1). Significant interactions of the chemicals with the hormone detection assays that occurred at non-cytotoxic doses were only observed for T after exposure to spironolactone, finasteride and danazol. However, when the data without correction for this interference was compared to the same results after subtraction of the concentration measured for interference at the respective dose, no significant impact on the overall trend/response could be observed (Figure 11.2).

As for the above described 12 core chemicals, the H295R Steroidogenesis Assay allowed distinguishing between inducers and inhibitors of different strength/potency for both T and E2 (Figures 11.3 – 11.6). Five of the 14 compounds (42%) tested negative in the assay. Fenarimol, flutamide, finasteride, danazol and dimethoate did not elicit significant dose dependent responses for T, and piperonyl butoxide, fenarimol, finasteride, dinitrophenol and spironolactone did not reveal any dose-dependent effects on E2 production. One exception was dinitrophenol, which was identified as a significant inhibitor of T at all doses tested. However, no dose-response trend was recognizable and the magnitude of the effect was weak. Therefore, it is possible that this response represents an artifact. As stated in the results section it is desirable to have at least two (preferably three) repeat experiments for each chemical due to uncertainties resulting from potential inter assay variation. As a consequence, the data presented here should be considered with care until results from the repeat experiments are available. Regardless of these remaining uncertainties, it could be demonstrated that the H295R Steroidogenesis Assay protocol successfully identified chemicals with unknown modes of interaction with sex steroid synthesis as inducers and inhibitors of T and E2 production. Some of the chemicals

identified as inhibitors showed a biphasic response with typically slight increases in hormone production up to concentrations of 0.1 to 1 μM with the exception of dinitrophenol, which revealed a slight decrease in E2 production at lesser concentrations. These changes were statistically significant for T and E2 after treatment with fenarimol, tricrecyl phosphate and dimethoate, and fenarimol, dinitrophenol and dimethoate, respectively. With the exception of E2 in the fenarimol and dinitrophenol exposure groups none of these changes exceeded 1.5-fold, and did not affect the final categorization of a chemical. It is hypothesized that these minor changes are likely to be a compensatory mechanism, reflecting the integrative nature of the H295R Steroidogenesis Assay rendering a more realistic assay with regard to the identification of potential in vivo inducers/inhibitors of T and E2 production.

The specificity of the assay could be demonstrated by the relatively great number of chemicals that tested negative for the interference with the production of either T or E2 or both hormones.

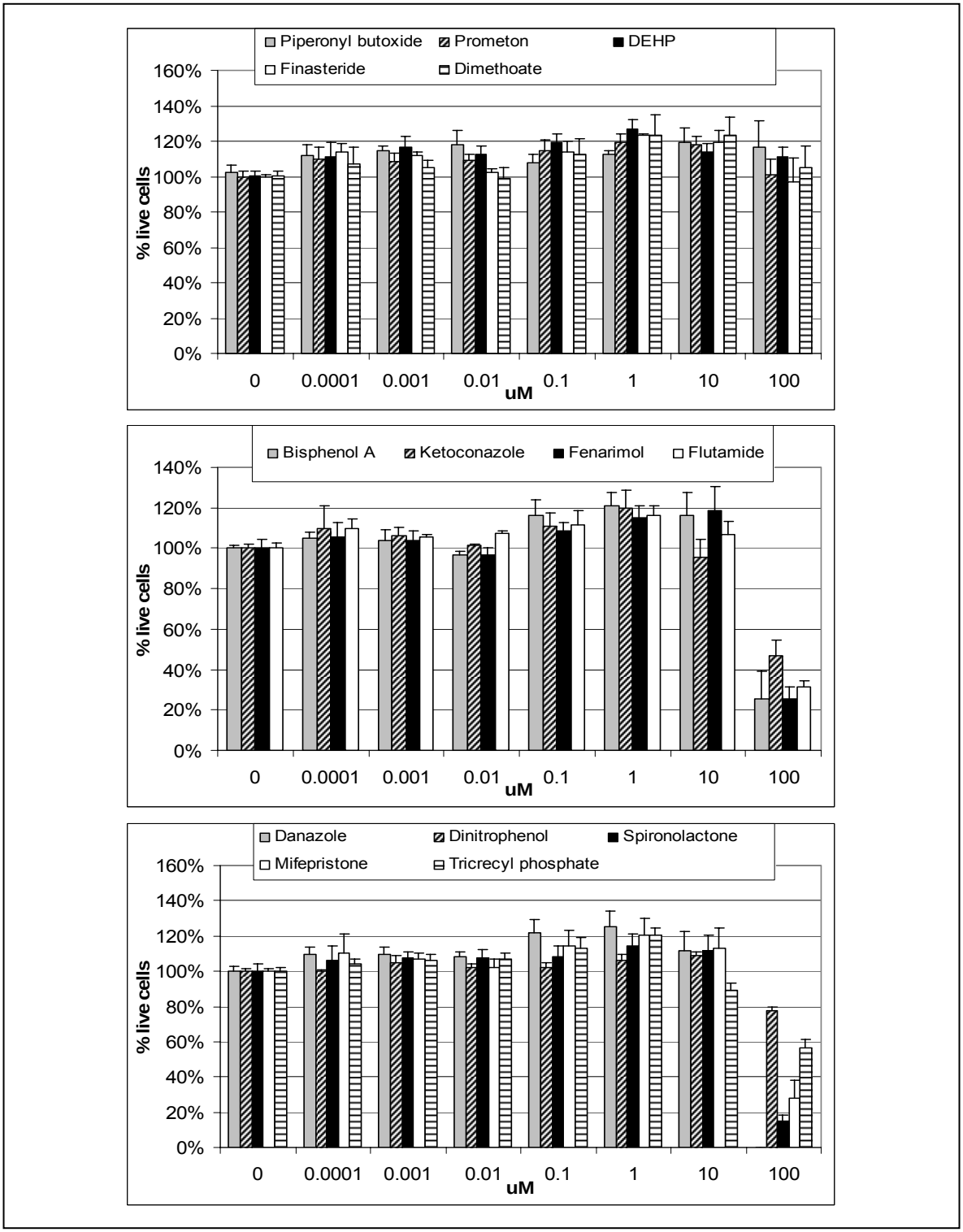


Figure 11.1: *Cell viability after exposure to piperonly butoxide, prometon, DEHP, flutamide and dimethoate. Cell viability is expressed relative to the solvent controls (SC = 100%) in each plate. Error bars = 1 x SD.*

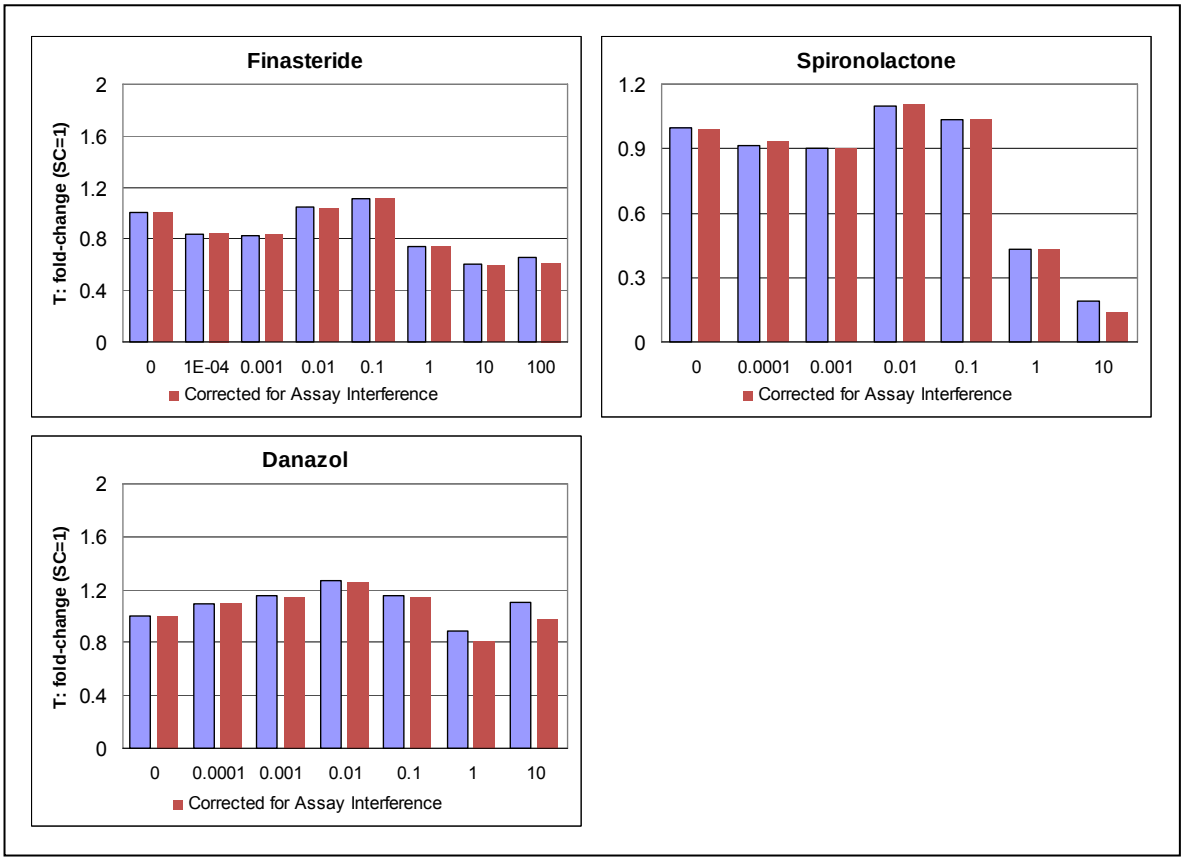


Figure 11.2: *. Changes in the concentrations of testosterone (T) relative to the solvent controls (SC=1) after exposure to spironolactone, finasteride and danazol after (red bars) and before (blue bars) subtraction of concentration at which the chemicals interfered with the hormone detection assay. Error bars = 1x standard deviation.*

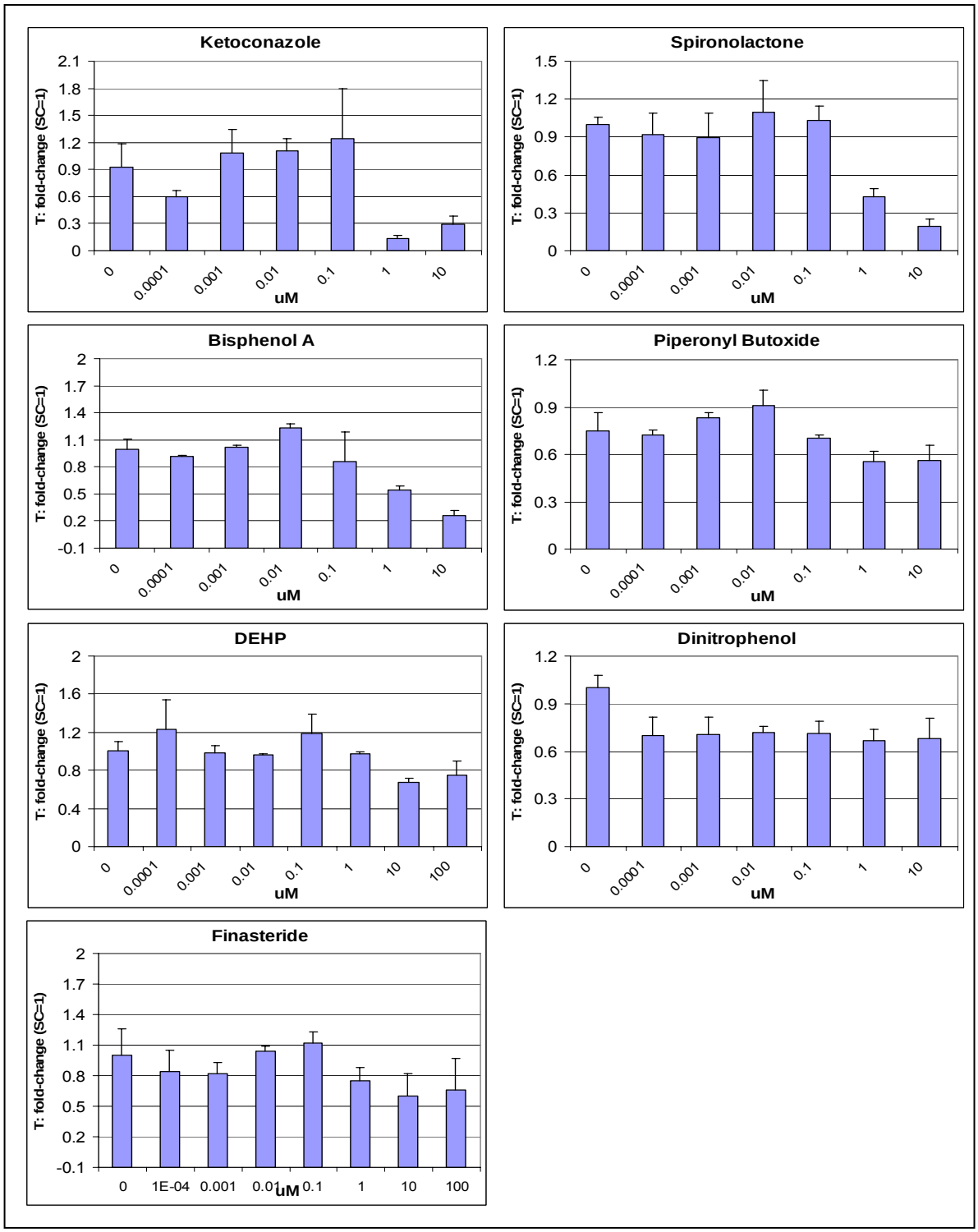


Figure 11.3: *Changes in the concentrations of testosterone (T) relative to the solvent controls (SC=1) after exposure to ketoconazole, spironolactone, bisphenol A, piperonyl butoxide, DEHP, dinitrophenol and finasteride. Error bars = 1x standard deviation.*

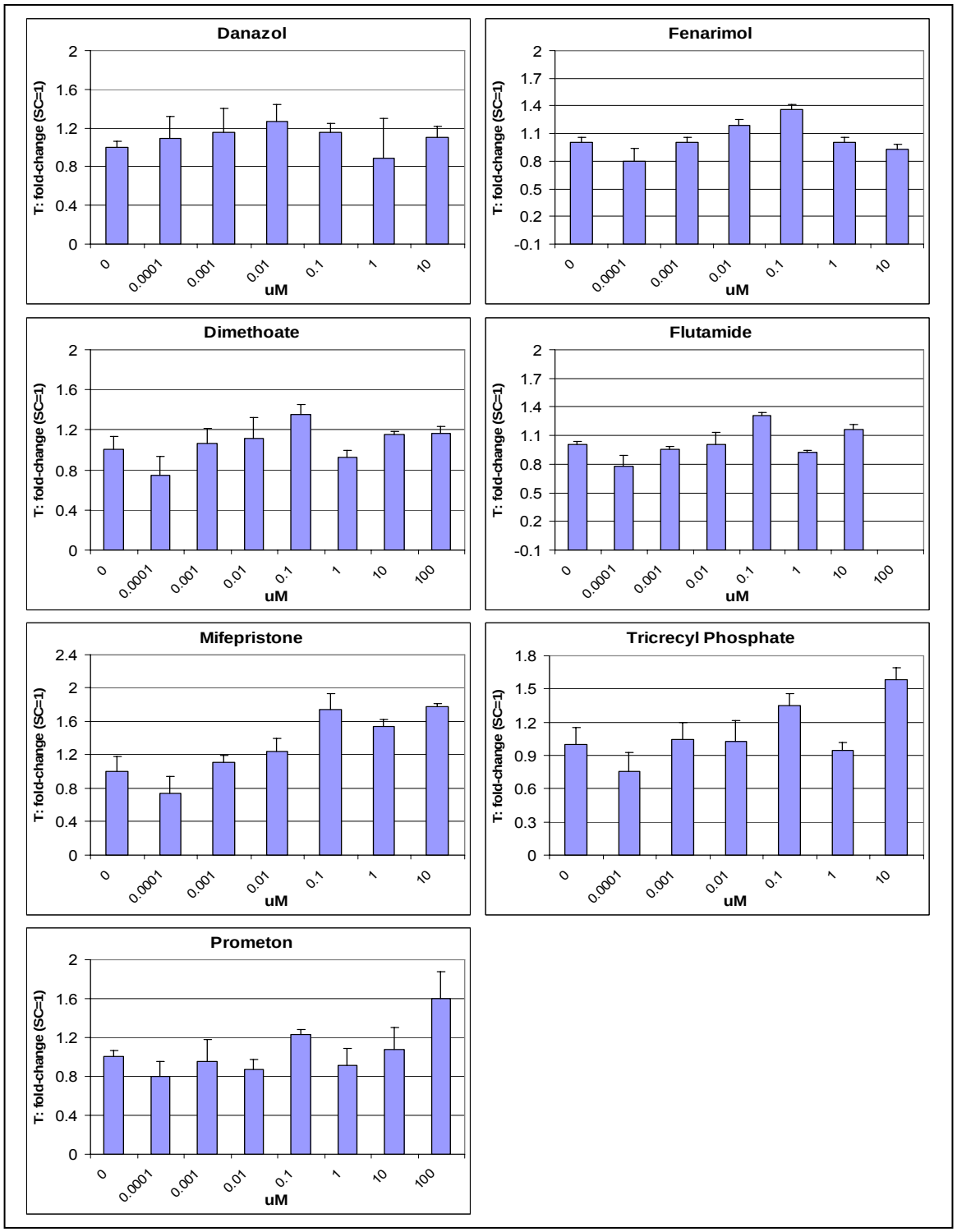


Figure 11.4: *Changes in the concentrations of testosterone (T) relative to the solvent controls (SC=1) after exposure to danazol, fenarimole, dimethoate, flutamide, mifepristone, tricrescyl phosphate and prometon. Error bars = 1x standard deviation.*

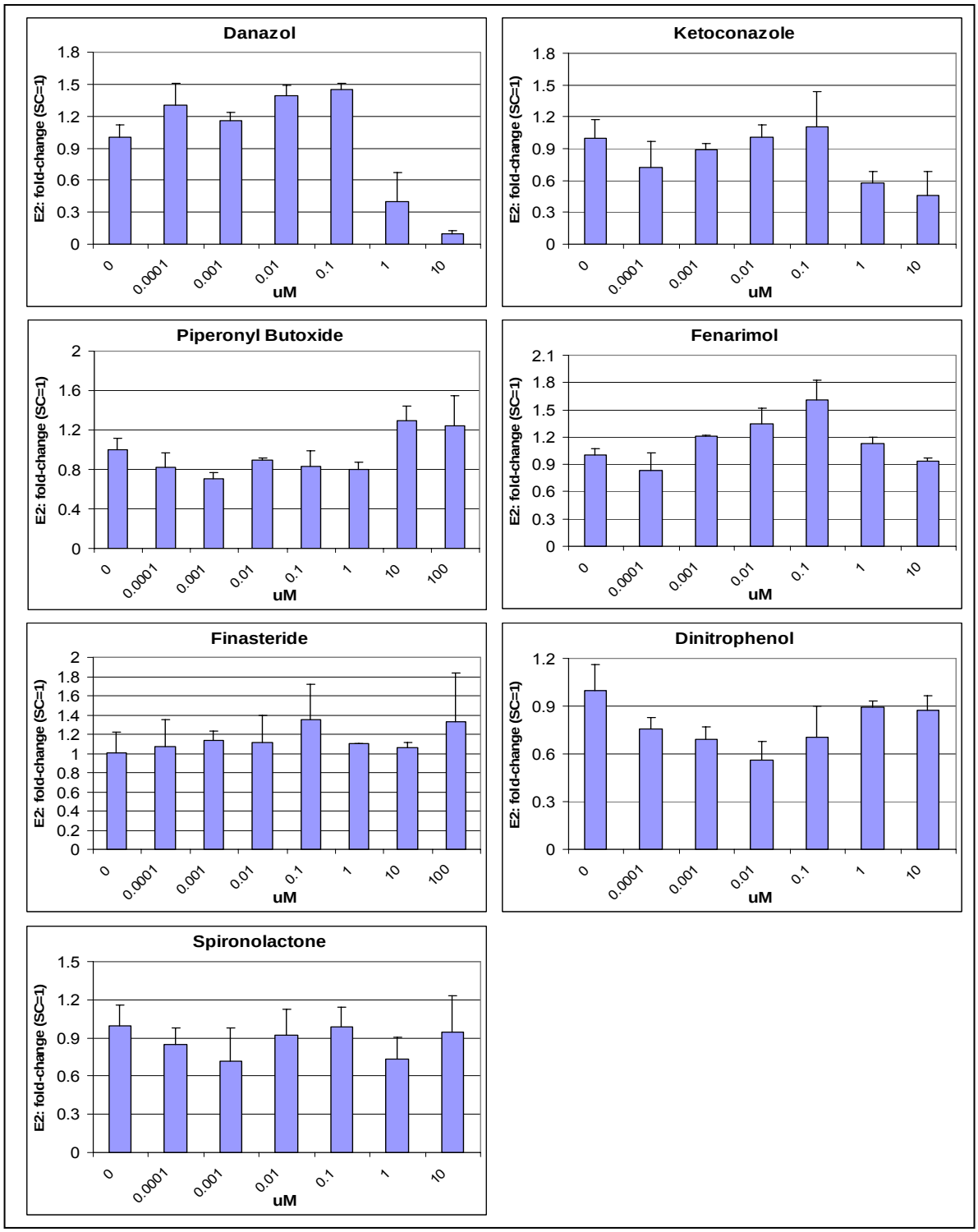


Figure 11.5: *Changes in the concentrations of estradiol (E2) relative to the solvent controls (SC=1) after exposure to danazol, ketoconazole, piperonyl butoxide, fenarimole, finasteride, dinitrophenol and spironolactone. Error bars = 1x standard deviation.*

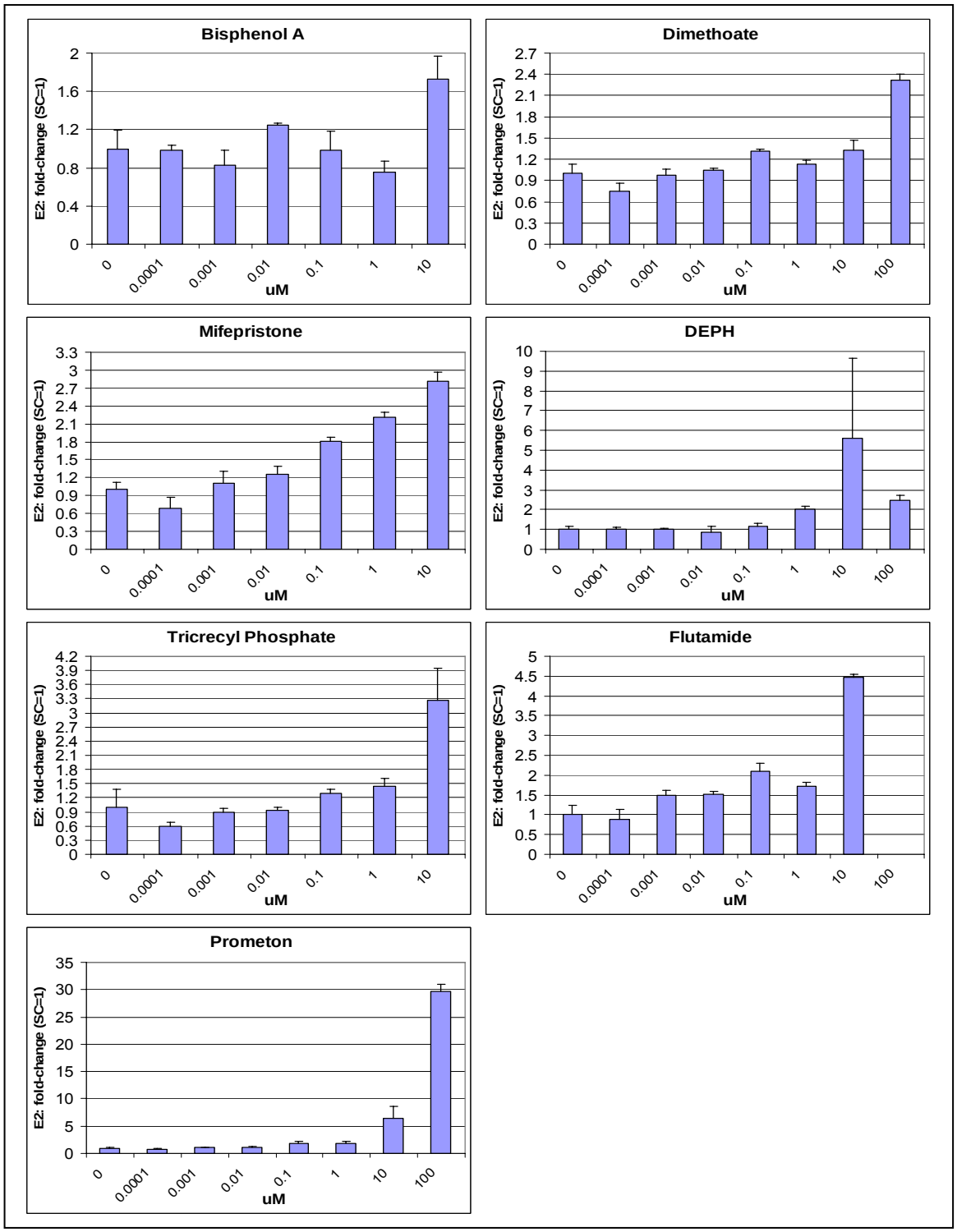


Figure 11.6: *Changes in the concentrations of estradiol (E2) relative to the solvent controls (SC=1) after exposure to bisphenol A, dimethoate, mifepristone, DEPH, tricrecy phosphate, flutamide and prometon. Error bars = 1x standard deviation.*

Table 11.1: *Least observed effect concentrations (LOECs [μM]; measured by Dunnett's test), strength and direction of change (arrows; see Table 10.1 for explanation), and chemical doses (μM) required to elicit 50% of the response caused by forskolin (inducers) and prochloraz (inhibitors) (PC50). Increasing intensities of yellow/red indicate increasing inhibiting strength/potency of a chemical; Increasing intensities of blue indicate increasing inducing strength/potency of a chemical. Light blue represents negative testing compound. .For chemicals where a PC50 was not achieved, a PCmax is shown with numbers in parentheses indicates the maximum percentage of induction or inhibition observed at the maximum dose tested (exception: E2 response after exposure to DEHP [a]).*

	Testosterone		
	LOEC	Max Change	PC50 or PCmax
Ketoconazole	1	⇩⇩⇩	0.61
Spironolactone	1	⇩⇩⇩	1.3
Bisphenol A	1	⇩⇩	4.0
Dinitrophenol	0.0001	⇩	1 (30%)
Piperonyl butoxide	1	⇩	10 (43%)
DEHP	10[a]	⇩	10 (21%)
Finasteride			
Dimethoate			
Fenarimol			
Danazol			
Flutamide			
Prometon	100	⇧	11
Tricrecyl phosphate	10	⇧	2.3
Mifepristone	0.1	⇧	0.0062
	Estradiol		
	LOEC	Max Change	PC50 or PCmax
Danazol	0.01	⇩⇩⇩	0.85
Ketoconazole	10	⇩⇩⇩	2.6
Piperonyl butoxide			
Fenarimol			
Finasteride			
Dinitrophenol			
Spironolactone			
Bisphenol A	10	⇧	100 (1.1%)
Dimethoate	10	⇧⇧	100 (5%)
Tricrecyl phosphate	10	⇧⇧	100 (7%)
Flutamide	10	⇧⇧⇧	100 (12%)
DEHP	10[a]	⇧⇧⇧	10 (6%)
Mifepristone	0.1	⇧⇧	100 (7%)
Prometon	10	⇧⇧⇧⇧	11

[a] considered because clear dose-response at all but the greatest dose

Inducers		**Inhibitors**	
	≤ 1 μM/⇧⇧⇧⇧		≤ 0.01 μM/⇩⇩⇩⇩
	≤10 μM/⇧⇧⇧		≤ 1 μM/⇩⇩⇩
	≤100 μM/⇧⇧		≤ 10 μM/⇩⇩
	Max dose (>20%)/⇧		≤ 100 μM/⇩
	Max dose (<20%)		Max dose (> 20%)

12 CONCLUSIONS AND SUMMARY

Considering the data obtained during the validation studies to date, the H295R Steroidogenesis protocol successfully detected all inducers and inhibitors of T (6 out of 7 laboratories) and E2 (5 out of 7 laboratories) production as well as negative chemicals. For each hormone, there were only two cases in which a laboratory did not confirm the trend that was observed at all other groups. In all situations such inconsistencies only occurred for weak inducers/inhibitors, and would have only in two cases resulted in a false negative response. However, even in these two cases a chemical would have been identified as a positive testing compound by the H295R Steroidogenesis Assay because there was a significant effect for the other hormone measured. In most of the cases (three out of four), these differences occurred at groups for which only data from one experiment for the chemical was available, indicating the need for replicate testing.

In addition, the assay succeeded identifying inducers and inhibitors of the production of either T or E2, or both hormones as well as negative chemicals from a list of 14 additional compounds with largely unknown modes of interaction with sex steroid synthesis. Furthermore, the assay reliably characterized both inducers and inhibitors based on the strength and potency of their effect. However, given the nature of the dosing regime that bracketed a wide range of concentrations in an attempt to screen chemicals that differed in potency and modes of action, there are some limitations that preclude the precise description of weak inducers and inhibitors. It is assumed that the H295R Steroidogenesis assay will be used as part of a Tier1 screening battery utilizing a number of vitro and in vivo assays, and thus, even in the rare occasions where there might be uncertainties as to how a chemical will be characterized basing on the data obtained with the H295R assay, these chemicals will be further evaluated in other tests. However, an alternative approach to better categorize weak inducers or inhibitor or steroidogenesis would be to conduct a second experiment with the chemicals of concern that use a more refined set of dose that would provide a in more detail the dose-response relationships.

A second issue that was observed during the validation studies was that some chemicals had the potential to interact directly with the hormone detection system utilized to measure hormone concentrations in the medium. However, for this portion of the validation study cross-reactivity of the assay system would have mischaracterized only one chemical as an inducer of E2 production. Currently, efforts are underway to develop a normalization approach that allows for the correction of measured hormone concentrations for possible cross-reactivities. Until such methods are available and have been validated,

the samples will have to be re-analyzed using alternative hormone detection methods that are not confounded by the chemical of interest. As stated above, the incidence of such interference by a chemical occurred only in rare occasions and is not considered to be of concern relative to the overall performance of the assay. This emphasizes, however, the importance of evaluating the cross-reactivity of a chemical prior to actually conducting the assay that will objectively evaluate its hormone altering potential.

Additional factors that have to be considered when evaluating chemicals using the H295R Steroidogenesis Assay are solubility and effects on cell viability. Chemicals that are not soluble at the highest concentration to be tested (typically 100µM) will need to be diluted to their solubility limit in whatever solvent (DMSO, ethanol, etc.) that is being used in the assay. Since non-soluble concentrations are not considered to be relevant to the specific mechanism of chemical interaction tested for by the H295R Steroidogenesis Assay, interaction with E2 and T production, they should not be used in evaluating the potency and efficacy of the chemical in the assay. Furthermore, although we acknowledge that chemicals that are not completely in solution may also have the potential to interact with cells or organisms, these effects are likely to be primarily associated with alterations of cell membranes altering their function and confounding the results of the assay, and not due to the specific endocrine effect to be tested here. Similarly, effects of chemicals on cell viability are to be considered as "general toxic" interactions that are not considered to be of relevance to the specific aim of this assay: identifying effects on the production of T and E2. The results of the validation study demonstrated that the H295R Steroidogenesis Assay protocol successfully identified both confounders, interference with the hormone detection assay and cytotoxicity, and thus, allowed to distinguish between "true inducers/inhibitors" and effects that were due to other factors such as cytotoxicity and hormone assay interference. A possible decision tree for the conduct of the H295R Steroidogenesis Assay considering the afore-discussed aspects has been proposed.

In this report, the assessment of different data evaluation approaches demonstrated the limitation of simply using data normalized to a solvent control in determining effective concentrations (EC20, EC50, etc) from potency curves. A superior approach for estrogen production was found that normalized the hormone data from each treatment dose relative to the response observed for the model compounds forskolin and prochloraz run in parallel on the QC-plate, the PC50 approach. This approach appears to be a more objective reflection of the true chemical response on steroidogenesis and gives a two dimensional response based on both the strength of the response and its potency (concentration at which the response is observed) and allows for a more appropriate categorization of chemicals. However, this approach is not as easily applied to inducers of T because, the model inducer

forskolin, which has been used in this and other studies, only caused a moderate induction (2- to 3-fold) of T. While forskolin still permitted the distinction between inducers of T of different strength this rendered the utility of forskolin in categorizing chemicals based on changes in T production, sub-optimal.

13 REFERENCES

Battelle (2005) Sliced Testis Assay Prevalidation:Baseline/Positive Control Study.

ECVAM (1995) Practical Aspects of the Validation of Toxicity Test Procedures: The Report and Recommendations of ECVAM Workshop No. 5. ATLA 23:129-147.

Endocrine Disruptor Methods Validation Advisory Committee (2005) Letter from the EDMVAC to Stephen L. Johnson, Administrator of EPA, May 5, 2005.

Gazdar AF, Oie HK, Shackleton CH, Chen, TR, Triche, TJ, Myers CE, Chrousos GP, Brennan MF, Stein CA., La Rocca RV (1990) Cancer Res 50:5488-5496.

Giesy JP, Hilscherova K, Newsted JL, Mehrle PM (2002) Optimization of the H295R cell line for use in evaluating toxicant induced effects of steroidogenesis. Progress Report to the US-EPA ORD Service Centre/NHEERL. July 2002.

Hartung T et al. (2004) ATLA 32, 467-472.

Hecker M, Newsted JL, Murphy MB, Higley EB, Jones PD, Wu R, Giesy JP (2006a) Toxicol Appl Pharmacol 217:114-124.

Hecker M, Giesy, JP, Jones PD, Higley EB, Newsted JL, Mehrle P (2006b) Influence of cell passage and freeze/thaw events on basal production of 17β-estradiol and testosterone by H295R cells. Interim Draft Report to the US-EPA. Contract code: GS-10F-0041L. September 2006.

Hecker M, Hollert H, Cooper R, Vinggaard A-M, Akahori Y, Murphy M, Nellemann C, Higley E, Newsted J, Wu R, Lam P, Laskey J, Buckalew A, Grund S, Nakai M, Timm G, Giesy J (2007a) Env Sci Pollut Res 14(Special Issue 1):23−30.

Hecker M, Giesy, JP, Jones PD, Higley EB, Newsted JL, Mehrle P (2007b) Standardization and refinement of the H295R cell based assay to identify chemical modulators of steroidogenesis and aromatase activity - Repeat of Inter-Laboratory Study. Report to the US-EPA. Contract code: GS-10F-0041L. March 2007.

Hilscherova K, Jones PD, Gracia T, Newsted JL, Zhang X, Sanderson JT, Yu RMK, Wu RSS, Giesy JP (2004) Toxicol Sci 81:78-89.

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

Kavlock RJ, Daston GP, De Rosa C, Fenner-Crisp P, Gray LE, Kaattari S, Lucier G, Luster M, Mac MJ, Maczka C, Miller R, Moore J, Rolland R, Scott G, Sheehan DM, Sinks T,, Tilson HA. (1996) Environ Health Perspect 104:715-740.

OECD (1996) Final Report of the OECD Workshop on the Harmonization of Validation and Acceptance Criteria for Alternative Toxicological Test Methods (Solna Report) as presented to the Seventh Meeting of the National Coordinators of the Test Guidelines Program, 18-19 September 1996. ENV/MC/ CHEM/TG(96)9.

OECD (1998) Report of the First Meeting of the OECD Endocrine Disrupter Testing and Assessment (EDTA) Task Force, 10-11 March 1998. ENV/MC/CHEM/RA(98)5

OECD (2002). Report of the Sixth Meeting of the Endocrine Disrupters Testing and Assessment Task Force, 24-25 June 2002. ENV/JM/TG/EDTA/M(2002).

OECD (2005). Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. Guidance Document No. 34. June 2005

Ohno S, Shinoda S, Toyoshima S, Nakazawa H, Makino T, Nakajin S (2002) J Steroid Biochem Molec Biol 80:355-363.

Puddefoot, J.R., Barker, S., Glover, H.R., Malouitre, S.D.M., Vinson, G.P. (2002) Non-competitive steroid inhibition of oestrogen receptor functions. Int J Cancer 101:17-22.

Rainey, W. E., Bird, I. M., Sawetawan, C., Hanley, N. A., Mccarthy, J. L., Mcgee, E. A., Wester, R., and Mason, J. I. (1993) J. Clin. Endocrinol. Metab. 77, 731-737.

Sanderson, J. T., Boerma, J., Lansbergen, G., and Van den Berg, M. (2002) Toxicol. Appl. Pharmacol. 182, 44-54.

Takeyoshi M. 2007. Prevalidation and Inter-laboratory Validation of Stably Transfected Transcriptional Activation Assay to Detect Estrogenic Activity: The Human Estrogen Receptor Alpha Mediated Reporter Gene Assay Using hER-HeLa-9903 Cell Line. Chemical Evaluation and Research Institute, Japan. October, 2007.

Villeneuve DL, Blankenship AL, Giesy JP (1998) Estrogen receptors - environmental xenobiotics. In: Denison MS, Helferich WG (eds) Toxicant-receptor interactions and modulation of gene expression. Lippincott-Raven Publishers, Philadelphia.

Wilson, V. S., Bobseine, K., Lambright, C. R., and Gray, L. E. (2002) Toxicol. Sci. 66, 69-81.

Wilson, V. S., Bobseine, K., and Gray, L. E. (2004) Toxicol. Sci. 81, 69-77.

14 APPENDICES

APPENDIX I – H295R Cell Culture Protocol

APPENDIX II – H295R Cell Exposure Protocol

APPENDIX III – Data Entry Sheet Template

APPENDIX IV – Raw data sheets Laboratory 1

APPENDIX IV – Raw data sheets Laboratory 2

APPENDIX IV – Raw data sheets Laboratory 3

APPENDIX IV – Raw data sheets Laboratory 4

APPENDIX IV – Raw data sheets Laboratory 5

APPENDIX IV – Raw data sheets Laboratory 6