Reproductive Toxicology 55 (2015) 11–19

Contents lists available at ScienceDirect

Reproductive Toxicology

journal homepage: www.elsevier.com/locate/reprotox

Evaluation of an alternative *in vitro* test battery for detecting reproductive toxicants in a grouping context

E. Dinant Kroese[a,*], Sieto Bosgra[a], Harrie E. Buist[a], Geertje Lewin[e], Sander C. van der Linden[f], Hai-yen Man[f], Aldert H. Piersma[b,c], Emiel Rorije[c], Sjors H.W. Schulpen[b], Michael Schwarz[d], Frederik Uibel[d], Barbara M.A. van Vugt-Lussenburg[f], Andre P.M. Wolterbeek[a], Bart van der Burg[f]

[a] TNO, Zeist, The Netherlands
[b] Institute for Risk Assessment Sciences, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands
[c] Center for Health Protection, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
[d] Institute of Experimental and Clinical Pharmacology and Toxicology, Department of Toxicology, University of Tübingen, Tübingen, Germany
[e] Fraunhofer Institute for Toxicology and Experimental Medicine (ITEM), Hannover, Germany
[f] BioDetection Systems BV, Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 1 May 2014
Received in revised form
17 September 2014
Accepted 1 October 2014
Available online 14 October 2014

Keywords:
Alternative model
Developmental toxicity
Read across
Grouping
CALUX
Embryonic stem cell test
Zebrafish embryotoxicity assay
ReProGlo

ABSTRACT

Previously we showed a battery consisting of CALUX transcriptional activation assays, the ReProGlo assay, and the embryonic stem cell test, and zebrafish embryotoxicity assay as 'apical' tests to correctly predict developmental toxicity for 11 out of 12 compounds, and to explain the one false negative [7]. Here we report on applying this battery within the context of grouping and read across, put forward as a potential tool to fill data gaps and avoid animal testing, to distinguish *in vivo* non- or weak developmental toxicants from potent developmental toxicants within groups of structural analogs. The battery correctly distinguished 2-methylhexanoic acid, monomethyl phthalate, and monobutyltin trichloride as non- or weak developmental toxicants from structurally related developmental toxicants valproic acid, mono-ethylhexyl phthalate, and tributyltin chloride, respectively, and, therefore, holds promise as a biological verification model in grouping and read across approaches. The relevance of toxicokinetic information is indicated.

1. Introduction

Testing strategies in toxicology are constantly being challenged, especially in the past decade when important regulatory changes occurred with the introduction of the EU legislation REACH [1]. This regulation enforces hazard and risk assessment of over 30 thousand preregistered chemicals before 2018. It is foreseen that this will require great numbers of test animals, especially for reproductive and developmental toxicity testing [2]. Therefore, innovative strategies that replace and/or reduce animal testing are urgently needed. Such strategies are also stimulated by REACH that prescribes animal testing only as a last resort [1].

Substantial efforts have already been undertaken to develop alternative methods for the assessment of reproductive and development toxicity [3,4]. However, none of these assays alone can cover the whole mammalian reproductive cycle due to its inherent complexity [3,4]. Therefore, recent studies have attempted to combine several *in vitro* assays into a test battery instead of applying individual assays. A feasibility study published by Schenk et al. [5] studied 10 compounds in a battery of 14 assays [5]. This battery was able to detect all reproductive toxicants for which the modes of action were actually represented in at least one of the assays. Another example is represented by the US EPA's ToxCast program, which used a large group of high throughput alternative assays to analyze the developmental toxicity of some 300 chemicals [6]. They showed >70% balanced accuracy of detecting developmental toxicants across some 650 assays reduced to a multivariate signature for chemicals identified as developmental toxicants in rat or rabbit guideline studies. This relatively low prediction potential might be

* Corresponding author at: TNO, PO Box 360, 3700 AJ, Zeist, The Netherlands. Tel.: +31 088 866 1508.
E-mail address: dinant.kroese@tno.nl (E.D. Kroese).

http://dx.doi.org/10.1016/j.reprotox.2014.10.003
0890-6238/

due to the fact that most assays applied were not intended to focus on developmental toxicity, and had only limited representation of assays for morphogenesis and differentiation. In a previous study of the European Framewok Programme 7 project ChemScreen, therefore, the zebrafish embryo test (ZET) and the embryonic stem cell test (EST) were included as more apical assays [7], to detect effects on development of a whole egg from fertilization until the hatching stage 72 h later [8], and to detect effects on cellular differentiation of cardiomyocytes [9], respectively. Also the medium throughput screen ReProGlo, monitoring interference with the WNT pathway [10], and assays for CYP17 and CYP19, to detect effects on steroidogenesis, enzymes essential for reproductive hormone homeostasis [11,12] were included. Finally, a panel of 24 high throughput CALUX assays were added that measure changes in activity of key transcription factors, varying from nuclear receptors, including reproductive hormone receptors, to transcription factors involved in cellular signaling [13,14]. The approach also encompassed toxicokinetic modeling to reveal whether effective *in vitro* concentrations observed in the battery are in the range expected from the *in vivo* reproductive toxicity data, in consideration of suggestions by Daston et al. [15]. As previously reported [7], this battery approach (including the toxicokinetic model) successfully identified eleven out of twelve compounds with varying mechanisms of action, with the unidentified compound, glufosinate ammonium, having a mechanism not covered by the battery. This result encourages to further optimize this battery into one ultimately able to detect all reprotoxic compounds.

The use and interpretation of battery results very much depends on the purpose of testing and the information that may already be available. In the absence of any *in vivo* test information relevant to potential reproductive toxicity and/or in the absence of any structural alerts pointing to such effects, the battery could be applied as a filter optimizing and/or reducing the testing of potential reproductive toxicants in animal studies [16]. The battery could also help by prioritizing chemicals for further investigation and/or by selecting candidate (pharmaceutical) compounds for further development [17]. On the other hand, when there are clear indications for potential reproductive toxicity based on close structural similarity of a query chemical to a reproductive toxicant, the battery could be used to confirm any reproductive toxicity, and to avoid any further *in vivo* studies. Ideally, this battery should also be capable to correctly distinguish reproductive toxicants from non-reproductive toxicants, even when both have structural similarity. The purpose of this investigation was to explore this discriminating capability of the battery *via* testing three groups consisting of structurally related chemicals, differing in their reproductive toxicity. Therefore, as a follow-up of our previous paper, we have tested two valproic acid (VPA) analogs: 2-ethylhexanoic acid (EHA), and 2-methylhexanoic acid (MHA), two analogs of monoethylhexyl phthalate (MEHP): monobenzyl phthalate (MBzP), and monomethyl phthalate MMP), and three organotin analogs of dioctyltin dichloride: tributyltin chloride, dibutyltin dichloride (TBTC), dibutyltin chloride dibutyltin dichloride (DBTC) and monobutyltin trichloride (MBTC). Of these analogs to be tested, MHA, and MMP were not considered reproductive toxicants, while MBTC is considered a weak reproductive toxicant, if at all. The outcome of this investigation will be discussed within the context of the purpose of the battery within an integrated testing strategy.

2. Materials and methods

2.1. General experimental set-up

As described in our previous manuscript [7], tests were selected at a special meeting of the ChemScreen consortium, based on the following criteria: (1) Relevance to known mechanisms and endpoints involved in reproductive toxicity. (2) Availability at partner institutes. (3) Overlap with ReProTect feasibility study [5], to allow comparison, and 4. with preference to higher throughput assays that can be automated, but also including more apical lower throughput assays. The battery consisted of the cardiac embryonic stem cell test (EST), the zebrafish embryotoxicity test (ZET), the ReProGlo assay (ReProGlo), and a panel of CALUX assays (see Table 2). The individual tests were performed, according to previously standardized protocols established in the various collaborating centers as further detailed below.

2.1.1. The embryonic stem cell test (EST)

Pluripotent mouse D3 embryonic stem cells (ESC; ATCC, Rockville, MD) were routinely subcultured every 2–3 days and grown as a monolayer in complete medium, consisting of DMEM (Gibco BRL, Gaithersburg, MD) supplemented with 20% fetal bovine serum (Hyclone, Logan, UT), 1% non-essential amino acids (Gibco BRL, Gaithersburg, MD), 1% penicillin/streptomycin (Gibco BRL, Gaithersburg, MD), 2 mM L-glutamine (Gibco BRL, Gaithersburg, MD) and 0.1 mM β-mercapto-ethanol (Sigma–Aldrich, Zwijndrecht, The Netherlands). Leukemia inhibitory factor (LIF; Chemicon, Temecula, CA) was added directly to the culture disk in a final concentration of 1000 units/ml. The cells were maintained in a humidified atmosphere at 37 °C and 5% CO_2. Embryoid body (EB) formation was used as the initial step for the EST differentiation assay. EB were obtained *via* hanging drop culture in complete medium without LIF [18]. In brief, stem cell suspensions (3.75×10^4 cells/ml) were placed on ice before the set up of the culture. Drops (20 μl) containing 750 cells were placed onto the inner side of the lid of a Petri dish filled with phosphate buffered saline (PBS; Gibco BRL, Gaithersburg, MD) and incubated at 37 °C, 90% relative humidity and 5% CO_2. After 3 days of hanging drop culture EB had formed and these were subsequently transferred to bacterial Petri dishes (Greiner Bio-one, Frickenhausen, Germany). On day 5, 24 EB were plated one per well into 24 well tissue culture plates (TPP, Trasadingen, Switzerland). Differentiation was determined microscopically at day 10 of differentiation by inspection of EB outgrowths into contracting myocardial cells. EB were considered as cardiomyocyte positive if at least one contracting focus was present. The number of positive EB was expressed as fraction of total EB examined. Data are expressed as follows: 'cytotox' values are IC50 (concentration inhibiting cell viability by 50%), and 'diff010' values are ID50 (concentration inhibiting differentiation by 50%), respectively, after 10 days of exposure in the assay. The test was considered positive if ID50 ≤ IC50, or if IC50 ≤ 100 μM ('≥4' in the table). If kinetic data were available for a compound, these cut-off values could be shifted up or downwards according to the calculated relevant *in vivo* plasma concentrations.

2.1.2. The zebrafish embryotoxicity test (ZET)

The ZET was performed as described previously [19]. In brief, fertilized batches of eggs with a fertilization rate of at least 90% were collected within 30–60 min after spawning using 800 μm mesh and were rinsed with MilliQ water to remove impurities. The fertilized eggs were then directly transferred into different Petri dishes containing the test compounds at selected concentrations. Subsequently, embryos at the 8–16 cells stage were selected and transferred to a 12-well plate containing 3 mL of the test medium per well. Embryos were kept in an incubator at 28.0 ± 1 °C with a photoperiod of 12 h light:12 h dark. Morphological changes were evaluated at 24, 48, 72 and 96 hpf, as described recently [19]. Experiments were considered valid if survival rates in controls were >90% [20]. The morphological scores for each experiment were normalized and expressed as a percentage compared to controls, combined and presented as mean ± standard error of the mean (SEM). The