Integrated *in vitro-in silico* models for predicting *in vivo* developmental toxicity

facilitating non-animal based safety assessment

Jochem Louisse

Thesis committee

Thesis supervisors

Prof. dr. ir. I.M.C.M. Rietjens
Professor of Toxicology
Wageningen University

Prof. dr. B.J. Blaauboer
Doerenkamp-Zbinden Chair on Alternatives to Animal Testing in Toxicological Risk Assessment
Institute for Risk Assessment Sciences, Utrecht University

Thesis co-supervisor

Dr. M. Verwei
Senior Scientist Pharmacokinetics
TNO, Zeist

Other members

Prof. dr. ir. J. Keijer, Wageningen University
Prof. dr. M. van den Berg, Utrecht University
Prof. dr. C.F.M. Hendriksen, Utrecht University
Ir. D.H. Waalkens-Berendsen, TNO Triskelion

This research was conducted under auspices of the Graduate School VLAG (Advanced studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences).

Integrated *in vitro-in silico* models for predicting *in vivo* developmental toxicity

facilitating non-animal based safety assessment

Jochem Louisse

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 25 May 2012
at 1.30 p.m. in the Aula.

Jochem Louisse

Integrated *in vitro-in silico* models for predicting *in vivo* developmental toxicity; facilitating non-animal based safety assessment

256 pages

Thesis Wageningen University, Wageningen, NL (2012)
With references, with summaries in Dutch and English

ISBN 978-94-6173-241-5

Propositions

1. Physiologically based kinetic modeling is essential for the implementation of *in vitro* toxicity assays as alternatives for *in vivo* toxicity testing in risk assessment practice (this thesis).
2. Differences in protein binding cannot be ignored, when extrapolating *in vitro* toxicity data to the *in vivo* situation (this thesis).
3. Toxicologists unjustly ignore Paracelsus's concept of dose, when dividing chemicals in toxicants and non-toxicants.
4. Effects of chemicals on the epigenome are currently underestimated in studies on toxicological modes of action.
5. The increasing influence of public opinion on risk management results in decisions that are no longer science based.
6. As the postulation of propositions is an attainment target for the PhD degree at Wageningen University, graduate schools should offer a related course.

Propositions belonging to the thesis, entitled

"Integrated *in vitro-in silico* models for predicting *in vivo* developmental toxicity; facilitating non-animal based safety assessment."

Jochem Louisse

Wageningen, 25 May 2012.

TABLE OF CONTENTS

CHAPTER 1

General Introduction

Chapter 1

General introduction and aim of the thesis

In the last decade, the European Commission implemented the REACH (Registration, Evaluation, Authorization and restriction of CHemicals) legislation, aiming at greater protection of human health and the environment towards the adverse effects of industrial chemicals. This REACH legislation aims at the harmonization of the testing requirements for so-called 'existing' and 'new' chemicals, the latter being the chemicals that have been marketed after a change in legislation in 1981 (Grindon and Combes, 2006). Since for many of these chemicals toxicity data are lacking, the implementation of the REACH legislation will require the generation of toxicity data for a large number of chemicals. Currently, these toxicity data originate from studies performed with laboratory animals, although the REACH legislation stresses the importance of the use of alternative testing strategies. It has been estimated that the number of laboratory animals required for REACH would amount to 3.9 million if the use of alternative methods would not be accepted by regulatory authorities (Van der Jagt *et al.*, 2004). In addition, it was estimated that with the expected acceptance scenario for alternative approaches, the total number would amount to 2.6 million. Furthermore, it was calculated that the majority of the animals will be used for the endpoints developmental toxicity (26% of total) and reproduction toxicity (37% of total) (Van der Jagt *et al.*, 2004). The developmental and reproductive toxicity testing requirements for chemicals under REACH depend on the chemical tonnage levels manufactured or imported in the European Union. At tonnage levels of ≥10 metric tons per year, toxicity testing according to OECD (Organisation for Economic Co-operation and Development) test guideline 421, describing a reproductive/developmental toxicity screening test or 422, describing a combined repeated dose toxicity study with reproductive/ developmental component, should be performed (Scialli, 2008). For the higher tonnage levels (≥100 and ≥1000 metric tons per year), data from toxicity studies according to OECD test guideline 414, describing a prenatal developmental toxicity test, in one or two species are required. Furthermore, a two-generation reproductive toxicity study (not identified by OECD test guideline number in REACH, but corresponding to OECD test guideline 416) in one species is required at these high tonnage levels. For chemicals manufactured or imported at tonnage levels of ≥100 metric tons per year, this study is only required if a 28-day or 90-day

study indicates adverse effects on reproductive organs (Scialli, 2008). According to the estimations of Höfer *et al.* (2004) the total numbers of animals needed for REACH would be more than the double amount estimated by Van der Jagt *et al.* (2004) when taking the offspring into account that is produced during these developmental toxicity studies. Not only for REACH chemicals, but also for pharmaceuticals, food additives and pesticides, regulatory agencies require safety studies performed in laboratory animals, including developmental toxicity studies. Therefore, the development, validation and application of reliable alternative assays for *in vivo* developmental toxicity studies are urgently needed in the safety assessments of REACH chemicals, pharmaceuticals, food additives and pesticides, contributing to the replacement, reduction and refinement (3Rs) of animal use in toxicological research.

Several *in vitro* assays have been reported for developmental toxicity, which may be useful in alternative testing strategies for *in vivo* developmental toxicity studies currently carried out in rats and rabbits. In these assays, isolated cells, tissues or whole embryos are cultured and exposed to the chemical of interest at increasing concentrations, after which the chemical-induced effects are assessed. The effects can be diverse, varying for example from gene or protein expression changes to morphological changes. The *in vitro* developmental toxicity assays reported in the literature are described in more detail in chapter 2 of the present thesis. Based on the outcomes of the *in vitro* assays, it is estimated whether developmental toxicity will be expected to occur *in vivo*.

The translation of *in vitro* toxicity data into relevant data for the *in vivo* situation is not straightforward. One of the differences between *in vitro* and *in vivo* toxicity assays, is that in *in vitro* assays, concentration-response curves are obtained, whereas in *in vivo* assays, dose-response curves are obtained. These dose-response curves are needed in toxicological risk assessment, to set a point of departure to derive safe exposure levels for humans. Therefore, in order to use *in vitro* toxicity data for risk assessment, concentration-response curves obtained *in vitro* need to be translated into *in vivo* dose-response curves (Blaauboer, 2010). This translation of *in vitro* concentration-response curves into *in vivo* dose-response curves is at present an important bottleneck and challenge for the use of *in vitro* toxicity assays in risk assessment practice. In theory, the translation of *in vitro* toxicity

concentrations into *in vivo* doses may be achieved using physiologically based kinetic (PBK) modeling using a reverse dosimetry approach (DeJongh *et al.*, 1999; Verwei *et al.*, 2006; Forsby and Blaauboer, 2007).

A PBK model is a set of mathematical equations that together describe the absorption, distribution, metabolism and excretion (ADME) characteristics of a chemical within an organism on the basis of physiological and anatomical parameters, physico-chemical parameters and kinetic parameters (Krishnan and Andersen, 2001). PBK models describe the fate of the chemical (and its metabolite (s)) in the body in time after a certain dose of the chemical via a certain exposure route (e.g. oral or dermal) in a certain species (e.g. rat or human). PBK models can be used to translate external dose levels of a chemical into internal blood or tissue concentrations of the chemical (and its metabolite(s)). Furthermore, PBK models can also be used the other way around, called reverse dosimetry, by translating internal blood or tissue concentrations of a chemical (and/or its metabolite(s)) into external dose levels of the chemical. Reverse dosimetry is illustrated in Figure 1 showing a hypothetical example of this approach. In this example a concentration-response curve obtained in an *in vitro* toxicity assay is set as internal concentrations in the PBK model (in blood in this case), followed by the calculation of the doses needed to reach these internal concentrations using the PBK model, resulting in a predicted dose-response curve for a toxicological endpoint.

The reverse dosimetry approach has been used to translate a single *in vitro* effect concentration (e.g. an EC_{50} value) into a single *in vivo* effect dose (e.g. a lowest observed adverse effect level (LOAEL)) (DeJongh *et al.*, 1999; Verwei *et al.*, 2006; Forsby and Blaauboer, 2007). However, the *in vitro* effect concentrations and the *in vivo* effect doses were chosen arbitrarily in previous reverse dosimetry work. In theory, a complete concentration-response curve obtained in an *in vitro* assay could be translated into a complete predicted *in vivo* dose-response curve, by translating all data obtained in the *in vitro* experiment to the *in vivo* situation. This would enable dose-response modeling on the predicted *in vivo* dose-response data, defining a predicted *in vivo* dose-response curve, which could be used in the risk assessment process to set a point of departure to derive safe exposure levels.

The aim of the present thesis is to provide the proof-of-principle that *in vitro*

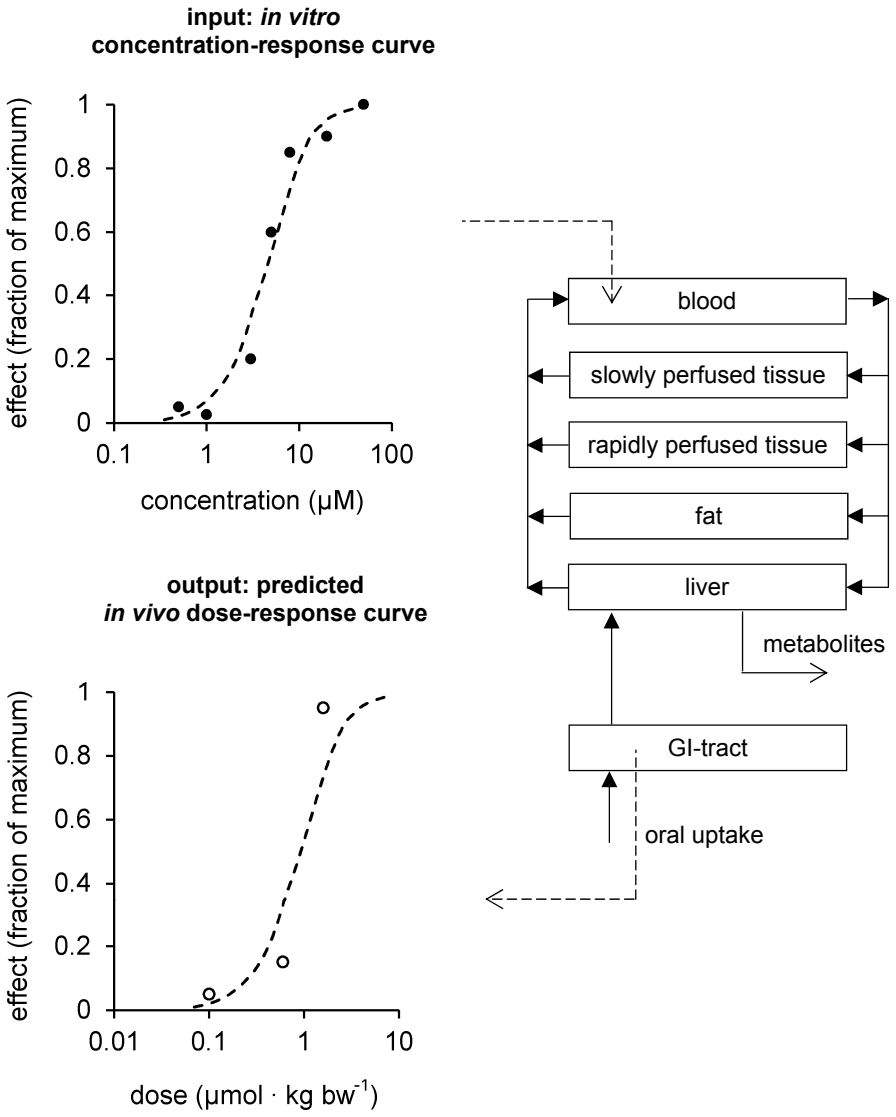

Figure 1. Example of the reverse dosimetry approach in which an *in vivo* dose-response curve is predicted for a certain (toxic) endpoint by setting the concentrations of an *in vitro* concentration-response curve as internal blood concentrations, followed by the calculation of the doses needed to reach these internal concentrations using the PBK model. The closed symbols represent the data obtained in the *in vitro* assay (hypothetical data), whereas the open symbols represent data from *in vivo* toxicity studies (hypothetical data), which are used to evaluate whether the approach reliably predicts *in vivo* dose-response curves.

concentration-response curves can be converted into *in vivo* dose-response curves suitable for risk assessment practice, using PBK modeling as a tool to apply reverse dosimetry. The toxicological endpoint for which this proof-of-principle will be developed is developmental toxicity, because, as outlined above, this type of toxicity studies is estimated to require more than one fourth of the animals expected to be needed under the REACH legislation. This indicates that reliable alternatives for this endpoint would contribute significantly to the 3Rs of animal use. The following sections introduce three important factors needed in this study, which include 1) an *in vitro* model to obtain *in vitro* concentration-response curves for a developmental toxicity endpoint, 2) the selection of chemicals to be studied and 3) the development of the PBK models that accurately describe the *in vivo* kinetics, which are needed for reverse dosimetry.

In vitro model to define in vitro concentration-response curves for developmental toxicity

An *in vitro* model is needed that can be used to determine relevant *in vitro* concentration-response curves to be translated into *in vivo* dose-response curves for developmental toxicity. The *in vitro* developmental toxicity assays including their readouts reported in the literature are described in more detail in chapter 2 of the present thesis. The *in vitro* assay that was chosen for the studies in the present thesis is the embryonic stem cell (ES-D3 cell) differentiation assay of the embryonic stem cell test (EST). The EST is the only validated *in vitro* developmental toxicity assay that does not require primary animal tissues (Genschow *et al.*, 2002; Genschow *et al.*, 2004). The implementation of this assay, when proven to provide reliable outcomes in an alternative testing strategy, would therefore result in the largest decrease of animal use. Further information on the EST is presented in chapter 2.

Selection of the chemicals to be studied

The chemicals to be studied in this proof-of-principle study needed to fulfill two criteria. These criteria were 1) the presence of *in vivo* kinetic data on the chemical in the literature, in order to be able to build and/or evaluate the PBK models developed and used for reverse dosimetry and 2) the presence of *in vivo* developmental toxicity data of the chemical in the literature, in order to evaluate

whether the ultimately predicted dose-response curves for developmental toxicity are in concordance with data obtained in *in vivo* developmental toxicity studies.

To select chemicals to be studied in the present thesis, chemicals reported to cause *in vivo* developmental toxicity were assessed for the two above mentioned criteria. These chemicals were obtained from the following sources: 1) chemicals already used in validation studies of *in vitro* developmental toxicity assays of the European Centre for the Validation of Alternative Methods (ECVAM) (Genschow *et al.*, 2002; Genschow *et al.*, 2004; Piersma *et al.*, 2004; Spielmann *et al.*, 2004), 2) a list of chemicals developed to be used for such validation studies (Brown *et al.*, 2002), 3) a list of chemicals developed for the European Framework Program project ReProTect which aimed at the development of *in vitro* test systems for reproduction toxicity and developmental toxicity (Pazos *et al.*, 2010) and 4) the 'Niet-limitatieve lijst reprotoxische stoffen' ('Non-restrictive list of reproductive toxicants') of the Dutch Ministry of Social Affairs and Employment (Staatscourant, 2001). The chemicals of these four sources fulfilling the two criteria in order to be possibly included in the present thesis are listed in Table 1. When possible, the chemicals were grouped in chemical categories sharing structural similarities. From this list of chemicals that were appropriate for this proof-of-principle study, only series of structurally related chemicals were selected. This third criterion enabled the assessment of relative differences within a series of chemicals, thereby providing an additional possibility for qualitative evaluation of the predictions made. Two series of model compounds were selected, being 1) a group of glycol ethers (Figure 2), which need to be bioactivated in order to cause developmental toxicity, which are hydrophilic, show low protein binding and have an unspecific mode of action and 2) a group of retinoids (Figure 3), containing several members that directly act as teratogens without bioactivation, which are hydrophobic, show high protein binding and have a specific mode of action.

Glycol ethers

Glycol mono-ethers contain both an ether and an alcohol group, making them widely applicable as solvents in cosmetics, printing inks, plastics, household and industrial cleaning products and textile dyes. Non-solvent applications include their use as anti-icing agents in jet fuel and as hydraulic system fluids. Most glycol ethers are of low acute toxicity. Glycol ethers are neither genotoxic nor

Table 1. Selection of chemicals that are appropriate for the proof-of-principle study fulfilling the two criteria defined, being 1) the presence of *in vivo* kinetic data on the chemical in the literature, in order to be able to build and/or evaluate the PBK models developed and used for reverse dosimetry and 2) the presence of *in vivo* developmental toxicity data of the chemical in the literature, in order to evaluate whether the ultimately predicted dose-response curves for developmental toxicity are in concordance with data obtained in *in vivo* developmental toxicity studies.

Chemical	Cas no.	Chemical	Cas no.
aromatic solvens		hydroxyurea	127-07-1
benzene	71-43-2	lovastatin	75330-75-5
ethylbenzene	100-41-4	methyl ethyl ketone	78-93-3
toluene	108-88-3	methylmercury	115-09-3
xylene	1330-20-7	mirex	2385-85-5
		nitrofen	1836-78-5
boric acid	10043-35-3	ochratoxin A	303-47-9
busulfan	55-98-1		
bromodichloromethane	75-27-4	**retinoids**	
carbon disulfide	75-15-0	retinol	68-26-8
carbon tetrachloride	56-23-5	all-*trans* -retinoic acid	302-79-4
cinnamic aldehyde	104-55-2	13-*cis* -retinoic acid	4759-48-2
dimethylacetamide	127-19-5	etretinate	54350-48-0
dimethylformamide	68-12-2	acitretin	55079-83-9
D-penicillamine	52-67-5		
ethanol	64-17-5	sodium arsenate	10048-95-0
ethylene thiourea	96-45-7	2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD)	1746-01-6
ethylhexanoic acid	149-57-5	theophylline	58-55-9
fumonisin B1	116355-83-0	valproic acid	99-66-1
		warfarin	81-81-2
glycol ethers			
ethylene glycol	107-21-1		
ethylene glycol monomethyl ether	109-86-4		
ethylene glycol monoethyl ether	110-80-5		
ethylene glycol monobutyl ether	111-76-2		
ethylene glycol monophenyl ether	122-99-6		

carcinogenic (ECETOC, 2005). The short chain ethylene based glycol ethers ethylene glycol monomethyl ether (EGME) and ethylene glycol monoethyl ether (EGEE) induce systemic toxicity after biotransformation to their alkoxyacetic acid metabolites, methoxyacetic acid (MAA) and ethoxyacetic acid (EAA) respectively (ECETOC, 2005; Brown et al., 1984; Giavini et al., 1993). MAA and EAA have been shown to induce bone marrow depression, testicular atrophy, immunotoxicity and

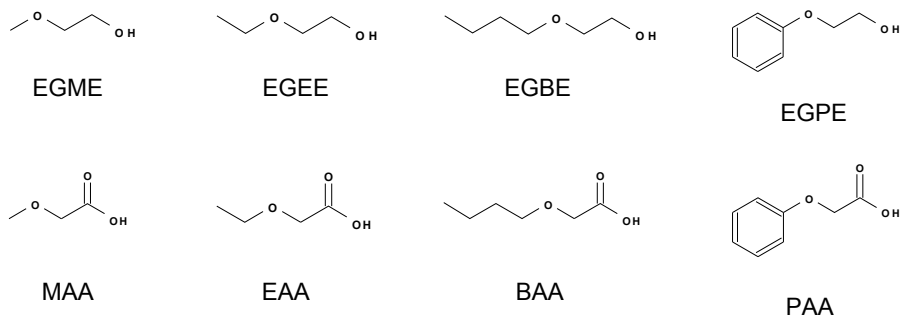

Figure 2. Chemical structures of glycol ethers and their alkoxyacetic acid metabolites studied in the present thesis. EGME: ethylene glycol monomethyl ether, EGEE: ethylene glycol monoethyl ether, EGBE: ethylene glycol monobutyl ether, EGPE: ethylene glycol monophenyl ether, MAA: methoxyacetic acid, EAA: ethoxyacetic acid, BAA: butoxyacetic acid, PAA: phenoxyacetic acid.

developmental toxicity in animals (ECETOC, 2005; Feuston et al., 1990; Hanley et al., 1984; Nagano et al., 1981; Wier et al., 1987). The longer chain ethylene based glycol ethers ethylene glycol monobutyl ether (EGBE) and ethylene glycol monophenyl ether (EGPE) cause red blood cell haemolysis in rat, induced by their alkoxyacetic acid metabolites butoxyacetic acid (BAA) and phenoxyacetic acid (PAA), respectively (ECETOC, 2005).

Retinoids

Chemicals that possess a chemical structure or have functional properties similar to retinol (vitamin A) are called retinoids (Figure 3). Retinol is required for physiological processes such as vision, reproduction, growth, cell differentiation, immune function and embryonic development (Collins and Mao, 1999). Most of the biological activities of the retinoids are believed to be mediated by various nuclear receptors of the steroid, thyroid, vitamin D and retinoid superfamily of receptors (Collins and Mao, 1999). The most studied retinoid receptors include two types of ligand-dependent transcription factors, the retinoic acid receptors (RARs) and the retinoid X receptors (RXRs), existing in different subtypes and isoforms. In general, the activity of retinoids is mediated by RAR-RXR heterodimers. For a proper physiological function, people require a certain amount of retinol, taken up via the diet. Retinol deficiency can lead to xerophthalmia, immunodeficiency and weight loss, whereas excessive retinol levels cause toxicity to the central nervous

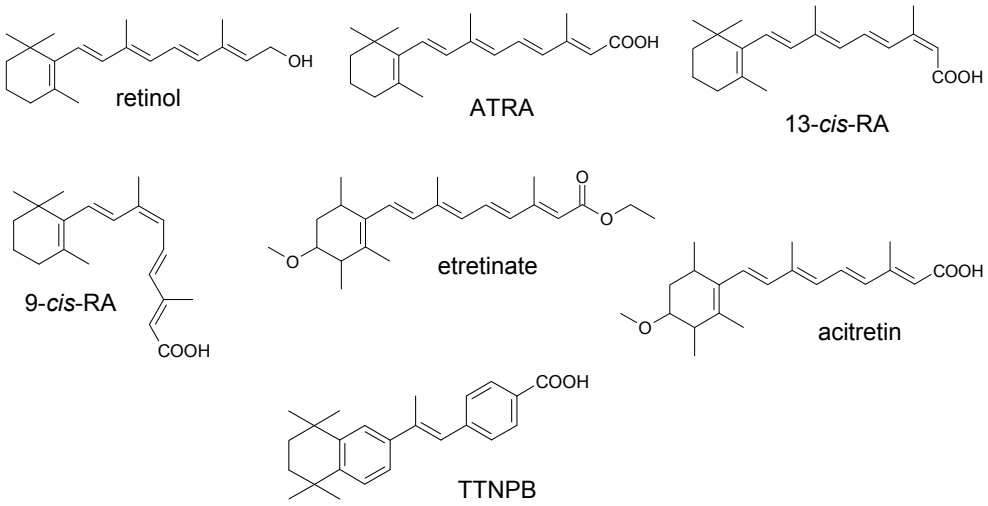

Figure 3. Chemical structures of retinoids studied in the present thesis. ATRA: all-*trans*-retinoic acid, 13-*cis*-RA: 13-*cis*-retinoic acid, 9-*cis*-RA: 9-*cis*-retinoic acid, TTNPB: (E)-4[2-(5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2-naphthalenyl)-1-propenyl]benzoic acid.

system, liver, bone and skin (Collins and Mao, 1999). Proper levels of vitamin A must be maintained for normal embryogenesis. Vitamin A deficiency and high retinoid intake induce adverse effects to the developing conceptus (Collins and Mao, 1999). People may also be exposed to retinoids via medication. The natural retinoids all-*trans*-retinoic acid (ATRA), 13-*cis*-retinoic acid (13-*cis*-RA) and 9-*cis*-retinoic acid (9-*cis*-RA) are used against acute promyolocytic leukemia (Tallman *et al.*, 1997), cystic acne (Kanigsberg and Desgroseilliers, 1983) and Kaposi sarcoma (Washenik *et al.*, 2000) respectively, whereas the synthetic retinoids etretinate and acitretin are used against psoriasis (Gollnick *et al.*, 1988). These retinoids have been shown to be teratogenic in a number of experimental animals, including mice, rats, guinea pigs, hamsters, rabbits, dogs, pigs, chickens and monkeys (Collins and Mao, 1999).

Development of PBK models

In order to translate the *in vitro* concentration-response curves obtained in the ES-D3 cell differentiation assay into *in vivo* dose-response curves with reverse dosimetry, PBK models describing the ADME processes of the chemical of interest

need to be developed. In a tutorial on PBK modeling in molecular nutrition and food research, the process of PBK model development has recently been explained (Rietjens *et al.*, 2011). The development of a PBK model can be divided in the following steps: 1) definition of a conceptual model, 2) translation of the conceptual model into a mathematical model by defining mathematical equations, 3) obtaining model parameter values, 4) solving the mathematical equations using appropriate software and 5) evaluation of model performance. Subsequently the PBK model can be applied to make predictions (step 6), such as predictions of interspecies differences in kinetics, enabling extrapolation of animal data to humans. The PBK model can also be used for reverse dosimetry, translating *in vitro* concentrations into *in vivo* doses, as was the application of PBK modeling in the present thesis. The six steps are described in more detail in the following sections.

Step 1: Defining a conceptual model: the basic structure of a PBK model

The basic structure of a PBK model assumes that an intact organism can be described as a set of basic compartments. In a PBK model these basic compartments can each be depicted as a single box. In the standard PBK modeling approach, it is assumed, as a generally accepted simplification, that a compartment is a single region of the body with a uniform chemical concentration. However, models using non-homogenous chemical distribution in compartments exist as well. These compartments may be a particular functional or anatomical portion of an organ, a single blood vessel with surrounding tissue, an entire discrete organ such as the liver or kidney, or a widely distributed tissue type such as fat or skin. As a first step in developing a chemical-specific PBK model one has to decide which organs are relevant for the ADME characteristics of that chemical and its biological effects. These compartments need to be specifically described, whereas all other organs can be lumped together in either a slowly or richly perfused tissue compartment. A reason to include a specific organ can be that this organ represents the site of uptake or elimination, a major site of metabolism or a target organ for the beneficial or adverse effect of the chemical of interest. However, if no or little data on the *in vivo* kinetics of the chemical itself or a structurally related chemical is available, one cannot choose beforehand which organs will be relevant for the *in vivo* AMDE characteristics. In that case, a generic PBK model may be used, which describes the kinetic processes in all organs of the body using default

descriptions of compartments and kinetic processes. To make these generic models chemical-specific, kinetic parameter values for the chemical of interest need to be determined (step 3).

Step 2: Translation of the conceptual model into a mathematical model by defining mathematical equations

The second step in PBK modeling is the translation of the conceptual model into a mathematical model, describing the kinetic processes in each compartment by differential equations. As an example, the kinetic processes of a certain chemical in the liver can be described as:

$$\begin{aligned} dAL/dt = \; & k_{uptake} * AGI(t) \\ & + QL * (CB - CL/PL) \\ & - V_{max} * CL/PL / (K_m + CL/PL) \end{aligned}$$

This equation describes that the change in the amount of the chemical (A) in the liver (L) in time (t), represented as dAL/dt, is a function of 1) the uptake of the chemical from the gastrointestinal tract directly into the liver, 2) the exchange of the chemical between blood and the liver and 3) the conversion of the chemical into its metabolite(s) due to biotransformation. The uptake from the gastrointestinal tract can be described by a first-order process, assuming passive absorption from the gastrointestinal tract to the liver compartment. This implies that uptake can be described by $k_{uptake} * AGI(t)$, with k_{uptake} (in h^{-1}) representing the first-order absorption rate constant and AGI(t) (in mol) being the amount of chemical remaining at a certain time point in the gastrointestinal tract, with AGI at time = 0 being equal to the oral dose. The exchange of the chemical between liver tissue and blood, equals the blood flow rate to the liver (QL in L/h) times the difference in chemical concentration between the blood entering the liver (CB) and the blood leaving the liver (CVL); thus QL(CB - CVL), assuming a perfusion-limited and not a diffusion-limited process (Krishnan and Andersen, 2001). This implies that the exchange of the chemical between liver tissue and blood are dominated by blood flow (perfusion) of the liver and not by diffusion. This equation is based on Fick's law of simple diffusion (dCt/dt = k*dC), which states that the flux of a chemical is proportional to its concentration gradient (Krishnan and Andersen, 2001). Assuming equilibrium between the chemical concentration in the blood

leaving the liver (CVL) and the blood concentration of the chemical in the liver (not bound to liver tissue), CVL in the equation equals CL/PL, with CL representing the total chemical concentration in the liver (bound and not bound to liver tissue; in mol/L) and PL representing the liver:blood partition coefficient of the chemical (Krewski *et al.*, 1994; Krishnan and Andersen, 2001; Clewell and Clewell, 2008). The metabolic conversion of the chemical in the liver can be modeled by a Michaelis-Menten equation leading to a contribution equal to - V_{max} * CL/PL / (K_m + CL/PL) with the negative sign reflecting that this contribution results in a decrease in the amount of the chemical in the liver and V_{max} and K_m representing the maximum rate and Michaelis–Menten constant for metabolism, respectively. The Michaelis–Menten equation expresses the concentration of the chemical in the liver as CL/PL, corresponding to the concentration of the chemical in the liver that is not bound to liver tissue, being available for metabolism.

If an available generic PBK model is used as starting-point for the PBK model development for the chemical(s) of interest, the mathematical equations describing the kinetic processes in the compartments do not have to be defined (step 2). In that case, one can immediately start with obtaining the model parameter values (step 3).

Step 3: Obtaining the model parameter values

In step 2, based on the basic structure of the PBK model, a set of mathematical equations is defined, one set for each compartment. For the further development of a PBK model for the chemical(s) of interest, in the next step model parameter values need to be obtained. The required model parameter values include values for 1) physiological and anatomical parameters (e.g. cardiac output, tissue volumes and tissue blood flows), 2) physicochemical parameters (e.g. tissue:blood partition coefficients) and 3) other kinetic parameters (e.g. kinetic constants for transport as well as for biotransformation reactions of the chemical and/or its metabolites) (Krewski *et al.*, 1994; Krishnan and Andersen, 2001; Clewell and Clewell, 2008). The values for physiological and anatomical parameters of a species can be obtained from the literature. Representative reference physiological parameter values for different laboratory animals as well as humans can, for instance, be obtained from Brown *et al.* (1997). Physicochemical parameter values (i.e. partition coefficients, needed to describe the distribution of the chemical over the body compartments)

may be obtained using *in vivo* data on chemical concentrations in blood and tissues of animals exposed to the chemical, or by using *in vitro* techniques (Gargas *et al.*, 1989; Artola-Garicano *et al.*, 2000). Also several *in silico* tools are described in the literature to estimate partition coefficients, based on physicochemical properties such as the chemical's LogP value (DeJongh *et al.*, 1997; Poulin and Theil, 2002; Berezhkovskiy, 2004). Values for kinetic parameters to describe absorption, metabolism and excretion are often estimated using *in vivo* kinetic data, by choosing these parameter values in such a way that the PBK model predictions fit the available *in vivo* kinetic datasets. However, these parameter values may also be obtained using *in vitro* techniques, such as *in vitro* incubations with liver enzymes in order to determine parameter values for metabolism (Clewell, 1993; Punt *et al.*, 2009). To limit the use of animals in PBK model development, *in vitro* and *in silico* approaches are preferred to determine PBK model parameter values.

Step 4: Solving the mathematical equations using appropriate software

When the mathematical model has been defined and the parameter values have been obtained, the next step involves solving the mass-balance differential equations and making simulations. Different software packages are available for the numerical integration of the ordinary differential equations of the PBK model, such as Berkeley Madonna (Macey and Oster, UC Berkeley, CA, USA), acslXtreme (AEgis Technologies Group) and MATLAB (The MathWorks). Integration algorithms that can be applied are, for instance, the Gear algorithm or Rosenbrock algorithm, which are both capable of handling "stiff" sets of differential equations, which most PBK models are (Chiu *et al.* 2007).

Step 5: Evaluation of model performance

When the differential equations have been solved, thereby calculating the concentrations of the chemical and its metabolites in the different compartments over time, subsequently the evaluation of the model performance against experimental *in vivo* kinetic data needs to be performed. This is important to test the validity of the model and thus the reliability of subsequent predictions. Furthermore, discrepancies between the model predictions and the experimental data may be informative and may for example indicate that an important kinetic process has been overlooked or its model parameter value imprecisely incorporated

in the model. If such situations are encountered, the nature of the discrepancies may indicate a new hypothesis and improvements for the model. In case a PBK model is developed for a new chemical for which no or limited *in vivo* kinetic data are available, this evaluation cannot be performed. In that case, one should take into account that the model predictions may be less certain.

The evaluation step also includes a sensitivity analysis. This analysis provides a quantitative evaluation of how the parameters of the model influence the model output. A typical way of performing a sensitivity analysis is the central difference method, which compares the magnitude in model output due to a defined change in input parameters. In general, input parameters are changed with 1 or 5%, with each parameter being analyzed individually, keeping the other parameters to their initial values. This analysis yields normalized sensitivity coefficients that correspond to the ratio of change in simulation output (e.g. tissue concentration or formation of metabolites) to change in parameter value. A normalized sensitivity coefficient of, for instance, 0.8 signifies that 1% change in the numerical value of the input parameter will result in 0.8% change in model output. The greater the absolute value of the normalized sensitivity coefficient, the more influence the parameter has on the model output. The sign of the sensitivity ratio indicates whether the model output is directly or inversely related to the parameter. In a sensitivity analysis normalized sensitivity coefficients are generally calculated for all parameters at selected dose levels, but only parameters that significantly influence the model output, having a sensitivity coefficient of, for instance, higher than 0.1 in absolute value, are generally regarded as parameters that influence the model output to a significant extent (US EPA, 2006; Chiu *et al.*, 2007). Based on the results of the sensitivity analysis, requirements to the accuracy of the parameter values can be set.

Step 6: Application of PBK modeling in reverse dosimetry

Once the performance of the model has been evaluated, simulations and predictions can be made. In the present thesis, PBK models were developed for reverse dosimetry in order to predict *in vivo* dose-response curves for developmental toxicity by converting concentration-response curves from the *in vitro* ES-D3 cell differentiation assay into relevant *in vivo* dose-response scenarios. With reverse dosimetry, the PBK model is used to translate internal blood or tissue

concentrations of a chemical (and/or its metabolite(s)) into external dose levels of the chemical, enabling *in vitro-in vivo* extrapolations (Figure 1). With these *in vitro-in vivo* extrapolations using PBK modeling with reverse dosimetry, the following aspects should be considered: 1) the relation between the *in vitro* toxic effect concentration and the predicted *in vivo* toxic effect dose, 2) the determination of the dose metric used to relate exposure to toxicity and 3) differences in free fraction of the chemical in the *in vitro* assay compared with the *in vivo* situation.

1) The relation between the in vitro toxic effect concentration and the predicted in vivo toxic effect dose

With the translation of *in vitro* concentrations into *in vivo* doses, different approaches can be used to relate an *in vitro* toxic effect to the predicted *in vivo* toxic effect. In previous work on predicting *in vivo* developmental toxicity by translating *in vitro* developmental toxicity data into *in vivo* dose levels with reverse dosimetry, developmental toxicity dose levels were predicted by translating EC_{50} values of the ES-D3 cell differentiation assay to the *in vivo* situation (Verwei *et al.*, 2006). The obtained predictions for *in vivo* developmental toxicity were related to LOAELs of the chemicals for developmental toxicity, as reported in the literature (Verwei *et al.*, 2006). In the present thesis, not a single outcome (such as an EC_{50} value) of the *in vitro* data is translated to the *in vivo* situation, but complete *in vitro* concentration-response curves are translated into complete *in vivo* dose-response curves. This enables the application of benchmark dose (BMD) modeling in the assessment of the data, taking into account all data of the *in vitro* concentration-response curve. The resulting predicted BMD values can then be compared with BMD values obtained from data on *in vivo* developmental toxicity reported in the literature, in order to evaluate the predictive value of the approach.

The relation between the readout in the *in vitro* assay and the predicted *in vivo* effect should be considered as well. In order to be useful to predict *in vivo* dose-response curves, the readout parameter (which is the inhibition of the differentiation of embryonic stem cells into functional cardiomyocytes in the present thesis) should be relevant and predictive for the *in vivo* toxicological endpoint (developmental toxicity). It is not intended in the present thesis to specifically predict dose levels affecting cardiac development during embryonic development. The inhibition of cardiac differentiation is rather used to represent a

sensitive *in vitro* readout parameter for developmental toxicity (Genschow *et al.*, 2004), which in the ideal situation could represent the most sensitive *in vivo* developmental toxicity endpoint.

2) The determination of the dose metric used to relate exposure to toxicity

Toxic effects of chemicals are related to chemical exposure. Generally, it is assumed that the toxic outcome of a chemical is dependent on either the maximal concentration (C_{max}) to which the *in vitro* or *in vivo* system is exposed, or the area under the concentration-time curve (AUC) of the chemical. The toxic effect of glycol ethers can be related to the C_{max} of the toxic alkoxyacetic acid metabolites (Sweeney *et al.*, 2001), whereas the toxic effects of retinoids can be better related to the AUC (Tzimas *et al.*, 1997). This asks for different reverse dosimetry approaches for these chemicals, which are discussed in the present thesis.

3) Differences in free fraction of the chemical in the in vitro assay compared with the in vivo situation

When the *in vitro* concentration-response curves are set as internal concentrations in the PBK model to predict the accompanying *in vivo* dose levels, one must take into account that for some chemicals, the free concentration in the culture medium will be higher than the free concentration in the blood (or tissue) with equal nominal concentrations. This is the case for protein-binding chemicals, because protein concentrations are higher in blood than in culture medium (Gülden *et al.*, 2006; Blaauboer, 2010). Because it is assumed that the available free fraction of a chemical causes the toxicity, corrections have to be made to take these differences in availability into account. These corrections made in the present thesis in the translation of the *in vitro* effect concentrations of the retinoid ATRA into *in vivo* doses provide an example of the importance of such a correction for protein binding.

Outline of thesis

In **chapter 1** of the thesis, background information on the topic is given and the aim of the thesis is presented, being the provision of the proof-of-principle that *in vitro* concentration-response curves can be converted into *in vivo* dose-response

curves for developmental toxicity, suitable for risk assessment, using PBK modeling with a reverse dosimetry approach. To reach this aim, an *in vitro* model to obtain *in vitro* concentration-response curves for developmental toxicity is needed, as well as the selection of the chemicals to be studied and the development of PBK models. **Chapter 2** gives an overview of the *in vitro* assays and readouts for developmental toxicity reported in the literature, which could be useful for the reverse dosimetry approach to predict *in vivo* developmental toxicity. It describes how obtained *in vitro* developmental toxicity data may be used to predict *in vivo* developmental toxicity and which factors hamper these *in vitro-in vivo* extrapolations.

Chapters 3, 4 and 5 deal with the first series of chemicals used in the present thesis: the glycol ethers. These chemicals are known to cause developmental toxicity via their alkoxyacetic acid metabolites. First, in **chapter 3** the mode of action of these metabolites is investigated by assessing whether developmental toxicity of glycol ethers may be caused by a reduction of the intracellular pH (pH_i) induced by their alkoxyacetic acid metabolites. To this end, mouse embryonic fibroblasts (Balb/c-3T3 cells) and mouse ES-D3 cells are exposed to increasing concentrations of the alkoxyacetic acid metabolites MAA, EAA, BAA and PAA of the glycol ethers EGME, EGEE, EGBE and EGPE, respectively (Figure 2). The effect concentrations decreasing the pH_i are related to effect concentrations inhibiting ES-D3 cell differentiation and the effect of amiloride, a selective inhibitor of the Na^+/H^+-antiporter, on sensitivity of the cells to MAA is characterised, to assess whether the alkoxyacetic acid-induced decrease in pH_i may be the cause of their inhibition of embryonic stem cell differentiation. **Chapter 4** presents the results of the exposure of differentiating ES-D3 cells to MAA, EAA, BAA and PAA. It also compares the *in vitro* outcomes of these glycol ether alkoxyacetic acid metabolites with literature data on the *in vivo* outcomes of the respective parent glycol ethers obtained in mice. In **chapter 5** the results are shown for the translation of the *in vitro* concentration-response curves of MAA, EAA, BAA and PAA obtained in the ES-D3 cell differentiation assay (Chapter 4) into *in vivo* dose-response curves of the parent glycol ethers (EGME, EGEE, EGBE and EGPE respectively) in rat and human, using the reverse dosimetry approach. To provide the proof-of principle that *in vitro* concentration-response curves for developmental toxicity can be converted into dose-response curves for *in vivo* developmental toxicity, predicted dose-response curves for *in vivo* developmental toxicity in rats are compared with

literature data on glycol ether-induced *in vivo* developmental toxicity in rats.

In chapter 6 and 7 of the thesis, results for the second selected series of chemicals, the retinoids, are presented. In **chapter 6**, the concentration-response curves for a group of retinoids (retinol, ATRA, 13-*cis*-RA, 9-*cis*-RA, acitretin, etretinate and TTNPB ((E)-4[2-(5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2-naphthalenyl)-1-propenyl]benzoic acid); Figure 3) in the ES-D3 cell differentiation assay are presented. The outcomes of these studies are compared with literature data on concentration-response effects in two other *in vitro* assays for developmental toxicity (the limb bud micromass test and the rat post implantation whole embryo culture test). Furthermore, the outcomes obtained in the ES-D3 cell differentiation assays are compared with literature data on dose-response effects of these retinoids obtained in rats, mice and rabbits *in vivo*. Moreover, it is investigated whether the retinoid-induced inhibition of ES-D3 cell differentiation is mediated via the nuclear retinoic acid receptors, which is the pathway proposed to be disturbed in retinoid-induced developmental toxicity *in vivo*. In **chapter 7** the *in vitro* concentration-response curve of the retinoid ATRA obtained in the ES-D3 cell differentiation assay (chapter 6) is translated into *in vivo* dose-response curves of ATRA in rat and human, using the reverse dosimetry approach. The PBK model developed for ATRA is solely based on kinetic parameter values derived using *in vitro* techniques. To assess whether *in vivo* developmental toxicity could be predicted for this chemical with a relatively complex mode of action, using the reverse dosimetry approach, predicted dose-response curves for *in vivo* developmental toxicity in rats are compared with literature data on ATRA-induced *in vivo* developmental toxicity in rats.

Finally, **chapter 8** summarizes and discusses the results of the present thesis and provides future perspectives. It discusses possible implications of the use of the reverse dosimetry approach for toxicological risk assessment and gives directions for future research in the field.

References

Artola-Garicano, E., Vaes, W. H., Hermens, J. L. (2000). Validation of negligible depletion solid-phase microextraction as a tool to determine tissue/blood partition coefficients for semivolatile and nonvolatile organic chemicals. *Toxicol Appl Pharmacol* **166**, 138-44.

Berezhkovskiy, L. M. (2004). Volume of distribution at steady state for a linear pharmacokinetic system with peripheral elimination. *J Pharm Sci* **93**, 1628-1640.

Blaauboer, B. J. (2010). Biokinetic modeling and in vitro-in vivo extrapolations. *J Toxicol Environ Health B Crit Rev* **3**, 242-52.

Brown, N. A., Holt, D., Webb, M. (1984). The teratogenicity of methoxyacetic acid in the rat. *Toxicol Lett* **22**, 93-100.

Brown, N. A. (2002). Selection of test chemicals for the ECVAM international validation study on in vitro embryotoxicity tests. European Centre for the Validation of Alternative Methods. *Altern Lab Anim* **30**, 177-98.

Brown, R. P., Delp, M. D., Lindstedt, S. L., Rhomberg, L. R., Beliles, R. P. (1997). Physiological parameter values for physiologically based pharmacokinetic models. *Toxicol Indust Health* **13**, 407-84.

Chiu, W. A., Barton, H. A., DeWoskin, R. S., Schlosser, P., Thomson C. M., Sonawane, B., Lipscomb, J. C., Krishnan, K. (2007). Evaluation of physiologically based pharmacokinetic models for use in risk assessment. *J Appl Toxicol* **27**, 218-37.

Clewell, H. J. 3rd. (1993). Coupling of computer modeling with in vitro methodologies to reduce animal usage in toxicity testing. *Toxicol Lett* **68**, 101-17.

Clewell, R. A., Clewell, H. J. 3rd. (2008). Development and specification of physiologically based pharmacokinetic models for use in risk assessment. *Regul Toxicol Pharmacol* **50**, 129-43.

Collins, M. D., Mao, G. E. (1999). Teratology of retinoids. *Annu Rev Pharmacol Toxicol* **39**, 399-430.

DeJongh, J., Verhaar, H. J., Hermens, J. L. (1997). A quantitative property-property relationship (QPPR) approach to estimate in vitro tissue-blood partition coefficients of organic chemicals in rats and humans. *Arch Toxicol* **72**, 17-25.

DeJongh, J., Nordin-Andersson, M., Ploeger, B. A., Forsby A. (1999). Estimation of systemic toxicity of acrylamide by integration of in vitro toxicity data with kinetic simulations. *Toxicol Appl Pharmacol* **158**, 261-68.

ECETOC. (2005). The toxicology of glycol ethers and its relevance to man. Technical Report 095, 4th ed. European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels, Belgium.

Feuston, M. H., Kerstetter, S. L., Wilson, P. D. (1990). Teratogenicity of 2-methoxyethanol applied as a single dermal dose to rats. *Fundam Appl Toxicol* **15**, 448-56.

Forsby, A., Blaauboer, B. J. (2007). Integration of in vitro neurotoxicity data with biokinetic modelling for the estimation of in vivo neurotoxicity. *Hum Exp Toxicol* **26**, 333-38.

Gargas, M. L., Burgess, R. J., Voisard, D. E., Cason, G. H., Andersen, M. E. (1989). Partition coefficients of low-molecular-weight volatile chemicals in various liquids and tissues. *Toxicol Appl Pharmacol* **98**, 87-99.

Giavini, N. A., Broccia, M. L., Menegola, E., Prati, M. (1993). Comparative in vitro study of the embryotoxic effects of three glycol ethers and their metabolites, the alkoxyacids. *Toxicol in Vitro* **7**, 777-84.

Genschow, E., Spielmann, H., Scholz, G., Seiler, A., Brown, N. A., Piersma, A. H., Brady, M., Clemann, N., Huuskonen, H., Paillard, F., Bremer, S., Becker, K. (2002). The ECVAM international validation study on in vitro embryotoxicity tests: results of the definitive phase and evaluation of prediction models. European Centre for the Validation of Alternative Methods. *Atern Lab Anim* **30**, 151-76.

Genschow, E., Spielmann, H., Scholz, G., Pohl, I., Seiler, A., Clemann, N., Bremer, S., Becker, K. (2004).

Validation of the embryonic stem cell test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* **32**, 209-44.

Gollnick, H., Bauer, R., Brindley, C., Orfanos, C. E., Plewig, G., Wokalek, H., Hoting, E. (1988). Acitretin versus etretinate in psoriasis. Clinical and pharmacokinetic results of a German multicenter study. *J Am Acad Dermatol* **19**, 458-68.

Grindon, C., Combes, R. (2008). Introduction to the EU REACH legislation. *Altern Lab Anim* **36** Suppl 1, 1 -6.

Gülden, M., Dierickx, P., Seibert, H. (2006). Validation of a prediction model for estimating serum concentrations of chemicals which are equivalent to toxic concentrations in vitro. *Toxicol In Vitro* **20**, 1114-24.

Hanley, T. R. Jr., Yano, B. L., Nitschke, K. D., John, J. A. (1984). Comparison of the teratogenic potential of inhaled ethylene glycol monomethyl ether in rats, mice, and rabbits. *Toxicol Appl Pharmacol* **75**, 409-22.

Höfer, T., Gerner, I., Gundert-Remy, U., Liebsch, M., Schulte, A., Spielmann, H., Vogel, R., Wettig, K. (2004). Animal testing and alternative approaches for the human health risk assessment under the proposed new European chemicals regulation. *Arch Toxicol* **78**, 549-64.

Kanigsberg, N., DesGroseilliers, J. P. (1983). Use of 13-cis-retinoic acid in cystic acne. *Can Med Assoc J* **129**, 224.

Krewski, D., Withey, J. R., Ku, L. F., Andersen, M. E. (1994). Applications of physiologic pharmacokinetic modeling in carcinogenic risk assessment. *Environ Health Perspect* **102**, 37-50.

Krishnan, K., Andersen, M. E. In: Hayes, A. W. (Ed.). (2001). Principles and Methods of Toxicology, Raven Press, NY, 193-241.

Nagano, K., Nakayama, E., Oobayashi, H., Yamada, T., Adachi, H., Nishizawa, T., Ozawa, H., Nakaichi, M., Okuda, H., Minami, K., Yamazaki, K. (1981). Embryotoxic effects of ethylene glycol monomethyl ether in mice. *Toxicology* **20**, 335-43.

Paini, A., Punt, A., Viton, F., Scholz, G., Delatour, T., Marin-Kuan, M., Schilter, B., van Bladeren, P. J., Rietjens, I. M. C. M. (2010). A physiologically based biodynamic (PBBD) model for estragole DNA binding in rat liver based on in vitro kinetic data and estragole DNA adduct formation in primary hepatocytes. *Toxicol Appl Pharmacol* **245**, 57-66.

Pazos, P., Pellizzer, C., Stummann, T. C., Hareng, L., Bremer, S. (2010). The test chemical selection procedure of the European Centre for the Validation of Alternative Methods for the EU Project ReProTect. *Reprod Toxicol* **30**, 161-99.

Piersma, A. H., Genschow, E., Verhoef, A., Spanjersberg, M. Q., Brown, N. A., Brady, M., Burns, A., Clemann, N., Seiler, A., Spielmann, H. (2004). Validation of the postimplantation rat whole-embryo culture test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* **32**, 275-307.

Poulin, P., Theil, F. P. (2002). Prediction of pharmacokinetics prior to in vivo studies. II. Generic physiologically based pharmacokinetic models of drug disposition. *J Pharm Sci* **91**, 1358-70.

Punt, A., Paini, A., Boersma, M. G., Freidig, A. P., Delatour, T., Scholz, G., Schilter, B., van Bladeren, P. J., Rietjens, I. M. C. M. (2009). Use of physiologically based biokinetic (PBBK) modeling to study estragole bioactivation and detoxification in humans as compared with male rats. *Toxicol Sci* **110**: 255-69.

Rietjens, I. M. C. M., Louisse, J., Punt, A. (2011). Tutorial on physiologically based kinetic modeling in molecular nutrition and food research. *Mol Nutr Food Res* **55**, 941-56.

Scialli, A.R. (2008). The challenge of reproductive and developmental toxicology under REACH. *Regul Toxicol Pharmacol* **51**, 244-50.

Spielmann, H., Genschow, E., Brown, N. A., Piersma, A. H., Verhoef, A., Spanjersberg, M. Q., Huuskonen, H., Paillard, F., Seiler, A. (2004). Validation of the rat limb bud micromass test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* **32**, 245-74.

Staatscourant (2001). Niet-limitatieve lijst reprotoxische stoffen. *Staatscourant* **104**, 10.

Sweeney, L. M., Tyler, T. R., Kirman, C. R., Corley, R. A., Reitz, R. H., Paustenbach, D. J., Holson, J. F., Whorton, M. D., Thompson, K. M., Gargas, M. L. (2001). Proposed occupational exposure limits for select ethylene glycol ethers using PBPK models and Monte Carlo simulations. *Toxicol Sci* **62**, 124-39.

Tallman, M. S., Andersen, J. W., Schiffer, C. A., Appelbaum, F. R., Feusner, J. H., Ogden, A., Shepherd, L., Willman, C., Bloomfield, C. D., Rowe, J. M., Wiernik, P. H. (1997). All-trans-retinoic acid in acute promyelocytic leukemia. *N Engl J Med* **337**, 1021-8.

Tzimas, G., Thiel, R., Chahoud, I., Nau, H. (1997). The area under the concentration-time curve of all-trans-retinoic acid is the most suitable pharmacokinetic correlate to the embryotoxicity of this retinoid in the rat. *Toxicol Appl Pharmacol* **143**, 436-44.

U.S. EPA. (2006). Approaches for the application of physiologically based pharmacokinetic (PBPK) models and supporting data in risk assessment. EPA/600/R-05/043F. National Center for Environmental Assessment: Washington, DC.

Van der Jagt, K., Munn, S. J., Tørsløv, J., de Bruijn, J. (2004). Alternative approaches can reduce the use of test animals under REACH. *Report EUR* **21405**.

Verwei, M., van Burgsteden, J. A., Krul, C. A. M., van de Sandt, J. J. M., Freidig, A. P. (2006). Prediction of in vivo embryotoxic effect levels with a combination of in vitro studies and PBPK modelling. *Toxicol Lett* **165**, 79-87.

Washenik, K., Clark-Loeser, L., Friedman-Kien, A. (2000). Kaposi's sarcoma. *N Engl J Med* **343**, 581-2.

Wier, P. J., Lewis, S. C., Traul, K. A. (1987). A comparison of developmental toxicity evident at term to postnatal growth and survival using ethylene glycol monoethyl ether, ethylene glycol monobutyl ether and ethanol. *Teratog Carcinog Mutagen* **7**, 55-64.

Chapter 2 is reprinted with permission from Informa Healthcare.

CHAPTER 2

Toward *in vitro* biomarkers for developmental toxicity and their extrapolation to the *in vivo* situation

Jochem Louisse, Miriam Verwei, Ruud A. Woutersen,
Bas J. Blaauboer, Ivonne M.C.M. Rietjens

Based on:
Expert Opinion On Drug Metabolism and Toxicology 8, 11-27 (2012).

Abstract

Introduction: Reliable *in vitro* and *in silico* assays as alternatives for *in vivo* developmental toxicity studies are urgently needed, for the replacement, reduction and refinement (3Rs) of animal use in toxicological research. Therefore, relevant biomarkers for *in vivo* developmental toxicity in *in vitro* assays are needed.

Areas covered: The present review gives an overview of alternative assays, as described in the literature, for *in vivo* developmental toxicity, including the effects (readouts) assessed in these assays. It is discussed how these data may be used to obtain relevant biomarkers for *in vivo* developmental toxicity, and how *in vitro* effect data can be translated to the *in vivo* situation using physiologically based kinetic (PBK) modeling.

Expert opinion: Relevance of readouts in *in vitro* developmental toxicity assays as predictive biomarkers for *in vivo* developmental toxicity should be evaluated by comparing the obtained *in vitro* effect concentrations with *in vivo* internal concentrations at dose levels causing developmental toxicity. Extrapolation of the *in vitro* effect concentrations to *in vivo* dose levels using PBK modeling (i.e., reverse dosimetry) is promising in its use to set points of departure for risk assessment, enabling the use of *in vitro* toxicity data in the safety assessment of chemicals.

Introduction

For the registration and safe use of chemicals, information on safety is required. It is estimated that more than 25% of the 3.9 million laboratory animals expected to be needed for safety testing for industrial chemicals under the European REACH (Registration, Evaluation, Authorization and restriction of CHemicals) legislation will be used for developmental toxicity studies (Van der Jagt *et al.*, 2004). Also for pharmaceuticals, food additives and pesticides, regulatory agencies require safety studies performed with laboratory animals, including developmental toxicity studies. The use of high numbers of laboratory animals for developmental toxicity testing in chemical safety assessment meets increasing ethical and economic constraints. Therefore, the development, validation and application of reliable *in vitro* and/or *in silico* assays as alternatives for *in vivo* developmental toxicity studies are urgently needed, contributing to the replacement, reduction and refinement (3Rs) of animal use in toxicological research. According to the test guidelines for developmental toxicity testing of the Organisation for Economic Co-operation and Development (OECD guidelines 414, 421, 422 or 426), the preferred rodent and non-rodent species for *in vivo* developmental toxicity testing are rat and rabbit, respectively. Several *in vitro* alternative assays for developmental toxicity testing have been developed, which may be useful in the early phase of drug discovery, in the development and early safety testing of new food additives and for setting priorities for toxicity testing of chemicals within the REACH framework. However, these assays are not capable yet of replacing or reducing the regulatory *in vivo* developmental toxicity studies. Consequently, a large number of laboratory animals are still needed for *in vivo* developmental toxicity testing.

The present review gives an overview of the alternative test systems for developmental toxicity testing that have been described in the literature. It focuses on how the readouts of these *in vitro* assays may serve as predictive biomarkers for *in vivo* developmental toxicity and describes how the data obtained may be used to predict the potency of a chemical to induce developmental toxicity *in vivo*.

Alternative assays for *in vivo* developmental toxicity testing

Several *in vitro* and *ex vivo* assays have been described in the literature, which

evaluate the developmental toxicity potencies of chemicals. Assays have been developed to be able to predict developmental toxicity not only for a broad range of chemicals covering diverse developmental toxicity endpoints *in vivo* (e.g., the embryonic stem cell test (EST) and the rat postimplantation whole embryo culture (WEC) test), but also for chemicals affecting specific phases of development, such as the development of the neuronal systems. Several readouts for developmental toxicity have been defined in these assays in order to obtain predictive biomarkers for the *in vivo* situation. The term 'biomarker' has been defined as 'a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention' (Biomarkers definitions working group, 2001). Applying this definition specifically to toxicology, Stachlewitz (2007) came up with the definition of a biomarker as 'something that can be objectively measured that will provide a reliable signal indicating the chemical being tested is about to have an adverse effect on a system'. Identified biomarkers can, for example, range from the level of gene expression changes to morphological changes. Many readouts in *in vitro* assays for developmental toxicity have been described in literature that may be useful as biomarkers (Tables 1 and 2). However, whether these readouts can be used as predictive biomarkers for *in vivo* developmental toxicity remains to be established.

The assays that have been developed as alternatives for *in vivo* developmental toxicity testing can roughly be divided into 1) cellular assays, using either primary cells or cells from continuous cell lines and 2) WEC assays.

Cellular assays

The European Centre for the Validation of Alternative Methods (ECVAM) has validated two cellular assays for developmental toxicity including their prediction models, which are the rat limb bud micromass (MM) test and the embryonic stem cell test (EST) (Genschow *et al.*, 2002). In addition, several other (embryonic stem) cell-based assays have been described in the literature, focusing on different *in vitro* readouts for developmental toxicity, which have so far not been validated by ECVAM. In the following sections, the cellular assays are discussed in more detail.

The MM test

In the MM test, primary cell cultures dissociated from limb buds and/or midbrains are used (Flint, 1983; Flint and Orton, 1984; Kistler, 1987). In the validated MM test protocol, only limb bud cells are used, which are derived from embryos obtained from pregnant rats at gestational day (GD)14. Limb buds are dissociated and plated in high density spots (MM cultures), which allows their differentiation into chondrocytes (Genschow *et al.*, 2002; Spielmann *et al.*, 2004). After 5 days of exposure to the chemical, the effect on growth and differentiation of limb bud cells is assessed. Only the EC_{50} values obtained in the differentiation assay are applied in the prediction model for the MM test, because no better predictions were made by using either the EC_{50} values obtained in the proliferation assay or the EC_{50} values for both proliferation and differentiation (Spielmann *et al.*, 2004). Since the strong embryotoxic chemical methotrexate was classified as non-embryotoxic, the MM test was not recommended for screening purposes (Spielmann *et al.*, 2004).

The EST

In the last decade of the previous century, the use of embryonic stem cells for the assessment of developmental toxicity potencies of chemicals was introduced (Newall and Beedles, 1994; Heuer *et al.*, 1997; Scholz *et al.*, 1999; Genschow *et al.*, 2000). These initiatives led to the development of the EST, which was validated with its prediction model by the ECVAM (Heuer *et al.*, 1997; Scholz *et al.*, 1999; Genschow *et al.*, 2004). The EST measures the potency of a chemical to inhibit cell growth of embryonic fibroblasts (3T3 cells) and embryonic stem cells (ES-D3 cells) (cell proliferation assays) and the potency of the chemical to inhibit ES-D3 cell differentiation into contracting cardiomyocytes (ES-D3 cell differentiation assay). The exposure duration is 10 days, in which the (exposure) medium is renewed two times (3 and 5 days after start). The acquired EC_{50} values in these assays are applied in a prediction model to predict the embryotoxic potency (non-, weak or strong embryotoxic) of the chemical. It was concluded in its validation study that the EST was ready for consideration for use in assessing the embryotoxic potentials of chemicals for regulatory testing (Genschow *et al.*, 2004).

Other cellular assays for developmental toxicity testing

Several *in vitro* readouts for developmental toxicity have been assessed in embryonic stem cell-based cultures, for example, the expression of specific genes for neuronal differentiation (Stumann *et al.*, 2007; Kuegler *et al.*, 2010; Theunissen *et al.*, 2010; Theunissen *et al.*, 2011; Zimmer *et al.*, 2011a; Zimmer *et al.*, 2011b), osteoblast differentiation (Zur Nieden *et al.*, 2010a; Zur Nieden *et al.*, 2010b) or endothelial differentiation (Fesstag *et al.*, 2007). These readouts may be useful biomarkers for adverse effects on specific developmental pathways, such as neurodevelopment, bone formation, or vasculogenesis and/or angiogenesis, respectively. An overview of the studies using different *in vitro* readouts in embryonic stem cell-based test systems, other than the inhibition of ES-D3 cell differentiation into contracting cardiomyocytes, is presented in Table 1. Appearance of a certain effect on exposure to a developmental toxicant does not necessarily mean that this effect will be a useful biomarker for developmental toxicity. For obtaining useful biomarkers in *in vitro* assays, it should be carefully assessed whether the effect is adverse, whether it is obtained at a relevant concentration and whether it is relevant and predictive for the *in vivo* toxicological endpoint.

In the last decade, the use of toxicogenomics for obtaining biomarkers has gained importance (Afshari *et al.*, 2011). Transcriptome analyses in embryonic stem cell cultures exposed to varying chemicals have been used to differentiate teratogens from non-teratogens by analyzing gene expression patterns in control cells, cells exposed to teratogens and cells exposed to non-teratogens (Van Dartel *et al.*, 2010a; Van Dartel *et al.*, 2010b; Van Dartel *et al.*, 2011a; Van Dartel *et al.*, 2011b; Van Dartel *et al.*, 2011c; Van Dartel and Piersma, 2011). Toxicogenomic approaches have also been proven useful to assess whether chemicals share the same mode of action (Osman *et al.*, 2010) and whether pathways affected on *in vivo* exposure are also affected in *in vitro* embryonic cell cultures, in order to demonstrate the predictive capacities of the *in vitro* assays (Stigson *et al.*, 2007).

Table 1. Readouts in embryonic stem cell-based assays for developmental toxicity, other than the inhibition of embryonic stem cell differentiation into contracting cardiomyocytes. (Table continues on following page.)

Test system	Readout(s)	Chemical(s)	Reference(s)
mouse embryonic stem cells (ES-D3)	mRNA and protein expression levels of several neuronal marker genes	methylmercury	Theunissen *et al.* (2010)
mouse embryonic stem cells (ES-D3)	transcriptome analyses	methylmercury	Theunissen *et al.* (2011)
mouse embryonic stem cells (CGR8)	mRNA expression levels of several neuronal marker genes	methylmercury	Zimmer *et al.* (2011a)
mouse embryonic stem cells (R1)	mRNA expression levels of mesodermal, endodermal and ectodermal genes	valproic acid, carbamazepine	Murabe *et al.* (2007a,b)
mouse embryonic stem cells (ES-D3)	alpha-fetoprotein enhancer induced green fluorescent protein (GFP) protein expression levels	diphenylhydantoin	Paparella *et al.* (2002)
mouse embryonic stem cells (ES-D3)	sarcomeric myosin heavy chain and alpha-actin protein expression levels	5-fluorouracil, all-*trans* -retinoic acid, penicillin G	Seiler *et al.* (2004)
mouse embryonic stem cells (ES-D3)	mRNA expression levels of visceral endodermal and mesodermal genes	5-fluorouracil	Pamies *et al.* (2010)
mouse embryonic stem cells (ES-D3)	mRNA expression levels of markers for mature neural cells, osteoblasts and chondrocytes	penicillin G, 5-fluorouracil, all-*trans* -retinoic acid, diphenylhydantoin, valproic acid, thalidomide,	Zur Nieden *et al.* (2004)
mouse embryonic stem cells (ES-D3)	mRNA expression levels of several neuronal marker genes	methylmercury	Stumann *et al.* (2009)
mouse embryonic stem cells (ES-D3)	transcriptome analyses	5-fluorouracil, monobutylphthalate, monomethylphthalate, monoethylhexylphthalate, flusilazole, hexaconazole, triadimefon, carbamazepine, methylmercury, methoxyacetic acid, valproic acid, all-*trans* -retinoic acid, warfarin, nitrofen, methotrexate, saccharine, penicillin G	Van Dartel *et al.* (2009) Van Dartel *et al.* (2010a,b) Van Dartel *et al.* (2011a-c)

Test system	Readout(s)	Chemical(s)	Reference(s)
mouse embryonic stem cells (ES-D3)	proteome analyses	monobutylphthalate	Osman *et al.* (2010)
mouse embryonic stem cells (ES-D3)	proteome analyses	warfarin, lovastatin	Groebe *et al.* (2010a)
mouse embryonic stem cells (ES-D3)	proteome analyses	dinoseb, ochratoxin-A, nitrofen, β-aminoproprionitril, metoclopramide, doxylamine succinate, D-penicillamine	Groebe *et al.* (2010b)
mouse embryonic stem cells (R1)	transcriptome analyses	valproic acid and valproic acid analogues	Jergil *et al.* (2011)
monkey embryonic stem cells (CMK-6)	mRNA expression of several embryonic stem cell differentiation marker genes	bisphenol A	Yamamoto *et al.* (2007)
human embryonic stem cells (H9)	transcriptome analyses	cytosine arabinoside	Jagtap *et al.* (2011)
human embryonic stem cells (PKU1.1)	cardiac specific mRNA markers	cyclophosphamide	Zhu *et al.* (2011)
human embryonic stem cells (WA09)	metabolome analyses	ascorbic acid, isoniazid, penicillin G, folic acid, levothyroxine, retinol, doxylamine, thiamine, aspirin, caffeine, diphenhydramine, dexamethasone, diphenylhydantoin, methotrexate, 5-fluorouracil, busulfan, cytosine arabinoside, hydroxyurea, all-*trans* -retinoic acid, thalidomide, valproic acid, amiodarone, rifampicin, carbamazepine, 13-*cis* -retinoic acid, cyclophosphamide	West *et al.* (2010)

WEC assays

Of the WEC assays, only the postimplantation rat WEC test, including its prediction model, has been validated by the ECVAM (Piersma *et al.*, 2004). Other *ex vivo* embryo culture assays include the frog embryo teratogenesis assay – Xenopus (FETAX) and the chicken embryotoxicity screening test (CHEST). Furthermore, several studies have described the use of zebrafish embryos for developmental toxicity testing. In most studies reported in the literature using WECs, morphological examination is used to determine the developmental toxicity potency of the chemical tested. However, other readouts, listed in Table 2, may also be useful as biomarkers for developmental toxicity in these WEC systems.

The postimplantation rat WEC test

Rat WECs have already been used for decades to assess developmental toxicity potencies of chemicals (Brown and Fabro, 1981; Schmid *et al.*, 1983; Klug *et al.*, 1985; Piersma *et al.*, 1995). This section describes the WEC test procedure as applied in the ECVAM validation study (Genschow *et al.*, 2002; Piersma *et al.*, 2004). In the WEC test, embryos obtained from pregnant rats at GD10 are cultured in 100% rat serum, containing the chemical, for 48 h. After this culture period, heart beat and yolk sac circulation are scored and measurements of yolk sac diameter, crown-rump length and head length are taken (Genschow *et al.*, 2002; Piersma *et al.*, 2004). Furthermore, morphological, developmental, functional and growth parameters are scored. The number and specificity of malformations are recorded. No observed effect concentrations (IC_{NOEC}), 50% effect concentrations (IC_{50}) and maximum effect concentrations (IC_{max}) are determined. In the first prediction model developed for the WEC test, the no observed effect concentrations on total morphological score (IC_{NOEC} TMS) and the IC_{50} values for inducing malformations (IC_{50mal}) were used for classification, whereas in a second prediction model, the IC_{NOEC} TMS values, the IC_{max} values for malformations and the IC_{50} values obtained in the 3T3 cell proliferation assay of the EST were used for classification (Genschow *et al.*, 2000; Genschow *et al.*, 2002; Piersma *et al.*, 2004). The second prediction model was demonstrated to be better in predicting the embryotoxicity classes (non-, weak or strong) of the chemicals. It was concluded that the WEC test was ready for consideration for use in assessing the embryotoxic potentials of chemicals for regulatory testing (Piersma *et al.*, 2004).

Table 2. Readouts in whole embryo culture assays for developmental toxicity other than morphological examinations.

Test system	Readout(s)	Chemical(s)	Reference(s)
frog embryos	heat shock protein mRNA expression levels	sodium arsenate	Gornatti *et al.* (2002)
zebrafish embryos	mRNA expression levels of nervous system genes	ethanol	Fan *et al.* (2010)
zebrafish embryos	transcriptome analyses	glycol ether alkoxyacetic acid metabolites, 1,2,4-triazoles	Hermsen *et al.* (2011a)
rat embryos	E-cadherin mRNA and protein expression levels	cadmium	Chen and Hales (1994)
rat embryos	heat shock protein expression levels	*N*-acetoxy-2-acetylaminofluorene, cadmium chloride, cyclophosphamide, sodium arsenate, sodium salicylate	Mirkes *et al.* (1994)
rat embryos	8-Isoprostaglandin F2alpha levels	diclofenac	Chan *et al.* (2002)
rat embryos	transcriptome analyses	caffeine, methylmercury, monobutylphthalate, methoxyacetic acid	Robinson *et al.* (2010)
rat embryos	proteome analyses	all-*trans*-retinoic acid	Luijten *et al.* (2010)

The FETAX

The FETAX is a developmental toxicity screening test conducted on *Xenopus laevis* mid-blastula-stage eggs over the organogenesis period (Bantle *et al.*, 1999; Fort and Paul, 2002; Mouche *et al.*, 2011). The teratogenic potential of a chemical is determined after a 96-h exposure to the chemical, by scoring the mortality and malformation rate for the larvae, as well as the larvae length (Mouche *et al.*, 2011). Subsequently, a teratogenic index (TI) is determined by dividing the obtained LC_{50} (50% embryo lethality) by the $EC_{50,\ malformation}$ (50% induction of malformed larvae). If the TI > 1.2, the chemical will be classified as positive, if the TI < 1.0, it will be classified as negative. Chemicals with a TI between 1.0 and 1.2 are classified as positive if growth is affected and morphogenesis and/or viability is disturbed, but only when the concentration affecting morphogenesis is lower than the concentration leading to significant mortality (Mouche *et al.*, 2011).

An interlaboratory validation study was performed to evaluate the repeatability and reliability of the FETAX to evaluate the potential teratogenic hazard for 12 chemicals. It was concluded that the FETAX proved to yield repeatable and reliable data, as long as technicians were adequately trained (Bantle *et al.*, 1999). The data from this validation study were later re-evaluated using characteristic malformation criteria, leading to an improved reliability, test endpoint precision and predictability (Fort and Paul, 2002).

The CHEST

In the CHEST, the developmental toxicity potential of chemicals is evaluated using chicken embryos. The method was first described by Jelinek (1977) and more recently by Davies and Freeman (1995). In the CHEST, fertilized chicken eggs are exposed to different doses of the chemical for different durations, after which the effects are assessed. The number of dead, malformed and growth-retarded fetuses is recorded, on the basis of which the developmental toxicity potency of the chemical is determined (Davies and Freeman, 1995). The CHEST protocol is divided into two phases: 1) determination of the toxic dose range in early administration time (24 h; CHEST I) and 2) determination of the teratogenic dose range covering also the late effects on embryo development (days 2, 3 and 4; CHEST II) (Davies and Freeman, 1995). Several studies have evaluated the CHEST

and similar test protocols, and it was concluded that the CHEST is a reproducible test system that delivers quantifiable data for evaluation (Adler *et al.*, 2011). The CHEST is being used in academia and industry for routine screening purposes and mechanistic studies (Adler *et al.*, 2011).

Developmental toxicity testing using zebrafish embryos

In recent years, the zebrafish (*Danio rero*) has gained popularity as a test species for developmental toxicity testing. Selderslaghs *et al.* (2009) determined TI values ($LC_{50}/EC_{50,\ malformation}$) in zebrafish embryos 24, 48, 72 and 144 h after fertilization and chemical exposure, and were able to distinguish four positive from two negative chemicals for developmental toxicity. However, separation of positives from negatives was not based on TI values, as is the case in the FETAX, but on the fact that for the two negative chemicals, no TI values could be determined, as no LC_{50} nor $EC_{50,\ malformation}$ was reached (Selderslaghs *et al.*, 2009). In a more recent study, this zebrafish assay was used to classify a larger group of 27 chemicals based on a prediction model using a TI cut-off value of 2. When comparing the classification obtained in this zebrafish assay with animal data, a sensitivity of 72% and a specificity of 100% were obtained for these 27 chemicals (Selderslaghs *et al.*, 2011). Hermsen *et al.* (2011b) developed a quantitative evaluation method to assess the development of the zebrafish embryo based on specific endpoints in time. In their zebrafish embryotoxicity test, the effect of chemicals on zebrafish development is assessed using a scoring system for development and teratogenic effects (Hermsen *et al.*, 2011b).

Zebrafish embryos have also been used to assess the neurodevelopmental toxicity potential of chemicals, based on locomotor activity studies (Selderslaghs *et al.*, 2010) and on gene expression patterns of nervous system genes (Fan *et al.*, 2010). Demicco *et al.* (2010) studied the effects of a series of pyrethroid insecticides in zebrafish embryos and concluded that the zebrafish might be an appropriate alternative model to study the mechanisms of neurodevelopmental toxicity and that it might be used to identify chemicals that should be further tested in mammalian models.

Translation of *in vitro* obtained developmental toxicity data to the *in vivo* situation

In order to use *in vitro* toxicity data for the estimation of *in vivo* toxic effects, the obtained *in vitro* toxicity data need to be translated to the *in vivo* situation. *In vivo* developmental toxicity may be predicted in several ways: the *in vitro* data can be used to 1) predict the developmental toxicity class to which a chemical belongs (e.g., non-, weak or strong embryotoxic) using a prediction model, 2) predict the relative developmental toxicity potency of a chemical based on the *in vitro* effect concentrations (usually within a group of structurally related chemicals) and 3) predict *in vivo* developmental toxicity dose levels by translating *in vitro* effect concentrations into *in vivo* effect doses.

Prediction of developmental toxicity classes using prediction models

Data from *in vitro* developmental toxicity studies are often used to divide chemicals in embryotoxicity or teratogenicity classes. This is done, for example, in the FETAX by separating non-teratogens from teratogens by deriving TI values (Mouche *et al.*, 2011). In the EST, the MM test and the WEC test, chemicals are divided into different embryotoxicity classes (non-, weak or strong embryotoxic) using prediction models (Genschow *et al.*, 2002). In these assays, the embryotoxic potency is predicted by applying the *in vitro* effect concentrations in prediction models. The prediction models of the EST, the MM test and the WEC test were developed by linear discriminant analysis, aiming at the creation of rules that assign a chemical to a toxicity class, taking into account the endpoints determined in the test (Genschow *et al.*, 2002). In the ECVAM validation studies, chemicals were correctly classified in 78, 70 and 80% of the experiments, using the *in vitro* data obtained in and the prediction models developed for the EST, the MM test and the WEC test, respectively (Genschow *et al.*, 2002). However, using the MM test, the strong embryotoxic chemical methotrexate was classified as non-embryotoxic. Therefore, the MM test was not recommended for screening purposes (Spielmann *et al.*, 2004). The predictability of the EST and the WEC test using their prediction models was further evaluated in later studies (Paquette *et al.*, 2008; Marx-Stoelting *et al.*, 2009; Thomson *et al.*, 2011). The usefulness of the prediction model of the EST was questioned, since only 2 of 13 chemicals that were tested in the European Framework Program project ReProTect were correctly classified

using the obtained *in vitro* data and the prediction model (Marx-Stoelting *et al.*, 2009). This might be due to the subjective classification of the chemicals over the three different embryotoxicity classes, possibly hampering the predictability by the prediction models (Marx-Stoelting *et al.*, 2009). It was suggested that the number of toxicity classes should be reduced to two: 1) non-embryotoxic chemicals and 2) embryotoxic chemicals (Marx-Stoelting *et al.*, 2009). When the WEC test was challenged with a group of pharmaceuticals not used in the ECVAM validation study, only 24 of 48 chemicals were correctly classified, confirming the concerns about the application of the prediction model to a more diverse chemical set than that used in the validation study (Thomson *et al.*, 2011).

If prediction models are to predict the developmental toxicity potencies of a set of chemicals, all relevant *in vivo* kinetic and dynamic characteristics playing a role in the developmental toxicity outcome of the chemicals should be addressed by the assay and the prediction model. Although some of the dynamic processes playing a role in the developmental toxicity outcome might be represented in the *in vitro* assay, important *in vivo* kinetic characteristics, which might greatly differ between the chemicals, are not taken into account. These characteristics include intestinal uptake, bioactivation to active metabolites (in the case of proteratogens), detoxification to inactive metabolites, placental transfer to the fetus and/or excretion into the urine. Since all of these *in vivo* kinetic processes may play a role in the levels of active molecules reaching the embryo, and hence the *in vivo* developmental toxicity potency of a chemical, it is probably impossible to make a prediction on the developmental toxicity potency of all chemicals applying *in vitro* effect concentrations obtained in a single prediction model.

Prediction of relative developmental toxicity potencies of structurally related chemicals using in vitro effect concentrations

The *in vivo* developmental toxicity potency of a chemical may also be estimated by using a so-called category approach, which may also be used in read-across approaches for developmental toxicity (Fabjan *et al.*, 2006; Enoch *et al.*, 2009). Read-across is based on the concept that data gaps of chemicals can be filled by reading across or interpolating or extrapolating (preferably interpolating) missing data values, if data exist for chemicals that are structurally related (Scialli, 2008). The grouping of chemicals for read-across may be facilitated by using structure-

activity relationship (SAR) models or by *in vitro* testing (Scialli, 2008). The relative *in vivo* developmental toxicity potencies of chemicals can be estimated, if *in vivo* toxicity data are available for structurally related chemicals, by evaluating the relative *in vitro* developmental toxicity potencies of these chemicals. Furthermore, based on the toxicity profile of the chemical with known *in vivo* toxicity, the most appropriate *in vitro* test could be selected representing the most sensitive developmental endpoints for a specific chemical category (Fabjan *et al.*, 2006). Recently, the use of quantitative SAR (QSAR) models for reproductive and developmental toxicity has been evaluated by the European Commission's Joint Research Centre (Piparo and Worth, 2010). It was concluded that relatively few models for this endpoint are present and that those available have limited applicability domains (Piparo and Worth, 2010).

Several studies have assessed the *in vitro* relative potencies of structurally related chemicals using the ES-D3 cell differentiation assay of the EST, for example, for a series of glycol ether alkoxyacetic acid metabolites (De Jong *et al.*, 2009) (Chapter 4), valproic acid analogues (De Jong *et al.*, 2011a; Riebeling *et al.*, 2011), retinoids (Louisse *et al.*, 2011) (Chapter 6) and 1,2,4-triazoles (De Jong *et al.*, 2011b). Also the WEC test has been used to predict the relative *in vivo* potencies of structurally related chemicals, for example, for a series of glycol ether alkoxyacetic acid metabolites (Giavini *et al.*, 1993), phthalates and their metabolites (Janer *et al.*, 2008) and 1,2,4-triazoles (De Jong *et al.*, 2011b). The potential of the zebrafish in a category approach has been shown for a series of glycol ether alkoxyacetic acid metabolites (Hermsen *et al.*, 2011a) and 1,2,4-triazoles (Hermsen *et al.*, 2011a; De Jong *et al.*, 2011b). These studies indicate the potential of these *in vitro* developmental toxicity assays for predicting relative *in vivo* potencies of structurally related chemicals. A disadvantage of this approach is that chemical-specific species-dependent differences in *in vivo* kinetics, which may affect the relative *in vivo* potencies, are not taken into account.

Prediction of in vivo developmental toxicity dose levels by translating in vitro effect concentrations into in vivo effect doses

In vitro toxic effect concentrations that have been obtained in a relevant assay can be translated into *in vivo* doses in order to make predictions for the *in vivo* situation (DeJongh *et al.*, 1999; Verwei *et al.*, 2006; Forsby and Blaauboer, 2007;

Louisse *et al.*, 2010 (Chapter 5)). It has previously been shown that *in vitro* effect concentrations obtained in the ES-D3 cell differentiation assay can be successfully extrapolated to *in vivo* dose levels causing developmental toxicity, using physiologically based kinetic (PBK) modeling (Verwei *et al.*, 2006; Louisse *et al.*, 2010 (Chapter 5)). A PBK model is a set of mathematical equations that together describe the absorption, distribution, metabolism and excretion (ADME) characteristics of a chemical within an organism on the basis of physiological and anatomical parameters, physicochemical parameters (partition coefficients) and other kinetic parameters (Krishnan and Andersen, 2001). PBK models describe the fate of the chemical (and its metabolite(s)) in the body in time after a certain dose of the chemical via a certain exposure route (e.g., oral or dermal) in a certain species (e.g., rat or man). Therefore, PBK models can be used to translate external dose levels of a chemical into internal tissue or blood concentrations of the chemical (and its metabolite(s)). PBK models can also be used for so-called reverse dosimetry, in which internal concentrations of the chemical (or of one of its metabolites) are set and the model is used to calculate which external dose levels are needed to reach these internal concentrations. This was done for glycol ethers for which *in vitro* toxicity data of the toxic glycol ether alkoxyacetic acid metabolites derived in the ES-D3 cell differentiation assay of the EST were used as input in the PBK model for the internal blood concentrations to calculate the corresponding dose levels of the parent glycol ethers (Louisse *et al.*, 2010) (Chapter 5). This enables the extrapolation of *in vitro* concentration-response curves to predicted *in vivo* dose-response curves for developmental toxicity (Louisse *et al.*, 2010). The dose-response curves that were predicted for rat were found to be in concordance with the developmental toxicity dose levels measured in reported *in vivo* rat studies (Figure 1) (Louisse *et al.*, 2010). This example provides a proof-of-principle that PBK modeling can be used to convert *in vitro* concentration-response curves obtained in *in vitro* assays into *in vivo* dose-response curves for developmental toxicity. Reverse dosimetry approaches using PBK modeling are also used in the United States Environmental Protection Agency's ToxCast program and the European Framework Program projects ESNATS and ChemScreen to predict *in vivo* effect levels based on *in vitro* effect concentrations.

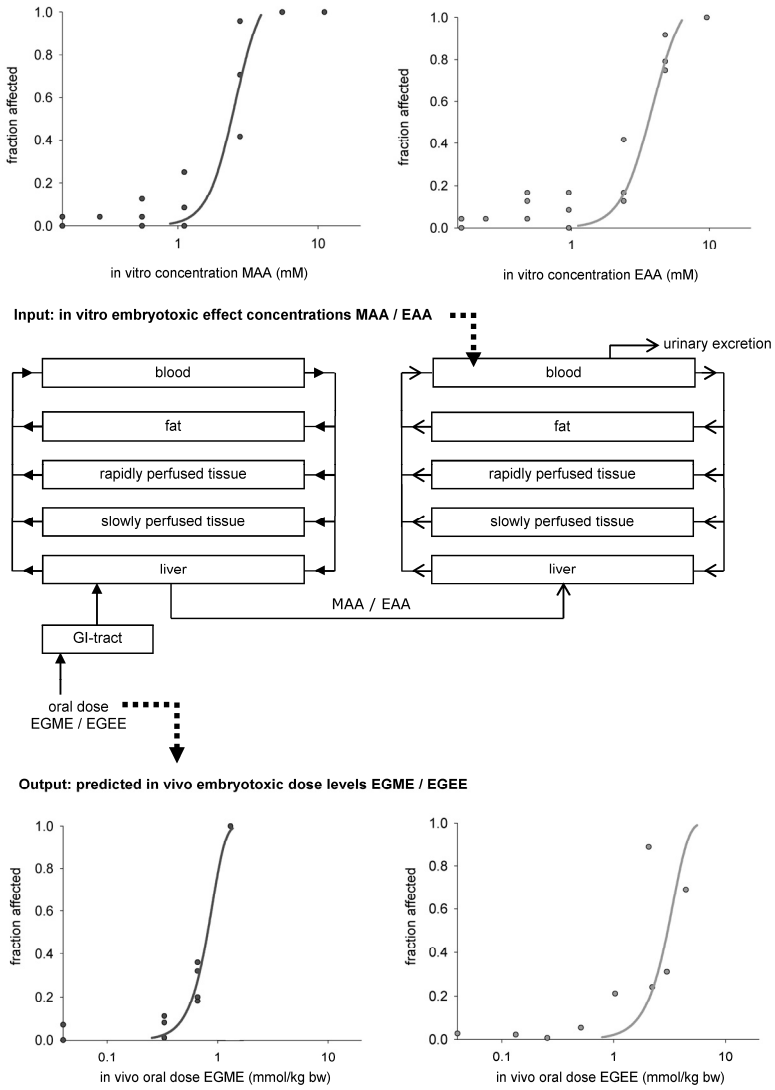

Figure 1. Extrapolation of *in vitro* concentration-response curves for the embryotoxic metabolites methoxyacetic acid (MAA) of ethylene glycol monomethyl ether (EGME) and ethoxyacetic acid (EAA) of ethylene glycol monoethyl ether (EGEE) in the ES-D3 cell differentiation assay to *in vivo* dose-response curves for developmental toxicity of EGME and EGEE using PBK modeling (Louisse *et al.*, 2010). Circles in upper graphs represent *in vitro* toxicity data of MAA and EAA obtained in the ES-D3 cell differentiation assay. The curves fitted to these *in vitro* data were translated into *in vivo* dose-response curves for developmental toxicity of EGME and EGEE using the PBK model. The circles in the lower graphs represent data reported in the literature on the fraction of affected embryos after EGME and EGEE exposure.

The translation of *in vitro* obtained effects to the *in vivo* situation can be hampered by several factors, including 1) the lack or the low expression levels of biotransformation enzymes in *in vitro* test systems compared with the *in vivo* situation, 2) the possible lack of relevant cells or tissues in the *in vitro* test system for the *in vivo* observed effects, 3) the lack of secondary adverse effects resulting from maternal or placental toxicity in the *in vitro* test system and 4) the lack of relevant data needed to construct a PBK model that reliably describes the *in vivo* kinetics, for extrapolating *in vitro* effect concentrations to *in vivo* effect doses. The following sections discuss these factors in more detail.

The lack or low expression levels of biotransformation enzymes

Proteratogens, for example cyclophosphamide and valpromide, need to be bioactivated in order to give rise to their developmental toxicity effects. In *in vitro* test systems, the expression levels of biotransformation enzymes might be too low to bioactivate the parent compounds (Coecke *et al.*, 2006). The addition of a bioactivation system might increase the *in vitro* developmental toxicity potency of a proteratogen. Hettwer *et al.* (2010) showed that the proteratogen cyclophosphamide was 70-fold more potent in the ES-D3 cell differentiation assay of the EST when the medium with cylophosphamide was pre-incubated with murine hepatocytes (preconditioned medium). However, no increase in the *in vitro* potency was obtained for the proteratogen valpromide, using the same experimental procedure, because the levels of the active metabolite (valproic acid) in the culture medium were too low to increase the toxicity (Hettwer *et al.*, 2010). Furthermore, using preconditioned medium will not increase a proteratogen's potency in a test system, when the toxic bioactivation products are unstable (Coecke *et al.*, 2006).

Bioactivation systems have also been added to WEC test systems, in order to improve the predictability for proteratogens. Oglesby *et al.* (1986) showed that cyclophosphamide at concentrations that did not induce embryotoxicity in cultured rat embryos induced embryotoxicity when rat embryos were co-cultured with rat, rabbit or hamster hepatocytes. Mitra *et al.* (1992) showed that the embryotoxic effects of ethylene dibromide were increased in a rat WEC test system, when purified rat liver glutathione S-transferases were added to the culture medium. In another study, it was shown that the embryotoxic potency of

phenytoin was increased in a mouse embryo culture system when embryos were co-cultured with hepatocytes from rats or rabbits, but not from mice (Ozolins *et al.*, 1995). The addition of a bioactivation system to the *in vitro* test system can also decrease the potency of proteratogens, as was shown for trichloroethylene and tetrachloroethylene in cultured rat embryos (Saillenfait *et al.*, 1995). The addition of hepatic microsomal fractions to the culture medium decreased the embryotoxic effects of trichloroethylene and tetrachloroethylene, whereas the main metabolites dichloro- and trichloroacetic acid themselves had a higher potency than the parent compounds (Saillenfait *et al.*, 1995). The addition of a biotransformation system has also been applied in zebrafish developmental toxicity assays. Busquet *et al.* (2008) showed that on addition of induced rat male microsomes and reduced nicotinamide adenine dinucleotide phosphate, the potency of the proteratogens cyclophosphamide and ethanol induced abnormalities in the fish embryos whereas the proteratogens alone did not show an effect.

The examples show that the potency of a proteratogen in an *in vitro* developmental toxicity assay does not always increase with the addition of a bioactivation system. Whether an increase or a decrease in the potency of a proteratogen will be observed with the addition of a bioactivation system, is dependent on the overall outcome of bioactivation and detoxification reactions and the concentration of the toxic metabolite achieved. Therefore, the *in vitro* potency of a proteratogen does not necessarily increase in a test system to which a bioactivation system has been added. This may be dependent on the species from which the bioactivation system originates, implicating that the use of human tissue-derived bioactivation systems may be the most relevant for making predictions for humans. Furthermore, one must take into account that the addition of cellular fractions as a bioactivation system can itself cause toxicity in cellular systems (Tan *et al.*, 1982) and WECs (Luijten *et al.*, 2008).

To investigate whether an unknown chemical could be a proteratogen, the metabolites formed could be identified (e.g., using hepatocytes in culture, or liver fractions (microsomes, S9) and co-factors) and tested (e.g., the main metabolites) in the *in vitro* test system. The *in vivo* formation and internal concentrations of the parent compound and its metabolites can then be estimated using PBK modeling. Subsequently, estimations of parent compound exposure levels resulting in *in vivo*

developmental toxicity can be made. Although metabolite identification (whether or not in combination with metabolite production and testing) is a standard procedure in the drug development process, this might be a too expensive and time-consuming process to apply in routine testing strategies for all chemicals.

The possible lack of relevant cells or tissues in the in vitro test system for the in vivo observed effects

The *in vitro* assays for developmental toxicity described in the literature have varying levels of biological complexity. If the developmental toxicant has an adverse effect on specific cell types, the toxicity might not be detected in an *in vitro* test system in which these cells are absent. This would suggest that WEC assays, which are closer to the *in vivo* situation than cellular assays, would be better test systems to detect adverse effects of a range of chemicals affecting different cell types and with different modes of action (Piersma, 2004; Piersma, 2006). In addition, a battery of cellular assays that would cover the relevant cell types that can be affected during development would detect a developmental toxicant that might not be detected by a single cellular assay (Piersma, 2006). Instead of using cellular assays in which specific cell types are present, embryonic stem cells, which can be differentiated into a mix of relevant cell types, could be used, enabling the assessment of effects on different cell types, and on the differentiation into different cell types, using multiple readouts in one test system. Using differentiating ES-D3 cells may be a useful tool to accomplish this, because although the readout in the classical ES-D3 cell differentiation assay is the inhibition of ES-D3 cell differentiation into contracting cardiomyocytes, cell types of all the three germ layers are present in these cultures (Doetschman *et al.*, 1985; Toumadje *et al.*, 2003).

Toxicogenomic tools are useful to get insight into the mechanism of toxicity of chemicals (Afshari *et al.*, 2011). These technologies may help to assess whether a certain assay contains the relevant cell types to exhibit a specific toxic effect, in order to assess the applicability domain of the test system. Van Dartel *et al.* (2009, 2010a,b, 2011a-c) analyzed gene expression patterns for ES-D3 cells differentiating into cardiomyocytes exposed to chemicals at concentrations affecting cardiomyocyte differentiation. These studies showed the ability to differentiate teratogens from non-teratogens by analysing gene expression patterns in exposed

differentiating embryonic stem cells. Toxicogenomic approaches are also useful to assess whether pathways affected by chemical exposure in the *in vivo* situation are also affected in the *in vitro* test system, in order to determine its predictive capacity (Stigson *et al.*, 2007). However, with the interpretation of toxicogenomic data, one must be aware that gene expression changes themselves are not necessarily adverse. A change in gene expression may reflect an adaptation of the cell or tissue to a situation, not resulting in an adverse effect. Furthermore, a chemical may also have an effect on protein levels or interactions in the cell, resulting in adverse effects, but not directly in a change in gene expression. Proteomic approaches may be useful in assessing effects on protein expression, post-translational modifications and protein interactions (Barrier and Mirkes, 2005). However, the low throughput of proteome analyses may be a major limitation.

The lack of secondary adverse effects resulting from maternal or placental toxicity in the in vitro test system

For some chemicals, developmental toxicity may occur only at dose levels causing maternal toxicity (Chernof *et al.*, 1989; Black *et al.*, 1992; Rogers *et al.*, 2005). Also toxicant-induced alterations in maternal physiology at doses that do not clearly indicate maternal toxicity (such as lethality and/or reduction in maternal body weight) may affect the conceptus's development (Rogers *et al.*, 2005). Furthermore, chemicals affecting the placenta and hence affecting the delivery of oxygen and nutrients to the conceptus may induce developmental toxicity (Maranghi *et al.*, 1998). If this maternal and/or placental toxicity occurs at lower concentrations than the concentrations that would directly affect the conceptus, the chemical's *in vivo* developmental toxicity potency would be underestimated, if effect data of the *in vitro* developmental toxicity assay are translated to the *in vivo* situation, without taking the maternal and/or placental toxicity effects into account. Since maternally induced developmental toxicity effects can be species-dependent (Rogers *et al.*, 2005), relevant human *in vitro* assays assessing maternal and placental toxicity could help in the prediction of secondary developmental toxicity effects of chemicals to humans.

The lack of relevant data needed to construct a PBK model that reliably describes the in vivo kinetics, for extrapolating in vitro effect concentrations to in vivo effect doses

For the extrapolation of *in vitro* effect concentrations to *in vivo* effect doses, PBK models are needed that reliably describe the *in vivo* kinetics. Definition of these PBK models requires robust data for model parameterization. Current PBK models often use chemical-specific parameter values, including tissue:blood partition coefficients and other kinetic parameter values, determined using *in vivo* kinetic data (Rietjens *et al.*, 2011). To limit the use of animals, *in vitro* and *in silico* approaches are preferred to determine these parameter values. However, if extrapolations have to be made for many chemicals, the generation of chemical-specific parameter values can be a labor- and time-consuming process (Rietjens *et al.*, 2011). Furthermore, for some kinetic processes, such as renal clearance, no *in vitro* method is currently available to determine the kinetic parameter values (Adler *et al.*, 2011). The use of QSAR approaches in predicting chemical-specific parameter values may be useful to accelerate the process of building PBK models and hence the use of PBK models for extrapolating *in vitro* effect concentrations to *in vivo* effect doses.

Validation processes of *in vitro* developmental toxicity assays

In order to be useful as reliable alternatives for *in vivo* developmental toxicity studies, *in vitro* developmental toxicity assays need to be validated for their predictive capacities. The validation processes that have been applied for *in vitro* developmental toxicity assays have been described by Piersma (2004). For developmental toxicity, the EST, the MM test and the WEC test have been validated according to the criteria of the ECVAM (Genschow *et al.*, 2002). Recently, Daston *et al.* (2010) reported on the limitations of previous approaches for validating *in vitro* assays for developmental toxicity testing and proposed a new approach of how screening assays for developmental toxicity should be validated in the future. One of the problems that occurred with previous validation studies is the categorization of chemicals into 'positives' (developmental toxicants) and 'negatives' (non-developmental toxicants) without notion of the dose (or concentration), implying that a non-developmental toxicant would never cause developmental toxicity effects, even at very high dose levels. However, also in developmental toxicology, Paracelsus's concept of dose applies (The dose makes the poison), which was restated by Karnofsky (1965) for teratology saying that 'with

the right dose, the right species, and the right timing, any chemical could be shown as a teratogen'. Therefore, Daston *et al.* (2010) defined a developmental toxicant as 'an exposure (agent at a stated internal concentration with stated timing) to the developing organism that leads to a permanent adverse effect' and a non-developmental toxicant as 'an exposure that does not cause permanent adverse effects'. Using these definitions and the internal concentration as key component permits a chemical to be considered as a developmental toxicant at one concentration and as a nondevelopmental toxicant at another concentration (Daston *et al.*, 2010). Thus, in future validation processes for *in vitro* developmental toxicity assays, a list of internal exposures at which *in vivo* developmental toxicity occurs and a list of internal exposures at which *in vivo* developmental toxicity does not occur should be used (Daston *et al.*, 2010). These 'values for internal exposure' could be obtained by assessing the *in vivo* kinetics at dose levels that do not cause developmental toxicity (e.g., at no observed adverse effect levels) and assessing the *in vivo* kinetics at dose levels that cause developmental toxicity (e.g., at lowest observed adverse effect levels). These data on internal concentrations may be acquired in *in vivo* kinetic studies, but should preferably be estimated by developing PBK models for these chemicals. In the validation process, these 'values for internal exposure' can then be compared with toxic effect concentrations obtained in *in vitro* developmental toxicity assays, to assess the predictive capacity of the assays. It must be noted that the concentrations that may cause the developmental toxicity may be dependent on the developmental stage of the conceptus, because physiological changes that occur during gestation may affect chemical availability and/or sensitivity of the conceptus. Therefore, the timing of exposure (both *in vitro* and *in vivo*) has to be taken into account. A single time point measurement may lead to misinterpretation of the obtained data sets. Unfortunately, little data are available on the *in vivo* internal concentrations of chemicals at developmentally toxic doses. If the *in vivo* internal fetal concentrations are to be used as the reference concentrations for the evaluation of *in vitro* developmental toxicity assays, the effect concentrations in different *in vitro* assays would be roughly the same if they would have the same sensitivity and predictive capacity. We evaluated this for the nominal effect concentrations (EC_{50} values) for the inhibition of ES-D3 cell differentiation and the induction of malformations (EC_{50} values) in the WEC test (Figure 2), reported in the ECVAM

validation studies (Genschow *et al.*, 2004; Piersma *et al.*, 2004). Figure 2 shows that the ES-D3 cell differentiation assay was for most chemicals more sensitive than the WEC test, especially for all-*trans*-retinoic acid, 5-fluorouracil and methylmercury chloride, indicated by 80-, 61- and 26-fold lower EC_{50} values, respectively. For the other chemicals, the EC_{50} values for both assays differ less than 10-fold (Figure 2). The results presented in Figure 2 indicate that although the EC_{50} values from the ES-D3 cell differentiation assay tend to be somewhat lower than those for the WEC test, there is a reasonable correlation (r = 0.92). From this it is concluded that the ES-D3 cell differentiation assay and the WEC test would have roughly the same predictive capacity for many of the tested chemicals, but not for all. Differences in

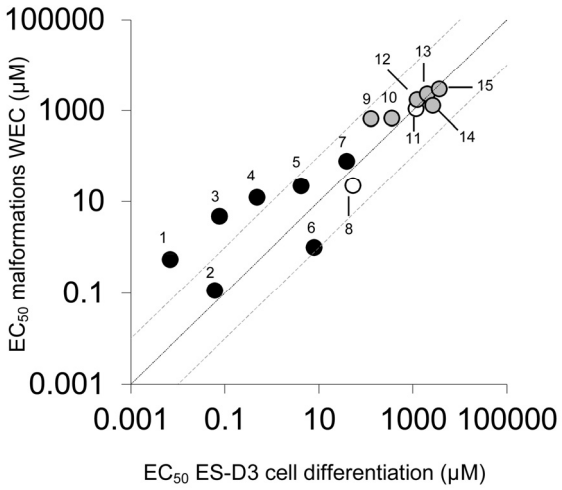

Figure 2. Average EC_{50} values for chemicals affecting ES-D3 cell differentiation (EC_{50} ES-D3 cell differentiation) compared with average EC_{50} values for chemicals inducing malformations in cultured rat embryos (EC_{50} malformations WEC). The short dashed lines indicate EC_{50} ES-D3 cell differentiation = EC_{50} malformations WEC. The long dashed line covers the area in which the EC_{50} ES-D3 cell differentiation values do not differ more than 10-fold from the EC_{50} malformations WEC values. Average EC_{50} values were calculated from data reported by Genschow *et al.* (2004) and Piersma *et al.* (2004). 1: All-*trans*-retinoic acid; 2: methotrexate; 3: 5-fluorouracil 4: methylmercury; 5: 5-bromo-2'-deoxyuridine; 6: 6-amino-nicotinamide; 7: hydroxyurea; 8: diphenhydramine hydrochloride; 9: pentyl-4-yn-valproic acid; 10: valproic acid; 11: acrylamide; 12: salicylic acid sodium salt; 13: methoxyacetic acid; 14: boric acid; 15: lithium chloride. Black circles: chemicals classified as strong embryotoxic in ECVAM validation studies. Grey circles: chemicals classified as weak embryotoxic in ECVAM validation studies. White circles: chemicals classified as non-embryotoxic in ECVAM validation studies.

the obtained effect concentrations of the same chemical in different *in vitro* test systems may be caused by several factors, including the following.

Differences in sensitivities of the test systems

Chemicals may induce developmental toxicity by affecting not only general processes, such as cell proliferation, cell viability and/or maintenance of intracellular pH, but also specific processes during development, such as neuron outgrowth or vasculogenesis. It can be expected that effect concentrations obtained in different *in vitro* assays for developmental toxicity will not differ considerably for chemicals with a nonspecific mode of action, whereas for chemicals with a more complex or specific mode of action, effect concentrations obtained in different *in vitro* assays can differ much more. If an *in vivo* affected process is not represented in the *in vitro* assay, effect concentrations will represent rather nonspecific toxicity appearing at higher concentrations than the concentrations that would affect the specific process. Thus, if a chemical exerts its effect via a specific mechanism, this may not be detected in an oversimplified test system.

Differences in exposure characteristics of the test protocols

In the WEC test, embryos are exposed for 2 days to the chemical, whereas in the ES-D3 cell differentiation assay, the total period of exposure is 10 days, in which the exposure medium is renewed twice. For all-*trans*-retinoic acid, for example, the developmental toxicity outcome is reported to be related to the area under the concentration-time curve, and not to the peak concentration (Tzimas *et al.*, 1997). Therefore, the longer exposure time in the ES-D3 cell differentiation assay might play a role in the lower obtained effect concentrations measured in this *in vitro* assay for all-*trans*-retinoic acid. Furthermore, constituents of culture media could influence the chemical's toxic potency. The high sensitivity of the ES-D3 cell differentiation assay for 5-fluorouracil could be due to the fact that 5-fluorouracil exerts its toxic effect by the formation of an inhibitory complex with thymidylate synthetase and folate, which blocks DNA synthesis, as was discussed before by Verwei *et al.* (2006). Because of the relatively high folate concentrations in ES-D3 cell culture medium compared with serum, more inhibitory complexes may be formed resulting in a higher potency of 5-FU in the ES-D3 cell differentiation assay than in the WEC test and in the *in vivo* situation (Verwei *et al.*, 2006).

Differences in culture media may also result in differences in the free concentration of the chemical. This will depend on the extent of protein binding of the chemicals, which may be species-dependent, and the different amounts of proteins present in the test systems (Gülden *et al.*, 2006; Blaauboer, 2010). These differences may also partly explain the higher sensitivity of the ES-D3 cell differentiation assay for all-*trans*-retinoic acid and methylmercury, compared with the WEC test, because the medium used for the WEC test (rat serum) contains higher levels of protein than the medium used for the ES-D3 cell differentiation assay (20% FCS). This also indicates that the use of different concentrations of serum in assays may lead to interlaboratory differences in nominal effect concentrations obtained. Therefore, free effect concentrations in an *in vitro* test system would preferably be compared with the free *in vivo* internal concentrations at developmental toxicity dose levels, instead of comparing nominal effect concentration in the *in vitro* assay with nominal concentrations *in vivo*.

Differences in biotransformation systems in the in vitro test systems

The expression of biotransformation enzymes may differ in the different test systems or may be completely absent. This can lead to differences in detoxification and/or bioactivation between test systems, and, therefore, differences in effect concentrations obtained.

Expert opinion: toward the prediction of relevant *in vivo* internal toxic effect concentrations using *in vitro* assays for developmental toxicity and their subsequent extrapolation to *in vivo* dose levels

For the development of *in vitro* developmental toxicity assays, knowledge on the mechanisms underlying developmental toxicity will be helpful to assess whether a certain *in vitro* assay would be suitable to detect a specific toxic effect. The use of toxicogenomic approaches is useful to assess whether in the *in vitro* assay similar pathways are affected as compared with the *in vivo* situation on exposure to a developmental toxicant (Stigson *et al.*, 2007). It is unlikely that a single *in vitro* assay can represent all the processes that may be affected by chemicals during *in vivo* developmental processes (Piersma, 2006). Therefore, the combination of a group of assays that together represent the processes that can be affected in the

developmental process is needed (Piersma, 2006). A chemical that is a neurodevelopmental toxicant, for example, may not have an effect at a relevant concentration in the ES-D3 cell differentiation assay, in which the effect on differentiation into cardiac cells is assessed. Because it is not known for a new chemical which developmental process it may affect, it should be tested in a battery of assays. The assay in which the lowest effect concentration is obtained for that chemical may be the most relevant (the first) process that will be affected *in vivo*. Using PBK modeling, it can then be estimated whether these concentrations are expected to be reached *in vivo* at certain exposure levels.

In validation processes, it should be evaluated whether the concentrations at which toxicity is observed are relevant considering the *in vivo* internal concentrations at developmental toxicity dose levels of the chemicals (Daston *et al.*, 2010). This would imply that instead of a list of standard chemicals classified as either developmental toxicants (weak or strong) or nondevelopmental toxicants, a list of internal exposures should be used in the evaluation and validation of *in vitro* developmental toxicity assays (Daston *et al.*, 2010). The comparison of the effect concentrations obtained in the *in vitro* assays with internal concentrations at developmentally toxic dose levels will give insight into which of the assays provide the most relevant toxic effect concentrations. Parameters such as dose (concentration), timing and duration of exposure should be taken into account in these comparisons. If an *in vitro* developmental toxicity assay proves to predict relevant *in vitro* effect concentrations (e.g., for a certain group of chemicals), its applicability domain can be determined and the assay can be used to predict the *in vivo* developmental toxicity potencies of the chemicals by the obtained *in vitro* ranking using a category approach. However, in such a category approach, important kinetic considerations, such as differences in absorption, detoxification and excretion, may be overlooked. These kinetic considerations are accounted for when the *in vitro* effect concentrations are translated into *in vivo* dose levels using PBK modeling, enabling the use of *in vitro* data in a quantitative manner, instead of using *in vitro* data only for hazard identification. The development of PBK models may, therefore, become an important factor in the process of the development of alternatives to animal experiments in toxicological research. This requires skill competence of researchers in programming and modeling, and the development of

in vitro or *in silico* tools that can be used to determine robust kinetic parameter values needed to construct these PBK models. With the combined *in vitro-in silico* approach, *in vitro* toxicity data may be used to set points of departure for risk assessment (like a $BMDL_{10}$ (lower limit of the 95% confidence interval on the benchmark dose at which a benchmark response equivalent to a 10% effect size (BMR_{10}) is reached (BMD_{10})), enabling derivation of safe exposure limits for humans using animal-free approaches.

References

Adler, S., Basketter, D., Creton, S., Pelkonen, O., van Benthem, J., Zuang, V., Andersen, K. E., Angers-Loustau, A., Aptula, A., Bal-Price, A., Benfenati, E., Bernauer, U., Bessems, J. G. M., Bois, F. Y., Boobis, A., Brandon, E., Bremer, S., Broschard, T., Casati, S., Coecke, S., Corvi, R., Cronin, M., Daston, G. P., Dekant, W., Felter, S., Grignard, E., Gundert-Remy, U., Heinonen, T., Kimber, I., Kleinjans, J., Komulainen, H., Kreiling, R., Kreysa, J., Leite, S. B., Loizou, G., Maxwell, G., Mazzatorta, P., Munn, S. J., Pfuhler, S., Phrakonkham, P., Piersma, A. H., Poth, A., Prieto, P., Repetto, G., Rogiers, V., Schoeters, G., Schwarz, M., Serafimova, R., Tähti, H., Testai, E., van Delft, J. H., van Loveren, H., Vinken, M., Worth, A., Zaldivar, J. M. (2011). Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Arch Toxicol* **85**, 367-485.

Afshari, C. A., Hamadeh, H. K., Bushel, P. R. (2011). The evolution of bioinformatics in toxicology: advancing toxicogenomics. *Toxicol Sci* **120**, S225-37.

Bantle, J. A., Finch, R. A., Fort, D. J., Stover, E. L., Hull, M., Kumsher-King, M., Gaudet-Hull, A. M. (1999). Phase III interlaboratory study of FETAX Part 3. FETAX validation using 12 chemicals with and without an exogenous metabolic activation system. *J Appl Toxciol* **19**, 447-72.

Barrier, M., Mirkes, P. E. (2005). Proteomics in developmental toxicology. *Reprod Toxicol* **19**, 291-304.

Biomarkers definitions working group (2001). Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* **69**, 89-95.

Blaauboer, B. J. (2010). Biokinetic modeling and in vitro-in vivo extrapolations. *J Toxicol Environ Health B Crit Rev* **3**, 242-52.

Black, D. L., Marks, T. A. (1992). Role of maternal toxicity in assessing developmental toxicity in animals: a discussion. *Regul Toxicol Pharmacol* **16**, 189-201.

Brown, N. A., Fabro, S. (1981). Quantitation of rat embryonic development in vitro: a morphological scoring system. *Teratology* **24**, 65-78.

Busquet, F., Nagel, R., Von Landenberg, F., Mueller, S. O., Huebler, N., Broschard, T. H. (2008). Development of a new screening assay to identify proteratogenic substances using zebrafish danio rerio embryo combined with exogenous mammalian metabolic activation system (m*DAR*T). *Toxicol Sci* **104**, 177-188.

Chan, L. Y., Chiu, P. Y., Siu, N. S., Wang, S. S., Lau, T. K. (2002). Diclofenac-induced embryotoxicity is associated with increased embryonic 8-isoprostaglandin F2alpha level in rat whole embryo culture. *Reprod Toxicol* **16**, 841-44.

Chen, B., Hales, B. F. (1994). Cadmium-induced rat embryotoxicity *in vitro* is associated with an

increased abundance of E-cadherin protein in the yolk sac. *Toxicol Appl Pharmacol* **128**, 293-301.

Chernof, N., Rogers, J. M., Kavlock, R. J. (1989). An overview of maternal toxicity and prenatal development: considerations for developmental toxicity hazard assessment. *Toxicology* **59**, 111-25.

Coecke, S., Ahr, H., Blaauboer, B. J., Bremer, S., Casati, S., Castell, J., Combes, R., Corvi, R., Crespi, C. L., Cunningham, M. L., Elaut, G., Eletti, B., Freidig, A. P., Gennari, A., Ghersi-Egea, J. F., Guillouzo, A., Hartung, T., Hoet, P., Ingelman-Sundberg, M., Munn, S. J., Janssens, W., Ladstetter, B., Leahy, D., Long, A., Meneguz, A., Monshouwer, M., Morath, S., Nagelkerke, F., Pelkonen, O., Ponti, J., Prieto, P., Richert, L., Sabbioni, E., Schaack, B., Steiling, W., Testai, E., Vericat, J. A., Worth, A. (2006). Metabolism: a bottleneck in *in vitro* toxicological test development. The report and recommendations of ECVAM workshop 54. *Altern Lab Anim* **34**, 49-84.

Daston, G. P., Chapin, R. E., Scialli, A. R., Piersma, A. H., Carney, E. W., Rogers, J. M., Friedman. J. M. (2010). A different approach to validating screening assays for developmental toxicity. *Birth Defects Res B Dev Reprod Toxicol* **89**, 526-30.

Davies, W. J., Freeman, S. J. (1995). Chick embryotoxicity screening test (CHEST I and II). *Methods Mol Biol* **43**, 307-10.

De Jong, E., Louisse, J., Verwei, M., Blaauboer, B. J., van der Sandt, J. J. M., Woutersen, R. A., Rietjens, I. M. C. M., Piersma, A. H. (2009). Relative developmental toxicity of glycol ether alkoxy acid metabolites in the embryonic stem cell test as compared to the *in vivo* potency of their parent compounds. *Toxicol Sci* **110**, 117-24.

De Jong, E., Doedée, A. M., Reis-Fernandes, M. A., Nau, H., Piersma, A. H. (2011a). Potency ranking of valproic acid analogues as to inhibition of cardiac differentiation of embryonic stem cells in comparison to their *in vivo* embryotoxicity. *Reprod Toxicol* **31**, 375-82.

De Jong, E., Barenys, M., Hermsen, S. A. B., Verhoef, A., Ossendorp, B. C., Bessems, J. G. M., Piersma, A. H. (2011b). Comparison of the mouse embryonic stem cell test, the rat whole embryo culture and the zebrafish embryotoxicity test as alternative methods for developmental toxicity testing of six 1,2,4-triazoles. *Toxicol Appl Pharmacol* **253**, 103-11.

DeJongh, J., Nordin-Andersson, M., Ploeger, B. A., Forsby A. (1999). Estimation of systemic toxicity of acrylamide by integration of in vitro toxicity data with kinetic simulations. *Toxicol Appl Pharmacol* **158**, 261-68.

DeMicco, A., Cooper, K. R., Richardson, J. R., White, L. A. (2010). Developmental neurotoxicity of pyrethroid insecticides in zebrafish embryos. *Toxicol Sci* **113**, 177-86.

Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W., Kemler, R. (1985). The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J Embryol Exp Morph* **87**, 27-45.

Enoch, S. J., Cronin, M. T. D., Madden, J. C., Hewitt, M. (2009). Formation of structural categories to allow for read-across for teratogenicity. *QSAR Comb Sci* **28**, 696-708.

Fabjan, E., Hulzebos, E., Mennes, W., Piersma, A. H. (2006). A category approach for reproductive effects of phthalates. *Crit Rev Toxicol* **36**, 695-726.

Fan, C. Y., Cowden, J., Simmons, S. O., Padilla, S., Ramabhadran, R. (2010). Gene expression changes in developing zebrafish as potential markers for rapid developmental neurotoxicity screening. *Neurotoxicol Teratol* **32**, 91-98.

Fesstag, M., Sehner, C., Steinberg, P., Viertel, B. (2007). An *in vitro* embryotoxicity assay based on the disturbance of the differentiation of murine embryonic stem cells into endothelial cells. II: testing of chemicals. *Toxicol In Vitro* **21**, 1631-40.

Flint, O. P. (1982). A micromass culture method for rat embryonic neural cells. *J Cell Sci* **61**, 247-62.

Flint, O. P., Orton, T. C. (1984). An *in vitro* assay for teratogens with cultures of rat embryo midbrain and limb bud cells. *Toxicol Appl Pharmacol* **76**, 383-95.

Forsby, A., Blaauboer, B. J. (2007). Integration of in vitro neurotoxicity data with biokinetic modelling for the estimation of in vivo neurotoxicity. *Hum Exp Toxicol* **26**, 333-38.

Fort, D. J., Paul, R. R. (2002). Enhancing the predictive validity of frog embryo teratogenesis assay-Xenopus (FETAX). *J Appl Toxicol* **22**, 185-91.

Genschow, E., Scholz, G., Brown, N., Piersma, A. H., Brady, M., Clemann, N., Huuskonen, H., Paillard, F., Bremer, S., Becker, K., Spielmann, H. (2000). Development of prediction models for three *in vitro* embryotoxicity tests in an ECVAM validation study. *In Vitr Mol Toxicol* **13**, 51-66.

Genschow, E., Spielmann, H., Scholz, G., Seiler, A., Brown, N. A., Piersma, A. H., Brady, M., Clemann, N., Huuskonen, H., Paillard, F., Bremer, S., Becker, K. (2002). The ECVAM international validation study on in vitro embryotoxicity tests: results of the definitive phase and evaluation of prediction models. European Centre for the Validation of Alternative Methods. *Atern Lab Anim* **30**, 151-76.

Genschow, E., Spielmann, H., Scholz, G., Pohl, I., Seiler, A., Clemann, N., Bremer, S., Becker, K. (2004). Validation of the embryonic stem cell test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* **32**, 209-44.

Giavini, E., Broccia, M. L., Menegola, E. , Prati, M. (1993). Comparitive in vitro study of the embryotoxic effects of three glycol ethers and their metabolites, the alkoxyacids. *Toxicol In Vitro* **7**, 777-84.

Gornati, R., Monetti, C., Vigetti, D., Bosisio, S., Sabbionim E., Bernardini, G., Prati, M. (2002). Arsenic toxicity and HSP70 expression in Xenopus laevis embryos. *Altern Lab Anim* **30**, 597-603.

Groebe, K., Hayess, K., Klemm-Manns, M., Schwall, G., Wozny, W., Steemans, M., Peters, A. K., Sastri, C., Jaeckel, P., Stegmann, W., Zengerling, H., Schöpf, R., Poznanovic, S., Stummann, T. C., Seiler, A., Spielmann, H., Schrattenholz, A. (2010a). Unexpected commons mechanistic pathways for embryotoxicity of warfarin and lovastatin. *Reprod Toxicol* **30**, 121-30.

Groebe, K., Hayess, K., Klemm-Manns, M., Schwall, G., Wozny, W., Steemans, M., Peters, A. K., Sastri, C., Jaeckel, P., Stegmann, W., Zengerling, H., Schöpf, R., Poznanovic, S., Stummann, T. C., Seiler, A., Spielmann, H., Schrattenholz, A. (2010b). Protein biomarkers for in vitro testing of embryotoxicity. *J Proteome Res* **9**, 5727-38.

Gülden, M., Dierickx, P., Seibert, H. (2006). Validation of a prediction model for estimating serum concentrations of chemicals which are equivalent to toxic concentrations in vitro. *Toxicol In Vitro* **20**, 1114-24.

Hermsen, S. A. B., Pronk, T. E., van den Brandhof, E. J., van der Ven, L. T. M., Piersma, A. H. (2011a). Chemical class-specific gene expression changes in the zebrafish embryo after exposure to glycol ether alkoxy acids and 1,2,4-triazole antifungals. *Reprod Toxicol* **32**, 245-52.

Hermsen, S. A. B., van den Brandhof, E. J., van der Ven, L. T. M., Piersma, A. H. (2011b). Relative embryotoxicity of two classes of chemicals in a modified zebrafish embryotoxicity test and comparison with their in vivo potencies. *Toxicol In Vitro* **25**, 745-53.

Hettwer, M., Reis-Fernandes, M. A., Iken, M., Ott, M., Steinberg, P., Nau, H. (2010). Metabolic activation capacity by primary hepatocytes expands the applicability of the embryonic stem cell test as alternative to experimental animal testing. *Reprod Toxicol* **30**, 113-20.

Heuer, J., Bremer, S., Pohl, I., Spielmann, H. (1997). Development of an in vitro embryotoxicity test using murine embryonic stem cell cultures. *Toxicol In Vitro* **7**, 551-56.

Jagtap, S., Meganathan, K., Gaspar, J., Wagh, V., Winkler, J., Hescheler, J., Sachinidis, A. (2011). Cytosine arabinoside induces ectoderm and inhibits mesoderm expression in human embryonic stem cells during multilineage differentiation. *Br J Pharmacol* **162**, 1743-56.

Janer, G., Verhoef, A., Piersma, A. H. (2008). Use of the rat postimplantation embryo culture to assess the embryotoxic potency within a chemical catergory and to identify toxic metabolites. *Toxicol In Vitro* **22**, 1797-805.

Jelinek, R. (1977). The chick embryotoxicity screening test (CHEST). In Neubert, D., Merker, H. J., Kwasigroch, T. E. (eds). *Methods in prenatal toxicology*. Thieme, Stuttgart, 381-86.

Jergil, M., Forsberg, M., Salter, H., Stockling, K., Gustafson, A. L., Dencker, L., Stigson, M. (2011). Short-time gene expression response to valproic acid and valproic acid analogs in mouse embryonic stem cells. *Toxicol Sci* **121**, 328-42.

Karnofsky, D. A. (1965). Drugs as teratogens in animals and man. *Annu Rev Pharmacol* **10**, 447-72.

Kistler, A. (1987). Limb bud cell cultures for estimating the teratogenic potential of chemicals. Validation of the test system with retinoids. *Arch Toxicol* **60**, 403-14.

Klug, S., Lewandowski, C., Neubert, D. (1985). Modification and standardization of the culture of early postimplantation embryos for toxicological studies. *Arch Toxicol* **58**, 84-8.

Krishnan, K., Andersen, M. E. (2001). Physiologically-based pharmacokinetic modeling and toxicology. In Hayes, A. W. (ed). *Principles and methods of toxicology*. Raven Press, NY, 193-241.

Kuegler, P. B., Zimmer, B., Waldmann, T., Baudis, B., Ilmjärv, S., Hescheler, J., Gaughwin, P., Brundin, P., Mundy, W., Bal-Price, A. K., Schrattenholz, A., Krause, K. H., van Thriel, C., Rao, M. S., Kadereit, S., Leist, M. (2010). Markers of murine embryonic and neural stem cells, neurons and astrocytes: reference points for developmental neurotoxicity testing. *ALTEX* **27**, 17-42.

Louisse, J., de Jong, E., van de Sandt, J. J. M., Blaauboer, B. J., Woutersen, R. A., Piersma, A. H., Rietjens, I. M. C. M., Verwei, M. (2011). The use of in vitro toxicity data and physiologically based kinetic modeling to predict dose-response curves for in vivo developmental toxicity of glycol ethers in rat and man. *Toxicol Sci* **118**, 470-84.

Louisse, J., Gönen, S., Rietjens, I. M. C .M., Verwei, M. (2011). Relative developmental toxicity potencies of retinoids in the embryonic stem cell test compared with their relative potencies in in vivo and two other in vitro assays for developmental toxicity. *Toxicol Lett* **203**, 1-8.

Luijten, M., Verhoef, A., Westerman, A., Piersma, A. H. (2008). Application of a metabolizing system as an adjunct to the rat whole embryo culture. *Toxicol In Vitro* **22**, 1332-36.

Luijten, M., van Beelen, V. A., Verhoef, A., Renkens, M. F., van Herwijnen, M. H., Westerman, A., van Schooten, F. J., Pennings, J. L., Piersma, A. H. (2010). Transcriptomics analysis of retinoic acid embryotoxicity in rat postimplantation whole embryo culture. *Reprod Toxicol* **30**, 333-40.

Maranghi, F., Macrì, C., Ricciardi, C., Stazi, A. V., Mantovani, A. (1998). Evaluation of the placenta: suggestions for a greater role in developmental toxicology. *Adv Exp Med Biol* **444**, 129-36.

Marx-Stoelting, P., Adriaens, E., Ahr, H. J., Bremer, S., Garthoff, B., Gelbke, H. P., Piersma, A. H., Pellizzer, C., Reuter, U., Rogiers, V., Schenk, B., Schwengberg, S., Seiler, A., Spielmann, H., Steemans, M., Stedman, D. B., Vanparys, P., Vericat, J. A., Verwei, M., van de Water, F., Weimer, M., Schwarz, M. (2009). A review of the implementation of the embryonic stem cell test (EST), the report and recommendation of an ECVAM/ReProTect workshop. *Atern Lab Anim* **37**, 313-28.

Mirkes, P. E., Doggett, B., Cornel, L. (1994). Induction of a heat shock response (HSP 72) in rat embryos exposed to selected chemical teratogens. *Teratology* **49**, 135-42.

Mitra, A., Hilbelink, D. R., Dwornikk, J. J., Kulkarni, A. (1992). Rat hepatic glutathione S-transferase-mediated embryotoxic bioactivation of ethylene dibromide. *Teratology* **46**, 439-46.

Mouche, I., Malesic, L., Gillardeaux, O. (2011). FETAX assay for evaluation of developmental toxicity. *Methods Mol Biol* **691**, 257-69.

Murabe, M., Yamauchi, J., Fujiwara, Y., Hiroyama, M., Sanbe, A., Tanoue, A. (2007a). A novel embryotoxic estimation method of VPA using ES cells differentiation system. *Biophys Biochem Res Commun* **352**, 164-69.

Murabe, M., Yamauchi, J., Fujiwara, Y., Miyamoto, Y., Hiroyama, M., Sanbe, A., Tanoue, A. (2007b). Estimation of the embryotoxic effect of CBZ using an ES cell differentiation system. *Biophys Biochem Res Commun* **356**, 739-44.

Newall, D. R., Beedles, K. E. (1994). The stem-cell test-a novel in vitro assay for teratogenic potential. *Toxicol In Vitro* **8**, 697-701.

Oglesby, L. A., Ebron, M. T., Beyer, P. E., Carver, B. D., Kavlock, R. J. (1986). Co-culture of rat embryos and hepatocytes: in vitro detection of a proteratogen. *Teratog Carcinog Mutagen* **6**, 129-38.

Osman, A. M., van Dartel, D. A. M., Zwart, E., Blokland, M., Pennings, J. L., Piersma, A. H. (2010). Proteome profiling of mouse embryonic stem cells to define markers for cell differentiation and embryotoxicity. *Reprod Toxciol* **30**, 322-32.

Ozolins, T. R., Oglesby, L. A., Wiley, M. J., Wells, P. G. (1995). In vitro murine embryotoxicity of cyclophosphamide in embryos co-cultured with maternal hepatocytes: development and application of a murine embryo-hepatocyte co-culture model. *Toxicology* **102**, 259-74.

Pamies, D., Vicente-Salar, N., Sogorb, M. A., Roche, E., Reig, J. A. (2010). Specific effect of 5-fluorouracil on alpha-fetoprotein gene expression during the in vitro mouse embryonic stem cell differentiation. *Int J Toxicol* **29**, 297-304.

Paparella, M., Kolossov, E., Fleischmann, B. K., Hescheler, J., Bremer, S. (2002). The use of quantitative image analysis in the assessment of in vitro embryotoxicity endpoints based on a novel embryonic stem cell clone with endoderm-related GFP expression. *Toxicol In Vitro* **16**, 589-97.

Paquette, J. A., Kumpf, S. W., Streck, R. D., Thomson, J. J., Chapin, R. E., Stedman, D. B. (2008). Assessment of the embryonic stem cell test and application and use in the pharmaceutical industry. *Birth Defects Res B Dev Reprod Toxicol* **83**, 104-11.

Piersma, A. H., Attenon, P., Bechter, R., Govers, M. J., Krafft, N., Schmid, B. P., Stadler, J., Verhoef. A., Verseil, C. (1995). Interlaboratory evaluation of embryotoxicity in the postimplantation rat embryo culture. *Reprod Toxicol* **9**, 275-80.

Piersma, A. H. (2004). Validation of alternative methods for developmental toxicity testing. *Toxicol Lett* **149**, 147-53.

Piersma, A. H., Genschow, E., Verhoef, A., Spanjersberg, M. Q., Brown, N. A., Brady, M., Burns. A., Clemann, N., Seiler, A., Spielmann, H. (2004). Validation of the postimplantation rat whole-embryo culture test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* **32**, 275-307.

Piersma, A. H. (2006). Alternative methods for developmental toxicity testing. *Basic Clin Pharmacol Toxicol* **98**, 427-31.

Piparo, E. L., Worth, A. (2010). Review of QSAR models and software tools for predicting developmental and reproductive toxicity. *Report EUR* **24522**.

Riebeling, C., Pirow, R., Becker, K., Buesen, R., Eikel, D., Kaltenhäuser, J., Meyer, F., Nau, H., Slawik, B.,

Visan, A., Volland, J., Spielmann, H., Luch, A., Seiler, A. (2011). The embryonic stem cell test as tool to assess structure-dependent teratogenicity: the case of valproic acid. *Toxicol Sci* **120**, 360-70.

Rietjens, I. M. C. M., Louisse, J., Punt, A. (2011). Tutorial on physiologically based kinetic modeling in molecular nutrition and food research. *Mol Nutr Food Res* **55**, 941-56.

Robinson, J. F., van Beelen, V. A., Verhoef, A., Renkens, M. F., Luijten, M., van Herwijnen, M. H., Westerman, A., Pennings, J. L., Piersma, A. H. (2010). Embryotoxicant-specific transcriptomic responses in rat postimplantation whole-embryo culture. *Toxicol Sci* **118**, 675-85.

Rogers, J. M., Chernogg, N., Keen, C. L., Daston, G. P. (2005). Evaluation and interpretation of maternal toxicity in Segment II studies: Issues, some answers, and data needs. *Toxicol Appl Pharmacol* **207**, S367-74.

Saillenfait, A. M., Langonné, I., Sabaté, J. P. (1995). Developmental toxicity of trichloroethylene, tetrachloroethylene and four of their metabolites in rat whole embryo culture. *Arch Toxicol* **70**, 71-82.

Schmid, B. P., Trippmacher, A., Bianchi, A. (1983). Validation of the whole-embryo culture method for in vitro teratogenicity testing. *Dev Toxicol Environ Sci* **11**, 563-6.

Scholz, G., Genschow, E., Pohl, I., Bremer, S., Paparella, M., Raabe, H., Southee, J., Spielmann, H. (1999). Prevalidation of the embryonic stem cell test (EST)-a new in vitro embryotoxicity test. *Toxicol In Vitro* **13**, 675-81.

Scialli, A. R. (2008). The challenge of reproductive and developmental toxicology under REACH. *Regul Toxicol Pharmacol* **51**, 244-50.

Seiler, A., Visan, A., Buesen, R., Genschow, E., Spielmann, H. (2004). Improvement of an in vitro stem cell assay for developmental toxicity: the use of molecular endpoints in the embryonic stem cell test. *Reprod Toxicol* **18**, 231-40.

Selderslaghs, I. W. T., van Rompay, A. R., de Coen, W., Witters, H. E. (2009). Development of a screening assay to identify teratogenic and embryotoxic chemicals using the zebrafish embryo. *Reprod Toxicol* **28**, 308-20.

Selderslaghs, I. W. T., Hooyberghs, J., de Coen, W., Witters, H. E. (2010). Locomotor activity in zebrafish embryos: a new method to assess developmental neurotoxicity. *Neurotoxicol Teratol* **32**, 460-71.

Selderslaghs, I. W. T., Blust, R., Witters, H. E. (2012). Feasibility study of the zebrafish assay as an alternative method to screen for developmental toxicity and embryotoxicity using a training set of 27 chemicals. *Reprod Toxicol* **33**, 142-54.

Spielmann, H., Genschow, E., Brown, N. A., Piersma, A. H., Verhoef, A., Spanjersberg, M. Q., Huuskonen, H., Paillard, F., Seiler, A. (2004). Validation of the rat limb bud micromass test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* **32**, 245-74.

Stachlewitz, R. F. (2007). Bringing biomarkers to the masses: an elegant approach to discover potential biomarkers of vascular injury. *Toxicol Sci* **100**, 5-6.

Stigson, M., Kultima, K., Jergil, M., Scholz, B., Alm, H., Gustafson, A. L., Dencker, L. (2007). Molecular targets and early response biomarkers for the prediction of developmental toxicity in vitro. *Altern Lab Anim* **35**, 335-42.

Stumann, T. C., Hareng, L., Bremer, S. (2007). Embryotoxicity hazard assessment of methylmercury and chromium using embryonic stem cells. *Toxicology* **242**, 130-43.

Tan, E. L., Schenley, R. L., Hsie, A. W. (1982). Microsome-mediated cytotoxicity to CHO cells. *Mutat Res*

103, 359-65.

Theunissen, P. T., Schulpen, S. H. W., van Dartel, D. A. M., Hermsen, S. A., van Schooten, F. J., Piersma, A. H. (2010). An abbreviated protocol for multilineage neural differentiation of murine embryonic stem cells and its perturbation by methyl mercury. *Reprod Toxicol* **29**, 383-92.

Theunissen, P. T., Pennings, J. L. A., Robinson J. F., Claessen, S. M., Kleinjans, J. C., Piersma, A. H. (2011). Time-response evaluation by transcriptomics of methylmercury effects on neural differentiation of murine embryonic stem cells. *Toxicol Sci* **122**, 437-47.

Thomson, J., Johnson, K., Chapin, R., Stedman, D., Kumpf, S., Ozolins, T. R. (2011). Not a walk in the park: the ECVAM whole embryo culture model challenged with pharmaceuticals and attempted improvements with random forest design. *Birth Defects Res B Dev Reprod Toxicol* **92**, 111-21.

Toumadje, A., Kusumoto, K., Parton, A., Mericko, P., Dowell, L., Ma, G., Chen, L., Barnes, D. W., Sato, J. D. (2003). Pluripotent differentiation in vitro of murine ES-D3 embryonic stem cells. *In vitro Cell Dev Biol Anim* **39**, 449-53.

Tzimas, G., Thiel, R., Chahoud, I., Nau, H. (1997). The area under the concentration-time curve of all-trans-retinoic acid is the most suitable pharmacokinetic correlate to the embryotoxicity of this retinoid in the rat. *Toxicol Appl Pharmacol* **143**, 436-44.

Van Dartel, D. A. M., Pennings, J. L. A., Hendriksen, P. J. M., van Schooten, F. J., Piersma, A. H. (2009). Early gene expression changes during embryonic stem cell differentiation into cardiomyocytes and their modulation by monobuthyl phthalate. *Reprod Toxicol* **27**, 93-102.

Van Dartel, D. A. M., Pennings, J. L. A., van Schooten, F. J., Piersma, A. H. (2010a). Transcriptomics-based identification of developmental toxicants through their interference with cardiomyocyte differentiation of embryonic stem cells. *Toxicol Appl Pharmacol* **243**, 420-28.

Van Dartel, D. A. M., Pennings, J. L. A., de la Fonteyne, L. J. J., van Herwijnen, M. H., van Schooten, F., J., Piersma, A. H. (2010b). Monitoring developmental toxicity in the embryonic stem cell test using differential gene expression of differentiation-related genes. *Toxicol Sci* **116**, 130-39.

Van Dartel, D. A. M., Pennings, J. L. A., de la Fonteyne, L. J. J., Brauers, K. J., Claessen, S., van Delft, J. H., Kleinjans, J. C., Piersma, A. H. (2011a). Evaluation of developmental toxicant identification using gene expression profiling in embryonic stem cell differentiation cultures. *Toxicol Sci* **119**, 126-34.

Van Dartel, D. A. M., Pennings, J. L. A., de la Fonteyne, L. J. J., Brauers, K. J., Claessen, S., van Delft, J. H., Kleinjans, J. C., Piersma, A. H. (2011b). Concentration-dependent gene expression responses to flusilazole in embryonic stem cell differentiation cultures. *Toxicol Appl Pharmacol* **251**, 110-18.

Van Dartel, D. A. M., Pennings, J. L. A., Robinson, J. F., Kleinjans, J. C., Piersma, A. H. (2011c). Discriminating classes of developmental toxicants using gene expression profiling in the embryonic stem cell test. *Toxicol Lett* **201**, 143-51.

Van Dartel, D. A. M., Piersma, A. H. (2011). The embryonic stem cell test combined with toxicogenomics as an alternative testing model for the assessment of developmental toxicity. *Reprod Toxicol* **32**, 235-44.

Van der Jagt, K., Munn, S. J., Tørsløv, J., de Bruijn, J. (2004). Alternative approaches can reduce the use of test animals under REACH. *Report EUR* **21405**.

Verwei, M., van Burgsteden, J. A., Krul, C. A. M., van de Sandt, J. J. M., Freidig, A. P. (2006). Prediction of in vivo embryotoxic effect levels with a combination of in vitro studies and PBPK modelling. *Toxicol Lett* **165**, 79-87.

West, P. R., Weir, A. M., Smith, A. M., Donley, E. L., Cezar, G. G. (2010). Predicting human developmental toxicity of pharmaceuticals using human embryonic stem cells and metabolomics.

Toxicol Appl Pharmacol **247**, 18-27.

Yamamoto, M., Tase, N., Okuno, T., Kondo, Y., Akiba, S., Shimozawa, N., Terao, K. (2007). Monitoring of gene expression in differentiation of embryoid bodies from cynomolgus monkey embryonic stem cells in the presence of bisphenol A. *J Toxicol Sci* **32**, 301-10.

Zhu, M., Zhao, J., Chen, G., Guan, L. (2001). Early embryonic sensitivity to cyclophosphamide in cardiac differentiation from human embryonic stem cells. *Cell Biol Int* **35**, 927-38.

Zimmer, B., Schildknecht, S., Kuegler, P. B., Tanavde, V., Kadereit, S., Leist, M. (2011a). Sensitivity of dopaminergic neuron differentiation from stem cells to chronic low-dose methylmercury exposure. *Toxicol Sci* **121**, 357-67.

Zimmer, B., Kuegler, P. B., Baudis, B., Genewsky, A., Tanavde, V., Koh, W., Tan, B., Waldmann, T., Kadereit, S., Leist, M. (2011b). Coordinated waves of gene expression during neuronal differentiation of embryonic stem cells as basis for novel approaches to developmental neurotoxicity testing. *Cell Death Differ* **18**, 383-95.

Zur Nieden, N. I., Davis, L. A., Rancourt, D. E. (2010a). Comparing three novel endpoints for developmental osteotoxicity in the embryonic stem cell test. *Toxicol Appl Pharmacol* **247**, 91-97.

Zur Nieden, N. I., Davis, L. A., Rancourt, D. E. (2010b). Monolayer cultivation of osteoprogenitors shortens duration of the embryonic stem cell test while reliably predicting developmental osteotoxicity. *Toxicology* **277**, 66-73.

Zur Nieden, N. I., Kempka, G., Ahr, H. J. (2004). Molecular multiple endpoint embryonic stem cell test-a possible approach to test for the teratogenic potential of chemicals. *Toxicol Appl Pharmacol* **194**, 257-69.

Chapter 3 is reprinted with permission from Elsevier.

CHAPTER 3

Decrease of intracellular pH as possible mechanism of developmental toxicity of glycol ether alkoxyacetic acid metabolites

Jochem Louisse, Yanqing Bai, Miriam Verwei,
Johannes J. M. van de Sandt, Bas J. Blaauboer,
Ivonne M.C.M. Rietjens

Based on:
Toxicology and Applied Pharmacology 245, 236-243 (2010).

Abstract

Developmental toxicity of glycol ethers is caused by their alkoxyacetic acid metabolites, but the mechanism underlying the developmental toxicity of these acid metabolites is so far not known. The present study investigates a possible mechanism underlying the developmental toxicity of glycol ether alkoxyacetic acid metabolites using the methoxyacetic acid (MAA) metabolite of ethylene glycol monomethyl ether as the model compound. The results obtained demonstrate an MAA-induced decrease of the intracellular pH (pH_i) of embryonic BALB/c-3T3 cells as well as of embryonic stem cells (ES-D3 cells), at concentrations that affect ES-D3 cell differentiation. These results suggest a mechanism for MAA-mediated developmental toxicity similar to the mechanism of developmental toxicity of the drugs valproic acid and acetazolamide (ACZ), known to decrease the pH_i *in vivo*, and therefore used as positive controls. The toxic alkoxyacetic acid metabolites ethoxyacetic acid, butoxyacetic acid and phenoxyacetic acid also caused an intracellular acidification of BALB/c-3T3 cells at concentrations that are known to inhibit ES-D3 cell differentiation. Two other developmentally toxic chemicals, all-*trans*-retinoic acid and 5-fluorouracil, did not decrease the pH_i of embryonic cells at concentrations that affect ES-D3 cell differentiation, pointing at a different mechanism of developmental toxicity of these chemicals. MAA and ACZ induced a concentration-dependent inhibition of ES-D3 cell differentiation, which was potentiated by amiloride, an inhibitor of the Na^+/H^+-antiporter, corroborating an important role of the pH_i in the mechanism of developmental toxicity of both chemicals. Together, the results presented indicate that a decrease of the pH_i may be the mechanism of developmental toxicity of the alkoxyacetic acid metabolites of the glycol ethers.

Introduction

Several members of the chemical group of glycol ethers are known for having adverse effects on the developing fetus in mice, rats and rabbits (Hanley *et al.*, 1984; Nelson *et al.*, 1984; Wier *et al.*, 1987). Most attention in developmental toxicity studies of glycol ethers has gone to ethylene glycol monomethyl ether (EGME), being the most potent member of this group of chemicals (Nelson *et al.*, 1984; Welsch, 2005). Developmental toxicity effects are not caused by EGME itself, but by its metabolite methoxyacetic acid (MAA) (Figure 1) (Brown *et al.*, 1984; Giavini *et al.*, 1993).

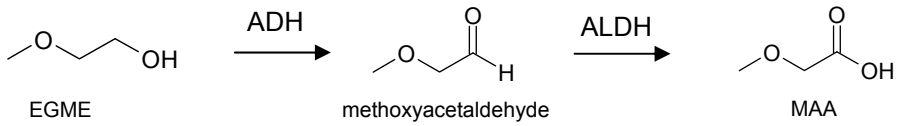

Figure 1. EGME is metabolized to its toxic metabolite MAA by alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (ALDH) (Miller *et al.*, 1983).

Although being the most potent glycol ether, EGME only causes developmental toxicity in rats when MAA plasma levels in the millimolar range are reached (Sweeney *et al.*, 2001). Also in *in vitro* developmental toxicity studies MAA concentrations in the millimolar range are needed to obtain adverse effects. For example, the BMC_5 (Benchmark concentration at which a 5% change occurs compared to the control) of MAA for the total morphological score in the whole embryo culture is 1.6 mM (Piersma *et al.*, 2008) and the BMC_{50} for MAA in the embryonic stem cell (ES-D3 cell) differentiation assay of the Embryonic Stem cell Test (EST) is 2.3–2.5 mM (De Jong *et al.*, 2009). Like several other chemicals that cause developmental toxicity such as valproic acid (VPA), ethylhexanoic acid and boric acid, MAA is a weak acid, suggesting that the acidic nature of MAA and related glycol ether alkoxyacetic acid metabolites might play a role in their induction of developmental toxicity. Nau and Scott (1986) showed that weak acids accumulate in the developing embryo because of its relatively high pH compared to maternal blood. They hypothesized that birth defects caused by these acids could be due to an alteration of the intracellular pH (pH_i) of embryonic cells (Nau and Scott, 1986). A decrease of the pH_i of embryonic cells by the drugs acetazolamide

(ACZ) (Scott *et al.*, 1990) and VPA (Scott *et al.*, 1997) has been linked to their induction of developmental toxicity in mice. ACZ inhibits carbonic anhydrase, causing indirectly an intracellular acidification (Scott *et al.*, 1990), whereas VPA-induced intracellular acidification could be due to the acidic nature of the chemical itself, carrying protons directly to the embryo (Scott *et al.*, 1997). When amiloride, a blocker of the Na^+/H^+-antiporter in the cell membrane, was co-administered with ACZ or VPA, an exacerbation in developmental toxicity was found (Scott *et al.*, 1990; Scott *et al.*, 1997). This was due to the inhibition of the Na^+/H^+-antiporter by amiloride, impeding the cells to transport the excess of H^+ ions out of the cell when being exposed to an acid overload (Scott *et al.*, 1990). The goal of the present study was to investigate the possible role of an effect on the pH_i in the mechanism of developmental toxicity induced by MAA and related glycol ether alkoxyacetic acid metabolites. This was done by measuring the changes in the pH_i of BALB/c-3T3 embryonic fibroblasts exposed to MAA and the alkoxyacetic acid metabolites of ethylene glycol monoethyl ether (EGEE), ethylene glycol monobutyl ether (EGBE) and ethylene glycol monophenyl ether (EGPE), being ethoxyacetic acid (EAA), butoxyacetic acid (BAA) and phenoxyacetic acid (PAA), respectively. Effects on the pH_i of BALB/c-3T3 cells by ACZ and VPA, both known to decrease the pH_i *in vivo*, were investigated as well. Also, the effect of MAA and ACZ on the pH_i of undifferentiated, differentiating and differentiated ES-D3 cells was measured. Furthermore, the effect of the Na^+/H^+-antiporter blocker amiloride on MAA-induced inhibition of ES-D3 cell differentiation was assessed, using ACZ as a positive control.

Materials and methods

Chemicals

ACZ, all-*trans*-retinoic acid (ATRA), amiloride, EAA, 5-fluorouracil (5-FU) and MAA were purchased from Sigma-Aldrich (Zwijndrecht, the Netherlands). BAA, dimethyl sulfoxide (DMSO) and VPA were purchased from Acros Organics (Geel, Belgium) and PAA was purchased from Fluka (Zwijndrecht, the Netherlands). 2',7'-Bis-(2-carboxyethyl)-5-(and-6)-carboxyfluorescein (BCECF), BCECF acetoxymethyl ester (BCECF-AM) and nigericin were purchased from Molecular Probes

(Invitrogen, Breda, the Netherlands). Ethanol, K_2HPO_4 and KH_2PO_4 were obtained from Merck (Darmstadt, Germany).

BALB/c-3T3 cell culture

The BALB/c-3T3 murine embryonic fibroblast cell line was cultured in 75 cm^2 cell culture flasks (Corning, Schiphol-Rijk, the Netherlands) at 37 °C in a humidified atmosphere of 5% CO_2 in Dulbecco's modified Eagle's medium (DMEM) (Invitrogen) supplemented with 10% fetal calf serum (HyClone-Perbio, Etten-Leur, the Netherlands) and 100 U/ml penicillin/100 µg/ml streptomycin (Invitrogen). Cells were subcultured three times a week, using trypsin-EDTA (Invitrogen) to detach cells.

Undifferentiated ES-D3 cell culture

The murine ES-D3 cell line was cultured according to the method described by De Smedt *et al.* (2008). Briefly, cells were cultured in cell culture Petri dishes (35 mm × 10 mm, Corning) at 37 °C in a humidified atmosphere of 5% CO_2 in DMEM supplemented with 20% fetal calf serum, 2 mM glutamine (Invitrogen), 50 U/ml penicillin/50 µg/ml streptomycin, 1% v/v non-essential amino acids (Invitrogen) and 0.1 mM β-mercaptoethanol (Sigma-Aldrich). To keep cells undifferentiated, 1000 U/ml murine leukemia inhibiting factor (mLIF) (ESGRO®, Chemicon International, Amsterdam, The Netherlands) was added to the culture medium. Cells were subcultured three times a week, using non enzymatic dissociation buffer (Sigma-Aldrich) to detach cells.

BCECF calibration curves

To assess the effects of the test chemicals on the pH_i, BCECF-AM was used. BCECF-AM passively diffuses through the cell membrane and is cleaved intracellularly by cellular esterases, resulting in the free acid BCECF, which is trapped inside the cell. BCECF fluorescence at λ_{ex}=490 nm, λ_{em}=535 nm is pH-dependent, whereas BCECF fluorescence at λ_{ex}=440 nm, λ_{em}=535 nm is not (Owen, 1992). The BCECF fluorescence ratio (λ_{ex}=490 nm, λ_{em}=535 nm)/(λ_{ex}=440 nm, λ_{em}=535 nm) is a measure for the pH_i: the higher the ratio, the higher the pH_i. The fluorescence ratio measured in the cells can be translated into actual pH_i values using a BCECF calibration curve. Two types of calibration curves were made: a series of calibration curves using increasing concentrations of the free acid BCECF

in potassium phosphate of different pH values and a calibration curve using BCECF -loaded BALB/c-3T3 cells subsequently incubated with potassium phosphate of different pH values and 10 µM nigericin to equilibrate the pH_i with the extracellular pH (Thomas *et al.*, 1979). To obtain calibration curves for the free acid BCECF, 0.1 M potassium phosphate solutions were prepared with a pH range from 5 to 8 and supplemented with BCECF at concentrations ranging from 0.05 to 1 µM. Solutions containing BCECF were added to a ViewPlate-96-well plate (Perkin Elmer, Groningen, the Netherlands; 200 µl/well). BCECF fluorescence at λ_{ex}=490 nm, λ_{em}=535 nm and λ_{ex}=440 nm, λ_{em}=535 nm was measured using a Varian Cary Eclipse Fluorescent Spectrophotometer (Varian, Middelburg, the Netherlands). To obtain a calibration curve for the BCECF-loaded BALB/c-3T3 cells, a cell suspension of BALB/c-3T3 cells was made and cells were washed with HBSS. The cells were incubated for 30 min with 1 µM BCECF-AM in HBSS and subsequently washed with HBSS. Then, cells were transferred to 0.1 M potassium phosphate of pH values ranging from 5 to 8 supplemented with 10 µM nigericin to final cell concentrations of $0.5 \cdot 10^6$ cells/ml. From these solutions, 200 µl was transferred to a ViewPlate-96-well plate. Subsequently, BCECF fluorescence was measured. The calibration curve in BALB/c-3T3 cells appeared to resemble the curve obtained at 1 µM BCECF (see results section), revealing that cellular uptake and hydrolysis of BCECF-AM appear to be efficient. The calibration curve obtained with BALB/c-3T3 cells was used to translate measured fluorescence ratios into actual pH_i values.

pH_i measurements in BALB/c-3T3 and undifferentiated ES-D3 cells

A cell suspension of either BALB/c-3T3 or ES-D3 cells was made and cells were washed with HBSS. Then, the cells were incubated for 30 min with 1 µM BCECF-AM in HBSS and subsequently washed with HBSS. Cells were diluted to a concentration of $1 \cdot 10^6$ cells/ml in HBSS of which 100 µl was transferred to the inner 60 wells of a ViewPlate-96-well plate. Subsequently, 100 µl exposure medium containing the test chemicals, added directly to HBSS (BAA, EAA, MAA), added from a 100 times concentrated stock solution in DMSO to HBSS (5-FU, ACZ, ATRA, PAA), or added from a 100 times concentrated stock solution in ethanol (VPA), was added to the cells (final solvent concentrations 0.5% v/v). After 30 and 60 min incubation, BCECF fluorescence at λ_{ex}=490 nm, λ_{em}=535 nm and λ_{ex}=440 nm, λ_{em}=535 nm was measured using a Varian Cary Eclipse Fluorescent

Spectrophotometer. Fluorescence measurements of cells exposed to ACZ were also performed after 2 and 4 h exposure.

ES-D3 cell differentiation assay

To study the effects of ACZ and MAA in the presence or absence of amiloride on the differentiation of ES-D3 cells into contracting cardiomyocytes, the differentiation assay described by De Smedt *et al.* (2008) was used, with small modifications. Briefly, on day 0, droplets of 20 µl cell suspension ($3.75 \cdot 10^4$ cells/ml) were seeded on the inside of the lid of a 96-well plate (Corning). Caps of Eppendorf tubes were placed at the corners of the well plate to elevate the lid in order to prevent direct contact of the hanging drops with the well plate. Wells were filled with 250 µl phosphate buffered saline (PBS, Invitrogen) to obtain the right humidity and to prevent evaporation of the hanging drops. On day 3, cell aggregates, called embryoid bodies (EBs), were formed and these were transferred to bacteriological petri dishes (Greiner Bio-one, Alphen a/d Rijn, the Netherlands). On day 5, EB structure and size were evaluated using light microscopy. Pictures of EBs were taken using an Olympus Color View IIIu camera connected to an Olympus CKX41 culture microscope (Olympus Soft Imaging Solutions, Münster, Germany). EBs were transferred to 24-well plates (Corning, 1 EB/well) and were incubated for another 5 days. On day 10, the number of wells containing contracting cardiomyocytes was determined using light microscopy. Exposure to test chemicals, added from 400 times concentrated stock solutions in DMSO to culture medium, started at day 0 and lasted for 10 days. Exposure medium was refreshed at day 3 and day 5. Solvent controls (0.25% DMSO v/v) were included in each experiment. A test was considered valid when at least 21 out of 24 wells of the blank (non-exposed cells) and the solvent control plate contained contracting cardiomyocytes. For each plate, the fraction of EBs differentiated into contracting cardiomyocytes was calculated.

pH_i measurements in differentiating EBs (day 5 EBs) and differentiated EBs (day 10 EBs)

To measure the effect of MAA on the pH_i of differentiating ES-D3 cells, the ES-D3 cell differentiation assay was carried out as described above with small modifications. The effect of ACZ and MAA on the pH_i was investigated using

differentiating EBs (day 5 EBs) and differentiated EBs (day 10 EBs). Day 5 EBs were washed with HBSS and incubated for 30 min with 1 µM BCECF-AM in HBSS. The BCECF-AM solution was removed after 30 min and EBs were washed with HBSS. Subsequently, 20 EBs/well were plated in a ViewPlate-96-well plate containing ACZ added from a 200 times concentrated stock solution in DMSO to HBSS or MAA added directly to HBSS. After 30 min incubation, BCECF fluorescence was measured in MAA-exposed cells and after 1, 2 and 4 h incubation BCECF fluorescence was measured in ACZ-exposed cells. To measure the pH_i of day 10 EBs, day 5 EBs that had not been used for pH_i measurements were plated in a ViewPlate -96-well plate. At day 10, the differentiated EBs were washed with HBSS and incubated for 30 min with 1 µM BCECF-AM in HBSS. The BCECF-AM solution was removed after 30 min and differentiated EBs were washed with HBSS. Subsequently, the differentiated EBs were exposed to ACZ added from a 200 times concentrated stock solution in DMSO to HBSS or MAA added directly to HBSS. After 30 min incubation, BCECF fluorescence was measured in MAA-exposed cells and after 1, 2 and 4 h incubation BCECF fluorescence was measured in ACZ-exposed cells.

Data analysis

Concentration-response curve fitting for the ES-D3 cell differentiation assay was performed using the United States Environmental Protection Agency (US EPA)'s benchmark dose software (BMDS) version 2.0. In order to test for significant differences in pH_i values of cells exposed to test chemicals compared to the control, a one-way ANOVA followed by a Dunnett's Multiple Comparison Test was used (Graphpad Prism 5). To test for significant effects of amiloride on differentiating EB size (day 5 EBs) and the fraction differentiated EBs (day 10 EBs), one-tailed Student's t-tests were used (Graphpad Prism 5).

Results

BCECF calibration curves

Figure 2 shows the measured fluorescence ratios (λ_{ex}=490 nm, λ_{em}=535 nm/ λ_{ex}=440 nm, λ_{em}=535 nm) obtained with the BCECF concentrations present in 0.1 M potassium phosphate with pH values ranging from 5 to 8 and the measured

fluorescence ratios of BCECF-loaded BALB/c-3T3 cells in 0.1 M potassium phosphate with pH values ranging from 5 to 8 and supplemented with 10 µM nigericin. A sigmoidal model was fitted to these data points to obtain BCECF calibration curves using Sigmaplot 9.01 (Figure 2). Figure 2 shows that pH-dependent BCECF calibration curves vary with the BCECF concentration. The calibration curve of 1 µM BCECF resembles the calibration curve made in the BCECF-loaded BALB/c-3T3 cells, revealing that cellular uptake and hydrolysis of BCECF-AM appear to be efficient and suggesting that intracellular BCECF concentrations are equivalent to or higher than the extracellular concentration of 1 µM BCECF-AM added to upload the cells. The calibration curve made in the BCECF-loaded BALB/c-3T3 cells was used to translate the measured fluorescence ratios obtained in the pH_i experiments performed in BALB/c-3T3 and ES-D3 cells into actual pH_i values.

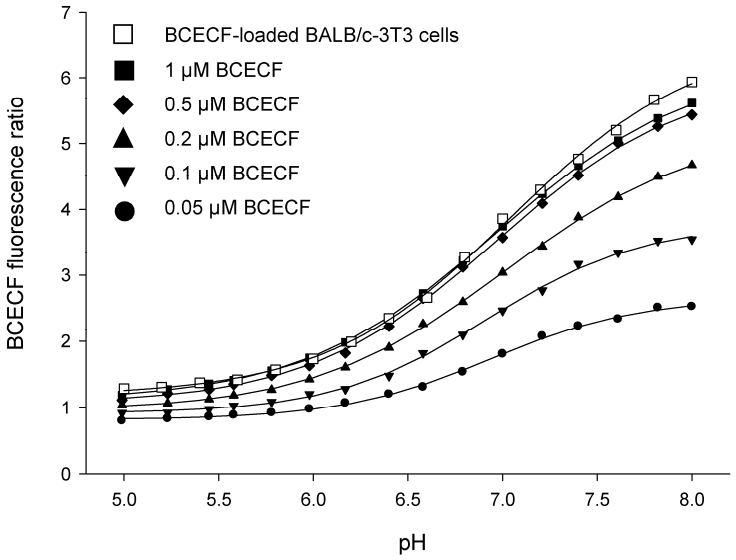

Figure 2. BCECF fluorescence ratios ((λ_{ex}=490 nm, λ_{em}=535 nm)/(λ_{ex}=440 nm, λ_{em}=535 nm)) of calibration curves obtained using increasing BCECF concentrations added to 0.1 M potassium phosphate of pH values ranging from 5 to 8 (closed symbols) and BCECF-loaded BALB/c-3T3 cells incubated in 0.1 M potassium phosphate with pH values ranging from 5 to 8 supplemented with 10 µM nigericin to equilibrate the pH_i with the extracellular pH (open symbol). The calibration curve made in BCECF-loaded BALB/c-3T3 cells was used to convert measured fluorescence ratios in both BALB/c-3T3 and ES-D3 cells into actual pH_i values.

ACZ, VPA and the alkoxyacetic acid metabolites of glycol ethers cause a decrease of the pH_i of BALB/c-3T3 cells

Figure 3 shows the concentration-dependent decrease of the pH_i of BALB/c-3T3 cells exposed to the developmentally toxic drugs ACZ and VPA. No effect on the pH_i of ACZ-exposed BALB/c-3T3 cells was found after 30 and 60 min exposure. A small but significant decrease of the pH_i by 5 mM ACZ was found after 2 h exposure and by 2 and 5 mM ACZ after 4 h exposure (Figure 3A). For VPA, no differences in effects on the pH_i were found between 30 and 60 min exposure (data shown for 30 min exposure in Figure 3B). A significant decrease of the pH_i was measured at VPA

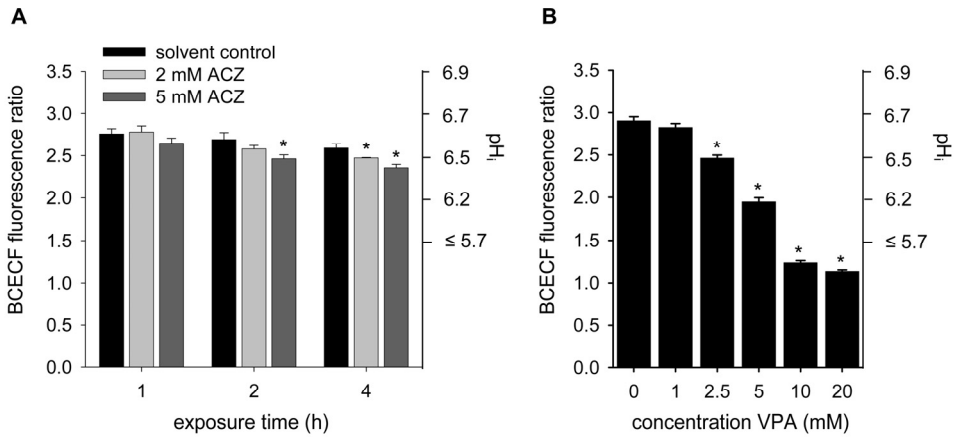

Figure 3. Decrease of pH_i of BALB/c-3T3 cells exposed for 1, 2 and 4 h to ACZ (A) and for 30 min to VPA (B) indicated by decreasing BCECF fluorescence ratio values (λ_{ex}=490 nm, λ_{em}=535 nm/λ_{ex}=440 nm, λ_{em}=535 nm). The fluorescence ratios were converted into pH_i values indicated on the Y-axis on the right using the calibration curve made for BCECF-loaded BALB/c-3T3 cells of Figure 2. Note that the conversion of BCECF fluorescence ratios into pH_i values does not follow a linear scale since the calibration curve is also non-linear. *Significantly different ($p<0.05$) from BCECF fluorescence ratio in cells of the solvent control.

concentrations of 2.5 mM and higher (Figure 3B). Also the developmentally toxic alkoxyacetic acid metabolites MAA, EAA, BAA and PAA caused a concentration-dependent decrease of the pH_i of these cells after 30 min exposure (Figure 4), which was the same after 60 min exposure (data not shown). Exposure of BALB/c-3T3 cells to 5-FU and ATRA up to concentrations of 10 µM and 1 µM respectively, did not change the pH_i of BALB/c-3T3 cells (Figure 5).

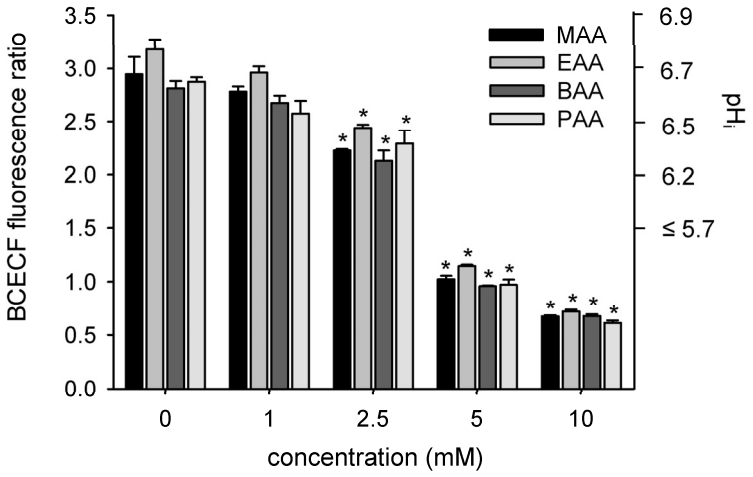

Figure 4. Decrease of pH_i of BALB/c-3T3 cells exposed for 30 min to MAA, EAA, BAA or PAA indicated by decreasing BCECF fluorescence ratio values (λ_{ex}=490 nm, λ_{em}=535 nm/λ_{ex}=440 nm, λ_{em}=535 nm). The fluorescence ratios were converted into pH_i values indicated on the Y-axis on the right using the calibration curve made for BCECF-loaded BALB/c-3T3 cells of Figure 2. Note that the conversion of BCECF fluorescence ratios into pH_i values does not follow a linear scale since the calibration curve is also non-linear. *Significantly different ($p<0.05$) from BCECF fluorescence ratio in cells of the solvent control.

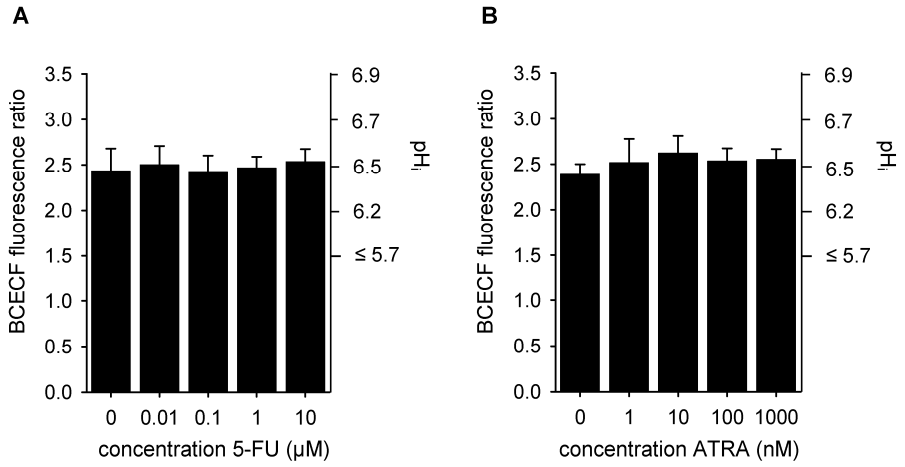

Figure 5. 5-FU (A) and ATRA (B) do not change the pH_i of BALB/c-3T3 cells as indicated by BCECF fluorescence ratio values (λ_{ex}=490 nm, λ_{em}=535 nm/λ_{ex}=440 nm, λ_{em}=535 nm). The fluorescence ratios were converted into pH_i values indicated on the Y-axis on the right using the calibration curve made for BCECF-loaded BALB/c-3T3 cells of Figure 2. Note that the conversion of BCECF fluorescence ratios into pH_i values does not follow a linear scale since the calibration curve is also non-linear.

MAA and ACZ cause a decrease of the pH_i of differentiating and differentiated ES-D3 cells

The effect of ACZ (Figure 6) and MAA (Figure 7) on the pH_i of undifferentiated, differentiating and differentiated ES-D3 cells was investigated. The concentrations of MAA and ACZ used are the same as the ones that were used in the ES-D3 cell differentiation assay (Figures 8 and 9). Figure 6A shows the results obtained for the undifferentiated ES-D3 cells exposed for 4 h to ACZ and also presents, for comparison, the data for the effect of ACZ on the pH_i of BALB/c-3T3 cells. The data show that the pH_i of undifferentiated ES-D3 cells is not affected by ACZ, whereas ACZ induces a small and significant decrease of the pH_i of BALB/c-3T3 cells (Figure 6A). ACZ caused a more pronounced decrease of the pH_i of differentiating ES-D3 cells (day 5 EBs, Figure 6B), and a small decrease in differentiated ES-D3 cells (day 10 EBs, Figure 6C). In day 5 EBs, a significant decrease of the pH_i was found after 2 and 4 h exposure to 1 and 2.5 mM ACZ (Figure 6B). In day 10 EBs, a significant decrease of the pH_i was found after 1 and 2 h exposure to 2.5 mM ACZ (Figure 6C).

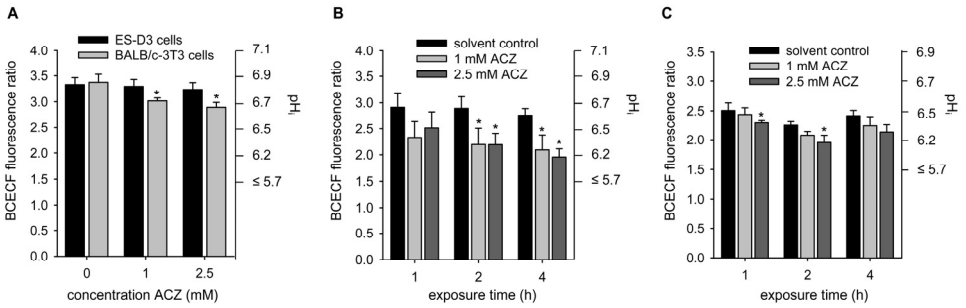

Figure 6. Decrease of pH_i of BALB/c-3T3 and ES-D3 cells exposed for 4 h to ACZ (A), day 5 differentiating (B) and day 10 differentiated (C) EBs exposed for 1, 2 and 4 h to ACZ indicated by decreasing BCECF fluorescence ratio values (λ_{ex}=490 nm, λ_{em}=535 nm/λ_{ex}=440 nm, λ_{em}=535 nm). The fluorescence ratios were converted into pH_i values indicated on the Y-axis on the right using the calibration curve made for BCECF-loaded BALB/c-3T3 cells of Figure 2. Note that the conversion of BCECF fluorescence ratios into pH_i values does not follow a linear scale since the calibration curve is also non-linear. *Significantly different ($p<0.05$) from BCECF fluorescence ratio in cells of the solvent control.

Figure 7A shows the results obtained for the effects of MAA on the pH_i of undifferentiated ES-D3 cells and also presents, for comparison, the data for the effect of MAA on the pH_i of BALB/c-3T3 cells. The data show that the decrease of the pH_i of ES-D3 cells by MAA is comparable to the MAA-induced decrease of the

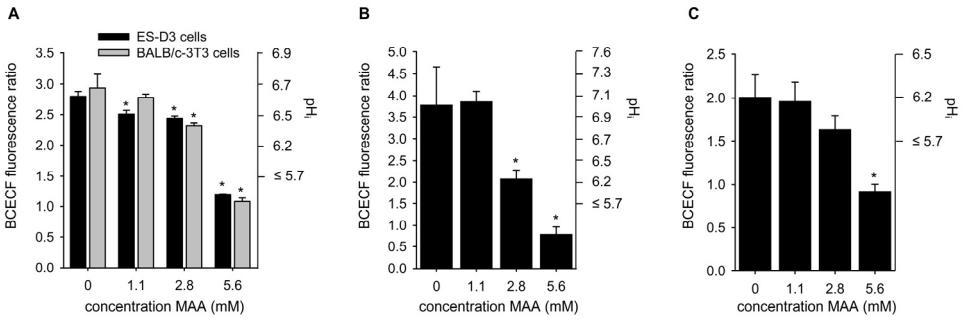

Figure 7. Decrease of pH_i of BALB/c-3T3 and ES-D3 cells (A), day 5 differentiating (B) and day 10 differentiated (C) EBs exposed for 30 min to MAA indicated by decreasing BCECF fluorescence ratio values (λ_{ex}=490 nm, λ_{em}=535 nm/λ_{ex}=440 nm, λ_{em}=535 nm). The fluorescence ratios were converted into pH_i values indicated on the right using the calibration curve made for BCECF-loaded BALB/c-3T3 cells of Figure 2. Note that the conversion of BCECF fluorescence ratios into pH_i values does not follow a linear scale since the calibration curve is also non-linear. *Significantly different ($p<0.05$) from BCECF fluorescence ratio in cells of the solvent control.

pH_i of BALB/c-3T3 cells. MAA also causes a concentration-dependent decrease of the pH_i of differentiating (day 5 EBs, Figure 7B) and differentiated (day 10 EBs, Figure 7C) ES-D3 cells. In day 5 EBs, a significant decrease of the pH_i was found at MAA concentrations of 2.8 and 5.6 mM, whereas in day 10 EBs, only a significant decrease of the pH_i was found at 5.6 mM.

MAA and ACZ decrease day 5 EB size and inhibit ES-D3 cell differentiation, both potentiated by the Na^+/H^+-antiporter blocker amiloride

Figure 8 shows the effects of ACZ and MAA in the absence or presence of 100 μM amiloride on the structure and size of differentiating ES-D3 cells (day 5 EBs). A decrease in EB-size was noticed for all the ACZ concentrations tested. When 100 μM amiloride was co-administered with ACZ, the EBs had a granular structure (Figure 8). For MAA, a decrease in EB-size was only noticed at the highest concentration tested (5.6 mM), whereas when 100 μM amiloride was co-administered with MAA, the EB-size had already decreased at a concentration of 1.1 mM (Figure 8).

The inhibition of ES-D3 cell differentiation into contracting cardiomyocytes by ACZ and MAA is shown in Figure 9. ACZ, the positive control for inducing developmental toxicity by causing intracellular acidification, inhibited ES-D3 cell

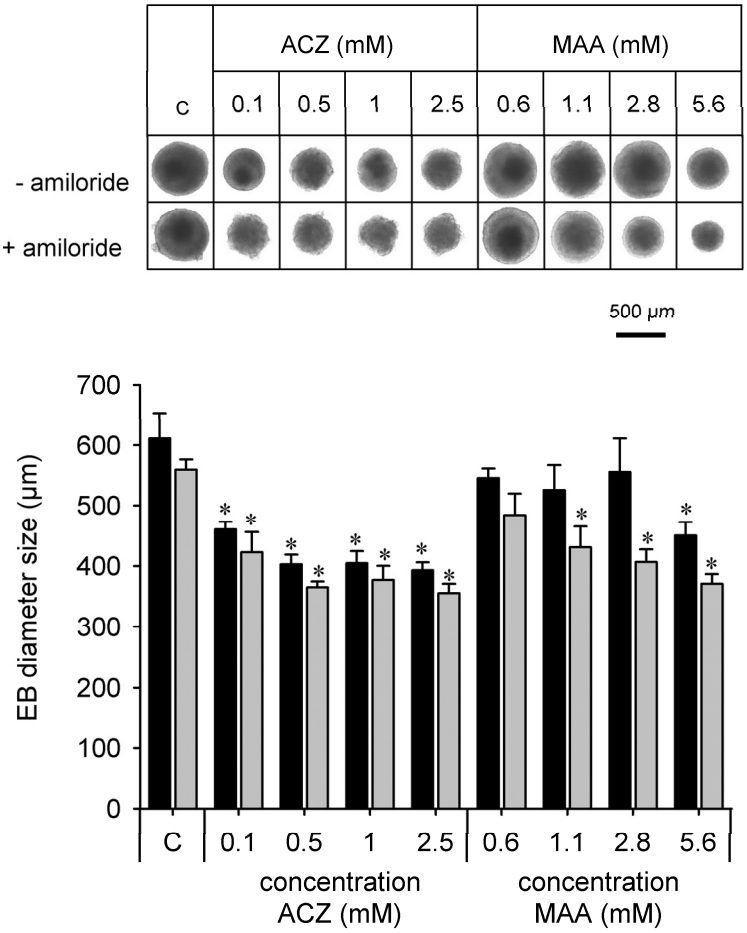

Figure 8. Decrease in EB size at day 5 after exposure to ACZ or MAA in the absence (black columns) or presence (grey columns) of 100 µM amiloride. C=solvent control. *Significantly smaller diameter compared to control ($p<0.05$).

differentiation with a BMC_{50} of 1.3 mM ($BMCL_{50}$: 1.2 mM; the lower limit of the 95% confidence interval on the BMC_{50}) (Figure 9A), whereas MAA inhibited ES-D3 cell differentiation with a BMC_{50} of 2.9 mM ($BMCL_{50}$: 2.6 mM) (Figure 9B). The effects of both ACZ and MAA were potentiated when 100 µM amiloride was co-administered, resulting in BMC_{50} values of 0.20 mM for ACZ ($BMCL_{50}$: 0.15 mM) and 1.3 mM for MAA ($BMCL_{50}$: 1.2 mM). Exposure to 100 µM amiloride alone did not affect ES-D3 cell differentiation (Figure 9).

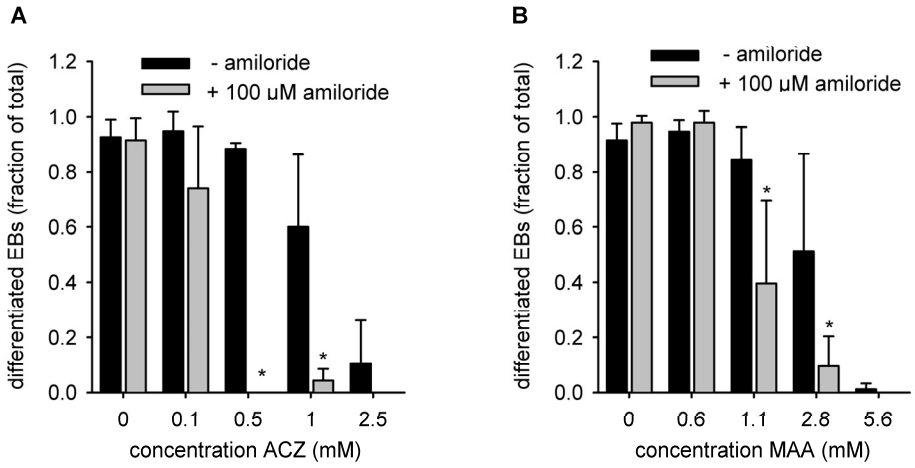

Figure 9. Inhibition of EB differentiation into contracting cardiomyocytes when being exposed to ACZ (A, n=3) or MAA (B, n=4) in the absence (black columns) or presence (grey columns) of 100 µM amiloride. *Significant difference ($p<0.05$) between absence and presence of 100 µM amiloride.

Discussion

The goal of the present study was to investigate whether the alkoxyacetic acid metabolites of glycol ethers, being their proximate developmental toxicants, affect the pH_i of embryonic cells, providing a hypothesis for the mechanism underlying glycol ether-induced developmental toxicity. The study shows that the glycol ether alkoxyacetic acid metabolites MAA, EAA, BAA and PAA, cause a concentration-dependent decrease of the pH_i of embryonic cells. Furthermore, the results obtained show that MAA-induced inhibition of ES-D3 cell differentiation is potentiated by the Na^+/H^+-antiporter blocker amiloride, impeding the cells to remove the excess of H^+-ions when being exposed to the acid overload. These results point to a role of the decrease of the pH_i in the MAA-induced inhibition of ES-D3 cell differentiation. Altogether, the results of the present study indicate that a decrease of the pH_i of embryonic cells is a possible mechanism by which the glycol ether alkoxyacetic acid metabolites induce developmental toxicity.

The developmental toxicants ACZ and VPA have been shown to induce developmental toxicity *in vivo* by causing a decrease in the pH_i of mouse embryos (Scott *et al.*, 1990; Scott *et al.*, 1997). The present study shows that these chemicals,

which were used as positive controls, also decrease the pH_i of embryonic cells *in vitro*. ACZ decreased the pH_i of embryonic cells at concentrations that inhibit ES-D3 cell differentiation into contracting cardiomyocytes (Figures 3A and 6, Table 1). The decrease of the pH_i caused by ACZ was dependent on the exposure time. This is probably due to the fact that ACZ is a carbonic anhydrase inhibitor, resulting in an indirect effect on the pH_i of the cells (Scott *et al.*, 1990). The decrease of the pH_i caused by ACZ was rather small in the BALB/c-3T3 cells compared to the decrease of the pH_i caused by the other positive control VPA. This might be due to the fact that ACZ and VPA have different mechanisms underlying the pH_i decrease, proceeding by respectively the inhibition of intracellular carbonic anhydrase (Scott *et al.*, 1990) and by carrying protons directly to the cell as suggested by Scott *et al.* (1997).

Table 1. Concentrations of chemicals inhibiting ES-D3 cell differentiation into contracting cardiomyocytes.

Chemical	Concentration inhibiting ES-D3 cell differentiation
ACZ	BMC_{50}: 2.2 mM[a]
VPA	ID_{50}: 0.3 - 0.5 mM[b]
MAA	BMC_{50}: 2.3 - 2.5 mM[c]
EAA	BMC_{50}: 2.9 - 3.9 mM[c]
BAA	BMC_{50}: 4.5 - 5.9 mM[c]
PAA	BMC_{50}: 4.6 - 7.8 mM[c]
5-FU	ID_{50}: 0.6 - 0.7 µM[b]
ATRA	ID_{50}: 1-27 nM[b]

[a] BMC_{50}: 50% inhibition of differentiation, determined in present study.
[b] ID_{50}: 50% inhibition of differentiation, determined by Genschow *et al.* (2004).
[c] BMC_{50}: 50% inhibition of differentiation, determined by De Jong *et al.* (2009).

Since no carbonic anhydrase enzyme activity or carbonic anhydrase protein expression has been detected in 3T3-L1 fibroblasts (Lynch *et al.*, 1993), the small decrease of the pH_i found in Balb/c-3T3 fibroblasts exposed to ACZ might be due to low expression of carbonic anhydrase in these cells. ACZ did not cause a decrease in the pH_i of undifferentiated ES-D3 cells at all, pointing at possibly no carbonic

anhydrase expression in undifferentiated ES-D3 cells (Figure 6A). However, a more pronounced pH_i decrease was found in differentiating ES-D3 cells (Figure 6B), suggesting that carbonic anhydrase is expressed during the differentiation of ES-D3 cells into cardiomyocytes. This suggestion is further supported by the finding that carbonic anhydrase is expressed in the embryonic heart, including the myocardium, during mouse development (Vuillemin and Pexieder, 1997).

A significant decrease of the pH_i by VPA was found at concentrations of 2.5 mM and higher, approximately 5 times higher than the concentrations that were found to inhibit ES-D3 cell differentiation into contracting cardiomyocytes (Genschow *et al.*, 2004) (Figure 3B, Table 1). Like ACZ, the alkoxyacetic acid metabolites MAA, EAA, BAA and PAA caused a decrease of the pH_i of embryonic cells at concentrations that inhibit the differentiation of ES-D3 cells into contracting cardiomyocytes (De Jong *et al.*, 2009) (Figure 4, Table 1).

The two other developmental toxicants that were tested in the present study, 5-FU and ATRA, did not change the pH_i of embryonic cells, even at concentrations more than 10 times higher than those that inhibit the differentiation of ES-D3 cells into contracting cardiomyocytes (Genschow *et al.*, 2004) (Figure 5, Table 1). These results point at a different mechanism of developmental toxicity of these chemicals, which are suggested to be thymidylate synthetase inhibition by 5-FU (Shuey *et al.*, 1994) and disturbance of retinoid signal transduction pathways by ATRA (Collins and Mao, 1999).

The role of a decrease of the pH_i by the glycol ether alkoxyacetic acid metabolites on the inhibition of ES-D3 cell differentiation was further investigated using MAA as a model compound and ACZ as a positive control. Like ACZ, MAA caused a decrease of the pH_i of differentiating and differentiated ES-D3 cells at similar concentrations as was found in BALB/c-3T3 cells (Figure 7). A decrease of the pH_i caused by ACZ and MAA in the ES-D3 cell differentiation assay occurred at concentrations affecting both the diameter size of day 5 EBs (Figure 8) and the ES-D3 cell differentiation into contracting cardiomyocytes (Figure 9), indicating that the ACZ- and MAA-induced decrease of the pH_i might affect EB formation and ES-D3 cell differentiation. This indication is strengthened by the findings that the Na^+/H^+-antiporter blocker amiloride potentiates the MAA-induced decrease in diameter size of day 5 EBs (Figure 8) and the ACZ- and MAA-induced inhibition of

ES-D3 cell differentiation into contracting cardiomyocytes (Figure 9), since amiloride is known for its inhibition of cellular recovery from an acid overload (Muallem and Loessberg, 1990; Phillips *et al.*, 2000; Scott *et al.*, 1990). However, whether amiloride can potentiate EGME-induced developmental toxicity *in vivo* as has been shown for both ACZ (Scott *et al.*, 1990) and VPA (Scott *et al.*, 1997) is not clear. Nelson *et al.* (1989) investigated the effects of amiloride on EGME-induced developmental toxicity. No conclusions about a possible potentiation of EGME-induced developmental toxicity by amiloride could be drawn, since the number of malformations was higher in the group to which amiloride was co-administered with EGME, whereas the number of resorptions was higher in the group that was exposed to EGME alone (Nelson *et al.*, 1989). It should be mentioned that amiloride was administered via the diet in the EGME study (Nelson *et al.*, 1989), whereas amiloride was administered subcutaneously in the ACZ study (Scott *et al.*, 1990). Therefore, it is not clear whether in the study of Nelson *et al.* (1989) amiloride plasma levels have been high enough to potentiate EGME-induced developmental toxicity, due to possible differences in amiloride kinetics upon oral versus subcutaneous administration.

Scott *et al.* (1994) suggested that the developmental toxicity potential of an acid depends on its 'intrinsic activity', which can be measured in an *in vitro* developmental toxicity test, and its kinetic properties *in vivo*, like half-life values in the maternal blood or in the embryo. Methylhexanoic acid and octanoic acid are for example rapidly eliminated from the blood and therefore not developmentally toxic *in vivo*, whereas ethylhexanoic acid and VPA are slowly eliminated from the blood, resulting in embryo malformations (Scott *et al.*, 1994). Differences in *in vitro* potencies of the glycol ether alkoxyacetic acid metabolites MAA, EAA, BAA and PAA have been observed in rat whole embryo cultures and in the EST (Giavini *et al.*, 1993; De Jong *et al.*, 2009), with MAA being the most potent followed by EAA, BAA and PAA respectively. Also differences in *in vivo* kinetics of MAA, EAA and BAA are reported, with BAA being the most rapidly excreted followed by EAA and MAA respectively (Ghanayem *et al.*, 1990; Aasmoe *et al.*, 1999). The differences in *in vitro* potencies and elimination rates of the alkoxyacetic acid metabolites are in concordance with the *in vivo* potency of their parent glycol ethers, with EGME being the most potent, followed by EGEE, EGBE and EGPE, respectively (Nagano *et*

al., 1981; Wier *et al.*, 1987; Heindel *et al.*, 1990; De Jong *et al.*, 2009). *In vivo* developmental toxicity effects of these glycol ethers are found at doses that result in peak plasma levels of the accompanying alkoxyacetic acid metabolites in the millimolar range (EGME/MAA (Sweeney *et al.*, 2001), EGEE/EAA (Doe, 1984; Gargas *et al.*, 2000), EGBE/BAA (Sleet *et al.*, 1989; Ghanayem *et al.*, 1990)), similar to the concentrations that affect the pH_i, as shown in the present study.

Altogether it is concluded that the alkoxyacetic acid metabolites of glycol ethers cause an intracellular acidification of embryonic cells at concentrations that inhibit embryonic stem cell differentiation. Since these concentrations are observed as peak plasma levels of animals exposed to developmentally toxic doses of the parent glycol ethers, it is suggested that intracellular acidification plays a role in the developmental toxicity caused by the alkoxyacetic acid metabolites of glycol ethers.

References

Aasmoe, L., Mathiesen, M., Sager, G. (1999). Elimination of methoxyacetic acid and ethoxyacetic acid in rat. *Xenobiotica* **29**, 417-24.

Brown, N. A., Holt, D., Webb, M. (1984). The teratogenicity of methoxyacetic acid in the rat. *Toxicol Lett* **22**, 93-100.

Collins, M. D., Mao, G. E. (1999). Teratology of retinoids. *Annu Rev Pharmacol Toxicol* **39**, 399-430.

De Jong, E., Louisse, J., Verwei, M., Blaauboer, B. J., van de Sandt, J. J. M., Woutersen, R. A., Rietjens, I. M. C. M., Piersma, A. H. (2009). Relative developmental toxicity of glycol ether alkoxy acid metabolites in the embryonic stem cell test as compared to the *in vivo* potency of their parent compounds. *Toxicol Sci* **110**, 117-24.

De Smedt, A., Steemans, M., de Boeck, M., Peters, A. K., van der Leede, B. J., van Goethem, F., Lampo, A., Vanparys, P. (2008). Optimisation of the cell cultivation methods in the embryonic stem cell test results in an increased differentiation potential of the cells into strong beating myocard cells. *Toxicol In Vitro* **22**, 1789-96.

Doe, J. E. (1984). Ethylene glycol monoethyl ether and ethylene glycol monoethyl ether acetate teratology studies. *Environ Health Perspect* **57**, 33-41.

Gargas, M. L., Tyler, T. R., Sweeney, L. M., Corley, R. A., Weitz, K. K., Mast, T. J., Paustenbach, D. J., Hays, S. M. (2000). A toxicokinetic study of inhaled ethylene glycol ethyl ether acetate and validation of a physiologically based pharmacokinetic model for rat and human. *Toxicol Appl Pharmacol* **165**, 63 -73.

Genschow, E., Spielmann, H., Scholz, G., Pohl, I., Seiler, A., Clemann, N., Bremer, S., Becker, K. (2004). Validation of the embryonic stem cell test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* **32**, 209-44.

Ghanayem, B. I., Sanders, J. M., Clark, A. M., Bailer, J., Matthews, H. B. (1990). Effects of dose, age, inhibition of metabolism and elimination on the toxicokinetics of 2- butoxyethanol and its

metabolites. *J Pharmacol Exp Ther* **253**, 136-43.

Giavini, E., Broccia, M. L., Menegola, E., Prati, M. (1993). Comparative in vitro study of the embryotoxic effects of three glycol ethers and their metabolites, the alkoxyacids. *Toxicol In Vitro* **7**, 777-84.

Hanley Jr., T. R., Yano, B. L., Nitschke, K. D., John, J. A. (1984). Comparison of the teratogenic potential of inhaled ethylene glycol monomethyl ether in rats, mice, and rabbits. *Toxicol Appl Pharmacol* **75**, 409-22.

Heindel, J. J., Gulati, D. K., Russell, V. S., Reel, J. R., Lawton, A. D., Lamb IV, J. C. (1990). Assessment of ethylene glycol monobutyl and monophenyl ether reproductive toxicity using a continuous breeding protocol in Swiss CD-1 mice. *Fundam Appl Toxicol* **15**, 683-96.

Lynch, C. J., Hazen, S. A., Horetsky, R. L., Carter, N. D., Dodgson, S. J. (1993). Differentiation-dependent expression of carbonic anhydrase II and III in 3T3 adipocytes. *Am J Physiol* **265,** C234-43.

Miller, R. R., Hermann, E. A., Langvardt, P. W., McKenna, M. J., Schwetz, B. A. (1983). Comparative metabolism and disposition of ethylene glycol monomethyl ether and propylene glycol monomethyl ether in male rats. *Toxicol Appl Pharmacol* **67**, 229-37.

Muallem, S., Loessberg, P. A. (1990). Intracellular pH-regulatory mechanisms in pancreatic acinar cells: II. Regulation of H^+ and HCO_3^- transporters by Ca^{2+}-mobilizing agonists. *Biol Chem* **265**, 12813-19.

Nagano, K., Nakayama, E., Oobayashi, H., Yamada, T., Adachi, H., Nishizawa, T., Ozawa, H., Nakaichi, M., Okuda, H., Minami, K., Yamazaki, K. (1981). Embryotoxic effects of ethylene glycol monomethyl ether in mice. *Toxicology* **20**, 335-43.

Nau, H., Scott Jr., W. J. (1986). Weak acids may act as teratogens by accumulating in the basic milieu of the early mammalian embryo. *Nature* **323**, 276-78.

Nelson, B. K., Setzer, J. V., Brightwell, W. S., Mathinos, P. R., Kuczuk, M. H., Weaver, T. E., Goad, P. T. (1984). Comparative inhalation teratogenicity of four glycol ether solvents and an amino derivative in rats. *Environ Health Perspect* **57**, 261-71.

Nelson, B. K., Vorhees, C. V., Scott Jr., W. J., Hastings, L. (1989). Effects of 2-methoxyethanol on fetal development, postnatal behavior, and embryonic intracellular pH of rats. *Neurotoxicol Teratol* **11**, 273 -84.

Owen, C. S. (1992). Comparison of spectrum-shifting intracellular pH probes 5'(and 6')- carboxy-10-dimethylamino-3-hydroxyspiro[7H-benzo[c]xanthene-7, 1'(3'H)-isobenzofuran]- 3'-one and 2', 7'-biscarboxyethyl-5(and 6)-carboxyfluorescein. *Anal Biochem* **204**, 65-71.

Piersma, A. H., Janer, G., Wolterink, G., Bessems, J. G. M., Hakkert, B. C., Slob, W. (2008). Quantitative extrapolation of in vitro whole embryo culture embryotoxicity data to developmental toxicity in vivo using the benchmark dose approach. *Toxicol Sci* **101**, 91-100.

Phillips, K. P., Léveillé, M. C., Claman, P., Baltz, J. M. (2000). Intracellular pH regulation in human preimplantation embryos. *Hum Reprod* **15**, 896-904.

Scott, W. J., Duggan, C. A., Schreiner, C. M., Collins, M. D. (1990). Reduction of embryonic intracellular pH: a potential mechanism of acetazolamide-induced limb malformations. *Toxicol Appl Pharmacol* **103**, 238-54.

Scott Jr., W. J., Collins, M. D., Nau, H. (1994). Pharmacokinetic determinants of embryotoxicity in rats associated with organic acids. *Environ Health Perspect* **102**, 97-101.

Scott Jr., W. J., Schreiner, C. M., Nau, H., Vorhees, C. V., Beliles, R. P., Colvin, J., McCandless, D. (1997). Valproate-induced limb malformations in mice associated with reduction of intracellular pH. *Reprod Toxicol* **11**, 483-93.

Shuey, D. L., Lau, C., Logsdon, T. R., Zucker, R. M., Elstein, K. H., Narotsky, M. G., Setzer, R. W., Kavlock, R. J., Rogers, J. M. (1994). Biologically based dose-response modeling in developmental toxicology: biochemical and cellular sequelae of 5-fluorouracil exposure in the developing rat. *Toxicol Appl Pharmacol* **126**, 129-44.

Sleet, R. B., Price, C. J., Marr, M. C., Morrissey, R. E., Schwetz, B. A. (1989). Teratologic evaluation of ethylene glycol monobutyl ether (cas no. 111-76-2) administered to Fischer-344 rats on either gestational days 9 through 11 or days 11 through 13. *NTP Report* **89-058**.

Sweeney, L. M., Tyler, T. R., Kirman, C. R., Corley, R. A., Reitz, R. H., Paustenbach, D. J., Holson, J. F., Whorton, M. D., Thompson, K. M., Gargas, M. L. (2001). Proposed occupational exposure limits for select ethylene glycol ethers using PBPK models and Monte Carlo simulations. *Toxicol Sci* **62**, 124-39.

Thomas, J. A., Buchsbaum, R. N., Zimniak, A., Racker, E. (1979). Intracellular pH measurements in Ehrlich ascites tumor cells utilizing spectroscopic probes generated in situ. *Biochem* **18**, 2210-18.

Vuillemin, M., Pexieder, T. (1997). Carbonic anhydrase II expression pattern in mouse embryonic and fetal heart. *Anat Embryol* **195**, 267-77.

Welsch, F. (2005). The mechanism of ethylene glycol ether reproductive and developmental toxicity and evidence for adverse effects in humans. *Toxicol Lett* **156**, 13-28.

Wier, P. J., Lewis, S. C., Traul, K. A. (1987). A comparison of developmental toxicity evident at term to postnatal growth and survival using ethylene glycol monoethyl ether, ethylene glycol monobutyl ether and ethanol. *Teratog Carcinog Mutagen* **7**, 55-64.

Chapter 4 is reprinted with permission from Oxford University Press.

CHAPTER 4

Relative developmental toxicity of glycol ether alkoxyacetic acid metabolites in the embryonic stem cell test as compared with the *in vivo* potency of their parent compounds

Esther de Jong, Jochem Louisse, Miriam Verwei, Bas J. Blaauboer, Johannes J.M. van de Sandt, Ruud A. Woutersen, Ivonne M.C.M. Rietjens, Aldert H. Piersma

Based on:
Toxicological Sciences 110, 117-124 (2009).

Abstract

The embryonic stem cell test (EST) has been proposed as an *in vitro* assay that might reduce animal experimentation in regulatory developmental toxicology. So far, evaluation of the EST was not performed using chemicals within distinct chemical classes. Evaluation within a distinct class of structurally related chemicals can define the usefulness of the assay for the chemical class tested. The aim of the present study was to evaluate the relative sensitivity of the EST for a selected series of homologous chemicals and to compare the data to the relative developmental toxicity of the chemicals *in vivo*. To this end a series of proximate developmentally toxic glycol ether alkoxyacetic acid metabolites was tested in the embryonic stem cell (ES-D3 cell) proliferation assay and the ES-D3 cell differentiation assay of the EST. All glycol ether alkoxyacetic acid metabolites tested showed a concentration-dependent inhibition of ES-D3 cell differentiation but not in ES-D3 cell proliferation. Methoxyacetic acid was the most potent chemical followed by ethoxyacetic acid, butoxyacetic acid and phenoxyacetic acid, respectively. The potency ranking of the chemicals in the ES-D3 cell differentiation assay corresponds with the available *in vivo* data. The relative differences between the potencies of the chemicals appeared more pronounced in the *in vivo* studies than in the ES-D3 cell differentiation assay. A possible explanation for this discrepancy could be the difference in the kinetics of the chemicals *in vivo* as compared with their *in vitro* kinetics. This study illustrates that the ES-D3 cell differentiation assay of the EST can be used to set priorities for developmental toxicity testing within classes of related chemicals.

Introduction

In the context of the European REACH (Registration, Evaluation, Authorization and restriction of CHemicals) legislation the evaluation of the developmental toxicity of a large number of chemicals is foreseen. Unfortunately, the current regulatory animal tests for the evaluation of developmental toxicity of chemicals require a high number of laboratory animals, are time-consuming and expensive (Van der Jagt *et al.*, 2004). Consequently, there is a great need to develop alternative approaches and methods (Schaafsma *et al.*, 2009). *In vitro* alternatives could accelerate the process of chemical safety testing compared with classical *in vivo* testing, also reducing costs and animal use.

Currently three *in vitro* developmental toxicity assays have been formally validated, namely the rat postimplantation whole embryo culture test, the rat limb bud micromass test and the embryonic stem cell test (EST) (Genschow *et al.*, 2004). Of these three assays, only the last does not require pregnant animals to obtain embryonic tissue. Embryonic stem cell lines have been derived from the inner cell mass of blastocysts and have the advantage of being able to differentiate into all cell types of the three primary germ layers (Martin, 1981). The EST uses this ability to study the effect of chemicals on the differentiation of embryonic stem cells into contracting cardiomyocytes as an indication of developmental toxicity. The assay has been validated under the supervision of the European Centre for the Validation of Alternative Methods (ECVAM) in four different laboratories using 20 different chemicals resulting in an overall accuracy of 78% (Genschow *et al.*, 2004). However, it remains to be answered whether a validation study using 20 chemicals of diverse nature is sufficient to make a definitive statement on the validity of the test system. In addition, the added value of the use of 3T3 cytotoxicity in the prediction model has been questioned (Spielmann *et al.*, 2006). The applicability domain of the EST in terms of the chemical classes for which the test makes reliable predictions is still not sufficiently defined to allow regulatory implementation and may vary with different classes of chemicals tested (Hartung *et al.*, 2004; Spielmann *et al.*, 2006). Therefore, the applicability of the assay in a so-called category approach, which assumes that the *in vitro* ranking of chemicals of a certain chemical category corresponds with the potency ranking of these chemicals *in vivo*, need to be established. Therefore, evaluation of the EST within a distinct

class of structurally related chemicals can define the usefulness of the assay for the chemical class under study. The aim of the present study was to evaluate the performance of the embryonic stem cell (ES-D3 cell) proliferation and the ES-D3 cell differentiation assay of the EST for a selected series of homologous chemicals. To this end a series of glycol ethers and their glycol ether alkoxyacetic acid metabolites (Figure 1) was tested in the EST in two independent laboratories.

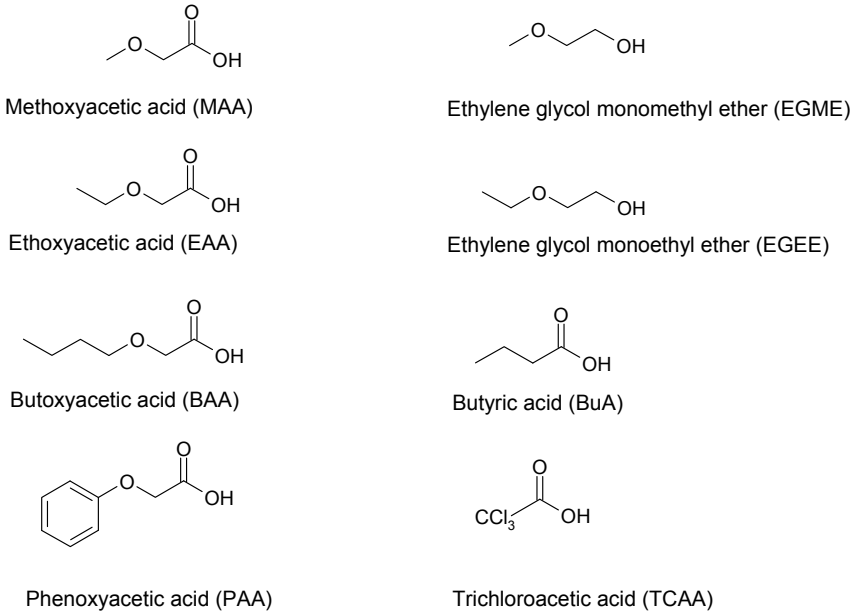

Figure 1. Chemical structure of the glycol ethers (EGME, EGEE), glycol ether alkoxyacetic acid metabolites (MAA, EAA, BAA, PAA) and two structural analogues (BuA, TCAA) included in the present study.

Results were compared among chemicals as well as between laboratories, and with available developmental toxicity data from *in vivo* animal studies. This class of chemicals was selected based on the availability of sufficient *in vivo* developmental toxicity data in order to allow comparison with the EST data. Also, these chemicals showed differences in embryotoxic potencies among the chemicals within the class, which facilitates the *in vitro-in vivo* comparison. Glycol ethers are widely used as solvents in cosmetics, printing inks, plastics, household and industrial cleaning products and textile dyes. Some of them have been shown to cause developmental toxicity through several routes of exposure in mice, rats and rabbits (Feuston *et al.*,

1990; Hanley *et al.*, 1984; Nagano *et al.*, 1981; Wier *et al.*, 1987). Frequently occurring malformations include spina bifida occulta, fused, small or missing digits and exencephaly. As the glycol ether alkoxyacetic acid metabolites have been identified as the proximate developmental toxicants in animal studies and in whole embryo cultures (Brown *et al.*, 1984; Giavini *et al.*, 1993), the glycol ether alkoxyacetic acid metabolites methoxyacetic acid (MAA), ethoxyacetic acid (EAA), butoxyacetic acid (BAA) and phenoxyacetic acid (PAA) were selected for the present study together with the parent glycol ethers of MAA and EAA, being ethylene glycol monomethyl ether (EGME) and ethylene glycol monoethyl ether (EGEE), respectively (Figure 1). The specificity of the effects of the glycol ether metabolites in the EST was further studied by comparison with the effects of two structurally similar non-glycol ether organic acids, butyric acid (BuA) and trichloroacetic acid (TCAA) (Figure 1).

To determine the applicability of the ES-D3 cell proliferation and ES-D3 cell differentiation assay of the EST for glycol ethers, the inhibition concentration-response data obtained for several glycol ether alkoxyacetic acid metabolites were compared with *in vivo* developmental toxicity data of their parent compounds reported in the literature, using the benchmark dose (BMD) approach. In the BMD approach a dose-response curve is fitted to determine the BMD for *in vivo* data and the benchmark concentration (BMC) for *in vitro* data, corresponding to the dose or concentration at which a certain benchmark response (BMR) is reached (Crump, 1984). The method has several principal advantages above the classical No Observed Adverse Effect Level (NOAEL) approach, as discussed elsewhere (Slob, 2002). One of the uncertainties in the BMD approach, as with any type of dose-response modeling, lies in the model chosen to mimic the "true" dose-response curve. The choice of the model fitted could influence the BMC values calculated and thereby influence the obtained ranking of the chemicals. Therefore, multiple models were used in the analysis to address this uncertainty. Another issue is the choice of the BMR used to determine the BMC values for the potency ranking. When concentration-response curves are not of a similar shape, different response levels could lead to a difference in the ranking of the chemicals based on the BMCs derived *in vitro*. This study investigated these uncertainties by deriving BMC values for multiple BMR values and with multiple types of dose-response models based on ES-D3 cell differentiation assay data for the glycol ether alkoxyacetic acid metabolites as the selected chemical class. In addition, the effect of interlaboratory

variation on chemical ranking was addressed by testing the same chemical set in two independent laboratories. Finally, the resultant ranking was compared with *in vivo* data to assess the predictive value of the EST for the developmental toxicity of glycol ethers.

Materials and methods

Cell line and culture conditions

The mouse embryonic stem cell line D3 (ATCC, Rockville, MD) was grown in Dulbecco's modified Eagle's medium (DMEM) (Invitrogen, Breda, The Netherlands) at 37 ^{o}C and 5% CO_2 and routinely passaged three times per week. Culture medium was supplemented with 20% fetal calf serum (Hyclone, Thermo-Fisher Scientific, Etten-Leur, The Netherlands), 2 mM glutamine (Invitrogen), 50 U/ml penicillin (Invitrogen), 50 µg/ml streptomycin (Invitrogen), 1% non-essential amino acids (Invitrogen), 0.1 mM β-mercaptoethanol (Sigma-Aldrich, Zwijndrecht, The Netherlands) and 1000 U/ml murine leukemia inhibiting factor (mLIF) (ESGRO®, Chemicon International, Amsterdam, The Netherlands).

Chemicals tested

All chemicals were obtained from Sigma-Aldrich unless stated otherwise. The glycol ether alkoxyacetic acid metabolites MAA, EAA, BAA (Tokyo Chemical Industries, Zwijndrecht, Belgium) and PAA were selected (Figure 1). Other test chemicals included the structural analogues TCAA and BuA and the parent glycol ethers EGME and EGEE. All chemicals were dissolved and diluted in DMSO (0.1% v/v final concentration in culture medium) with the exception of MAA, EAA, EGME and EGEE which were dissolved and diluted directly in culture medium. 5-Fluorouracil, dissolved in phosphate buffered saline (PBS, Invitrogen), was used as a positive control in the ES-D3 cell differentiation assay (Genschow *et al.*, 2004).

ES-D3 cell proliferation assay

The ES-D3 cell proliferation assay was carried out in Lab 1. To determine the potencies of the chemicals in inducing inhibition of ES-D3 cell proliferation, ES-D3 cells were seeded in 96-well plates at 500 cells per well in routine culture medium and allowed to attach for 2 h before chemical exposure at equal concentrations as

used in the differentiation assay. The cells were subsequently cultured for 5 days at 37 °C and 5% CO_2, with a medium renewal containing the corresponding chemical for exposure on day 3. On day 5, CellTiter-blue (Promega, Leiden, The Netherlands) was added to each well and incubated for 2 h. After the incubation period the fluorescence was measured using a FLUOstar spectrofluorometer (FLUOstar Optima, BMG Labtech, de Meeren, The Netherlands) at $\lambda_{excitation}$ = 544 nm and $\lambda_{emission}$ = 590 nm. Three independent experiments were performed for each chemical.

ES-D3 cell differentiation assay

The ES-D3 cell differentiation assay for cardiac differentiation was carried out in 2 labs as previously described using culture medium in the absence of mLIF (Genschow *et al.*, 2004; De Smedt *et al.*, 2008). Briefly, on day 0 of the assay, droplets of 20 µl of an ES-D3 cell suspension ($3.75 \cdot 10^4$ cells/ml) were placed onto the inner side of the lid of a 10-cm Petri dish (Greiner Bio-one, Alphen a/d Rijn, the Netherlands) containing 5 ml of phosphate buffered saline (PBS, Invitrogen) and incubated at 37 °C and 5% CO_2 for three days, in lab 1. In lab 2, droplets of 20 µl of an ES-D3 cell suspension ($3.75 \cdot 10^4$ cells/ml) were seeded on the inside of the lid of a 96-well plate (Corning, Schiphol-Rijk, the Netherlands). Caps of Eppendorf tubes were placed at the corners of the well plate to elevate the lid in order to prevent direct contact of the hanging drops with the well plate. Wells were filled with 250 µl PBS to obtain the right humidity and to prevent evaporation of the hanging drops. On day 3, the resulting embryoid bodies (EBs) were transferred to bacteriological Petri dishes (Greiner). On day 5, the EBs were transferred to 24-well plates (lab 1: TPP, Trasadingen, Switzerland; lab 2: Corning). On day 10, the number of wells containing contracting cardiomyocytes was determined by light microscopy. For each plate, the fraction of EBs differentiated into contracting cardiomyocytes was calculated. Cells were exposed to the test chemical in concentrations up to 1 mg/ml from day 0 onwards with a medium renewal containing the corresponding chemical for exposure on day 3 and day 5. Solvent controls were included in each experiment. Tests were accepted for further analysis if the solvent control plate contained at least 21 wells containing contracting cardiomyocytes out of 24 EBs incubated. Three independent experiments were done for each of the chemicals.

BMC and BMD derivation

In vitro data. The results from the EST were analyzed using the BMD approach, meaning that a concentration-response curve was fitted to determine the concentration associated with a predefined BMR. For each of the chemicals a 50% change in the number of contracting cardiomyocytes was selected as the BMR to calculate the benchmark concentrations for the differentiation ($BMCd_{50}$). BMD values at a BMR of 25 and 10% were also determined to study the effect of the choice of the response level on the ranking of the chemicals within the chemical class. For the concentration-response curves for chemical-induced inhibition of ES-D3 cell proliferation, a 50% inhibition of cell proliferation was selected as the BMR to calculate the benchmark concentrations for inhibition of cell proliferation ($BMCp_{50}$). For the continuous data from this assay a model was selected according to the procedure previously described by Slob (2002) in which a family of concentration-response models are fitted together. The goodness-of-fit was determined by the log-likelihoods of each model. The final model selected was the model with the lowest number of parameters which gave the best significant fit. For the quantal data from the differentiation assay several models with a statistically equal goodness-of-fit for the data were used to determine the effect of the model choice on the results. Models fitted included Logistic, Probit, Weibull and Gamma. The BMC calculations were performed using PROAST software (Slob, 2002).

In vivo data. For each of the chemicals a literature search was performed for *in vivo* developmental toxicity of the parent glycol ethers and their alkoxyacetic acid metabolites. For the glycol ether alkoxyacetic acid metabolites, which are considered as the proximate toxicants (Brown *et al.*, 1984; Giavini *et al.*, 1993), no relevant *in vivo* studies were available. Search terms included the chemical name together with combinations of the search terms teratogen, teratogenic, teratogenicity, development, developmental toxicity, malformation and embryotoxicity. Only studies containing multiple exposure times, multiple doses, and an oral exposure route were included in the data analysis for each of the glycol ethers. Fetal body weight and incidence of malformations were selected as the *in vivo* endpoints. The BMDs were calculated using dose-response modeling as done with the *in vitro* dataset for continuous data. When required, dose-response

modeling using constrained fit parameters was applied. The BMR for the BMD for each chemical was defined as a 10% decrease in fetal weight and a 10% additional incidence of malformations. These effect levels were selected because they could be estimated within each of the selected studies and could be distinguished from the background variation. This approach has previously been used by Piersma *et al.* (2008). The BMD calculations were performed using PROAST software (Slob, 2002).

Results

ES-D3 cell proliferation assay

ES-D3 cell proliferation assays were performed to evaluate the potencies of the chemicals on the inhibition of ES-D3 cell proliferation (Figure 2). The concentration-response curves obtained were used to calculate the $BMCp_{50}$ values corresponding to a 50% inhibition of ES-D3 cell proliferation (Table 1). MAA was the most potent of the glycol ether alkoxyacetic acid metabolites, reducing the ES-D3 cell proliferation with a $BMCp_{50}$ of 6.7 mM. The non-glycol ethers TCAA and BuA resulted in lower $BMCp_{50}$ values of 4.9 mM and 0.9 mM, respectively, reflecting higher toxic potencies of these chemicals than of the glycol ether alkoxyacetic acid metabolites tested. The alkoxyacetic acid metabolite BAA and the parent compounds EGME and EGEE did not cause an inhibition of ES-D3 cell proliferation.

Table 1. $BMCp_{50}$ values for inhibition of ES-D3 cell proliferation by four glycol ether alkoxyacetic acid metabolites and two structural analogues.

Chemical	$BMCp_{50}$ (mM) lab 1
MAA	6.7 (6.3 - 7.0)
EAA	12 (11 - 13)
BAA	n.d.
PAA	9.7 (8.6 - 12)
BuA	0.9 (0.8 - 0.9)
TCAA	4.9 (4.8 - 5.1)

The 90% confidence interval is given between brackets.
n.d.: not detected.

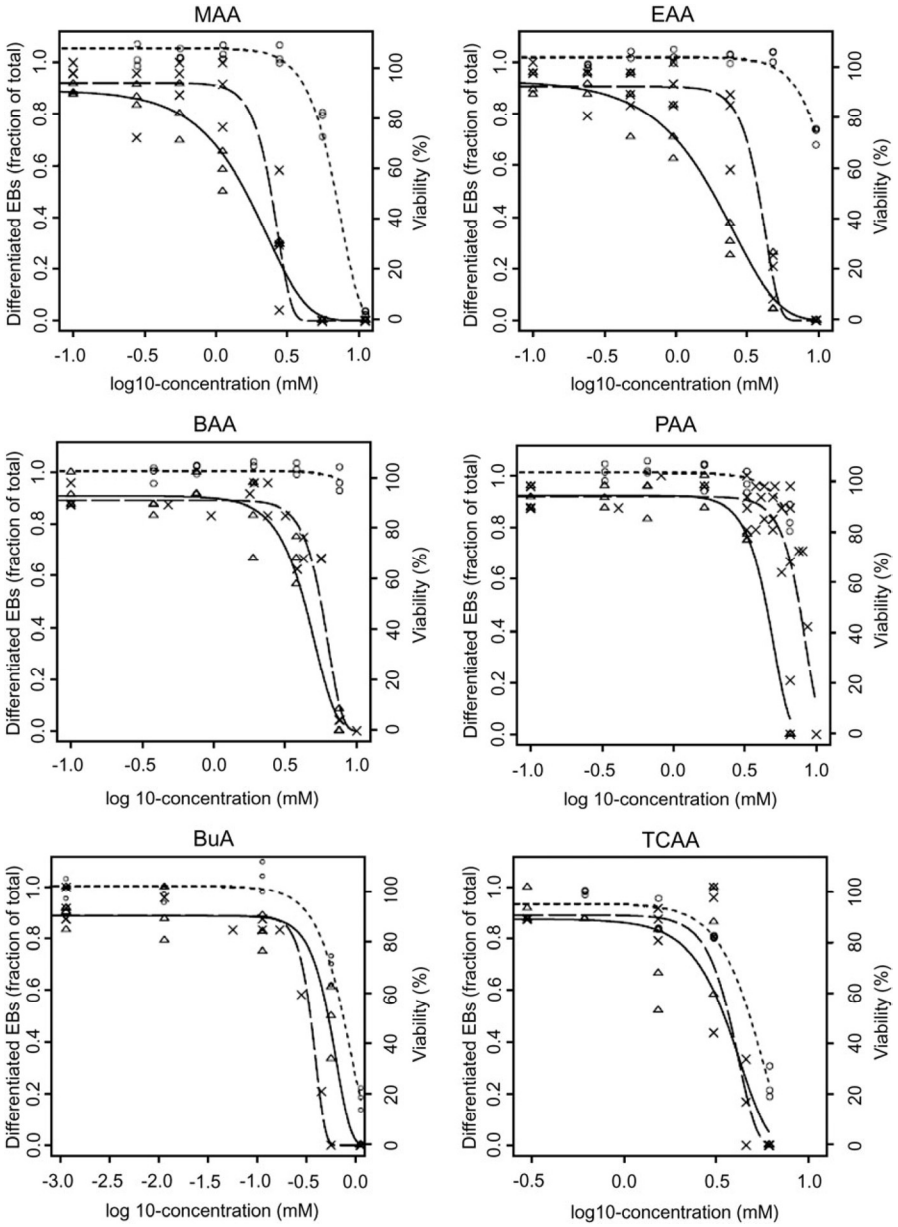

Figure 2. Concentration-dependent effects of glycol ether alkoxyacetic acid metabolites and two structural analogues on ES-D3 cell proliferation in lab 1 (------) and on ES-D3 cell differentiation into contracting cardiomyocytes in lab 1 (——) and lab 2 (— —).

ES-D3 cell differentiation assay

The effect of each of the chemicals on the differentiation of ES-D3 cells into contracting cardiomyocytes was evaluated in two independent laboratories (Figure 2). The Weibull model gave the best fit for the dataset obtained, based on the log-likelihood of the fitted curves and was used to determine the $BMCd_{50}$ values for both labs corresponding to a 50% decline in the fraction of contracting cardiomyocytes in comparison to the solvent control (Table 2).

Table 2. $BMCd_{25}$ and $BMCd_{50}$ values for inhibition of ES-D3 cell differentiation by four glycol ether alkoxyacetic acid metabolites and two structural analogues.

Chemical	$BMCd_{25}$ (mM) lab 1	$BMCd_{50}$ (mM) lab 1	$BMCd_{25}$ (mM) lab 2	$BMCd_{50}$ (mM) lab 2
MAA	1.7 (2.5 - 2.0)	2.3 (2.1 - 2.6)	2.0 (1.9 - 2.2)	2.5 (2.3 - 2.7)
EAA	2.1 (1.8 - 2.5)	2.9 (2.6 - 3.2)	3.2 (2.9 - 3.5)	3.9 (3.7 - 4.2)
BAA	3.4 (3.0 - 3.8)	4.5 (4.1 - 5.0)	4.8 (4.5 - 5.2)	5.9 (5.5 - 6.2)
PAA	3.7 (3.4 - 4.0)	4.6 (4.6 - 5.0)	6.4 (6.1 - 6.8)	7.8 (7.5 - 8.2)
BuA	0.4 (0.4 - 0.5)	0.6 (0.5 - 0.6)	0.3 (0.2 - 0.3)	0.4 (0.3 - 0.4)
TCAA	2.8 (2.4 - 3.1)	3.8 (3.4 - 4.1)	3.2 (2.9 - 3.4)	3.9 (3.7 - 4.1)

The 90% confidence interval is given between brackets. $BMCd_{25}$ = 25% reduction of the fraction of wells containing contracting cardiomyocytes; $BMCd_{50}$ = 50% reduction of the fraction of wells containing contracting cardiomyocytes.

The results show that all the glycol ether alkoxyacetic acid metabolites exert a concentration-dependent inhibition of ES-D3 cell differentiation into contracting cardiomyocytes. MAA was the most potent of the four glycol ether alkoxyacetic acid metabolites followed by EAA, BAA and PAA, respectively (Table 2). In both laboratories, BuA had a strong effect on the differentiation of ES cells into contracting cardiomyocytes with a $BMCd_{50}$ value four times lower than the most potent glycol ether alkoxyacetic acid metabolite MAA. Unlike the glycol ether alkoxyacetic acid metabolites, both parent compounds tested (EGME and EGEE) did not inhibit differentiation (data not shown), corroborating that the metabolites are the proximate developmental toxicants *in vivo* after exposure to the parent compounds. Therefore, the potency of the glycol ether alkoxyacetic acid metabolites in the EST was compared with *in vivo* toxicity data of the glycol ethers. For all chemicals, the ES-D3 cell differentiation assay was more sensitive than the

ES-D3 cell proliferation assay. This is indicated by the differences between the $BMCp_{50}$ and the $BMCd_{50}$ values, which amount to 2.9-fold for MAA, 4.1-fold for EAA, ≥ 2-fold for BAA, 2.1-fold for PAA, 1.5-fold for BuA and 1.3-fold for TCAA. Because all alkoxyacetic acid metabolites were active in the ES-D3 cell differentiation assay, but not in the ES-D3 proliferation assay, at the concentrations tested, the data obtained in the ES-D3 cell differentiation were used for comparison with *in vivo* toxicity data of the parent glycol ethers.

The interlaboratory reproducibility of the results was determined by comparing the $BMCd_{50}$ values obtained at the two labs (Figure 3). Interlaboratory variation is represented by the R^2, where the R^2 can range between 0 (no correlation) and 1 (equal values). The R^2 of 0.89 indicates a good correlation between the two laboratories whereas the $BMCd_{50}$ values were on average about 1.3-fold higher in Lab 2 than in Lab 1.

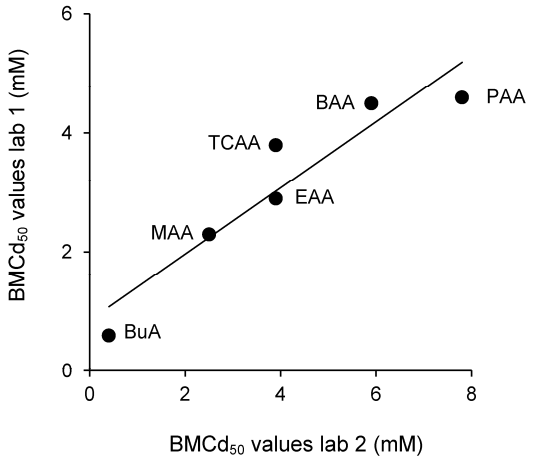

Figure 3. Scatter plot for the interlaboratory comparison of $BMCd_{50}$ values of MAA, EAA, BAA, PAA, BuA and TCAA in the ES-D3 cell differentiation assay of the EST.

Choice of BMR and dose-response model

For the *in vitro* data set a 50% decline in differentiated EBs was arbitrarily selected as the BMR for the BMCd derivation to compare the results of the different chemicals. The ranking of the different chemicals might differ with the magnitude of the BMR chosen. This possibility was investigated by also calculating the $BMCd_{25}$

values for each of the chemicals at both labs (Table 2). The choice for $BMCd_{25}$ instead of $BMCd_{50}$ values did not affect the ranking within the group of chemicals. Furthermore, using an even lower BMR of 10% and using the $BMCd_{10}$ values for comparison did also not affect the ranking of the chemicals (data not shown). In addition, the effect of the choice of the fitted model was evaluated by calculating the $BMCd_{50}$ values with different models for each of the chemicals (Table 3). For the $BMCd_{50}$ values it was shown that the choice of model did not significantly affect the values obtained. Furthermore, the ranking of the chemicals was not affected when using different models. Similarly, the choice of model did not significantly affect the $BMCd_{25}$ or $BMCd_{10}$ values and did not modify their ranking (data not shown).

Table 3. Effect of fitted model on $BMCd_{50}$ values for inhibition of ES-D3 cell differentiation by four glycol ether alkoxyacetic acid metabolites and two structural analogues in two laboratories.

Chemical	Gamma	LogLogistic	LogProbit	Weibull
Lab 1				
MAA	2.4 (2.1 - 2.7)	2.5 (2.3 - 2.7)	2.5 (2.3 - 2.7)	2.3 (2.1 - 2.6)
EAA	2.7 (2.4 - 3.0)	2.4 (2.2 - 2.6)	2.6 (2.4 - 2.8)	2.9 (2.6 - 3.2)
BAA	4.5 (4.1 - 4.9)	4.5 (4.1 - 4.8)	4.5 (4.2 - 4.9)	4.5 (4.1 - 5.0)
PAA	4.6 (4.3 - 5.0)	4.6 (4.3 - 4.9)	4.6 (4.3 - 4.9)	4.6 (4.6 - 5.0)
BuA	0.6 (0.5 - 0.6)	0.6 (0.5 - 0.6)	0.6 (0.5 - 0.6)	0.6 (0.5 - 0.6)
TCAA	3.9 (3.6 - 4.3)	4.0 (3.7 - 4.4)	4.0 (3.7 - 4.3)	3.8 (3.4 - 4.1)
Lab 2				
MAA	2.5 (2.3 - 2.6)	2.3 (2.1 - 2.5)	2.5 (2.3 - 2.6)	2.5 (2.3 - 2.7)
EAA	3.7 (3.4 -4.0)	3.4 (3.1 - 3.8)	3.6 (3.3 - 4.0)	3.9 (3.7 - 4.2)
BAA	5.7 (5.4 - 6.1)	5.5 (5.1 - 5.9)	5.7 (5.3 - 6.1)	5.9 (5.5 - 6.2)
PAA	7.7 (7.4 - 8.1)	8.2 (7.6 -8.9)	7.7 (7.4 - 8.1)	7.8 (7.5 - 8.2)
BuA	0.3 (0.3 - 0.4)	0.3 (0.3 - 0.4)	0.3 (0.3 - 0.4)	0.4 (0.3 - 0.4)
TCAA	3.8 (3.6 - 4.0)	3.4 (3.2 - 3.7)	3.8 (3.5 - 4.0)	3.9 (3.7 - 4.1)

The 90% confidence interval is given between brackets.

BMD derivation from in vivo data

Table 4 provides the details of the animal studies used to establish the BMD values. Studies were selected to represent the most uniform study design available between the different studies with respect to the animal species used, exposure route and exposure duration. For all chemicals a mouse developmental toxicity study with the parent glycol ethers was available with exposure from gestational day (GD) 7 or 8 to GD14 by oral exposure with the exception of ethylene glycol monophenyl ether (EGPE). For EGPE, a mouse study with exposure from premating day 7 to birth was available, which was included in the present study. BMD values were determined that correspond with a 10% reduction in fetal body weight and with a 10% additional incidence of malformations. EGME was the most potent embryotoxic chemical followed by EGEE. Ethylene glycol monobutyl ether (EGBE) exposure led to an increase in malformations and a reduction of fetal body weight with BMD values slightly higher than those of EGEE. However, these effects occurred only at EGBE doses that were maternally toxic, and thus they may be in part secondary to maternal toxicity. In the available *in vivo* study for EGPE, a slight decrease in fetal weight only occurred after exposure to a relatively high dose of 4 g/kg bw/day indicating that this glycol ether has the lowest developmental toxicity potency within the series.

Discussion

The aim of the present study was to assess the usefulness of the ES-D3 cell proliferation assay and the ES-D3 cell differentiation assay of the EST in category approaches to predict *in vivo* developmental toxicity within series of glycol ethers and their glycol ether alkoxyacetic acid metabolites. The present study showed that not all glycol ether alkoxyacetic acid metabolites inhibited ES-D3 cell proliferation at the concentrations tested, whereas all inhibited the differentiation of ES-D3 cells into contracting cardiomyocytes in a concentration-dependent manner. MAA appeared to be the most potent ES-D3 cell differentiation-inhibiting chemical, followed by EAA, BAA and PAA, respectively. Because all alkoxyacetic acid metabolites were active in the ES-D3 cell differentiation assay, but not in the ES-D3 proliferation assay, at the concentrations tested, the $BMCd_{50}$ values obtained in the

Table 4. Developmental toxicity of glycol ethers in mice.

Chemical	Mice strain	Route	Exposure	Dose (mg·kg bw^{-1}·day $^{-1}$)	Reference	BMD_{10} fetal weight (mmol·kg bw^{-1}·day^{-1})	BMD_{10} malformation (mmol·kg bw^{-1}·day^{-1})
EGME (MAA)	JCL-ICR	G	GD7-14	0, 31, 63, 125, 250	Nagano *et al.* (1981)	1.3[a]	0.8[a]
EGEE (EAA)	CD-1	G	GD8-14	0, 1000, 1800, 2600, 3400, 4200	Wier *et al.* (1987)	4.9[a]	14[b]
EGBE (BAA)	CD-1	G	GD8-14	0, 350, 650, 1000, 1500, 2000	Wier *et al.* (1987)	12[a]	15[b]
EGPE (PAA)	CD-1	D	7PM-B	0, 400, 2000, 4000	Heindel *et al.* (1990)	31[a]	n.a.

BMD_{10} fetal weight: dose (mmol · kg bw^{-1} · day^{-1}) at 10% decrease in fetal body weight; BMD_{10} malformation: dose (mmol · kg bw^{-1} · day^{-1}) at a 10% increase in malformations; G: gastric incubation; D: diet; GD: gestational day; PM: premating; B: birth; n.a.: data not available.

[a] BMDs were calculated from a model dose-response.

[b] BMDs were calculated from a model dose-response using constrained fit parameters.

ES-D3 cell differentiation were compared with *in vivo* developmental toxicity data for the corresponding parent glycol ethers, based on the BMD_{10} values for fetal weight and malformations derived from the *in vivo* data for the parent glycol ethers.

The *in vitro* potency ranking of the glycol ether alkoxyacetic acid metabolites was found to be the same as the *in vivo* potency ranking based on the BMD_{10} values of the parent glycol ethers. However, the relative differences between the *in vitro* $BMCd_{50}$ levels of the four chemicals were smaller than the relative differences in BMD_{10} levels of the chemicals found *in vivo*. In view of the observed effects on malformations in the animal studies, EGME (the parent compound of MAA) is 17-fold more potent than EGEE (the parent compound of EAA). In contrast, there was only a 1.4-fold difference between MAA and EAA in the *in vitro* study. In addition, the $BMCd_{50}$ values for BAA were twofold higher in the *in vitro* ES-D3 cell differentiation assay compared with those for MAA, whereas in the *in vivo* animal studies there is about a 19-fold difference between the BMD_{10} values for EGME and EGBE (the parent compounds of MAA and BAA, respectively). With respect to the *in vivo* BMD_{10} values it should be mentioned that EGBE only induced malformations at maternally toxic doses. EGBE has been shown to induce severe hemolytic anemia following gavage and dermal exposure (Ghanayem *et al.*, 2001). In general, a possible role of maternal toxicity in the mediation of developmental effects should be kept in mind, but a useful *in vitro* correlate is not easily defined.

The differences between the relative potencies *in vitro* and *in vivo* may be explained by differences in the kinetics of these chemicals in the *in vivo* model as compared with the *in vitro* model system. Differences in *in vivo* kinetics of the glycol ethers and their metabolites are reflected by their half-life values. After intravenous injection MAA was shown to have a low elimination rate with a half-life in blood ranging between 14 and 19 h, whereas EAA and BAA displayed a shorter half-life ranging between 7.6 and 10 h for EAA and ranging between 1.5 and 3.2 h for BAA (Aasmoe and Aarbakke, 1997; Aasmoe *et al.*, 1999; Ghanayem *et al.*, 1990). This indicates that MAA is eliminated more slowly than EAA and BAA, resulting in higher plasma levels, which may explain the relatively higher embryotoxic potency of its parent glycol ether *in vivo*. The *in vitro* system does not simulate the *in vivo* kinetic processes and, therefore, the *in vitro* levels should be extrapolated to *in vivo* levels by integrating information on *in vivo* kinetics of the

chemicals. To overcome these issues, physiologically based kinetic modeling may prove to be a way to ultimately link the *in vitro* $BMCd_{50}$ values to BMD_{10} values from *in vivo* studies (Piersma *et al.*, 2008; Verwei *et al.*, 2006). Work on such *in vitro-in silico* models for the developmental toxicity of glycol ethers is presented in chapter 5.

The two structurally related non-glycol ether chemicals, TCAA and BuA, also caused a concentration-dependent inhibition of ES-D3 cell differentiation into cardiomyocytes. These obtained effect concentrations, however, were close to the obtained effect concentrations for the inhibition of ES-D3 cell proliferation, indicating that the effect on proliferation and differentiation may have been caused by a cytotoxic effect of the chemicals. This observation demonstrates that it is informative to compare $BMCd_{50}$ values with $BMCp_{50}$ values of the test chemicals for the interpretation of results obtained with the EST.

In the present study possible sources of uncertainties influencing the results were investigated as well. First, we have studied interlaboratory variation. The high correlation between the $BMCd_{50}$ values of the two laboratories with an R^2 of 0.89 demonstrates the interlaboratory reproducibility of the ES-D3 cell differentiation assay. The 1.3-fold difference in the sensitivity of the ES-D3 cell differentiation assay observed between laboratories in $BMCd_{50}$ values may be due to subtle differences in culture conditions. However, we show that this difference did not interfere with the overall conclusions in regard to the relative potency of the glycol ether metabolites. Furthermore, the results have shown that the selection of different models for the curve fitting did not affect the overall outcome. Finally, changing the arbitrarily chosen *in vitro* BMR of a 50% decline in the number of contracting cardiomyocytes to a 25 or 10% decline did not affect the overall ranking of the potency of the chemicals either. These findings could be expected on the basis of the data structure in our experiments, because with an adequate distribution of test concentrations over the entire concentration-response range, the concentration-response models will deliver similar fits and $BMCd_{50}$ values, as illustrated here.

In conclusion, the ES-D3 cell differentiation assay of the EST is a useful tool for the identification of the relative embryotoxic potencies of glycol ethers as a chemical class, by testing their alkoxyacetic acid metabolites. However, the

different kinetic processes in the ES-D3 cell differentiation assay as compared with the kinetic processes in the *in vivo* situation may affect the interpretation and should be taken into account in the quantitative extrapolation of the *in vitro* results to the *in vivo* situation. Combining the obtained concentration-response curves with a physiologically based kinetic model describing the *in vivo* kinetics of the glycol ethers could lead to a more accurate link of the *in vitro* results from the EST to the data from the *in vivo* studies. To the best of our knowledge this is the first study to evaluate the EST for embryotoxic potencies within a series of related chemicals. This study shows that the ES-D3 cell differentiation assay of the EST can be used to set priorities for developmental toxicity testing within series of related chemicals. Prior knowledge on major metabolites of the parent compound of interest can focus the choice of derivatives to be tested in the EST. A similar approach could be applied to other chemical classes to further define the applicability domain of the EST in terms of the groups of chemicals for which the test can be used and to set priorities in future animal testing for developmental toxicity. This approach can be useful in regulatory settings such as the European chemicals legislation REACH, which anticipates significant additional animal testing while at the same time promoting the use of alternatives wherever appropriate (Van der Jagt *et al.*, 2004).

References

Aasmoe, L., Aarbakke, J. (1997). Gender difference in the elimination of 2-methoxyethanol, methoxyacetic acid and ethoxyacetic acid in rat. *Xenobiotica* **27**, 1237-44.

Aasmoe, L., Mathiesen, M., Sager, G. (1999). Elimination of methoxyacetic acid and ethoxyacetic acid in rat. *Xenobiotica* **29**, 417-44.

Brown, N. A., Holt, D., Webb, M. (1984). The teratogenicity of methoxyacetic acid in the rat. *Toxicol Lett* **22**, 93-100.

Crump, K. S. (1984). A new method for determining allowable daily intakes. *Fundam Appl Toxicol* **4**, 854-71.

De Smedt, A., Steemans, M., de Boeck, M., Peters, A. K., van der Leede, B. J., van Goethem, F., Lampo, A., Vanparys, P. (2008). Optimisation of the cell cultivation methods in the embryonic stem cell test results in an increased differentiation potential of the cells into strong beating myocard cells. *Toxicol In Vitro* **22**, 1789-96.

Feuston, M. H., Kerstetter, S. L., Wilson, P. D. (1990). Teratogenicity of 2-methoxyethanol applied as a single dermal dose to rats. *Fundam Appl Toxicol* **15**, 448-56.

Genschow, E., Spielmann, H., Scholz, G., Pohl, I., Seiler, A., Clemann, N., Bremer, S., Becker, K. (2004).

Validation of the embryonic stem cell test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* **32**, 209-44.

Ghanayem, B. I., Long, P. H., Ward, S. M., Chanas, B., Nyska, M., Nyska, A. (2001). Hemolytic anemia, thrombosis, and infarction in male and female F344 rats following gavage exposure to 2-butoxyethanol. *Exp Toxicol Pathol* **53**, 97-105.

Ghanayem, B. I., Sanders, J. M., Clark, A. M., Bailer, J., Matthews, H. B. (1990). Effects of dose, age, inhibition of metabolism and elimination on the toxicokinetics of 2-butoxyethanol and its metabolites. *J Pharmacol Exp Ther* **253**, 136-43.

Giavini, N. A., Broccia, M. L., Menegola, E., Prati, M. (1993). Comparative in vitro study of the embryotoxic effects of three glycol ethers and their metabolites, the alkoxyacids. *Toxicol in Vitro* **7**, 777-84.

Hanley, T. R. Jr., Yano, B. L., Nitschke, K. D., John, J. A. (1984). Comparison of the teratogenic potential of inhaled ethylene glycol monomethyl ether in rats, mice, and rabbits. *Toxicol Appl Pharmacol* **75**, 409-22.

Hartung, T., Bremer, S., Casati, S., Coecke, S., Corvi, R., Fortaner, S., Gribaldo, L., Halder, M., Hoffmann, S., Roi, A. J., Prieto, P., Sabbioni, E., Scott, L., Worth A., Zuang, V. (2004). A modular approach to the ECVAM principles on test validity. *Altern Lab Anim* **32**, 467-72.

Heindel, J. J., Gulati, D. K., Russell, V. S., Reel, J. R., Lawton, A. D., Lamb, J. C. (1990). Assessment of ethylene glycol monobutyl and monophenyl ether reproductive toxicity using a continuous breeding protocol in Swiss CD-1 mice. *Fundam Appl Toxicol* **15**, 683-96.

Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci USA* **78**, 7634-8.

Nagano, K., Nakayama, E., Oobayashi, H., Yamada, T., Adachi, H., Nishizawa, T., Ozawa, H., Nakaichi, M., Okuda, H., Minami, K., Yamazaki, K. (1981). Embryotoxic effects of ethylene glycol monomethyl ether in mice. *Toxicology* **20**, 335-43.

Piersma, A. H., Janer, G., Wolterink, G., Bessems, J. G. M., Hakkert, B. C., Slob, W. (2008). Quantitative extrapolation of in vitro whole embryo culture embryotoxicity data to developmental toxicity in vivo using the benchmark dose approach. *Toxicol Sci* **101**, 91-100.

Schaafsma, G., Kroese, E. D., Tielemans, E. L., van de Sandt, J. J. M., van Leeuwen, C. J. (2009). REACH, non-testing approaches and the urgent need for a change in mind set. *Regul Toxicol Pharmacol* **53**, 70-80.

Slob, W. (2002). Dose-response modeling of continuous endpoints. *Toxicol Sci* **66**, 298-312.

Spielmann, H., Seiler, A., Bremer, S., Hareng, L., Hartung, T., Ahr, H., Faustman, E., Haas, U., Moffat, G. J., Nau, H., Vanparys, P., Piersma, A. H., Sintes, J. R., Stuart, J. (2006). The practical application of three validated in vitro embryotoxicity tests. The report and recommendations of an ECVAM/ZEBET workshop (ECVAM workshop 57). *Altern Lab Anim* **34**, 527-38.

Van der Jagt, K., Munn, S. J., Tørsløv, J., de Bruijn, J. (2004). Alternative approaches can reduce the use of test animals under REACH. *Report EUR* **21405**.

Verwei, M., van Burgsteden, J. A., Krul, C. A., van de Sandt, J. J. M., Freidig, A. P. (2006). Prediction of in vivo embryotoxic effect levels with a combination of in vitro studies and PBPK modelling. *Toxicol Lett* **165**, 79-87.

Wier, P. J., Lewis, S. C., Traul, K. A. (1987). A comparison of developmental toxicity evident at term to postnatal growth and survival using ethylene glycol monoethyl ether, ethylene glycol monobutyl ether and ethanol. *Teratog Carcinog Mutagen* **7**, 55-64.

Chapter 5 is reprinted with permission from Oxford University Press.

CHAPTER 5

The use of *in vitro* toxicity data and physiologically based kinetic modeling to predict dose-response curves for *in vivo* developmental toxicity of glycol ethers in rat and human

Jochem Louisse, Esther de Jong, Johannes J.M. van de Sandt, Bas J. Blaauboer, Ruud A. Woutersen, Aldert H. Piersma, Ivonne M.C.M. Rietjens, Miriam Verwei

Based on:
Toxicological Sciences 118, 470-484 (2010).

Chapter 5

Abstract

At present, regulatory assessment of systemic toxicity is almost solely carried out using animal models. The European Commission's REACH legislation stimulates the use of animal-free approaches to obtain information on the toxicity of chemicals. *In vitro* toxicity tests provide *in vitro* concentration-response curves for specific target cells, whereas *in vivo* dose-response curves are regularly used for human risk assessment. The present study shows an approach to predict *in vivo* dose-response curves for developmental toxicity by combining *in vitro* toxicity data and *in silico* kinetic modeling. A physiologically based kinetic (PBK) model was developed, describing the kinetics of four glycol ethers and their developmentally toxic alkoxyacetic acid metabolites in rat and human. *In vitro* toxicity data of these metabolites derived in the embryonic stem cell test were used as input in the PBK model to translate *in vitro* concentration-response curves into predicted *in vivo* dose-response curves for developmental toxicity of the parent glycol ethers in rat and human. The predicted dose-response curves for rat were found to be in concordance with the developmental toxicity dose levels measured in reported *in vivo* rat studies. Therefore, predicted dose-response curves for rat could be used to set a point of departure for deriving safe exposure limits in human risk assessment. Combining the *in vitro* toxicity data with a human PBK model allows the prediction of dose-response curves for human developmental toxicity. This approach could therefore provide a means to reduce the need for animal testing in human risk assessment practices.

Introduction

The implementation of the European REACH (Registration, Evaluation, Authorization and restriction of CHemicals) legislation will lead to the evaluation of the toxicity of a large number of chemicals. The European Commission estimated the number of laboratory animals required for REACH to amount to 3.9 million if the use of alternative methods is not accepted by regulatory authorities. In addition, they indicated that with the expected acceptance scenario for alternative approaches, the total number will amount to 2.6 million (Van der Jagt *et al.*, 2004). Taking into account the offspring produced during the studies, a number of 9 million laboratory animals was estimated (Höfer *et al.*, 2004). Hartung and Rovida (2009) calculated an even higher number of laboratory animals required for REACH, which was contradicted by the European Chemical Agency (ECHA, 2009). Irrespective of the actual number of laboratory animals required for REACH, the development of validated and accepted *in vitro* and *in silico* approaches is urgently needed. Because it is expected that more than 25% of the laboratory animals needed for REACH will be used for developmental toxicity studies (Van der Jagt *et al.*, 2004), *in vitro* and *in silico* alternatives for developmental toxicity studies could contribute substantially to the reduction of animal use.

Three alternative methods for *in vivo* developmental toxicity tests have been scientifically validated so far by the European Centre for the Validation of Alternative Methods (ECVAM), which are the rat post implantation whole-embryo culture test, the rat limb bud micromass test and the embryonic stem cell test (EST) (Genschow *et al.*, 2002). Only the EST does not require live animals because a mouse embryonic stem cell line is used. De Jong *et al.* (2009) showed that the embryonic stem cell (ES-D3 cell) differentiation assay of the EST, which assesses the effects of chemicals on the differentiation of mouse embryonic stem cells into contracting cardiomyocytes, can be used to rank the potency of chemicals within a series of alkoxyacetic acid metabolites formed from glycol ethers. The alkoxyacetic acid metabolites of the glycol ethers, and not the parent glycol ethers themselves, have been identified as the proximate developmental toxicants of these chemicals (Brown *et al.*, 1984; Giavini *et al.*, 1993). The *in vitro* potencies of the developmentally toxic glycol ether alkoxyacetic acid metabolites, as measured in the ES-D3 cell differentiation assay, were found to correspond with the

developmental toxicity potency of the corresponding glycol ethers *in vivo* (De Jong *et al.*, 2009). However, the ES-D3 cell differentiation assay provides *in vitro* concentration-response curves, whereas for human risk assessment *in vivo* dose-response curves are often required. These *in vitro* data should therefore be translated into *in vivo* data by taking into account *in vivo* kinetics (Verwei *et al.*, 2006).

The goal of the present study was to integrate *in vitro* toxicity data and *in silico* kinetic modeling as an approach to predict dose-response curves for developmental toxicity in both rat and human, thereby providing a basis for the use of *in vitro* toxicity data in human risk assessment. To this end, four glycol ethers (ethylene glycol monomethyl ether (EGME), ethylene glycol monoethyl ether (EGEE), ethylene glycol monobutyl ether (EGBE) and ethylene glycol monophenyl ether (EGPE)) were used as model chemicals belonging to one chemical class, but showing differences in *in vivo* developmental toxicity potencies. To take the *in vivo* kinetics into consideration, the *in vitro* data derived from the ES-D3 cell differentiation assay were used as input in a physiologically based kinetic (PBK) model, translating *in vitro* effect concentrations into predicted *in vivo* developmental toxicity dose levels, as described by Verwei *et al.* (2006).

The predicted developmental toxicity dose levels are based on the assumption that concentrations of the toxic alkoxyacetic acid metabolites that cause an inhibition of embryonic stem cell differentiation *in vitro* will also cause toxic effects in the developing embryo *in vivo*. By using the ES-D3 cell differentiation assay data as input for the blood concentration in the PBK model, the model allows the calculation of the *in vivo* dose levels of the parent glycol ethers that will lead to these blood concentrations of the toxic alkoxyacetic acid metabolites *in vivo*. By calculating this dose level for each concentration tested in the ES-D3 cell differentiation assay, *in vitro* concentration-response curves for the toxic alkoxyacetic acid metabolites can be converted into predicted *in vivo* dose-response curves for the parent glycol ethers. This dose-response curve could be used to set a point of departure, like a $BMDL_{10}$ (lower limit of the 95% confidence interval on the benchmark dose at which a benchmark response equivalent to a 10% effect size (BMR_{10}) is reached (BMD_{10})), for deriving safe exposure levels for humans. Therefore, the combined *in vitro-in silico* approach provides a platform to use *in*

vitro toxicity data for risk assessment practices, thereby contributing to the reduction of animal use in chemical risk assessment for developmental toxicity endpoints.

Materials and methods

Chemicals

EGPE, ethylene glycol monopropyl ether, HCl and phenoxyacetic acid (PAA) were purchased from Sigma-Aldrich (Zwijndrecht, the Netherlands). Ethyl acetate and methanol were obtained from Biosolve (Valkenswaard, the Netherlands) and sodium sulfate (Na_2SO_4) from Merck (Darmstadt, Germany). Ammonium formate was purchased from Fluka (Zwijndrecht, the Netherlands).

In vitro-in silico approach to predict dose-response curves for in vivo developmental toxicity in rat and human

The *in vitro-in silico* approach to predict developmental toxicity dose levels in rats, as proposed by Verwei *et al.* (2006), was used to estimate dose-response curves for the *in vivo* developmental toxicity of the glycol ethers EGME, EGEE, EGBE and EGPE in both rat and human. To this end the following steps were followed in the present study, which are described in more detail in the following sections: 1) development of a PBK model for *in vivo* glycol ether kinetics in rat and human, 2) evaluation of the PBK model, 3) determination of *in vitro* effect concentrations in the ES-D3 cell differentiation assay to be used as input for the PBK model, 4) PBK model-based prediction of dose-response curves for *in vivo* developmental toxicity in rat and human and 5) evaluation of the potential of the *in vitro-in silico* approach to predict dose-response curves for *in vivo* developmental toxicity in the rat.

Development of a PBK model for in vivo glycol ether kinetics in rat and human

A generic PBK model describing the kinetics of the glycol ethers EGME, EGEE, EGBE and EGPE and their alkoxyacetic acid metabolites methoxyacetic acid (MAA), ethoxyacetic acid (EAA), butoxyacetic acid (BAA) and PAA, respectively, in rat and human was developed, using the model of EGME and MAA of Gargas *et al.*

(2000a) as starting point. A schematic representation of the model is shown in Figure 1.

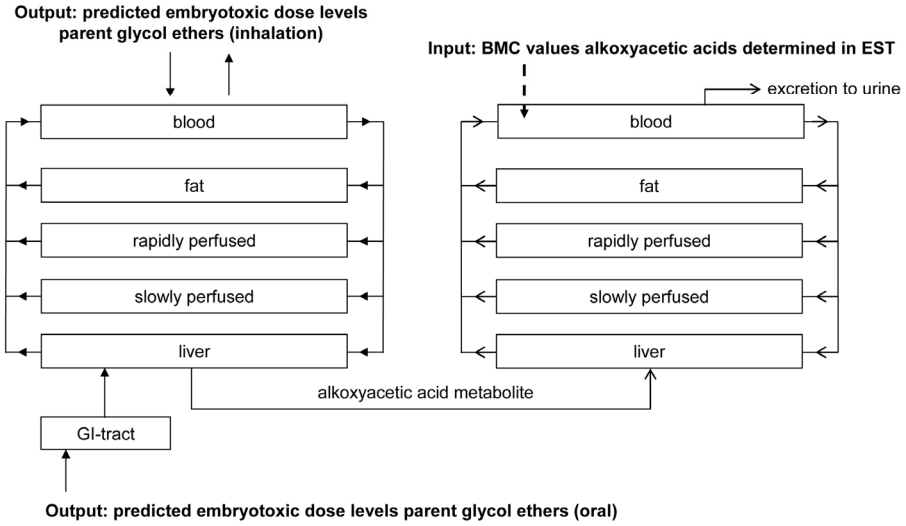

Figure 1. Schematic representation of the PBK model of glycol ethers and their alkoxyacetic acid metabolites in rat and human. The PBK model is based on the PBK model of EGME and MAA described by Gargas *et al.* (2000a). GI tract: gastrointestinal tract.

Physiological parameter values for the rat and human model describing *in vivo* kinetics after inhalation exposure were taken from Gargas *et al.* (2000a) (Table 1) and represent the physiological parameter values at the beginning of pregnancy. In the present study, the PBK model was adjusted to also describe *in vivo* kinetics after oral exposure. Oral uptake of the glycol ethers was described by an oral absorption constant (k_a) fitted to the *in vivo* plasma concentrations of EGME (Hays *et al.*, 2000) and EGPE (present study) in rats that were orally dosed. Partition coefficients of the glycol ethers and their alkoxyacetic acid metabolites for rat and human tissues were estimated using the equation described by Berezhkovskiy (2004) (Table 2). Biotransformation kinetics of EGME, EGEE and EGBE determined in rat and human hepatocytes by Green *et al.* (1996) were scaled to a whole liver by assuming a number of $128 \cdot 10^6$ hepatocytes/g rat liver (Seglen, 1976) and $99 \cdot 10^6$ hepatocytes/g human liver (Barter *et al.*, 2007).

Table 1. Physiological parameter values used in the PBK model for glycol ethers taken from Gargas *et al.* (2000a).

Physiological parameters	Values	
	rat	human
body weight (kg)	0.25	60
tissue volume (percentage of body weight)		
liver	4.0	2.4
fat	10.1	27.6
rapidly perfused tissue	6.1	3.7
slowly perfused tissue	65	48.7
blood	5.9	5.9
cardiac output ($L \cdot h^{-1} \cdot kg\ bw^{-0.74}$)	14.0	19.2
tissue blood flows (percentage of cardiac output)		
liver:hepatic artery	3.4	3.3
liver:portal vein	21.6	20.7
fat	14.2	9.5
rapidly perfused tissue	45.8	47.5
slowly perfused tissue	15.0	19.0
alveolar ventilation ($L \cdot h^{-1} \cdot kg\ bw^{-0.74}$)	14.0	15.3

No data on the biotransformation kinetics of EGPE to PAA are available in the literature. Because the biotransformation kinetics of EGME, EGEE and EGBE are alike, it is expected that they are close to those of EGPE. Because the properties that might describe the biotransformation kinetics of EGPE, such as logP and VanderWaals volume (Chang *et al.*, 2009; Soffers *et al.*, 2001), are closest to those of EGBE (Table 3), the biotransformation kinetics of EGBE were used in the EGPE PBK model. Literature studies describing the kinetics of EGME (Gargas *et al.*, 2000a; Hays *et al.*, 2000), EGEE acetate (EGEEA) (Gargas *et al.*, 2000b) and EGBE (Ghanayem *et al.*, 1990) in rats were used to obtain the parameter values for urinary excretion of MAA, EAA and BAA, respectively, in rats by using a urinary excretion constant (K_{ex}) in the model fitted to the *in vivo* plasma concentrations of the alkoxyacetic acid metabolites (Table 2). Because no *in vivo* EGPE kinetic study in rats was found in the literature, an *in vivo* kinetic study with a limited number of rats was carried out to determine the parameter values for urinary excretion of PAA

Table 2. Chemical-specific parameter values used in the PBK model for glycol ethers.

Partition coefficients[a]	**liver:blood**	**fat:blood**	**rapidly perfused tissue:blood**	**slowly perfused tissue:blood**	**blood:air**
EGME	0.76	0.37	0.76	0.80	32800[b]
MAA	0.76	0.13	0.76	0.80	
EGEE	0.77	0.67	0.77	0.80	22093[b]
EAA	0.76	0.13	0.76	0.79	
EGBE	0.73	3.7	0.73	0.72	7965[b]
BAA	0.64	0.11	0.64	0.67	
EGPE	0.85	8.4	0.85	0.79	5651[c]
PAA	0.64	0.11	0.64	0.67	

Biotransformation kinetics	**rat**		**human**	
	apparent K_m (mM)	apparent V_{max} (nmol · h^{-1} · 10^6 hepatocytes^{-1})	apparent K_m (mM)	apparent V_{max} (nmol · h^{-1} · 10^6 hepatocytes^{-1})
EGME → MAA[d]	6.3	1511	1.7	61.3
EGEE → EAA[d]	6.6	1519	1.2	70.8
EGBE → BAA[d]	0.9	741	0.9	113
EGPE → PAA[e]	0.9	741	0.9	113

Excretion kinetics [f]	**rat**	**human**
	K_{ex} (L · h^{-1})	K_{ex} (L · h^{-1})
MAA	0.0045	0.18
EAA	0.025	0.30
BAA	0.060	5.0
PAA	0.080	2.0

[a] Estimated using *in silico* method described by Berezhkovskiy (2004).
[b] Determined by Johanson and Dynésius (1988).
[c] Estimated based on linear correlation between LogP values of EGME, EGEE and EGBE, and the logarithm of their blood:air partition coefficients.
[d] Data from biotransformation studies in hepatocytes (Green *et al.*, 1996).
[e] Assumed to be similar as for EGBE (see the "Materials and Methods" section).
[f] Fitted to *in vivo* plasma concentrations of alkoxyacetic acid metabolites (MAA: Hays *et al.*, 2000; Gargas *et al.*, 2000a; EAA: Gargas *et al.*, 2000b; BAA: Ghanayem *et al.*, 1990; PAA: data from present study) for the rat model and to *in vivo* urinary excretion rates (MAA: Groeseneken *et al.*, 1989; EAA: Groeseneken *et al.*, 1986; BAA: Jones and Cocker, 2003; Kezic *et al.*, 2004) or *in vivo* plasma concentrations (2,4-D: Sauerhoff *et al.*, 1977, used for excretion rate of PAA) for the human model.

Table 3. LogP values and VanderWaals volumes of glycol ethers.

Chemical	LogP[a]	VanderWaals volume[b] (Å^3)
EGME	-0.61	88
EGEE	-0.22	106
EGBE	0.84	143
EGPE	1.19	153

[a] LogP values calculated using ChemBioDraw Ultra 12.0.
[b] VanderWaals volumes calculated using Spartan 04 for Windows version 1.0.3.

in rats. Model parameter values for urinary excretion of MAA, EAA and BAA in human were estimated using *in vivo* kinetic studies of EGME (Groeseneken *et al.*, 1989), EGEE (Groeseneken *et al.*, 1986) and EGBE (Jones and Cocker, 2003; Kezic *et al.*, 2004), respectively, in human, by using a K_{ex} in the model fitted to the *in vivo* urinary excretion rates of the alkoxyacetic acid metabolites (Table 2). Because no study on human EGPE kinetics is described in the literature, the urinary excretion of PAA was determined by fitting the K_{ex} to *in vivo* plasma concentrations of 2,4-dichlorophenoxyacetic acid (2,4-D) of human orally exposed to 2,4-D (Sauerhoff *et al.*, 1977) (Table 2). The present study shows that PAA's plasma half-life in the rat is almost equal to 2,4-D's plasma half-life in rat (see the "Discussion" section). Therefore, it was assumed that PAA's plasma half-life in human will be the same as 2,4-D's plasma half-life in human, making it possible to use the *in vivo* kinetic data for 2,4-D in human to estimate the K_{ex}, which is directly related to the plasma half-life, for the human PAA model.

Evaluation of the PBK model

To evaluate the PBK model describing *in vivo* kinetics of the glycol ethers, *in vivo* kinetic studies for rat and human reported in literature were used (Table 4). The doses, exposure routes and exposure duration used in the studies were applied in the PBK model simulation, to evaluate the model (Table 4). For the rat EGEE model, an *in vivo* kinetic study with EGEEA was used. No *in vivo* kinetic studies of EGPE were present in the literature. Therefore, *in vivo* EGPE kinetics in rats were determined in the present study, as described in the following. No kinetic data of EGPE in human were available, so for the human EGPE PBK model no evaluation could be carried out.

Table 4. Published kinetic studies used to build and/or evaluate the PBK model.

Species	Chemical	Reference	Exposure route	Exposure	Dose(s)	Data shown
rat	EGME / MAA	Hays *et al.* (2000)	oral	once	3.3 mmol·kg bw^{-1}	Figure 2A
	EGME / MAA	Gargas *et al.* (2000a)	inhalation	5 days, 6h·day^{-1}	10, 50 ppm	Figure 2B
	EGEE (EGEEA) / EAA	Gargas *et al.* (2000b)	inhalation	5 days, 6h·day^{-1}	50, 100 ppm	Figure 2C
	EGBE / BAA	Ghanayem *et al.* (1990)	intravenous	once	0.3, 0.5, 1.0 mmol·kg bw^{-1}	Figure 2D
	EGPE / PAA	present study	oral	once	1.1, 3.3 mmol·kg bw^{-1}	Figure 2E
human	EGME / MAA	Groeseneken *et al.* (1989)	inhalation	4h	5 ppm	Figure 3A
	EGEE / EAA	Groeseneken *et al.* (1986)	inhalation	4h	3, 6, 9 ppm	Figure 3B
	EGBE / BAA	Jones and Cocker (2003)	inhalation	2h	20 ppm	Figure 3C
	EGBE / BAA	Kezic *et al.* (2004)	inhalation	0.5h	19 ppm	Figures 3C,D

EGPE kinetic study in rat. The experimental protocol of the EGPE kinetic study was approved by the Animal Welfare Committee of the National Institute of Public Health and the Environment (RIVM). Female Sprague-Dawley rats (254.4 ± 5.2 g) at approximately 13 weeks of age were obtained from Harlan Laboratories (Horst, the Netherlands). Rats were maintained in a temperature-, humidity- and light cycle-controlled facility for 1 week prior to exposure. Feed (SDS, Witham, Essex, UK) and water were provided *ad libitum*. Two groups of 4 rats per group were exposed to either 1.1 mmol · kg body weight (bw)$^{-1}$ EGPE or 3.3 mmol · kg bw^{-1} EGPE by oral gavage. Blood samples of 150 µl were taken from the tail vein prior to EGPE exposure and at 5, 15 and 30 min and 1, 2, 4, 6, 8 and 24 h after dosing and were stored in heparin-coated tubes (Sarstedt, Nümbrecht, Germany). Samples were centrifuged at 1300 g for 10 min. The acquired serum samples were stored at -20 °C.

EGPE and PAA analysis of plasma samples using gas chromatography-mass spectrometry and liquid chromatography-mass spectrometry. To determine the EGPE and PAA plasma concentrations, 20 µl plasma was added to 20 µl of a 0.5% (v/v) HCl solution. Then, 1 ml ethyl acetate, containing the internal standard ethylene glycol monopropyl ether, was added and samples were vortexed thoroughly. Sodium sulfate was added as a drying agent, after which samples were vortexed and centrifuged. The supernatant was used for EGPE analysis using gas chromatography-mass spectrometry (GC-MS) and PAA analysis using liquid chromatography-mass spectrometry (LC-MS).

EGPE was analyzed with an Agilent HP 5972 GC-MS system (Hewlett-Packard, Palo Alto, CA) equipped with a 60 m × 0.25 mm × 0.5 µm Stabilwax-DA column (Restek; Interscience B.V., Breda, the Netherlands). The oven temperature was initially maintained for 5 min at 60 °C after injection and then increased in steps of 15 °C · min^{-1} to 250 °C, which was held for 15 min. Samples of 1 µl were injected using a programmed temperature vaporization splitless mode and helium as carrier gas with a constant flow rate of 1.5 ml · min^{-1}. A positive electron ionization method was used for EGPE and ethylene glycol monopropyl ether (internal standard) detection. The ratio of EGPE (m/z = 94) and ethylene glycol monopropyl ether (m/z = 73) was used for quantification.

LC-MS analysis was carried out with an HP Agilent 1100 system (Hewlett-Packard) equipped with a Waters Atlantis C18 T3 5-µm column, 150 × 3.0 mm (Waters,

Etten-Leur, the Netherlands), coupled to a Waters Quattro Premier XE mass spectrometer (Waters). Aliquots of 5 µl were injected. The flow rate was 0.3 ml · min^{-1}. A gradient was made using a 5 mM ammonium formate in methanol/H_2O (20/80, v/v) solution (solution A) and a 5 mM ammonium formate in methanol/ H_2O (90/10, v/v) solution (solution B). A linear gradient was applied from 0 to 100% solution B over 5 min, after which solution B was kept at 100% for another 5 min, lowered to 0% in 0.5 min, and equilibrated at these initial conditions for 8.5 min. A positive electrospray ionization mode was used for mass spectrometrical analysis. Sample analysis was carried out by the multiple reaction monitoring scan mode. The m/z transition 150.77 → 92.60 was used for PAA quantification.

Data analysis of EGPE kinetic study in rats. The plasma concentrations of EGPE and PAA determined in the EGPE kinetic study were used to calculate EGPE's and PAA's half-lives in plasma after oral dosing of 1.1 or 3.3 mmol EGPE · kg bw^{-1}, using Kinetica 4.2.

Sensitivity analysis. A sensitivity analysis was carried out to identify the key parameters highly influencing the model output (peak blood concentrations of alkoxyacetic acid metabolites). The effect of a 5% increase in parameter value was evaluated by calculating normalized sensitivity coefficients (SC = (C' - C)/(P' - P) · (P/C)) as was done by Evans and Andersen (2000) (C = initial value of model output, C' = value of model output resulting from 5% increase in parameter value, P = initial parameter value, P' = 5% increased parameter value). This 5% change in parameter value is theoretical and does not necessarily reflect realistic variations. Sensitivity analyses were carried out for the rat and human models using a single oral exposure of 1 mmol · kg bw^{-1} and an 8-h inhalation exposure of 50 ppm.

Determination of in vitro effect concentrations in the ES-D3 cell differentiation assay

The *in vitro* developmental toxicity effect data derived in two different labs in our previous study (De Jong *et al.*, 2009) were used to calculate benchmark concentrations (BMCs) at which 1-99% (BMC_1 to BMC_{99}) of the embryoid bodies (aggregations of embryonic stem cells) did not differentiate into contracting cardiomyocytes, using United States Environmental Protection Agency (EPA)'s benchmark dose software (BMDS) version 2.0 (Weibull model). These data were used as input in the PBK model to predict dose-response curves for *in vivo*

developmental toxicity of the glycol ethers in rat and human, based on the *in vitro* ES-D3 cell differentiation assay data of both labs.

PBK model-based prediction of dose-response curves for in vivo developmental toxicity in rat and human

To predict the dose-response curves for *in vivo* developmental toxicity of the glycol ethers, derived BMCs for the inhibition of embryonic stem cell differentiation by the alkoxyacetic acids were used as input for the peak blood concentrations in the PBK model (Figure 1). The model was then used to calculate the corresponding dose levels of the parent glycol ethers leading to these concentrations in the blood, resulting in the predicted dose-response curves for *in vivo* developmental toxicity. For rat, dose-response curves were predicted using identical exposure regimens (route of exposure and exposure duration) in the PBK model simulations as applied in the *in vivo* developmental toxicity studies (Table 5). For human, dose-response curves were predicted for single- and repeated-exposure regimens, to investigate the effect of repeated dosing. For repeated dosing, a 5-day exposure regimen was used (8 h · day^{-1} for inhalation exposure), representing a working week exposure.

Evaluation of the potential of the in vitro-in silico approach to predict dose-response curves for in vivo developmental toxicity in rat

To evaluate the potential of the *in vitro-in silico* approach to predict dose-response curves for *in vivo* developmental toxicity, the predicted dose-response curves for the rat were compared with data obtained in *in vivo* developmental toxicity studies listed in the ECETOC Technical Report 095 (ECETOC, 2005). Developmental toxicity endpoints described include resorptions, malformations, fetal deaths and decrease in fetal bw. The *in vivo* data were used to derive $BMDL_{10}$ values on the basis of dose-response analyses (BMDS version 2.0), which can be used as points of departure in human risk assessment practices (Barlow *et al.*, 2009). These $BMDL_{10}$ values derived from *in vivo* data were compared to predicted $BMDL_{10}$ values, which were acquired by translating the *in vitro* $BMCL_{10}$ (lower limit of the 95% confidence interval on the BMC_{10}) values into dose levels using the rat PBK model. No human glycol ether developmental toxicity data were available. Consequently, predicted dose-response curves for developmental toxicity in human

Table 5. Comparison of $BMDL_{10}$ values determined from *in vivo* developmental toxicity studies and predicted $BMDL_{10}$ values for developmental toxicity using the *in vitro-in silico* approach with *in vitro* $BMCL_{10}$ values as input (symbols shown correspond with symbols used in Figure 5). (Table continues on following page.)

Chemical	Reference	Exposure route	Days of exposure	Doses	Critical endpoint	Measured $BMDL_{10}$	Predicted $BMDL_{10}$		Figure	Symbol
							Lab 1	Lab 2		
EGME	Toraason *et al.* (1985)	oral	GD7-13	0.3, 0.6, 1.1 mmol·kg bw^{-1}·day^{-1}	cardiac malformations	0.46 mmol·kg bw^{-1}·day^{-1}	0.12 mmol·kg bw^{-1}·day^{-1}	0.34 mmol·kg bw^{-1}·day^{-1}	5A	●
					resorptions	0.52 mmol·kg bw^{-1}·day^{-1}				▲
					fetal bw decrease	0.14 mmol·kg bw^{-1}·day^{-1}				♦
EGME	Doe *et al.* (1983)	inhalation	GD6-17, 6h·day^{-1}	100, 300 ppm	fetal deaths	n.d.			5E	○
EGEE	Nelson *et al.* (1984)	inhalation	GD7-15, 7h·day^{-1}	50, 100, 200 ppm	skeletal malformations	11 ppm	21 ppm	64 ppm	5E	♦
					visceral malformations	41 ppm				■
					resorptions	49 ppm				●
					fetal bw decrease	37 ppm				▲
EGEE	Goad and Cranmer (1984)	oral	GD7-15	2.2 mmol·kg bw^{-1}·day^{-1}	cardiac malformations	n.d.			5B	♦
EGEE	Stenger *et al.* (1971)	oral	GD1-21	0.1, 0.3, 0.5, 1.0, 2.1 mmol·kg bw^{-1}·day^{-1}	skeletal malformations	0.80 mmol·kg bw^{-1}·day^{-1}	0.28 mmol·kg bw^{-1}·day^{-1}	1.3 mmol·kg bw^{-1}·day^{-1}	5B	●
EGEE	Chester *et al.* (1986)[a]	oral	GD1-21	2.3, 3.0, 4.4 mmol·kg bw^{-1}·day^{-1}	fetal deaths	n.d.			5B	▲
EGEE	Doe (1984)	inhalation	GD6-15, 6h·day^{-1}	10, 50, 250 ppm	skeletal malformations[b]	66 ppm	62 ppm	280 ppm	5F	Δ
					visceral malformations[b]	n.d.				○
					fetal bw decrease	n.d.				◊
					post implantation loss	n.d.				□
EGEE	Andrew and Hardin (1984)	inhalation	GD1-18, 7h·day^{-1}	200, 770 ppm	resorptions	n.d.			5F	●
					cardiac malformations	n.d.				■
					reduced ossification[b]	n.d.				♦
					supernummary ribs[b]	n.d.				▲

Chemical	Reference	Exposure route	Days of exposure	Doses	Critical endpoint	Measured $BMDL_{10}$	Predicted $BMDL_{10}$		Figure	Symbol
							Lab 1	Lab 2		
EGBE	Sleet *et al.* (1989)	oral	GD9-11	1.3, 2.5[c] mmol·kg bw^{-1}·day^{-1}	resorptions	1.4 mmol·kg bw^{-1}·day^{-1}	2.6 mmol·kg bw^{-1}·day^{-1}	3.1 mmol·kg bw^{-1}·day^{-1}	5C	●
					skeletal malformations	n.d.				▲
			GD11-13	1.3, 2.5[c], 5.1[c] mmol·kg bw^{-1}·day^{-1}	resorptions	n.d.				♦
					skeletal malformations	3.6 mmol·kg bw^{-1}·day^{-1}	2.6 mmol·kg bw^{-1}·day^{-1}	3.1 mmol·kg bw^{-1}·day^{-1}		■
EGBE	Tyl *et al.* (1984)	inhalation	GD6-15, 6h·day^{-1}	25, 50, 100, 200[c] ppm	fetal deaths	n.d.			5G	○
					total malformations	n.d.				Δ
EGBE	Nelson *et al.* (1984)	inhalation	GD7-15, 7h·day^{-1}	150, 200 ppm	resorptions	n.d.			5G	●
					fetal bw decrease	n.d.				▲
					skeletal malformations	n.d.				♦
					visceral malformations	n.d.				■
EGPE	Unilever Research (1984)	subcutaneous	GD6-15	0.6, 1.3, 2.5[c] mmol·kg bw^{-1}·day^{-1}	total malformations	n.d.			5D	●
					fetal deaths	1.0 mmol·kg bw^{-1}·day^{-1}	2.4 mmol·kg bw^{-1}·day^{-1}	6.2 mmol·kg bw^{-1}·day^{-1}		▲

n.d.: $BMDL_{10}$ value could not be determined because of either no effect or unsuitability of data to carry out benchmark dose analysis.

[a] Data taken from ECETOC Technical Report 095 (ECETOC, 2005).

[b] Minor anomalies according to authors.

[c] Dose at which maternal toxicity including hemolysis was observed.

could not be compared with data on developmental toxicity of glycol ethers in human.

Results

Development of a PBK model for in vivo glycol ether kinetics in rat and human

The parameter values for oral uptake (k_a) were fitted to *in vivo* kinetic data for EGME and EGPE, as described in the "Materials and Methods" section, and were found to be 4 h^{-1} in both cases. Therefore, in all models, a k_a of 4 h^{-1} was used for oral uptake of the parent glycol ethers. The parameter values for urinary excretion of the alkoxyacetic acid metabolites (K_{ex}), obtained by fitting these parameter values to *in vivo* kinetic data as described in the "Materials and Methods" section, are shown in Table 2. For the rat model, the K_{ex} value is the lowest for MAA, followed by that for EAA, BAA and PAA, respectively. For the human model, the K_{ex} value is the lowest for MAA, followed by that for EAA, PAA and BAA, respectively.

Evaluation of the PBK model

Figure 2 shows the PBK-model predicted plasma concentrations of EGME/MAA (Figures 2A and 2B), EGEE/EAA (Figure 2C), EGBE/BAA (Figure 2D) and EGPE/PAA (Figure 2E) in the rat and the measured plasma levels from *in vivo* rat kinetic studies. The plasma concentrations of EGPE and PAA after EGPE exposure were measured in the present study. The results of the PBK simulations show that the predicted plasma concentrations are close to the measured plasma concentrations of the four glycol ethers and their alkoxyacetic metabolites measured in rat. Figures 3A-C show the PBK-model predicted urinary excretion rates of MAA (Figure 3A), EAA (Figure 3B) and BAA (Figure 3C) in human and the measured urinary excretion rates from *in vivo* human kinetic studies. Figure 3D shows the prediction of the EGBE plasma concentrations in human and the measured EGBE plasma concentrations as measured in an *in vivo* human kinetic study. The results show that the model predicted urinary excretion rates and the predicted EGBE plasma levels are within a factor of 10 compared with the measured urinary excretion rates of MAA, EAA and BAA and the EGBE plasma concentrations measured in human.

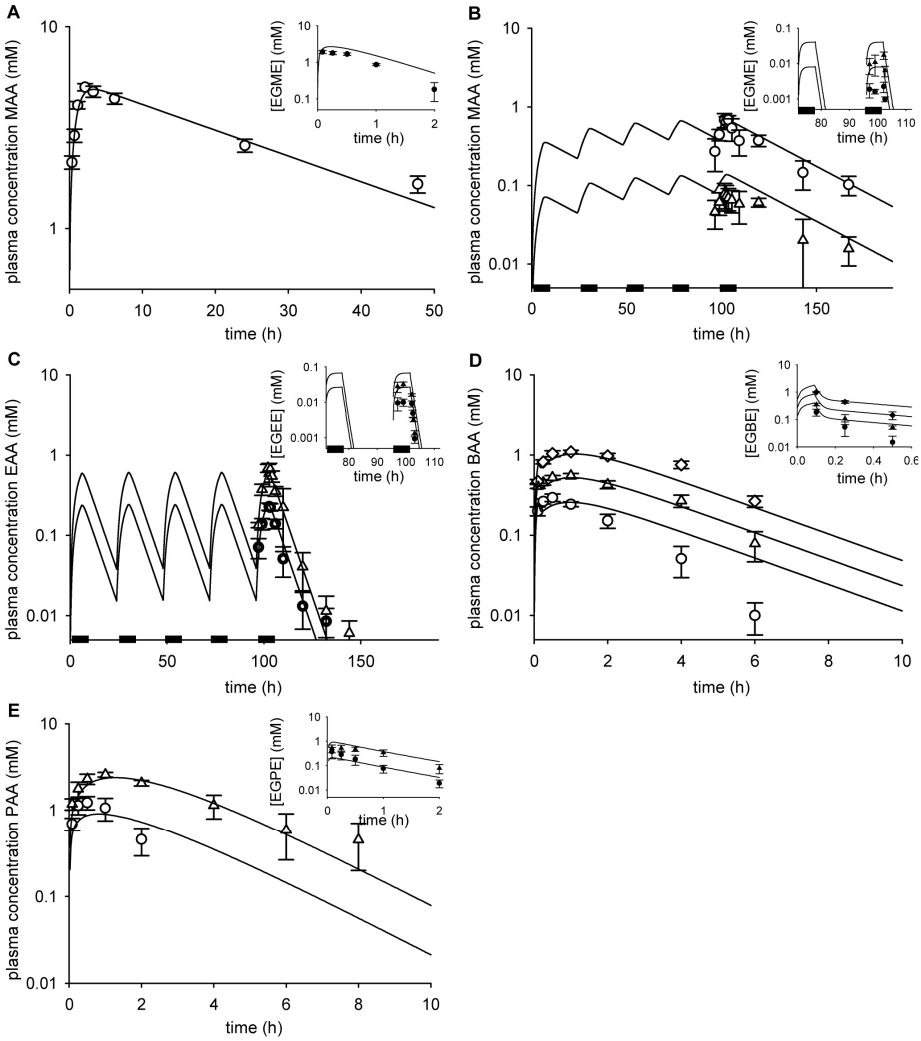

Figure 2. Measured and PBK-model predicted plasma concentrations of parent glycol ethers and alkoxyacetic acid metabolites in rats exposed to EGME (A, B), EGEE (C), EGBE (D) and EGPE (E). The main graphs show the plasma alkoxyacetic acid concentrations (open symbols), whereas the insets show the plasma parent glycol ether concentrations (closed symbols). The curves represent the PBK model predictions and the symbols with error bars (= SD) the measured *in vivo* data. (A) Oral exposure to 3.3 mmol EGME · kg bw^{-1} (Hays *et al.*, 2000); (B) 5 days (6 h · day^{-1}) of inhalation exposure to 10 (triangles) or 50 (circles) ppm EGME (Gargas *et al.*, 2000a). Thick lines on x-axis indicate time interval of glycol ether exposure. Plasma concentrations were determined on days 5, 6 and 7; (C) 5 days (6 h · day^{-1}) of inhalation exposure to 50 (circles) or 100 (triangles) ppm EGEE (Gargas *et al.*, 2000b). Thick lines on x-axis indicate time interval of glycol ether exposure. Plasma concentrations were determined on days 5, 6 and 7; (D) intravenous exposure to 0.26 (circles), 0.53 (triangles), or 1.1 (diamonds) mmol EGBE · kg bw^{-1} (Ghanayem *et al.*, 1990); (E) oral exposure to 1.1 (circles) or 3.3 (triangles) mmol EGPE · kg bw^{-1} (present study).

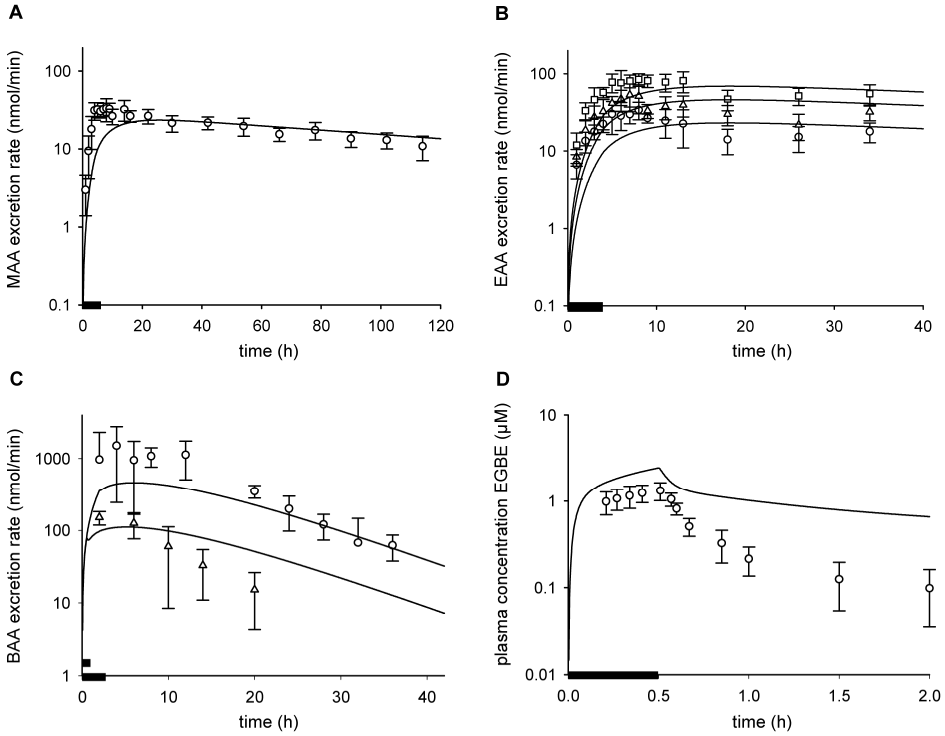

Figure 3. Measured and PBK-model predicted urinary excretion rates of MAA (A), EAA (B) and BAA (C) in human after inhalation exposure to EGME, EGEE and EGBE, respectively, and measured and predicted EGBE plasma concentrations (D) in human after inhalation exposure to EGBE. The curves represent the PBK model predictions and the symbols with error bars (= SD) the measured *in vivo* data. (A) 4-h inhalation exposure to 5 ppm EGME (Groeseneken *et al.*, 1989); (B) 4-h inhalation exposure to 3 (circles), 6 (triangles), or 9 (squares) ppm EGEE (Groeseneken *et al.*, 1986); (C) 2-h inhalation exposure to 20 ppm EGBE (Jones and Cocker, 2003; circles) and 0.5-h inhalation exposure to 19 ppm EGBE (Kezic *et al.*, 2004; triangles); (D) 0.5-h inhalation exposure to 19 ppm EGBE (Kezic *et al.*, 2004). Thick lines on x-axis indicate time interval of glycol ether exposure.

In general, prediction of *in vivo* kinetics of the glycol ethers by the PBK models was found to be somewhat better in rat than in human (Figures 2 and 3).

EGPE kinetics in rat. The plasma concentrations of EGPE and PAA in rats exposed to 1.1 and 3.3 mmol EGPE · kg bw^{-1} are shown in Figure 2E. Maximum EGPE plasma concentrations measured were already reached 5 min after exposure, and were 0.4 mM after exposure to 1.1 mmol EGPE · kg bw^{-1} and 0.5 mM after exposure to 3.3 mmol EGPE · kg bw^{-1}. The maximum PAA plasma concentration of 1.2 mM obtained upon exposure to 1.1 mmol EGPE · kg bw^{-1}, was reached 30 min after

exposure, whereas the maximal PAA plasma concentration of 2.6 mM obtained upon exposure to 3.3 mmol EGPE · kg bw^{-1}, was reached 60 min after exposure. EGPE was not detected in the plasma 4 h after exposure to both doses. The half-lives of EGPE in plasma were 0.4 and 0.9 h after exposure to 1.1 and 3.3 mmol EGPE · kg bw^{-1}, respectively. PAA was not detected 4 h after exposure to 1.1 mmol EGPE · kg bw^{-1} and 24 h after exposure to 3.3 mmol EGPE · kg bw^{-1}. The half-life of PAA in the plasma was 1.0 h in the rats that were exposed to 1.1 mmol EGPE · kg bw^{-1} and 2.6 h in the rats that were exposed to 3.3 mmol EGPE · kg bw^{-1}.

Sensitivity analysis. Figure 4 shows the normalized sensitivity coefficients of the most sensitive model parameters for peak blood concentrations of the alkoxyacetic acids in the rat and human PBK model for EGME/MAA (single exposure). The results of the sensitivity analyses for the PBK models for the other glycol ethers were similar to those of the EGME PBK models (data not shown). The sensitivity analysis revealed that the most sensitive physiological parameters for peak blood

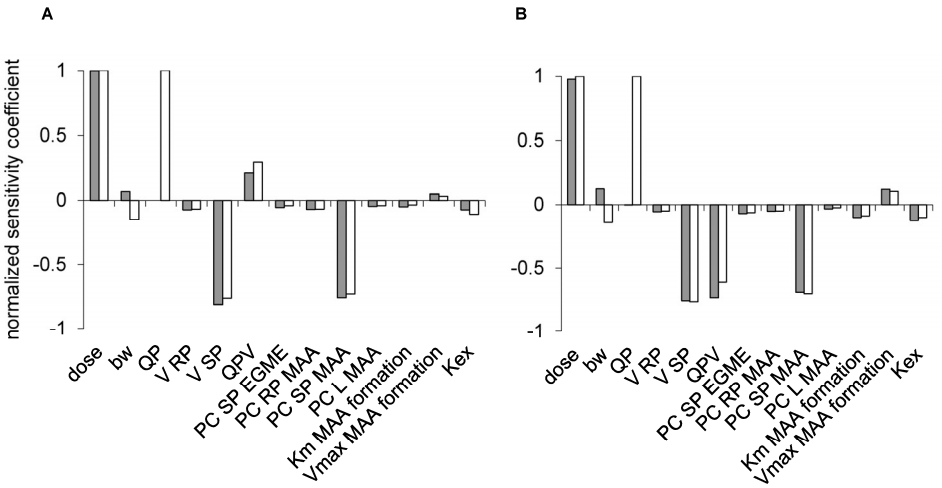

Figure 4. Sensitivity of the predicted peak MAA plasma concentrations to different model parameters of the rat (A) and human (B) PBK models for EGME and MAA, after oral exposure to 1 mmol EGME · kg bw^{-1} (grey bars) or inhalation exposure of 8 h to 50 ppm EGME (white bars). All model parameters with normalized sensitivity coefficients smaller than -0.03 or larger than 0.03 are shown. Dose: oral or inhaled dose EGME; bw: body weight; QP: alveolar ventilation rate; V RP: volume richly perfused tissue; V SP: volume slowly perfused tissue; QPV: blood flow through portal vein; PC SP EGME: partition coefficient EGME slowly perfused tissue:blood; PC RP MAA: partition coefficient MAA rapidly perfused tissue:blood; PC SP MAA: partition coefficient MAA slowly perfused tissue:blood; PC L MAA: partition coefficient MAA liver:blood; Km MAA formation: Michaelis-Menten constant for biotransformation of EGME to MAA; Vmax MAA formation: maximum rate of MAA formation; K_{ex}: MAA excretion rate.

concentrations of the alkoxyacetic acids are the alveolar ventilation rate (QP; with inhalation exposure), the blood flow through the portal vein (QPV) and the volume of the slowly perfused tissue compartment (V SP). The most sensitive chemical-specific parameters for peak blood concentrations of the alkoxyacetic acids appeared to be the slowly perfused tissue:blood partition coefficient of MAA (PC SP MAA) and the K_{ex}. All SCs for the remaining model parameters are between -0.2 and 0.2.

Comparison of predicted and observed dose-response curves for in vivo developmental toxicity in rat

Figure 5 shows the predicted dose-response curves of the glycol ethers EGME, EGEE, EGBE and EGPE in rat, presenting also the measured developmental toxicity data as derived from literature. The *in vitro* concentration-response curves (at BMC_1 to BMC_{99} values) for the alkoxyacetic acid metabolites in the ES-D3 cell differentiation assay of the EST (De Jong *et al.*, 2009) translated into *in vivo* dose-response curves are shown in Figure 5A (EGME), Figure 5B (EGEE) and Figure 5C (EGBE) for oral exposure; in Figure 5D (EGPE) for subcutaneous exposure; and in Figure 5E (EGME), Figure 5F (EGEE), Figure 5G (EGBE) and Figure 5H (EGPE) for inhalation exposure. Each figure shows two predicted dose-response curves, derived from the *in vitro* ES-D3 cell differentiation assay data obtained in each of the labs in our previous study (De Jong *et al.*, 2009). The individual data points in the graphs (Figure 5) represent developmental toxicity data from *in vivo* rat studies as reported in the literature (for studies used, see Table 5). For the predicted dose-response curves, the dose regimens used in the corresponding *in vivo* studies were applied in the PBK model simulations. In Figures 5E-G, *in vivo* data of inhalation studies are shown in which different exposure regimens have been used, that is, differences in the number of exposure days and differences in exposure duration per day (either 6 or 7 $h \cdot day^{-1}$). Predicted dose-response curves in these figures are shown for the studies using a 7 $h \cdot day^{-1}$ exposure regimen (Figures 5E-H). When applying a 6 $h \cdot day^{-1}$ exposure regimen in the PBK model simulations, predicted dose-response curves are slightly shifted to the right (data not shown). The developmental toxicity data from reported *in vivo* rat studies were used to derive $BMDL_{10}$ values that could be used for deriving safe exposure limits for human risk assessment. These $BMDL_{10}$ values derived from *in vivo* data are presented in Table

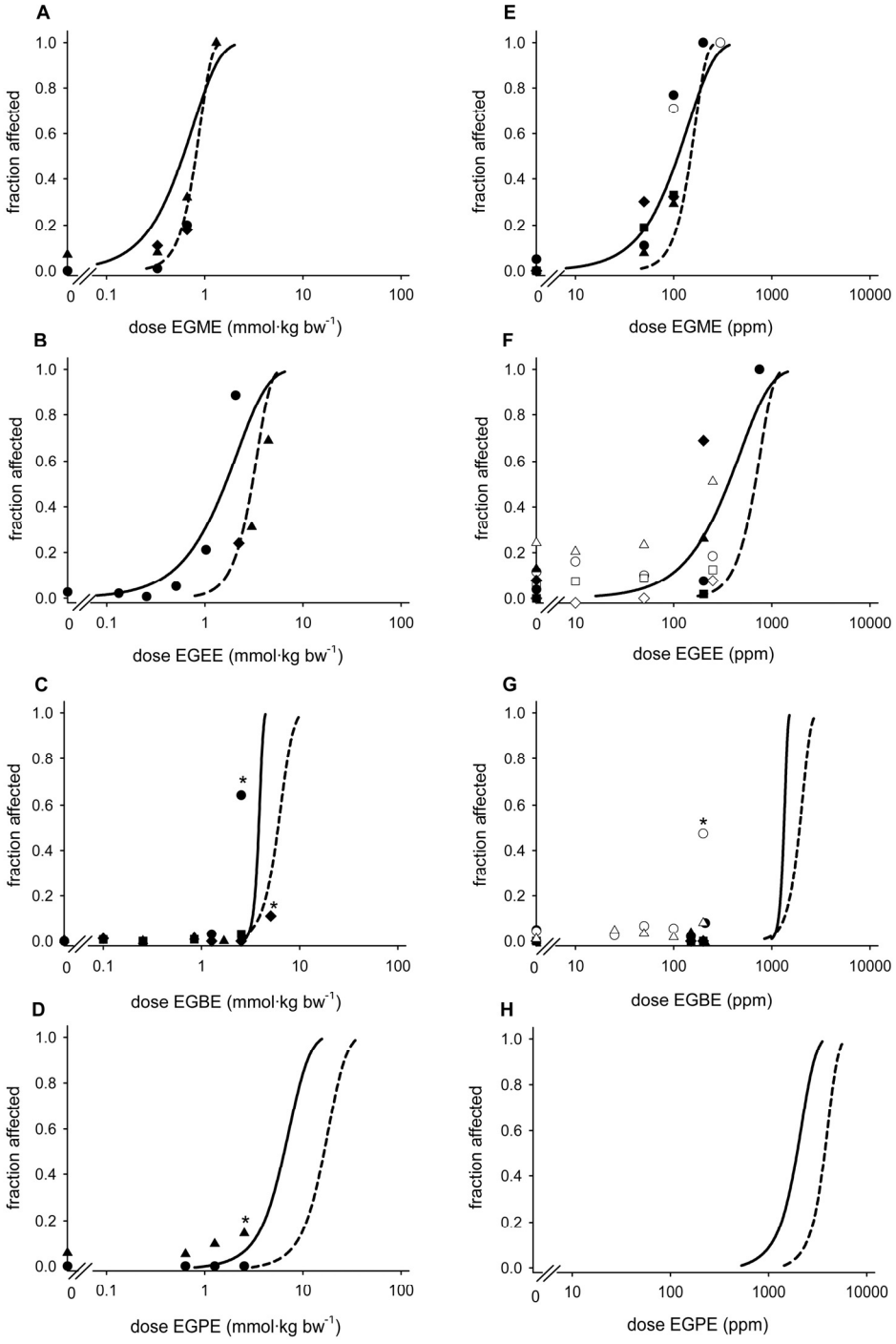
A
B
C
D
E
F
G
H
fraction affected
dose EGME (mmol·kg bw-1)
dose EGEE (mmol·kg bw-1)
dose EGBE (mmol·kg bw-1)
dose EGPE (mmol·kg bw-1)
dose EGME (ppm)
dose EGEE (ppm)
dose EGBE (ppm)
dose EGPE (ppm)

Figure 5. Predicted dose-response curves for developmental toxicity of glycol ethers EGME (A and E), EGEE (B and F), EGBE, (C and G) and EGPE (D and H) after repeated oral (A-C), subcutaneous (D), or inhalation (E-H) exposure in the rat. The individual data points represent measured developmental toxicity data taken from *in vivo* developmental toxicity studies as reported in the literature. The developmental toxicity endpoints represented by the symbols are shown in Table 5. For all endpoints, except for fetal bw decrease, the fraction of affected embryos or fetuses was calculated. For the endpoint fetal bw decrease, the average decrease in fetal bw of fetuses in the exposed groups was calculated as fraction of the average bw of fetuses of the control group. The curves represent the predicted dose-response curves for developmental toxicity and are acquired by translating *in vitro* effect concentrations ranging from BMC_1 up to BMC_{99} from laboratory 1 (straight line) or laboratory 2 (dashed line) into *in vivo* dose levels using the PBK model. "*" indicates a dose at which maternal toxicity including hemolysis was observed.

5, together with the predicted $BMDL_{10}$ values, which were acquired by translating the *in vitro* $BMCL_{10}$ values into *in vivo* dose levels using the rat PBK model. When comparing *in vivo*-derived and *in vivo*-predicted $BMDL_{10}$ values for each of the chemicals studied, differences in the range of 0.2- to 6-fold were found (Table 5). No $BMDL_{10}$ value for developmental toxicity could be derived for the EGBE inhalation exposure study in rats. However, when comparing the individual data points of the *in vivo* studies and the predicted dose-response curve (Figure 5G), a somewhat larger discrepancy between the predicted dose-response curve and the reported developmental toxicity dose levels was found. In rats exposed to 200 ppm EGBE, 47% of the fetuses died (Tyl *et al.*, 1984) (Figure 5G), whereas the predicted BMD_{47} value for developmental toxicity amounted to 1300 (predicted dose-response curve 1) and 1900 (predicted dose-response curve 2) ppm EGBE, resulting in approximately a 7- to 10-fold difference in predicted versus reported developmental toxicity dose levels.

Prediction of in vivo dose-response curves for developmental toxicity in human

The predicted dose-response curves for *in vivo* developmental toxicity of the glycol ethers EGME, EGEE and EGPE in human are shown in Figure 6 for a single-exposure (black lines) and for a repeated-exposure regimen (grey lines). For repeated exposure, a 5-day exposure regimen was chosen, representing daily exposure during a working week (8 h · day^{-1} for inhalation exposure). No dose-response curves could be predicted for EGBE because the maximum BAA plasma concentration reached using the human PBK model was lower than the BAA concentrations that affect ES-D3 cell differentiation. The *in vitro* concentration-response curves (BMC_1 to BMC_{99} values) for the alkoxyacetic acid metabolites in

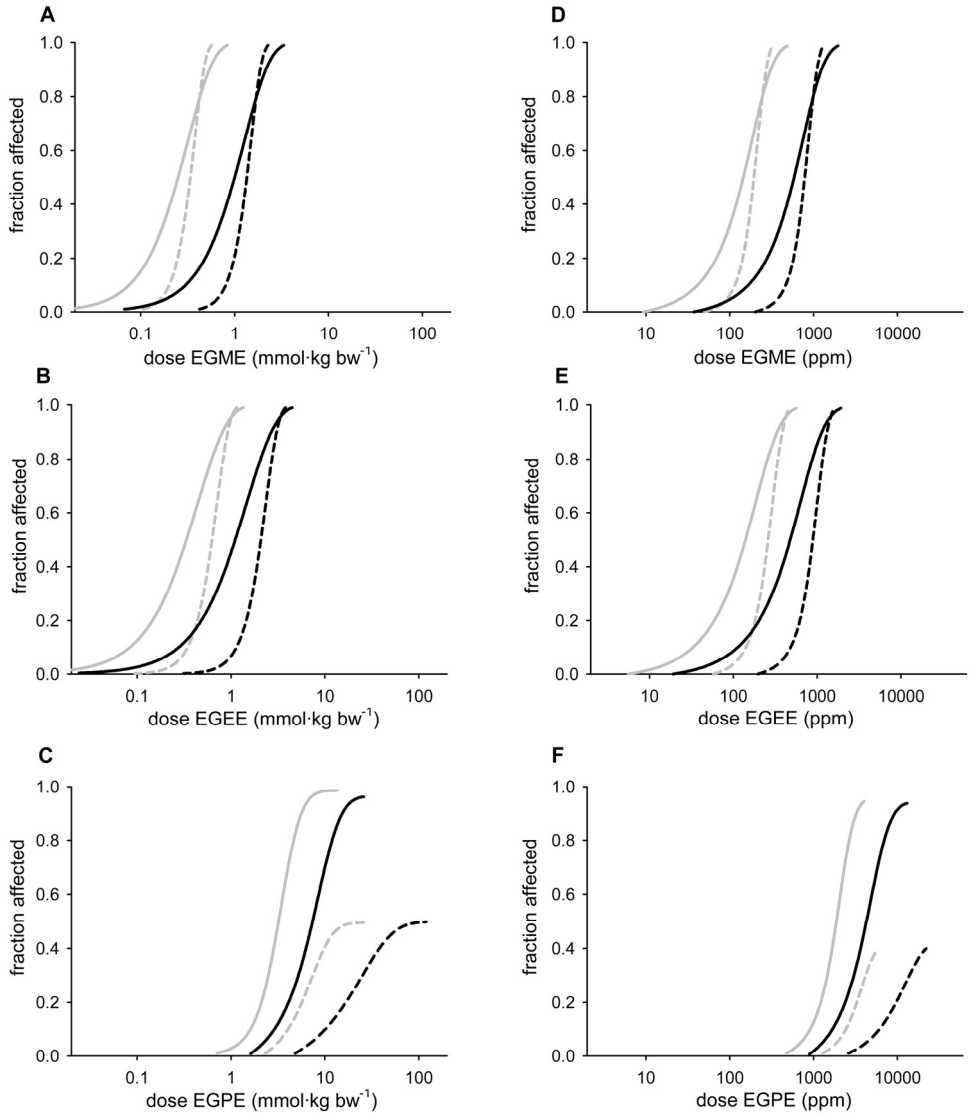

Figure 6. Predicted dose-response curves for developmental toxicity of glycol ethers EGME (A and D), EGEE (B and E) and EGPE, (C and F) after oral (A-C) or inhalation (D-F) exposure in human. The curves represent the predicted dose-response curves for developmental toxicity for a 1-day (black lines) or a 5-day (grey lines) exposure and are acquired by translating *in vitro* effect concentrations ranging from BMC_1 up to BMC_{99} from laboratory 1 (straight lines) or laboratory 2 (dashed lines) into *in vivo* dose levels using the PBK model.

the ES-D3 cell differentiation assay (De Jong *et al.*, 2009) translated into *in vivo* dose-response curves for human are shown in Figures 6A-C for oral exposure and in Figures 6D-F for inhalation exposure. For inhalation exposure, which is a more relevant exposure route for glycol ethers than oral exposure, $BMDL_{10}$ values for developmental toxicity were predicted for a 1- or 5-day exposure (Table 6). These $BMDL_{10}$ values were predicted by translating *in vitro* $BMCL_{10}$ values into *in vivo* dose levels using the human PBK model. The results show that upon repeated exposure, predicted $BMDL_{10}$ values decrease compared with a single exposure.

Table 6. Predicted $BMDL_{10}$ values for developmental toxicity in human exposed for 1 or 5 days to glycol ethers via inhalation ($8\ h \cdot day^{-1}$) using the *in vitro-in silico* approach (*in vitro* toxicity data were taken from two different laboratories (De Jong *et al.*, 2009) leading to different predicted $BMDL_{10}$ values per chemical).

Chemical	Lab 1		Lab 2	
	1-day exposure (ppm)	5-day exposure (ppm)	1-day exposure (ppm)	5-day exposure (ppm)
EGME	110	28	330	83
EGEE	81	24	370	110
EGBE	n.d.	n.d.	n.d.	n.d.
EGPE	1300	720	4800	2000

n.d.: $BMDL_{10}$ could not be determined because $BMCL_{10}$ concentration of BAA could not be reached using the human PBK model.

Discussion

The aim of the present study was to investigate the feasibility of using a combination of *in vitro* toxicity data and PBK modeling to predict *in vivo* dose-response curves for developmental toxicity, thereby providing a platform to use *in vitro* toxicity data not only qualitatively (presence or absence of effect) but also quantitatively (prediction of dose-response curves). This study shows that our predicted $BMDL_{10}$ values for *in vivo* developmental toxicity of glycol ethers in rat differed a factor of 0.2-6 from the measured $BMDL_{10}$ values determined from reported *in vivo* developmental toxicity studies. Predicted $BMDL_{10}$ values might in the future replace measured $BMDL_{10}$ values determined in *in vivo* developmental toxicity studies and could be used as points of departure in risk assessment to derive safe exposure levels for humans. This may be achieved by applying

uncertainty factors for interspecies variation, as also currently used in risk assessment based on animal toxicity studies. In addition, one might even use the human PBK models to obtain dose-response curves and $BMDL_{10}$ values for human, thereby eliminating the need for an uncertainty factor for interspecies differences in kinetics.

From the *in vivo* developmental toxicity studies, $BMDL_{10}$ values were derived for diverse developmental toxicity endpoints (i.e., resorptions, fetal bw decrease, malformations and fetal deaths), whereas the predicted $BMDL_{10}$ values were derived for one *in vitro* developmental toxicity readout parameter (i.e., inhibition of embryonic stem cell differentiation into contracting cardiomyocytes). Although the ES-D3 cell differentiation assay determines the concentrations of a chemical that inhibit cardiomyocyte differentiation, the present study was not intended to specifically predict dose levels affecting cardiac development during embryonic development. The inhibition of cardiac differentiation was rather used to represent a sensitive readout parameter for *in vitro* developmental toxicity (Genschow *et al.*, 2004), which in the ideal situation could represent the most sensitive *in vivo* developmental toxicity endpoint. However, one might argue that other developmental toxicity endpoints such as developmental neurotoxicity might be less well predicted with the use of the *in vitro* effect concentrations that inhibit cardiomyocyte differentiation. Therefore, for predicting $BMDL_{10}$ values for other developmental toxicity endpoints such as developmental neurotoxicity, an *in vitro* assay focusing on this specific endpoint (Breier *et al.*, 2010) might improve the approach to an even further extent. The type of observed glycol ether-induced malformations will probably depend on the developmental stage at which the embryo is exposed to the glycol ethers (and their alkoxyacetic acid metabolites). Horton *et al.* (1985) showed, for example, that exposure of mice to EGME on gestational day 7 (GD7) resulted in exencephaly, whereas paw anomalies dominated when mice were exposed later in pregnancy (GD11). Because the present study did not intend to predict the developmental toxicity for a specific developmental toxicity endpoint, no critical window of glycol ether exposure was defined for the predictions.

The predicted dose-response curves for developmental toxicity were acquired by using the peak plasma concentrations of the alkoxyacetic acid metabolites as input

in the PBK models. Sweeney *et al.* (2001) evaluated for different EGME developmental toxicity studies in rat and mice the relation between the MAA peak plasma concentration and the percentage of malformed fetuses, and the relation between the average daily area under the blood concentration-time curve (AUC) of MAA and the percentage of malformed fetuses. No preferred dose metric (MAA peak plasma concentration or average daily AUC of MAA) was obtained. For the translation of *in vitro* data into *in vivo* data in the present study, alkoxyacetic acid peak concentrations were selected as the preferred dose metric to be used as input values for the PBK models.

In the *in vivo* developmental toxicity studies with EGBE and EGPE, developmental toxicity was only observed at dose levels that also caused maternal toxicity, including hemolysis (Sleet *et al.*, 1989; Tyl *et al.*, 1984; Unilever Research, 1984). *In vitro* studies showed that hemolytic effects of BAA are already found at 1 mM (lowest concentration tested; Ghanayem *et al.*, 1989), indicating that hemolysis occurs at lower BAA concentrations than the concentrations inhibiting embryonic stem cell differentiation (De Jong *et al.*, 2009). Because maternal toxicity was not taken into account in the present study, in which we aimed to predict developmental toxicity, this may explain why our predicted developmental toxicity dose levels are higher than the developmental toxicity dose levels measured in *in vivo* studies, due to developmental toxicity as a secondary effect of maternal toxicity.

De Jong *et al.* (2009) showed that the ranking of the potencies of the glycol ether alkoxyacetic acid metabolites in the ES-D3 cell differentiation assay of the EST is the same as that of the potencies of their parent glycol ethers *in vivo*. However, differences in *in vitro* potencies were small (at maximum threefold) compared with differences in *in vivo* potencies for developmental toxicity (at maximum 24-fold) (De Jong *et al.*, 2009). The present study shows that the predicted differences in potencies of the glycol ethers improve, when combining the *in vitro* toxicity data with PBK models simulating *in vivo* kinetics (Table 7). The physiological parameter values used in the rat and human PBK models are representative for the beginning of pregnancy (Gargas *et al.*, 2000a). During pregnancy, some physiological parameter values will change, for example, the size of some tissue compartments. It is not expected that this would have a large influence on the model outcome of the

Table 7. Measured and predicted $BMDL_{10}$ values for glycol ether-induced malformations in rat and relative potency of these glycol ethers in inducing malformations as measured *in vivo* and predicted *in vivo* (*in vitro* toxicity data were taken from two different laboratories (De Jong *et al.*, 2009) leading to different predicted $BMDL_{10}$ values per chemical).

Chemical	Measured $BMDL_{10}$ for malformations (mmol·kg bw^{-1}·day^{-1})	Predicted $BMDL_{10}$ (mmol·kg bw^{-1}·day^{-1})	
		Lab 1	Lab 2
EGME (MAA)	0.46	0.12	0.34
EGEE (EAA)	0.8	0.28	1.3
EGBE (BAA)	3.6	2.6	3.1
EGPE (PAA)	> 2.5	2.4	6.2
Relative potency			
EGME (MAA)	1.0	1.0	1.0
EGEE (EAA)	1.7	2.3	3.8
EGBE (BAA)	7.8	22	9.1
EGPE (PAA)	> 5.4	20	18

present models, because the sensitivity analysis of the model revealed that the influence of these parameters, including richly perfused tissue (fetus) (Gargas *et al.*, 2000a), is low (Figure 4). Although the approach used assumes that *in vitro* developmentally toxic concentrations have to reach the embryo to result in developmental toxicity, it was not necessary to include a specific fetal compartment in the model, because fetal MAA and EAA concentrations were found to be identical to maternal plasma concentrations in rats exposed to EGME (Gargas *et al.*, 2000a) and EGEEA (Gargas *et al.*, 2000b), respectively. It was assumed that this is also true for BAA and PAA. We aimed to build our PBK model based on parameter values derived from *in silico* (partition coefficients) or *in vitro* (metabolism) methods. The parameter values for EGPE metabolism were assumed to be equal to those of EGBE based on comparable LogP values and VanderWaals volumes (Table 3), which are important descriptors in Quantitative Structure-Activity Relationships (QSARs) for metabolism (Chang *et al.*, 2009; Soffers *et al.*, 2001). Given the outcomes of the sensitivity analysis, which revealed that metabolic parameters were of limited influence, and considering the model outcomes of the EGPE model (Figure 2F), it seems to be justified to use the kinetic parameter values for EGBE metabolism in the EGPE model. The parameter values for the oral uptake (k_a) of the glycol ethers and the parameter values for urinary excretion of the

alkoxyacetic acid metabolites (K_{ex}) were acquired using *in vivo* kinetic data. The sensitivity analysis showed that the k_a and the K_{ex} were not sensitive parameters for the prediction of peak plasma levels of the alkoxyacetic acids (SC k_a = 0, SC K_{ex} = -0.1). Although the K_{ex} is not a sensitive parameter, differences in peak plasma concentrations of MAA, EAA, BAA and PAA reached are to a large extent the result of the large differences in K_{ex} values (at maximum 18-fold for rat and 11-fold for human), which are directly related to the plasma half-lives of the alkoxyacetic acid metabolites. The present study shows that PAA has a shorter plasma half-life in rats (1.0-2.6 h) than BAA (1.5-3.2 h; Ghanayem *et al.*, 1990), EAA (7.6-10 h; Aasmoe and Aarbakke, 1997; Aasmoe *et al.*, 1999) and MAA (14-19 h; Aasmoe and Aarbakke, 1997; Aasmoe *et al.*, 1999). This indicates that the developmental toxicity potencies of the glycol ethers EGME, EGEE, EGBE and EGPE in rats are related to the plasma half-lives of their alkoxyacetic acid metabolites, as we suggested before (De Jong *et al.*, 2009). The plasma half-life of PAA, as measured in rats in the present study, was comparable with that of 2,4-D in rats (1.1-2.1 h; Timchalk, 2004). Because no human data on the excretion kinetics of PAA are available in the literature, the excretion data of 2,4-D in human were therefore assumed to be suitable to derive the parameter value for PAA excretion for the human PBK model. However, because the K_{ex} seems to play an important role in determining differences in *in vivo* developmental toxicity potencies of the glycol ethers, an inaccurate estimation of this parameter value would result in incorrectly predicted developmental toxicity dose levels.

No dose-response curves for developmental toxicity of EGBE could be predicted for human. Due to the high excretion rate of BAA in the human model, a steady-state level of BAA is reached (Figure 7), which is lower than the BAA concentrations that decrease the ES-D3 cell differentiation *in vitro* by 10%. The predicted dose-response curve for developmental toxicity of EGPE for human, using the *in vitro* ES-D3 cell differentiation assay data of laboratory 2, does not reach a 100% of affected fetuses (Figures 6C and 6F). This is due to the steady-state level that is reached for PAA (Figure 7), which equals the BMC_{50} value of PAA derived from the ES-D3 cell differentiation assay data of laboratory 2.

The extrapolation of *in vitro* toxicity data with human PBK models may predict safe dose levels for human, without the need for applying an uncertainty factor for

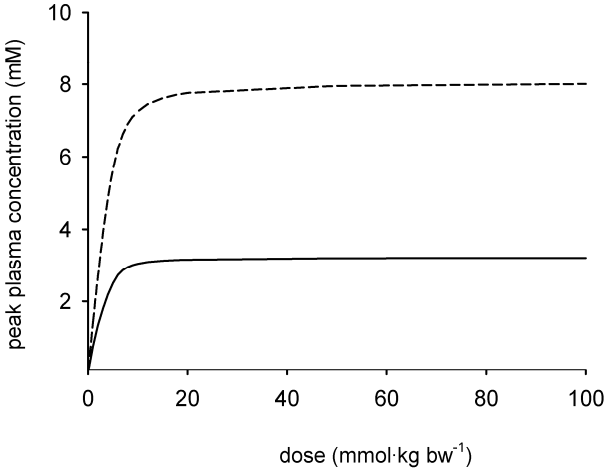

Figure 7. Predicted peak plasma concentrations of BAA (solid line) and PAA (dashed line) in man upon five subsequent daily exposures to EGBE and EGPE respectively, at doses ranging from 0 to 100 mmol · kg bw^{-1}.

interspecies differences in kinetics. With the use of probabilistic PBK models, in which distributions rather than single points are used for parameter values, one might even take the variability of *in vivo* kinetics in the human population into account (chapter 7). This would enable a better prediction of dose-response curves for *in vivo* developmental toxicity for humans and eliminate the need not only for an uncertainty factor for interspecies differences in kinetics, but also for an uncertainty factor for intraspecies differences in kinetics. However, predicted developmental toxicity dose levels for humans cannot be verified with *in vivo* human data. Consequently, possible interspecies differences in dynamics might be overlooked. Therefore, one might want to keep the uncertainty factor for these differences in dynamics.

Altogether, the results of the present study show that our predicted dose-response curves for glycol ether-induced developmental toxicity in rats, based on ES-D3 cell differentiation assay data in combination with PBK modeling, are in good concordance with glycol ether developmental toxicity dose levels reported in literature. The differences between our predicted $BMDL_{10}$ values and the $BMDL_{10}$ values derived from *in vivo* studies for glycol ethers are small considering the experimental differences in $BMDL_{10}$ values that might be derived from *in vivo*

(developmental) toxicity studies. The predicted $BMDL_{10}$ values could be used as points of departure in risk assessment practices, thereby contributing to the reduction of animal use in the risk assessment of these chemicals. It is concluded that the combined *in vitro-in silico* approach, after further development and evaluation, may contribute to a science-based risk assessment of chemicals, using no or reduced numbers of laboratory animals.

Acknowledgements

The authors would like to thank Ans Soffers for helping with the calculations of LogP values and VanderWaals volumes, and Martijn Vermeulen for helping with the GC-MS and LC-MS analyses.

References

Aasmoe, L., Aarbakke, J. (1997). Gender difference in the elimination of 2-methoxyethanol, methoxyacetic acid and ethoxyacetic acid in rat. *Xenobiotica* **27**, 1237-44.

Aasmoe, L., Mathiesen, M., Sager, G. (1999). Elimination of methoxyacetic acid and ethoxyacetic acid in rat. *Xenobiotica* **29**, 417-24.

Andrew, F. D., Hardin, B. D. (1984). Developmental effects after inhalation exposure of gravid rabbits and rats to ethylene glycol monoethyl ether. *Environ Health Perspect* **57**, 13-23.

Barlow, S., Chesson, A., Collins, J. D., Flynn, A., Hardy, A., Jany, K., Knaap, A., Kuiper, H., Larsen, J., Lovell, D., Le Neindre, P., Schans, J., Schlatter, J., Silano, V., Skerfving, S., Vannier, P. (2009). Guidance of the Scientific Committee on a request from EFSA on the use of the benchmark dose approach in risk assessment. *The EFSA Journal* **1150**, 1-72.

Barter, Z. E., Bayliss, M. K., Beaune, P. H., Boobis, A. R., Carlile, D. J., Edwards, R. J., Houston, J. B., Lake, B. G., Lipscomb, J. C., Pelkonen, O. R., Tucke, G. T., Rostami-Hodjegan, A. (2007). Scaling factors for the extrapolation of in vivo metabolic drug clearance from in vitro data: reaching a consensus on values of human microsomal protein and hepatocellularity per gram of liver. *Curr Drug Metab* **8**, 33-45.

Berezhkovskiy, L. M. (2004). Volume of distribution at steady state for a linear pharmacokinetic system with peripheral elimination. *J Pharm Sci* **93**, 1628-40.

Breier, J. M., Gassmann, K., Kayser, R., Stegeman, H., de Groot, D., Fritsche, E., Shafer, T. J. (2010). Neural progenitor cells as models for high-throughput screens of developmental neurotoxicity: State of the science. *Neurotoxicol Teratol* **32**, 4-15.

Brown, N. A., Holt, D., Webb, M. (1984). The teratogenicity of methoxyacetic acid in the rat. *Toxicol Lett* **22**, 93-100.

Chang, C., Duignan, D. B., Johnson, K. D., Lee, P. H., Cowan, G. S., Giffors, E. M., Stankovic, C. J., Lepsy, C. S., Stoner, C. L. (2009). The development and validation of a computational model to predict rat liver microsomal clearance. *J Pharm Sci* **98**, 2857-67.

Chester, A., Hull, J., Andrew, F. (1986). Lack of teratogenic effect after ethylene glycol monoethyl ether (EGEE) in rats via drinking water. *Teratology* **33**, C57 (Abstract).

De Jong, E., Louisse, J., Verwei, M., Blaauboer, B. J., van de Sandt, J. J. M., Woutersen, R. A., Rietjens, I. M. C. M., Piersma, A. H. (2009). Relative developmental toxicity of glycol ether alkoxy acid metabolites in the embryonic stem cell test as compared to the in vivo potency of their parent compounds. *Toxicol Sci* **110**, 117-24.

Doe, J. E., Samuels, D. M., Tinston, D. J., de Silva Wickramaratne, G. A. (1983). Comparative aspects of the reproductive toxicology by inhalation in rats of ethylene glycol monomethyl ether and propylene glycol monomethyl ether. *Toxicol Appl Pharmacol* **69**, 43-7.

Doe, J. E. (1984). Ethylene glycol monoethyl ether and ethylene glycol monoethyl ether acetate teratology studies. *Environ Health Perspect* **57**, 33-41.

ECETOC (2005). The Toxicology of Glycol Ethers and its Relevance to Man (Fourth Edition) ECETOC *Technical Report* **095**.

ECHA (2009). New study on the number of test animals for REACH. *Press release ECHA/PR/09/11, 28 August 2009.*

Evans, M. V., Andersen, M. E. (2000). Sensitivity analysis of a physiological model for 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD): assessing the impact of specific model parameters on sequestration in liver and fat in the rat. *Toxicol Sci* **54**, 71-80.

Gargas, M. L., Tyler, T. R., Sweeney, L. M., Corley, R. A., Weitz, K. K., Mast, T. J., Paustenbach, D. J., Hays, S. M. (2000a). A toxicokinetic study of inhaled ethylene glycol monomethyl ether (2-ME) and validation of a physiologically based pharmacokinetic model for the pregnant rat and human. *Toxicol Appl Pharmacol* **165**, 53-62.

Gargas, M. L., Tyler, T. R., Sweeney, L. M., Corley, R. A., Weitz, K. K., Mast, T. J., Paustenbach, D. J., Hays, S. M. (2000b). A toxicokinetic study of inhaled ethylene glycol ethyl ether acetate and validation of a physiologically based pharmacokinetic model for rat and human. *Toxicol Appl Pharmacol* **165**, 63-73.

Genschow E., Spielmann H., Scholz G., Seiler A., Brown N. A., Piersma A. H., Brady M., Clemann N., Huuskonen H., Paillard F., Bremer, S., Becker, K. (2002). The ECVAM international validation study on in vitro embryotoxicity tests: results of the definitive phase and evaluation of prediction models. European Centre for the Validation of Alternative Methods. *Atern Lab Anim* **30**, 151-76.

Genschow, E., Spielmann, H., Scholz, G., Pohl, I., Seiler, A., Clemann, N., Bremer, S., Becker, K. (2004). Validation of the embryonic stem cell test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* **32**, 209-44.

Ghanayem, B. I., Burka, L. T., Matthews, H. B. (1989). Structure-activity relationships for the in vitro hematotoxicity of N-alkoxyacetic acids, the toxic metabolites of glycol ethers. *Chem Biol Interact* **70**, 339-52.

Ghanayem, B. I., Sanders, J. M., Clark, A. M., Bailer, J., Matthews, H. B. (1990). Effects of dose, age, inhibition of metabolism and elimination on the toxicokinetics of 2-butoxyethanol and its metabolites. *J Pharmacol Exp Ther* **253**, 136-43.

Giavini, E., Broccia, M. L., Menegola, E., Prati, M., 1993. Comparative in vitro study of the embryotoxic effects of three glycol ethers and their metabolites, the alkoxyacids. *Toxicol In Vitro* **7**, 777-84.

Goad, P. T., Cranmer, J. M. (1984). Gestation period sensitivity of ethylene glycol monoethyl ether in rats. *Toxicologist* **4**, 87.

Green, C. E., Gordon, G. R., Cohen, P. M., Nolen, H. W., Peters, J. H., Tyson, C. A. (1996). In vitro

metabolism of glycol ethers by human and rat hepatocytes. *Occup Hyg* **2**, 67-75.

Groeseneken, D., Veulemans, H., Masschelein, R. (1986). Urinary excretion of ethoxyacetic acid after experimental human exposure to ethylene glycol monoethyl ether. *Br J Ind Med* **43**, 615-19.

Groeseneken, D., Veulemans, H., Masschelein, R., van Vlem, E. (1989). Experimental human exposure to ethylene glycol monomethyl ether. *Int Arch Occup Environ Health* **61**, 243-47.

Hartung, T., Rovida, C. (2009). Chemical regulators have overreached. Nature **460**, 1080-1.

Hays, S. M., Elswick, B. A., Blumenthal, G. M., Welsch, F., Conolly, R. B., Gargas, M. L. (2000). Development of a physiologically based pharmacokinetic model of 2-methoxyethanol and 2-methoxyacetic acid disposition in pregnant rats. *Toxicol Appl Pharmacol* **163**, 67-74.

Höfer, T., Gerner, I., Gundert-Remy, U., Liebsch, M., Schulte, A., Spielmann, H., Vogel, R., Wettig, K. (2004). Animal testing and alternative approaches for the human health risk assessment under the proposed new European chemicals regulation. *Arch Toxicol* **78**, 549-64.

Horton, V. L., Sleet, R. B., John-Greene, J. A., Welsch, F. (1985). Developmental phase-specific and dose-related teratogenic effects of ethylene glycol monomethyl ether in CD-1 mice. *Tox Appl Pharmacol* **80**, 108-18.

Johanson, G., Dynésius, B. (1988). Liquid/air partition coefficients of six commonly used glycol ethers. *Br J Ind Med* **45**, 561-64.

Jones, K., Cocker, J. (2003). A human exposure study to investigate biological monitoring methods for 2-butoxyethanol. *Biomarkers* **8**, 360-70.

Kezic, S., Meuling, W. J. A., Jakasa, I. (2004). Free and total urinary 2-butoxyacetic acid following dermal and inhalation exposure to 2-butoxyethanol in human volunteers. *Int Arch Occup Environ Health* **77**, 580-86.

Nelson, B. K., Setzer, J. V., Brightwell, W. S., Mathinos, P. R., Kuczuk, M. H., Weaver, T. E., Goad, P. T. (1984). Comparative inhalation teratogenicity of four glycol ether solvents and an amino derivative in rats. *Environ Health Perspect* **57**, 261-71.

Sauerhoff, M. W., Braun, W. H., Blau, G. E., Gehring, P. J. (1977). The fate of 2,4-dichlorophenoxyacetic acid (2,4-D) following oral administration to man. *Toxicology* **8**, 3-11.

Seglen, P. O. (1976). Preparation of isolated rat liver cells. *Methods Cell Biol* **13**, 29-83.

Sleet, R. B., Price, C. J., Marr, M. C., Morrissey, R. E., Schwetz, B. A. (1989). Teratologic evaluation of ethylene glycol monobutyl ether (CAS no. 111-76-2) administered to Fisher-344 rats on either gestational days 9 through 11 or days 11 through 13. *NTP Report* **89-058**.

Soffers, A. E. M. F., Boersma, M. G., Vaes, W. H. J., Vervoort, J., Tyrakowska, B., Hermens, J. L. M., Rietjens, I. M. C. M. (2001). Computer-modeling-based QSARs for analyzing experimental data on biotransformation and toxicity. *Toxicol In Vitro* **15**, 539-51.

Stenger, E. G., Aeppli, L., Müller, D., Peheim, E., Thomann, P. (1971). Toxicology of ethyleneglycol-monoethyl ether. *Arzneimittelforschung* **21**, 880-85.

Sweeney, L. M., Tyler, T. R., Kirman, C. R., Corley, R. A., Reitz, R. H., Paustenbach, D. J., Holson, J. F., Whorton, M. D., Thompson, K. M., Gargas, M. L. (2001). Proposed occupational exposure limits for select ethylene glycol ethers using PBPK models and Monte Carlo simulations. *Toxicol Sci* **62**, 124-39.

Timchalk, C. (2004). Comparative inter-species pharmacokinetics of phenoxyacetic acid herbicides and related organic acids; Evidence that the dog is not a relevant species for evaluation of human health risk. *Toxicology* **200**, 1-19.

Toraason, M., Stringer, B., Stober, P., Hardin, B. D. (1985). Electrocardiographic study of rat fetuses exposed to ethylene glycol monomethyl ether (EGME). *Teratology* **32**, 33-9.

Tyl, R. W., Millicovsky, G., Dodd, D. E., Pritts, I. M., France, K. A., Fisher, L. C. (1984). Teratologic evaluation of ethylene glycol monobutyl ether in Fischer 344 rats and New Zealand white rabbits following inhalation exposure. *Environ Health Perspect* **57**, 47-68.

Unilever Research (1984). Teratogenicity of phenoxyethanol by subcutaneous injection in Colworth Wistar rats. *Research report* **PES 841023**.

Van der Jagt, K., Munn, S. J., Tørsløv, J., de Bruijn, J. (2004). Alternative approaches can reduce the use of test animals under REACH. *Report EUR* **21405**.

Verwei, M., van Burgsteden, J. A., Krul, C. A. M., van de Sandt, J. J. M., Freidig, A. P. (2006). Prediction of in vivo embryotoxic effect levels with a combination of in vitro studies and PBPK modelling. *Toxicol Lett* **165**, 79-87.

Chapter 6 is reprinted with permission from Elsevier.

CHAPTER 6

Relative developmental toxicity potencies of retinoids in the embryonic stem cell test compared with their relative potencies in *in vivo* and two other *in vitro* assays for developmental toxicity

Jochem Louisse, Süleyman Gönen, Ivonne M.C.M. Rietjens, Miriam Verwei

Based on:
Toxicology Letters 203, 1-8 (2011).

Abstract

The present study determines the relative developmental toxicity potencies of retinoids in the embryonic stem cell (ES-D3 cell) differentiation assay of the embryonic stem cell test, and compares the outcomes with their relative potencies in *in vivo* and two other *in vitro* assays for developmental toxicity. The results reveal that the potency ranking obtained in the ES-D3 cell differentiation assay is similar to the reported potency rankings in the two other *in vitro* assays for developmental toxicity. TTNPB ((E)-4[2-(5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2-naphthalenyl)-1-propenyl]benzoic acid) was the most potent retinoid, whereas etretinate and retinol had the lowest potency. All-*trans*-retinoic acid, 13-*cis*-retinoic acid, 9-*cis*-retinoic acid and acitretin showed an intermediate potency. *In vivo* potency rankings of the developmental toxicity of retinoids appear to be dependent on the species and/or exposure regimens used. The obtained *in vitro* potency ranking does not completely correspond with the *in vivo* potency rankings, although TTNPB is correctly predicted to be the most potent and retinol the least potent congener. The lack of *in vivo* kinetic processes in the ES-D3 cell differentiation assay might explain the deviating potency predictions of some retinoids. Therefore, knowledge on the chemical-specific species-dependent *in vivo* kinetics is essential when using *in vitro* toxicity data for the estimation of *in vivo* developmental toxicity potencies within series of related chemicals.

Introduction

For the registration of industrial chemicals under REACH (Registration, Evaluation, Authorization and restriction of CHemicals), information on safety is required. It is estimated that more than 25% of the 3.9 million laboratory animals needed for safety testing for REACH will be used for developmental toxicity studies (Van der Jagt *et al.*, 2004). Also for pharmaceuticals, food additives and pesticides, regulatory agencies require safety studies performed in laboratory animals, including developmental toxicity studies. Because of the high number of laboratory animals used for developmental toxicity testing, the development, validation and application of *in vitro* and *in silico* assays that can contribute to the replacement, reduction and refinement (3Rs) of experimental animal studies, is urgently needed. Currently, *in vitro* assays are considered to be useful in the development and early safety testing of pharmaceuticals and food additives and for setting priorities for toxicity testing of chemicals within the REACH framework.

For developmental toxicity, three *in vitro* assays have been scientifically validated so far, which are the rat limb bud micromass (MM) test, the postimplantation rat whole-embryo culture (WEC) test and the embryonic stem cell test (EST) (Genschow *et al.*, 2002). Of these three assays, the EST, using embryonic stem cells, is the only test in which no primary animal tissues are required. The EST measures the potency of the test chemical to inhibit cell proliferation of embryonic fibroblasts (BALB/c-3T3 cells) and embryonic stem cells (ES-D3 cells) (cell proliferation assays), and the potency of the test chemical to inhibit ES-D3 cell differentiation into contracting cardiomyocytes (ES-D3 cell differentiation assay). The acquired EC_{50} values in these assays are applied in a prediction model to predict the embryotoxic potency (non-, weak or strong embryotoxic) of the chemical. The EST, including its prediction model, was validated under the supervision of the European Centre for the Validation of Alternative Methods (ECVAM) and was able to predict the developmental toxicity potencies of twenty different chemicals with an overall accuracy of 78% (Genschow *et al.*, 2004). The usefulness of the prediction model is questioned however, since only two of thirteen chemicals that were tested in the European Framework Program project ReProTect were correctly classified using the prediction model (Marx-Stoelting *et al.*, 2009). It is also not known whether the prediction model is able to classify the

developmental toxicity potencies of chemicals within a chemical class. The ES-D3 cell differentiation assay, used as standalone assay, has been shown to be able to predict the *in vivo* potency ranking of four glycol ethers based on the *in vitro* toxic effect concentrations of their toxic alkoxyacetic acid metabolites (De Jong *et al.*, 2009). This makes the differentiation assay of the EST a useful tool to set priorities for developmental toxicity testing within a series of glycol ethers (De Jong *et al.*, 2009). However, it remains to be established whether the differentiation assay of the EST is also able to correctly predict the relative developmental toxicity potencies of chemicals belonging to other chemical classes.

The aims of the present study were 1) to compare the potency ranking of retinoids obtained in the *in vitro* ES-D3 cell differentiation assay of the EST with reported potency rankings of these retinoids in two other *in vitro* assays for developmental toxicity, the MM test and the WEC test and 2) to determine the applicability of the ES-D3 cell differentiation assay to predict the relative *in vivo* developmental toxicity potencies of retinoids, in order to investigate the potential of the ES-D3 cell differentiation assay as a tool to set priorities for developmental toxicity testing within a series of retinoids.

To this end, the relative potencies of retinol, all-*trans*-retinoic acid (ATRA), 13-*cis*-retinoic acid (13-*cis*-RA), 9-*cis*-retinoic acid (9-*cis*-RA), etretinate, acitretin and TTNPB ((E)-4[2-(5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2-naphthalenyl)-1-propenyl]benzoic acid) (Figure 1) were determined in the ES-D3 cell differentiation assay and compared with their relative potencies in the two other *in vitro* test systems as reported in the literature. Furthermore, the relative potencies of these retinoids in the ES-D3 cell differentiation assay were also compared with their relative potencies derived from reported *in vivo* studies. The *in vitro* and *in vivo* potencies of the retinoids were determined using the benchmark dose (BMD) approach, using the BMC_{50} values (benchmark concentration at which a benchmark response of 50% (BMR_{50}) is reached) for the *in vitro* developmental toxicity potency ranking, and the BMD_{10} values (benchmark dose at which the BMR_{10} is reached) for the *in vivo* developmental toxicity potency ranking. Finally, the retinoids were also tested in the ES-D3 cell differentiation assay in the presence of the retinoic acid receptor (RAR) antagonist Ro 41-5253 (Figure 1), to investigate whether the retinoid-induced inhibition of cardiomyocyte differentiation was RAR-mediated.

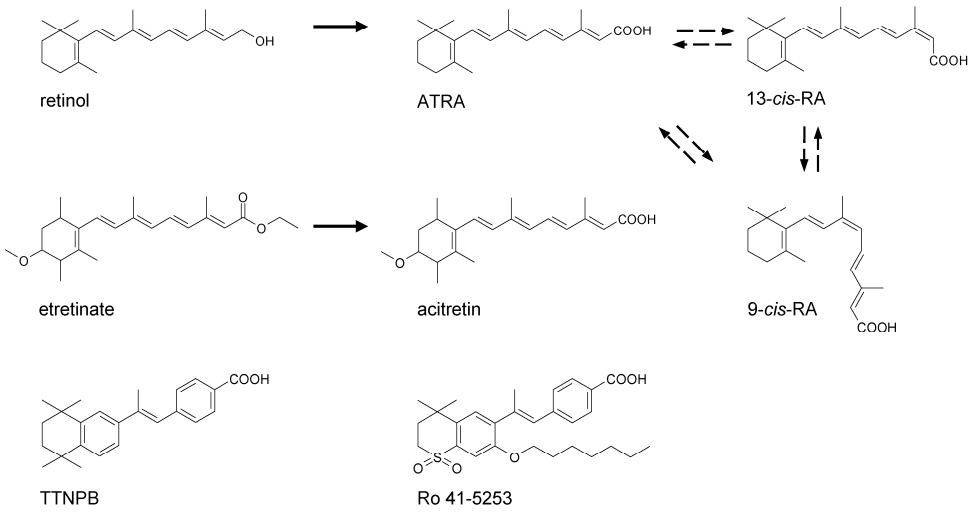

Figure 1. Chemical structure of retinoids used in the present study. Straight arrow = metabolism; dashed arrow = isomerization. ATRA: all-*trans*-retinoic acid; 13-*cis*-RA: 13-*cis*-retinoic acid; 9-*cis*-RA: 9-*cis*-retinoic acid; TTNPB: (E)-4[2-(5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2-naphthalenyl)-1-propenyl] benzoic acid.

Materials and methods

Chemicals

9-*Cis*-RA, 13-*cis*-RA, acitretin, ATRA, retinol and TTNPB were purchased from Sigma-Aldrich (Zwijndrecht, The Netherlands) and etretinate from Sequoia Research Products (Berkshire, UK). Ro 41-5253 was purchased from Enzo Life Sciences (Raamsdonksveer, The Netherlands) and dimethyl sulfoxide (DMSO) from Acros Organics (Geel, Belgium).

ES-D3 cell culture

The murine ES-D3 cell line was a kind gift from Johnson & Johnson (Beerse, Belgium), and was cultured according to the method described by De Smedt *et al.* (2008). Briefly, cells were cultured in cell culture Petri dishes (35 mm × 10mm, Corning, Schiphol-Rijk, The Netherlands) at 37 °C in a humidified atmosphere of 5% CO_2 in Dulbecco's Modified Eagle Medium (Invitrogen, Breda, The Netherlands) supplemented with 20% (v/v) fetal calf serum (BioWhittaker, Verviers, Belgium), 2 mM glutamine (Invitrogen), 50 U/ml penicillin/50 µg/ml

streptomycin (Invitrogen), 1% (v/v) non-essential amino acids (Invitrogen) and 0.1 mM β-mercaptoethanol (Sigma-Aldrich). To keep cells undifferentiated, 1000 U/ml murine leukemia inhibiting factor (ESGRO®, Chemicon International, Amsterdam, The Netherlands) was added to the culture medium. Cells were subcultured three times a week, using non-enzymatic dissociation buffer (Sigma-Aldrich) to detach cells.

ES-D3 cell differentiation assay

To determine the concentrations affecting ES-D3 cell differentiation into contracting cardiomyocytes, the ES-D3 cell differentiation assay as described by De Smedt *et al.* (2008) was used, with small modifications as described by Louisse *et al.* (2010a). Briefly, on day 0, droplets of 20 µl cell suspension ($3.75 \cdot 10^4$ cells/ml) were seeded on the inside of the lid of a 96-well plate. Caps of Eppendorf tubes were placed at the corners of the well plate to elevate the lid in order to prevent direct contact of the hanging drops with the well plate. Wells were filled with 250 µl phosphate buffered saline (Invitrogen) to obtain the right humidity and to prevent evaporation of the hanging drops. On day 3, cell aggregates, called embryoid bodies (EBs), were formed and these were transferred to bacteriological petri dishes (Greiner Bio-one, Alphen a/d Rijn, The Netherlands). On day 5, EBs were transferred to 24-well plates (Corning, 1 EB/well) and were incubated for another 5 days. On day 10, the number of wells containing contracting cardiomyocytes was determined using light microscopy. Exposure to test chemicals, added from a 400 times concentrated stock solution in DMSO to culture medium, started at day 0 and lasted for 10 days. Exposure medium was refreshed at day 3 and day 5. Solvent controls (0.25% DMSO, v/v) were included in each experiment. A test was considered valid when at least 21 out of 24 wells of the blank (non-exposed cells) and the solvent control plate contained contracting cardiomyocytes. For each plate, the number of EBs differentiated into contracting cardiomyocytes was determined. All experimental handlings, except for the determination of the number of wells containing contracting cardiomyocytes, were carried out under yellow light, in order to prevent photodegradation of the retinoids. Each retinoid was tested in triplicate, resulting in three *in vitro* datasets per retinoid.

Data analysis

In vitro data. The *in vitro* datasets were used to derive BMC_{50} values on the basis of concentration-response analysis using the United States Environmental Protection Agency (US EPA)'s benchmark dose software (BMDS) version 2.1. For data analysis, the number of non-differentiated EBs and the total number of EBs, in both the solvent control plate and the exposure plates, were used. All models for dichotomous data were used to describe the data. Goodness-of-fit of the models was assessed based on the P-values, the scaled residuals and the graphical displays acquired. Based on the goodness-of-fit of the models, for each individual dataset it was determined which model was used to derive a BMC_{50} value. Each dataset was analyzed separately, resulting in three BMC_{50} values per chemical. The average BMC_{50} values were used for the *in vitro* potency ranking. Figures of concentration-response effects were made using SigmaPlot 9.01. Concentration-response curves were included in these figures using a 3-parameter logistic model. These curves were not used for determination of BMC_{50} values. In order to test for significant effects of Ro 41-5253 on the inhibition of cardiomyocyte differentiation by the retinoids, one-tailed Student's t-tests were used (Graphpad Prism 5).

In vivo data. Developmental toxicity studies using oral dosing of the retinoids in rat, mouse and rabbit reported in the literature, were used to determine the *in vivo* developmental toxicity potency ranking of the retinoids. The studies used are listed in Table 2. Data on malformations were used for dose-response analysis using BMDS version 2.1. For data analysis, the number of malformed fetuses and the total number of fetuses in both the control groups and the exposure groups were used. All models for dichotomous data were used to describe the *in vivo* developmental toxicity data. Goodness-of-fit of the models was assessed based on the P-values, the scaled residuals and the graphical displays acquired. Based on the goodness-of-fit of the models, for each individual dataset it was determined which model was used to derive a BMD_{10} value. The acquired BMD_{10} values were used for the *in vivo* developmental toxicity potency ranking. If present, data for total malformations were used for BMD analysis. If no data on total malformations were available, the obtained BMD_{10} value for cleft palate incidence was used for the potency ranking.

Results

Retinoids cause a concentration-dependent decrease in ES-D3 cell differentiation which is counteracted by the RAR-antagonist Ro 41-5253

All retinoids tested caused a concentration-dependent inhibition of ES-D3 cell differentiation into contracting cardiomyocytes (Figure 2). The BMC_{50} values derived from these data are presented in Table 1, including the resulting potency ranking of the retinoids in the ES-D3 cell differentiation assay. The most potent retinoid in inhibiting ES-D3 cell differentiation was TTNPB (Figure 2, Table 1). When comparing BMC_{50} values, TTNPB is two orders of magnitude more potent than ATRA, 13-*cis*-RA, 9-*cis*-RA and acitretin and four orders of magnitude more potent than retinol and etretinate (Table 1). *In vitro* potency differences between ATRA, 13-*cis*-RA, 9-*cis*-RA and acitretin are small (at most 3.6-fold), as is the *in vitro* potency difference between retinol and etretinate (1.9-fold) (Table 1).

Figure 3 shows the effects of ATRA and acitretin at their BMC_{50} concentrations, and retinol, 13-*cis*-RA, 9-*cis*-RA, etretinate and TTNPB at their BMC_{90} concentrations in the absence or presence of 5 μM of the RAR-antagonist Ro 41-5253. The results show that the retinoid-induced inhibition of ES-D3 cell differentiation is counteracted by the RAR-antagonist Ro 41-5253 (Figure 3).

Relative developmental toxicity potencies of retinoids in the MM test and the WEC test

The relative developmental toxicity potencies of the retinoids in the MM test and the WEC test as reported in the literature are shown in Table 1. In the MM test, a potency ranking of TTNPB > acitretin > ATRA > 13-*cis*-RA > etretinate was obtained by Kistler (1987). Similar results were obtained in the MM test by Reiners *et al.* (1988) and Kochhar *et al.* (1995) (Table 1).

Steele *et al.* (1987) obtained a potency ranking of ATRA ≈ 13-*cis*-RA ≈ acitretin > etretinate in the WEC test, and Cicurel and Schmid (1988) found a ranking of TTNPB > ATRA > 13-*cis*-RA in this assay. A higher potency of ATRA compared with 13-*cis*-RA was also found by others in the WEC test (Tesh, 1988; Klug *et al.*, 1989). Acitretin was reported to be more potent than etretinate (Bechter and Hall, 1987) and ATRA to be more potent than retinol (Morriss and Steele, 1977) in the WEC test (Table 1).

Table 1. Potency rankings of retinoids in the ES-D3 cell differentiation assay, the MM test and the WEC test. Potency ranking for the ES-D3 cell differentiation assay and MM test were based on BMC_{50} values (±SD) and EC_{50} values, respectively. For the WEC test studies, no EC_{50} values were reported, and the potency ranking was based on the morphological score as determined by the authors mentioned.

ES-D3 cell differentiation assay												
Present study	TTNPB	>	ATRA	>	13-*cis*-RA	>	9-*cis*-RA	>	acitretin	>	etretinate	>	retinol
BMC_{50} (nM)	0.17 (±0.026)		3.3 (±1.8)		5.0 (±3.3)		10 (±8.6)		12 (±6.3)		1500 (±610)		2800 (±1200)
MM test												
Kistler (1987)[a]	TTNPB	>	acitretin	>	ATRA	>	13-*cis*-RA	>	etretinate				
EC_{50} (nM)	0.06		50		80		400		3000				
Reiners *et al.* (1988)[b]	acitretin	>	ATRA	>	13-*cis*-RA								
EC_{50} (nM)	37		50		60								
Kochhar *et al.* (1995)[b]	ATRA	>	9-*cis*-RA	>	13-*cis*-RA								
EC_{50} (nM)	15		31		500								
WEC test												
Morris and Steele (1977)	ATRA	>	retinol										
Bechter and Hall (1987)	acitretin	>	etretinate										
Steele *et al.* (1987)	ATRA	≈	13-*cis*-RA	≈	acitretin	>	etretinate						
Cicurel and Schmid (1988)	TTNPB	>	ATRA	>	13-*cis*-RA								
Tesh (1988)	ATRA	>	13-*cis*-RA										
Klug *et al.* (1989)	ATRA	>	13-*cis*-RA										

[a] Rat limb bud cells used.
[b] Mouse limb bud cells used.

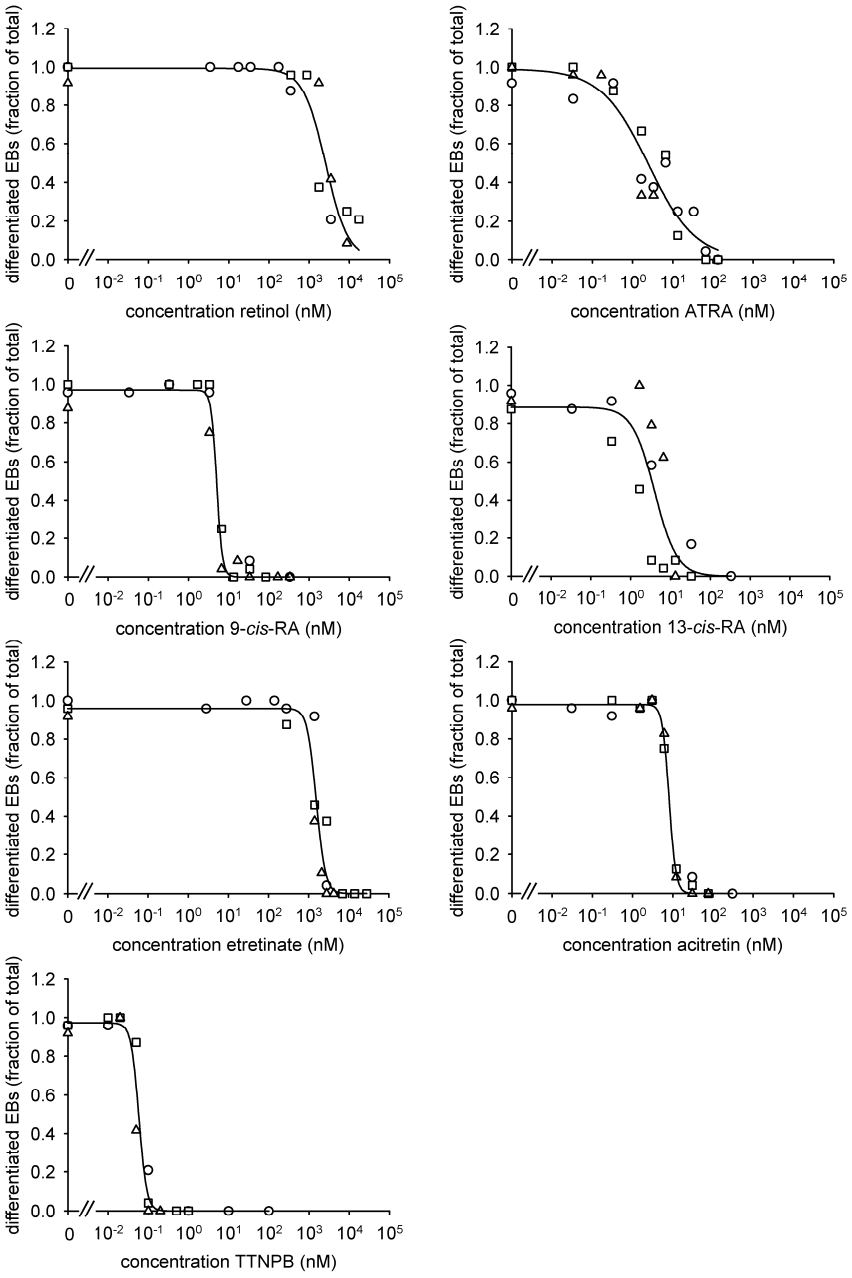

Figure 3. Retinoid-induced inhibition of ES-D3 cell differentiation. All retinoids were tested in triplicate in the ES-D3 cell differentiation assay, indicated by different symbols.

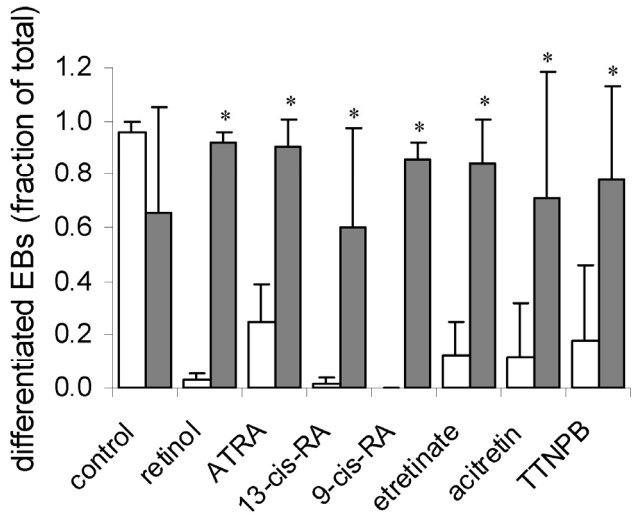

Figure 3. Counteraction of Ro 41-5253 on retinoid-induced inhibition of ES-D3 cell differentiation. ES-D3 cells were exposed to the retinoids in the absence (white bars) or presence (grey bars) of 5 μM Ro 41-5253 (n = 3-5).

* Significant difference ($p < 0.05$) between absence and presence of 5 μM Ro 41-5253.

Relative developmental toxicity potencies of retinoids in vivo

Table 2 shows the results of the BMD analysis of the *in vivo* developmental toxicity data as reported in the literature and the resulting potency rankings of the retinoids *in vivo*. Potency rankings were made for different species and/or exposure regimens for total malformations. In some of the studies, no data on total malformations were reported. In these cases the BMD_{10} values obtained for data on cleft palate incidence were used for the potency ranking. If data were available on both total malformations and cleft palate incidence, both a BMD_{10} value for total malformations and a BMD_{10} value for cleft palate incidence were obtained (Table 2). Using the BMD_{10} values presented in Table 2, different rankings were obtained, dependent on the species and/or exposure regimen used. For rats exposed on gestational day (GD) 9 or GD10 (Table 2, ranking A) the potency ranking was TTNPB > ATRA > retinol. In rats exposed on GD10-11 (Table 2, ranking B) ATRA was found to have a higher potency than acitretin. The obtained potency ranking for mice exposed on GD11 (Table 2, ranking C) was ATRA ≈ etretinate > 9-*cis*-RA > 13-*cis*-RA, whereas in mice exposed on GD7-16 (Table 2, ranking D) the potency

Table 2. Calculated BMD_{10} values for retinoid-induced total malformations and cleft palate incidence determined from reported *in vivo* developmental toxicity studies. In all studies, animals were exposed using gastric intubation. (Table continues on following page.)

Chemical	Animal	Day(s) of exposure	Doses (mg·kg bw^{-1}·day^{-1})	Reference	BMD_{10} (µmol·kg bw^{-1}·day^{-1})	
					total malformations	cleft palate incidence
A: rats exposed on GD9 or GD10						
TTNPB	Fü-albino rat	GD9	0, 0.003, 0.01, 0.03, 0.1	Kistler *et al.* (1990)	0.071	n.d.
ATRA	Wistar rat	GD10	0, 20, 30, 50, 80, 100	Tembe *et al.* (1996)	-	76
retinol	Wistar rat	GD10	0, 50, 100, 200, 500	Tembe *et al.* (1996)	-	400
13-*cis*-RA	Wistar rat	GD10	0, 100, 200, 500, 1000	Tembe *et al.* (1996)	-	n.d.
B: rats exposed on GD10-11						
ATRA	Wistar rat	GD10-11	0, 6.25, 12.5, 25, 50, 100	Turton *et al.* (1992)	5.6	n.d.
acitretin	Wistar rat	GD10-11	0, 25, 50	Turton *et al.* (1992)	67	n.d.
etretinate	Wistar rat	GD10-11	0, 25, 50	Turton *et al.* (1992)	n.d.	n.d.
13-*cis*-RA	Wistar rat	GD10-11	0, 100, 200	Turton *et al.* (1992)	n.d.	n.d.
C: mice exposed on GD11						
ATRA	CD-1 mouse	GD11	0, 10, 25, 50, 100	Kochhar *et al.* (1995)	-	19
ATRA	CD-1 mouse	GD11	0, 2.5, 10, 30, 60, 100	Campbell *et al.* (2004)	18	45
etretinate	CD-1 mouse	GD11	0, 10, 25, 100, 200	Kochhar *et al.* (1989)	26	-
9-*cis*-RA	CD-1 mouse	GD11	0, 10, 25, 50, 100	Kochhar *et al.* (1995)	-	79
13-*cis*-RA	CD-1 mouse	GD11	0, 75, 150, 300, 600, 900	Kwasigroch and Bullen (1991)	-	140

Chemical	Animal	Day(s) of exposure	Doses (mg·kg bw^{-1}·day^{-1})	Reference	BMD_{10} (µmol·kg bw^{-1}·day^{-1})	
					total malformations	cleft palate incidence
D: mice exposed on GD7-16						
TTNPB	Fü-albino mouse	GD7-16	0, 0.0003, 0.001, 0.003, 0.01	Kistler *et al.* (1987)	0.0092	-
acitretin	Fü-albino mouse	GD7-16	0, 1, 3, 10	Kistler *et al.* (1987)	6.6[a]	n.d.
13-*cis*-RA	Fü-albino mouse	GD7-16	0, 10, 30, 100, 300	Kistler *et al.* (1987)	320	-
E: rabbits exposed on GD7-19 or GD6-18						
TTNPB	Swiss hare rabbit	GD7-19	0, 0.0003, 0.001, 0.003	Kistler *et al.* (1987)	0.0027	-
acitretin	Swiss hare rabbit	GD7-19	0, 0.2, 0.6, 2	Kistler and Hummler (1985)	2.5	3.6
ATRA	Swiss hare rabbit	GD6-18	0, 0.7, 2, 6	Tzimas *et al.* (1994)	21	-
13-*cis*-RA	Swiss hare rabbit	GD6-18	0, 3, 7.5, 10	Tzimas *et al.* (1994)	26	-

n.d.: BMD_{10} value could not be determined because of inadequate data.
'–': no data available on this endpoint.
[a] BMD_{10} value for total skeletal malformations is given.

ranking was TTNPB > acitretin > 13-*cis*-RA. Finally, the potency ranking in rabbits exposed on GD7-19 or GD6-18 (Table 2, ranking E) was TTNPB > acitretin > ATRA ≈ 13-*cis*-RA.

TTNPB was the most potent retinoid *in vivo* in all species, indicated by BMD_{10} values that are three or more orders of magnitude lower than the BMD_{10} values for the other retinoids. For etretinate and 9-*cis*-RA, BMD_{10} values could only be derived for mice exposed on GD11. The potency of etretinate was found to be in the same range of that of ATRA (between 1.4-fold lower and 1.7-fold higher), whereas the potency of 9-*cis*-RA was between 1.8- and 4.4-fold lower than that of ATRA (Table 2). For retinol, a BMD_{10} value could only be derived for rats exposed on GD10. The potency of retinol was 5.3-fold lower than that of ATRA and 5600-fold lower than that of TTNPB. The *in vivo* potency ranking of ATRA, acitretin and 13-*cis*-RA was dependent on the species and/or exposure regimen used. Acitretin was found to be 12-fold less potent than ATRA in rats exposed on GD10-11, whereas acitretin (exposed on GD7-19) was found to be 8.4-fold more potent than ATRA (exposed on GD6-18) in rabbits. When comparing *in vivo* potencies of ATRA and 13-*cis*-RA, large differences were found between the derived BMD_{10} values in mice exposed on GD11 (between 3.1- and 7.8-fold), whereas in rabbits exposed on GD6-18, ATRA and 13-*cis*-RA had almost the same potency (1.2-fold difference).

Discussion

The aims of the present study were 1) to compare the potency ranking of retinoids in the ES-D3 cell differentiation assay of the EST with reported potency rankings of these retinoids in the MM test and the WEC test and 2) to determine the applicability of the ES-D3 cell differentiation assay to predict the relative *in vivo* developmental toxicity potencies of retinoids, in order to investigate the potential of the ES-D3 cell differentiation assay as a tool to set priorities for developmental toxicity testing within a series of retinoids.

All retinoids tested showed a concentration-dependent inhibition of ES-D3 cell differentiation. The obtained *in vitro* potency ranking, based on the derived BMC_{50} values, is in general in concordance with the potency rankings of the retinoids in the MM test and the WEC test, as reported in the literature (Table 1), that is

TTNPB being the most potent, etretinate and retinol being the least potent, and ATRA, 9-*cis*-RA, 13-*cis*-RA and acitretin having an intermediate potency. However, the relative potencies of the retinoids with intermediate potency (ATRA, 9-*cis*-RA, 13-*cis*-RA and acitretin) is slightly different in the test systems (Table 1). Acitretin was 1.4-1.6-fold more potent than ATRA in the MM test, whereas in the present study ATRA was 3.6-fold more potent than acitretin (Table 1). In the WEC test, no clear difference in potency between ATRA and acitretin was observed. When comparing the EC_{50} values of ATRA and 13-*cis*-RA in the MM test, Kistler (1987) and Kochhar *et al.* (1995) obtained large differences of respectively 5- and 33-fold, whereas Reiners *et al.* (1988) obtained a difference of only 1.2-fold (Table 2). In the present study, the difference between the BMC_{50} values of ATRA and 13-*cis*-RA was 1.5-fold (Table 1). When interpreting the results of the ATRA isomers, it must be noted that during exposure, isomerization to other isomers can take place. This has been shown to happen for ATRA and 13-*cis*-RA in the WEC test (Klug *et al.*, 1989), resulting in a combined exposure to the different retinoic acid isomers. Therefore, a specific exposure to one of the retinoic acid isomers is not possible, to some extent hampering the interpretation of the results for the individual retinoic acid isomers.

To determine the applicability of the ES-D3 cell differentiation assay to predict the developmental toxicity potency ranking of retinoids *in vivo*, the *in vitro* developmental toxicity potency ranking was compared with the *in vivo* developmental toxicity potency rankings, based on BMD_{10} values derived from reported *in vivo* developmental toxicity studies. If possible, BMD_{10} values were obtained for total malformations. If these data were not presented in the studies used, data on cleft palate incidence were used, in order to include as many retinoid congeners as possible in the *in vivo* potency rankings. It must be noted that different BMD_{10} values for total malformations compared with BMD_{10} values for cleft palate incidence might be expected using an identical exposure regimen. This is reflected for example in the 2.5-fold higher BMD_{10} value for cleft palate incidence than the BMD_{10} value for total malformations, obtained from the data of Campbell *et al.* (2004) for mice exposed on GD11 to ATRA (Table 2). On the other hand, the BMD_{10} value for cleft palate incidence obtained from the data of Kochhar *et al.* (1995) is only 1.1-fold higher than the BMD_{10} value for total malformations obtained

from the data of Campbell *et al.* (2004) (Table 2), suggesting that valid *in vivo* potency rankings could be obtained, based on BMD_{10} values derived for total malformations and/or cleft palate incidence.

The results show that different *in vivo* developmental toxicity potency rankings were obtained among different species and exposure regimens (Table 2). Differences in dynamics and kinetics among species might play a role in the differences in *in vivo* developmental toxicity potency rankings. In all species, TTNPB was the most potent congener *in vivo*, which was correctly predicted by the ES-D3 cell differentiation assay. The retinoids ATRA, 9-*cis*-RA and acitretin showed intermediate potencies *in vivo*, which was also found in the ES-D3 cell differentiation assay (Table 1). The lowest potency, based on the derived BMD_{10} values, was obtained for retinol (Table 2), as was also correctly predicted by the ES-D3 cell differentiation assay. However, although no BMD_{10} value could be determined for total malformations or cleft palate incidence for 13-*cis*-RA in rats, it must be noted that a lower potency of 13-*cis*-RA compared to that of retinol was reported for rats being exposed on GD10 (Tembe *et al.*, 1996).

Etretinate was one of the least potent retinoids in the ES-D3 cell differentiation assay (Table 1), whereas it was found to be relatively potent *in vivo* (Table 2). Reiners *et al.* (1988) compared the *in vitro* (MM test) and *in vivo* (mouse) potencies of ATRA, etretinate and etretinate's metabolite acitretin. A relatively low potency of etretinate was obtained *in vitro*, whereas it was relatively potent *in vivo*. The high *in vivo* potency is due to the metabolism of etretinate to acitretin, leading to high concentrations of acitretin in the fetuses (Reiners *et al.*, 1988). The lack or low expression level of this bioactivation system in the cells used in the ES-D3 cell differentiation assay, in the limb bud cells used in the MM test, and in the embryos used in the WEC test, might explain the misclassification of etretinate as a retinoid with a relatively low teratogenic potency. Etretinate's potency was increased in the WEC test and the MM test when carboxylic ester hydrolase was added to the test system (Bechter and Hall, 1987; Kochhar *et al.*, 1989), resulting in the formation of etretinate's metabolite acitretin (Kochhar *et al.*, 1989), corroborating the need of acitretin formation in etretinate-induced developmental toxicity.

The relatively low potency of 13-*cis*-RA in rats and mice was not found in rabbits, having only a 1.2-fold lower potency than ATRA (Table 2). This corresponds with

the results of the ES-D3 cell differentiation assay, in which 13-*cis*-RA was 1.5-fold less potent than ATRA (Table 1). The low potency of 13-*cis*-RA compared with ATRA in rats and mice *in vivo* is ascribed to a lower placental transfer of 13-*cis*-RA than that of ATRA (Collins *et al.*, 1994; Tzimas *et al.*, 1995). Although the placental transfer of 13-*cis*-RA is also lower than that of ATRA in rabbits (Tzimas *et al.*, 1994), both retinoids have almost the same developmental toxicity potencies in rabbits. This might be due to a slower excretion of 13-*cis*-RA than of ATRA in this species (Tzimas *et al.*, 1994). The role of *in vivo* kinetics in the interspecies differences in *in vivo* developmental toxicity potencies of ATRA and 13-*cis*-RA is summarized by Nau (2001).

These examples show that the knowledge of *in vivo* kinetic processes (absorption, distribution, metabolism and excretion) is indispensable when interpreting results of *in vitro* toxicity experiments for estimating *in vivo* potency rankings. When combining *in vitro* toxicity data with information on *in vivo* kinetics, better estimations of *in vivo* toxicity potencies or *in vivo* toxic dose levels can be made. Previously, we have shown that with the use of physiologically based kinetic (PBK) models, which describe the *in vivo* kinetics of chemicals, *in vitro* toxic effect concentrations obtained in the ES-D3 cell differentiation assay can be extrapolated to *in vivo* dose levels (Verwei *et al.*, 2006; Louisse *et al.*, 2010b). Chapter 7 presents such a combined *in vitro-in silico* model for the prediction of developmental toxicity dose levels of ATRA in rat and human.

Retinoid-induced developmental toxicity effects are suggested to be RAR-mediated. Disruption of ATRA-induced pathways might result in disruption of the embryonic development resulting in malformations (Collins and Mao, 1999). Ali-Khan and Hales (2006) showed that retinol-induced limb defects in mouse limbs (*ex vivo*) are prevented by the RAR-antagonist BMS 453. Furthermore, Eckhardt and Schmitt (1994) showed that Ro 40-6055-induced developmental toxicity in the MM test was counteracted by the RAR-antagonist Ro 41-5253. The present study shows that Ro 41-5253 also counteracts the retinoid-induced inhibition of ES-D3 cell differentiation, suggesting that the RAR is involved in the *in vitro* developmental toxicity effects induced by the retinoids in the ES-D3 cell differentiation assay. This indicates that the ES-D3 cell differentiation assay contains a relevant pathway for retinoid-induced developmental toxicity, making it

a useful tool for assessing developmental toxicity potencies of retinoids and possible other RAR-agonists.

Altogether, it is concluded that the ES-D3 cell differentiation assay results in a similar potency ranking of retinoids as obtained in the MM test and the WEC test, that is TTNPB being the most potent retinoid, ATRA, 13-*cis*-RA, 9-*cis*-RA and acitretin showing an intermediate potency, and etretinate and retinol being the least potent. Therefore, it can be seen as an equally adequate *in vitro* method for predicting *in vivo* developmental toxicity potencies as the other *in vitro* assays for developmental toxicity. Moreover, because the ES-D3 cell differentiation assay does not use primary animal tissues, it can even be seen as a better *in vitro* method, leading to a larger reduction of animal use in developmental toxicity testing. However, although TTNPB was correctly classified as the most potent retinoid by the ES-D3 cell differentiation assay and retinol as the least potent, not all retinoids were correctly ranked on their *in vivo* potencies. This shows that the ES-D3 cell differentiation assay is not sufficient as standalone for predicting *in vivo* developmental toxicity potencies of retinoids. The deviating ranking could be explained by the lack of *in vivo* kinetic processes in the *in vitro* assay, which play a role in the developmental toxicity potencies of these retinoids *in vivo*. Therefore, knowledge on chemical-specific species-dependent *in vivo* kinetics is essential when using *in vitro* toxicity data for estimation of *in vivo* developmental toxicity potencies within series of related chemicals.

Acknowledgments

The authors would like to thank Han van de Sandt and Bas Blaauboer for their stimulating contributions within the framework of the project.

References

Ali-Khan, S. E., Hales, B. F. (2006). Retinoid receptor antagonists alter the pattern of apoptosis in organogenesis stage mouse limbs. *Toxicol Sci* **90**, 208-20.

Bechter, R., Hall, L. (1987). In vitro embryotoxicity and teratogenic potential of etretinate and etretin. *Toxicol In Vitro* **1**, 225-31.

Campbell Jr., J. L., Smith, M. A., Fisher, J. W., Warren, D. A. (2004). Dose-response for retinoic acid-

induced forelimb malformations and cleft palate: a comparison of computerized image analysis and visual inspection. *Birth Defects Res B Dev Reprod Toxicol* **71**, 289-95.

Collins, M. D., Mao, G. E. (1999). Teratology of retinoids. *Annu Rev Pharmacol Toxicol* **39**, 399-430.

Collins, M. D., Tzimas, G., Hummler, H., Bürgin, H., Nau, H. (1994). Comparative teratology and transplacental pharmacokinetics of all-trans-retinoic acid, 13-cis-retinoic acid, and retinyl palmitate following daily administrations in rats. *Toxicol Appl Pharmacol* **127**, 132-144.

Cicurel, L., Schmid, B. P. (1988). Postimplantation embryo culture for the assessment of the teratogenic potential and potency of chemicals. *Experientia* **44**, 833-40.

De Jong, E., Louisse, J., Verwei, M., Blaauboer, B. J., van de Sandt, J. J. M., Woutersen, R. A., Rietjens, I. M. C. M., Piersma, A. H. (2009). Relative developmental toxicity of glycol ether alkoxy acid metabolites in the embryonic stem cell test as compared to the *in vivo* potency of their parent compounds. *Toxicol Sci* **110**, 117-24.

De Smedt, M., Steemans, M., de Boeck, A. K., Peters, B. J., van der Leede, F., van Goethem, A., Lampo, Vanparys, P. (2008). Optimisation of the cell cultivation methods in the embryonic stem cell test results in an increased differentiation potential of the cells into strong beating myocard cells. *Toxicol In Vitro* **22**, 1789-96.

Eckhardt, K., Schmitt, G. (1994). A retinoic acid receptor alpha antagonist counteracts retinoid teratogenicity in vitro and reduced incidence and/or severity of malformations in vivo. *Toxicol Lett* **70**, 299-308.

Genschow, E., Spielmann, H., Scholz, G., Seiler, A., Brown, N. A., Piersma, A. H., Brady, M., Clemann, N., Huuskonen, H., Paillard, F., Bremer, S., Becker, K. (2002). The ECVAM international validation study on in vitro embryotoxicity tests: results of the definitive phase and evaluation of prediction models. European Centre for the Validation of Alternative Methods. *Atern Lab Anim* **30**, 151-76.

Genschow, E., Spielmann, H., Scholz, G., Pohl, I., Seiler, A., Clemann, N., Bremer, S., Becker, K. (2004). Validation of the embryonic stem cell test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* **32**, 209-44.

Kistler, A., Hummler, H. (1985). Teratogenesis and reproductive safety evaluation of the retinoid etretin (Ro 10-1670). *Arch Toxicol* **58**, 50-56.

Kistler, A. (1987). Limb bud cell cultures for estimating the teratogenic potential of chemicals. Validation of the test system with retinoids. *Arch Toxicol* **60**, 403-14.

Kistler, A., Galli, B., Howard, W. B. (1990). Comparative teratogenicity of three retinoids: the arotinoids Ro 13-7410, Ro-13-6298 and Ro 15-1570. *Arch Toxicol* **64**, 43-8.

Klug, S., Creech Kraft, J., Wildi, E., Merker, H. J., Persaud, T. V. N., Nau, H., Neubert, D. (1989). Influence of 13-cis and all-trans retinoic acid on rat embryonic development in vitro: correlation with isomerisation and drug transfer to the embryo. *Arch Toxicol* **63**, 185-92.

Kochhar, D. M., Penner, J. D., Minutella, L. M. (1989). Biotransformation of etretinate and developmental toxicity of etretin and other aromatic retinoids in teratogenesis bioassays. *Drug Met Dispos* **17**, 618-24.

Kochhar, D. M., Jiang, H., Penner, J. D., Heyman, R. A. (1995). Placental transfer and developmental effects of 9-cis-retinoic acid in mice. *Teratology* **51**, 257-65.

Kwasigroch, T. E., Bullen, M. (1991). Effects of isotretinoin (13-cis-retinoic acid) on the development of mouse limbs in vivo and in vitro. *Teratology* **44**, 605-16.

Louisse, J., Bai, Y., Verwei, M., van de Sandt, J. J. M., Blaauboer, B. J., Rietjens, I. M. C. M. (2010a).

Decrease of intracellular pH as possible mechanism of embryotoxicity of glycol ether alkoxyacetic acid metabolites. *Toxicol Appl Pharmacol* **245**, 236-43.

Louisse, J., de Jong, E., van de Sandt, J. J. M., Blaauboer, B. J., Woutersen, R. A., Piersma, A. H., Rietjens, I. M. C. M., Verwei, M. (2010b). The use of in vitro toxicity data and physiologically based kinetic modeling to predict dose-response curves for in vivo developmental toxicity of glycol ethers in rat and man. *Toxicol Sci* **118**, 470-84.

Marx-Stoelting, P., Adriaens, E., Ahr, H. J., Bremer, S., Garthoff, B., Gelbke, H. P., Piersma, A. H., Pellizzer, C., Reuter, U., Rogiers, V., Schenk, B., Schwengberg, S., Seiler, A., Spielmann, H., Steemans, M., Stedman, D. B., Vanparys, P., Vericat, J. A., Verwei, M., van de Water, F., Weimer, M., Schwarz, M. (2009). A review of the implementation of the embryonic stem cell test (EST), the report and recommendation of an ECVAM/ReProTect workshop. *Atern Lab Anim* **37**, 313-28.

Morriss, G. M., Steele, C. E. (1977). Comparison of the effects of retinol and retinoic acid on postimplantation rat embryos in vitro. *Teratology* **15**, 109-19.

Nau, H. (2001). Teratogenicity of isotretinoin revisited: species variation and the role of all-trans-retinoic acid. *J Am Acad Dermatol* **45**, S183-87.

Reiners, J., Löfberg, B., Creech Kraft, J., Kochhar, D. M., Nau, H. (1988). Transplacental pharmacokinetics of teratogenic doses of etretinate and other aromatic retinoids in mice. *Reprod Toxicol* **2**, 19-29.

Steele, C. E., Marlow, R., Turton, J., Hicks, R. M. (1987). In-vitro teratogenicity of retinoids. *Br J Exp Pathol* **68**, 215-23.

Tembe, E. A., Honeywell, R., Buss, N. E., Renwick, A. G. (1996). All-trans-retinoic acid in maternal plasma and teratogenicity in rats and rabbits. *Toxicol Appl Pharmacol* **141**, 456-72.

Tesh, J. M. (1988). The application of whole-embryo culture to new product development. *Toxicol In Vitro* **2**, 189-94.

Turton, J. A., Willars, G. B., Haselden, J. N., Ward, S. J., Steele, C. E., Hicks, R. M. (1992). Comparative teratogenicity of nine retinoids in the rat. *Int J Exp Pathol* **73**, 551-63.

Tzimas, G., Bürgin, H., Collins, M. D., Hummler, H., Nau, H. (1994). The high sensitivity of the rabbit to the teratogenic effects of 13-cis-retinoic acid (isotretinoin) is a consequence of prolonged exposure of the embryo to 13-cis-retinoic acid and 13-cis-4-oxo-retinoic acid, and not isomerization to all-trans-retinoic acid. *Arch Toxciol* **68**, 119-28.

Tzimas, G., Collins, M. D., Nau, H. (1995). Developmental stage-associated differences in the transplacental distribution of 13-cis- and all-trans-retinoic acid as well as their glucuronides in rats and mice. *Toxicol Appl Pharmacol* **133**, 91-101.

Van der Jagt, K., Munn, S. J., Tørsløv, J., de Bruijn, J. (2004). Alternative approaches can reduce the use of test animals under REACH. *Report EUR* **21405**.

Verwei, M., van Burgsteden, J. A., Krul, C. A. M., van de Sandt, J. J. M., Freidig, A. P. (2006). Prediction of in vivo embryotoxic effect levels with a combination of in vitro studies and PBPK modelling. *Toxicol Lett* **165**, 79-87.

CHAPTER 7

Prediction of *in vivo* developmental toxicity of all-*trans*-retinoic acid by combining *in vitro* toxicity data and physiologically based kinetic modeling

Jochem Louisse, Bas J. Blaauboer, Ivonne M.C.M. Rietjens, Miriam Verwei

In preparation.

Chapter 7

Abstract

The use of high numbers of laboratory animals for toxicity testing in chemical safety assessment meets increasing ethical and economic constraints. The development, validation and application of reliable alternatives for *in vivo* toxicity testing is therefore urgently needed. In order to use toxicity data obtained from *in vitro* assays for risk assessment, *in vitro* concentration-response curves need to be translated into *in vivo* dose response-curves. This can be achieved by reverse dosimetry using physiologically based kinetic (PBK) modeling. The present study provides a proof-of-principle of this combined *in vitro-in silico* approach by translating *in vitro* concentration-response curves of the retinoid all-*trans*-retinoic acid (ATRA) obtained in the embryonic stem cell (ES-D3 cell) differentiation assay of the embryonic stem cell test (EST), into *in vivo* dose-response curves for developmental toxicity using a PBK model that is solely based on kinetic model parameter values derived using *in vitro* techniques. The results reveal that dose levels of ATRA causing *in vivo* developmental toxicity were predicted within one order of magnitude compared with reported *in vivo* developmental toxicity data from the literature. This indicates for the first time the feasibility of using only *in vitro* toxicity and kinetic data for defining an *in vivo* dose-response curve suitable for risk assessment practice.

Introduction

For the registration of chemicals, such as pharmaceuticals, industrial chemicals, food additives and pesticides, regulatory agencies require safety studies performed in laboratory animals. Increased use of laboratory animals is foreseen for safety evaluations because of the European Commission's REACH (Registration, Evaluation, Authorization and restriction of CHemicals) legislation. It is expected that more than half of the laboratory animals needed for REACH will be used for reproductive and developmental toxicity testing (Van der Jagt *et al.*, 2004). The use of these high numbers of laboratory animals for toxicity testing meets increasing ethical and economic constraints, indicating that the development, validation and application of reliable alternatives for *in vivo* studies is urgently needed. Currently, *in vitro* assays are considered to be useful in the development and early safety testing of pharmaceuticals, new food additives and pesticides, and for setting priorities for toxicity testing of industrial chemicals within the REACH framework. However, at present *in vitro* toxicity data are hardly used in chemical risk assessment. One of the reasons is that concentration-response curves are obtained in *in vitro* assays, whereas *in vivo* dose-response curves are required in the hazard characterization process. Therefore, *in vitro* toxic effect concentrations need to be translated into *in vivo* toxic dose levels, in order to increase the acceptance and implementation of *in vitro* assays in toxicological risk assessment strategies.

The translation of *in vitro* toxicity data into *in vivo* effect doses (DeJongh *et al.*, 1999; Verwei *et al.*, 2006a; Forsby and Blaauboer, 2007; Rotroff *et al.*, 2010; Wetmore *et al.*, 2012) or *in vivo* dose-response curves (Paini *et al.*, 2010; Louisse *et al.*, 2010) can be achieved using physiologically based kinetic (PBK) modeling with a reverse dosimetry approach. With reverse dosimetry, *in vitro* concentration-response data are set as internal concentrations in a PBK model, after which the PBK model calculates the doses that are needed to reach these internal concentrations (in the blood or target tissue) in the *in vivo* situation. This translation results in predicted *in vivo* dose-response data (Figure 1).

We have recently shown that with this reverse dosimetry approach, *in vivo* dose-response curves for the developmental toxicity of a group of glycol ethers can be reliably predicted based on *in vitro* concentration-response curves obtained in the embryonic stem cell (ES-D3 cell) differentiation assay of the embryonic stem cell

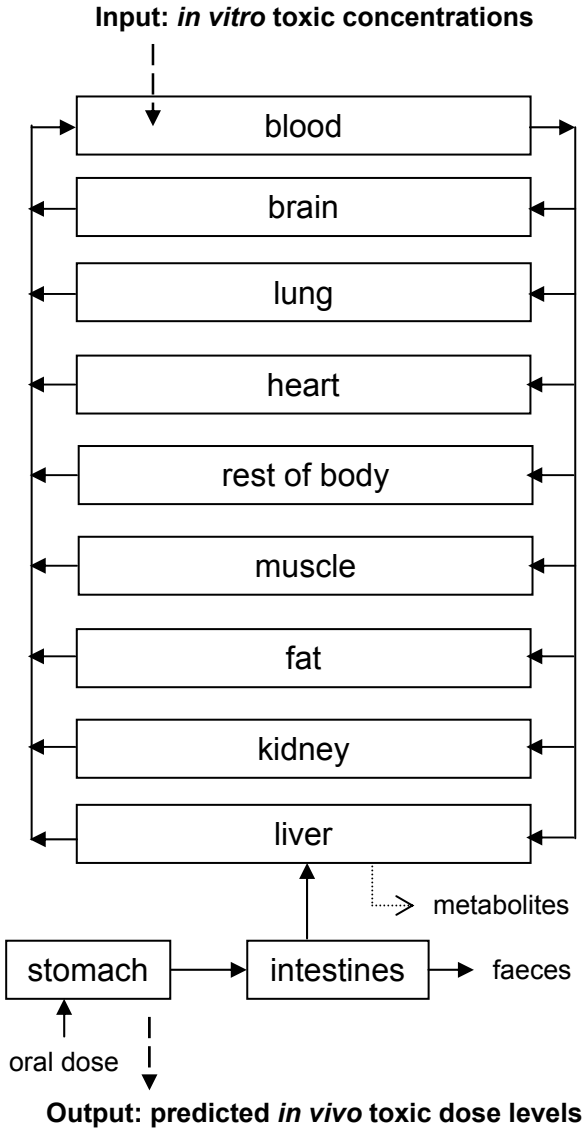

Figure 1. Graphic representation of the PBK model developed for ATRA to extrapolate the *in vitro* concentration-response curve of ATRA determined in the ES-D3 cell differentiation assay to predicted dose-response curves for ATRA-induced *in vivo* developmental toxicity in rat and human.

test (EST) (Louisse *et al.*, 2010). To further extend the potential of the approach to predict *in vivo* developmental toxicity dose levels, examples with chemicals from

other chemical categories are needed. In addition, PBK models, which are currently often built using *in vivo* kinetic data (Rietjens *et al.*, 2011), should instead be preferentially based on parameter values obtained using only *in vitro* and/or *in silico* techniques to increase the contribution to the replacement, reduction and refinement (3Rs) of animal use in toxicological research.

The aim of the present study was to assess the feasibility of predicting *in vivo* developmental toxicity dose levels of the retinoid all-*trans*-retinoic acid (ATRA) with reverse dosimetry, using a PBK model that is based on kinetic parameter values derived using *in vitro* techniques only. ATRA is the major metabolite of retinol (vitamin A) and is required for various processes in the body (Collins and Mao, 1999). Proper levels of vitamin A must be maintained for normal embryogenesis, whereas vitamin A deficiency and high retinoid intake induce adverse effects to the developing conceptus (Collins and Mao, 1999). To predict dose-response curves for *in vivo* ATRA-induced developmental toxicity, the reverse dosimetry approach using PBK modeling was applied. In the present study, the following steps were followed: I) the determination of an *in vitro* concentration-response curve for ATRA-induced developmental toxicity, II) the development of a PBK model for ATRA kinetics in rat and human, III) the evaluation of the PBK model predictions for *in vivo* ATRA kinetics, IV) the translation of the *in vitro* concentration-response curve into predicted *in vivo* dose-response curves for rat and human using the PBK model for reverse dosimetry and V) the comparison of the predicted dose-response curves for ATRA-induced developmental toxicity with data on *in vivo* ATRA-induced developmental toxicity in rats, in order to evaluate the predictive value of the approach.

The results obtained show that the reverse dosimetry approach, combining *in vitro* toxicity data and *in silico* PBK modeling, is a promising tool to predict *in vivo* dose-response curves for ATRA-induced developmental toxicity. To the best of our knowledge, this is the first proof-of-principle that *in vivo* dose-response curves for a systemic endpoint can be predicted without the need of *in vivo* studies. The predicted dose-response curves obtained can be used to set points of departure to derive safe exposure levels for humans, pointing at the potential of the approach to increase the acceptance and implementation of *in vitro* assays in toxicological risk assessment strategies.

Materials and methods

Chemicals

ATRA and taurocholate were purchased from Sigma-Aldrich (Zwijndrecht, the Netherlands) and [11,12-^{3}H(N)]-ATRA, [1-methyl-^{14}C]-caffeine and D-[1-^{3}H(N)]-mannitol from Perkin-Elmer (Groningen, the Netherlands). Polydimethylsiloxane (PDMS) was obtained from General Electrics Plastics BV (Bergen op Zoom, the Netherlands) and hexane from Biosolve (Valkenswaard, the Netherlands). Ethanol, KOH and HEPES were purchased from Merck Chemicals (Darmstadt, Germany).

Reverse dosimetry approach to translate in vitro concentration-response curves into in vivo dose-response curves for developmental toxicity

I) Determination of in vitro developmental toxicity in the ES-D3 cell differentiation assay

The *in vitro* developmental toxicity data of ATRA in the ES-D3 cell differentiation assay were taken from our previous study (Louisse *et al.*, 2011). The data of three independent experiments were combined to one data set, by calculating per concentration the total number of non-differentiated embryoid bodies (EBs) as a fraction of the total number of EBs (sample size) used for that concentration. The data were used for BMD analysis using the models for quantal data of the PROAST software, developed by the Dutch National Institute for Public Health and the Environment (RIVM, Bilthoven, the Netherlands), to determine Benchmark concentrations for the inhibition of ES-D3 cell differentiation.

II) Development of a PBK model for ATRA kinetics in rat and human

The following steps were followed to develop a PBK model for ATRA: 1) definition of a conceptual model, 2) translation of the conceptual model into a mathematical model by defining mathematical equations, 3) defining the model parameter values and 4) solving the mathematical equations using appropriate software. Subsequently, the model predictions were evaluated to determine whether the model can be used for reverse dosimetry.

The conceptual model defined for ATRA is shown in Figure 1 (step 1). The kinetic processes for this model were translated into a mathematical model for Berkeley Madonna software (UC Berkeley, CA, USA) (step 2). The processes of stomach

emptying (rat, liquid: 15 min (Reilly *et al.*, 1990), human, liquid: 15 min (Sun *et al.*, 1988)) and the small intestinal transition (rat: 1.5h, human: 4h (Davies and Morris, 1993)) are included in the model. Values for anatomical and physiological parameters were taken from literature (Brown *et al.*, 1997, Table 1), whereas values for chemical-specific kinetic model parameters were determined in the present study (step 3). These kinetic parameter values were obtained using data derived using only *in vitro* techniques and include parameter values for 1) intestinal absorption, 2) partitioning into tissues (i.e. distribution) and 3) metabolic clearance (i.e. metabolism and subsequent excretion). The mathematical equations were subsequently solved (step 4) using Berkeley Madonna software.

Table 1. Anatomical and physiological parameter values used in PBK model. Values were taken from Brown *et al.* (1997).

Physiological parameters	**Values**	
	rat	**human**
tissue volumes (percentage of body weight)		
liver	2.7	2.6
muscle	44	40
kidney	0.54	0.40
heart	0.24	0.50
brain	0.42	2.0
lung	0.37	0.80
fat	10	21
rest of body	31	23
blood	5.9	7.9
cardiac output ($L \cdot h^{-1} \cdot kg\ bw^{-0.74}$)	14.0	19.2
tissue blood flows (percentage of cardiac output)		
liver	17.4	22.7
muscle	27.8	19.1
kidney	14.1	17.5
heart	5.1	4.0
brain	2.0	11.4
lung	2.1	2.5
fat	7.0	5.2
rest of body	24.5	17.6

Intestinal absorption. Intestinal ATRA uptake values were determined by using apparent permeability (P_{app}) coefficients ((amount transported (μmol) · $time^{-1}$ (s))/(surface area (cm^2) · concentration applied (mM)) for transport across human intestinal colorectal adenocarcinoma (CaCo-2) monolayers. To this end, CaCo-2 cells (P7-P9, obtained from ATCC) were grown on Costar 12-well transwell plate inserts (Corning; membrane pore size: 0.4 μm, surface area: 1.12 cm^2) by adding 0.5 mL of a $2.24 \cdot 10^5$ cell suspension in culture medium (DMEM (Invitrogen, Breda, the Netherlands, cat. no. 61965), supplemented with 10% FCS (BioWhittaker, Verviers, Belgium), 100 μM NEAA (Invitrogen) and 50 μg/ml gentamycin (Invitrogen)) to each transwell plate insert (apical compartment) and 1.5 mL of culture medium to the basolateral compartment. Cells were grown in a humidified atmosphere (5% CO_2) at 37°C. Culture medium was refreshed every two to three days. Transport studies were performed 20 days post-seeding. For transport studies, the apical compartment was filled at t=0 with 0.65 mL 10 mM HEPES in HBSS (Invitrogen) (pH 6.5) containing the test chemicals ([11,12-^{3}H(N)]-ATRA (1 nM – 10 μM), [1-methyl-^{14}C]-caffeine (10 μM, control for transcellular transport) or D-[1-^{3}H(N)]-mannitol (10 μM, control for paracellular transport)), in the absence or presence of 10 mM taurocholate. The basolateral compartment was filled with 1.5 mL HBSS (pH 7.4) in the absence or presence of 0.6 mM BSA (Sigma-Aldrich). Samples from the apical compartment (0.15 mL) were taken at t=0 min and t=20 min. Samples from the basolateral compartment (0.5 mL) were taken at t=10 min (followed by the addition of 0.5 mL HBSS to retain a volume of 1.5 mL) and t=20 min. Initial experiments using varying incubation times showed that transport velocity was linear until 30 minutes (data not shown). At t=20 min, the transwell filters with cells were washed with HBSS and incubated in 0.5 mL 80% 1.5 M KOH/20% ethanol. Radioactivity was measured after adding 10 mL of OptiPhase HiSafe (Perkin Elmer) to the samples and the filters with cells, using a Tri-Carb Liquid Scintillation Counter (Perkin Elmer). The calculated amount in the three compartments (apical compartment, basolateral compartment and filter plus cells) were used for mass balance calculations.

The rate of ATRA transfer can be calculated using these P_{app} coefficients, as was done before for folic acid by Verwei *et al.* (2006b). Because P_{app} coefficients ((amount transported/time)/(surface area × concentration applied) are values

describing the transport velocity (amount per time), depending on the surface area and the concentration applied, the transfer rate across the intestinal wall can be obtained by multiplying the P_{app} coefficient by the surface area of the intestine and the luminal concentration in the intestine. In the present PBK model this resulted in: intestinal transfer rate ATRA ($\mu mol \cdot h^{-1}$) = P_{app} coefficient ($cm \cdot h^{-1}$) × absorptive surface area (cm^2) × luminal concentration of ATRA (mM). The absorptive surface area of the rat intestines (94 cm^2) was calculated assuming a radius of 0.18 cm (Zimmerman *et al.*, 2001) and a small intestine length of 83 cm (McConnell *et al.*, 2008). The absorptive surface area of the human intestines (25800 cm^2) was taken from Verwei *et al.* (2006b). The luminal concentration of ATRA is calculated by the model by dividing the amount of ATRA (µmol) by the small intestinal volume (mL). The small intestinal volume was calculated to be 8.4 mL for rat (radius: 0.18 cm, length: 83 cm) and 9 L for human (radius: 2.5 cm (Kararli, 1995), length: 460 cm (Hossenpour and Behdad, 2008)).

Tissue:blood partition coefficients. Tissue:blood partition coefficients were calculated by dividing obtained tissue:buffer (PBS, Invitrogen) partition coefficients by obtained blood:buffer partition coefficients. The tissue:buffer and blood:buffer partition coefficients were determined using a recently developed SPME-based approach developed by the Netherlands Organization for Applied Scientific Research (TNO, Zeist, the Netherlands) (Vaes, 2010). In this method, the chemical of interest is able to distribute over three different phases (compartments), which are i) a buffer phase, ii) a serum (or tissue homogenate) phase and iii) a coating phase. It is assumed that the chemical will distribute over these phases according to the following equation:

$$1)\ M_t = C_b * V_b + C_s * V_s + C_c * V_c$$

in which M_t = total mass, C_b = concentration in buffer, V_b = volume buffer, C_s = concentration in serum (or tissue homogenate), V_s = volume serum (or tissue homogenate), C_c = concentration in coating and V_c = volume coating. The concentration of the chemical in the coating can be described as:

$$2)\ C_c = C_b * K_c$$

in which C_c = concentration in coating, C_b = concentration in buffer and K_c = coating:buffer partition coefficient. The concentration of the chemical in the serum

can be described as:

$$3)\ C_s = C_b * K_s$$

in which C_s = concentration in serum (or tissue homogenate), C_b = concentration in buffer and K_s = serum (or tissue homegenate):buffer partition coefficient.

These equations can be rewritten as:

$$4)\ A_c = M_t / (1 + ((V_b + K_s * V_s) / (V_c * K_c)))$$

in which A_c = amount in coating, M_t = total mass, V_b = volume buffer, K_s = serum (or tissue homogenate):buffer partition coefficient, V_s = volume serum (or tissue homogenate), V_c = volume coating and K_c = coating:buffer partition coefficient.

In the experimental set-up, 96-well plates (Flexible 96, Clear 96-well Flexible PET microplates, round bottom (Perkin Elmer)) were coated with increasing volumes of polydimethylsiloxane (PDMS) (V_c). This was achieved by adding 200 µL of varying PDMS concentrations dissolved in hexane to the wells, after which the hexane was allowed to evaporate. Increasing concentrations of serum (Sigma-Aldrich) (or tissue homogenates (tissues obtained from a supernumerary 2-months -old male Wistar rat of the animal facility)), obtained by dilution in buffer, were spiked with [11,12-^{3}H(N)]-ATRA, after which 200 µL of these solutions were added per (coated) well. The serum (or tissue homogenate) concentration can be described as a volume of serum and a volume of buffer (together amounting to 200 µL) resulting in varying serum (or tissue homogenates) volumes (V_s) and buffer volumes (V_b). It was assumed that ATRA in blood binds only to serum components, assuming that blood consists of two third of plasma (Lee and Blaufox, 1985), implicating that pure blood was considered as 1.5 times diluted plasma. Per 96-well plate, a maximum of 96 situations can be created with different combinations of V_c, V_b and V_s. Plates are put overnight at an orbital shaker to allow equilibration of the chemical over the different phases. Subsequently, per well, 50 µL of the medium was removed from the wells and transferred to a non-coated 96-well plate, and the remaining medium from the coated 96-well plate was removed and the plate was washed. To both plates, 200 µL of OptiPhase HiSafe was added per well, after which the plates were covered with TopSeal®-A-96-well microplate press-on adhesive sealing film (Perkin Elmer). Radioactivity measurements were

performed using a 1450-023 MicroBeta TriLux Microplate Scintillation and Luminescence Counter (Perkin Elmer) to obtain A_c (radioactivity in coating) and M_t (radioactivity in coating + radioactivity of transferred medium (times four, because one fourth of medium is used for radioactivity counting of medium)). From equation 4, V_b, V_s and V_c are varying (chosen) values and A_c and M_t are determined by radioactivity measurements. The remaining values for K_s and K_c are unknown. These are estimated using the solver function of Microsoft Excel by solving equation 4 for the maximal 96 different combinations of V_b, V_s and V_c, resulting in an estimated K_s and K_c. The free fraction in pure serum (or tissue homegenate) can be obtained by using the K_s (free fraction = 1 / (K_s + 1)).

Metabolic clearance. Data on the *in vitro* kinetic conversion of ATRA by microsomes as reported in the literature (Table 2) were used to determine the model parameter values for metabolic clearance. The reported apparent K_m values were applied to the model, whereas the apparent V_{max} values were scaled to a complete organ assuming a microsomal protein yield of 35 mg/g rat liver (Medinsky *et al.*, 1994), 7 mg/g rat kidney (Beierschmitt and Weiner, 1986), 20 mg/g rat lung (Medinsky *et al.*, 1994), 4 mg/g rat brain (Ravindranath and Anandatheerthavarada, 1990) and 32 mg/g human liver (Barter *et al.*, 2007). For the rat model, kinetic constants obtained using liver, kidney, lung and brain fractions were used. For the human model, kinetic constants on the formation of 4-OH-ATRA were obtained using cytochrome P450 (CYP) enzyme kinetics (CYP2C8, CYP3A4, CYP3A5, CYP3A7, CYP26A1) and kinetic constants on the formation of all-*trans*-retinoyl β-glucuronide were obtained using uridine 5'-diphospho-glucuronosyltransferase (UGT) enzyme kinetics (UGT2B7). 4-OH-ATRA levels *in vivo* are low, probably due to a rapid oxidation of the C4-hydroxylated ATRA (Marchetti *et al.*, 1997), resulting in ATRA's main oxidized metabolite 4-oxo-ATRA. The expression levels of CYP and UGT enzymes in the human population is variable. Expression levels of these enzymes in human livers were incorporated in the model using data reported in the literature (Naraharisetti *et al.*, 2010; Lin *et al.*, 2002; Sim *et al.*, 2005; Tay *et al.*, 2010, Zaya *et al.*, 2006). A Monte Carlo simulation for a virtual population of 1000 individuals was performed, in which in each simulation, the expression levels of the CYP and UGT enzymes were randomly taken from a log-normal distribution. The expression levels of the enzymes were

Table 2. Kinetic data on *in vitro* conversion of ATRA reported in literature used to determine PBK model parameter values for hepatic clearance.

Species	*In vitro* system	Cofactor	Product	Apparent K_m (μM)	Apparent V_{max}	Reference
rat	liver microsomes	NADPH	total metabolites	0.51	100[a]	Ahmad *et al.* (2000)
	kidney microsomes	NADPH	total metabolites	0.48	170[a]	Ahmad *et al.* (2000)
	lung microsomes	NADPH	total metabolites	0.50	1.0[a]	Ahmad *et al.* (2000)
	brain microsomes	NADPH	total metabolites	1.14	15[a]	Ahmad *et al.* (2000)
	liver microsomes	UDP-GlcUA	all-*trans* -retinoyl β-glucuronide	418	650[a]	Little *et al.* (1997)
human	CYP2C8	NADPH	4-hydroxy-ATRA	13.4	4.8[b]	Thatcher *et al.* (2010)
	CYP3A4	NADPH	4-hydroxy-ATRA	19.4	4.0[b]	Thatcher *et al.* (2010)
	CYP3A5	NADPH	4-hydroxy-ATRA	11.1	4.9[b]	Thatcher *et al.* (2010)
	CYP3A7	NADPH	4-hydroxy-ATRA	11.3	2.3[b]	Thatcher *et al.* (2010)
	CYP26A1	NADPH	4-hydroxy-ATRA	0.0094	11[b]	Lutz *et al.* (2009)
	UGT2B7	UDP-GlcUA	all-*trans* -retinoyl β-glucuronide	1.3	0.52[c]	Samokyszyn *et al.* (2000)

[a] pmol · min^{-1} · mg microsomal $protein^{-1}$
[b] pmol · min^{-1} · pmol CYP $enzyme^{-1}$
[c] pmol · min^{-1} · mg UGT2B7 $enzyme^{-1}$

assumed to vary independently. For this simulation, the mean μ_w and SD σ_w describing the log-normal distribution of each enzyme expression level value were derived from the data reported in the literature using the following equations:

$$\mu_w = \ln(\mu_x/\sqrt{(1+CV_x^2)})$$

and

$$\sigma_w = \ln(1+CV_x^2)$$

where μ_x is the mean and CV_x is the coefficient of variation of the non-log-transformed enzyme expression level values as reported in the literature (Zhang *et al.*, 2007; Punt *et al.*, 2010).

III) PBK model evaluation

Model predictions were evaluated by comparing model predicted ATRA blood concentrations with ATRA blood concentrations from *in vivo* kinetic studies reported in the literature (Table 3). A sensitivity analysis was performed to identify the key parameters highly influencing the model output (peak blood concentrations and blood AUC). The effect of a 5% increase in parameter value was evaluated by calculating normalized sensitivity coefficients (SC = (C' - C) / (P' - P) × (P / C)) as described by Evans and Andersen (2000) (C = initial value of model output, C' = value of model output resulting from 5% increase in parameter value, P = initial parameter value, P' = 5% increased parameter value). For the sensitivity analysis of the human model, the expression levels of the biotransformation enzymes were set at their median expression level.

IV) Translation of the in vitro concentration-response curve into predicted in vivo dose-response curves

To predict the dose-response curves for *in vivo* developmental toxicity of ATRA, *in vitro* data for the ATRA-induced inhibition of ES-D3 cell differentiation were used as input in the blood compartment of the PBK model and translated into *in vivo* oral dose levels. Because it is assumed that the free fraction of the chemical will cause ATRA's toxicity, corrections were made to take the differences in the free fraction in the *in vitro* compared to the *in vivo* situation into account. To this end, the free fractions for the different situations (culture medium, rat blood and human blood) were determined using the approach described in the section on

Table 3. Published kinetic studies used to evaluate PBK model-based predictions of ATRA blood levels over time. *In vivo* kinetic data are shown in Figures 4 and 5.

Species	Sex	Weight	Dose(s) (mg·kg bw^{-1})	Exposure route	Kinetic data shown in Figure	Reference
rat (Sprague-Dawley)	male	315 g	0.015, 0.25, 5	intravenous	4A	Swanson *et al.* (1981)
rat (Sprague-Dawley)	n.a.	360 g	2.8, 6.9, 14	intravenous	4B	Shelley *et al.* (1982)
rat (Wistar)	male	250-300 g	8	intravenous	4C	El Mansouri *et al.* (1995)
rat (Wistar)	female	n.a. (250 g assumed)	6	oral	4D	Collins *et al.* (1995)
rat (Sprague-Dawley)	n.a.	360 g	14	oral	4E	Shelley *et al.* (1982)
rat (Wistar)	male	250-300 g	20	intrajejunal	4F	El Mansouri *et al.* (1995)
human	n.a.	n.a. (70 kg assumed)	2	oral	5A	Adamson *et al.* (1993)
human	n.a.	n.a. (70 kg assumed)	4	oral	5B	Adamson *et al.* (1993)

n.a.: information not available

tissue:blood partition coefficients, after which corrections factors were calculated.

Predictions of dose-response curves for *in vivo* developmental toxicity were made by relating the toxic effect to 1) the C_{max} in blood, as was done before for the glycol ethers (Louisse *et al.*, 2010) and 2) the AUC in blood. To relate the toxic effect to the AUC in blood, the concentration-response data were translated into AUC-response data, by multiplying the applied concentration with the assay time (Daston *et al.*, 2010), assuming an unchanged concentration of the test chemicals in the exposure medium over time. The model was then used to calculate the corresponding dose levels of ATRA leading to these ATRA AUC values in the blood compartment, resulting in the predicted dose-response data for *in vivo* developmental toxicity. Subsequently, BMD modeling was applied using PROAST software to obtain the predicted *in vivo* dose-response curves for developmental toxicity.

V) Evaluation of the predictive value of the approach

To evaluate the potential of the *in vitro-in silico* approach to predict *in vivo* developmental toxicity of ATRA, the predicted dose-response curves for the rat were compared with data obtained in *in vivo* developmental toxicity studies reported in the literature as listed in Table 4. Developmental toxicity endpoints reported in these studies include several types of malformations. Furthermore, the *in vivo* data were used to derive BMD_{10} values on the basis of dose-response analyses using PROAST software. These BMD_{10} values derived from *in vivo* data were compared with predicted BMD_{10} values, which were acquired by performing dose-response modeling on the predicted dose-response data using PROAST software.

Results

I) In vitro concentration-response curve for ATRA-induced developmental toxicity

In vitro concentration-response data on ATRA-induced inhibition of ES-D3 cell differentiation were used as input data for the reverse dosimetry approach. The data reported for three independent ES-D3 cell differentiation studies (Louisse *et al.*, 2011) were combined to one dataset to which concentration-response modeling

Table 4. *In vivo* developmental toxicity studies used to evaluate predicted dose-response curves for ATRA-induced developmental toxicity in rats. In all studies, rats were exposed using gastric intubation at GD10. Symbols shown correspond with symbols used in Figure 7. BMD_{10} values calculated from these data are shown in Figure 9.

Rat strain	Doses applied ($mg \cdot kg\ bw^{-1}$)	Critical endpoint	Figure	Symbol	Reference
Sprague-Dawley	0, 3, 6, 12	malformations head	7A	■	Wise *et al.* (2010)
		malformations torso and limb	7A	▲	
		extra vertebra	7A	●	
Wistar	0, 20, 30, 50, 80, 100	cleft palate	7B	■	Tembe *et al.* (1996)
		microtia	7B	▲	
Wistar (RORO)	0, 6, 12, 18	external malformations cranium	7C	■	Bürgin and Schmitt (2003)
		visceral malformations palate/mouth	7C	▲	
		skeletal malformations skull	7C	●	

was applied using benchmark dose (BMD) analysis. Figure 2 shows the *in vitro* data and the 95% lower and upper confidence limits of the concentration-response curve fitted to these data, which were subsequently used as input (as internal concentrations) in the PBK model to predict *in vivo* dose-response curves (step IV).

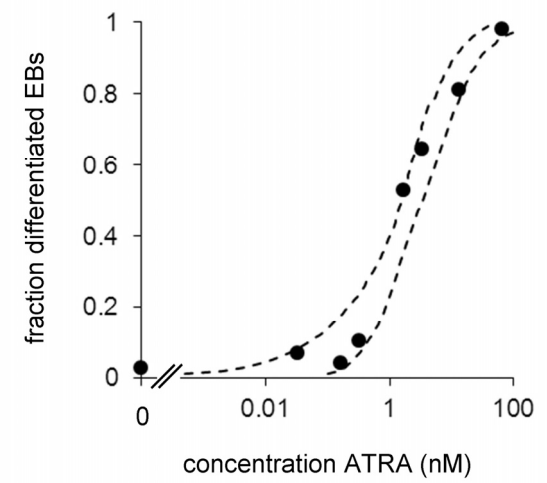

Figure 2. Concentration-response data of ATRA-induced inhibition of ES-D3 cell differentiation based on data from Louisse *et al.* (2011). The circles represent the combined data obtained in three independent ES-D3 cell differentiation assays. The dashed lines represent the 95% lower (left curve) and the 95% upper (right curve) confidence limit of the curve fitted to the data.

II) Development of a PBK model for ATRA kinetics in rat and human

To translate the *in vitro* concentration-response curve into *in vivo* dose-response curves, a PBK model describing the *in vivo* kinetics of ATRA is required. The conceptual basis for this model is shown in Figure 1, which was converted into mathematical equations. Subsequently, the parameter values determining the ADME processes, such as values for physiological and anatomical parameters, tissue:blood partition coefficients and other kinetic parameters were defined. The physiological and anatomical parameter values used for the model were obtained from the literature (Table 1). The kinetic parameter values were obtained using only *in vitro* techniques and include parameter values for 1) intestinal absorption, 2) partitioning into tissues (i.e. distribution) and 3) metabolic clearance (i.e. metabolism and subsequent excretion).

Parameter value for intestinal absorption. PBK model parameter values for intestinal absorption were estimated by studying the transfer rate of ATRA across CaCo-2 cells cultured in a transwell system. From these studies, apparent permeability (P_{app}) coefficients are obtained, representing the transport rate as a function of the intestinal surface area and the concentration applied (expressed in $cm \cdot s^{-1}$). To assess the effect of typical *in vivo* conditions, ATRA's transport characteristics were studied in the *in vitro* system using different medium conditions in the apical compartment (representing the intestinal cavity) and the basolateral compartment (representing portal blood). The presence of the bile acid taurocholate in the apical compartment decreases the P_{app} coefficient, whereas the presence of albumin in the basolateral compartment increases the P_{app} coefficient (Figure 3). The P_{app} coefficient for ATRA is not dependent on the concentration applied in the apical compartment (Figure 3), indicating that this value can be used to calculate the intestinal absorption ($amount \cdot time^{-1}$) for all concentrations applied. The average P_{app} coefficient ($20 \cdot 10^{-6}$ $cm \cdot s^{-1}$ equivalent to $5.6 \cdot 10^{-9}$ $cm \cdot h^{-1}$) obtained using physiological conditions (taurocholate in the apical compartment and albumin in the basolateral compartment) was converted to obtain the PBK model parameter value for intestinal absorption (in $ml \cdot h^{-1}$). This parameter value was obtained by multiplying the converted P_{app} coefficient by the surface area of the intestinal lumen (see materials and methods for calculations).

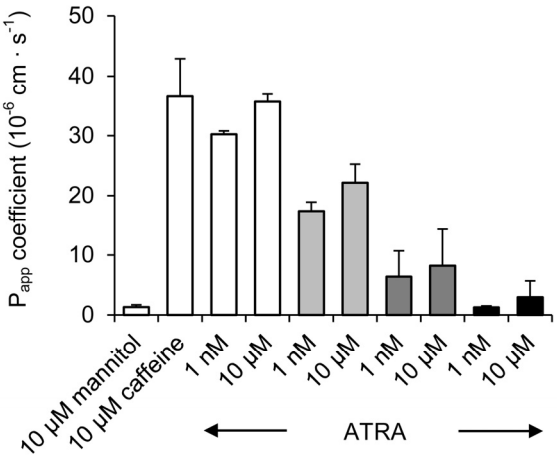

Figure 3. Apparent permeability (P_{app}) coefficients for mannitol, caffeine and ATRA transport across monolayers of CaCo-2 cells cultured in a transwell system. Mannitol and caffeine were used as controls for paracellular and transcellular transport, respectively. White bars: apical: HBSS, basolateral: 0.6 mM BSA in HBSS; light grey bars: apical: 10 mM taurocholate in HBSS, basolateral: 0.6 mM BSA in HBSS; dark grey bars: apical: HBSS, basolateral: HBSS; black bars: apical: 10 mM taurocholate in HBSS, basolateral: HBSS.

Parameter values for tissue:blood partition coefficients. To predict the distribution of ATRA over different tissues in the body, tissue:blood partition coefficients needed to be determined. These were obtained by determining tissue:buffer partition coefficients and blood:buffer partition coefficients using a recently developed SPME-based method (Vaes, 2010). Tissue:blood partition coefficients were obtained by dividing the tissue:buffer partition coefficients (Table 5) by the blood:buffer partition coefficient (measured to be 325). Table 5 shows the obtained tissue:buffer partition coefficients (with olive oil representing fat tissue) and the calculated tissue:blood partition coefficients. The rest-of-body:blood partition coefficient in the model was set at 1. The tissue:blood partition coefficients determined for rat were also used for the human PBK model.

Parameter values for metabolic clearance. The excretion of ATRA was assumed to result from metabolic clearance only. This was based on previous observations that the amount of reported unchanged ATRA excreted in urine, bile and faeces is negligible (Swanson *et al.*, 1981; Hänni *et al.*, 1976; Li *et al.*, 1996). To obtain model

Table 5. Measured rat tissue:buffer partition coefficients and calculated rat tissue:blood partition coefficients.

Tissue	Tissue:buffer partition coefficient	Tissue:blood partition coefficient[a]
fat (olive oil)	2200	6.8
lung	1280	3.9
brain	4460	14
heart	640	2.0
kidney	1810	5.6
muscle	1120	3.5
liver	1160	3.6

[a] determined by dividing the tissue:buffer partition coefficient by the blood:buffer partition coefficient (i.e. 325).

parameter values for hepatic clearance, data on kinetics for the *in vitro* conversion of ATRA by liver microsomes reported in the literature were used (Table 2). For rats, kinetic constants (apparent V_{max} and K_m values) obtained *in vitro* using liver microsomes were available from the literature (Table 2). The apparent V_{max} values of the reactions were scaled to a rat liver assuming a microsomal protein yield of 35 mg/g rat liver (Medinsky *et al.*, 1994), yielding model parameter values for hepatic clearance. For rats also kinetic constants obtained using kidney, lung and brain microsomes were available from the literature (Table 2). These data were incorporated in the model assuming a microsomal protein yield of 7 mg/g rat kidney (Beierschmitt and Weiner, 1986), 20 mg/g rat lung (Medinsky *et al.*, 1994) and 4 mg/g rat brain (Ravindranath and Anandatheerthavarada, 1990), respectively.

For humans, kinetic constants obtained *in vitro* using specific biotransformation enzymes were available from the literature (Table 2). The data on apparent V_{max} values of the reactions were scaled to a human liver assuming a liver microsomal protein yield of 32 mg/g human liver (Barter *et al.*, 2007). The expression levels of the biotransformation enzymes in human liver microsomes have been reported in the literature for different individuals (Naraharisetti *et al.*, 2010; Lin *et al.*, 2002; Sim *et al.*, 2005; Tay *et al.*, 2010; Zaya *et al.*, 2006). These data were used to determine the distribution of the expression levels of these enzymes in human

livers. When assuming that these data would be representative for the human population, predictions for a virtual human population can be made, using Monte Carlo simulations. These simulations were used in step III to assess interindividual differences in ATRA kinetics in the human population.

III) PBK model evaluation

To evaluate the PBK model predictions of ATRA kinetics, PBK model predictions for ATRA blood concentrations in time were compared with ATRA blood concentrations from reported *in vivo* kinetic studies (Figures 4 and 5). Table 3 gives an overview of the *in vivo* kinetic studies available for evaluation of the predictions

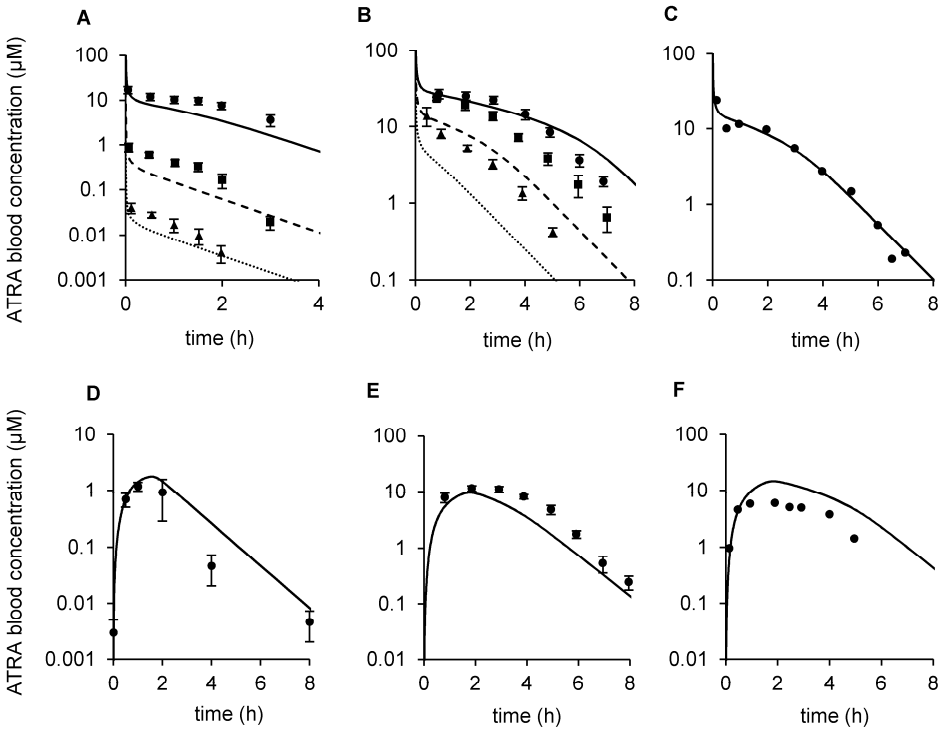

Figure 4. ATRA blood concentrations in rats after intravenous (A, B, C) or oral (D, E, F) dosing. Symbols represent average blood levels (SDs are depicted by error bars) from *in vivo* studies reported in the literature (for references for the respective curves A-F see Table 3). Lines represent model predicted blood concentrations. ATRA dose levels for the different curves were as follows: A) 0.015 (triangles, ·····), 0.25 (squares, – – –) or 5 (circles, —) mg · kg bw^{-1}. B) 2.8 (triangles, ·····), 6.9 (squares, – – –) or 14 (circles, —) mg · kg bw^{-1}. C) 8 mg · kg bw^{-1}, D) 6 mg · kg bw^{-1}, E) 14 mg · kg bw^{-1} F) 20 mg · kg bw^{-1}.

made by the in this study developed PBK models for rat and human including their exposure regimens. Comparison of the reported *in vivo* data for rat with predictions made by the rat PBK model reveals that the PBK model reliably predicts ATRA blood concentration-time curves in rat (Figure 4) after intravenous (Figures 4A-C) and oral (Figures 4D-F) exposure to diverse doses of ATRA.

For the evaluation of the human PBK model, a kinetic study was used in which data on blood concentrations for six individuals were presented (Adamson *et al.*, 1993). The data from these six individuals reflect the large interindividual variation in ATRA blood levels in humans, amounting to more than two orders of magnitude at certain time points (Figure 5). This figure also shows the blood concentrations predicted by the PBK model for the 1st, the 5th, the 50th, the 95th and the 99th percentile of a virtual population of 1000 people, using a Monte Carlo simulation. Most data points on the blood concentrations of the 6 individuals that were dosed with 2 mg · kg bw^{-1} fall within the 5th and the 95th percentile of the blood concentrations predicted for the virtual population (Figure 5A). When the human individuals were dosed with 4 mg · kg bw^{-1}, most data points on the blood concentrations fall within the 1st and the 95th percentile of the blood concentrations modeled for the virtual population (Figure 5B).

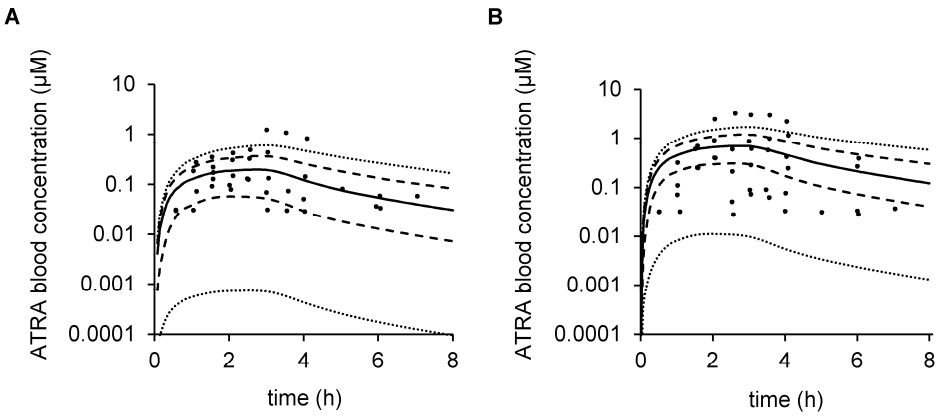

Figure 5. ATRA blood concentrations in human after oral dosing. Symbols represent blood concentrations from individuals taken from an *in vivo* kinetic study reported in the literature (Adamson *et al.*, 1993). Straight lines represent the median of the model predicted blood concentrations, whereas dashed lines represent the 5th and the 95th percentiles and dotted lines the 1st and 99th percentiles of the blood concentrations predicted for a virtual population, using a Monte Carlo simulation. ATRA dose levels for the different curves were as follows: A) 2 mg · kg bw^{-1}, B) 4 mg · kg bw^{-1}.

To further evaluate the PBK models a sensitivity analysis was performed to identify the key parameters that influence the model output to the largest extent. This analysis was carried out for the rat and human model, using doses of 0.1, 1, 10 and 100 mg ATRA · kg bw^{-1}. This analysis reveals that predicted ATRA blood concentrations (C_{max} and AUC) are most dependent on the parameter value for intestinal absorption, the muscle:blood partition coefficient, and the parameters for hepatic clearance in both rat and human (Figure 6). The analysis also revealed

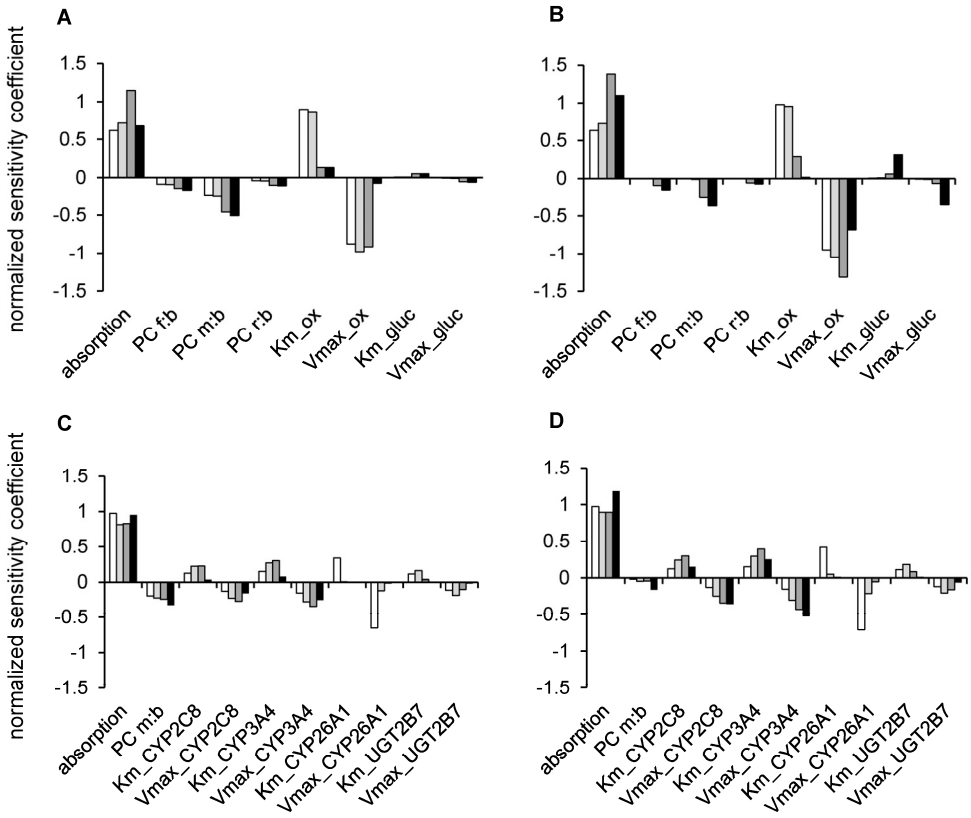

Figure 6. Normalized sensitivity coefficients (NSCs) for kinetic model parameters for rat (A, B) and human (C, D) PBK models for ATRA blood C_{max} values (A, C) and AUC values (B, D) at different oral doses. White bars: 0.1 mg · kg bw^{-1}; light grey bars: 1 mg · kg bw^{-1}; dark grey bars: 10 mg · kg bw^{-1}; black bars: 100 mg · kg bw^{-1}. NSCs are shown for model parameters for which at least at one of the dose metrics (C_{max}, AUC) at one of the doses the NSC > |0.1|. Absorption: intestinal absorption, PC f:b: partition coefficient fat:blood, PC m:b: partition coefficient muscle:blood, PC r:b: partition coefficient rest of body:blood, Km_ox: K_m oxidation, Vmax_ox: V_{max} oxidation, Km_gluc: K_m glucuronidation, Vmax_gluc: V_{max} glucuronidation, Km_X: K_m enzyme X, Vmax_X: V_{max} enzyme X.

that the output of the rat PBK model is not sensitive to the parameters for metabolism in kidney, lung and brain, indicating a negligible role of extrahepatic clearance in total ATRA clearance.

IV) Translation of the in vitro concentration-response curve into predicted in vivo dose-response curves, and V) the evaluation of the predictive value of the approach

After obtaining the *in vitro* concentration-response curve and developing the PBK model needed for reverse dosimetry, the concentrations of the *in vitro* concentration-response curve are used as input in the blood compartment of the PBK model and translated into *in vivo* dose levels, defining predicted *in vivo* dose-response data for ATRA-induced developmental toxicity. Applying dose-response modeling on the obtained predicted dose-response data results in the predicted *in vivo* dose-response curve. In order to make these translations, the following aspects need prior consideration: 1) the possible correction needed for the difference in the free fraction of ATRA in the *in vitro* compared with the *in vivo* situation and 2) the choice of the dose metric used (maximal chemical concentration achieved (i.e. the C_{max}) or the area under the chemical concentration time curve (i.e. the AUC)) to correlate exposure to toxicity.

ATRA has a high binding affinity for proteins and lipids (Smith *et al.*, 1973), which leads to differences in the free fraction of ATRA in the *in vitro* (culture medium) compared to the *in vivo* (blood) situation (Gülden *et al.*, 2006; Blaauboer, 2010). Because it is assumed that the toxic effect is caused by the free fraction of ATRA, corrections need to be made for differences in the free fraction in the *in vitro* situation (culture medium) compared to the *in vivo* situation (rat or human blood). This was achieved by multiplying the *in vitro* concentration-response curve (Figure 2) with a correction factor, which was obtained by dividing the free fraction in culture medium by the free fraction in blood, which were determined using the SPME-based method. The correction factors amounted to 3.5 for rat blood and 6.0 for human blood, reflecting that the protein-bound fraction *in vivo* is higher than that in culture medium.

Generally, it is assumed that the (toxic) effect of a chemical can be best related to either the C_{max} or the AUC. For the prediction of dose-response curves for ATRA -induced developmental toxicity, the predictions for rat were based on the C_{max} or

on the AUC reached in the blood compartment of the PBK model. To calculate *in vitro* AUC values needed for the extrapolations based on relating the toxicity to the AUC, the concentration-response curve (Figure 2) was multiplied by the exposure duration (i.e. 10 days). *In vitro* concentrations were translated into *in vivo* doses using the rat PBK model, after which BMD modeling was applied to obtain the predicted *in vivo* dose-response curves for developmental toxicity. Figure 7 presents the 95% lower and upper confidence limits of these dose-response curves predicted based on either the C_{max} or the AUC. This figure also includes data points reported for *in vivo* developmental toxicity of ATRA in rats. Comparison of the reported experimental data with the two sets of predicted dose-response curves reveals that predictions based on relating the toxic effect to the AUC show a better concordance with data from reported *in vivo* developmental toxicity studies in rat than predictions based on relating the toxic effect to the C_{max} (Figure 7).

Based on the evaluation of the predictions for rats, predictions for ATRA-induced developmental toxicity in humans were made by relating the

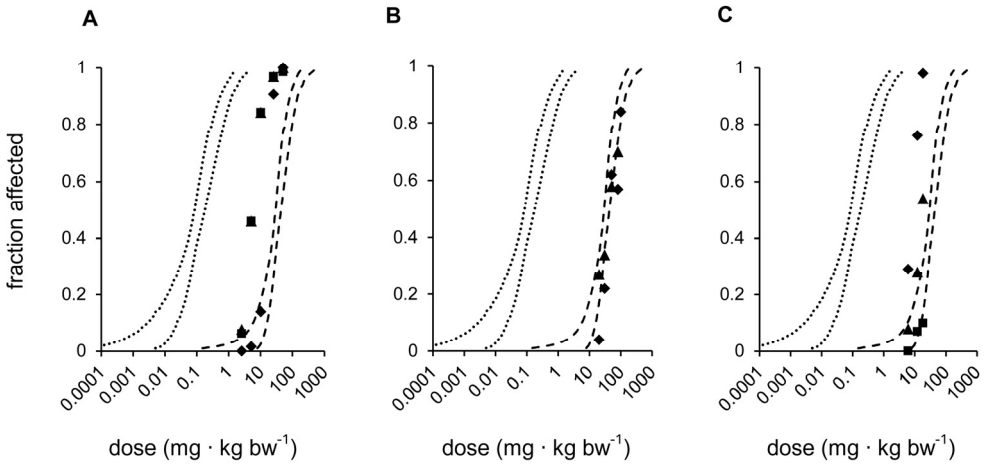

Figure 7. *In vivo* developmental toxicity of ATRA in rats. Symbols represent data from *in vivo* studies reported in the literature (A: Wise *et al.*, 2010, B: Tembe *et al.*, 1995, C: Bürgin and Schmitt, 2003). Endpoints represented by the symbols used are listed in Table 4. Lines represent the predicted dose-response curve by extrapolating the *in vitro* data (Figure 2) to *in vivo* dose levels using the rat PBK model. Subsequently, BMD modeling was applied to obtain the predicted *in vivo* dose-response curves for developmental toxicity. Both the 95% lower and upper confidence limits of the predicted dose-response curves are presented. Dotted lines represent the predictions made by relating the developmental toxicity effect to the C_{max}, whereas dashed lines represent the predictions made by relating the developmental toxicity effect to the AUC.

developmental toxicity effect to the AUC. Dose-response curves for ATRA-induced developmental toxicity in humans were predicted for a virtual population of 1000 individuals, using a Monte Carlo simulation. The predicted dose-response curve for the 5th, the 50th and the 95th percentile of this virtual population is shown in Figure 8, indicating a large predicted interindividual variation in sensitivity to this chemical in the human population.

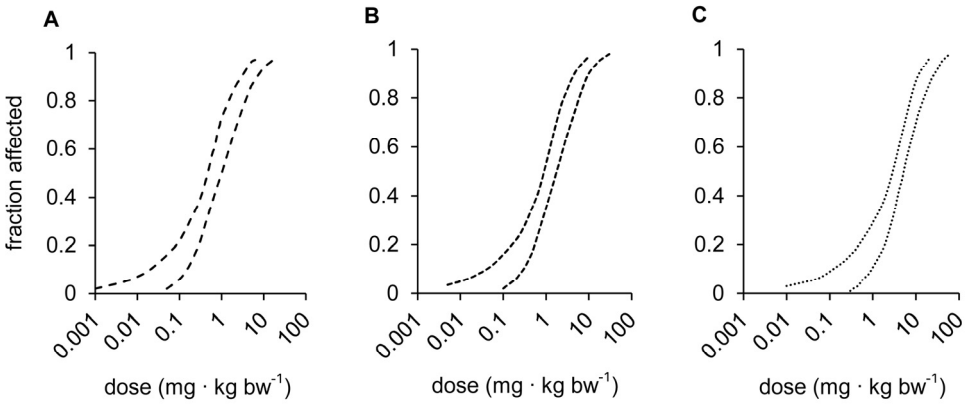

Figure 8. Predicted *in vivo* developmental toxicity of ATRA in human. Dose-response curves for ATRA-induced developmental toxicity in humans were predicted by extrapolating the *in vitro* data (Figure 2) to *in vivo* dose levels using the human PBK model. Subsequently, BMD modeling was applied to obtain the predicted *in vivo* dose-response curves for developmental toxicity. Both the 95% lower and upper confidence limits of the predicted dose-response curves are presented. Predictions were made for a virtual population of 1000 individuals, using a Monte Carlo simulation. The predicted dose-response curves for the 5th (A), the 50th (B) and the 95th (C) percentile of this virtual population are shown. Predictions are made by relating the developmental toxicity effect to the AUC.

Finally, the reverse dosimetry approach was used to predict BMD_{10} values (benchmark doses at which a benchmark response equivalent to a 10% effect size (BMR_{10}) is reached). The lower limit of the 95% confidence interval on the BMD_{10} ($BMDL_{10}$) can be used as point of departure in human risk assessment practices (Barlow *et al.*, 2009). BMD_{10} values were predicted by extrapolating the concentrations of the *in vitro* concentration-response curve (Figure 2) to *in vivo* dose levels (obtained by relating the toxic effects to the AUC) after which the resulting dose-response data were used for BMD analysis. The predicted BMD_{10} value for ATRA-induced developmental toxicity in rat thus obtained was compared with BMD_{10} values determined from data from reported *in vivo* developmental toxicity studies (Table 4), which were shown to be within the same order of

magnitude (Figure 9). The predicted $BMDL_{10}$ value for ATRA-induced developmental toxicity in rats amounted to 6 mg · kg bw^{-1}. For humans, $BMDL_{10}$ values were predicted in the same way as for rat, for a virtual population of 1000 individuals, using a Monte Carlo simulation. The $BMDL_{10}$ values for ATRA-induced developmental toxicity for the 5^{th}, 50^{th} and 95^{th} percentile of this virtual population amounted to 0.02, 0.05 and 0.1 mg · kg bw^{-1}, indicating a higher sensitivity of humans than of rats to ATRA-induced developmental toxicity.

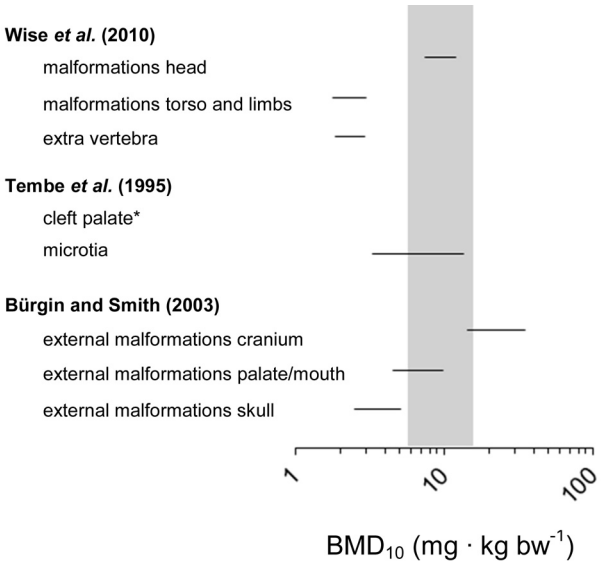

Figure 9. BMD_{10} values in rats for ATRA-induced developmental toxicity. BMD_{10} values were determined for data from *in vivo* studies reported in the literature as shown in Table 4 and Figure 7. The lines cover the BMD_{10} values for different ATRA-induced malformations ranging from the $BMDL_{10}$ (lower limit of the 95% confidence interval on the BMD_{10}) to the $BMDU_{10}$ (upper limit of the 95% confidence interval on the BMD_{10}). The grey area represents the predicted BMD_{10} (ranging from the $BMDL_{10}$ to the $BMDU_{10}$).

* BMD_{10} value could not be determined for this endpoint.

Discussion

The aim of the present study was to assess the feasibility of predicting *in vivo* developmental toxicity dose levels of the retinoid all-*trans*-retinoic acid (ATRA) with reverse dosimetry, using a PBK model that is based on kinetic parameter values derived using *in vitro* techniques only. To this end, *in vitro* concentration-

response curves of ATRA-induced inhibition of ES-D3 cell differentiation were translated into *in vivo* dose-response curves using PBK modeling with a reverse dosimetry approach. The results obtained reveal that the predicted BMD_{10} value for ATRA-induced developmental toxicity in rats is within one order of magnitude compared with BMD_{10} values for ATRA-induced developmental toxicity derived from reported *in vivo* developmental toxicity studies from the literature. To the best of our knowledge, this provides the first proof-of-principle indicating that toxic dose levels for this complex systemic endpoint can be accurately predicted based on *in vitro* kinetic and toxicity data only. Therefore, it indicates the feasibility of the use of this combined *in vitro-in silico* approach in toxicological risk assessment strategies.

For the extrapolation of the *in vitro* concentration-response curve of ATRA in the ES-D3 cell differentiation assay to *in vivo* dose-response curves, a PBK model describing the *in vivo* kinetics of ATRA needed to be developed. The PBK model parameter values for kinetic processes (intestinal uptake, tissue partitioning and metabolic clearance) are often obtained using *in vivo* kinetic data, but were in the present study solely derived using *in vitro* techniques. ATRA clearance was described by metabolism in the liver compartment and not via urinary and biliary excretion. This was based on previous observations that the amount of reported unchanged ATRA excreted in urine, bile and faeces is negligible (Swanson *et al.*, 1981; Hänni *et al.*, 1976; Li *et al.*, 1996). Predictions for ATRA concentrations in rat blood were in general in good concordance with the ATRA blood concentrations reported in the literature upon intravenous and oral intake (Figures 3 and 4). Deviations are in all cases less than one order of magnitude and no consistency in under- or overestimation is noticed, which may reflect variation between the *in vivo* kinetic studies. This indicates that the PBK model developed using solely *in vitro*-derived kinetic parameter values can be reliably used for reverse dosimetry.

Moreover, the human PBK model reliably predicts ATRA blood concentrations in time upon oral intake. Reported blood concentrations in humans are highly variable between individuals (Adamson *et al.*, 1993; Figure 6). These interindividual differences were also predicted by the PBK model, using a Monte Carlo simulation for a virtual population of 1000 individuals with varying expression levels of biotransformation enzymes in the liver as reported in the literature (Naraharisetti

et al., 2010; Lin *et al.*, 2002; Sim *et al.*, 2005; Tay *et al.*, 2010, Zaya *et al.*, 2006). Most reported blood concentrations in humans fall within the 1st and the 99th percentile of the PBK-model based predicted blood concentrations for the simulated population (Figure 6). The fact that certain *in vivo* data points fall outside the 99th percentile of the PBK-model based predicted blood concentrations for the simulated population indicates that the human PBK model may slightly underestimate blood concentrations. Considering the results of the sensitivity analysis, this may be due to the overestimation of partitioning into muscle tissue, an overestimation of hepatic clearance, or an underestimation of intestinal absorption, because of the high normalized sensitivity coefficients obtained for these model parameters (Figures 5C and 5D).

Since it was concluded that both the rat and human PBK model reliably predict ATRA blood concentrations in time, the PBK models were used for reverse dosimetry, translating the *in vitro* concentration-response curve into *in vivo* dose-response curves for ATRA-induced developmental toxicity. The obtained predicted dose-response curve for ATRA-induced developmental toxicity in rats showed that predicted developmental toxicity dose levels of ATRA were closer to the *in vivo* rat data when predictions were based on relating the toxic effect to the AUC, than when predictions were based on relating the toxic effect to the C_{max}. This is in line with the study reported by Tzimas *et al.* (1997), who found a better correlation when correlating various developmental toxicity endpoints to the AUC (correlation coefficients of 0.90-0.98) than when correlating the developmental toxicity endpoints to the C_{max} (correlation coefficients of 0.22-0.43). To predict *in vivo* developmental toxicity by relating the developmental toxicity effects to the AUC, *in vitro* AUC values were obtained by multiplying the applied concentration in the ES -D3 cell differentiation assay with the assay time, as suggested by Daston *et al.* (2010). By doing so, it is assumed that the concentration of the test chemical in the exposure medium remains constant over time and that the sensitivity of the differentiating stem cells during the differentiation period is constant as well. The influence of possible differences in sensitive periods on the extrapolation of *in vitro* developmental toxicity data to the *in vivo* situation needs further research, but is beyond the scope of the present study. The fact, however, that deviations between predicted and observed dose-response curves were limited, suggests that the approach taken provides a reasonable way of estimating the AUC.

Since predicted dose-response curves for rats were best related to the dose metric AUC, dose-response curves for ATRA-induced developmental toxicity in humans were based on relating the toxic effect to the AUC as well. Because the human PBK model developed allowed the prediction of the *in vivo* kinetics for a virtual population using Monte Carlo simulations, dose-response curves for ATRA-induced developmental toxicity could be predicted for this virtual population as well. This resulted in dose-response curves for different individuals, ranging from relative sensitive to relative insensitive individuals. Presenting different percentiles (e.g. the 5^{th}, the 50^{th} and the 95^{th}) of the dose-response curves for this virtual population reveals the distribution of sensitivity to ATRA-induced developmental toxicity, based on interindividual difference in kinetics (Figure 8).

To evaluate whether the presented reverse dosimetry approach can be used to determine points of departure for risk assessment, BMD_{10} values were predicted by extrapolating the concentrations of the *in vitro* concentration-response curve to *in vivo* dose levels after which the resulting dose-response data were used for BMD analysis. The predicted BMD_{10} value for rat thus obtained was shown to be in concordance with BMD_{10} values determined from reported *in vivo* developmental toxicity studies described in the literature, indicated by the less than one order of magnitude difference between the predicted and the *in vivo* BMD_{10} values (Figure 9). Considering the experimental differences in BMD_{10} values derived from the different *in vivo* studies, the differences between our predicted BMD_{10} value and the BMD_{10} values derived from *in vivo* studies are small (Figure 9). This indicates that the reverse dosimetry approach could be used in the risk assessment process to set safe exposure limits by applying uncertainty factors for interspecies and intraspecies differences in kinetics and dynamics. No such factors need to be applied when predictions are made using the developed human PBK model, because these result in BMD_{10} values for humans. The predicted $BMDL_{10}$ for the median of the virtual human population amounted to 0.05 mg · kg bw^{-1}, which is 120-fold lower than the $BMDL_{10}$ predicted for rats (i.e. 6 mg · kg bw^{-1}). This indicates that the default uncertainty factor for interspecies differences in kinetics (amounting to √10) for the extrapolation of rats to humans would be too low for ATRA. To protect also the relative sensitive individuals in the human population, the $BMDL_{10}$ for the 5^{th} percentile of the population (i.e. 0.02 mg · kg bw^{-1}) might be used to set safe exposure limits. The $BMDL_{10}$ for the 5^{th} percentile of the population

is 2.5-fold lower than the $BMDL_{10}$ for the 50^{th} percentile of the population. This indicates that the default uncertainty factor for intraspecies differences in kinetics (amounting to √10) for the extrapolation of the 'average' human to relatively sensitive individuals would be sufficient. The obtained results indicate that using a reverse dosimetry approach with the human ATRA PBK model, might provide a better prediction of *in vivo* ATRA-induced developmental toxicity in humans, than the classical approach translating toxicity data obtained in rats to humans using default uncertainty factors for inter- and intraspecies differences.

Although the approach presented is promising to be used in the toxicological risk assessment of chemicals, there are some limitations. Predictions were made for single exposures in the present study. Making predictions for repeated exposures to ATRA is complex, because ATRA plasma levels decrease upon repeated dosing (up to 9-fold in rats) (Collins *et al.*, 1995). This is probably due to increased expression levels of biotransformation enzymes (Collins *et al.*, 1995; Tay *et al.*, 2010). In order to make predictions for ATRA-induced developmental toxicity after repeated dosing, the change in expression levels of biotransformation enzymes should be incorporated in the PBK model, to account for these changes in clearance. This phenomenon needs further consideration in future reverse dosimetry work. Another aspect to consider in future reverse dosimetry approaches for developmental toxicity is the placental transfer of chemicals. In the present study, predictions of dose-response curves were based on the *in vitro* concentration-response data as input in the blood compartment of the PBK model, as was done before for the glycol ethers (Louisse *et al.* 2010). No specific compartment for the embryo was incorporated in the PBK model, because it has been reported that the embryo AUC of ATRA only slightly differs from the maternal plasma AUC of ATRA upon single dosing (Collins *et al.*, 1995). For the natural retinoid 13-*cis*-retinoic acid and the synthetic retinoid CD394 however, the embryonic concentrations are reported to be 10-fold (Nau, 2001) and 6-fold (Sass *et al.*, 1995), respectively, lower than the maternal plasma concentrations. Therefore, the assumption of equal maternal and embryo levels cannot be made for all chemicals, indicating that *in vitro* placental transport models are needed in order to predict the placental transfer of these chemicals to the fetus.

To the best of our knowledge, the present study is the first proof-of-principle showing that *in vivo* dose-response curves for developmental toxicity can be accurately predicted without the need of *in vivo* studies. It shows the feasibility of the reverse dosimetry approach, integrating *in vitro* and *in silico* techniques, to predict *in vivo* developmental toxicity dose levels. This is indicated by the less than one order of magnitude difference between the predicted BMD_{10} value for ATRA-induced developmental toxicity and the BMD_{10} values determined from reported *in vivo* developmental toxicity studies described in the literature. This difference is small considering the experimental differences in BMD_{10} values derived from the *in vivo* developmental toxicity studies (Figure 9). Furthermore, the present study uncovered the species differences in ATRA kinetics, resulting in the prediction of a 120-fold higher sensitivity of humans to ATRA-induced developmental toxicity than of rats. This indicates that the default uncertainty factor for interspecies differences in kinetics to determine safe exposure levels for humans (amounting to $\sqrt{10}$) would be too low for ATRA. Therefore, reverse dosimetry using human PBK models may be the best tool to predict human toxicity dose levels to be used to set safe exposure levels for humans. However, since the assay used to define the *in vitro* concentration-response curves in the present study is based on mouse embryonic stem cells, possible interspecies differences in dynamics might be overlooked. Therefore, an uncertainty factor for these differences in dynamics would still be needed. If the *in vitro* model used to determine the *in vitro* concentration-response curves would be based on human cells instead of mouse cells, no uncertainty factor for interspecies differences in dynamics would be needed, leaving only the need for uncertainty factors accounting for human interindividual differences in dynamics. Altogether, it is concluded that the combined *in vitro-in silico* approach has the potential to be used in toxicological risk assessment strategies, thereby contributing to the 3Rs of animal use.

Acknowledgments

The authors would like to thank Wouter Vaes, Jaap Jan Stevenhagen and Sieto Bosgra for their help with the binding studies, Ans Punt for her help with the Monte Carlo analyses and Harvey Clewell for the discussions about retinoic acid kinetics.

References

Adamson, P. C., Pitot, H. C., Balis, F. M., Rubin, J., Murphy, R. F., Poplack, D. G. (1993). Variability in the oral bioavailability of all-trans-retinoic acid. *J Natl Cancer Inst* **85**, 993-6.

Ahmad, M., Ahmadi, M., Nicholls, P. J., Smith, H. J. (2002). In-vitro metabolism of retinoic acid by different tissues from male rats. *J Pharm Pharmacol* **52**, 511-5.

Barlow, S., Chesson, A., Collins, J. D., Flynn, A., Hardy, A., Jany, K., Knaap, A., Kuiper, H., Larsen, J., Lovell, D., Le Neindre, P., Schans, J., Schlatter, J., Silano, V., Skerfving, S., Vannier, P. (2009). Guidance of the Scientific Committee on a request from EFSA on the use of the benchmark dose approach in risk assessment. *The EFSA Journal* **1150**, 1-72.

Barter, Z. E., Bayliss, M. K., Beaune, P. H., Boobis, A. R., Carlile, D. J., Edwards, R. J., Houston, J. B., Lake, B. G., Lipscomb, J. C., Pelkonen, O. R., Tucker, G. T., Rostami-Hodjegan. A. (2007). Scaling factors for the extrapolation of in vivo metabolic drug clearance from in vitro data: reaching a consensus on values of human microsomal protein and hepatocellularity per gram of liver. *Curr Drug Metab* **8**, 33-45.

Blaauboer, B. J. (2010). Biokinetic modeling and in vitro-in vivo extrapolations. *J Toxicol Environ Health B Crit Rev* **3**, 242-52.

Beierschmitt, W. P., Weiner, M. (1986). Age-related changes in renal metabolism of acetaminophen in male fisher 344 rats. *Age* 9, 7-13.

Berezhkovskiy, L. M. (2004). Volume of distribution at steady state for a linear pharmacokinetic system with peripheral elimination. *J Pharm Sci* **93**, 1628-1640.

Bürgin, H., Schmitt, G. (2003). Comparison of the embryo-fetal toxicity of all-trans retinoic acid in two strains of Wistar rat.

Brown, R. P., Delp, M. D., Lindstedt, S. L., Rhomberg, L. R., Beliles, R. P. (1997). Physiological parameter values for physiologically based pharmacokinetic models. *Toxicol Ind Health* **13**, 407-84.

Collins, M. D., Mao, G. E. (1999). Teratology of retinoids. *Annu. Rev. Pharmacol. Toxicol.* **39**, 399-430.

Collins, M. D., Tzimas, G., Bürgin, H., Hummler, H., Nau, H. (1995). Single versus multiple dose administration of all-trans-retinoic acid during organogenesis: differential metabolism and transplacental kinetics in rat and rabbit. *Toxicol Appl Pharmacol* **130**, 9-18.

Daston, G. P., Chapin, R. E., Scialli, A. R., Piersma, A. H., Carney, E. W., Rogers, J. M., Friedman. J. M. (2010). A different approach to validating screening assays for developmental toxicity. *Birth Defects Res B Dev Reprod Toxicol* **89**, 526-30.

Davies, B., Morris, T. (1993). Physiological parameters in laboratory animals and humans. *Pharm Res* **10**, 1093-5.

DeJongh, J., Nordin-Andersson, M., Ploeger, B. A., Forsby A. (1999). Estimation of systemic toxicity of acrylamide by integration of in vitro toxicity data with kinetic simulations. *Toxicol Appl Pharmacol* **158**, 261-68.

El Mansouri, S., Tod, M., Leclerq, M., Petitjean, O., Perret, G., Porthault, M. (1995). Time- and dose-dependent kinetics of all-trans-retinoic acid in rats after oral or intravenous administration(s). *Drug Metab Dispos* **23**, 227-31.

Evans, M. V., Andersen, M. E. (2000). Sensitivity analysis of a physiological model for 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD): assessing the impact of specific model parameters on sequestration in liver and fat in the rat. *Toxicol Sci* **54**, 71-80.

Forsby, A., Blaauboer, B. J. (2007). Integration of in vitro neurotoxicity data with biokinetic modelling for the estimation of in vivo neurotoxicity. *Hum Exp Toxicol* **26**, 333-38.

Gülden, M., Dierickx, P., Seibert, H. (2009). Validation of a prediction model for estimating serum concentrations of chemicals which are equivalent to toxic concentrations in vitro. *Toxicol In vitro* **20**, 1114-24.

Hänni, R., Bigler, F., Meister, W., Englert, G. (1976). Isolation and identification of three urinary metabolites of retinoic acid in the rat. *Helv Chim Acta* **59**, 2221-7.

Hosseinpour, M., Behdad, A. (2008). Evaluation of small bowel measurement in alive patients. *Surg Radio Anat* **30**, 653-5.

Kararli T. T. (1995). Comparison of the gastrointestinal anatomy, physiology, and biochemistry of humans and commonly used laboratory animals. *Biopharm Drug Dispos* **16**, 351-380.

Lee, H. B., Blaufox, M. D. (1985). Blood volume in the rat. *J Nucl Med* **26**, 72-6.

Li, S., Barua, A. B., Huselton, C. A. (1996). Quantification of retinoyl-beta-glucuronides in rat urine by reversed-phase high-performance liquid chromatography with ultraviolet detection. *J Chromatogr B Biomed Appl* **683**, 155-62.

Lin, Y. S., Dowling, A. L. S., Quigley, S. D., Farin, F. M., Zhang, J., Lamba, J., Schuetz, A. G., Thummel, K. E. (2002). Co-regulation of CYP3A4 and CYP3A5 and contribution to hepatic and intestinal midazolam metabolism. *Mol Pharmacol* **62**, 162-72.

Little, J. M., Lehman, P. A., Nowell, S., Samokyszyn. V., Radominska, A. (1997). Glucuronidation of all-trans-retinoic acid and 5,6-epoxy-all-trans-retinoic acid. Activation of rat liver microsomal UDP-glucuronosyltransferase activity by alamethicin. *Drug Metab Dispos* **25**, 5-11.

Louisse, J., de Jong, E., van de Sandt, J. J. M., Blaauboer, B. J., Woutersen, R. A., Piersma, A. H., Rietjens, I. M. C. M., Verwei, M. (2010). The use of *in vitro* toxicity data and physiologically based kinetic modeling to predict dose-response curves for *in vivo* developmental toxicity of glycol ethers in rat and man. *Toxicol Sci* **118**, 470-484.

Louisse, J., Gönen, S., Rietjens, I. M. C. M., Verwei, M. (2011). Relative developmental toxicity potencies of retinoids in the embryonic stem cell test compared with their relative potencies in in vivo and two other in vitro assays for developmental toxicity. *Toxicol Lett* **203**, 1-8.

Lutz, J. D., Dixit, V., Yeung, C. K., Dickmann, L. J., Zelter, A., Thatcher, J. E., Nelson, W. L., Isoherranen. N. (2009). Expression and functional characterization of cytochrome P450 26A1, a retinoic acid hydroxylase. *Biochem Pharmacol* **77**, 258-68.

Marchetti, M. N., Sampol, E., Bun, H., Scoma, H., Lacarelle, B., Durand, A. (1997). In vitro metabolism of three major isomers of retinoic acid in rats. Intersex and interstrain comparison. *Drug Metab Dispos* **25**, 637-46.

McConnell, E. L., Basit, A. W., Murdan, S. (2008). Measurements of rat and mouse gastrointestinal pH, fluid and lymphoid tissue, and implications for in-vivo experiments. *J Pharm Pharmacol* **60**, 63-70.

Medinsky, M. A., Leavens, T. L., Csanady, G. A., Gargas, M. L., Bond, J. A. (1994). In vivo metabolism of butadiene by mice and rats: a comparison of physiological model predictions and experimental data. *Carcinogenesis* **15**, 1329-1340.

Naraharisetti, S. B., Lin, Y. S., Rieder, M. J., Marciante, K. D., Psaty, B. M., Thummel, K. E., Totah, R. A. (2010). Human liver expression of CYP2C8: Gender, Age, and Genotype Effects. *Drug Metab Dispos* **38**, 889-893.

Nau, H. (2001). Teratogenicity of isotretinoin revisited: species variation and the role of all-*trans*-

retinoic acid. *J Am Acad Dermatol* **45**, S183-S187.

Paini, A., Punt, A., Viton, F., Scholz, G., Delatour, T., Marin-Kuan, M., Schilter, B., van Bladeren, P. J., Rietjens, I. M. C. M. (2010). A physiologically based biodynamic (PBBD) model for estragole DNA binding in rat liver based on in vitro kinetic data and estragole DNA adduct formation in primary hepatocytes. *Toxicol Appl Pharmacol* **245**, 57-66.

Punt, A., Jeurissen, S., Boersma, M. G., Delatour, T., Scholz, G., Schilter, B., van Bladeren P. J., Rietjens, I. M C. M. (2010). Evaluation of human interindividual variation in bioactivation of estragole using physiologically based biokinetic modeling. *Toxicol Sci* **113**, 337-48.

Ravindranath, V., Anandatheerthavarada H. K. (1990). Preparation of brain microsomes with cytochrome P450 activity using calcium aggregation method. *Anal Biochem* **187**, 310-3.

Reilly, J. A. Jr., Forst, C. F., Quigley, E. M., Rikkers, L. F. (1990). Gastric emptying of liquids and solids in the portal hypertensive rat. *Dig Dis Sci* **35**, 781-6.

Rietjens, I. M. C. M., Louisse, J., Punt, A. (2011). Tutorial on physiologically based kinetic modeling in molecular nutrition and food research. *Mol Nutr Food Res* **55**, 941-56.

Rotroff, D. M., Wetmore, B. A., Dix, D. J., Ferguson, S. S., Clewell, H. J., Houck, K. A., Lecluyse, E. L., Andersen, M. E., Judson, R. S., Smith, C. M., Sochaski, M. A., Kavlock, R. J., Boellmann, F., Martin, M. T., Reif, D. M., Wambaugh, J. F., Thomas, R. S. (2010). Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening. *Toxicol Sci* **117**, 348-58.

Samokyszyn, V. M., Gall, W. E., Zawada, G., Freyaldenhoven, M. A., Chen, G., Mackenzie, P. I., Tephly, T. R., Radominska-Pandya, A. (2000). 4-Hydroxyretinoic acid, a novel substrate for human liver microsomal UDP-glucuronosyltransferase(s) and recombinant UGT2B7. *J Biol Chem* **275**, 6908-14.

Sass, J. O., Hartmann, J., Chahoud, I., Shroot, B., Nau, H. (1995). Transplacental pharmacokinetics of a synthetic retinoid which is not bound by mouse embryonic cellular retinoic acid-binding protein. *Toxicol Lett* **75**, 159-68.

Shelley, R. S., Jun, H. W., Price, J. C., Cadwallader, D. E. (1982). Blood level studies of all-trans- and 13-cis-retinoic acids in rats using different formulations. *J Pharm Sci* **71**, 904-7.

Sim, S. C., Edwards, R. J., Boobis, A. R., Ingelman-Sundberg, M. (2005). CYP3A7 protein expression is high in a fraction of adult human livers and partially associated with the CYP3A7*1C allele. *Pharmacogenet Genomics* **15**, 625-31.

Smith, J. E., Milch, P. O., Muto, Y., Goodman, D. S. (1973). The plasma transport and metabolism of retinoic acid in the rat. *Biochem J* **132**, 821-7.

Sun, W. M., Houghton, L. A., Read, N. W., Grundy, D. G., Johnson, A. G. (1988). Effect of meal temperature on gastric emptying of liquids in man. *Gut* **29**, 302-5.

Swanson, B. N., Frolik, C. A., Zaharevitz, D. W., Roller, P. P., Sporn, M. B. (1981). Dose-dependent kinetics of all-trans-retinoic acid in rats. Plasma levels and excretion into bile, urine, and faeces. *Biochem Pharmacol* **30**, 107-13.

Tay, S., Dickmann, L., Dixit, V., Isoherranen, N. (2010). A comparison of the roles of peroxisome proliferator-activated receptor and retinoic acid receptor on CYP26 regulation. *Mol Pharmacol* **77**, 218-27.

Tembe, E. A., Honeywell, R., Buss, N. E., Renwick, A. G. (1996). All-trans-retinoic acid in maternal plasma and teratogenicity in rats and rabbits. *Toxicol Appl Pharmacol* **141**, 456-472.

Thatcher, J. E., Zelter, A., Isoherranen, N. (2010). The relative importance of CYP26A1 in hepatic clearance of all-trans retinoic acid. *Biochem Pharmacol* **80**, 903-12.

Tzimas, G., Thiel, R., Chahoud, I., Nau, H. (1997). The area under the concentration-time curve of all-trans-retinoic acid is the most suitable pharmacokinetic correlate to the embryotoxicity of this retinoid in the rat. *Toxicol Appl Pharmacol* **143**, 436-44.

Van der Jagt, K., Munn, S. J., Tørsløv, J., de Bruijn, J. (2004). Alternative approaches can reduce the use of test animals under REACH. *Report EUR* **21405**.

Vaes, W. (2010). Assay system for determining binding of hydrophobic drugs. Patent no. WO/2010/117276.

Verwei, M., van Burgsteden, J. A., Krul, C. A. M., van de Sandt, J. J. M., Freidig, A. P. (2006a). Prediction of in vivo embryotoxic effect levels with a combination of in vitro studies and PBPK modelling. *Toxicol Lett* **165**, 79-87.

Verwei, M., Freidig, A. P., Havenaar, R., Groten, J. P. (2006b). Predicted serum folate concentrations based on in vitro studies and kinetic modeling are consistent with measured folate concentrations in humans. *J Nutr* **136**, 3074-8.

Wetmore, B. A., Wambaugh, J. F., Ferguson, S. S., Sochaski, M. A., Rotroff, D. M., Freeman, K., Clewell, H. J. 3rd, Dix, D. J., Andersen, M. E., Houck, K. A., Allen, B., Judson, R. S., Singh, R., Kavlock, R. J., Richard, A. M., Thomas, R. S. (2012). Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. *Toxicol Sci* **125**, 157-74.

Wise, L. D., Xue, D., Winkelmann, C. T. (2010). Micro-computed tomographic evaluation of fetal skeletal changes induced by all-trans-retinoic acid in rats and rabbits. *Birth Defects Res B Dev Reprod Toxicol* **89**, 408-17.

Zaya, M. J., Hines, R. N., Stevens, J. C. (2006). Epirubin glucuronidation and UGT2B7 developmental expression. *Drug Metabol Dispos* **34**, 2097-2101.

Zhang, X., Tsang, A. M., Okino, M. S., Power, F. W., Knaak, J. B., Harrison, L. S., Dary, C. C. (2007). A physiologically based pharmacokinetic/pharmacodynamic model for carbofuran in Sprague-Dawley rats using the exposure-related dose estimating model. *Toxicol Sci* **100**, 345-59.

Zimmerman, C. L., Han, S., Wiedmann, T. S. (2001). The absorption of retinoic acids from the gastrointestinal tract is dependent upon chemical structure. *Cancer Chemother Pharmacol* **47**, 27-33.

CHAPTER 8

Summary, general discussion, future perspectives and conclusions

Summary

The aim of the present thesis is to provide the proof-of-principle that *in vitro* concentration-response curves can be converted into *in vivo* dose-response curves, using physiologically based kinetic (PBK) modeling with a reverse dosimetry approach, in order to provide alternative methods for *in vivo* studies in toxicological risk assessment. Developmental toxicity was chosen as toxicological endpoint for this proof-of-principle study, since these type of toxicity studies are estimated to require more than one fourth of the animals expected to be needed under the European REACH (Registration, Evaluation, Authorization and restriction of CHemicals) legislation (Van der Jagt *et al.*, 2004). Also for pharmaceuticals, food additives and pesticides, regulatory agencies require safety studies performed in laboratory animals, including developmental toxicity studies. Therefore, the application of reliable alternatives for *in vivo* developmental toxicity studies would contribute significantly to the replacement, reduction and refinement (3Rs) of animal use in toxicological risk assessment.

Chapter 1 of the present thesis provides background information introducing the topic, and defines the aim of the thesis. It describes important factors that are needed to develop the concept of translating *in vitro* concentration-response curves into predicted *in vivo* dose-response curves for developmental toxicity, using PBK modeling with a reverse dosimetry approach. Briefly, these factors include 1) an *in vitro* model to obtain *in vitro* concentration-response curves for developmental toxicity, 2) the selection of chemicals to be studied and 3) the development of the PBK models needed for reverse dosimetry.

The *in vitro* model used was the embryonic stem cell (ES-D3 cell) differentiation assay of the validated embryonic stem cell test (EST) (Genschow *et al.*, 2004). In this assay, the effect of the chemical of interest on ES-D3 cell differentiation into functional (contracting) cardiomyocytes is assessed. The chemical-induced inhibition of ES-D3 cell differentiation is the *in vitro* developmental toxicity endpoint assessed. The ES-D3 cell differentiation assay is the only validated *in vitro* developmental toxicity assay for which no primary animal tissues are required (Genschow *et al.*, 2002; Genschow *et al.*, 2004). Therefore, alternatives for *in vivo* developmental toxicity studies using this *in vitro* assay would result in the largest contribution to the 3Rs of animal use for developmental toxicity testing.

To provide the proof-of-principle that the obtained *in vitro* concentration-response curves in the ES-D3 cell differentiation assay can be converted into *in vivo* dose-response curves using PBK modeling with a reverse dosimetry approach, appropriate model compounds needed to be selected fulfilling certain criteria. First, a list of chemicals was put together containing chemicals meeting two criteria, being 1) the presence of *in vivo* kinetic data on the chemical in the literature, in order to be able to build and/or evaluate the PBK models developed and used for reverse dosimetry and 2) the presence of *in vivo* developmental toxicity data of the chemical in the literature, in order to evaluate whether the ultimately predicted dose-response curves for developmental toxicity are in concordance with data obtained in *in vivo* developmental toxicity studies. From this list of appropriate chemicals that could be used for the research, only series of structurally related chemicals were selected. This third criterion enabled the assessment of relative differences within a series of chemicals, thereby providing an additional possibility for qualitative evaluation of the predictions made. Two series of model chemicals were selected, being 1) a group of glycol ethers, which need to be bioactivated in order to cause developmental toxicity, which are hydrophilic, show low protein binding and have an unspecific mode of action and 2) a group of retinoids, containing several members that directly act as teratogens without bioactivation, which are hydrophobic, show high protein binding and have a specific mode of action.

In order to translate the *in vitro* concentration-response curves obtained in the ES-D3 cell differentiation assay into *in vivo* dose-response curves with reverse dosimetry, PBK models describing the ADME (Absorption, Distribution, Metabolism and Excretion) processes of the chemicals of interest needed to be developed. The process of PBK model development is described in chapter 1 and can be divided in the following steps: 1) definition of a conceptual model, 2) translation of the conceptual model into a mathematical model by defining mathematical equations, 3) determination of the model parameter values, 4) solving the mathematical equations using appropriate software and 5) evaluation of the model performance. Subsequently, the model can be used to make predictions (step 6), e.g. to predict dose-dependent inter- or intraspecies differences in kinetics, or, as was done in the present thesis, to translate internal concentrations set as input in the model into predicted external dose levels (i.e. reverse

dosimetry).

In **chapter 2**, *in vitro* assays are presented that have been described in the literature for the endpoint developmental toxicity. These assays can be divided in 1) cellular assays, using either primary cells or cells from continuous cell lines and 2) whole embryo culture (WEC) assays. An overview is given of the varying readout parameters, ranging from gene or protein expression changes to cell death or *in vitro* malformations, that have been assessed in these *in vitro* assays. Chapter 2 also describes how *in vitro* developmental toxicity data may be used, according to the literature, to make predictions for the *in vivo* situation. This can be done in several ways: the *in vitro* data may be used to 1) predict the developmental toxicity class to which a chemical belongs (e.g., non-, weak or strong embryotoxic) using a prediction model, 2) predict the relative developmental toxicity potency of a chemical based on the *in vitro* effect concentrations (usually within a group of structurally related chemicals) and 3) predict *in vivo* developmental toxicity dose levels by translating *in vitro* effect concentrations into *in vivo* effect doses, using PBK modeling. However, the prediction of *in vivo* developmental toxicity based on *in vitro* data can be over- or underestimated due to various reasons. These include: 1) the lack or the low expression levels of biotransformation enzymes in *in vitro* test systems compared with the *in vivo* situation, 2) the possible lack of relevant cells or tissues in the *in vitro* test system for the effects observed *in vivo*, 3) the lack of secondary adverse effects resulting from maternal or placental toxicity in the *in vitro* test system and 4) the lack of relevant kinetic data needed to construct a PBK model that reliably describes the *in vivo* kinetics, for the quantitative extrapolation of *in vitro* effect concentrations to *in vivo* effect doses.

The first group of chemicals selected to develop the proof-of-principle for translating *in vitro* concentration-response curves into *in vivo* dose-response curves using PBK modeling with reverse dosimetry, was a group of glycol ethers. The glycol ethers used were ethylene glycol monomethyl ether (EGME), ethylene glycol monoethyl ether (EGEE), ethylene glycol monobutyl ether (EGBE) and ethylene glycol monophenyl ether (EGPE). These glycol ethers are *in vivo* mainly metabolized to their alkoxyacetic acid metabolites, being methoxyacetic acid (MAA), ethoxyacetic acid (EAA), butoxyacetic acid (BAA) and phenoxyacetic acid (PAA), respectively, which are considered to be responsible for the developmental

toxicity of the glycol ethers (ECETOC, 2005). It was shown in the validation studies of the EST and the rat postimplantation WEC test, that MAA's toxic effect concentrations are in the milimolar range (Genschow *et al.*, 2004; Piersma *et al.*, 2004). *In vivo* studies show that MAA's peak plasma concentrations at doses causing developmental toxicity in rats and mice *in vivo*, are in the millimolar range as well (Sweeney *et al.*, 2001). Considering MAA's high *in vitro* and *in vivo* effect concentrations and its acidic nature, it was hypothesized that the alkoxyacetic acid metabolite-induced developmental toxicity may be caused by the induction of an intracellular pH (pH_i) decrease of embryonic cells, resulting in an inhibition of cell growth and differentiation, thereby disturbing embryonic development. To further investigate this mode of action of the alkoxyacetic acid metabolites of the glycol ethers, the effects of MAA on the pH_i of ES-D3 cells and embryonic fibroblasts (Balb/c-3T3 cells) were assessed in **chapter 3** of the present thesis. In this study, the chemicals acetazolamide (ACZ) and valproic acid (VPA) were used as positive controls because for these chemicals a relation between a chemical-induced pH_i decrease in the conceptus and the developmental toxicity outcome has been established previously in *in vivo* studies (Scott *et al.*, 1990; Scott *et al.*, 1997). The results obtained showed that MAA, ACZ and VPA cause a pH_i decrease at concentrations that induce inhibition of ES-D3 cell differentiation, indicating that the pH_i decrease caused by these chemicals is related to the inhibition of cell differentiation. Two other developmental toxicants, all-*trans*-retinoic acid and 5-fluorouracil, did not decrease the pH_i of embryonic cells at concentrations that affect ES-D3 cell differentiation, pointing at a different mode of action of these chemicals, as also described in the literature (Collins and Mao, 1999; Shuey *et al.*, 1994). MAA and ACZ induced a concentration-dependent inhibition of ES-D3 cell differentiation, which was potentiated by amiloride, an inhibitor of the Na^+/H^+-antiporter, corroborating an important role of the pH_i on the mechanism of developmental toxicity of both chemicals. The results presented in chapter 3 indicate that a decrease of the pH_i is likely to be the mode of developmental toxicity of the alkoxyacetic acid metabolites of the glycol ethers and corroborate the role of the alkoxyacetic acid metabolites in the mode of action for the developmental toxicity of the parent glycol ethers. Confirmed by this mode of action, the *in vitro* concentration-response curves (chapter 4) to be converted into *in vivo* dose-response curves (chapter 5) were made with the toxic alkoxyacetic acid

metabolites instead of with the parent glycol ethers, considering that subsequent PBK modeling (chapter 5) would take into account the bioactivation of the glycol ethers to their alkoxyacetic acid metabolites.

Chapter 4 presents the concentration-response curves of MAA, EAA, BAA, PAA, EGME and EGEE in the ES-D3 cell differentiation assay and compares the outcomes with literature data on the *in vivo* developmental toxicity effects of EGME, EGEE, EGBE and EGPE in mice. All glycol ether alkoxyacetic acid metabolites tested showed a concentration-dependent inhibition of cardiomyocyte differentiation, with MAA as the most potent chemical followed by EAA, BAA and PAA, respectively. Both parent glycol ethers tested (EGME and EGEE, up to 10 mM) did not inhibit ES-D3 cell differentiation, corroborating that the metabolites are the proximate developmental toxicants *in vivo*. The potencies of the glycol ether alkoxyacetic acid metabolites in the ES-D3 cell differentiation assay were compared with *in vivo* toxicity data of the glycol ethers in mice. The potency ranking of the chemicals in the ES-D3 cell differentiation assay corresponded with the available *in vivo* data on their parent glycol ether's potencies (EGME > EGEE > EGBE > EGPE). The relative differences between the potencies of the chemicals appeared more pronounced in the *in vivo* studies than in the ES-D3 cell differentiation assay. A possible explanation for this discrepancy can be the absence of *in vivo* kinetic processes in the *in vitro* test system, which are known to differ between these chemicals in the *in vivo* situation.

These differences in *in vivo* kinetics can be described in PBK models to be used for the translation of the *in vitro* effect concentrations into *in vivo* effect doses. **Chapter 5** describes the results of applying this reverse dosimetry approach for the glycol ethers EGME, EGEE, EGBE and EGPE in rat and human. A PBK model was developed, describing the kinetics of the four glycol ethers and their developmentally toxic alkoxyacetic acid metabolites in rat and human. The *in vitro* concentration-response curves of these metabolites in the ES-D3 cell differentiation assay (chapter 4) were used as input in the PBK model to translate *in vitro* concentration-response curves into predicted *in vivo* dose-response curves for developmental toxicity of the parent glycol ethers in rat and human. The predicted dose-response curves for rat were found to be in concordance (within one order of magnitude) with the developmental toxicity dose levels measured in

different *in vivo* rat developmental toxicity studies. The results thus obtained in chapter 5 provide the first proof-of-principle that *in vitro* concentration-response curves can be converted into *in vivo* dose-response curves suitable for risk assessment practice, using PBK modeling with a reverse dosimetry approach. Therefore, it was concluded that this approach, after further evaluation, may contribute to a science-based risk assessment of chemicals, using no or reduced numbers of laboratory animals.

The second group of chemicals used to assess the reliability of predicting *in vivo* dose-response curves for *in vivo* developmental toxicity by using the reverse dosimetry approach, was a group of retinoids. Retinoids are chemicals that possess a chemical structure or functional properties similar to retinol (vitamin A). The retinoids investigated were retinol, all-*trans*-retinoic acid (ATRA), 13-*cis*-retinoic acid (13-*cis*-RA), 9-*cis*-retinoic acid (9-*cis*-RA), etretinate, acitretin and TTNPB ((E) -4[2-(5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2-naphthalenyl)-1-propenyl] benzoic acid). Proper levels of vitamin A must be maintained for normal embryogenesis, whereas vitamin A deficiency and high retinoid intake may induce adverse effects to the developing conceptus (Collins and Mao, 1999). Several retinoids have been shown to be teratogenic in a number of experimental animals, including mice, rats, guinea pigs, hamsters, rabbits, dogs, pigs, chickens and monkeys (Collins and Mao, 1999). Retinoids have previously been shown to be active in the limb bud micromass test in the nanomolar to micromolar range (Kistler, 1987), indicating the higher *in vitro* potency of these chemicals compared with the glycol ether alkoxyacetic acid metabolites. Furthermore, whereas the mechanism of developmental toxicity of the glycol ethers is likely to be unspecific (induction of pH_i decrease), the mechanism of retinoid-induced developmental toxicity is more complex (disturbance of intracellular retinoid signal transduction pathways (Collins and Mao, 1999)).

Chapter 6 presents the concentration-response curves of retinol, ATRA, 13-*cis*-RA, 9-*cis*-RA, etretinate, acitretin and TTNPB in the ES-D3 cell differentiation assay and compares the outcomes of their relative potencies with the relative potencies detected in two other *in vitro* assays for developmental toxicity (the limb bud micromass test and the rat postimplantation WEC test) and with the relative potencies in *in vivo* (mouse, rat and rabbit) developmental toxicity studies. The

results reveal that the potency ranking obtained in the ES-D3 cell differentiation assay is similar to the reported potency rankings in the two other *in vitro* assays for developmental toxicity. TTNPB was the most potent retinoid, whereas etretinate and retinol had the lowest potency. ATRA, 13-*cis*-RA, 9-*cis*-RA and acitretin showed an intermediate potency. *In vivo* potency rankings of the developmental toxicity of retinoids appear to be dependent on the species and/or exposure regimens used. The obtained *in vitro* potency ranking did not completely correspond with the *in vivo* potency rankings, although TTNPB was correctly predicted to be the most potent and retinol the least potent congener. The differences between the kinetics in the ES-D3 cell differentiation assay and the kinetic processes *in vivo* may explain the deviating potency predictions of some retinoids. Therefore, it was concluded that knowledge on species-dependent chemical-specific *in vivo* kinetics is essential when using *in vitro* toxicity data for the estimation of *in vivo* developmental toxicity potencies within series of related chemicals.

In **chapter 7**, the reverse dosimetry approach was applied for the retinoid ATRA, translating the *in vitro* concentration-response curve for the inhibition of ES-D3 cell differentiation into *in vivo* dose-response curves, using a PBK model describing ATRA's kinetics in rat and human. Whereas for the PBK models developed for the glycol ethers (chapter 5) kinetic data from *in vivo* studies were required to derive certain model parameter values (i.e. for clearance of the alkoxyacetic acid metabolites), the PBK model for ATRA was solely based on parameter values derived using *in vitro* techniques. Predictions of dose-response curves for *in vivo* developmental toxicity were made by relating the toxic effect to 1) the maximal concentration (C_{max}) in blood, as was done before for the glycol ethers (chapter 5) and 2) the area under the concentration-time curve (AUC) in blood. The predicted developmental toxicity dose levels of ATRA were closer to the *in vivo* rat data when predictions were based on relating the toxic effect to the AUC, than when predictions were based on relating the toxic effect to the C_{max}. This is in line with results reported in the literature, showing that the developmental toxicity outcome of ATRA *in vivo* is best correlated with the dose metric AUC (Tzimas *et al.*, 1997). The predicted dose-response curves for rat were found to be in concordance (within one order of magnitude) with the developmental toxicity dose levels measured in different *in vivo* rat developmental toxicity studies. It is important to stress that an essential addition of this study to the previous reverse dosimetry

work on glycol ethers (chapter 5) is, that no kinetic data from *in vivo* studies were needed to derive the model parameter values for the ATRA PBK model. This study thus provides the first proof-of-principle that *in vivo* dose-response relationships for this systemic endpoint can be predicted using reverse dosimetry with a PBK model that was parameterized without *in vivo* data. The predicted dose-response curves can be used to set points of departure for deriving safe exposure levels for humans, pointing at the potential of the approach as a tool in toxicological risk assessment strategies.

General discussion

The present thesis shows the proof-of-principle that *in vitro* concentration-response curves can be translated into predicted *in vivo* dose-response curves using PBK modeling with a reverse dosimetry approach, which may increase the acceptance of the use of *in vitro* toxicity data in risk assessment practice. Using this approach, the important bottleneck that *in vitro* toxicity data for systemic endpoints can only be used qualitatively can be overcome, since the predicted dose-response curves can be used quantitatively by determining points of departure for the risk assessment. Reverse dosimetry approaches have previously been used to translate a single *in vitro* effect concentration (e.g. an EC_{50} value) into a single *in vivo* effect dose (e.g. a lowest observed adverse effect level (LOAEL)) (DeJongh *et al.*, 1999; Verwei *et al.*, 2006; Forsby and Blaauboer, 2007). However, the *in vitro* effect concentrations and the *in vivo* effect doses were arbitrarily chosen in these studies. By translating all data obtained in the *in vitro* experiment to the *in vivo* situation enables dose-response modeling on the predicted *in vivo* dose-response data, defining a predicted *in vivo* dose-response curve, which can be used in the risk assessment process. The present thesis has shown the feasibility of this approach for chemicals belonging to two groups of chemicals with diverse chemical properties and mechanisms of developmental toxicity, indicating that the approach may be feasible for chemicals of diverse nature.

In the following sections, the possibilities and limitations of the use of *in vitro-in vivo* extrapolations using PBK modeling with a reverse dosimetry approach for risk assessment are discussed. Recommendations for future research to further develop

and validate and ultimately implement the combined *in vitro-in silico* approach are put forward. The topics to be specifically considered when discussing the results of the present thesis are 1) the choice of the *in vitro* assay for developmental toxicity, 2) developments in the field of PBK modeling, 3) additional factors to be taken into account when translating *in vitro* concentrations into *in vivo* doses and 4) future perspectives and conclusions for the use of *in vitro-in vivo* extrapolations using PBK modeling with a reverse dosimetry approach for risk assessment.

The choice of an in vitro toxicity assay for developmental toxicity

In the present thesis, the ES-D3 cell differentiation assay of the EST was used as *in vitro* assay to determine *in vitro* toxic effect concentrations for the endpoint developmental toxicity (Genschow *et al.*, 2002; Genschow *et al.*, 2004). The EST is the only validated *in vitro* assay for which no primary animal tissues are required. The validation study carried out by the European Centre for the Validation of Alternative Methods (ECVAM) showed consistency in the obtained toxic effect concentrations for 20 chemicals in four different laboratories (Genschow *et al.*, 2004), pointing at the reproducibility of the test system. The readout parameter in the ES-D3 cell differentiation assay is the chemical-induced inhibition of ES-D3 cell differentiation into functional (contracting) cardiomyocytes. This readout parameter may be subjective and not representative for chemicals that induce malformations in the *in vivo* situation different than those on the heart. Therefore, several studies have investigated the effect of chemicals on other readouts in differentiating embryonic stem cells, such as on the expression of specific genes for neuronal differentiation (Stumann *et al.*, 2007; Kuegler *et al.*, 2010; Theunissen *et al.*, 2010; Theunissen *et al.*, 2011; Zimmer *et al.*, 2011a; Zimmer *et al.*, 2011b), osteoblast differentiation (Zur Nieden *et al.*, 2010a; Zur Nieden *et al.*, 2010b) or endothelial differentiation (Fesstag *et al.*, 2007). These readouts may be useful to predict adverse effects on specific developmental pathways, such as neurodevelopment, bone formation, or vasculogenesis and/or angiogenesis, respectively. Furthermore, it has been proposed that with transcriptome approaches the sensitivity of the embryonic stem cell-based test system may possibly increase (Van Dartel *et al.*, 2011). However, with the interpretation of toxicogenomic data, one must be aware that gene expression changes *per se* are not necessarily adverse. A change in gene expression may reflect an adaptation of the

cell or tissue to a situation, not resulting in an adverse effect. Van Dartel *et al.* (2011) compared the effect of the chemical flusilazole in the 'classical' ES-D3 cell differentiation assay with the effects of a 24-hour exposure of differentiating ES-D3 cells to this chemical on gene expression changes. Although the sensitivity using the transcriptomic approach had slightly increased (Van Dartel *et al.*, 2011), it is not straightforward to use the obtained concentration-response curves for deriving adverse effect concentrations, because, especially at low concentrations, the gene expression changes may reflect an adaptation of the cell, not resulting in adverse effects. Since the chemical-induced inhibition of ES-D3 cell differentiation into functional cardiomyocytes is considered as adverse, the 'classical' ES-D3 cell differentiation assay may provide relevant *in vitro* adverse effect concentrations for reverse-dosimetry, in order to predict *in vivo* toxic dose levels. Although the ES-D3 cell differentiation assay determines the concentrations of a chemical that inhibit cardiomyocyte differentiation, it was not intended in the present thesis to specifically predict dose levels affecting cardiac development during embryonic development. The inhibition of cardiac differentiation was rather used to represent a sensitive *in vitro* readout parameters for developmental toxicity (Genschow et al., 2004), which in the ideal situation could represent the most sensitive *in vivo* readout parameter for the endpoint developmental toxicity. One might argue that other developmental toxicity endpoints such as developmental neurotoxicity or developmental immunotoxicity might be less well predicted with the use of the *in vitro* effect concentrations that inhibit cardiomyocyte differentiation. Therefore, for other developmental toxicity endpoints such as developmental neurotoxicity or developmental immunotoxicity, it should be assessed whether *in vivo* developmental toxicity dose levels are better predicted based on *in vitro* concentration-response curves obtained in assays focusing on these specific endpoints (e.g. Breier *et al.*, 2010; Mainali and Tew, 2004) than when dose levels are predicted based on *in vitro* concentration-response curves obtained in the classical ES-D3 cell differentiation assay.

The *in vitro* assays for developmental toxicity described in the literature have varying levels of biological complexity, ranging from cellular assays (e.g. the ES-D3 cell differentiation assay or the limb bud micromass test) to WEC assays (e.g. the rat postimplantation WEC assay). If the developmental toxicant has an adverse effect on specific cell types, the toxicity may not be detected in an *in vitro* test

system in which these cells are absent. This would suggest that WEC assays, which are closer to the *in vivo* situation than cellular assays, would be better test systems to detect adverse effects of a range of chemicals affecting different cell types and with different modes of action (Piersma, 2004; Piersma, 2006). In addition, a battery of cellular assays covering the relevant cell types that can be affected during development could be used to detect a developmental toxicant that might not be detected by a single cellular assay (Piersma, 2006). Instead of using cellular assays in which specific cell types are present, embryonic stem cells, which can be differentiated into a mix of relevant cell types, are useful to assess chemical-induced effects on different cell types, and on the differentiation into these different cell types, using multiple readouts in one test system. Differentiating ES-D3 cells may be a useful tool to accomplish this, because although the readout in the classical ES-D3 cell differentiation assay is the inhibition of ES-D3 cell differentiation into contracting cardiomyocytes, cell types of all the three germ layers are present in these cultures (Doetschman *et al.*, 1985; Toumadje *et al.*, 2003).

To assess whether the three ECVAM-validated *in vitro* developmental toxicity assays (the ES-D3 cell differentiation assay, the limb bud micromass (MM) test, and the rat postimplantation WEC test), which have different levels of biological complexity, differ in their sensitivity to two of the developmental toxicants studied in the present thesis, MAA and ATRA, the reported effect concentrations for these chemicals in these assays were compared. To this end, the averages of the reported EC_{50} values in the ECVAM validation studies (Genschow *et al.*, 2004; Piersma *et al.*, 2004; Spielmann *et al.*, 2004) were calculated (Table 1). Table 1 shows that these average EC_{50} values for MAA differ at most 4-fold between the assays. The average EC_{50} values for ATRA differ almost up to 80-fold between the assays (Table 1). At first sight, this may indicate large differences in the sensitivity of the test systems to ATRA and much smaller, if any, differences in the sensitivity of the test systems to MAA. However, a chemical can be more freely available in one test system compared with another, depending on the composition of the culture medium and the physico-chemical properties of the chemical. This phenomenon may play a role in the large differences in the obtained nominal effect concentrations for the hydrophobic ATRA.

The main factor that may influence the free fraction of a hydrophobic chemical *in vitro* is the percentage of serum that is present in the assay medium. The percentages serum applied in the three *in vitro* developmental toxicity assays are shown in Table 1. The solid phase microextraction (SPME) method as described in chapter 7 was used to determine the free fraction of a chemical in culture medium with increasing serum concentrations. This measurements show that with increasing serum percentages (v/v) in the medium, the free fraction of ATRA decreases (Figure 1). Differences in free fractions have to be taken into account in *in vitro-in vivo* extrapolations, as discussed in chapter 7, but should also be considered when assessing differences in the sensitivity of *in vitro* test systems. The

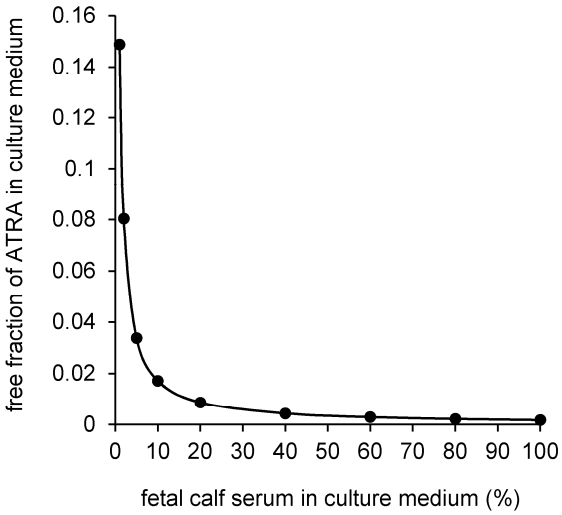

Figure 1. Free fraction of ATRA in cell culture medium with increasing concentrations of serum. Data are obtained using SPME method described in chapter 7.

graph presented in Figure 1 was used to determine the free EC_{50} values in the three assays (Table 1). These free EC_{50} values for ATRA in the three *in vitro* developmental toxicity assays now only differ at maximum 19-fold (Table 1), resulting in a 4-fold decrease in relative difference, than when comparing nominal EC_{50} values. Still, the almost 20-fold difference in the sensitivity of the test systems to ATRA remains large. Since the AUC is the dose metric that best correlates to the

developmental toxicity outcome of ATRA *in vivo*, *in vitro* AUC values can best be used in the comparison of assay sensitivity to ATRA. To this end, the free EC_{50} values were multiplied with the assay times, in order to calculate *in vitro* $EAUC_{50}$ values (Daston *et al.*, 2010). Comparison of these *in vitro* AUC values indicates that the differences in the sensitivity of the test systems are at maximum only 5-fold (Table 1). Table 1 also presents the same analysis for MAA. Since it is reported that MAA does not bind to proteins (O'Flaherty *et al.*, 1995), the free EC_{50} values are expected to be the same as the nominal EC_{50} values. Differences in the assay sensitivity to MAA increase 2-fold to an overall difference of 8-fold when the *in vitro* developmental toxicity outcomes are based on *in vitro* AUC values (Table 1). Altogether, these findings indicate that the three *in vitro* developmental toxicity assays generally have the same sensitivity to MAA and ATRA, when the developmental toxicity outcome is related to the *in vitro* C_{max} for MAA and to the *in vitro* AUC for ATRA. Given that the ES-D3 cell differentiation assay is the only validated *in vitro* developmental toxicity assay for which no primary animal tissues are required, it may be considered as the best available *in vitro* model for the largest contribution to the 3Rs of animal use in toxicological studies.

To contribute even more to the 3Rs of animal use, serum-free culture methods should be developed and implemented, as the use of serum requires high numbers of (fetal) calves (Jochems *et al.*, 2002; Van der Valk *et al.*, 2004). Recently, a serum-free culture method for ES-D3 cells and ES-D3 cell differentiation has been developed (Riebeling *et al.*, 2011). Another advantage of serum-free test systems is a better standardization, because batch-to-batch variation of serum can highly influence the outcome of the assay. Differentiation processes (e.g. of embryonic stem cells) are very sensitive to medium conditions, reflected by the fact that with certain batches of serum, ES-D3 cells do not differentiate into functional cardiomyocytes (data not shown). Therefore, the development of serum-free culture protocols for *in vitro* (developmental) toxicity assays is important for test system standardization to increase the reproducibility of test outcomes. However, FCS is an important source of proteins in culture medium, and one must be aware that the absence of proteins in test systems may also have limitations. For example, it was shown for ATRA that the presence of proteins is required for its solubilization and stabilization in culture medium, indicated by irreproducible results in *in vitro* studies if proteins were absent (Klaassen *et al.*, 1999). Therefore,

Table 1. *In vitro* effect concentrations of MAA and ATRA in different *in vitro* developmental toxicity assays. Average EC_{50} values (± SD) were calculated from values reported in the ECVAM validation studies.

Assay	Readout parameter	Reported EC_{50} in assay	Serum conditions culture medium	Calculated free EC_{50} in assay[d]	Exposure duration	Calculated free $EAUC_{50}$ in assay[e]
MAA						
ES-D3 cell differentiation assay	inhibition of ES-D3 cell differentiation	2.0 (±0.87) mM[a]	20% FCS	2.0 (±0.87) mM	10 days	480 (±210) mM·h
rat postimplantation WEC assay	induction of malformations	2.3 (±1.5) mM[b]	100% rat serum	2.3 (±1.5) mM	2 days	110 (±72) mM·h
limb bud micromass test	inhibition of limb bud cell differentiation	7.9 (±2.2) mM[c]	5% FCS	7.9 (±2.2) mM	5 days	950 (±260) mM·h
ATRA						
ES-D3 cell differentiation assay	inhibition of ES-D3 cell differentiation	6.8 (±8.8) nM[a]	20% FCS	0.059 (±0.076) nM	10 days	14 (±18) nM·h
rat postimplantation WEC assay	induction of malformations	540 (±290) nM[b]	100% rat serum	1.1 (±0.59) nM	2 days	53 (±18) nM·h
limb bud micromass test	inhibition of limb bud cell differentiation	17 (±9.4) nM[c]	5% FCS	0.58 (±0.32) nM	5 days	70 (±39) nM·h

[a] Averages calculated from reported data of ECVAM validation study (Genschow *et al.*, 2004).
[b] Averages calculated from reported data of ECVAM validation study (Piersma *et al.*, 2004).
[c] Averages calculated from reported data of ECVAM validation study (Spielmann *et al.*, 2004).
[d] Calculated by correcting the nominal EC_{50} values for the free fraction (Figure 1).
[e] Calculated by multiplying the free EC_{50} values by the assay duration time.

serum free culture media should contain a certain amount of proteins to enable testing of hydrophobic and relatively unstable chemicals.

Another aspect to consider in the choice of an *in vitro* assay is the origin of the biological material. Many of the cellular *in vitro* developmental toxicity assays are based on mouse embryonic stem cells (Chapter 2, Table 1). In theory, human embryonic stem cell-based test systems may provide better predicted concentration-response effects for humans than the often used mouse embryonic stem cell-based test systems. There may be possible interspecies differences in dynamics in these cell systems, although the impact of these interspecies differences in dynamics on the obtained effect concentrations needs to be established. The ethical constraints and the relatively difficult culture methods met with the use of human embryonic stem cells may hamper their application in testing strategies. Comparison of effect concentrations of model compounds in human and mouse embryonic stem cell-based test systems is needed in order to determine which test system would be the most appropriate to apply in a testing strategy to estimate toxic dose levels for humans using the reverse dosimetry approach. Furthermore, the comparison of these data may also be useful to get insight in interspecies differences, which may be useful for the extrapolation of toxicity data obtained in laboratory animals to humans in present risk assessment strategies.

PBK model development

The development of PBK models accurately describing *in vivo* kinetics is indispensable in the reverse dosimetry approach to translate *in vitro* concentration-response curves into *in vivo* dose-response curves. PBK models are needed that reliably describe the *in vivo* kinetics. Development of these PBK models requires robust data for parameterization. Current PBK models often use chemical-specific parameter values describing the ADME processes, determined using *in vivo* kinetic data (Rietjens *et al.*, 2011). To limit the use of animals, *in vitro* and *in silico* approaches are preferred to determine these parameter values whenever possible. The parameter values for blood clearance of the toxic alkoxyacetic acid metabolites in the glycol ether PBK model needed to be determined using *in vivo* kinetic data (chapter 5). For the ATRA PBK model, all the kinetic model parameter values were determined using *in vitro* input data only. This was possible assuming that only

ATRA contributes to the developmental toxicity outcome, and that ATRA itself is only cleared by the liver. For a chemical for which information on the kinetics is lacking, the role of urinary (or biliary) clearance on the total clearance of the toxic moieties (parent compound and/or metabolite) is not known, indicating that these processes cannot beforehand be ignored.

If PBK models should be built using parameter values obtained using *in vitro*/*in silico* methods, additional effort in the development and evaluation of these methods is needed. Recently, an overview was made for the non-animal based methods that are available for obtaining kinetic parameter values (Adler *et al.*, 2011). It was concluded that for most kinetic data, non-animal based methods are available or present at an advanced development stage. However, no appropriate non-animal based methods are available for predicting renal and biliary excretion (Adler *et al.*, 2011). Therefore, the development of *in vitro* methods to estimate parameter values for renal and biliary clearance is of utmost importance. Furthermore, for the endpoint developmental toxicity, experimental models to estimate parameter values for placental transport of chemicals are of particular importance. Although the embryo levels of the chemicals used in the present thesis were reported to be close to the maternal blood levels, this is not necessarily the case for all chemicals. For certain chemicals, the placenta may acts as a barrier and affect the concentrations that will be reached in the conceptus, and with that the developmental toxicity outcome of the chemical. The use of *ex vivo* perfused placenta studies may be useful for the estimation of these parameter values for placental transport. However, to carry out these studies, one is dependent on the availability of (fresh) human placentas, which is a limitation of the approach. Cell culture models studying transport of chemicals across human placental cell layers mimicking the placental barrier may provide an easier system to obtain PBK model parameter values for placental transport (Van der Aa *et al.*, 1998). To enter the fetal circulation, chemicals need to cross a layer of trophoblast cells and the endothelial cells of the fetal blood vessels (Van der Aa *et al.*, 1998). Human choriocarcinoma (BeWo) cells representing the human trophoblast and human umbilical vein endothelial cells (HUVECs) representing the fetal endothelial barrier, can be cultured on transwell inserts, mimicking the placental barrier (Bode *et al.*, 2006; Kazakoff *et al.*, 1995). These *in vitro* cell systems may be useful to provide PBK

model parameter values for placental transfer of chemicals, in order to predict fetal exposure levels. Furthermore, *in silico* models predicting placental transfer may be used to estimate parameter values for placental transfer of chemicals and resulting fetal exposure (Hewitt *et al.*, 2007).

Another aspect that needs further consideration in PBK model development is the effect of repeated dosing on the *in vivo* kinetic characteristics. These effects need to be incorporated in PBK models, when predictions are to be made for repeated chemical exposure, if chemical exposure affects the expression levels of metabolizing enzymes. For glycol ethers, repeated dosing does not affect their *in vivo* kinetics. Therefore, prediction of *in vivo* kinetics of glycol ethers upon repeated dosing was feasible with the developed PBK model, as shown in chapter 5. For ATRA however, biotransformation enzymes are induced upon exposure (Collins *et al.*, 1995; Tay *et al.*, 2010). It is therefore not possible to make predicted dose-response curves for ATRA-induced developmental toxicity after repeated dosing if the effects of ATRA exposure on biotransformation enzyme expression levels are neglected. Therefore, only studies using single exposures were used for the evaluation of the predicted dose-response curves for ATRA-induced developmental toxicity in the present thesis (chapter 7). In order to make predictions for ATRA-induced developmental toxicity after repeated dosing, the change in expression levels of metabolizing enzymes should be incorporated in the model, to account for resulting changes in hepatic clearance. These required input data for the model should preferably be obtained using *in vitro* techniques. However, at present, data obtained from induction studies of biotransformation enzymes (in for example hepatocytes) is only usable for qualitative considerations, whereas these studies are not suitable to make quantitative predictions for the *in vivo* situation (Coecke *et al.*, 2006). This indicates that adequate *in vitro* models have to be further optimized and validated in order to make them suitable for making quantitative predictions on enzyme induction in the *in vivo* situation.

A major limitation of the use of PBK models in risk assessment is that the development of PBK models is a labor- and time-consuming process, and only applicable for data-rich chemicals for specific PBK models. In order to accelerate the process of PBK model development, generic PBK models are being developed (e.g. in the European Framework Program ChemScreen project and the United

States Environmental Protection Agency (US EPA)'s ToxCast program) for which step 1 (definition of the conceptual model) and step 2 (translation of the conceptual model into a mathematical model by defining mathematical equations) are made generic for chemicals belonging to a specific chemical category, or even for chemicals of diverse nature that do not belong to a specific chemical category. In that case no conceptual and no mathematical model need to be developed, implicating that 'only' the appropriate chemical-specific values for the model parameters defined in the generic PBK model need to be determined. Still, the determination of some model parameter values may be a labor- and time-consuming step. Therefore, the use of *in silico* modeling, such as quantitative structure activity relationship (QSAR) modeling should be considered for the prediction of model parameter values. At present *in silico* models are often used for the prediction of tissue:blood partition coefficients for PBK models. For these predictions the LogP value of the chemical and sometimes information on the free fraction of the chemical in blood is needed (DeJongh *et al.*, 1997; Poulin and Theil, 2002; Berezhkovskiy, 2004). For very hydrophobic chemicals these *in silico* models may result in large overestimations in predicted tissue:blood partition coefficients, as is the case for ATRA (data not shown). Therefore, the applicability domains of these *in silico* models should be taken into account and the potential of *in silico* modeling for estimating values for kinetic model parameters remains to be proven.

Translation of in vitro concentrations into in vivo doses

The translation of *in vitro* concentration-response curves obtained in *in vitro* assays into *in vivo* dose-response curves is not a straightforward process. Initially two important questions should be answered to be able to reliably predict *in vivo* dose-response curves:

Is the toxic outcome of chemical exposure related to the dose metric maximum concentration (C_{max}), or to the dose metric area under the concentration-time curve (AUC)?

Is the toxic effect resulting from chemical exposure caused by the parent compound itself and/or by (one of) its metabolite(s)?

Is the toxic outcome of chemical exposure related to the dose metric C_{max}, or to the dose metric AUC?

The extrapolation of the *in vitro* concentration-response curves for the alkoxyacetic acid metabolites in the ES-D3 cell differentiation assay to predicted dose-response curves for *in vivo* developmental toxicity of the parent glycol ethers was based on relating the developmental toxicity outcome of these chemicals to the C_{max} of the alkoxyacetic acid metabolites reached upon parent glycol ether exposure. This resulted in accurate predictions of *in vivo* dose-response curves for the rat. For ATRA however, predicted dose-response curves for *in vivo* developmental toxicity better corresponded to *in vivo* toxicity data when the toxic effect was related to the dose metric AUC, than when the toxic effect was related to the dose metric C_{max} (Chapter 7). These findings for ATRA are in line with the study reported by Tzimas *et al.* (1997), who found a better correlation when correlating various *in vivo* developmental toxicity endpoints to the AUC (correlation coefficients of 0.90-0.98) than when correlating the *in vivo* developmental toxicity endpoints to the C_{max} (correlation coefficients of 0.22-0.43).

To predict *in vivo* developmental toxicity by relating the toxic effects to the AUC, *in vitro* AUC values can be obtained by multiplying the applied *in vitro* effect concentration with the assay time (Daston *et al.*, 2010). It should be taken into account that the concentration of the test chemical in the exposure medium may not remain constant over time, resulting in overestimations of *in vitro* (calculated) AUC values (Figure 2). Therefore, the *in vitro* toxic potency may be underestimated and thus also the *in vivo* toxic potency, with the translation of the *in vitro* data to the *in vivo* situation. Furthermore, the sensitivity of the test system may not be constant during assay time either. This may especially be an issue in 'dynamic assays', such as the ES-D3 cell differentiation assay, for which during the whole culture period (10 days) different cell types with possible differences in sensitivity to chemicals are present. Theoretically, it may therefore be possible that for a certain chemical, the assay is sensitive from the start of the assay until day 10, whereas for a second or a third chemical the assay may be sensitive from the start of the assay until day 3 or day 5.

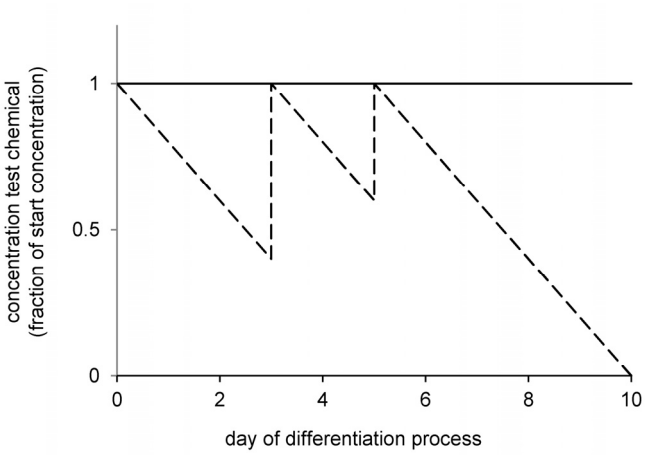

Figure 2. Hypothetical concentration of test chemicals in the ES-D3 cell differentiation assay for a chemical that is stable for the whole test duration (straight line) and for a chemical that is unstable (dashed line). Medium, including test chemical, is replaced at day 3 and day 5 of the assay.

In the *in vivo* situation, the conceptus may also differ in sensitivity to chemicals during different developmental stages. Differences in sensitive periods in the *in vitro* ES-D3 cell differentiation might reflect the sensitive periods in the *in vivo* developing conceptus. If so, the ES-D3 cell differentiation assay might be used to predict sensitive periods of the conceptus to a chemical in the *in vivo* situation. Van Dartel *et al.* (2009) exposed differentiating ES-D3 cells to varying test chemicals during different time frames of the differentiation assay, to assess the effect of the exposure period and exposure duration on cardiac differentiation. In general, it was shown that a longer exposure period resulted in lower observed effect concentrations. For different chemicals, differences in sensitive periods were observed, which may be related to differences in their modes of action (Van Dartel *et al.*, 2009).

In order to assess whether the toxic effect of a chemical is best related to the C_{max} or the AUC, or even another dose metric such as an AUC above a certain threshold, the effect of exposure duration on the test outcome could be investigated. In a first attempt, this was done for MAA and ATRA in the ES-D3 cell differentiation assay (Figure 3). To this end, EBs were exposed from day 3 for 4, 8, 24 or 48 hours to 1, 10 or 100 mM MAA (Figure 3A) or to 1, 10 or 100 nM ATRA (Figure 3B). Figure 3 shows that MAA-induced inhibition of ES-D3 cell differentiation is hardly dependent on

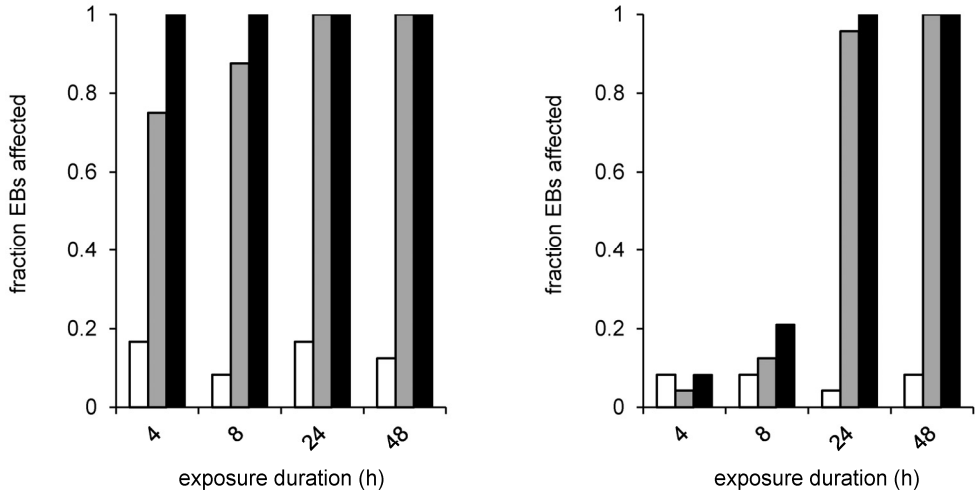

Figure 3. Effect of exposure duration of MAA (A) and ATRA (B) on ES-D3 cell differentiation into contracting cardiomyocytes. Cells were exposed to 1 mM MAA (A, white bars), 10 mM MAA (A, grey bars), 100 mM MAA (A, black bars), 1 nM ATRA (B, white bars), 10 nM ATRA (B, grey bars) or 100 nM ATRA (B, black bars). Exposure started when embryoid bodies (EBs) were cultured in petri dishes (day 3) and lasted 4, 8, 24 or 48 h.

the exposure duration, whereas the potency of ATRA clearly increases with longer exposure durations. This may indicate that MAA-induced toxicity can be best related to the C_{max}, whereas ATRA-induced toxicity can be best related to the AUC. Performing the ES-D3 cell differentiation assay by using varying exposure durations may therefore be useful to predict whether toxic effects are related to the C_{max} or to the AUC.

Is the toxic effect resulting from chemical exposure caused by the parent compound itself and/or by (one of) its metabolite(s)?

The (toxic) effect of a chemical may be caused by the parent chemical itself and/or by (one of) its metabolite(s). Chemicals that cause developmental toxicity by (one of) their metabolite(s) are called proteratogens, which need to be bioactivated in order to induce toxicity. In *in vitro* test systems, the expression levels of biotransformation enzymes may be too low to convert the parent compound to a significant extent (Coecke *et al.*, 2006). If this is the case, the parent compound will be less potent in the *in vitro* test system than its toxic metabolite(s). If for proteratogens the concentration-response curve obtained for the parent chemical

in the *in vitro* assay is translated into *in vivo* dose levels using reverse dosimetry, it can be expected that the predicted toxic doses are underestimating the *in vivo* potency of the chemical. An example of the implications of this underestimation of predicted developmental toxicity dose levels is shown in Figure 4 for the glycol ether EGME. It shows that the translation of the concentration-response curve for MAA-induced inhibition of ES-D3 cell differentiation to the *in vivo* situation accurately predicts measured *in vivo* EGME-induced developmental toxicity as reported in the literature. On the other hand, when the concentration-response curve for EGME-induced inhibition of ES-D3 cell differentiation is translated to the *in vivo* situation, EGME's developmental toxicity would be largely underestimated (Figure 4). This indicates that the *in vivo* developmental toxicity is caused by EGME's metabolite MAA and not by EGME itself, as reported in the literature.

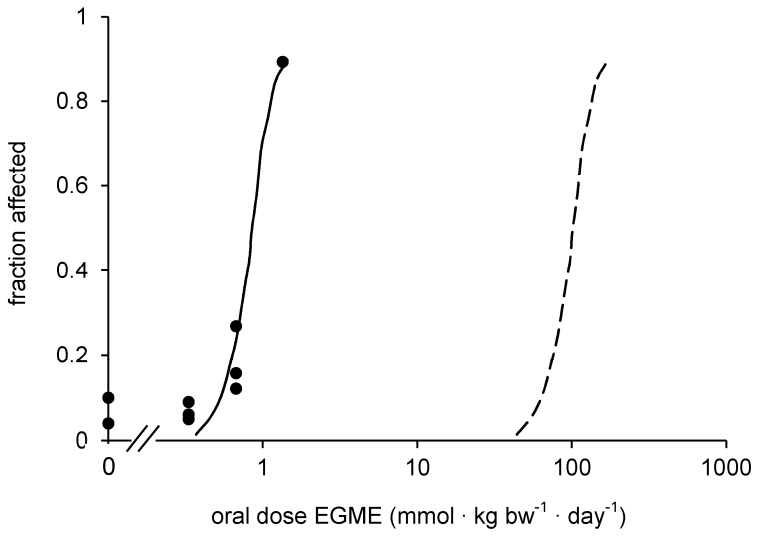

Figure 4. Predicted dose-response curves for EGME-induced developmental toxicity based on the translation of the concentration-response curve for MAA-induced (straight line) and EGME-induced (dashed line) inhibition of ES-D3 cell differentiation using PBK modeling with reverse dosimetry. The toxic outcome was related the dose metric C_{max} of MAA and EGME, respectively. Data on MAA-induced inhibition of ES-D3 cell differentiation were taken from De Jong *et al.* (2009), whereas data on EGME-induced inhibition of ES-D3 cell differentiation were taken from Verwei *et al.* (2006). The circles represent measured *in vivo* data on EGME-induced developmental toxicity reported in the literature (Toraason *et al.*, 1985).

The addition of a metabolic system may increase the *in vitro* potency, but may also decrease the potency of the proteratogen. Whether an increase or a decrease in the potency will be observed with the addition of a metabolic system is dependent on the overall outcome of bioactivation and detoxification reactions and the concentration of the toxic metabolite(s) achieved. Furthermore, the addition of cellular fractions as a bioactivation system can itself cause toxicity in cellular systems (Tan *et al.*, 1982) and WECs (Luijten *et al.*, 2008). To investigate whether a new chemical is expected to be a proteratogen, the metabolites formed may be identified (e.g., using hepatocytes in culture, or liver fractions (microsomes, S9) and co-factors) and tested (e.g., the main metabolites) in the *in vitro* test system. The *in vivo* formation of the metabolites can be described in the PBK model, after which parent compound exposure levels resulting in *in vivo* developmental toxicity can be predicted using reverse dosimetry, as was done in the present thesis for the glycol ethers. Although metabolite identification (whether or not in combination with metabolite production and testing) is a standard procedure in the drug development process, this may be a too expensive and time-consuming process to apply in testing strategies for REACH chemicals.

Future perspectives and conclusions for the use of the combined *in vitro-in silico* approach in risk assessment

The present thesis provides the proof-of-principle that *in vitro* concentration-response curves can be converted into *in vivo* dose-response curves for developmental toxicity, using PBK modeling with a reverse dosimetry approach. In order to increase the confidence of using the approach for risk assessment, more proof-of-principle studies with other chemicals for developmental toxicity and more studies for other endpoints are needed. At present, a few examples using reverse dosimetry for predicting *in vivo* toxicity are reported in the literature (Punt *et al.*, 2011). Reverse dosimetry-based predictions have been made for a limited number of chemicals for the toxicological endpoints acute and repeated dose toxicity, neurotoxicity, genotoxicity and developmental toxicity (Punt *et al.*, 2011). Reverse dosimetry approaches using PBK modeling are also being used in the US EPA's ToxCast program and the European Framework Program projects ESNATS and ChemScreen to predict *in vivo* effect levels based on *in vitro* effect

concentrations. As more examples of the proof-of-principle for other chemicals and other endpoints will be provided, it can be considered to apply the approach in testing strategies for toxicological risk assessment. All the more, because the approach fits well within the toxicity testing strategies as envisioned in the in 2007 released report *'Toxicity Testing in the 21st century: A Vision and a Strategy'* by the United States National Academy of Sciences (NAS) (NRC, 2007). This report envisions a future in which all routine toxicity testing will be conducted using *in vitro* assays by evaluating cellular responses on chemical exposure. In the newly proposed risk assessment strategy, it will be estimated based on *in vitro* models and PBK modeling at which exposure levels adverse cellular perturbations will occur in the *in vivo* situation (Andersen and Krewski, 2009). PBK modeling using a reverse dosimetry approach can highly contribute to these estimations.

The present thesis used reverse dosimetry by translating a complete *in vitro* concentration-response curve into a complete *in vivo* dose-response curve using PBK models accurately describing the *in vivo* kinetics of the chemical of interest. The predicted dose-response curves may be used to set points of departure for the risk assessment to determine acceptable daily intake values (ADIs). However, this approach using complex PBK models may be too time-consuming to apply for all the chemicals that have not been evaluated yet, and may therefore not be suitable to use for all chemicals. Therefore, a pre-selection of chemicals using simple generic PBK models as used in the ToxCast program may be used to identify the chemicals of highest concern to prioritize for further testing. The ToxCast program, initiated within the EPA, aims at the development of methods for prioritizing (environmental) chemicals for further screening and testing. In the approach used, the test chemical is tested in a battery of more than 500 *in vitro* test assays, after which the potency of the chemical in each positive assay is summarized using AC_{50} (concentration at 50% of maximum activity) or lowest effective concentration (LEC) values, depending on the type of concentration-response data that have been collected for the assay. *In vitro* assays are performed to estimate hepatic metabolic clearance and plasma protein binding in humans. These *in vitro* kinetic data are incorporated in the generic PBK model used for the *in vitro-in vivo* extrapolations to calculate a daily human oral dose, called the oral equivalent dose, that would be required to produce an *in vivo* blood concentration of the chemical equivalent to the *in vitro* AC_{50} or LEC value (Rotroff *et al.*, 2010; Wetmore *et al.*, 2012). If data on

exposure levels are available, the estimated oral equivalent dose is compared with human oral exposure estimates to assess whether an *in vivo* effect would be expected (Rotroff *et al.*, 2010; Wetmore *et al.*, 2012). If so, these chemicals may become of priority for further risk assessment.

The chemicals that would be prioritized for further risk assessment, could be assessed in more detail using reverse dosimetry with more refined *in vitro* toxicity assays and PBK models, as was done in the present thesis. These PBK models could be used to accurately predict *in vivo* toxic dose levels for sensitive individuals within the human population, by describing interindividual differences in kinetics in the model using Monte Carlo simulations (chapter 7). The predicted dose-response curves could be used to set points of departure to determine safe exposure levels. If PBK modeling with a reverse dosimetry approach, translating *in vitro* concentration-response curves into *in vivo* dose-response curves, would be used for human risk assessment, only human PBK models need to be developed, implicating that no uncertainty factor for interspecies differences in kinetics would be needed. If human interindividual differences in kinetics are incorporated in the PBK model (chapter 7), no uncertainty factor for these interindividual differences in kinetics need to be applied either. Furthermore, if the *in vitro* model used to determine the *in vitro* concentration-response curves would be based on human cells instead of mouse cells, no uncertainty factor for interspecies differences in dynamics would be needed, leaving only the need for uncertainty factors accounting for human interindividual differences in dynamics. On the other hand, an uncertainty factor may be introduced for uncertainties in the accuracy of predicted dose-response effects using the reverse dosimetry approach. For the chemicals used in the present thesis, differences in predicted dose-response effects and dose-response effects measured in reported *in vivo* rat studies were within one order of magnitude, indicating that an uncertainty factor of 10 may be used for uncertainties in the predictions. However, more data are needed to determine an appropriate value for this uncertainty factor.

In conclusion, the present thesis shows the feasibility of translating concentration-response curves obtained in *in vitro* toxicity assays into predicted *in vivo* dose-response curves, which can be used for risk assessment practice. The work presented in this thesis provides the first proof that *in vivo* developmental

toxicity dose levels can be accurately predicted by combining only *in vitro* toxicity data and *in silico* PBK modeling, using a reverse dosimetry approach. This work may therefore contribute to an increased acceptance and future implementation of *in vitro* toxicity data in risk assessment practice and may even be considered as the most important approach to achieve the objectives of the 3Rs of animal use in toxicological risk assessment.

References

Adamson, P. C., Pitot, H. C., Balis, F. M., Rubin, J., Murphy, R. F., Poplack, D. G. (1993). Variability in the oral bioavailability of all-trans-retinoic acid. *J Natl Cancer Inst* **85**, 993-6.

Adler, S., Basketter, D., Creton, S., Pelkonen, O., van Benthem, J., Zuang, V., Andersen, K. E., Angers-Loustau, A., Aptula, A., Bal-Price, A., Benfenati, E., Bernauer, U., Bessems, J. G. M., Bois, F. Y., Boobis, A., Brandon, E., Bremer, S., Broschard, T., Casati, S., Coecke, S., Corvi, R., Cronin, M., Daston, G. P., Dekant, W., Felter, S., Grignard, E., Gundert-Remy, U., Heinonen, T., Kimber, I., Kleinjans, J., Komulainen, H., Kreiling, R., Kreysa, J., Leite, S. B., Loizou, G., Maxwell, G., Mazzatorta, P., Munn, S. J., Pfuhler, S., Phrakonkham, P., Piersma, A. H., Poth, A., Prieto, P., Repetto, G., Rogiers, V., Schoeters, G., Schwarz, M., Serafimova, R., Tähti, H., Testai, E., van Delft, J. H., van Loveren, H., Vinken, M., Worth, A., Zaldivar, J. M. (2011). Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Arch Toxicol* **85**, 367-485.

Andersen, M. E., Krewski, D. (2009). Toxicity testing in the 21st century: bringing the vision to life. *Toxicol Sci* **107**: 324-330.

Berezhkovskiy, L. M. (2004). Volume of distribution at steady state for a linear pharmacokinetic system with peripheral elimination. *J Pharm Sci* **93**, 1628-1640.

Bode, C. J., Jin, H., Rytting, E., Silverstein, P. S., Young, A. M., Audus, K. L. (2006). In vitro models for studying trophoblast transcellular transport. *Methods Mol Med* **122**, 225-39.

Breier, J. M., Gassmann, K., Kayser, R., Stegeman, H., de Groot, D., Fritsche, E., Shafer, T. J. (2010). Neural progenitor cells as models for high-throughput screens of developmental neurotoxicity: State of the science. *Neurotoxicol Teratol* **32**, 4-15.

Coecke, S., Ahr, H., Blaauboer, B. J., Bremer, S., Casati, S., Castell, J., Combes, R., Corvi, R., Crespi, C. L., Cunningham, M. L., Elaut, G., Eletti, B., Freidig, A. P., Gennari, A., Ghersi-Egea, J. F., Guillouzo, A., Hartung, T., Hoet, P., Ingelman-Sundberg, M., Munn, S. J., Janssens, W., Ladstetter, B., Leahy, D., Long, A., Meneguz, A., Monshouwer, M., Morath, S., Nagelkerke, F., Pelkonen, O., Ponti, J., Prieto, P., Richert, L., Sabbioni, E., Schaack, B., Steiling, W., Testai, E., Vericat, J. A., Worth, A. (2006). Metabolism: a bottleneck in *in vitro* toxicological test development. The report and recommendations of ECVAM workshop 54. *Altern Lab Anim* **34**, 49-84.

Collins, M. D., Mao, G. E. (1999). Teratology of retinoids. *Annu Rev Pharmacol Toxicol* **39**, 399-430.

Collins, M. D., Tzimas, G., Bürgin, H., Hummler, H., Nau, H. (1995). Single versus multiple dose administration of all-trans-retinoic acid during organogenesis: differential metabolism and transplacental kinetics in rat and rabbit. *Toxicol Appl Pharmacol* **130**, 9-18.

Daston, G. P., Chapin, R. E., Scialli, A. R., Piersma, A. H., Carney, E. W., Rogers, J. M., Friedman. J. M. (2010). A different approach to validating screening assays for developmental toxicity. *Birth Defects Res B Dev Reprod Toxicol* **89**, 526-30.

De Jong, E., Louisse, J., Verwei, M., Blaauboer, B. J., van de Sandt, J. J. M., Woutersen, R. A., Rietjens, I. M. C. M., Piersma, A. H. (2009). Relative developmental toxicity of glycol ether alkoxy acid metabolites in the embryonic stem cell test as compared to the in vivo potency of their parent compounds. *Toxicol Sci* **110**, 117-24.

DeJongh, J., Nordin-Andersson, M., Ploeger, B. A., Forsby A. (1999). Estimation of systemic toxicity of acrylamide by integration of in vitro toxicity data with kinetic simulations. *Toxicol Appl Pharmacol* **158**, 261-68.

DeJongh, J., Verhaar, H. J., Hermens, J. L. (1997). A quantitative property-property relationship (QPPR) approach to estimate in vitro tissue-blood partition coefficients of organic chemicals in rats and humans. *Arch Toxicol* **72**, 17-25.

Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W., Kemler, R. (1985). The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J Embryol Exp Morph* **87**, 27-45.

ECETOC. (2005). The toxicology of glycol ethers and its relevance to man. Technical Report 095, 4th ed. European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels, Belgium.

Fesstag, M., Sehner, C., Steinberg, P., Viertel, B. (2007). An *in vitro* embryotoxicity assay based on the disturbance of the differentiation of murine embryonic stem cells into endothelial cells. II: testing of chemicals. *Toxicol In Vitro* **21**, 1631-40.

Forsby, A., Blaauboer, B. J. (2007). Integration of in vitro neurotoxicity data with biokinetic modelling for the estimation of in vivo neurotoxicity. *Hum Exp Toxicol* **26**, 333-38.

Genschow, E., Spielmann, H., Scholz, G., Pohl, I., Seiler, A., Clemann, N., Bremer, S., Becker, K. (2004). Validation of the embryonic stem cell test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* **32**, 209-44.

Genschow, E., Spielmann, H., Scholz, G., Seiler, A., Brown, N. A., Piersma, A. H., Brady, M., Clemann, N., Huuskonen, H., Paillard, F., Bremer, S., Becker, K. (2002). The ECVAM international validation study on in vitro embryotoxicity tests: results of the definitive phase and evaluation of prediction models. European Centre for the Validation of Alternative Methods. *Atern Lab Anim* **30**, 151-76.

Jochems, C. E., van der Valk, J. B., Stafleu, F. R., Baumans V. (2002). The use of fetal bovine serum: ethical or scientific problem? *Altern Lab Anim* **30**, 219-27.

Hewitt, M., Madden, J. C., Rowe, P. H., Cronin, M. T. (2007). Structure-based modelling in reproductive toxicology: (Q)SARs for the placental barrier. *SAR QSAR Environ Res* **18**, 57-76.

Kazakoff, P. W., McGuire, T. R., Hoie, E. B., Cano, M., Iversen, P. L. (1995). An in vitro model for endothelial permeability: assessment of monolayer integrity. *In Vitro Cell Dev Biol Anim* **31**, 846-52.

Kistler, A. (1987). Limb bud cell cultures for estimating the teratogenic potential of chemicals. Validation of the test system with retinoids. *Arch Toxicol* **60**, 403-14.

Klaassen, I., Brakenhoff, R. H., Smeets, S. J., Snow, G. B., Braakhuis, B. J. (1999). Considerations for in vitro retinoid experiments: importance of protein interaction. *Biochim Biophys Acta* **1427**, 265-75.

Kuegler, P. B., Zimmer, B., Waldmann, T., Baudis, B., Ilmjärv, S., Hescheler, J., Gaughwin, P., Brundin, P., Mundy, W., Bal-Price, A. K., Schrattenholz, A., Krause, K. H., van Thriel, C., Rao, M. S., Kadereit, S., Leist, M. (2010). Markers of murine embryonic and neural stem cells, neurons and astrocytes: reference points for developmental neurotoxicity testing. *ALTEX* **27**, 17-42.

Luijten, M., Verhoef, A., Westerman, A., Piersma, A. H. (2008). Application of a metabolizing system as an adjunct to the rat whole embryo culture. *Toxicol In Vitro* **22**, 1332-36.

Mainali, A. S., Tew, J. G. (2004). Dexamethasone selectively inhibits differentiation of cord blood stem cell derived-dendritic cell (DC) precursors into immature DCs. *Cell Immunol* **232**, 127-36.

National Research Council (NRC). (2007). Toxicity testing of the 21st century: a vision and a strategy. National Academy Press, Washington, DC.

O'Flaherty, E. J., Nau, H., McCandless, D., Beliles, R. P., Schreiner, C. M., Scott, W. J. Jr. (1995). Physiologically based pharmacokinetics of methoxyacetic acid: dose-effect considerations in C57BL/6 mice. *Teratology* **52**, 79-89.

Piersma, A. H. (2004). Validation of alternative methods for developmental toxicity testing. *Toxicol Lett* **149**, 147-53.

Piersma, A. H. (2006). Alternative methods for developmental toxicity testing. *Basic Clin Pharmacol Toxicol* **98**, 427-31.

Piersma, A. H., Genschow, E., Verhoef, A., Spanjersberg, M. Q., Brown, N. A., Brady, M., Burns, A., Clemann, N., Seiler, A., Spielmann, H. (2004). Validation of the postimplantation rat whole-embryo culture test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* **32**, 275-307.

Poulin, P., Theil, F. P. (2002). Prediction of pharmacokinetics prior to in vivo studies. II. Generic physiologically based pharmacokinetic models of drug disposition. *J Pharm Sci* **91**, 1358-70.

Punt, A., Schiffelers, M. J., Jean Horbach, G., van de Sandt, J. J. M., Groothuis, G. M., Rietjens, I. M. C. M., Blaauboer, B. J. (2011). Evaluation of research activities and research needs to increase the impact and applicability of alternative testing strategies in risk assessment practice. *Regul Toxicol Pharmacol* **61**, 105-14.

Riebeling, C., Schlechter, K., Buesen, R., Spielmann, H., Luch, A., Seiler, A. (2011). Defined culture medium for stem cell differentiation: applicability of serum-free conditions in the mouse embryonic stem cell test. *Toxicol In Vitro* **25**, 914-21.

Rietjens, I. M. C. M., Louisse, J., Punt, A. (2011). Tutorial on physiologically based kinetic modeling in molecular nutrition and food research. *Mol Nutr Food Res* **55**, 941-56.

Rotroff, D. M., Wetmore, B. A., Dix, D. J., Ferguson, S. S., Clewell, H. J., Houck, K. A., Lecluyse, E. L., Andersen, M. E., Judson, R. S., Smith, C. M., Sochaski, M. A., Kavlock, R. J., Boellmann, F., Martin, M. T., Reif, D. M., Wambaugh, J. F., Thomas, R. S. (2010). Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening. *Toxicol Sci* **117**, 348-58.

Scott, W. J., Duggan, C. A., Schreiner, C. M., Collins, M. D. (1990). Reduction of embryonic intracellular pH: a potential mechanism of acetazolamide-induced limb malformations. *Toxicol Appl Pharmacol* **103**, 238-54.

Scott, W. J. Jr., Schreiner, C. M., Nau, H., Vorhees, C. V., Beliles, R. P., Colvin, J., McCandless, D. (1997). Valproate-induced limb malformations in mice associated with reduction of intracellular pH. *Reprod Toxicol* **11**, 483-93.

Shuey, D. L., Lau, C., Logsdon, T. R., Zucker, R. M., Elstein, K. H., Narotsky, M. G., Setzer, R. W., Kavlock, R. J., Rogers, J. M. (1994). Biologically based dose-response modeling in developmental toxicology: biochemical and cellular sequelae of 5-fluorouracil exposure in the developing rat. *Toxicol Appl Pharmacol* **126**, 129-44.

Spielmann, H., Genschow, E., Brown, N. A., Piersma, A. H., Verhoef, A., Spanjersberg, M. Q., Huuskonen, H., Paillard, F., Seiler, A. (2004). Validation of the rat limb bud micromass test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* **32**, 245 -74.

Stumann, T. C., Hareng, L., Bremer, S. (2007). Embryotoxicity hazard assessment of methylmercury and chromium using embryonic stem cells. *Toxicology* **242**, 130-43.

Sweeney, L. M., Tyler, T. R., Kirman, C. R., Corley, R. A., Reitz, R. H., Paustenbach, D. J., Holson, J. F., Whorton, M. D., Thompson, K. M., Gargas, M. L. (2001). Proposed occupational exposure limits for select ethylene glycol ethers using PBPK models and Monte Carlo simulations. *Toxicol Sci* **62**, 124-39.

Tan, E. L., Schenley, R. L., Hsie, A. W. (1982). Microsome-mediated cytotoxicity to CHO cells. *Mutat Res* **103**, 359-65.

Tay, S., Dickmann, L., Dixit, V., Isoherranen, N. (2010). A comparison of the roles of peroxisome proliferator-activated receptor and retinoic acid receptor on CYP26 regulation. *Mol Pharmacol* **77**, 218-27.

Theunissen, P. T., Pennings, J. L. A., Robinson J. F., Claessen, S. M., Kleinjans, J. C., Piersma, A. H. (2011). Time-response evaluation by transcriptomics of methylmercury effects on neural differentiation of murine embryonic stem cells. *Toxicol Sci* **122**, 437-47.

Theunissen, P. T., Schulpen, S. H. W., van Dartel, D. A. M., Hermsen, S. A. B., van Schooten, F. J., Piersma, A. H. (2010). An abbreviated protocol for multilineage neural differentiation of murine embryonic stem cells and its perturbation by methyl mercury. *Reprod Toxicol* **29**, 383-92.

Toraason, M., Stringer, B., Stober, P., Hardin, B. D. (1985). Electrocardiographic study of rat fetuses exposed to ethylene glycol monomethyl ether (EGME). *Teratology* **32**, 33-9.

Toumadje, A., Kusumoto, K., Parton, A., Mericko, P., Dowell, L., Ma, G., Chen, L., Barnes, D. W., Sato, J. D. (2003). Pluripotent differentiation in vitro of murine ES-D3 embryonic stem cells. *In vitro Cell Dev Biol Anim* **39**, 449-53.

Tzimas, G., Thiel, R., Chahoud, I., Nau, H. (1997). The area under the concentration-time curve of all-trans-retinoic acid is the most suitable pharmacokinetic correlate to the embryotoxicity of this retinoid in the rat. *Toxicol Appl Pharmacol* **143**, 436-44.

Van Dartel, D. A. M., Pennings, J. L. A., de la Fonteyne, L. J. J., Brauers, K. J., Claessen, S., van Delft, J. H., Kleinjans, J. C., Piersma, A. H. (2011). Concentration-dependent gene expression responses to flusilazole in embryonic stem cell differentiation cultures. *Toxicol Appl Pharmacol* **251**, 110-18.

Van Dartel, D. A. M., Zeijen, N. J., de la Fonteyne, L. J. J., van Schooten, F. J., Piersma, A. H. (2009). Disentangling cellular proliferation and differentiation in the embryonic stem cell test, and its impact on the experimental protocol. *Reprod Toxicol* **28**, 254-61.

Van Dartel, D. A. M., Piersma, A. H. (2011). The embryonic stem cell test combined with toxicogenomics as an alternative testing model for the assessment of developmental toxicity. *Reprod Toxicol* **32**, 235-44.

Van der Aa, E. M., Peereboom-Stegeman, J. H., Noordhoek, J., Gribnau, F. W., Russel, F. G. (1998). Mechanisms of drug transfer across the human placenta. *Pharm World Sci* **20**, 139-48.

Van der Valk, J., Mellor, D., Brands, R., Fischer, R., Gruber, F., Gstraunthaler, G., Hellebrekers, L., Hyllner, J., Jonker, F. H., Prieto, P., Thalen, M., Baumans, V. (2004). The humane collection of fetal bovine serum and possibilities for serum-free cell and tissue culture. *Toxicol In Vitro* **18**, 1-12.

Van der Jagt, K., Munn, S. J., Tørsløv, J., de Bruijn, J. (2004). Alternative approaches can reduce the use of test animals under REACH. *Report EUR* **21405**.

Verwei, M., van Burgsteden, J. A., Krul, C. A. M., van de Sandt, J. J. M., Freidig, A. P. (2006). Prediction of in vivo embryotoxic effect levels with a combination of in vitro studies and PBPK modelling. *Toxicol Lett* **165**, 79-87.

Wetmore, B. A., Wambaugh, J. F., Ferguson, S. S., Sochaski, M. A., Rotroff, D. M., Freeman, K., Clewell, H. J. 3rd, Dix, D. J., Andersen, M. E., Houck, K. A., Allen, B., Judson, R. S., Singh, R., Kavlock, R. J., Richard, A. M., Thomas, R. S. (2012). Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. *Toxicol Sci* **125**, 157-74.

Zimmer, B., Kuegler, P. B., Baudis, B., Genewsky, A., Tanavde, V., Koh, W., Tan, B., Waldmann, T., Kadereit, S., Leist, M. (2011). Coordinated waves of gene expression during neuronal differentiation of embryonic stem cells as basis for novel approaches to developmental neurotoxicity testing. *Cell Death Differ* **18**, 383-95.

Zimmer, B., Schildknecht, S., Kuegler, P. B., Tanavde, V., Kadereit, S., Leist, M. (2011). Sensitivity of dopaminergic neuron differentiation from stem cells to chronic low-dose methylmercury exposure. *Toxicol Sci* **121**, 357-67.

Zur Nieden, N. I., Davis, L. A., Rancourt, D. E. (2010a). Comparing three novel endpoints for developmental osteotoxicity in the embryonic stem cell test. *Toxicol Appl Pharmacol* **247**, 91-97.

Zur Nieden, N. I., Davis, L. A., Rancourt, D. E. (2010b). Monolayer cultivation of osteoprogenitors shortens duration of the embryonic stem cell test while reliably predicting developmental osteotoxicity. *Toxicology* **277**, 66-73.

CHAPTER 9

Samenvatting, toekomstperspectieven en conclusies

Samenvatting

Het doel van dit proefschrift is om het 'proof-of-principle' te verschaffen dat *in vitro* concentratie-respons curves kunnen worden omgezet naar *in vivo* dosis-respons curves door gebruik te maken van fysiologisch gebaseerd kinetisch (PBK)-modelleren met een omgekeerde-dosimetrie-aanpak, om alternatieve methoden te verschaffen voor *in vivo* studies in de toxicologische risicobeoordeling. Als toxicologisch eindpunt voor deze proof-of-principle-studie is ontwikkelings-toxicologie gekozen, omdat er geschat wordt dat meer dan een kwart van de verwachte aantallen dieren die nodig zijn voor de Europese REACH (Registratie, Evaluatie, Autorisatie en restrictie van CHemicaliën)-wetgeving gebruikt zal worden voor dit type studies (Van der Jagt *et al.*, 2004). Ook voor farmaceutica, voedseladditieven en pesticiden vereisen regelgevende instanties veiligheidsstudies, inclusief ontwikkelingstoxiciteitsstudies, in dieren. De toepassing van betrouwbare alternatieven voor *in vivo* ontwikkelings-toxiciteitsstudies zou daarom een significante bijdrage leveren aan de vervanging, vermindering en verfijning (3V's) van diergebruik in de toxicologische risicobeoordeling.

Hoofdstuk 1 van dit proefschrift geeft achtergrondinformatie waarmee het onderwerp wordt geïntroduceerd en definieert het doel van het proefschrift. Het beschrijft belangrijke factoren die nodig zijn om het concept te ontwikkelen waarbij *in vitro* concentratie-respons curves worden vertaald naar voorspelde *in vivo* dosis-respons curves voor ontwikkelingstoxiciteit, door gebruik te maken van PBK-modellering met een omgekeerde-dosimetrie-aanpak. In het kort omvatten deze factoren 1) een *in vitro* model om *in vitro* concentratie-respons curves voor ontwikkelingstoxiciteit te verkrijgen, 2) de selectie van de te bestuderen chemicaliën en 3) de ontwikkeling van PBK-modellen die nodig zijn voor omgekeerde dosimetrie.

Het *in vitro* model dat is gebruikt is de embryonale stamcel (ES-D3 cel) differentiatietest van de gevalideerde embryonale stamceltest (EST) (Genschow *et al.*, 2004). In deze assay wordt het effect van de stof van interesse op de differentiatie van ES-D3 cellen naar functionele (contracterende) hartspiercellen vastgesteld. De door de chemicalie geïnduceerde remming van ES-D3-celdifferentiatie is het *in vitro* eindpunt voor ontwikkelingstoxicitieit. De ES-D3-celdifferentiatietest is de enige gevalideerde *in vitro* ontwikkelingstoxiciteitsassay

waarvoor geen primair dierweefsel nodig is (Genschow *et al.*, 2002; Genschow *et al.*, 2004). Alternatieven voor *in vivo* ontwikkelingstoxiciteit die gebruik maken van deze *in vitro* assay zouden daarom resulteren in de grootste bijdrage aan de 3V's van diergebruik voor het testen voor ontwikkelingstoxiciteit.

Om het 'proof-of-principle' te verschaffen dat de verkregen *in vitro* concentratie-respons curves in de ES-D3-celdifferentiatietest kunnen worden omgezet naar *in vivo* dosis-respons curves, door gebruik te maken van PBK-modellering met een omgekeerde-dosimetrie-aanpak, moesten geschikte modelstoffen worden geselecteerd die aan bepaalde criteria voldoen. Eerst werd een lijst chemicaliën samengesteld die aan twee criteria voldoen, welke zijn 1) de aanwezigheid van *in vivo* kinetische data van de stof in de literatuur, om het bouwen en/of evalueren van de PBK-modellen mogelijk te maken, die ontwikkeld en gebruikt worden voor omgekeerde dosimetrie, en 2) de aanwezigheid van *in vivo* ontwikkelings-toxiciteitsdata van de stof in de literatuur, om te evalueren of de uiteindelijk voorspelde dosis-respons curves voor ontwikkelingstoxiciteit in overeenstemming zijn met data verkregen in *in vivo* ontwikkelingstoxiciteitsstudies. Van deze lijst van geschikte chemicaliën werden alleen structureel verwante chemicaliën geselecteerd. Dit derde criterium maakte de beoordeling van de relatieve verschillen binnen een groep chemicaliën mogelijk, wat daarbij een extra mogelijkheid voor een kwalitatieve evaluatie van de voorspellingen verschaft. Twee series modelstoffen werden gelecteerd, welke zijn 1) een groep glycol ethers, die gebioactiveerd moeten worden om ontwikkelingstoxiciteit te veroorzaken, die hydrofiel zijn, een lage mate van eiwitbinding vertonen en een niet-specifiek werkingsmechanisme hebben, en 2) een groep retinoïden, waarvan verscheidene stoffen direct werken als teratogenen zonder bioactivatie, die hydrofoob zijn, een hoge mate van eiwitbinding vertonen en die een specifiek werkingsmechanisme hebben.

Om de *in vitro* concentratie-respons curves verkregen in de ES-D3-celdifferentiatietest te vertalen naar *in vivo* dosis-respons curves met omgekeerde dosimetrie moesten PBK-modellen worden ontwikkeld die de ADME (Absorptie, Distributie, Metabolisme en Excretie)-processen van de stoffen van interesse beschrijven. Het proces van PBK-modelontwikkeling is beschreven in hoofdstuk 1 en kan worden opgedeeld in de volgende stappen: 1) het definiëren van een conceptueel model, 2) het vertalen van het conceptuele model naar een wiskundig model door het definiëren van wiskundige vergelijkingen, 3) het bepalen van de

waardes voor de modelparameters, 4) het oplossen van de wiskundige vergelijkingen met geschikte software en 5) het evalueren van de modelprestaties. Vervolgens kan het model worden gebruikt om voorspellingen te maken (stap 6), bijvoorbeeld om dosis-afhankelijke verschillen in kinetiek tussen of binnen soorten te voorspellen, of, zoals gedaan is in dit proefschrift, om interne concentraties, ingesteld als input in het model, te vertalen naar voorspelde externe doses (dat wil zeggen omgekeerde dosimetrie).

In **hoofdstuk 2** worden *in vitro* testen gepresenteerd die zijn beschreven in de literatuur voor het eindpunt ontwikkelingstoxiciteit. Deze testen kunnen worden opgedeeld in 1) cellulaire testen die gebruik maken van primaire cellen of cellen van continue cellijnen, en 2) hele-embryo-kweek (WEC)-testen. Er wordt een overzicht gegeven van de verscheidene uitleesparameters, variërend van veranderingen in gen- of eiwitexpressie tot celdood of *in vitro* malformaties, die zijn onderzocht in deze *in vitro* testen. Hoofdstuk 2 beschrijft ook hoe *in vitro* ontwikkelingstoxiciteitsdata zouden kunnen worden gebruikt volgens de literatuur om voorspellingen te maken voor de *in vivo* situatie. Dit kan op verschillende manieren worden gedaan: de *in vitro* data zou kunnen worden gebruikt om 1) de ontwikkelingstoxiciteitsklasse waartoe een stof behoort (bijvoorbeeld niet, zwak of sterk embryotoxisch) te voorspellen met behulp van een predictiemodel, 2) de relatieve ontwikkelingstoxiciteitspotentie te voorspellen van een stof gebaseerd op de *in vitro* effectconcentraties (meestal binnen een groep van structureel verwante chemicaliën), en 3) *in vivo* ontwikkelingstoxiciteitsdoses te voorspellen door *in vitro* effectconcentraties te vertalen naar *in vivo* effectdoses met gebruik van PBK-modellering. De voorspelling van *in vivo* ontwikkelingstoxiciteit gebaseerd op *in vitro* data kan om verschillende redenen worden over- of onderschat. Deze omvatten: 1) het gebrek aan, of de lage expressie van biotransformatie-enzymen in *in vitro* testsystemen vergeleken met de *in vivo* situatie, 2) het mogelijke gebrek aan relevante cellen of weefsels in het *in vitro* testsysteem voor de *in vivo* waargenomen effecten, 3) het gebrek aan secundaire schadelijke effecten als gevolg van toxische effecten in de moeder of in de placenta, in het *in vitro* testsysteem en 4) het gebrek aan relevante kinetiekdata die nodig zijn om een PBK model te bouwen dat betrouwbaar de *in vivo* kinetiek beschrijft, voor de kwantitatieve extrapolatie van de *in vitro* effectconcentraties naar *in vivo* effectdoses.

De eerste groep chemicaliën die is geselecteerd om het 'proof-of-principle' te ontwikkelen om *in vitro* concentratie-respons curves te vertalen naar *in vivo* dosis-respons curves door gebruik te maken van PBK-modellering met omgekeerde dosimetrie, was een groep glycolethers. De gebruikte glycolethers waren ethyleen glycol monomethyl ether (EGME), ethyleen glycol monoethyl ether (EGEE), ethyleen glycol monobutyl ether (EGBE) en ethyleen glycol monophenyl ether (EGPE). De glycolethers worden *in vivo* vooral gemetaboliseerd naar hun alkoxyazijnzuurmetabolieten, welke respectievelijk methoxyazijnzuur (MAA), ethoxyazijnzuur (EAA), butoxyazijnzuur (BAA) en phenoxyazijnzuur (PAA) zijn, die verantwoordelijk worden geacht voor de ontwikkelingstoxiciteit van de glycolethers (ECETOC, 2005). In de validatiestudies van de EST en de rat postimplantatie WEC-test werd aangetoond dat de toxische effectconcentraties van MAA zich in de millimolair-range bevinden. *In vivo* studies laten zien dat de piek-plasmaconcentraties van MAA bij doses die bij ratten en muizen *in vivo* ontwikkelingstoxiciteit geven, zich ook in de millimolair-range bevinden (Sweeney *et al.*, 2001). Gezien de hoge *in vitro* en *in vivo* effectconcentraties van MAA en zijn zure aard, werd er gehypothetiseerd dat de door de alkoxyazijnzuurmetaboliet-geïnduceerde ontwikkelingstoxiciteit kan worden veroorzaakt door het induceren van een intracellulaire pH (pH_i)-afname van embryonale cellen, wat resulteert in een remming van de celgroei en celdifferentiatie, waardoor de embryonale ontwikkeling wordt verstoord. Om dit werkingsmechanisme van de alkoxyazijnzuurmetabolieten van de glycolethers verder te onderzoeken, werden de effecten van MAA op de pH_i van ES-D3 cellen en embryonale fibroblasten (Balb/c-3T3 cellen) onderzocht in **hoofdstuk 3** van dit proefschrift. In deze studie werden de chemicaliën acetazolamide (ACZ) en valproïnezuur (VPA) gebruikt als positieve controle, omdat voor deze stoffen een relatie tussen een chemisch-geïnduceerde daling van de pH_i van de conceptus en de veroorzaakte ontwikkelingstoxiciteit al eerder is vastgesteld in *in vivo* studies (Scott *et al.*, 1990; Scott *et al.*, 1997). De verkregen resultaten toonden aan dat MAA, ACZ en VPA een pH_i-afname veroorzaken bij concentraties die remming van ES-D3 celdifferentiatie induceren, wat aangeeft dat de door deze stoffen veroorzaakte pH_i-afname is gerelateerd aan de remming van celdifferentiatie. Twee andere ontwikkelingstoxische stoffen, all-*trans*-retinoïnezuur en 5-fluorouracil, hebben geen invloed op de pH_i van embryonale cellen bij concentraties die invloed hebben op de ES-D3 celdifferentiatie, wat wijst op een ander werkingsmechanisme van deze stoffen, zoals ook beschreven staat in de literatuur (Collins and Mao, 1999; Shuey *et al.*, 1994). MAA en ACZ induceerden een

concentratie-afhankelijke remming van de ES-D3 celdifferentiatie, die werd versterkt door amiloride, een remmer van de Na^+/H^+-antiporter, wat een belangrijke rol voor de pH_i in het ontwikkelingstoxiciteitsmechanisme van beide chemicaliën bevestigt. De resultaten die gepresenteerd zijn in hoofdstuk 3 geven aan dat een afname van de pH_i waarschijnlijk de wijze is waarop de alkoxyazijnzuurmetabolieten van de glycolethers ontwikkelingstoxiciteit veroorzaken, en bevestigen de rol van de alkoxyazijnzuurmetabolieten in het ontwikkelingstoxiciteitsmechanisme van de moeder-glycolethers. Bevestigd door dit werkingsmechanisme zijn de *in vitro* concentratie-respons curves (hoofdstuk 4), die moeten worden omgezet naar *in vivo* dosis-respons curves, gemaakt met de toxische azijnzuurmetabolieten in plaats van met de moeder-gycolethers, in ogenschouw nemende dat latere PBK-modellering (hoofdstuk 5) rekening zou houden met de bioactivatie van de gycolethers tot hun alkoxyazijnzuurmetabolieten.

Hoofdstuk 4 presenteert de concentratie-respons curves van MAA, EAA, BAA, PAA, EGME en EGEE in de ES-D3 celdifferentiatietest en vergelijkt de uitkomsten met literatuurgegevens over de *in vivo* ontwikkelingstoxiciteitseffecten van EGME, EGEE, EGBE en EGPE in muizen. Alle geteste alkoxyazijnzuurmetabolieten van de glycolethers lieten een concentratie-afhankelijke remming van hartspierceldifferentiatie zien waarbij MAA het meest potent was, gevolgd door respectievelijk EAA, BAA en PAA. Beide moeder-glycolethers die werden getest (EGME en EGEE, tot 10 mM) remden de ES-D3 celdifferentiatie niet, wat daarmee bevestigt dat de metabolieten de ontwikkelingstoxische stoffen *in vivo* zijn. De potenties van de alkoxyazijnzuurmetabolieten van de glycol ethers in de ES-D3 celdifferentiatietest werden vergeleken met de *in vivo* toxiciteit van de glycolethers in muizen. De potentierangschikking van de stoffen in de ES-D3 celdifferentiatietest kwam overeen met de beschikbare *in vivo* gegevens over de potenties van de moeder-glycolethers (EGME > EGEE > EGBE > EGPE). De relatieve verschillen tussen de potenties van de stoffen bleken geprononceerder in de *in vivo* studies dan in de ES-D3 celdifferentiatietest. Een mogelijke verklaring voor dit verschil kan de afwezigheid van *in vivo* kinetische processen in het *in vitro* testsysteem zijn, waarvan bekend is dat deze verschillen in de *in vivo* situatie voor deze stoffen.

Deze verschillen in *in vivo* kinetiek kunnen worden beschreven in PBK-modellen die worden gebruikt voor de vertaling van de *in vitro* effectconcentraties naar *in vivo* effectdoses. **Hoofdstuk 5** beschrijft de resultaten van de toepassing van deze omgekeerde-dosimetrie-aanpak voor de glycolethers EGME, EGEE, EGBE en EGPE

in de rat en de mens. Een PBK model dat de kinetiek van de vier glycolethers en hun ontwikkelingstoxische alkoxyazijnzuurmetabolieten beschrijft in de rat en de mens werd ontwikkeld. De *in vitro* concentratie-respons curves van deze metabolieten in de ES-D3 celdifferentiatietest (Hoofdstuk 4) werden gebruikt als input in het PBK-model om *in vitro* concentratie-respons curves te vertalen naar voorspelde *in vivo* dosis-respons curves voor de toxiciteit van de moeder-glycolethers in de rat en de mens. De voorspelde dosis-respons curves voor ratten bleken in overeenstemming (binnen een orde van grootte) met de ontwikkelingstoxiciteitsdosisniveaus gemeten in verscheidene *in vivo* ontwikkelingstoxiciteitsstudies in de rat. De aldus verkregen resultaten in hoofdstuk 5 vormen het eerste 'proof-of-principle' dat *in vitro* concentratie-respons curves kunnen worden omgezet naar *in vivo* dosis-respons curves met behulp van PBK modellering met een omgekeerde-dosimetrie-aanpak, die geschikt zijn voor gebruik in de risico-evaluatie. Daarom werd geconcludeerd dat deze aanpak, na verdere evaluatie, kan bijdragen aan een op wetenschap gebaseerde risicobeoordeling van chemische stoffen, door geen gebruik te maken van proefdieren, of slechts van kleine aantallen.

De tweede groep chemicaliën die werd gebruikt om de betrouwbaarheid te beoordelen van het voorspellen van *in vivo* dosis-respons curves voor ontwikkelingstoxiciteit met behulp van de omgekeerde-dosimetrie-aanpak, was een groep retinoïden. Retinoïden zijn chemicaliën die een chemische structuur of functionele eigenschappen hebben die vergelijkbaar zijn met die van retinol (vitamine A). De onderzochte retinoïden zijn retinol, all-*trans*-retinoïnezuur (ATRA), 13-*cis*-retinoïnezuur (13-*cis*-RA), 9-*cis*-retinoïnezuur (9-*cis*-RA), etretinaat, acitretine en TTNPB ((E) -4 [2 - (5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2-naftalenyl)-1-propenyl] benzoëzuur). Voor een normale embryogenese moeten de juiste niveaus van vitamine A worden gehandhaafd, terwijl een vitamine A-tekort en een hoge retinoïdeninname schadelijke effecten kan veroorzaken aan de ontwikkelende conceptus (Collins en Mao, 1999). Van verscheidene retinoïden is aangetoond dat ze teratogeen zijn in een aantal proefdieren, waaronder muizen, ratten, cavia's, hamsters, konijnen, honden, varkens, kippen en apen (Collins en Mao, 1999). Van retinoïden is eerder laten zien dat ze actief zijn in de nanomolaire tot micromolaire range in de 'limb bud micromass'-kweektest (Kistler, 1987), wat de grotere *in vitro* potentie van deze stoffen in vergelijking met de alkoxyazijnzuurmetabolieten van de glycolethers laat zien. Verder is het mechanisme van retinoïden-geïnduceerde ontwikkelingstoxiciteitseffecten

complexer (verstoring van de intracellulaire retinoïde-signaaltransductie-pathways (Collins en Mao, 1999) dan het waarschijnlijk niet-specifieke ontwikkelings-toxiciteitsmechanisme van de glycolethers (inductie van pH_i-afname).

Hoofdstuk 6 presenteert de concentratie-respons curves van retinol, ATRA, 13-*cis*-RA, 9-*cis*-RA, etretinaat, acitretine en TTNPB in de ES-D3 celdifferentiatietest en vergelijkt de uitkomsten van hun relatieve potenties met de relatieve potenties die zijn gevonden in twee andere *in vitro* testen voor ontwikkelingstoxiciteit (de limb bud micromass-kweektest en de rat postimplantatie WEC-test) en met de relatieve potentie in *in vivo* (muis, rat en konijn) ontwikkelingstoxiciteitsstudies. De resultaten laten zien dat de potentieranking verkregen in de ES-D3 celdifferentiatietest vergelijkbaar is met de potentierankings zoals gerapporteerd voor de twee andere *in vitro* testen voor ontwikkelingstoxiciteit. TTNPB was de meest potente retinoïde, terwijl etretinaat en retinol de laagste potentie hadden. ATRA, 13-*cis*-RA, 9-*cis*-RA en acitretine toonden een gemiddelde potentie. *In vivo* potentierankings van de ontwikkelingstoxiciteit van retinoïden lijken afhankelijk te zijn van de diersoort en / of de gebruikte blootstellingsregimes. De verkregen *in vitro* potentieranking komt niet volledig overeen met de *in vivo* potentieranking, maar TTNPB werd correct als de meest potente en retinol als de minst potente congeneer voorspeld. De verschillen tussen de kinetiek van de ES-D3 celdifferentiatietest en de kinetische processen *in vivo* zouden de afwijkende potentievoorspellingen voor sommige retinoïden kunnen verklaren. Daarom werd geconcludeerd dat kennis van soort-afhankelijke stof-specifieke *in vivo* kinetiek van essentieel belang is bij het gebruik van *in vitro* toxiciteitsgegevens voor de schatting van *in vivo* ontwikkelingstoxiciteitspotenties binnen reeksen van verwante chemicaliën.

In **hoofdstuk 7** werd de omgekeerde-dosimetrie-aanpak toegepast voor de retinoïde ATRA, door de *in vitro* concentratie-respons curve voor de remming van ES-D3 celdifferentiatie te vertalen naar *in vivo* dosis-respons curves, met behulp van een PBK model dat de kinetiek van ATRA beschrijft voor de rat en de mens. Terwijl voor de PBK-modellen ontwikkeld voor de glycolethers (hoofdstuk 5) kinetische gegevens van *in vivo* studies nodig waren voor het afleiden van bepaalde modelparameterwaarden (d.w.z. voor de klaring van de alkoxyazijnzuur-metabolieten), was het PBK-model voor ATRA uitsluitend gebaseerd op parameterwaarden afgeleid met behulp van *in vitro* technieken. Voorspellingen van de dosis-respons curves voor *in vivo* ontwikkelingstoxiciteit werden gemaakt door

het toxische effect te relateren aan 1) de maximale concentratie (C_{max}) in het bloed, zoals eerder gedaan voor de glycolethers (hoofdstuk 5) en 2) de oppervlakte onder de concentratie-tijd curve (AUC) in het bloed. De voorspelde ontwikkelingstoxiciteitsdoses voor ATRA waren dichter bij de *in vivo* rat data wanneer voorspellingen waren gebaseerd door het toxische effect te relateren aan de AUC, dan wanneer voorspellingen waren gebaseerd door het toxische effect te relateren aan de C_{max}. Dit is in overeenstemming met resultaten uit de literatuur, waaruit blijkt dat de effecten op de ontwikkelingstoxiciteitsuitkomst van ATRA *in vivo* het best is gecorreleerd met de dosismetriek AUC (Tzimas *et al.*, 1997). De voorspelde dosis-respons curves voor ratten bleken in overeenstemming (binnen een orde van grootte) met ontwikkelingstoxiciteitsdosisniveaus gemeten in verscheidene *in vivo* ontwikkelingstoxiciteitsstudies in de rat. Het is belangrijk om te benadrukken dat een essentiële aanvulling van deze studie op het eerdere omgekeerde-dosimetrie-werk aan de glycolethers (hoofdstuk 5) is, dat er geen kinetische gegevens van *in vivo* studies nodig waren voor het afleiden van de modelparameterwaarden voor het PBK-model van ATRA. Deze studie verschaft dus het eerste 'proof-of-principle' dat *in vivo* dosis-respons-relaties voor dit systemische eindpunt kunnen worden voorspeld met behulp van omgekeerde dosimetrie met een PBK model dat is geparameterizeerd zonder *in vivo* gegevens. De voorspelde dosis-respons curves kunnen worden gebruikt om startpunten vast te stellen voor het afleiden van veilige blootstellingsniveaus voor de mens, wat wijst op de mogelijkheden van de aanpak als een instrument in toxicologische risicobeoordelingsstrategieën.

Toekomstperspectieven en conclusies voor het gebruik van de gecombineerde *in vitro-in silico* aanpak in de risicobeoordeling

Dit proefschrift verschaft het 'proof-of-principle' dat *in vitro* concentratie-respons curves kunnen worden omgezet naar *in vivo* dosis-respons curves voor ontwikkelingstoxiciteit, met behulp van fysiologisch gebaseerd kinetisch (PBK)-modelleren met een omgekeerde-dosimetrie-aanpak. Om het vertrouwen van het gebruik van de aanpak voor de risico-evaluatie te verhogen, zijn meer 'proof-of-principle'-studies met andere ontwikkelingstoxische stoffen en meer onderzoeken naar andere eindpunten nodig. Op dit moment zijn een aantal voorbeelden van het voorspellen van *in vivo* toxiciteit met behulp van omgekeerde dosimetrie

gerapporteerd in de literatuur (Punt *et al.*, 2011). Voor een beperkt aantal chemicaliën zijn voorspellingen gebaseerd op omgekeerde dosimetrie voor de toxicologische eindpunten acute en herhaalde blootstellingstoxiciteit, neurotoxiciteit, genotoxiciteit en ontwikkelingstoxiciteit gemaakt (Punt *et al.*, 2011). Omgekeerde-dosimetrie-benaderingen met behulp van PBK-modellen worden ook gebruikt in het ToxCast-programma van de 'Environmental Protection Agency' (EPA) van de Verenigde Staten en in de Europese kaderprogrammaprojecten 'ESNATS' en 'ChemScreen' om *in vivo* effectdoses te voorspellen op basis van *in vitro* effectconcentraties. Naarmate er meer voorbeelden van het 'proof-of-principle' voor andere chemicaliën en andere eindpunten zullen worden verstrekt, kan het worden overwogen om de aanpak toe te passen in teststrategieën voor toxicologische risico-evaluaties. Des te meer omdat de aanpak goed binnen de toxiciteitsteststrategieën past zoals voorzien in het in 2007 uitgebrachte rapport '*Toxiciteitstesten in de 21e eeuw: een visie en een strategie*' door de Amerikaanse 'National Academy of Sciences' (NAS) (NRC, 2007). Dit rapport voorziet een toekomst waarin alle routine toxiciteitstesten met behulp van *in vitro* testen zullen worden uitgevoerd, door het evalueren van cellulaire reacties op blootstelling aan chemicaliën. In de nieuw voorgestelde risico-evaluatie-strategie, zal op basis van *in vitro* modellen en PBK-modelleren worden geschat bij welke blootstellingsniveaus schadelijke cellulaire verstoringen zullen optreden in de *in vivo* situatie (Andersen en Krewski, 2009). PBK-modelleren met behulp van een omgekeerde-dosimetrie-aanpak kan een belangrijke bijdrage leveren aan deze schattingen.

Dit proefschrift gebruikt omgekeerde dosimetrie door het vertalen van een complete *in vitro* concentratie-respons curve naar een complete *in vivo* dosis-respons curve met PBK-modellen die nauwkeurig de *in vivo* kinetiek beschrijven van de stof van interesse. De voorspelde dosis-respons curves kunnen worden gebruikt om startpunten vast te stellen voor de risicobeoordeling om aanvaardbare dagelijkse inname (ADI)-waarden te bepalen. De aanpak met behulp van complexe PBK-modellen kan echter te tijdrovend zijn om toe te passen voor alle chemicaliën die nog niet zijn geëvalueerd, en kan daarom niet geschikt zijn om te gebruiken voor alle chemicaliën. Daarom zou een voorselectie van chemicaliën kunnen worden gemaakt met behulp van eenvoudige algemene PBK-modellen zoals gebruikt in ToxCast, om chemicaliën van de grootste zorg te identificeren die verder onderzocht moeten worden. Het ToxCast programma, geïnitieerd binnen de EPA, richt zich op de

ontwikkeling van methoden om (milieu-)chemicaliën voor verdere screening en testen te prioriteren. In de gebruikte aanpak wordt de teststof getest in een batterij van meer dan 500 *in vitro* testen, waarna de potentie van de stof in elke positieve test wordt samengevat met behulp van de waarden voor de AC_{50} (concentratie op 50% van de maximale activiteit) of de laagste effectieve concentratie (LEC), afhankelijk van het type concentratie-respons gegevens dat verzameld is in de test. Daarnaast worden *in vitro* testen uitgevoerd om hepatische metabole klaring en plasma-eiwitbinding bij mensen te schatten. Deze *in vitro* kinetische gegevens worden opgenomen in het algemene PBK-model voor de *in vitro-in vivo* extrapolaties om een dagelijkse humane orale toediening te berekenen die nodig zou zijn om een *in vivo* bloedspiegel van de stof te bereiken die gelijk is aan de *in vitro* AC_{50}- of LEC-waarde, welke de orale equivalente dosis wordt genoemd (Rotroff *et al.*, 2010; Wetmore *et al.*, 2012). Indien gegevens over blootstelling beschikbaar zijn, wordt de geschatte orale equivalente dosis vergeleken met de geschatte menselijke orale blootstelling om te beoordelen of een *in vivo* effect kan worden verwacht (Rotroff *et al.*, 2010; Wetmore *et al.*, 2012). Als dat zo is, kunnen deze stoffen worden geprioriteerd voor verdere risicobeoordeling.

De chemicaliën die zouden worden geprioriteerd voor verdere risico-evaluatie, zouden in groter detail kunnen worden beoordeeld met behulp van omgekeerde dosimetrie met meer verfijnde *in vitro* toxiciteitstesten en PBK-modellen, zoals beschreven in dit proefschrift. Deze PBK-modellen kunnen worden gebruikt om nauwkeurig *in vivo* toxische doses te voorspellen voor de gevoelige individuen binnen de menselijke bevolking, door het beschrijven van interindividuele verschillen in kinetiek in het model met behulp van Monte Carlo-simulaties (hoofdstuk 7). De voorspelde dosis-respons curves kunnen worden gebruikt om startpunten vast te stellen om veilige blootstellingsniveaus te bepalen. Als PBK-modelleren met een omgekeerde-dosimetrie-aanpak, waarbij *in vitro* concentratie-respons curves worden vertaald naar *in vivo* dosis-respons curves, zou worden gebruikt voor menselijke risico-evaluaties, hoeven alleen PBK-modellen voor de mens te worden ontwikkeld, wat impliceert dat er geen onzekerheidsfactor voor verschillen in kinetiek tussen soorten nodig zou zijn. Als de menselijke interindividuele verschillen in kinetiek worden opgenomen in het PBK-model (hoofdstuk 7), zou er ook geen onzekerheidsfactor voor deze interindividuele verschillen in kinetiek hoeven worden toegepast. Als het *in vitro* model voor het bepalen van *in vitro* concentratie-respons curves zou zijn gebaseerd op menselijke

cellen in plaats van muizencellen, zou bovendien geen onzekerheidsfactor voor verschillen in dynamiek tussen soorten nodig zijn, zodat alleen onzekerheidsfactoren voor menselijke interindividuele verschillen in dynamiek nodig zouden zijn. Aan de andere kant zou een onzekerheidsfactor kunnen worden ingevoerd voor onzekerheid in de nauwkeurigheid van voorspelde dosis-respons effecten bij de omgekeerde-dosimetrie-benadering. Voor de chemicaliën die zijn gebruikt in dit proefschrift verschillen de voorspelde dosis-respons-effecten en de dosis-respons effecten gemeten in gerapporteerde *in vivo* studies in ratten minder dan een orde van grootte, wat aangeeft dat een onzekerheidsfactor van 10 zou kunnen worden gebruikt voor onzekerheden in de voorspellingen. Er zijn echter meer gegevens nodig om een geschikte waarde voor deze onzekerheid te bepalen.

Al met al toont dit proefschrift aan dat het haalbaar is om concentratie-respons curves verkregen in *in vitro* toxiciteitstesten te vertalen naar voorspelde *in vivo* dosis-respons curves, welke kunnen worden gebruikt in de risicobeoordeling. Het gepresenteerde werk in dit proefschrift levert het eerste bewijs dat *in vivo* ontwikkelingstoxiciteitsdoses nauwkeurig kunnen worden voorspeld door het combineren van enkel *in vitro* toxiciteitsgegevens en *in silico* PBK-modellering, door gebruik te maken van een omgekeerde-dosimetrie-aanpak. Dit werk kan daarom bijdragen aan een verhoogde acceptatie en de toekomstige implementatie van *in vitro* toxiciteitsgegevens in de risicobeoordeling en kan zelfs worden beschouwd als de belangrijkste aanpak om de doelstellingen van de 3 V's van diergebruik in de toxicologische risicobeoordeling te bereiken.

Referenties

Andersen, M. E., Krewski, D. (2009). Toxicity testing in the 21st century: bringing the vision to life. *Toxicol Sci* **107**: 324-330.

Collins, M. D., Mao, G. E. (1999). Teratology of retinoids. *Annu Rev Pharmacol Toxicol* **39**, 399-430.

ECETOC. (2005). The toxicology of glycol ethers and its relevance to man. Technical Report 095, 4th ed. European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels, Belgium.

Genschow, E., Spielmann, H., Scholz, G., Pohl, I., Seiler, A., Clemann, N., Bremer, S., Becker, K. (2004). Validation of the embryonic stem cell test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* **32**, 209-44.

Genschow, E., Spielmann, H., Scholz, G., Seiler, A., Brown, N. A., Piersma, A. H., Brady, M., Clemann, N., Huuskonen, H., Paillard, F., Bremer, S., Becker, K. (2002). The ECVAM international validation study on in vitro embryotoxicity tests: results of the definitive phase and evaluation of prediction models. European Centre for the Validation of Alternative Methods. *Atern Lab Anim* **30**, 151-76.

National Research Council (NRC). (2007). Toxicity testing of the 21st century: a vision and a strategy. National Academy Press, Washington, DC.

Punt, A., Schiffelers, M. J., Jean Horbach, G., van de Sandt, J. J. M., Groothuis, G. M., Rietjens, I. M. C. M., Blaauboer, B. J. (2011). Evaluation of research activities and research needs to increase the impact and applicability of alternative testing strategies in risk assessment practice. *Regul Toxicol Pharmacol* **61**, 105-14.

Rotroff, D. M., Wetmore, B. A., Dix, D. J., Ferguson, S. S., Clewell, H. J., Houck, K. A., Lecluyse, E. L., Andersen, M. E., Judson, R. S., Smith, C. M., Sochaski, M. A., Kavlock, R. J., Boellmann, F., Martin, M. T., Reif, D. M., Wambaugh, J. F., Thomas, R. S. (2010). Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening. *Toxicol Sci* **117**, 348-58.

Scott, W. J., Duggan, C. A., Schreiner, C. M., Collins, M. D. (1990). Reduction of embryonic intracellular pH: a potential mechanism of acetazolamide-induced limb malformations. *Toxicol Appl Pharmacol* **103**, 238-54.

Scott, W. J. Jr., Schreiner, C. M., Nau, H., Vorhees, C. V., Beliles, R. P., Colvin, J., McCandless, D. (1997). Valproate-induced limb malformations in mice associated with reduction of intracellular pH. *Reprod Toxicol* **11**, 483-93.

Shuey, D. L., Lau, C., Logsdon, T. R., Zucker, R. M., Elstein, K. H., Narotsky, M. G., Setzer, R. W., Kavlock, R. J., Rogers, J. M. (1994). Biologically based dose-response modeling in developmental toxicology: biochemical and cellular sequelae of 5-fluorouracil exposure in the developing rat. *Toxicol Appl Pharmacol* **126**, 129-44.

Sweeney, L. M., Tyler, T. R., Kirman, C. R., Corley, R. A., Reitz, R. H., Paustenbach, D. J., Holson, J. F., Whorton, M. D., Thompson, K. M., Gargas, M. L. (2001). Proposed occupational exposure limits for select ethylene glycol ethers using PBPK models and Monte Carlo simulations. *Toxicol Sci* **62**, 124-39.

Tzimas, G., Thiel, R., Chahoud, I., Nau, H. (1997). The area under the concentration-time curve of all-trans-retinoic acid is the most suitable pharmacokinetic correlate to the embryotoxicity of this retinoid in the rat. *Toxicol Appl Pharmacol* **143**, 436-44.

Van der Jagt, K., Munn, S. J., Tørsløv, J., de Bruijn, J. (2004). Alternative approaches can reduce the use of test animals under REACH. *Report EUR* **21405**.

Wetmore, B. A., Wambaugh, J. F., Ferguson, S. S., Sochaski, M. A., Rotroff, D. M., Freeman, K., Clewell, H. J. 3rd, Dix, D. J., Andersen, M. E., Houck, K. A., Allen, B., Judson, R. S., Singh, R., Kavlock, R. J., Richard, A. M., Thomas, R. S. (2012). Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. *Toxicol Sci* **125**, 157-74.

DANKWOORD
ACKNOWLEDGEMENTS

Dankwoord / Acknowledgements

Er zijn veel mensen die een rol hebben gespeeld bij de totstandkoming van mijn proefschrift. Ik ben daarvoor veel dank verschuldigd aan mijn begeleiders, collega's, familie en vrienden.

Ivonne, bedankt dat je me de mogelijkheid gaf om dit project te doen. Bedankt voor je altijd snelle feedback en constructieve commentaar op abstracts, manuscripten en presentaties. Ik wil je ook bedanken voor de mogelijkheden die je me hebt gegeven om naast mijn eigen onderzoek ook bij andere projecten betrokken te zijn en ervaring op te doen in het geven van onderwijs. Ik heb heel veel van je geleerd en prettig met je samengewerkt. Ik zie uit naar onze toekomstige samenwerking!

Miriam, bedankt voor de ruimte die je me gaf tijdens dit project. Bedankt voor je inhoudelijke input en je goede en prettige begeleiding op TNO. Ik heb onze overleggen, waar we lekker konden brainstormen, altijd zeer gewaardeerd. Bedankt voor je geduld als ik weer eens mijn resultaten of hypotheses tijdens een bila op papier begon te zetten. Ook veel dank voor de mogelijkheid die je me hebt gegeven om deel te nemen aan de ESNATS summer schools. Misschien heb ik daar mijn nieuwe baan in Italië wel aan te danken!

Bas, bedankt voor je bijdrage in dit project. Je enthousiasme voor het *in vitro*-toxicologieonderzoek zijn aanstekelijk en erg belangrijk om *in vitro* testing daadwerkelijk toe te gaan passen in de toxicologische risicobeoordeling. Ik heb je opbouwende kritiek in onze projectoverleggen altijd zeer gewaardeerd.

Han, bedankt voor je bijdrage aan het eerste deel van dit proefschrift. Bedankt voor je goede ideeën voor het vormgeven van het glycoletherartikel in samenwerking met het RIVM. Ik wil je ook bedanken voor je begeleiding tijdens het zwangerschapsverlof van Miriam. Ik heb je begeleiding en commentaar altijd als zeer constructief ervaren.

Ik wil Esther en Aldert van het RIVM bedanken voor de prettige samenwerking bij het glycoletherwerk beschreven in hoofdstuk 4 en 5 van dit proefschrift. Ruud, ook jij bedankt voor je bijdrage aan dit werk.

Ik heb mijn AIO-jaren als erg leuk en leerzaam ervaren. Dit is voor een groot deel te danken aan m'n collega's van de leerstoelgroep Toxicologie en van TNO/TNO Triskelion.

Het eerste deel van mijn project heb ik uitgevoerd op TNO. Marloes, bedankt voor de eerste opvang op TNO. Ik voelde me gelijk thuis! Mariska, dank voor je hulp bij het aanleren van de EST en alle andere hulp die je op het lab hebt geboden. Ik heb op TNO een aantal keer van kamer gewisseld en daarmee veel verschillende kamergenoten gehad. Bedankt voor de gezelligheid, discussies en

gesprekken Marcel, Marijn, Evita, Sieto, Nicole, Bianca en Astrid. Dank ook aan de andere TNO'ers Anna, Andreas, Astrid, Bert, Birol, Cyrille, Dimitri, Florence, Freek, Harm, Heleen, Irene, Jaap Jan, Jan, Jeroen, Lisette, Martijn, Mary-lène, Michelle, Nico, Rianne, Richard, Truus, Tyrone, Wilfred, Wouter, *et al.* voor jullie hulp bij mijn project en/of de gezelligheid tijdens werk, koffiepauzes en TNO-uitjes.

The second part of my PhD was mostly spent at the Division of Toxicology. I would like to thank all my TOX-colleagues for the good time I have had there! Thank you Irene and Gré for helping with ordering a variety of chemicals and labware and for your help with all kind of administrative stuff. Thanks Laura, Marelle, Ans S., Hans and Bert for your technical assistance and other help. Bert, thank you for fixing the HPLC so many times! Laura, thanks for the nice collaboration and company on the lab! My roomies Ana, Alicia and Hequn, thank you for the nice company, discussions and fun in room 1.024. Also many thanks to all the other (former) TOX-colleagues Agata, Ala', Alexandros, Ans P., Arif, Barae, David, Elise, Elton, Erryana, Gerrit, Henriëtte, Henrique, Jac, Jaime, Jonathan, Judith, Karsten, Letty, Linda, Marije, Merel, Myrthe, Niek, Nynke, Reiko, Rungnapa, Samantha, Si, Sourav, Suzanne, Tinka, Walter, Wasma, Wiratno, *et al.*. It has been a great pleasure working with you!

Several students were involved in my project. Thank you Yanqing, Süleyman, Inge, Hequn, Martine and Nadir for the nice collaboration and all the work you did! I enjoyed supervising you all and wish you the best in your future career!

Diane en Rob, bedankt voor jullie goede vriendschap! Ik ben blij dat jullie mijn paranimfen willen zijn en naast me op het podium zullen staan. Ik kijk uit naar alle leuke dingen die we in de toekomst nog zullen gaan doen!

Thijs, bedankt voor het ontwerpen van de omslag van mijn proefschrift en de uitnodiging. Ik vind het erg mooi geworden!

Naast werken op het lab en schrijven aan artikelen bleef er gelukkig nog genoeg tijd over voor de broodnodige ontspanning. Ik wil iedereen van 16C bedanken voor de leuke tijd tijdens m'n studie- en AIO-tijd op deze gezellige afdeling. Ik hoop dat we nog lang contact met elkaar zullen houden! Daarnaast ook dank aan alle korfballers van Débaldéruit. Leuk dat we nog steeds contact hebben en elkaar geregeld zien. Laten we het komende Débaldéruitje in Italië houden! Daarnaast ook dank aan de korfballers van SSS. Ik ben blij dat ik elke week mijn ontspanning bij SSS heb kunnen vinden en zo ook veel niet-WUR-Wageningers heb leren kennen. Ik zal het gaan missen in Italië...

Tot slot wil ik mijn lieve vrienden en familie bedanken. Bedankt voor jullie interesse en steun in de afgelopen jaren. Ik vind het erg leuk dat velen van jullie bij de verdediging van mijn proefschrift zullen zijn!

Jochem

ABOUT THE AUTHOR

Curriculum Vitae

Jochem Louisse werd geboren op 2 juni 1982 in Linschoten. Na het behalen van zijn VWO-diploma aan het Kalsbeek College te Woerden in 2000 is hij in 2001 begonnen aan zijn studie Biologie aan de Wageningen Universiteit te Wageningen. Tijdens zijn MSc-studie heeft hij 3 onderzoeksprojecten uitgevoerd, waarvan de eerste werd uitgevoerd op de leerstoelgroep Toxicologie van Wageningen Universiteit. Hier deed hij onderzoek naar de *in vitro* en *in vivo* hormoon-verstorende effecten van stoffen uit verpakkingsmaterialen van voedsel. Zijn tweede onderzoeksproject voerde hij uit op de leerstoelgroep Moleculaire Farmacologie en Toxicologie van het Radboud Universitair Medisch Centrum in Nijmegen, waarbij hij het effect van ontstekingsmediatoren onderzocht op de expressie van het transporteiwit P-glycoproteine in niercellen *in vitro*. Hij sloot zijn MSc-studie af met een onderzoeksproject aan de groep Neurobiologie, Fysiologie en Gedrag van de University of California Davis in Californië, Verenigde Staten, waar hij heeft gewerkt aan de ontwikkeling van een transgene klauwkikker (*Xenopus laevis*) die een schildklierhormoon-responsief element bevat met daaraan een luciferase-gen gekoppeld. Na het behalen van zijn MSc-diploma heeft hij op de leerstoelgroep Toxicologie van de Wageningen Universiteit gewerkt aan de identificatie van stoffen uit basilicum die de bioactivatie van de in basilicum voorkomende kankerverwekkende stof estragol remmen. In december 2007 trad Jochem in dienst als assistent in opleiding (AIO) op de leerstoelgroep Toxicologie van de Wageningen Universiteit. Het onderzoek beschreven in dit proefschrift werd uitgevoerd in samenwerkingsverband tussen de leerstoelgroep Toxicologie van de Wageningen Universiteit, TNO en het Institute for Risk Assessment Sciences van de Universiteit Utrecht. Vanaf juni 2012 zal Jochem werken als postdoc op het Institute for Health and Consumer Protection van het Joint Research Centre van de Europese Commissie in Ispra, Italië.

List of publications

Louisse, J., Verwei, M., Woutersen, R. A., Blaauboer, B. J., Rietjens, I. M. C. M. (2012). Toward in vitro biomarkers for developmental toxicity and their extrapolation to the in vivo situation. *Expert Opinion on Drug Metabolism and Toxicology* **8**, 11-27.

Rietjens, I. M. C. M., Louisse, J., Punt, A. (2011). Educational overview on physiologically based kinetic (PBK) modeling in molecular nutrition and food research. *Molecular Nutrition and Food Research* **55**, 941-956.

Louisse, J., Gönen, S., Rietjens, I. M. C. M., Verwei, M. (2011). Relative developmental toxicity potencies of retinoids in the embryonic stem cell test compared with their relative potencies in *in vivo* and two other *in vitro* assays for developmental toxicity. *Toxicology Letters* **203**, 1-8.

Louisse, J., de Jong, E., van de Sandt, J. J. M., Blaauboer, B. J., Woutersen, R. A., Piersma, A. H., Rietjens, I. M. C. M., Verwei, M. (2010). The use of in vitro toxicity data and physiologically based kinetic modeling to predict dose-response curves for in vivo developmental toxicity of glycol ethers in rat and man. *Toxicological Sciences* **118**, 470-484.

Louisse, J., Bai, Y., Verwei, M., van de Sandt, J. J. M., Blaauboer, B. J., Rietjens, I. M. C. M. (2010). Decrease of intracellular pH as possible mechanism of embryotoxicity of glycol ether alkoxyacetic acid metabolites. *Toxicology and Applied Pharmacology* 245, 236-243.

De Jong, E., Louisse, J., Verwei, M., Blaauboer, B. J., van de Sandt, J. J. M., Woutersen, R. A., Rietjens, I. M. C. M., Piersma, A. H. (2009). Relative developmental toxicity of glycol ether alkoxy acid metabolites in the embryonic stem cell test as compared with the in vivo potency of their parent compounds. *Toxicological Sciences* **110**, 117-124.

Alhusainy, W., Paini, A., Punt, A., Louisse, J., Spenkelink, A., Vervoort, J., Delatour, T., Scholz, G., Schilter, B., Adams, T., van Bladeren, P. J., Rietjens, I. M. C. M. (2010). Identification of nevadensin as an important herb-based constituent inhibiting estragole bioactivation and physiology-based biokinetic modeling of its possible in vivo effect. *Toxicology and Applied Pharmacology* **245**, 179-190.

Heemskerk, S., Peters, J. G. P., Louisse, J., Sagar, S., Russel, F. G. M., Masereeuw, R. (2010) Regulation of P-glycoprotein in renal proximal tubule epithelial cells by LPS and TNF-alpha. *Journal of Biomedical Biotechnology* 525180.

Ter Veld, M. G. R., Schouten, B., Louisse, J. van Es, D. S., van der Saag, P. T., Rietjens, I. M. C. M., Murk, A. J. (2006). Estrogenic potency of food-packaging-associated plasticizers and antioxidants as detected in ERalpha and ERbeta reporter gene cell lines. *Journal of Agricultural and Food Chemistry* **54**, 4407-4416.

Overview of completed training activities

Courses

PBPK Modeling and Risk Assessment Workshop, Hamner Institute for Health Sciences (2008)
Reproductive Toxicology, Postdoctoral Education in Toxicology (PET) (2008)
Toxicological Risk Assessment, PET (2008)
Organ Toxicology, PET (2008)
Ecotoxicology, PET (2009)
Pathobiology, PET (2009)
PBPK Modeling Course, Division of Toxicology, Wageningen University (WU) (2009)
Mutagenesis and Carcinogenesis, PET (2010)
Epidemiology, PET (2010)
Medical, Forensic and Regulatory Toxicology, PET (2010)

Meetings

15th International Congress on In Vitro Toxicology (2008), Stockholm, Sweden (oral presentation)
7th World Congress on Alternatives & Animal Use in the Life Sciences (2009), Rome, Italy (poster)
50th annual meeting of the American Society of Toxicology (2011), Washington DC, USA (poster)
51st annual meeting of the American Society of Toxicology (2012), San Francisco, USA (oral presentation)

General courses

Summer school EU FP7 ESNATS project (2009), Zermatt, Switzerland
Risk Communication, PET (2010)
'Durf, leer en presenteer' Presentation and media training, Pauline van Aken (2011)
Summer school EU FP7 ESNATS project (2011), Sithonia, Greece

Optionals

Research in progress presentations of PhD students and staff of the Division of Toxicology, WU
Research discussion presentations of MSc students of the Division of Toxicology, WU
Literature discussion presentations at the Division of Toxicology, WU
Research discussion presentations of TNO
Dutch Society of Toxicology PhD student symposia (2008, 2009, 2010, 2011)
PhD trip Switzerland and Italy, Division of Toxicology (2011)

Approved by the graduate school VLAG

The studies described in this thesis were conducted within a collaborative project of Wageningen University, TNO and Utrecht University. The research was financially supported by The Netherlands Organization for Health Research and Development (ZonMw; project no. 114000088).

Financial support from Wageningen University, TNO, Stichting Stimuleringsfonds voor Alternatieven voor Dierproeven and the J.E. Jurriaanse Stichting for printing this thesis is gratefully acknowledged.

The thesis was printed by GVO drukkers en vormgevers B.V. / Ponsen & Looijen, Ede, The Netherlands.

Cover design: Thijs Louisse