

# The OECD validation program of the H295R steroidogenesis assay: Phase 3. Final inter-laboratory validation study

Markus Hecker · Henner Hollert · Ralph Cooper · Anne Marie Vinggaard · Yumi Akahori · Margaret Murphy · Christine Nellemann · Eric Higley · John Newsted · John Laskey · Angela Buckalew · Stefanie Grund · Sibylle Maletz · John Giesy · Gary Timm

Received: 6 May 2010 / Accepted: 8 September 2010 / Published online: 3 October 2010  
© Springer-Verlag 2010

## Abstract

**Background, goals, and scope** In response to increasing concerns regarding the potential of chemicals to interact with the endocrine system of humans and wildlife, various national and international programs have been initiated with the aim to develop new guidelines for the screening and testing of these chemicals in vertebrates. Here, we report on the validation of an in vitro assay, the H295R steroidogen-

esis assay, to detect chemicals with the potential to inhibit or induce the production of the sex steroid hormones testosterone (T) and 17 $\beta$ -estradiol (E2) in preparation for the development of an Organization for Economic Cooperation and Development (OECD) test guideline.

**Methods** A previously optimized and pre-validated protocol was used to assess the potential of 28 chemicals of diverse structures and properties to validate the H295R steroidogen-

---

Responsible editor: Philippe Garrigues

---

M. Hecker  
ENTRIX, Inc.,  
Saskatoon SK S7N 5B3 Saskatchewan, Canada

M. Hecker (✉) · E. Higley · J. Giesy  
Department of Veterinary Biomedical Sciences and Toxicology  
Centre, University of Saskatchewan,  
Saskatoon, Saskatchewan, Canada  
e-mail: mhecker@entrix.com

H. Hollert · S. Maletz  
Department of Ecosystem Analysis, Institute for Environmental  
Research, RWTH Aachen University,  
52074 Aachen, Germany

R. Cooper · A. Buckalew  
Endocrinology Branch, RTD, NHEERL, ORD,  
US Environmental Protection Agency,  
Research Triangle Park, NC, USA

A. M. Vinggaard · C. Nellemann  
Department of Toxicology and Risk Assessment,  
National Food Institute, Technical University of Denmark,  
Soborg, Denmark

Y. Akahori  
Chemicals Evaluation and Research Institute,  
Chemicals Assessment Center,  
Saitama, Japan

M. Murphy · J. Giesy  
Department of Biology and Chemistry,  
City University of Hong Kong,  
Kowloon, Hong Kong, SAR, China

J. Newsted  
ENTRIX, Inc.,  
Okemos, MI 48864, USA

J. Laskey  
Senior Environmental Employment Program,  
National Caucus on Black Aged,  
Research Triangle Park, NC, USA

S. Grund  
Department of Zoology, University of Heidelberg,  
69120 Heidelberg, Germany

G. Timm  
Office of Science Coordination and Policy, US Environmental  
Protection Agency,  
Washington, DC, USA

esis assay. These chemicals are comprised of known endocrine-active chemicals and “negative” chemicals that were not expected to have effects on the targeted endpoints, as well as a number of test chemicals with unknown modes of action at the level of the steroidogenic pathway. A total of seven laboratories from seven countries participated in this effort. In addition to effects on hormone production, confounding factors, such as cell viability and possible direct interference of test substances with antibody-based hormone detection assays, were assessed. Prior to and during the conduct of exposure experiments, each laboratory had to demonstrate that they were able to conduct the assay within the margin of predefined performance criteria.

**Results** With a few exceptions, all laboratories met the key quality performance parameters, and only 2% and 7% of all experiments for T and E2, respectively, were excluded due to exceedance of these parameters. Of the 28 chemicals analyzed, 13 and 14 tested affected production of T and E2, respectively, while 11 and 8 did not result in significant effects on T and E2 production, respectively. Four and six chemicals produced ambiguous results for effects on T and E2 production, respectively. However, four of these cases each for T and E2 were associated with only one laboratory after a personnel change occurred. Significant interference of test chemicals with some of the antibody-based hormone detection systems occurred for four chemicals. Only one of these chemicals, however, significantly affected the ability of the detection system to categorize the chemical as affecting E2 or T production.

**Discussion and conclusions** With one exception, the H295R steroidogenesis assay protocol successfully identified the majority of chemicals with known and unknown modes of interaction as inducers or inhibitors of T and E2 production. Thus it can be considered a reliable screen for chemicals that can alter the production of sex steroid hormones. One of the remaining limitations associated with the H295R steroidogenesis assay protocol is the relatively small basal production of E2 and its effect on quantifying the decreased production of this hormone with regard to the identification of weak inhibitors. An initial comparison of the data produced in this study with those from *in vivo* studies from the literature demonstrated the potential of the H295R steroidogenesis assay to identify chemicals affecting hormone homeostasis in whole organisms. Particularly promising was the lack of any false negatives during the validation and the very low number of false positives (1 out of 28 chemicals for each T and E2).

**Perspectives** Based on the results obtained during this validation study and the accordingly revised test protocols, an OECD draft test guideline was developed and submitted to the OECD working group of the national coordinators of the test guidelines program (WNT) for comments in December 2009.

**Keywords** Endocrine disruption · Estradiol · H295R · Hormone production · Transferability · Sex steroid · Steroidogenesis · Testosterone · Validation · OECD

## 1 Introduction

In response to increasing concerns regarding the potential of chemicals to interact with the endocrine system of humans and wildlife, various national and international programs have been initiated with the aim of developing new guidelines for screening and testing of these chemicals in vertebrates (OECD 1998, 2002; EDSTAC 1998). One of the leading and first nationally, legally binding programs was the endocrine disruptor screening program (EDSP) of the US Environmental Protection Agency (EPA), which employs a battery of *in vitro* and *in vivo* screening assays to assess the endocrine disrupting potential of a chemical. Specifically, the US Congress included a provision in the Food Quality Protection Act of 1996 adding section 408 to the Federal Food, Drug, and Cosmetic Act (FFDCA). 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 USC 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 was charged with the duty to provide EPA with recommendations on how to implement its 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 as well as the originally mandated effects relating to estrogen. EPA accepted the EDSTAC’s recommendations for a two-tier screening program (EPA 1998). Recognizing the global relevance of the issue of endocrine disruption, in addition to the US EPA activities, the Organization for Economic Cooperation and Development (OECD) initiated a high-priority activity in 1998 to revise existing, and to develop new, test guidelines for the screening and testing of potential endocrine disrupting chemicals. The OECD conceptual framework for testing and assessment of potential endocrine disrupting chemicals comprises five levels, each level corresponding to a different level of biological complexity (OECD 2002).

The objective of the steroidogenic screening assay is to detect substances that would disrupt estradiol and testosterone production. The steroidogenic assay is intended to identify xenobiotics that have as their target site(s) components that comprise the biochemical pathway begin-

ning with the sequence of reactions occurring after the gonadotropin hormone receptors [follicle-stimulating hormone receptor (FSHR) and luteinizing hormone receptor (LHR)] up through the production of the terminal sex 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 or pituitary gland or on storage or release of sex steroid hormones. Based on the objectives described above, the most promising assay for use as a screen would be a relatively fast, inexpensive, technically simple assay that identifies substances that alter sex steroid hormone production due to direct effects on the enzymes or other endogenous components of the steroidogenic pathway.

One of the assays recommended by EDSTAC as a Tier 1 screen was an *in vitro* rodent minced testis assay screen to detect chemicals with the potential to disrupt steroid hormone production (EDSTAC 1998). Despite its long history of use, the rodent minced testis assay had not been optimized at the time that it was recommended by EDSTAC. EPA conducted a series of studies to optimize the assay and evaluate its suitability to serve a function in the EDSP testing battery. Preliminary inter-laboratory studies exhibited large variability within and among laboratories (Battelle 2005). However, the seemingly insurmountable problem of assessing cytotoxicity specific to Leydig cells led EPA's advisory committee to recommend that EPA abandon further work on the minced testis assay (EDMVAC 2005). As a consequence, there was a need for a less variable and more reliable *in vitro* test system(s) as an alternative to the minced testis assay. One assay that offered promise with regard to the characterization of inducers and/or inhibitors of sex steroid production was the H295R steroidogenesis assay (Hecker et al. 2006; Hecker and Giesy 2008).

Development and standardization of the H295R steroidogenesis assay as a screen for the evaluation of the effects of chemicals on the synthesis of T and E2 has been conducted in a multistep process. The results of the assay optimization process and the pre-validation efforts undertaken to date have been reported previously (Hecker et al. 2006, 2007). After initial development of the assay, US EPA presented a progress report on the development of the H295R assay to an OECD committee and invited member countries to join the USA in its further standardization and validation. This invitation was accepted by laboratories in Japan, Denmark, Germany, Hong Kong, and Korea.

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 US Interagency Coordinating Committee on Validation of Alternative Methods (ICCVAM 1997). Here, the results of an inter-laboratory study that was part of the final validation of a H295R protocol in accordance with the OECD guidelines are presented.

Using three model chemicals tested by five independent laboratories, an inter-laboratory pre-validation study was conducted to develop the H295R steroidogenesis assay protocol (Hecker et al. 2007). These studies indicated that the H295R test protocol was capable of characterizing the effect of chemicals on the production of T and E2. The goal of the present project was to further validate the H295R steroidogenesis assay by assessing the transferability, flexibility, and applicability of an improved and revised protocol ([http://www.epa.gov/endo/pubs/assayvalidation/h295r\\_pr.htm](http://www.epa.gov/endo/pubs/assayvalidation/h295r_pr.htm)) across several laboratories using an extended test set of 28 chemicals selected and approved by the OECD Validation and Management Group for Non-Animal Testing (VMG NA).

## 2 Materials and methods

### 2.1 Study protocol

Based on the results obtained during the initial pre-validation studies (Hecker et al. 2006), a standardized H295R steroidogenesis assay protocol was developed (<http://www.oecd.org/dataoecd/56/11/44285292.pdf>). In brief, cells were to be cultured under standard cell culture conditions as described in the H295R steroidogenesis assay protocol (<http://www.oecd.org/dataoecd/56/11/44285292.pdf>) for a minimum of four to five passages to ensure sufficient basal E2 production (cell age was not to exceed ten passages). The assay was then performed in 24-well culture plates (Hecker et al. 2007). Cells were seeded at a density of approximately 200,000 to 300,000 cells/ml, and after an acclimation period of 24 h cells were exposed for 48 h to seven concentrations between 0.0001 and 100  $\mu$ M of the test chemical in triplicate. In parallel, a plate in which cells were exposed to a known inhibitor (prochloraz) and inducer (forskolin) of hormone production was run as a quality control (QC) measure. At the end of the exposure period, the medium was removed from each well, and hormones were extracted using ethyl ether (note: one laboratory did not conduct extraction; in this case, the medium was directly used in the assay; Table 1). Cell viability in each well was analyzed immediately after

**Table 1** Type of hormone detection assay and extraction used by the participating laboratories

Lab	Hormone detection system	Extraction
1	ELISA, LC-MS <sup>a</sup>	Ether
2	ELISA	Ether
3	RIA	No
4	ELISA	Ether
6	ELISA	Ether

ELISA Enzyme linked immunoassay, LC-MS liquid chromatography mass spectroscopy, RIA radio immunoassay

<sup>a</sup> QC plate hormone results were confirmed using LC-MS analysis

removal of medium by means of the MTT assay (Mosman 1983) or the Live/Dead<sup>®</sup> variability assay (Invitrogen, Carlsbad, CA, USA). All concentrations, where cell viability was less than or equal to 80%, were excluded from the data analysis. Concentrations of hormones in medium were measured using commercially available hormone detection kits (Table 1). Responses measured by means of antibody-based assays in the QC plate experiments were confirmed by instrumental techniques [liquid chromatography mass spectroscopy (LC-MS)] at Lab 1 following the method described by Chang et al. (2010; data not shown). Each experiment was repeated three times with exception of Labs 1 and 3, where one and two replicate experiments were conducted per chemical, respectively.

Laboratories were required to demonstrate competence in performing all of the procedures that are part of the H295R steroidogenesis assay prior to testing chemicals (Table 2). The QC that was part of the actual conduct of the assay to allow for the evaluation of the assay performance during each experiment also served as a benchmark for determining laboratory competence prior to the initiation of chemical testing.

Prior to initiation of the actual exposure experiments, each chemical was tested for potential interference with the hormone detection system used. This was of particular relevance for antibody-based assays such as enzyme-linked immunoassays (ELISAs) and radio immunoassays (RIAs) because it has been previously shown that some chemicals can interfere with these tests (Shapiro and Page 1976; Puddefoot et al. 2002; Villeneuve, personal communication).

## 2.2 Participating laboratories

A total of seven laboratories from the USA, Denmark, Germany, Japan, Hong Kong, and Canada, each with different levels of experience in conducting the H295R steroidogenesis assay, were invited to participate in this validation study. Inclusion of laboratories with different levels of proficiency in conducting the assay was essential to evaluate the completeness of the test protocols and their transferability. Each laboratory was assigned a random code number (1–7) as part of the study. However, part way through the study, two of the seven laboratories decided to cease their participation in the validation studies. Thus, with the exception of the QC exposure data, only the data for the remaining five laboratories that completed the validation studies is presented (Labs 1, 2, 3, 4, and 6).

## 2.3 Selection and testing of chemicals

A total of 28 chemicals were selected in this study to validate the H295R steroidogenesis assay as a screen for potential effects of endocrine-disrupting chemicals on the production of T and E2 (Supplemental Materials). These chemicals were selected based on their known or suspected endocrine activity, or lack thereof, and included inhibitors and inducers of different potencies as well as positive and negative controls. Where possible, the test set of chemicals

**Table 2** Performance criteria to be met by each laboratory during experiments

System	Parameter	Comparison to/between	T	E2
Hormone detection system	Sensitivity	Detectable fold decrease relative to SC	≥2-fold	≥2-fold
	Precision	CV among replicate measures (absolute concentrations) of the same well for SCs	≤25%	≤25%
Cell assay	Basal hormone production in SCs	Fold greater than LOQ of hormone detection system	≥5-fold	≥2.5-fold
	Precision (SCs)	CV among absolute concentrations of replicate wells	≤30%	≤30%
	Sensitivity (induction @ 10 μM forskolin)	Fold greater than SC	≥2-fold	≥7.5-fold
	Sensitivity (inhibition @ 3 μM prochloraz)	Fold less than SC	≥0.5-fold	≥0.5-fold

Induction and inhibition refer to the relative change in hormone production after exposure to 10 μM forskolin or 3 μM prochloraz, respectively, in the QC plates

CV Coefficient of variation (%), LOQ limit of quantification, SC solvent control

was harmonized with those used in other steroidogenesis assays currently under development or in validation [e.g., the Registration, Evaluation, Authorisation and Restriction of Chemical substances (REACH) program].

Prior to the start of the validation studies, all chemicals were pre-analyzed by the lead laboratory (Lab 1). To reduce the workload for individual laboratories, each of the other groups tested a total of 17 to 18 chemicals. Each chemical set consisted of a “core group” of 12 chemicals that were tested in parallel by all laboratories. In addition, three laboratories plus the lead laboratory conducted assays on a different set of five or six chemicals selected from the 16 chemicals that did not comprise the core chemicals. That is, the 16 non-core chemicals were divided into three subgroups of five to six chemicals, and each chemical subgroup was tested by one laboratory [total number of laboratories=4, so that with two laboratories (lead and one test lab) testing five to six different chemicals, all 16 of the non-core chemicals were analyzed].

#### 2.4 Statistical methods

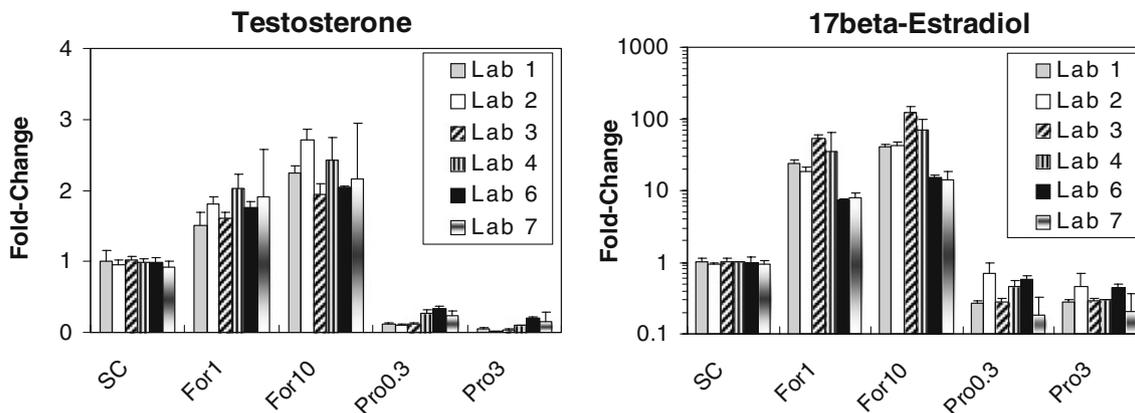
All data were expressed as mean±standard error of the mean (SEM). To examine the relative changes in hormone production, results were normalized to the mean solvent control (SC) value for each assay (i.e., each 24-well plate of cells used to test a given chemical), and results were expressed as percent change relative to the SC. Prior to conducting statistical analyses, the assumptions of data normality and variance of homogeneity were evaluated. Normality was evaluated using standard probability plots or the Shapiro–Wilk’s test. If the data were normally distributed or approximated a normal distribution, differences between chemical treatments and SCs were analyzed

using one-way analysis of variance (ANOVA) followed by a two-sided Dunnett’s test. If data were not normally distributed, the Kruskal–Wallis test followed by the Mann–Whitney *U* test were used. Data analysis was conducted using pooled replicate experiments. All statistical analyses were conducted using SYSTAT 11 (SYSTAT Software, Point Richmond, CA) Differences were considered significant at  $p < 0.05$ .

### 3 Results and discussion

#### 3.1 Laboratory performance assessment

With a few exceptions, all of the laboratories met the key quality performance parameters for conducting the H295R assay protocol (Table 2; Fig. 1). However, at Laboratory 2, there was a greater increase (forskolin) and a lesser decrease (prochloraz) in T concentrations when compared to the other laboratories. Furthermore, at one laboratory, there were instances when decreases in E2 or T production could not be measured due to low basal hormone production (Table 3). In addition, in rare occasions, there was an increase in variation among replicate wells such that the data could not be used. However, this only occurred at one laboratory during a single experiment [Lab 4; chemicals: letrozole, paraben, molinate, Ethylene dimethanesulfonate (EDS); Experiment 1], where the average coefficient of variation (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 was used for the data evaluation. However, it should be emphasized that these were rare events that did not impact the overall validity and utility of data produced during these studies. Overall, only



**Fig. 1** 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). For1=1 μM Forskolin; For10=10 μM Forskolin; Pro0.3=0.3 μM Prochloraz; Pro3=3 μM

Prochloraz. Error bars=1×standard deviation. Bars represent means of four independent experiments. (Lab 5: only T data from two experiments.)

**Table 3** Lowest observed effect concentrations (LOECs; measured by Dunnett's or Mann–Whitney  $U$  test<sup>mu</sup>) and strength and direction of change ( $\downarrow = >0.5$ -fold;  $\downarrow\downarrow = 0.5$ -fold to  $>0.25$ -fold;  $\downarrow\downarrow\downarrow = 0.25$ -fold to  $>0.1$ -fold;  $\downarrow\downarrow\downarrow\downarrow \leq 0.1$ -fold;  $\uparrow = <2$ -fold;  $\uparrow\uparrow = 2$ -fold to  $<4$ -fold;  $\uparrow\uparrow\uparrow = 4$ -fold to  $<20$ -fold;  $\uparrow\uparrow\uparrow\uparrow \geq 20$ -fold) for testosterone (T) and estradiol (E2) after exposure to the 12 core chemicals

	Lab 1 <sup>a</sup>		Lab 2		Lab 3		Lab 4		Lab 6	
	LOEC [ $\mu\text{g/ml}$ ]	Max change								
Fold change (testosterone)										
Prochloraz	0.0001	$\downarrow\downarrow\downarrow$	0.1	$\downarrow\downarrow\downarrow$	0.0001	$\downarrow\downarrow\downarrow$	0.01	$\downarrow\downarrow\downarrow$	0.01	$\downarrow\downarrow$
Aminoglutethimide	100 <sup>d</sup>	$\downarrow\downarrow$	100 <sup>d</sup>	$\downarrow\downarrow$	10	$\downarrow\downarrow$	100 <sup>d</sup>	$\downarrow\downarrow$	100 <sup>d</sup>	$\downarrow\downarrow$
Letrozole	100 <sup>d</sup>	$\downarrow\downarrow$	100 <sup>d</sup>	$\downarrow\downarrow$	100 <sup>a,d</sup>	$\downarrow\downarrow$	100 <sup>d</sup>	$\downarrow\downarrow$	100 <sup>d</sup>	$\downarrow\downarrow$
Nonoxynol-9	10 <sup>c,d</sup>	$\downarrow$	10 <sup>c,d</sup>	$\downarrow$	nd <sup>e</sup>		10 <sup>c,d</sup>	$\downarrow$	10 <sup>c,d</sup>	$\downarrow\downarrow$
Molinate	nd		nd		100	$\downarrow$	Nd		nd	
Benomyl	nd		nd		Nd		nd <sup>mu</sup>		nd	
EDS	Nd									
HCG	Nd									
Paraben	10	$\uparrow$	nd		1	$\uparrow$	nd		nd	
Atrazine	100 <sup>d</sup>	$\uparrow$	1	$\uparrow$	100 <sup>d</sup>	$\uparrow$	nd		nd	
Forskolin	10	$\uparrow\uparrow$	1	$\uparrow\uparrow$	1	$\uparrow$	1	$\uparrow\uparrow$	1	$\uparrow$
Trilostane	<b>0.1<sup>mu</sup></b>	$\uparrow\uparrow\uparrow$	<b>0.01<sup>mu</sup></b>	$\uparrow\uparrow\uparrow$	<b>1<sup>mu</sup></b>	$\uparrow\uparrow\uparrow\uparrow$	<b>1<sup>mu</sup></b>	$\uparrow\uparrow\uparrow\uparrow$	<b>0.01<sup>mu</sup></b>	$\uparrow\uparrow\uparrow\uparrow$
Fold change (estradiol)										
Letrozole	0.001	$\downarrow\downarrow$	0.001	$\downarrow\downarrow$	0.0001 <sup>mu</sup>	$\downarrow\downarrow$	0.01	$\downarrow\downarrow$	0.01	$\downarrow\downarrow$
Prochloraz	0.1	$\downarrow\downarrow$	1	$\downarrow$	0.1	$\downarrow\downarrow$	1	$\downarrow\downarrow$	0.1	$\downarrow\downarrow$
Aminoglutethimide	100 <sup>d</sup>	$\downarrow\downarrow$	10 <sup>mu</sup>	$\downarrow\downarrow$	10	$\downarrow\downarrow$	100 <sup>b,d</sup>	$\downarrow\downarrow$	100 <sup>d</sup>	$\downarrow\downarrow$
Benomyl	Nd		nd		nd <sup>a</sup>		nd <sup>a</sup>		nd	
EDS	Nd									
Nonoxynol-9	Nd									
HCG	Nd		nd	$\uparrow$	nd <sup>a</sup>		nd <sup>a</sup>		nd	
Paraben	nd <sup>mu</sup>	$\uparrow\uparrow\uparrow$	10	$\uparrow$	10 <sup>mu</sup>	$\uparrow\uparrow$	nd	$\uparrow\uparrow$	nd	
Molinate	100 <sup>d</sup>	$\uparrow\uparrow\uparrow\uparrow$	100 <sup>d</sup>	$\uparrow$	100 <sup>d mu</sup>	$\uparrow\uparrow$	100 <sup>d mu</sup>	$\uparrow\uparrow$	100 <sup>d</sup>	$\uparrow\uparrow$
Atrazine	10	$\uparrow\uparrow\uparrow$	1 <sup>mu</sup>	$\uparrow\uparrow\uparrow$	1 <sup>mu</sup>	$\uparrow\uparrow\uparrow\uparrow$	10 <sup>mu</sup>	$\uparrow\uparrow\uparrow$	0.1	$\uparrow$
Forskolin	0.01 <sup>mu</sup>	$\uparrow\uparrow\uparrow\uparrow$	0.1 <sup>mu</sup>	$\uparrow\uparrow\uparrow\uparrow$	0.1 <sup>mu</sup>	$\uparrow\uparrow\uparrow\uparrow$	0.1 <sup>mu</sup>	$\uparrow\uparrow\uparrow\uparrow$	0.01 <sup>mu</sup>	$\uparrow\uparrow$
Trilostane	<b>1<sup>mu</sup></b>	$\uparrow\uparrow\uparrow\uparrow$	<b>100</b>	$\uparrow$	<b>0.1<sup>mu</sup></b>	$\uparrow\uparrow\uparrow\uparrow$	<b>1<sup>mu</sup></b>	$\uparrow\uparrow\uparrow\uparrow$	<b>1<sup>mu</sup></b>	$\uparrow\uparrow\uparrow$

Ranges refer to maximum values measured in repeated experiments

nd Not detectable, — chemical not analyzed, *Data in bold* uncertainty due to interference of the antibody based hormone detection system with the test chemical

<sup>a</sup> Only one experiment was conducted or considered for data evaluation

<sup>b</sup> Not statistically significant;  $p=0.051$

<sup>c</sup> Greatest concentration cytotoxic

<sup>d</sup> Effects occurred at greatest non-cytotoxic concentration; no dose–response

<sup>e</sup> Cytotoxicity observed at concentration at which effects occurred at other laboratories=10

2% or 7% of all experiments for T and E2, respectively, were excluded due to exceedance of permitted variation.

Relative changes in the production of T and E2 after exposure to forskolin and prochloraz in the QC plates were comparable both within and among laboratories (Fig. 1), indicating that the H295R steroidogenesis assay functioned similarly at all laboratories. Coefficients of variation for relative changes measured after exposure to forskolin and prochloraz were between 12% and 13% and between 44%

and 77%, respectively, for T, and between 62% and 73% and 31% and 55%, respectively, for E2. There were no significant decreases in cell viability between any of the different treatment groups (results not shown).

### 3.2 Core chemical exposure experiments

There were chemical-specific differences in the response of T production after exposure of H295R cells to the 12 core

chemicals (Table 3). With a few exceptions, the observed chemical-specific responses of T production were comparable among laboratories and could be grouped into three different types of effects: inducers, inhibitors, and negative reference chemicals. Among the inducers, exposure to trilostane resulted in the greatest fold changes (>10-fold induction) in T concentration when compared to SCs. The least fold changes were observed for the atrazine exposures where induction of T production was less than 1.5-fold with the exception of Lab 2, at which maximum induction was 2.4-fold. No effect on T production was observed after exposure to atrazine at Lab 6. Exposure to prochloraz resulted in a greater than 15-fold reduction of T production at the greatest concentration tested (100  $\mu\text{M}$ ) at all laboratories with the exception of Lab 4 where an up to 4.5-fold reduction was observed. The greater LOEC reported for Lab 2 is likely a function of the relatively great variation among replicate experiments at 0.01 M (CV=35%). It is unclear why T production by cells was more sensitive to the exposure with prochloraz at Labs 1 and 3. However, a concentration-dependent response was observed starting at 0.01 M, which is similar to the response patterns at the other labs. Therefore, it cannot be excluded that the significant reduction at 0.0001 and 0.001 M represents an artifact. Exposure to the other inhibitors resulted in less than fourfold changes in T production. When chemicals exhibited a less than 1.5-fold change in T production, they were categorized as negatives. This threshold was defined based on the average variation observed across all laboratories among replicate experiments. Some of these negative chemicals could have been categorized as inhibitors in individual cases (molinate: Lab 4; benomyl: Lab 1). However, even in situations where inhibition was observed at an individual laboratory, changes were always less than twofold and typically were not concentration-dependent. For instance, exposure to nonoxynol-9 resulted in a decrease in T concentrations at non-cytotoxic concentrations at two of five laboratories for which data was available. Relative to the SCs, inhibition of T production at Lab 1 was 29% (1  $\mu\text{M}$ ), while at Lab 2, it was 47% (10  $\mu\text{M}$ ). However, it should be noted that, at Lab 2, exposure to 10  $\mu\text{M}$  nonoxynol-9 resulted in an average increase in cell viability (138% viable cells relative to the SCs), 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 by any of the other groups. The greatest letrozole concentration resulted in a significant decrease in T at all laboratories.

Significant differences in E2 production were observed for H295R cells exposed to the 12 core chemicals (Table 3). The direction of the effect for each chemical was comparable among laboratories (Table 3). Three chemicals inhibited E2 concentrations (letrozole, prochloraz, and

aminoglutethimide), while human chorionic gonadotrophin (HCG), EDS, benomyl, and nonoxynol-9 did not elicit any clear (>1.5-fold) effects at non-cytotoxic concentrations. For inducers of E2 production, the magnitude of the response ranged between 20-fold or greater (forskolin) to <threefold (paraben) than SCs. The most potent inducer of E2 production was forskolin. Exposure to forskolin resulted in increases in E2 production at concentrations greater or equal to 0.1  $\mu\text{M}$ , while exposure to other inducers typically did not reveal effects at concentrations less than 1  $\mu\text{M}$ . While responses for E2 after exposure to atrazine appeared to be greater or equal to two orders of magnitude more sensitive than at the majority of the other labs no concentration-dependent response pattern occurred up to 1  $\mu\text{M}$ . In fact, increases in E2 concentrations did not follow a concentration response at lesser concentrations and were very small (1.16-fold greater than SCs). The most potent inhibitors were letrozole and prochloraz, exposure to which resulted in marked reductions of E2 at concentrations greater 0.001 and 0.1  $\mu\text{M}$ , respectively. The exception to this pattern was exposure to letrozole at Lab 6, for which significant reductions occurred at concentrations greater than 0.01  $\mu\text{M}$ . Exposure to aminoglutethimide, in contrast, only caused a clear reduction in E2 concentrations at the greatest concentration tested. Variation between laboratories did not exceed twofold for a given concentration with the exception of trilostane.

### 3.3 Supplemental chemical exposure experiments

For the additional 16 chemicals, the H295R steroidogenesis assay was able to categorize inducers and inhibitors of T and E2 (Table 4). Five (31%) chemicals tested negative for T production and included di-(2-ethylhexyl) phthalate (DEHP), dimethoate, flutamide, glyphosate, and prometon. Four (25%) chemicals tested negative for E2 and included glyphosate, dinitrophenol, piperonyl butoxide, and spiro-nolactone. One exception was dinitrophenol, which was identified as a significant inhibitor of T at all concentrations tested at the 1st lab. However, changes in T were not concentration-dependent and the magnitude of the effect was weak (inhibition did not exceed 0.67-fold relative to the SC at any given exposure concentration). Therefore, it is possible that this response represents an artifact. Some of the chemicals identified as inhibitors of T showed a biphasic response where slight increases in hormone production were observed at concentrations of up to 1  $\mu\text{M}$ . However, with the exception of genistein, none of these changes exceeded 1.5-fold. Compared to the 12 core chemicals, there was greater variation among the responses observed at different laboratories for the 16 supplemental chemicals. Approximately 19% and 31% of the chemicals showed a significant response for T and E2, respectively, at

**Table 4** Lowest observed effect concentrations (LOECs; measured by Dunnett's test) and strength and direction of change ( $\downarrow$ =>0.5-fold;  $\downarrow\downarrow$ =0.5-fold to >0.25-fold;  $\downarrow\downarrow\downarrow$ =0.25-fold to >0.1-fold;  $\downarrow\downarrow\downarrow\downarrow$ =≤0.1-fold;  $\uparrow$ =<2-fold;  $\uparrow\uparrow$ =2-fold to <fold;  $\uparrow\uparrow\uparrow$ =4-fold to <20-fold;  $\uparrow\uparrow\uparrow\uparrow$ ≥20-fold) observed for the 16 test chemicals

	LOEC [ $\mu\text{g/ml}$ ]		Max change	
	1st Lab <sup>a</sup>	2nd Lab <sup>d</sup>	1st Lab	2nd Lab
<b>Testosterone</b>				
Ketoconazole	1	1	$\downarrow\downarrow\downarrow$	$\downarrow\downarrow\downarrow$
Genistein	10	10	$\downarrow\downarrow$	$\downarrow\downarrow\downarrow$
Finasteride	10	100 <sup>c</sup>	$\downarrow\downarrow$	$\downarrow\downarrow$
Bisphenol A	10	10	$\downarrow\downarrow$	$\downarrow$
Dinitrophenol	0.0001	100 <sup>c</sup>	$\downarrow$	$\downarrow\downarrow$
Piperonyl butoxide	10	10	$\downarrow$	$\downarrow$
Spironolactone	1	1	$\downarrow\downarrow\downarrow$	$\downarrow\downarrow$
Fenarimol	nd	10	Nd	$\downarrow\downarrow$
Danazol	nd	nd	Nd	nd
DEHP	nd	nd	Nd	nd
Dimethoate	nd	nd	Nd	nd
Flutamide	nd	nd	Nd	nd
Glyphosate	nd	nd	Nd	nd
Prometon	nd	nd	Nd	nd
Tricrescyl phosphate	10	nd	$\uparrow$	nd
Mifepristone	0.1	nd	$\uparrow$	nd
<b>Estradiol</b>				
Danazol	1	10	$\downarrow\downarrow\downarrow$	$\downarrow\downarrow$
Ketoconazole	10	10	$\downarrow\downarrow$	$\downarrow\downarrow$
Fenarimol	nd	1	nd	$\downarrow\downarrow$
Finasteride	nd	100 <sup>c</sup>	nd	$\downarrow$
Glyphosate	nd	nd	nd	nd
Dinitrophenol	nd	nd	nd	nd
Spironolactone	nd	nd	nd	nd
Piperonyl butoxide	nd	nd	nd	nd
Dimethoate	10	nd	$\uparrow\uparrow$	nd
Flutamide	10	nd	$\uparrow\uparrow$	nd
Tricrescyl phosphate	10	nd	$\uparrow\uparrow\uparrow$	nd
Bisphenol A	10	1	$\uparrow\uparrow$	$\uparrow\uparrow$
DEHP	1 <sup>b</sup>	1	$\uparrow\uparrow$	$\uparrow\uparrow$
Mifepristone	0.1	1	$\uparrow\uparrow$	$\uparrow\uparrow$
Prometon	100 <sup>c</sup>	100 <sup>c</sup>	$\uparrow\uparrow\uparrow\uparrow$	$\uparrow\uparrow$
Genistein	10	10	$\uparrow\uparrow\uparrow\uparrow$	$\uparrow\uparrow\uparrow\uparrow$

Chemicals tested at the second laboratories (2nd Lab) were as follows: Lab 2: Piperonyl butoxide, ketoconazole, prometon, DEHP, flutamide, and danazol; Lab 3: Bisphenol A, fenarimol, genistein, finasteride, and dinitrophenol; and Lab 4: Spironolactone, mifepristone, tricrescyl phosphate, dimethoate, and glyphosate

nd Not detectable

<sup>a</sup> Lead laboratory (Lab 1)

<sup>b</sup> Considered because there was a clear concentration–response at all but the greatest concentration

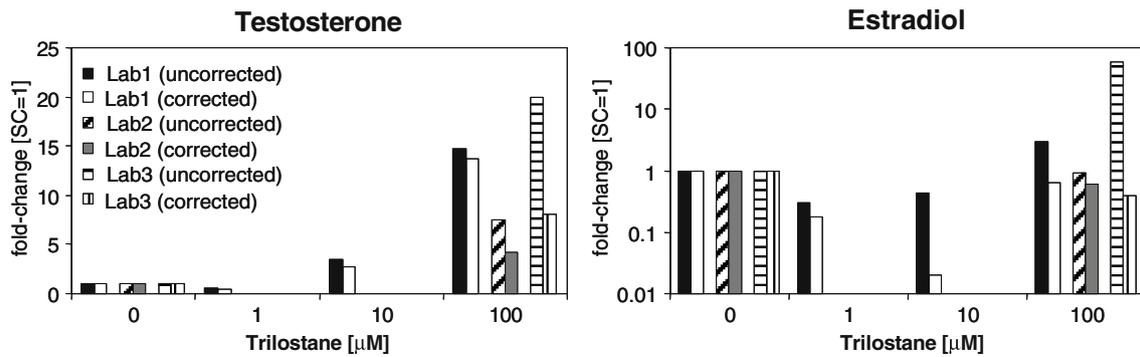
<sup>c</sup> Effects occurred at greatest non-cytotoxic concentration; no dose–response

<sup>d</sup> Participating laboratory (Labs 2, 3, and 4)

only one of the two laboratories where they were tested (Table 4). These were fenarimol, finasteride, dimethoate, flutamide, and tricrescyl phosphate for E2, and fenarimol, mifepristone, and tricrescyl phosphate for T. It is unclear what the bases for these differences are, but it should be noted that in four out of the eight cases where such incongruencies were observed (E2: dimethoate, tricrescyl phosphate; T: mifepristone, tricrescyl phosphate), they were associated with one group (Lab 4). In all four cases, these chemicals were identified as inducers by Lab 1, while no statistically significant effects were reported by the other testing group. Also, at the same laboratory, some of the cell viability data revealed no effects where significant decreases were observed at Lab 1 (tricrescyl phosphate and spironolactone). This result indicates that there may have been some issues related to dosing. Finally, basal E2 production measured by Lab 4 was approximately three-to-four times greater than that measured by Lab 1 (~200 vs. ~50 pg/ml), indicating that cells were at a suboptimal (late) passage when used for the experiment. This further supports the need for stringent conditions regarding the age of the cells, which should not be used beyond passage 10. When excluding this group, the data obtained at different laboratories for T and E2 did not match for one and three chemicals, respectively.

#### 3.4 Confounding factors—interference with hormone detection assays

The analysis of cross-reactivity of each chemical with the antibodies of the immunoassays used at most of the laboratories revealed interaction with a few chemicals at the greatest concentrations tested. A large interaction of the E2 immunoassay with trilostane (up to 100% of the overall response measured at the greatest test concentration) was observed at all laboratories with the exception of Lab 2. Similarly, a less pronounced cross-reactivity of trilostane was also reported for the T antibodies (up to 60% of overall response at the greatest concentration tested). However, since at most of the laboratories only the greatest chemical concentration was evaluated, an adjustment of the concentration–response curves could not be performed. However, an attempt to correct for the interaction with the antibodies at this greatest concentration (greatest three concentrations for Lab 1) indicated that while the induction of E2 after exposure to trilostane is likely to be solely due to this cross-reactivity, the induction of T could not be explained alone by this factor (Fig. 2). Similar interactions of trilostane with hormone detection systems have been also observed by other authors (Shapiro and Page 1976; Puddefoot et al. 2002; Villeneuve, personal communication). In addition, nonoxynol-9, paraben, and prochloraz also interacted with the E2 immunoassays. However, since the cross-reactivity



**Fig. 2** Changes in the concentrations of testosterone (T) and estradiol (E2) relative to the solvent controls (SC=1) after exposure to trilostane with (corrected) and without (uncorrected) adjustment of

final hormone concentration for interference with the antibody-based hormone detection system. Bars represent average responses of one (Lab 1), two (Lab 3), and three (Lab 2) independent experiments

of prochloraz, paraben, and nonoxynol-9 at the greatest concentrations tested were either low or these concentrations were excluded due to marked cytotoxicity, this factor had no effect on the interpretation of the results. Significant interactions of the chemicals with the hormone detection assays that occurred at non-cytotoxic concentrations were only observed for T after exposure to spironolactone, finasteride, and danazol at Lab 1 and for E2 after exposure to genistein at Labs 1 and 3. When uncorrected data for spironolactone, finasteride, and danazol were compared to the data corrected for this interference, significant impacts on the overall trend/response were not observed (data not shown). Similarly, while genistein interference with the E2 ELISA antibodies reduced the magnitude of the response by approximately 30%, it did not change the overall trend of the response. However, further analyses are required to address possible uncertainties resulting from the interference of a test chemical with the hormone detection system utilized.

### 3.5 Predictive power and accuracy of H295R steroidogenesis assay

In addition to the ability of an assay to produce reliable and transferable results as assessed in this validation effort, the potential of data obtained with an *in vitro* test, such as the H295R steroidogenesis assay, to be predictive of effects at higher organizational levels, such as organisms, is one of the key parameters relevant to its use as a screening tool. Comparisons of the *in vivo* and *in vitro* effects of prochloraz, ketoconazole, fenarimol, prometon, and aminoglutethimide have been made previously (Hecker et al. 2006; Villeneuve et al. 2007), and the findings reported in this study were similar to those reported by these authors. In brief, while not necessarily directly predictive of the direction of the responses *in vivo*, the H295R always captured an effect if there was an alteration in hormone profiles *in vivo*.

A comparison of the effects of E2 inducers observed in the H295R validation studies and the findings of *in vivo* studies showed that the results were comparable for six out of ten chemicals tested: atrazine (Wetzel et al. 1994; Spano et al. 2004), mifepristone (Fassett et al. 2008; Wang et al. 1994), danazol (Peters et al. 1980), tricresyl phosphate (Latendresse et al. 1995), flutamide (Andrews et al. 2000), and genistein (Harrison et al. 1999; Table 5). The results obtained with H295R cells for inhibitors of E2 production corresponded to the findings of *in vivo* studies for five out of six chemicals studied: letrozole (Kumru et al. 2007), aminoglutethimide (Berman and Laskey 1993; Monteiro et al. 2000), prochloraz (Vinggaard et al. 2005; Brande-Lavridsen et al. 2008), ketoconazole (Monteiro et al. 2000), and fenarimol (Ankley et al. 2005; Table 5). In only three cases were there opposite trends among results for E2 production obtained with the H295R steroidogenesis assay and *in vivo* tests. Exposure to DEHP, prometon, and bisphenol A *in vivo* resulted in an inhibition (Davis et al. 1994) and no effect (Villeneuve et al. 2006; Yamasaki et al. 2002) on E2 concentrations, respectively, while all three chemicals caused a significant increase in E2 *in vitro* in the present study. However, the increase of E2 concentrations observed with the H295R cells for prometon may have been an indicator for the decrease in the expression of secondary sex characteristics observed in male fish (Villeneuve et al. 2007). In the case of bisphenol A, the lack of response in the *in vivo* studies is likely due to the administration route of BPA, which was via gavage. Previous studies have reported that orally administered bisphenol A has very low bioavailability and is rapidly excreted (Pottenger et al. 2000). Three chemicals that tested negative for E2 effects *in vitro* (H295R), namely, benomyl, dimethoate, and glyphosate, also did not cause any changes in serum E2 concentrations *in vivo* (Spencer et al. 1996; Rawlings et al. 1998; Soso et al. 2007). No studies describing *in vivo* effects on the production of E2 were

**Table 5** Comparison of data obtained with the H295R steroidogenesis assay (this study) with in vivo data

Chemical	Testosterone		Estradiol		Species	References
	H295R	In vivo	H295R	In vivo		
Aminoglutethimide	↓	↑	↓	↓	Fish; rat	Berman and Laskey 1993; Monteiro et al. 2000
Atrazine	↑/–	–	↑	↑	Fish; rat	Spano et al. 2004; Wetzel et al. 1994
Benomyl	–	–	–	–	Rat	Carter and Laskey 1982; Spencer et al. 1996
Bisphenol A	↑	–	↓	–	Rat	Yamasaki et al. 2002
Butyl paraben	–	–	↑	n.d.	Rat	Taxvig et al. 2008
Danazol	–	n.d.	↑	↑	Human	Murakami et al. 1993; Peters et al. 1980
DEHP	–	–	↑	↓	Rat	Davis et al. 1994; Noriega et al. 2009
Dimethoate	–	n.d.	–	–	Sheep	Rawlings et al. 1998
Dinitrophenol	–	n.d.	–	n.d.	–	–
EDS	–	n.d.	–	n.d.	–	–
Fenarimol	↓	↑	↓	↓	Fish	Ankley et al. 2005
Finasteride	–	n.d.	–	n.d.	Amphibian	–
Flutamide	–	↑	↑	↑	Rat	Mikkilä et al. 2006; Andrews et al. 2000
Forskolin	↑	n.d.	↑	n.d.	–	–
Genistein	↓	↓	↑	↑	Monkey; rat	Harrison et al. 1999; Ohno et al. 2003
Glyphosate	–	–	–	–	Fish	Soso et al. 2007
HCG	–	– <sup>b</sup>	–	– <sup>b</sup>	– <sup>b</sup>	– <sup>b</sup>
Ketoconazole	↓	↓	↓	↓	Fish; rat	Monteiro et al. 2000; O'Connor et al. 2002
Letrozole	↓	↑	↓	↓	Rat	Kumru et al. 2007
Mifepristone	↑	↑	↑	↑	Human	Fassett et al. 2008; Wang et al. 1994
Molinate	↑	n.d. <sup>c</sup>	–	–	Rat	Ellis et al. 1998
Nonoxynol-9	–	n.d.	–	n.d.	–	–
Piperonyl butoxide	↓	n.d.	–	n.d.	–	–
Prochloraz	↓	↓	↓	↓	Amphibian; rat	Vinggaard et al. 2005; Brande-Lavridsen et al. 2008
Prometon	–	–	↑	– <sup>d</sup>	Fish	Villeneuve et al. 2006
Spironolactone	↓	↓	–	n.d. <sup>d</sup>	Amphibian; rat	Canosa and Ceballos 2001; Yamasaki et al. 2004
Tricrescyl phosphate	↑/–	n.d.	↑	↑	Rat	Latendresse et al. 1995
Trilostane	↑ <sup>a</sup>	↓/–	– <sup>a</sup>	–	Fish; human; rat	Jungmann et al. 1983; Villeneuve et al. 2006

↑ Increase, ↓ Decrease, – No effect, *n.d.* no data/study available

<sup>a</sup> Corrected for cross-reactivity

<sup>b</sup> H295R assay only capture effects downstream of LH/FSH

<sup>c</sup> Data not considered because of inconclusive results

<sup>d</sup> Identified as a reproductive toxicant in vivo

found for the other chemicals tested. However, given the general toxic properties of chemicals, such as nonoxynol-9 (spermicide), EDS (cytotoxicant to Leydig cells; Cooper and Jackson 1970; Kerr et al. 1985), and dinitrophenol (metabolic poison uncoupling oxidative phosphorylation), no specific interactions with the steroidogenic pathway at non-cytotoxic concentrations would be expected.

In general, effects on T production were less consistent when the results obtained with H295R cells were compared to those of in vivo studies (Table 5). Only one chemical of the five found to be inducers of T production in the cells, mifepristone, showed a similar trend in vivo (Wang et al.

1994), while three of the seven inhibitors (prochloraz: Vinggaard et al. 2005; Brande-Lavridsen et al. 2008, ketoconazole: O'Connor et al. 2002; Monteiro et al. 2000, genistein: Ohno et al. 2003) revealed comparable trends between the results of the validation studies and previously reported in vivo data. However, five of the test chemicals demonstrated conflicting trends between the results obtained with the H295R cells and those from in vivo studies: the inducers atrazine (Wetzel et al. 1994; Spano et al. 2004) and trilostane (Jungmann et al. 1983), the inhibitors letrozole (Kumru et al. 2007) and aminoglutethimide (Berman and Laskey 1993; Monteiro et al. 2000), and

the negative chemical bisphenol A (Yamasaki et al. 2002). As previously discussed for E2, the lack of response reported for bisphenol A in vivo was likely a function of low bioavailability and rapid excretion due to the form of administration (oral; Pottenger et al. 2000). With the exception of flutamide, 5 of the 11 chemicals that tested negative for changes in T production in the H295R steroidogenesis assay were also reported as causing no significant alterations in T concentrations in vivo: flutamide (Mikkilä et al. 2006), glyphosate (Soso et al. 2007), DEHP (Noriega et al. 2009), benomyl (Carter and Laskey 1982), and molinate (Ellis et al. 1998). For flutamide, a significant induction in T production was reported in rats in vivo (Andrews et al. 2000). Information on the effects of the other chemicals on production of T in vivo could not be found. The reason for the increased number of chemicals showing discrepancies between in vivo studies and the current work in the production of T as opposed to that of E2 is likely due to the intermediate role of T in the steroidogenesis pathway, which makes it possible that changes in T can be better compensated by the cells than those in E2.

Overall, no chemical was falsely characterized as having no effect by the H295R steroidogenesis assay based on its known mechanism of action with the exception of T production after exposure to flutamide. However, this chemical would have been flagged due to a comparable in vivo/in vitro effect on E2. There were no studies describing the effects of the model inducer forskolin on hormone homeostasis in vivo. However, considering the rapid metabolism of forskolin by an organism no marked effects would be expected. Overall, these results indicate that, while not necessarily always directly predictive of a specific type of response in an organism, the H295R assay system always flagged a chemical as a potential disruptor of steroidogenic processes. Furthermore, there were only two chemicals for which both in vivo and in vitro data were available that would have been wrongly characterized as either inducers or inhibitors of each hormone tested by the H295R steroidogenesis assay [atrazine (Wetzel et al. 1994) and bisphenol A (Yamasaki et al. 2002)] for T induction and prometon (Villeneuve et al. 2006) and bisphenol A (Yamasaki et al. 2002) for E2 inhibition].

#### 4 Conclusions

It was demonstrated that, with one exception, the H295R steroidogenesis assay protocol successfully identified the majority of chemicals with known and unknown modes of interaction with the production of T and E2. The results obtained in the current study confirm the findings reported for H295R cells by Hecker et al. (2006) as well as effects

described in other in vitro and in vivo studies (discussed in Hecker et al. 2006 and this manuscript) for a broad range of chemicals. One of the remaining limitations associated with the H295R steroidogenesis assay protocol is the relatively low basal production of E2 and its effect on quantifying the decreased production of this hormone with regard to the identification of weak inhibitors. To address this uncertainty, there should be further efforts aimed at increasing basal hormone production, e.g., by altering the cells or test protocols without affecting the potential of the cells to detect inducers of E2 production. Furthermore, most of the variation observed among laboratories in this study was likely due to changes in test practices and personnel during the course of this validation study. To address similar issues in the future, a number of additional performance criteria were included into the test protocols. These include the addition of a proficiency test that is required of each laboratory that plans to start using the assay or that has undergone changes in personnel, and the flexible protocols for refinement of the spacing of test chemical concentrations to enable the description of more precise concentration–response relationships. An initial comparison of H295R data from this study to in vivo studies from the literature demonstrated the potential of the H295R steroidogenesis assay to identify chemicals affecting hormone homeostasis in whole organisms. Particularly promising was the lack of any false negatives during the validation. Furthermore, the very low number of chemicals giving false positives represents an important aspect of this bioassay since it confirms the specificity of the test and will help avoid unnecessary additional testing. Future studies, including a larger number of chemicals with different structures and properties as well as comparison to parallel studies with whole organisms, should be conducted to confirm the predictive power of the H295R steroidogenesis assay for in vivo scenarios.

#### 5 Future perspectives

Based on the results obtained during this validation study and the accordingly revised test protocols, an OECD draft test guideline has been developed and submitted to the OECD working group of the national coordinators of the test guidelines program (WNT) for comments in December 2009 (<http://www.oecd.org/dataoecd/56/11/44285292.pdf>). Once accepted, this test guideline will replace the current H295R steroidogenesis assay protocol of US EPA's EDSP. Implementation of similar testing strategies for endocrine disruptors is currently discussed in the context of other chemical screening programs, such as REACH, but no definite decisions have been made as regards these to date. Furthermore, the H295R steroidogenesis assay has been

shown to be a valuable tool for the characterization of the endocrine potential of effluents and environmental samples (Kase et al. 2009).

**Acknowledgements** We thank the OECD VMG NA as well as the OECD secretariat for their support, encouragement, and guidance. Furthermore, we acknowledge many helpful discussions and manuscript review by Dr. Jerome Goldman, Endocrinology Branch, NHEERL, US EPA, Research Triangle Park, NC. Funding for this project was provided by US EPA, ORD Service Center/NHEERL (contract # GS-10F-0041L), a grant from the Western Economic Diversification Canada (project #s 6971 and 6807), and the Nordic Chemical Group (project # 2008-06-05). Conclusions drawn in this manuscript neither constitute nor necessarily reflect US EPA policy regarding the test chemicals and methods.

## References

- Andrews P, Freyberg A, Hartmann E, Eiben R, Loof I, Schmidt U, Temerowski M, Becka M (2000) Feasibility and potential gains of enhancing the subacute rat study protocol (OECD test guideline no. 407) by additional parameters selected to determine endocrine modulation. A pre-validation study to determine endocrine-mediated effects of the antiandrogenic drug flutamide. *Arch Toxicol* 75:65–73
- Ankley GT, Jensen KM, Durhan EJ, Makynen EA, Butterworth BC, Kahl MD, Villeneuve DL, Linnum A, Gray LE, Cardon M, Wilson VS (2005) Effects of two fungicides with multiple modes of action on reproductive endocrine function in the fathead minnow (*Pimephales promelas*). *Toxicol Sci* 86:300–308
- Battelle (2005) Sliced testis assay prevalidation: baseline/positive control study
- Berman E, Laskey JW (1993) Altered steroidogenesis in whole-ovary and adrenal culture in cycling rats. *Reprod Toxicol* 7:349–358
- Brande-Lavridsen N, Christensen-Dalsgaard J, Korsgaard B (2008) Effects of prochloraz and ethinylestradiol on sexual development in *Rana temporaria*. *J Exp Zool* 309A:389–398
- Canosa LF, Ceballos NR (2001) Effects of different steroid-biosynthesis inhibitors on the testicular steroidogenesis of the toad *Bufo arenarum*. *J Comp Physiol* 171: 519–26
- Carter SD, Laskey JW (1982) Effects of benomyl on reproduction in the male rat. *Toxicol Lett* 11:87–94
- Chang H, Wan Y, Naile J, Zhang X, Wiseman S, Hecker M, Hu J, Lam MHW, Giesy JP, Jones PD (2010) Simultaneous quantification of multiple classes of phenolic compounds in blood plasma by liquid chromatography–electrospray tandem mass spectrometry. *J Chromatogr A* 1217:506–513
- Cooper ERA, Jackson H (1970) Comparative effects of methylene, ethylene and propylene dimethane–sulphonates on the male rat reproductive system. *J Reprod Fertil* 23:103–108
- Davis BJ, Maronpot RR, Heindel JJ (1994) Di-(2-ethylhexyl) phthalate suppresses estradiol and ovulation in cycling rats. *Toxicol Appl Pharmacol* 128:216–223
- ECVAM (1995) Practical aspects of the validation of toxicity test procedures: the report and recommendations of ECVAM Workshop No. 5. *ATLA* 23:129–147
- EDMVAC (2005) Letter from the EDMVAC to Stephen L. Johnson, Administrator of EPA, May 5, 2005
- EDSTAC (1998) Endocrine disruptor screening and testing advisory committee final report 1998. U.S. Environmental Protection Agency. <http://www.epa.gov/opptintr/opptendo/finalrpt.htm>
- Ellis MK, Richardson AG, Foster JR, Smith FM, Widdowson PS, Farnworth MJ, Moore RB, Pitts MR, Wickramaratne GAdS (1998) The reproductive toxicity of molinate and metabolites to the male rat: effects on testosterone and sperm morphology. *Toxicol Appl Pharmacol* 151:22–32
- Fassett MJ, Lachelin GCL, McGarrigle HHG, Wing DA (2008) Alterations in saliva steroid hormone levels after oral mifepristone administration in women with pregnancies of greater 41 weeks' gestation. *Reprod Sci* 15:394–399
- Harrison RM, Phillippi PP, Swan KF, Henson MC (1999) Effect of genistein on steroid hormone production in the pregnant rhesus monkey. *Exp Biol Med* 222:78–84
- Hecker M, Giesy JP (2008) Novel trends in endocrine disruptor testing: the H295R steroidogenesis assay to identify inducers and inhibitors of hormone production. *Anal Biochem Chem* 390:287–291
- Hecker M, Newsted JL, Murphy MB, Higley EB, Jones PD, Wu R, Giesy JP (2006) Human adrenocarcinoma (H295R) cell system for rapid testing of chemical effects on steroidogenesis. *Toxicol Appl Pharmacol* 217:114–124
- Hecker M, Hollert H, Cooper R, Vinggaard A-M, Akahori Y, Murphy M, Nellesmann C, Higley E, Newsted J, Wu R, Lam P, Laskey J, Buckalew A, Grund S, Nakai M, Timm G, Giesy J (2007) The OECD validation program of the H295R steroidogenesis assay for the identification of in vitro inhibitors and inducers of testosterone and estradiol production: Phase 2. Inter-laboratory pre-validation studies. *Environ Sci Pollut Res* 14(Special Issue 1):23–30
- 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
- Jungmann E, Althoff PH, Balzer-Kuna S, Magnet W, Rottmann-Kuhnke U, Sprey R, Schwedes U, Usadel KH, Schöffling K (1983) The inhibiting effect of trilostane on testosterone synthesis. Hormonal and morphological alterations induced by subchronic trilostane treatment in rats and healthy volunteers. *Arzneimittelforschung* 33:754–756
- Kase R, Kunz P, Gerhardt A (2009) Identifikation geeigneter Nachweismöglichkeiten von homronaktiven und reproduktionstoxischen Wirkungen in aquatischen Oekosystemen. *Umweltwiss Schadt Forsch*. doi:10.1007/s12302-009-0072-2, Published online: 16 June 2009
- Kerr JB, Donachie K, Rommerts FFG (1985) Selective destruction and regeneration of rat Leydig cells in vivo: a new method for the study of seminiferous tubular–interstitial tissue interaction. *Cell Tissue Res* 242:145–156
- Kumru S, Yildiz AA, Yilmaz B, Sandal S, Gurates B (2007) Effects of aromatase inhibitors letrozole and anastrozole on bone metabolism and steroid hormone levels in intact female rats. *Gynecol Endocrinol* 23:556–561
- Latendresse JR, Brooks CL, Capen CC (1995) Toxic effects of butylated triphenyl phosphate-based hydraulic fluid and tricresyl phosphate in female F344 rats. *Vet Pathol* 32:394–402
- Mikkilä TFM, Toppari J, Paranko J (2006) Effects of neonatal exposure to 4-tert-octylphenol, diethylstilbestrol, and flutamide on steroidogenesis in infantile rat testis. *Toxicol Sci* 91:456–466
- Monteiro PRR, Reis-Henriques MA, Coimbra J (2000) Polycyclic aromatic hydrocarbons inhibit in vitro ovarian steroidogenesis in the flounder (*Platichthys flesus* L.). *Aquat Toxicol* 48:549–559
- Mosman T (1983) Rapid colometric assay for growth and survival: application to proliferation and cytotoxicity. *J Immunol Methods* 65:55–63
- Murakami K, Nakagawa T, Yamashiro G, Araki K, Akasofu K (1993) Adrenal steroids in serum during danazol therapy, taking into account cross-reactions between danazol metabolites and serum androgens. *Endocr J* 40:659–64
- Noriega NC, Howdeshell KL, Furr J, Lambright CR, Wilson VS, Gray LE Jr (2009) Pubertal administration of DEHP delays puberty,

- suppresses testosterone production, and inhibits reproductive tract development in male Sprague–Dawley and Long–Evans rats. *Toxicol Sci* 111:163–178
- O'Connor JC, Frame SR, Ladies GS (2002) Evaluation of a 15-day screening assay using intact male rats for identifying steroid biosynthesis inhibitors and thyroid modulators. *Toxicol Sci* 69:79–91
- 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) OECD conceptual framework for the testing and assessment of endocrine disrupting chemicals, in annex 2 of: OECD guideline for the testing of chemicals no. 440, the uterotrophic bioassay in rodents: a short-term screening test for oestrogenic properties (2007). Available at <http://titania.sourceoecd.org/v1=1549790/c1=28/nw=1/rpsv/cw/vhosts/oecdjournals/1607310x/v1n4/contp1-1.htm>
- Ohno S, Nakajima Y, Inoue K, Nakazawa H, Nakajin S (2003) Genistein administration decreases serum corticosterone and testosterone levels in rats. *Life Sci* 74:733–742
- Peters F, Reck G, Zimmermann G, Breckwoldt M (1980) The effect of danazol on the pituitary function, thyroid function, and mastodynia. *Arch Gynecol* 230:3–8
- Pottenger LH, Domoradzki JY, Markham DA, Hansen SC, Cagan SZ, Waechter JM Jr (2000) The relative bioavailability and metabolism of bisphenol A in rats is dependent upon the route of administration. *Toxicol Sci* 54:3–18
- Puddefoot JR, Barker S, Glover HR, Malouitre SDM, Vinson GP (2002) Non-competitive steroid inhibition of oestrogen receptor functions. *Int J Cancer* 101:17–22
- Rawlings NC, Cook SJ, Waldbillig D (1998) Effects of the pesticides carbofuran, chlorpyrifos, dimethoate, lindane, triallate, trifluralin, 2, 4-D, and pentachlorophenol on the metabolic endocrine and reproductive endocrine system in ewes. *J Toxicol Environ Health A* 54:21–36
- Shapiro R, Page LB (1976) Interference by 2, 3-dimercapto-1-propanol (BAL) in angiotensin I radioimmunoassay. *J Lab Clin Med* 2:22–31
- Soso AB, Barcellos LJG, Ranzani-Paiva MJ, Kreutz LC, Quevedo RM, Anziliero D, Lima M, da Silva LB, Ritter F, Bedin AC, Finco JA (2007) Chronic exposure to sub-lethal concentration of a glyphosate-based herbicide alters hormone profiles and affects reproduction of female Jundi'a (*Rhamdia quelen*). *Environ Toxicol Pharm* 23:308–313
- Spano L, Tyler CR, van Aerle R, Devos P, Mandiki SNM, Silvestre F, Thom c J-P, Kestemont P (2004) Effects of atrazine on sex steroid dynamics, plasma vitellogenin concentration and gonad development in adult goldfish (*Carassius auratus*). *Aquat Toxicol* 66:369–379
- Spencer F, Chi L, Zhu M-X (1996) Effect of benomyl and carbendazim on steroid and molecular mechanisms in uterine decidual growth in rats. *J Appl Toxicol* 16:211–214
- Taxvig C, Vinggaard AM, Hass U, Axelstad M, Boberg J, Reimer Hansen P, Frederiksen H, Nellemann C (2008) Do Parabens have the ability to interfere with steroidogenesis? *Tox Sci* 106:206–213
- Villeneuve DL, Murphy MB, Kahl MD, Jensen KM, Butterworth BC, Durhan EJ, Linnum A, Leino RL, Curtis LR, Giesy JP, Ankley GT (2006) Evaluation of the methoxy-triazine herbicide prometon, using a short-term fathead minnow reproduction test and a suite of in vitro bioassays. *Environ Toxicol Chem* 25:2143–2153
- Villeneuve DL, Ankley GT, Makynen EA, Blake LS, Greene KJ, Higley EB, Newsted JL, Giesy JP, Hecker M (2007) Comparison of fathead minnow ovary explant and H295R cell-based steroidogenesis assays for identifying endocrine-active chemicals. *Ecotoxicol Environ Saf* 68:20–32
- Vinggaard AM, Christiansen S, Laier P, Poulsen ME, Breinholt V, Jarfelt K, Jacobsen H, Dalgaard M, Nellemann C, Hass U (2005) Perinatal exposure to the fungicide prochloraz feminizes the male rat offspring. *Toxicol Sci* 85:886–897
- Wang J-D, Shi W-L, Zhang G-Q, Bai X-M (1994) Tissue and serum levels of steroid hormones and RU 486 after administration of mifepristone. *Contraception* 49:245–253
- Wetzel LT, Luempert LG III, Breckenridge CB, Tisdell MO, Stevens JT (1994) Chronic effects of atrazine on estrus and mammary tumor formation in female Sprague–Dawley and Fischer 344 rats. *J Toxicol Environ Health* 43:169–182
- Yamasaki K, Sawaki M, Noda S, Imatanaka N, Takatsuki M (2002) Subacute oral toxicity study of ethynylestradiol and bisphenol A, based on the draft protocol for the 'Enhanced OECD test guideline no. 407. *Arch Toxicol* 76:65–74
- Yamasaki K, Sawaki M, Noda S, Muroi T, Takakura S, Mitoma H, Sakamoto S, Nakai M, Yakabe Y (2004) Comparison of the Hershberger assay and androgen receptor binding assay of twelve chemicals. *Toxicology* 195:177–186