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. Blauboer
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 \textit{in vitro-in silico} models for predicting \textit{in vivo} developmental toxicity \\
facilitating non-animal based safety assessment \\

Jochem Louisse
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
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>General introduction</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>Toward <em>in vitro</em> biomarkers for developmental toxicity and their extrapolation to the <em>in vivo</em> situation</td>
<td>31</td>
</tr>
<tr>
<td>3</td>
<td>Decrease of intracellular pH as possible mechanism of developmental toxicity of glycol ether alkoxycetic acid metabolites</td>
<td>67</td>
</tr>
<tr>
<td>4</td>
<td>Relative developmental toxicity of glycol ether alkoxycetic acid metabolites in the embryonic stem cell test as compared with the <em>in vivo</em> potencies of their parent compounds</td>
<td>89</td>
</tr>
<tr>
<td>5</td>
<td>The use of <em>in vitro</em> toxicity data and physiologically based kinetic modeling to predict dose-response curves for <em>in vivo</em> developmental toxicity of glycol ethers in rat and human</td>
<td>109</td>
</tr>
<tr>
<td>6</td>
<td>Relative developmental toxicity potencies of retinoids in the embryonic stem cell test compared with their relative potencies in <em>in vivo</em> and two other <em>in vitro</em> assays for developmental toxicity</td>
<td>143</td>
</tr>
<tr>
<td>7</td>
<td>Prediction of <em>in vivo</em> developmental toxicity of all-trans-retinoic acid by combining <em>in vitro</em> toxicity data and physiologically based kinetic modeling</td>
<td>165</td>
</tr>
<tr>
<td>8</td>
<td>Summary, general discussion, future perspectives and conclusions</td>
<td>201</td>
</tr>
<tr>
<td>9</td>
<td>Samenvatting, toekomstperspectieven en conclusies</td>
<td>233</td>
</tr>
<tr>
<td></td>
<td>Dankwoord / Acknowledgements</td>
<td>247</td>
</tr>
<tr>
<td></td>
<td>About the author</td>
<td>251</td>
</tr>
</tbody>
</table>
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 (Blauboeer, 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 \textit{in vivo} doses may be achieved using physiologically based kinetic (PBK) modeling using a reverse dosimetry approach (DeJongh \textit{et al.}, 1999; Verwei \textit{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 \textit{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 \textit{in vitro} effect concentration (e.g. an EC\textsubscript{50} value) into a single \textit{in vivo} effect dose (e.g. a lowest observed adverse effect level (LOAEL)) (DeJongh \textit{et al.}, 1999; Verwei \textit{et al.}, 2006; Forsby and Blaauboer, 2007). However, the \textit{in vitro} effect concentrations and the \textit{in vivo} effect doses were chosen arbitrarily in previous reverse dosimetry work. In theory, a complete concentration-response curve obtained in an \textit{in vitro} assay could be translated into a complete predicted \textit{in vivo} dose-response curve, by translating all data obtained in the \textit{in vitro} experiment to the \textit{in vivo} situation. This would enable dose-response modeling on the predicted \textit{in vivo} dose-response data, defining a predicted \textit{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 \textit{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
Chapter 1

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.

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Cas no.</th>
<th>Chemical</th>
<th>Cas no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>aromatic solvents</td>
<td>/1-43-2</td>
<td>hydroxyurea</td>
<td>127-07-1</td>
</tr>
<tr>
<td>benzene</td>
<td>/5-71-8</td>
<td>lovastatin</td>
<td>/5336-70-5</td>
</tr>
<tr>
<td>ethylbenzene</td>
<td>100-41-4</td>
<td>methyl ethyl ketone</td>
<td>78-83-3</td>
</tr>
<tr>
<td>toluene</td>
<td>106-88-3</td>
<td>methylmercury</td>
<td>115-09-3</td>
</tr>
<tr>
<td>xylene</td>
<td>1330-20-7</td>
<td>xirax</td>
<td>2385-85-5</td>
</tr>
<tr>
<td>boric acid</td>
<td>10043-35-3</td>
<td>nitraten</td>
<td>1636-78-5</td>
</tr>
<tr>
<td>butylbenzene</td>
<td>55-98-1</td>
<td>ochratoxin A</td>
<td>303-47-9</td>
</tr>
<tr>
<td>bromodichloromethane</td>
<td>75-27-4</td>
<td>retinooids</td>
<td></td>
</tr>
<tr>
<td>carbon disulfide</td>
<td>75-15-0</td>
<td>retinol</td>
<td>68-26-8</td>
</tr>
<tr>
<td>carbon tetrachloride</td>
<td>56-23-5</td>
<td>all-trans-retinoic acid</td>
<td>302-79-4</td>
</tr>
<tr>
<td>cinnamic aldehyde</td>
<td>104-55-2</td>
<td>13-cis-retinoic acid</td>
<td>4759-46-2</td>
</tr>
<tr>
<td>dimethylacetamide</td>
<td>127-19-5</td>
<td>etretinate</td>
<td>54350-48-0</td>
</tr>
<tr>
<td>dimethylformamide</td>
<td>69-12-2</td>
<td>acitretin</td>
<td>55079-83-9</td>
</tr>
<tr>
<td>D-penicillamine</td>
<td>52-67-6</td>
<td>sodium arsenate</td>
<td>10048-05-0</td>
</tr>
<tr>
<td>ethanol</td>
<td>64-17-5</td>
<td>sodium arsenate</td>
<td>10048-05-0</td>
</tr>
<tr>
<td>ethylene thioarsine</td>
<td>96-45-7</td>
<td>2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD)</td>
<td>1746-01-6</td>
</tr>
<tr>
<td>ethylhexanoic acid</td>
<td>149-57-5</td>
<td>theophyline</td>
<td>58-55-9</td>
</tr>
<tr>
<td>fumonisin D1</td>
<td>116355-03-0</td>
<td>valeric acid</td>
<td>92-66-1</td>
</tr>
<tr>
<td>glycol ethers</td>
<td></td>
<td>warfarin</td>
<td>81-81-2</td>
</tr>
<tr>
<td>ethylene glycol</td>
<td>107-21-1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ethylene glycol monomethyl ether</td>
<td>109-86-4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ethylene glycol monooctyl ether</td>
<td>110-80-5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ethylene glycol monoethyl ether</td>
<td>111-76-2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ethylene glycol mononaphyl ether</td>
<td>122-99-6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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
General introduction

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
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, 9-cis-retinoic acid, TTNPB: (E)-4[2-(5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2-naphthalenyl)-1-propenyl]benzoic acid.

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.

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
Chapter 1

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:

\[
\frac{dA_L}{dt} = k_{\text{uptake}} \cdot AGI(t) \\
+ QL \cdot (CB - CL/PL) \\
- V_{\text{max}} \cdot CL/PL / (K_m + CL/PL)
\]

This equation describes that the change in the amount of the chemical (A) in the liver (L) in time (t), represented as \(dA_L/dt\), 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 \(k_{\text{uptake}} \cdot AGI(t)\), with \(k_{\text{uptake}}\) (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 (\(dC_t/dt = k^*dC\)), 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 \( -\frac{V_{\text{max}} \cdot \text{CL/PL}}{K_m + \text{CL/PL}} \) with the negative sign reflecting that this contribution results in a decrease in the amount of the chemical in the liver and \( V_{\text{max}} \) 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 \( \text{CL/PL} \), 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 \textit{in vitro-in vivo} extrapolations (Figure 1). With these \textit{in vitro-in vivo} extrapolations using PBK modeling with reverse dosimetry, the following aspects should be considered: 1) the relation between the \textit{in vitro} toxic effect concentration and the predicted \textit{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 \textit{in vitro} assay compared with the \textit{in vivo} situation.

1) The relation between the \textit{in vitro} toxic effect concentration and the predicted \textit{in vivo} toxic effect dose

With the translation of \textit{in vitro} concentrations into \textit{in vivo} doses, different approaches can be used to relate an \textit{in vitro} toxic effect to the predicted \textit{in vivo} toxic effect. In previous work on predicting \textit{in vivo} developmental toxicity by translating \textit{in vitro} developmental toxicity data into \textit{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 \textit{in vivo} situation (Verwei et al., 2006). The obtained predictions for \textit{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 \textit{in vitro} data is translated to the \textit{in vivo} situation, but complete \textit{in vitro} concentration-response curves are translated into complete \textit{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 \textit{in vitro} concentration-response curve. The resulting predicted BMD values can then be compared with BMD values obtained from data on \textit{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 \textit{in vitro} assay and the predicted \textit{in vivo} effect should be considered as well. In order to be useful to predict \textit{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 \textit{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\textsubscript{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\textsubscript{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
Chapter 1

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
Chapter I


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.


Chapter 1


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.


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)
Chapter 2

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 EC50 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 EC50 values obtained in the proliferation assay or the EC50 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 EC50 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).
Chapter 2

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.)

<table>
<thead>
<tr>
<th>Test system</th>
<th>Readout(s)</th>
<th>Chemical(s)</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mouse embryonic stem cells (ES-D3)</td>
<td>mRNA and protein expression levels of several neuronal marker genes</td>
<td>methylmercury</td>
<td>Theunissen et al. (2010)</td>
</tr>
<tr>
<td>mouse embryonic stem cells (ES-D3)</td>
<td>transcriptome analyses</td>
<td>methylmercury</td>
<td>Theunissen et al. (2011a)</td>
</tr>
<tr>
<td>mouse embryonic stem cells (CGR8)</td>
<td>mRNA expression levels of neuronal marker genes</td>
<td>methylmercury</td>
<td>Zimmer et al. (2011a)</td>
</tr>
<tr>
<td>mouse embryonic stem cells (R1)</td>
<td>mRNA expression levels of myodermal, endodermal and ectodermal genes</td>
<td>valproic acid, carbamazepine</td>
<td>Murabo et al. (2007a,b)</td>
</tr>
<tr>
<td>mouse embryonic stem cells (ES-D2)</td>
<td>alpha-tetrahydrocannabinol induced green fluorescent protein (GFP) protein expression levels</td>
<td>aprenin</td>
<td>Paparrilia et al. (2002)</td>
</tr>
<tr>
<td>mouse embryonic stem cells (ES-D3)</td>
<td>sarcomeric myosin heavy chain and alpha-actin protein expression levels</td>
<td>5-fluorouracil, all-trans-retinoic acid, prostaglandin G</td>
<td>Seiler et al. (2004)</td>
</tr>
<tr>
<td>mouse embryonic stem cells (ES-D3)</td>
<td>mRNA expression levels of visceral endodermal and mesodermal genes</td>
<td>5-fluorouracil</td>
<td>Ramies et al. (2010)</td>
</tr>
<tr>
<td>mouse embryonic stem cells (ES-D3)</td>
<td>mRNA expression levels of markers for mature neural cells, osteoblasts and chondrocytes</td>
<td>pentocillin G, 5-fluorouracil, all-trans-retinoic acid, diphenhydantoin, valproic acid, trancisomide,</td>
<td>Zuber Nieden et al. (2004)</td>
</tr>
<tr>
<td>mouse embryonic stem cells (ES-D3)</td>
<td>mRNA expression levels of several neuronal marker genes</td>
<td>methylmercury</td>
<td>Stumman et al. (2009)</td>
</tr>
<tr>
<td>mouse embryonic stem cells (ES-D3)</td>
<td>transcriptome analyses</td>
<td>5-fluorouracil, monobutylphthalate, monomethylphthalate, monoethylhexylphthalate, fluoxazol, hexaconazole, triadimefon, carbamazepine, methylmercury, methoxyacetic acid, valproic acid, all-trans-retinoic acid, warfarin, norten, methotrexate, saccharine, pentocillin G</td>
<td>Van Dantel et al. (2009)</td>
</tr>
<tr>
<td>Test system</td>
<td>Readout(s)</td>
<td>Chemical(s)</td>
<td>Reference(s)</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-----------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>mouse embryonic stem cells (ES-D3)</td>
<td>proteome analyses</td>
<td>monobutylyphthalate</td>
<td>Osman et al. (2010)</td>
</tr>
<tr>
<td>mouse embryonic stem cells (ES-D3)</td>
<td>proteome analyses</td>
<td>warfarin, lovastatin</td>
<td>Grobe et al. (2010a)</td>
</tr>
<tr>
<td>mouse embryonic stem cells (ES-D3)</td>
<td>proteome analyses</td>
<td>dinoseb, ochratoxin-A, nitrofen, β-aminopropionitrile, metotrexate,</td>
<td>Grobe et al. (2010b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dioxylamine succinate, D-penicillamine</td>
<td></td>
</tr>
<tr>
<td>mouse embryonic stem cells (R1)</td>
<td>transcriptome analyses</td>
<td>valproic acid and valproic acid analogues</td>
<td>Jeryl et al. (2011)</td>
</tr>
<tr>
<td>monkey embryonic stem cells (CMK-6)</td>
<td>mRNA expression of several embryonic stem cell differentiation marker genes</td>
<td>bisphenol A</td>
<td>Yamamoto et al. (2007)</td>
</tr>
<tr>
<td>human embryonic stem cells (H9)</td>
<td>transcriptome analyses</td>
<td>cytosine arabinoside</td>
<td>Jagtap et al. (2011)</td>
</tr>
<tr>
<td>human embryonic stem cells (PKU1.1)</td>
<td>cardiac specific mRNA markers</td>
<td>cyclophosphamide</td>
<td>Zhu et al. (2011)</td>
</tr>
<tr>
<td>human embryonic stem cells (WA99)</td>
<td>metabolome analyses</td>
<td>ascorbic acid, isoniazid, penicillin G, folic acid, levotiroxine, retinol,</td>
<td>West et al. (2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dioxylamine, thiamine, aspin, caffeine, depren逃离mine, dexamethasone,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>diphenylhydantoin, methotrexate, 5-fluouracil, busulfan, cytosine arabinoside,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>hydroxyurea, all-trans-retinoic acid, thalidomide, valproic acid, amiodarone,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>rifampicin, vorbarbazine, 13cis-retinoic acid, cyclophosphamide</td>
<td></td>
</tr>
</tbody>
</table>
In vitro biomarkers for developmental toxicity

**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\textsubscript{NOEC}), 50% effect concentrations (IC\textsubscript{50}) and maximum effect concentrations (IC\textsubscript{max}) are determined. In the first prediction model developed for the WEC test, the no observed effect concentrations on total morphological score (IC\textsubscript{NOEC TMS}) and the IC\textsubscript{50} values for inducing malformations (IC\textsubscript{50mal}) were used for classification, whereas in a second prediction model, the IC\textsubscript{NOEC TMS} values, the IC\textsubscript{max} values for malformations and the IC\textsubscript{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.

<table>
<thead>
<tr>
<th>Test system</th>
<th>Readout(s)</th>
<th>Chemical(s)</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>frog embryos</td>
<td>heat shock protein mRNA expression levels</td>
<td>sodium arsenate</td>
<td>Gornatt et al. (2002)</td>
</tr>
<tr>
<td>zebrafish embryos</td>
<td>mRNA expression levels of nervous system genes</td>
<td>ethanol</td>
<td>Fan et al. (2010)</td>
</tr>
<tr>
<td>zebrafish embryos</td>
<td>transcriptome analyses</td>
<td>glycol ether alkoxyacetic acid metabolites, 1,2,4-triazoles</td>
<td>Hermans et al. (2011a)</td>
</tr>
<tr>
<td>rat embryos</td>
<td>E-cadherin mRNA and protein expression levels</td>
<td>cadmium</td>
<td>Chen and Hales (1994)</td>
</tr>
<tr>
<td>rat embryos</td>
<td>heat shock protein expression levels</td>
<td>N-acetoxy-2-acetylaminofluorene, sodium arsenate, sodium salicylate</td>
<td>Mikes et al. (1994)</td>
</tr>
<tr>
<td>rat embryos</td>
<td>8-isoprostaglandin F2alpha levels</td>
<td>diclofenac</td>
<td>Chan et al. (2002)</td>
</tr>
<tr>
<td>rat embryos</td>
<td>transcriptome analyses</td>
<td>caffeine, methylmercury, monobutylphthalate, methoxyacetic acid</td>
<td>Robinson et al. (2010)</td>
</tr>
<tr>
<td>rat embryos</td>
<td>proteome analyses</td>
<td>all-trans-retinoic acid</td>
<td>Luján et al. (2011)</td>
</tr>
</tbody>
</table>
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$_{50,\text{malformation}}$ (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 rerio) 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 Blaauwboer, 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.
In vitro biomarkers for developmental toxicity

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 (Louise 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.

In summary, the in vitro data of MAA and EAA were extrapolated to in vivo dose-response curves for developmental toxicity of EGME and EGEE using PBK modeling. The in vivo data represent the fraction of affected embryos after exposure to EGME and EGEE.
Chapter 2

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 cyclophosphamide 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 (EC₅₀ values) for the inhibition of ES-D₃ cell differentiation and the induction of malformations (EC₅₀ 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<sub>50</sub> values, respectively. For the
other chemicals, the EC<sub>50</sub> values for both assays differ less than 10-fold (Figure 2).
The results presented in Figure 2 indicate that although the EC<sub>50</sub> 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](image_url)

Figure 2. Average EC<sub>50</sub> values for chemicals affecting ES-D3 cell differentiation (EC<sub>50</sub> ES-D3 cell
differentiation) compared with average EC<sub>50</sub> values for chemicals inducing malformations in cultured
rat embryos (EC<sub>50</sub> malformations WEC). The short dashed lines indicate EC<sub>50</sub> ES-D3 cell differentiation
= EC<sub>50</sub> malformations WEC. The long dashed line covers the area in which the EC<sub>50</sub> ES-D3 cell
differentiation values do not differ more than 10-fold from the EC<sub>50</sub> malformations WEC values. Average
EC<sub>50</sub> 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: penty1-4-y1-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).
Chapter 2

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; Blauboer, 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\textsubscript{10} (lower limit of the 95% confidence interval on the benchmark dose at which a benchmark response equivalent to a 10% effect size (BMR\textsubscript{10}) is reached (BMD\textsubscript{10})), enabling derivation of safe exposure limits for humans using animal-free approaches.

References


In vitro biomarkers for developmental toxicity
Chapter 2


**In vitro biomarkers for developmental toxicity**


Chapter 2


Riebeling, C., Pirow, R., Becker, K., Buesen, R., Eikel, D., Kaltenhäuser, J., Meyer, F., Nau, H., Slawik, B.,
In vitro biomarkers for developmental toxicity


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.


In vitro biomarkers for developmental toxicity

Toxicol Appl Pharmacol 247, 18-27.


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:
Chapter 3

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 (pHi) 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, 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, 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, in the mechanism of developmental toxicity of both chemicals. Together, the results presented indicate that a decrease of the pH, may be the mechanism of developmental toxicity of the alkoxyacetic acid metabolites of the glycol ethers.
Glycol ether-induced decrease of the pH

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).

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 BMC₅ (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 BMC₅₀ 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 alkoxycetic 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 (pHᵢ) of embryonic cells (Nau and Scott, 1986). A decrease of the pHᵢ of embryonic cells by the drugs acetazolamide

Figure 1. EGME is metabolized to its toxic metabolite MAA by alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (ALDH) (Miller et al., 1983).
Chapter 3

(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 $\text{Na}^+/\text{H}^+$-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 $\text{Na}^+/\text{H}^+$-antiporter by amiloride, impeding the cells to transport the excess of $\text{H}^+$ 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 $\text{Na}^+/\text{H}^+$-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). $\text{Bis}(2\text{-carboxyethyl})$-$\text{5-(and-6)}$-carboxyfluorescein (BCECF), BCECF acetoxymethyl ester (BCECF-AM) and nigericin were purchased from Molecular Probes.
Glycol ether-induced decrease of the pH

(Invitrogen, Breda, the Netherlands). Ethanol, K$_2$HPO$_4$ and KH$_2$PO$_4$ 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 $\lambda_{ex}=490$ nm, $\lambda_{em}=535$ nm is pH-dependent, whereas BCECF fluorescence at $\lambda_{ex}=440$ nm, $\lambda_{em}=535$ nm is not (Owen, 1992). The BCECF fluorescence ratio ($\lambda_{ex}=490$ nm, $\lambda_{em}=535$ nm)/($\lambda_{ex}=440$ nm, $\lambda_{em}=535$ 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 \( \mu \text{M} \) nigericin to equilibrate the pH 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 \( \mu \text{M} \). Solutions containing BCECF were added to a ViewPlate-96-well plate (Perkin Elmer, Groningen, the Netherlands; 200 \( \mu \text{l/well} \)). BCECF fluorescence at \( \lambda_{ex}=490 \text{ nm} \), \( \lambda_{em}=535 \text{ nm} \) and \( \lambda_{ex}=440 \text{ nm} \), \( \lambda_{em}=535 \text{ nm} \) 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 \( \mu \text{M} \) 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 \( \mu \text{M} \) nigericin to final cell concentrations of 0.5 \( \cdot 10^6 \text{ cells/ml} \). From these solutions, 200 \( \mu \text{l} \) 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 \( \mu \text{M} \) 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 values.

**pH 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 \( \mu \text{M} \) BCECF-AM in HBSS and subsequently washed with HBSS. Cells were diluted to a concentration of \( 1 \cdot 10^6 \text{ cells/ml} \) in HBSS of which 100 \( \mu \text{l} \) was transferred to the inner 60 wells of a ViewPlate-96-well plate. Subsequently, 100 \( \mu \text{l} \) 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 \( \lambda_{ex}=490 \text{ nm} \), \( \lambda_{em}=535 \text{ nm} \) and \( \lambda_{ex}=440 \text{ nm} \), \( \lambda_{em}=535 \text{ nm} \) 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 · 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<sub>i</sub> measurements in differentiating EBs (day 5 EBs) and differentiated EBs (day 10 EBs)**

To measure the effect of MAA on the pH<sub>i</sub> 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<sub>i</sub> 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, of day 10 EBs, day 5 EBs that had not been used for pH 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, 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 (λex=490 nm, λem=535 nm/ λex=440 nm, λem=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 pHi experiments performed in BALB/c-3T3 and ES-D3 cells into actual pHi values.

Figure 2. BCECF fluorescence ratios ((λex=490 nm, λem=535 nm)/(λex=440 nm, λem=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 pHi 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 pHi 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 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).
Glycol ether-induced decrease of the pH<sub>i</sub>

Figure 4. Decrease of pH<sub>i</sub> of BALB/c-3T3 cells exposed for 30 min to MAA, EAA, BAA or PAA indicated by decreasing BCECF fluorescence ratio values (λ<sub>ex</sub>=490 nm, λ<sub>em</sub>=535 nm; λ<sub>ex</sub>=440 nm, λ<sub>em</sub>=535 nm). The fluorescence ratios were converted into pH<sub>i</sub> 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<sub>i</sub> 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<sub>i</sub> of BALB/c-3T3 cells as indicated by BCECF fluorescence ratio values (λ<sub>ex</sub>=490 nm, λ<sub>em</sub>=535 nm; λ<sub>ex</sub>=440 nm, λ<sub>em</sub>=535 nm). The fluorescence ratios were converted into pH<sub>i</sub> 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<sub>i</sub> values does not follow a linear scale since the calibration curve is also non-linear.
MMA 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 ($\lambda_{ex}=490$ nm, $\lambda_{em}=535$ nm/$\lambda_{ex}=440$ nm, $\lambda_{em}=535$ 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
pHi of BALB/c-3T3 cells. MAA also causes a concentration-dependent decrease of the pHi 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 pHi was found at MAA concentrations of 2.8 and 5.6 mM, whereas in day 10 EBs, only a significant decrease of the pHi 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...
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).
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-D$_3$ 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-D$_3$ 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\textsubscript{i} of embryonic cells in vitro. ACZ decreased the pH\textsubscript{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\textsubscript{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\textsubscript{i} of the cells (Scott et al., 1990). The decrease of the pH\textsubscript{i} caused by ACZ was rather small in the BALB/c-3T3 cells compared to the decrease of the pH\textsubscript{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\textsubscript{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.

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Concentration inhibiting ES-D3 cell differentiation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACZ</td>
<td>BMC\textsubscript{50}: 2.2 mM\textsuperscript{a}</td>
</tr>
<tr>
<td>VPA</td>
<td>ID\textsubscript{50}: 0.3 - 0.5 mM\textsuperscript{b}</td>
</tr>
<tr>
<td>MAA</td>
<td>BMC\textsubscript{50}: 2.3 - 2.5 mM\textsuperscript{a}</td>
</tr>
<tr>
<td>EAA</td>
<td>BMC\textsubscript{50}: 2.9 - 3.9 mM\textsuperscript{a}</td>
</tr>
<tr>
<td>BAA</td>
<td>BMC\textsubscript{50}: 4.5 - 5.9 mM\textsuperscript{a}</td>
</tr>
<tr>
<td>PAA</td>
<td>BMC\textsubscript{50}: 4.6 - 7.8 mM\textsuperscript{a}</td>
</tr>
<tr>
<td>5-FU</td>
<td>ID\textsubscript{50}: 0.6 - 0.7 \textmu M\textsuperscript{c}</td>
</tr>
<tr>
<td>ATRA</td>
<td>ID\textsubscript{50}: 1.27 nM\textsuperscript{c}</td>
</tr>
</tbody>
</table>

\textsuperscript{a} BMC\textsubscript{50}: 50% inhibition of differentiation, determined in present study.
\textsuperscript{b} ID\textsubscript{50}: 50% inhibition of differentiation, determined by Genschow et al. (2004).
\textsuperscript{c} BMC\textsubscript{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\textsubscript{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\textsubscript{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
Chapter 3

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
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, 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
Chapter 3


Glycol ether-induced decrease of the pH.
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:
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 ChEicals) 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.,
Glycol ethers in embryonic stem cell test

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 °C and 5% CO₂ 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% CO₂, 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 λ<sub>excitation</sub> = 544 nm and λ<sub>emission</sub> = 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 · 10⁴ 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% CO₂ for three days, in lab 1. In lab 2, droplets of 20 µl of an ES-D3 cell suspension (3.75 · 10⁴ 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.
**Chapter 4**

**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 (BMC<sub>d50</sub>). 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 (BMC<sub>p50</sub>). 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 BMCp₅₀ 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 BMCp₅₀ of 6.7 mM. The non-glycol ethers TCAA and BuA resulted in lower BMCp₅₀ 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. BMCp₅₀ values for inhibition of ES-D3 cell proliferation by four glycol ether alkoxyacetic acid metabolites and two structural analogues.

<table>
<thead>
<tr>
<th>Chemical</th>
<th>BMCp₅₀ (mM)</th>
<th>lab 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAA</td>
<td>6.7 (6.3 - 7.0)</td>
<td></td>
</tr>
<tr>
<td>EAA</td>
<td>12 (11 - 13)</td>
<td></td>
</tr>
<tr>
<td>BAA</td>
<td>n.d.</td>
<td></td>
</tr>
<tr>
<td>PAA</td>
<td>9.7 (8.6 - 12)</td>
<td></td>
</tr>
<tr>
<td>BuA</td>
<td>0.9 (0.8 - 0.9)</td>
<td></td>
</tr>
<tr>
<td>TCAA</td>
<td>4.9 (4.8 - 5.1)</td>
<td></td>
</tr>
</tbody>
</table>

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 (-- --).
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 BMCd50 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

### Table 2. BMCd25 and BMCd50 values for inhibition of ES-D3 cell differentiation by four glycol ether alkoxyacetic acid metabolites and two structural analogues.

<table>
<thead>
<tr>
<th>Chemical</th>
<th>BMCd25 (mM) lab 1</th>
<th>BMCd50 (mM) lab 1</th>
<th>BMCd25 (mM) lab 2</th>
<th>BMCd50 (mM) lab 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAA</td>
<td>1.7 (2.5 - 2.0)</td>
<td>2.3 (2.1 - 2.6)</td>
<td>2.0 (1.9 - 2.2)</td>
<td>2.5 (2.3 - 2.7)</td>
</tr>
<tr>
<td>EAA</td>
<td>2.1 (1.8 - 2.5)</td>
<td>2.9 (2.6 - 3.2)</td>
<td>3.2 (2.9 - 3.5)</td>
<td>3.9 (3.7 - 4.2)</td>
</tr>
<tr>
<td>BAA</td>
<td>3.4 (3.0 - 3.8)</td>
<td>4.5 (4.1 - 5.0)</td>
<td>4.8 (4.5 - 5.2)</td>
<td>5.9 (5.5 - 6.2)</td>
</tr>
<tr>
<td>PAA</td>
<td>3.7 (3.4 - 4.0)</td>
<td>4.6 (4.6 - 5.0)</td>
<td>6.4 (6.1 - 6.8)</td>
<td>7.8 (7.5 - 8.2)</td>
</tr>
<tr>
<td>BuA</td>
<td>0.4 (0.4 - 0.5)</td>
<td>0.6 (0.5 - 0.6)</td>
<td>0.3 (0.2 - 0.3)</td>
<td>0.4 (0.3 - 0.4)</td>
</tr>
<tr>
<td>TCAA</td>
<td>2.8 (2.4 - 3.1)</td>
<td>3.6 (3.4 - 4.1)</td>
<td>3.2 (2.9 - 3.4)</td>
<td>3.9 (3.7 - 4.1)</td>
</tr>
</tbody>
</table>

The 90% confidence interval is given between brackets. BMCd25 = 25% reduction of the fraction of wells containing contracting cardiomyocytes; BMCd50 = 50% reduction of the fraction of wells containing contracting cardiomyocytes.
Chapter 4

ES-D3 cell proliferation assay. This is indicated by the differences between the BMCP<sub>50</sub> and the BMC<sub>d50</sub> 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 BMC<sub>d50</sub> values obtained at the two labs (Figure 3). Interlaboratory variation is represented by the R<sup>2</sup>, where the R<sup>2</sup> can range between 0 (no correlation) and 1 (equal values). The R<sup>2</sup> of 0.89 indicates a good correlation between the two laboratories whereas the BMC<sub>d50</sub> 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 BMC<sub>d50</sub> values of MAA, EAA, BAA, PAA, BuA and TCAA in the ES-D3 cell differentiation assay of the EST.](image)

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 BMC<sub>d</sub> 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 BMC<sub>d25</sub>...
values for each of the chemicals at both labs (Table 2). The choice for \( \text{BMCd}_{25} \) instead of \( \text{BMCd}_{50} \) values did not affect the ranking within the group of chemicals. Furthermore, using an even lower BMR of 10% and using the \( \text{BMCd}_{10} \) 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 \( \text{BMCd}_{50} \) values with different models for each of the chemicals (Table 3). For the \( \text{BMCd}_{50} \) 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 \( \text{BMCd}_{25} \) or \( \text{BMCd}_{10} \) values and did not modify their ranking (data not shown).

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

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Lab 1</th>
<th>Gamma</th>
<th>LogLogistic</th>
<th>LogProbit</th>
<th>Weibull</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAA</td>
<td>2.4 (2.1 - 2.7)</td>
<td>2.5 (2.3 - 2.7)</td>
<td>2.5 (2.3 - 2.7)</td>
<td>2.3 (2.1 - 2.6)</td>
<td></td>
</tr>
<tr>
<td>EAA</td>
<td>2.7 (2.4 - 3.0)</td>
<td>2.4 (2.2 - 2.6)</td>
<td>2.6 (2.4 - 2.8)</td>
<td>2.9 (2.6 - 3.2)</td>
<td></td>
</tr>
<tr>
<td>DAA</td>
<td>4.5 (4.1 - 4.9)</td>
<td>4.5 (4.1 - 4.8)</td>
<td>4.5 (4.2 - 4.9)</td>
<td>4.5 (4.1 - 5.0)</td>
<td></td>
</tr>
<tr>
<td>PAA</td>
<td>4.6 (4.3 - 5.0)</td>
<td>4.6 (4.3 - 4.9)</td>
<td>4.6 (4.3 - 4.9)</td>
<td>4.6 (4.6 - 5.0)</td>
<td></td>
</tr>
<tr>
<td>BuA</td>
<td>0.6 (0.5 - 0.6)</td>
<td>0.6 (0.5 - 0.6)</td>
<td>0.6 (0.5 - 0.6)</td>
<td>0.6 (0.5 - 0.6)</td>
<td></td>
</tr>
<tr>
<td>TCAA</td>
<td>3.9 (3.6 - 4.3)</td>
<td>4.0 (3.7 - 4.4)</td>
<td>4.0 (3.7 - 4.3)</td>
<td>3.8 (3.4 - 4.1)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Lab 2</th>
<th>Gamma</th>
<th>LogLogistic</th>
<th>LogProbit</th>
<th>Weibull</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAA</td>
<td>2.5 (2.3 - 2.6)</td>
<td>2.3 (2.1 - 2.5)</td>
<td>2.5 (2.3 - 2.6)</td>
<td>2.5 (2.3 - 2.7)</td>
<td></td>
</tr>
<tr>
<td>EAA</td>
<td>3.7 (3.4 - 4.0)</td>
<td>3.4 (3.1 - 3.8)</td>
<td>3.6 (3.3 - 4.0)</td>
<td>3.9 (3.7 - 4.2)</td>
<td></td>
</tr>
<tr>
<td>BAA</td>
<td>5.7 (5.4 - 6.1)</td>
<td>5.5 (5.1 - 5.9)</td>
<td>5.7 (5.3 - 6.1)</td>
<td>5.9 (5.5 - 6.2)</td>
<td></td>
</tr>
<tr>
<td>PAA</td>
<td>7.7 (7.4 - 8.1)</td>
<td>8.2 (7.6 - 8.9)</td>
<td>7.7 (7.4 - 8.1)</td>
<td>7.8 (7.5 - 8.2)</td>
<td></td>
</tr>
<tr>
<td>BuA</td>
<td>0.3 (0.3 - 0.4)</td>
<td>0.3 (0.3 - 0.4)</td>
<td>0.3 (0.3 - 0.4)</td>
<td>0.4 (0.3 - 0.4)</td>
<td></td>
</tr>
<tr>
<td>TCAA</td>
<td>3.8 (3.6 - 4.0)</td>
<td>3.4 (3.2 - 3.7)</td>
<td>3.8 (3.5 - 4.0)</td>
<td>3.9 (3.7 - 4.1)</td>
<td></td>
</tr>
</tbody>
</table>

The 90% confidence interval is given between brackets.
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 BMCd50 values obtained in the
Table 4. Developmental toxicity of glycol ethers in mice.

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Mice strain</th>
<th>Route</th>
<th>Exposure</th>
<th>Dose (mg·kg bw⁻¹·day⁻¹)</th>
<th>Reference</th>
<th>BMD₉₀ fetal weight (mmol·kg bw⁻¹·day⁻¹)</th>
<th>BMD₉₀ malformation (mmol·kg bw⁻¹·day⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGME (MAA)</td>
<td>JCL-ICR</td>
<td>G</td>
<td>GD7-14</td>
<td>0, 31, 63, 125, 250</td>
<td>Nagano et al. (1981)</td>
<td>1.3⁹</td>
<td>0.8⁹</td>
</tr>
<tr>
<td>EGEE (EAA)</td>
<td>CD-1</td>
<td>G</td>
<td>GD6-14</td>
<td>0, 1000, 1800, 2600, 3400, 4200</td>
<td>Wier et al. (1987)</td>
<td>4.9⁹</td>
<td>14⁹</td>
</tr>
<tr>
<td>t-EGEE (BAA)</td>
<td>CD-1</td>
<td>G</td>
<td>GD8-14</td>
<td>0, 350, 850, 1000, 1500, 2000</td>
<td>Wier et al. (1987)</td>
<td>12⁹</td>
<td>15⁹</td>
</tr>
<tr>
<td>EGPE (PAA)</td>
<td>CD-1</td>
<td>D</td>
<td>7PM B</td>
<td>0, 400, 2000, 4000</td>
<td>Heindel et al. (1990)</td>
<td>31⁹</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

BMD₀ fetal weight: dose (mmol·kg bw⁻¹·day⁻¹) at 10% decrease in fetal body weight; BMD₀ 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.

⁹BMDs 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* BMC$_{50}$ 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 BMC$_{50}$ 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
Glycol ethers in embryonic stem cell test

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
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.


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. Blauboer, Ruud A. Woutersen, Aldert H. Piersma, Ivonne M.C.M. Rietjens, Miriam Verwei

Based on:
Toxicological Sciences 118, 470-484 (2010).
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 \textit{in vivo} (De Jong \textit{et al.}, 2009). However, the ES-D3 cell differentiation assay provides \textit{in vitro} concentration-response curves, whereas for human risk assessment \textit{in vivo} dose-response curves are often required. These \textit{in vitro} data should therefore be translated into \textit{in vivo} data by taking into account \textit{in vivo} kinetics (Verwei \textit{et al.}, 2006).

The goal of the present study was to integrate \textit{in vitro} toxicity data and \textit{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 \textit{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 \textit{in vivo} developmental toxicity potencies. To take the \textit{in vivo} kinetics into consideration, the \textit{in vitro} data derived from the ES-D3 cell differentiation assay were used as input in a physiologically based kinetic (PBK) model, translating \textit{in vitro} effect concentrations into predicted \textit{in vivo} developmental toxicity dose levels, as described by Verwei \textit{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 \textit{in vitro} will also cause toxic effects in the developing embryo \textit{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 \textit{in vivo} dose levels of the parent glycol ethers that will lead to these blood concentrations of the toxic alkoxyacetic acid metabolites \textit{in vivo}. By calculating this dose level for each concentration tested in the ES-D3 cell differentiation assay, \textit{in vitro} concentration-response curves for the toxic alkoxyacetic acid metabolites can be converted into predicted \textit{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_{\text{90}} (lower limit of the 95\% confidence interval on the benchmark dose at which a benchmark response equivalent to a 10\% effect size (BMR_{\text{90}}) is reached (BMD_{\text{90}})), for deriving safe exposure levels for humans. Therefore, the combined \textit{in vitro-in silico} approach provides a platform to use \textit{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 (Na₂SO₄) 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.*
Chapter 5

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

<table>
<thead>
<tr>
<th>Physiological parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rat</td>
</tr>
<tr>
<td>body weight (kg)</td>
<td>0.26</td>
</tr>
<tr>
<td>tissue volume (percentage of body weight)</td>
<td></td>
</tr>
<tr>
<td>liver</td>
<td>4.0</td>
</tr>
<tr>
<td>fat</td>
<td>10.1</td>
</tr>
<tr>
<td>rapidly perfused tissue</td>
<td>6.1</td>
</tr>
<tr>
<td>slowly perfused tissue</td>
<td>65</td>
</tr>
<tr>
<td>blood</td>
<td>5.9</td>
</tr>
<tr>
<td>cardiac output ($L \cdot h^{-1} \cdot kg \cdot bw^{0.74}$)</td>
<td>14.0</td>
</tr>
<tr>
<td>tissue blood flows (percentage of cardiac output)</td>
<td></td>
</tr>
<tr>
<td>liver, hepatic artery</td>
<td>3.4</td>
</tr>
<tr>
<td>liver, portal vein</td>
<td>21.6</td>
</tr>
<tr>
<td>fat</td>
<td>14.2</td>
</tr>
<tr>
<td>rapidly perfused tissue</td>
<td>45.8</td>
</tr>
<tr>
<td>slowly perfused tissue</td>
<td>15.0</td>
</tr>
<tr>
<td>alveolar ventilation ($L \cdot h^{-1} \cdot kg \cdot bw^{0.74}$)</td>
<td>14.0</td>
</tr>
</tbody>
</table>
Table 2. Chemical-specific parameter values used in the PBK model for glycol ethers.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>EGME</td>
<td>0.76</td>
<td>0.37</td>
<td>0.76</td>
<td>0.00</td>
<td>32800&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>MAA</td>
<td>0.76</td>
<td>0.13</td>
<td>0.76</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>EGEE</td>
<td>0.77</td>
<td>0.67</td>
<td>0.77</td>
<td>0.80</td>
<td>22093&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>EAA</td>
<td>0.76</td>
<td>0.13</td>
<td>0.76</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>EGBE</td>
<td>0.73</td>
<td>3.7</td>
<td>0.73</td>
<td>0.72</td>
<td>7965&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>BAA</td>
<td>0.64</td>
<td>0.11</td>
<td>0.64</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>EGPE</td>
<td>0.55</td>
<td>6.4</td>
<td>0.55</td>
<td>0.79</td>
<td>5001&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>PAA</td>
<td>0.64</td>
<td>0.11</td>
<td>0.64</td>
<td>0.67</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Biotransformation kinetics</th>
<th>Rat</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGME → MAA&lt;sup&gt;b&lt;/sup&gt;</td>
<td>8.3</td>
<td>1511</td>
</tr>
<tr>
<td>EGEE → EAA&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6.6</td>
<td>1519</td>
</tr>
<tr>
<td>EGBE → BAA&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.9</td>
<td>741</td>
</tr>
<tr>
<td>EGPE → PAA&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.9</td>
<td>741</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Excretion kinetics&lt;sup&gt;f&lt;/sup&gt;</th>
<th>Rat</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAA</td>
<td>0.0045</td>
<td>0.18</td>
</tr>
<tr>
<td>EAA</td>
<td>0.023</td>
<td>0.30</td>
</tr>
<tr>
<td>BAA</td>
<td>0.060</td>
<td>5.0</td>
</tr>
<tr>
<td>PAA</td>
<td>0.060</td>
<td>2.0</td>
</tr>
</tbody>
</table>

<sup>a</sup> Estimated using in silico method described by Berezhkovskiy (2004).

<sup>b</sup> Determined by Johanson and Dynesius (1988).

<sup>c</sup> Estimated based on linear correlation between LogP values of EGME, EGEE and EGBE, and the logarithm of their blood:air partition coefficients.

<sup>d</sup> Data from biotransformation studies in hepatocytes (Green et al., 1996).

<sup>e</sup> Assumed to be similar as for EGBE (see the “Materials and Methods” section).

<sup>f</sup> 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: Groeseneiken et al., 1989; EAA: Groeseneiken 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.
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>
<thead>
<tr>
<th>Chemical</th>
<th>LogP$^a$</th>
<th>VanderWaals volume$^b$ ($Å^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGME</td>
<td>-0.61</td>
<td>88</td>
</tr>
<tr>
<td>EGEE</td>
<td>-0.22</td>
<td>106</td>
</tr>
<tr>
<td>EGBE</td>
<td>0.84</td>
<td>143</td>
</tr>
<tr>
<td>EGPE</td>
<td>1.19</td>
<td>153</td>
</tr>
</tbody>
</table>


$^b$ VanderWaals volumes calculated using Spartan 04 for Windows version 1.0.3.
Table 4. Published kinetic studies used to build and/or evaluate the PBK model.

<table>
<thead>
<tr>
<th>Species</th>
<th>Chemical</th>
<th>Reference</th>
<th>Exposure route</th>
<th>Exposure</th>
<th>Dose(s)</th>
<th>Data shown</th>
</tr>
</thead>
<tbody>
<tr>
<td>rat</td>
<td>EGME / MAA</td>
<td>Hays et al. (2000)</td>
<td>oral</td>
<td>once</td>
<td>3.3 mmol·kg bw$^{-1}$</td>
<td>Figure 2A</td>
</tr>
<tr>
<td></td>
<td>EGME / MAA</td>
<td>Gargas et al. (2000a)</td>
<td>inhalation</td>
<td>5 days, 6h·day$^{-1}$</td>
<td>10, 50 ppm</td>
<td>Figure 2B</td>
</tr>
<tr>
<td></td>
<td>EGEE (EGEEA) / EAA</td>
<td>Gargas et al. (2000b)</td>
<td>inhalation</td>
<td>5 days, 6h·day$^{-1}$</td>
<td>50, 100 ppm</td>
<td>Figure 2C</td>
</tr>
<tr>
<td></td>
<td>EGBE / BAA</td>
<td>Ghanayem et al. (1990)</td>
<td>intravenous</td>
<td>once</td>
<td>0.3, 0.5, 1.0 mmol·kg bw$^{-1}$</td>
<td>Figure 2D</td>
</tr>
<tr>
<td></td>
<td>EGPE / FAA</td>
<td>present study</td>
<td>oral</td>
<td>once</td>
<td>1.1, 3.3 mmol·kg bw$^{-1}$</td>
<td>Figure 2E</td>
</tr>
<tr>
<td>human</td>
<td>EGME / MAA</td>
<td>Groeseneken et al. (1989)</td>
<td>inhalation</td>
<td>4h</td>
<td>5 ppm</td>
<td>Figure 3A</td>
</tr>
<tr>
<td></td>
<td>EGEE / EAA</td>
<td>Groeseneken et al. (1986)</td>
<td>inhalation</td>
<td>4h</td>
<td>3, 6, 9 ppm</td>
<td>Figure 3B</td>
</tr>
<tr>
<td></td>
<td>EGBE / BAA</td>
<td>Jones and Cocker (2003)</td>
<td>inhalation</td>
<td>2h</td>
<td>20 ppm</td>
<td>Figure 3C</td>
</tr>
<tr>
<td></td>
<td>EGBE / BAA</td>
<td>Kezic et al. (2004)</td>
<td>inhalation</td>
<td>0.5h</td>
<td>19 ppm</td>
<td>Figures 3C, D</td>
</tr>
</tbody>
</table>
**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 Ci8 T3 5-µm column, 150 × 3.0 mm (Waters,
Chapter 5

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\(_2\)O (20/80, v/v) solution (solution A) and a 5 mM ammonium formate in methanol/H\(_2\)O (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 BMDLₙ₀ 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 BMDLₙ₀ values derived from in vivo data were compared to predicted BMDLₙ₀ values, which were acquired by translating the in vitro BMCLₙ₀ (lower limit of the 95% confidence interval on the BMCₙ₀) 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<sub>10</sub> values determined from <em>in vivo</em> developmental toxicity studies and predicted BMDL<sub>10</sub> values for developmental toxicity using the <em>in vitro-in silico</em> approach with <em>in vitro</em> BMCL<sub>10</sub> values as input (symbols shown correspond with symbols used in Figure 5). (Table continues on following page.)

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Reference</th>
<th>Exposure route</th>
<th>Days of exposure</th>
<th>Doses</th>
<th>Critical endpoint</th>
<th>Measured BMDL&lt;sub&gt;10&lt;/sub&gt;</th>
<th>Predicted BMDL&lt;sub&gt;10&lt;/sub&gt;</th>
<th>Figure</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGME</td>
<td>Torssen et al. (1983)</td>
<td>oral</td>
<td>GD7-15</td>
<td>0.3, 0.6, 1.1 mg/kg bw&lt;sup&gt;2&lt;/sup&gt; day&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>cardiac malformations</td>
<td>0.46 mg/kg bw&lt;sup&gt;2&lt;/sup&gt; day&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>0.12 mg/kg bw&lt;sup&gt;2&lt;/sup&gt; day&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>0.34 mg/kg bw&lt;sup&gt;2&lt;/sup&gt; day&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>5A</td>
</tr>
<tr>
<td>EGME</td>
<td>Dye et al. (1984)</td>
<td>inhalation</td>
<td>GD6-17, 6th day&lt;sup&gt;2&lt;/sup&gt;</td>
<td>100, 300 ppm</td>
<td>fetal deaths</td>
<td>nd</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EGEE</td>
<td>Nelson et al. (1984)</td>
<td>inhalation</td>
<td>GD7-15, 7th day&lt;sup&gt;2&lt;/sup&gt;</td>
<td>50, 100, 200 ppm</td>
<td>skeletal malformations</td>
<td>11 ppm</td>
<td>21 ppm</td>
<td>64 ppm</td>
<td>5E</td>
</tr>
<tr>
<td>EGEE</td>
<td>Osterman et al. (1984)</td>
<td>oral</td>
<td>GD7-15</td>
<td>2.2 mg/kg bw&lt;sup&gt;2&lt;/sup&gt; day&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>cardiac malformations</td>
<td>nd</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>EGEE</td>
<td>Gomberg et al. (1977)</td>
<td>oral</td>
<td>GD1-21</td>
<td>0.1, 0.3, 1.0, 2.1 mg/kg bw&lt;sup&gt;2&lt;/sup&gt; day&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>skeletal malformations</td>
<td>0.03 mg/kg bw&lt;sup&gt;2&lt;/sup&gt; day&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>0.28 mg/kg bw&lt;sup&gt;2&lt;/sup&gt; day&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>1.2 mg/kg bw&lt;sup&gt;2&lt;/sup&gt; day&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>69</td>
</tr>
<tr>
<td>EGEE</td>
<td>Chester et al. (1986)&lt;sup&gt;2&lt;/sup&gt;</td>
<td>oral</td>
<td>GD7-15</td>
<td>2.3, 3.0, 4.4 mg/kg bw&lt;sup&gt;2&lt;/sup&gt; day&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>fetal deaths</td>
<td>nd</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>EGEE</td>
<td>Dye (1984)</td>
<td>inhalation</td>
<td>GD6-15, 6th day&lt;sup&gt;2&lt;/sup&gt;</td>
<td>10, 30, 25 ppm</td>
<td>skeletal malformations&lt;sup&gt;2&lt;/sup&gt;</td>
<td>66 ppm</td>
<td>62 ppm</td>
<td>256 ppm</td>
<td>5F</td>
</tr>
<tr>
<td>EGEE</td>
<td>Anders and Hamel (1984)</td>
<td>inhalation</td>
<td>GD7-15, 7th day&lt;sup&gt;2&lt;/sup&gt;</td>
<td>200, 770 ppm</td>
<td>resorptions</td>
<td>nd</td>
<td></td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

<sup>1</sup> Includes cardiovascular abnormalities.
<sup>2</sup> Includes resorptions.
Reverse dosimetry glycol ethers

**EGEE**

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Reference</th>
<th>Exposure route</th>
<th>Days of exposure</th>
<th>Doses</th>
<th>Critical endpoint</th>
<th>Measured BMDL&lt;sub&gt;10&lt;/sub&gt;</th>
<th>Predicted BMDL&lt;sub&gt;10&lt;/sub&gt;</th>
<th>Figure</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Blain et al. (1989)</td>
<td>oral</td>
<td>GD9-11</td>
<td>1.3, 2.5 mg kg&lt;sup&gt;-1&lt;/sup&gt; day&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>n.d.</td>
<td>1.4 mg kg&lt;sup&gt;-1&lt;/sup&gt; day&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>2.6 mg kg&lt;sup&gt;-1&lt;/sup&gt; day&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>3.1 mg kg&lt;sup&gt;-1&lt;/sup&gt; day&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>TC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GD11-1C</td>
<td>1.3, 2.5, 3.1 mg kg&lt;sup&gt;-1&lt;/sup&gt; day&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
<td>▲</td>
</tr>
<tr>
<td></td>
<td>Tyl et al. (1994)</td>
<td>inhalation</td>
<td>GD6-1, 3, 9-Log&lt;sup&gt;10&lt;/sup&gt;</td>
<td>20, 30, 100, 300 ppm</td>
<td>n.d.</td>
<td>3.6 mg kg&lt;sup&gt;-1&lt;/sup&gt; day&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>2.6 mg kg&lt;sup&gt;-1&lt;/sup&gt; day&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>3.1 mg kg&lt;sup&gt;-1&lt;/sup&gt; day&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>▲</td>
</tr>
<tr>
<td></td>
<td>Unilever Research (1994)</td>
<td>subcutaneous</td>
<td>GD6-15</td>
<td>0.5, 1, 3, 5 mg kg&lt;sup&gt;-1&lt;/sup&gt; day&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>n.d.</td>
<td>1.0 mg kg&lt;sup&gt;-1&lt;/sup&gt; day&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>2.4 mg kg&lt;sup&gt;-1&lt;/sup&gt; day&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>6.2 mg kg&lt;sup&gt;-1&lt;/sup&gt; day&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>▲</td>
</tr>
</tbody>
</table>

n.d.: BMDL<sub>10</sub> 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.
Chapter 5

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).
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; Km C: 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 Kex. 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, to BMC_{50} 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 · day^{-1}). Predicted dose-response curves in these figures are shown for the studies using a 7 h · day^{-1} exposure regimen (Figures 5E-H). When applying a 6 h · 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
Reverse dosimetry glycol ethers

Diagram A: Fraction affected vs dose EGME (mmol kg bw⁻¹)
Diagram B: Fraction affected vs dose EGEE (mmol kg bw⁻¹)
Diagram C: Fraction affected vs dose EGBE (mmol kg bw⁻¹)
Diagram D: Fraction affected vs dose EGPE (mmol kg bw⁻¹)
Diagram E: Fraction affected vs dose EGME (ppm)
Diagram F: Fraction affected vs dose EGEE (ppm)
Diagram G: Fraction affected vs dose EGBE (ppm)
Diagram H: Fraction affected vs dose EGPE (ppm)
5, together with the predicted BMDL_{10} values, which were acquired by translating the \textit{in vitro} BMCL_{10} values into \textit{in vivo} dose levels using the rat PBK model. When comparing \textit{in vivo}-derived and \textit{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 \textit{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.

\textit{Prediction of \textit{in vivo} dose-response curves for developmental toxicity in human}

The predicted dose-response curves for \textit{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 \textit{in vitro} concentration-response curves (BMC\textsubscript{c} to BMC\textsubscript{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 up to BMC_{99} from laboratory 1 (straight lines) or laboratory 2 (dashed lines) into in vivo dose levels using the PBK model.
Chapter 5

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 · 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).

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Lab 1</th>
<th>Lab 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-day exposure (ppm)</td>
<td>5-day exposure (ppm)</td>
</tr>
<tr>
<td>EGME</td>
<td>110</td>
<td>28</td>
</tr>
<tr>
<td>EGEE</td>
<td>81</td>
<td>24</td>
</tr>
<tr>
<td>EGPE</td>
<td>1300</td>
<td>720</td>
</tr>
</tbody>
</table>

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 \textit{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 \textit{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 \textit{in vitro} developmental toxicity (Genschow \textit{et al.}, 2004), which in the ideal situation could represent the most sensitive \textit{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 \textit{in vitro} effect concentrations that inhibit cardiomyocyte differentiation. Therefore, for predicting BMDL\(_{10}\) values for other developmental toxicity endpoints such as developmental neurotoxicity, an \textit{in vitro} assay focusing on this specific endpoint (Breier \textit{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 \textit{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
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 \textit{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 \textit{et al.}, 1990), EAA (7.6-10 h; Aasmoe and Aarbakke, 1997; Aasmoe \textit{et al.}, 1999) and MAA (14-19 h; Aasmoe and Aarbakke, 1997; Aasmoe \textit{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 \textit{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 \textit{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 \textit{in vitro} by 10%. The predicted dose-response curve for developmental toxicity of EGPE for human, using the \textit{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 \textit{in vitro} toxicity data with human PBK models may predict safe dose levels for human, without the need for applying an uncertainty factor for
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 \textit{in vivo} kinetics in the human population into account (chapter 7). This would enable a better prediction of dose-response curves for \textit{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 \textit{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\textsubscript{10} values and the BMDL\textsubscript{10} values derived from \textit{in vivo} studies for glycol ethers are small considering the experimental differences in BMDL\textsubscript{10} values that might be derived from \textit{in vivo}
(developmental) toxicity studies. The predicted BMDL\textsubscript{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 \textit{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**


Chapter 5


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.


Reverse dosimetry glycol ethers


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 \textit{in vivo} and two other \textit{in vitro} assays for developmental toxicity

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

Based on: \textit{Toxicology Letters} 203, 1-8 (2011).
Chapter 6

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 \textit{in vivo} and two other \textit{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 \textit{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-\textit{trans}-retinoic acid, 13-\textit{cis}-retinoic acid, 9-\textit{cis}-retinoic acid and acitretin showed an intermediate potency. \textit{In vivo} potency rankings of the developmental toxicity of retinoids appear to be dependent on the species and/or exposure regimens used. The obtained \textit{in vitro} potency ranking does not completely correspond with the \textit{in vivo} potency rankings, although TTNPB is correctly predicted to be the most potent and retinol the least potent congener. The lack of \textit{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 \textit{in vivo} kinetics is essential when using \textit{in vitro} toxicity data for the estimation of \textit{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.
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 × 10 mm, Corning, Schiphof-Rijk, The Netherlands) at 37 °C in a humidified atmosphere of 5% CO₂ 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 · 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 handleings, 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.
Chapter 6

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<sub>50</sub> 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<sub>50</sub> 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<sub>50</sub> concentrations, and retinol, 13-cis-RA, 9-cis-RA, etretinate and TTNPB at their BMC<sub>90</sub> 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.

<table>
<thead>
<tr>
<th>ES-D3 cell differentiation assay</th>
<th>Present study</th>
<th>TTNPE</th>
<th>ATRA</th>
<th>13-cis-RA</th>
<th>9-cis-RA</th>
<th>acetretin</th>
<th>etretinate</th>
<th>retinol</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMC_{50} (nM)</td>
<td>0.17 (±0.026)</td>
<td>3.3 (±1.8)</td>
<td>5.0 (±3.3)</td>
<td>10 (±8.6)</td>
<td>12 (±6.3)</td>
<td>1500 (±610)</td>
<td>2800 (±1200)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MM test</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kistler (1987)^a</td>
<td>TTNPE</td>
<td>actretin</td>
<td>ATRA</td>
<td>13-cis-RA</td>
<td>etretinate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EC_{50} (nM)</td>
<td>0.06</td>
<td>50</td>
<td>80</td>
<td>400</td>
<td>3000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reiners et al. (1988)^b</td>
<td>actretin</td>
<td>ATRA</td>
<td>13-cis-RA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EC_{50} (nM)</td>
<td>37</td>
<td>50</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kochhar et al. (1990)^b</td>
<td>ATRA</td>
<td>9-cis-RA</td>
<td>13-cis-RA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EC_{50} (nM)</td>
<td>15</td>
<td>31</td>
<td>500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>WEC test</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Morris and Steele (1977)</td>
<td>ATRA</td>
<td>retinol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bechter and Hall (1987)</td>
<td>acetretin</td>
<td>etretinate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steele et al. (1987)</td>
<td>ATRA</td>
<td>13-cis-RA</td>
<td>acetretin</td>
<td>etretinate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cittarel and Schmid (1988)</td>
<td>TTNPE</td>
<td>ATRA</td>
<td>13-cis-RA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tseh (1988)</td>
<td>ATRA</td>
<td>13-cis RA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kug et al. (1989)</td>
<td>ATRA</td>
<td>13-cis-RA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^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.
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 BMD₁₀ 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 BMD₁₀ value for total malformations and a BMD₁₀ value for cleft palate incidence were obtained (Table 2). Using the BMD₁₀ 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.)

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Animal</th>
<th>Day(s) of exposure</th>
<th>Doses (mg/kg bw(^{-1})day(^{-1}))</th>
<th>Reference</th>
<th>BMD(_{10}) ((\mu)mol/kg bw(^{-1})day(^{-1}))</th>
<th>total malformations</th>
<th>cleft palate incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: rats exposed on GD9 or GD10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I NPH</td>
<td>F(_{u})-albino rat</td>
<td>GD9</td>
<td>0, 0.003, 0.01, 0.03, 0.1</td>
<td>Kaster et al. (1990)</td>
<td>0.071</td>
<td>n.d.</td>
<td></td>
</tr>
<tr>
<td>ATRA</td>
<td>Wistar rat</td>
<td>GD10</td>
<td>0, 20, 30, 50, 80, 100</td>
<td>Tembe et al. (1990)</td>
<td>n.d.</td>
<td>n.d.</td>
<td></td>
</tr>
<tr>
<td>retinol</td>
<td>Wistar rat</td>
<td>GD10</td>
<td>0, 50, 100, 200, 500</td>
<td>Tembe et al. (1996)</td>
<td>n.d.</td>
<td>n.d.</td>
<td></td>
</tr>
<tr>
<td>13-cis-RA</td>
<td>Wistar rat</td>
<td>GD10</td>
<td>0, 100, 200, 500, 1000</td>
<td>Tembe et al. (1996)</td>
<td>n.d.</td>
<td>n.d.</td>
<td></td>
</tr>
<tr>
<td>B: rats exposed on GD10-11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATRA</td>
<td>Wistar rat</td>
<td>GD10-11</td>
<td>0, 6.25, 12.5, 25, 50, 100</td>
<td>Turton et al. (1992)</td>
<td>5.6</td>
<td>n.d.</td>
<td></td>
</tr>
<tr>
<td>13-cis-RA</td>
<td>Wistar rat</td>
<td>GD10-11</td>
<td>0, 100, 200</td>
<td>Tembe et al. (1992)</td>
<td>n.d.</td>
<td>n.d.</td>
<td></td>
</tr>
<tr>
<td>C: mice exposed on GD11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATRA</td>
<td>CD-1 mouse</td>
<td>GD11</td>
<td>0, 10, 25, 50, 100</td>
<td>Kochhar et al. (1995)</td>
<td>n.d.</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>ATRA</td>
<td>CD-1 mouse</td>
<td>GD11</td>
<td>0, 2.5, 10, 30, 50, 100</td>
<td>Campbell et al. (2004)</td>
<td>10</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>etretinate</td>
<td>CD-1 mouse</td>
<td>GD11</td>
<td>0, 10, 25, 100, 200</td>
<td>Kochhar et al. (1999)</td>
<td>76</td>
<td>n.d.</td>
<td></td>
</tr>
<tr>
<td>13-cis-RA</td>
<td>CD-1 mouse</td>
<td>GD11</td>
<td>0, 75, 150, 300, 600, 900</td>
<td>Kwassigroh and Bullen (1991)</td>
<td>n.d.</td>
<td>140</td>
<td></td>
</tr>
</tbody>
</table>
Retinoids in embryonic stem cell test

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Animal</th>
<th>Day(s) of exposure</th>
<th>Doses (mg kg bw⁻¹ day⁻¹)</th>
<th>Reference</th>
<th>( \text{BMD}_\alpha ) (( \mu \text{mol} \text{kg bw}^{-1} \text{day}^{-1} ))</th>
<th>( \text{BMD}_\alpha ) value for total skeletal malformations</th>
<th>cleft palate incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>D: mice exposed on GD7-16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNPB</td>
<td>Fj-abino mouse</td>
<td>GD7-16</td>
<td>0, 0.0003, 0.001, 0.003, 0.01</td>
<td>Kistler et al. (1987)</td>
<td>0.0092</td>
<td>-</td>
<td>n.d.</td>
</tr>
<tr>
<td>acitretin</td>
<td>Fj-abino mouse</td>
<td>GD7-16</td>
<td>0, 1, 3, 10</td>
<td>Kistler et al. (1987)</td>
<td>6.6*</td>
<td>n.d.</td>
<td></td>
</tr>
<tr>
<td>13-cri-RA</td>
<td>Fj-abino mouse</td>
<td>GD7-16</td>
<td>0, 10, 30, 100, 300</td>
<td>Kistler et al. (1987)</td>
<td>320</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>E: rabbits exposed on GD7-19 or GD6-18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNPBR</td>
<td>Swiss hare rabbit</td>
<td>GD7-19</td>
<td>0, 0.0003, 0.001, 0.003</td>
<td>Kistler et al. (1987)</td>
<td>0.0027</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>acitretin</td>
<td>Swiss hare rabbit</td>
<td>GD7-19</td>
<td>0, 0.2, 0.6, 2</td>
<td>Kistler and Hummier (1985)</td>
<td>2.5</td>
<td>3.6</td>
<td></td>
</tr>
<tr>
<td>ATRA</td>
<td>Swiss hare rabbit</td>
<td>GD6-18</td>
<td>0, 0.7, 2, 6</td>
<td>Tzimas et al. (1994)</td>
<td>21</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>13-cri-RA</td>
<td>Swiss hare rabbit</td>
<td>GD6-18</td>
<td>0, 3, 7,5, 10</td>
<td>Tzimas et al. (1991)</td>
<td>26</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

n.d.: \( \text{BMD}_\alpha \) value could not be determined because of inadequate data.

*: no data available on this endpoint.

* \( \text{BMD}_\alpha \) 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\textsubscript{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\textsubscript{10} values, was obtained for retinol (Table 2), as was also correctly predicted by the ES-D3 cell differentiation assay. However, although no BMD\textsubscript{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
Chapter 6

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 Blaauuboer for their stimulating contributions within the framework of the project.

References


Chapter 6


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. Blauboer, Ivonne M.C.M. Rietjens, Miriam Verwei

In preparation.
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 Blauboer, 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.
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 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 \( \text{[}^{11,12-3}\text{H(N)}\text{]} \)-ATRA, \( \text{[}^{1-}\text{methyl-14C}] \)-caffeine and \( \text{D-[}^{1-3}\text{H(N)}\text{]} \)-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, Biltoven, 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).

<table>
<thead>
<tr>
<th>Physiological parameters</th>
<th>Values</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rat</td>
<td>human</td>
</tr>
<tr>
<td>tissue volumes (percentage of body weight)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>liver</td>
<td>2.7</td>
<td>2.6</td>
</tr>
<tr>
<td>muscle</td>
<td>44</td>
<td>40</td>
</tr>
<tr>
<td>kidney</td>
<td>0.54</td>
<td>0.40</td>
</tr>
<tr>
<td>heart</td>
<td>0.24</td>
<td>0.50</td>
</tr>
<tr>
<td>brain</td>
<td>0.42</td>
<td>2.0</td>
</tr>
<tr>
<td>lung</td>
<td>0.37</td>
<td>0.80</td>
</tr>
<tr>
<td>fat</td>
<td>10</td>
<td>21</td>
</tr>
<tr>
<td>rest of body</td>
<td>31</td>
<td>23</td>
</tr>
<tr>
<td>blood</td>
<td>5.9</td>
<td>7.9</td>
</tr>
<tr>
<td>cardiac output (L h⁻¹ kg⁻¹ kg⁻0.75)</td>
<td>14.0</td>
<td>19.2</td>
</tr>
<tr>
<td>tissue blood flows (percentage of cardiac output)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>liver</td>
<td>17.4</td>
<td>22.7</td>
</tr>
<tr>
<td>muscle</td>
<td>27.8</td>
<td>19.1</td>
</tr>
<tr>
<td>kidney</td>
<td>14.1</td>
<td>17.5</td>
</tr>
<tr>
<td>heart</td>
<td>5.1</td>
<td>4.0</td>
</tr>
<tr>
<td>brain</td>
<td>2.0</td>
<td>11.4</td>
</tr>
<tr>
<td>lung</td>
<td>2.1</td>
<td>2.5</td>
</tr>
<tr>
<td>fat</td>
<td>7.0</td>
<td>5.2</td>
</tr>
<tr>
<td>rest of body</td>
<td>24.5</td>
<td>17.6</td>
</tr>
</tbody>
</table>
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 · 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\text{mol} \cdot \text{h}^{-1}) = P_{app} \text{ coefficient (cm} \cdot \text{h}^{-1}) \times \text{absorptive surface area (cm}^2 \times \text{luminal concentration of ATRA (mM). The absorptive surface area of the rat intestines (94 cm}^2 \text{) 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 \text{) 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) $M_t = C_b \cdot V_b + C_s \cdot V_s + C_c \cdot V_c$

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

2) $C_c = C_b \cdot K_c$

in which $C_c = \text{concentration in coating, } C_b = \text{concentration in buffer and } K_c = \text{coating:buffer partition coefficient. The concentration of the chemical in the serum$
can be described as:

\[ C_s = C_b \times K_s \]

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

These equations can be rewritten as:

\[ A_c = M_t / (1 + ((V_b + K_s \times V_s) / (V_c \times K_c))) \]

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 \([11,12-^3H(N)]\)-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
Reverse dosimetry all-trans-retinoic acid

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 homogenate) 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.

<table>
<thead>
<tr>
<th>Species</th>
<th>In vitro system</th>
<th>Cofactor</th>
<th>Product</th>
<th>Apparent ( K_m ) (µM)</th>
<th>Apparent ( V_{max} )</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>rat</td>
<td>liver microsomes</td>
<td>NADPH</td>
<td>total metabolites</td>
<td>0.51</td>
<td>100(^a)</td>
<td>Ahmad et al. (2000)</td>
</tr>
<tr>
<td></td>
<td>kidney microsomes</td>
<td>NADPH</td>
<td>total metabolites</td>
<td>0.48</td>
<td>170(^b)</td>
<td>Ahmad et al. (2000)</td>
</tr>
<tr>
<td></td>
<td>lung microsomes</td>
<td>NADPH</td>
<td>total metabolites</td>
<td>0.50</td>
<td>1.0(^d)</td>
<td>Ahmad et al. (2000)</td>
</tr>
<tr>
<td></td>
<td>brain microsomes</td>
<td>NADPH</td>
<td>total metabolites</td>
<td>1.14</td>
<td>15(^e)</td>
<td>Ahmad et al. (2000)</td>
</tr>
<tr>
<td></td>
<td>liver microsomes</td>
<td>UDP-GlcU</td>
<td>all-trans-retinoyl B-glucuronide</td>
<td>418</td>
<td>0.50(^a)</td>
<td>Littin et al. (1997)</td>
</tr>
<tr>
<td>human</td>
<td>CYP2C0</td>
<td>NADPH</td>
<td>4-hydroxy-ATRA</td>
<td>13.4</td>
<td>4.8(^a)</td>
<td>Thatcher et al. (2010)</td>
</tr>
<tr>
<td></td>
<td>CYP3A4</td>
<td>NADPH</td>
<td>4-hydroxy-ATRA</td>
<td>18.4</td>
<td>4.0(^a)</td>
<td>Thatcher et al. (2010)</td>
</tr>
<tr>
<td></td>
<td>CYP3A5</td>
<td>NADPH</td>
<td>4-hydroxy-ATRA</td>
<td>11.1</td>
<td>4.9(^a)</td>
<td>Thatcher et al. (2010)</td>
</tr>
<tr>
<td></td>
<td>CYP3A7</td>
<td>NADPH</td>
<td>4-hydroxy-ATRA</td>
<td>11.3</td>
<td>2.3(^a)</td>
<td>Thatcher et al. (2010)</td>
</tr>
<tr>
<td></td>
<td>CYP26A1</td>
<td>NADPH</td>
<td>4-hydroxy-ATRA</td>
<td>0.0094</td>
<td>11(^a)</td>
<td>Lutz et al. (2009)</td>
</tr>
<tr>
<td></td>
<td>UGT2B7</td>
<td>UDP-GlcU</td>
<td>all-trans-retinoyl B-glucuronide</td>
<td>1.3</td>
<td>0.52(^a)</td>
<td>Samokyszyw et al. (2000)</td>
</tr>
</tbody>
</table>

\(^a\) pmol · min\(^{-1}\) · mg microsomal protein
\(^b\) pmol · min\(^{-1}\) · pmol CYP enzyme
\(^c\) pmol · min\(^{-1}\) · mg UGT2B7 enzyme

assumed to vary independently. For this simulation, the mean $\mu_w$ and SD $\sigma_w$ 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 $\mu_x$ 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 ($SC = (C' - C) / (P' - P) \times (P / C)$) 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.

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

*n.a.: information not available*
Reverse dosimetry all-trans-retinoic acid
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.

<table>
<thead>
<tr>
<th>Rat strain</th>
<th>Doses applied (mg/kg bw$^{-1}$)</th>
<th>Critical endpoint</th>
<th>Figure</th>
<th>Symbol</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sprague-Dawley</td>
<td>0, 3, 6, 12</td>
<td>malformations head</td>
<td>?A</td>
<td>■</td>
<td>Wise et al. (2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>malformations torso and limb</td>
<td>?A</td>
<td>▲</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>extra vertebra</td>
<td>?A</td>
<td>●</td>
<td></td>
</tr>
<tr>
<td>Wistar</td>
<td>0, 20, 30, 50, 80, 100</td>
<td>cleft palate</td>
<td>?B</td>
<td>■</td>
<td>Tembe et al. (1996)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>microsia</td>
<td>?B</td>
<td>▲</td>
<td></td>
</tr>
<tr>
<td>Wistar (RCRO)</td>
<td>0, 6, 12, 18</td>
<td>external malformations cranium</td>
<td>?C</td>
<td>■</td>
<td>Bürgin and Schmitt (2003)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>visceral malformations palate/mouth</td>
<td>?C</td>
<td>▲</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>skeletal malformations skull</td>
<td>?C</td>
<td>●</td>
<td></td>
</tr>
</tbody>
</table>

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](image)

**Figure 2.** Concentration-response data of ATRA-induced inhibition of ES-D3 cell differentiation based on data from Louise 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).
Chapter 7

Figure 3. Apparent permeability ($P_{\text{app}}$) 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.

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Tissue:buffer partition coefficient</th>
<th>Tissue:blood partition coefficient(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>fat (olive oil)</td>
<td>7770</td>
<td>6.4</td>
</tr>
<tr>
<td>lung</td>
<td>1280</td>
<td>3.9</td>
</tr>
<tr>
<td>brain</td>
<td>4460</td>
<td>14</td>
</tr>
<tr>
<td>heart</td>
<td>640</td>
<td>2.0</td>
</tr>
<tr>
<td>kidney</td>
<td>1810</td>
<td>5.6</td>
</tr>
<tr>
<td>muscle</td>
<td>1120</td>
<td>3.5</td>
</tr>
<tr>
<td>liver</td>
<td>1160</td>
<td>3.6</td>
</tr>
</tbody>
</table>

\(^a\) 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 \textit{in vitro} conversion of ATRA by liver microsomes reported in the literature were used (Table 2). For rats, kinetic constants (apparent \(V_{\text{max}}\) and \(K_m\) values) obtained \textit{in vitro} using liver microsomes were available from the literature (Table 2). The apparent \(V_{\text{max}}\) values of the reactions were scaled to a rat liver assuming a microsomal protein yield of 35 mg/g rat liver (Medinsky \textit{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 \textit{et al.}, 1994) and 4 mg/g rat brain (Ravindranath and Anandatheerthavarada, 1990), respectively.

For humans, kinetic constants obtained \textit{in vitro} using specific biotransformation enzymes were available from the literature (Table 2). The data on apparent \(V_{\text{max}}\) values of the reactions were scaled to a human liver assuming a liver microsomal protein yield of 32 mg/g human liver (Barter \textit{et al.}, 2007). The expression levels of the biotransformation enzymes in human liver microsomes have been reported in the literature for different individuals (Naraharisetti \textit{et al.}, 2010; Lin \textit{et al.}, 2002; Sim \textit{et al.}, 2005; Tay \textit{et al.}, 2010; Zaya \textit{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](imageurl)  
**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.
Reverse dosimetry all-trans-retinoic acid

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 99th 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.
Chapter 7

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
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$_{\text{max}}$) 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$_{\text{max}}$ or the AUC. For the prediction of dose-response curves for ATRA-induced developmental toxicity, the predictions for rat were based on the C$_{\text{max}}$ 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<sub>max</sub> 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<sub>max</sub> (Figure 7).

Based on the evaluation of the predictions for rats, predictions for ATRA-induced developmental toxicity in humans were made by relating the
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 $5^{th}$, the $50^{th}$ and the $95^{th}$ 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.

![Dose-response curves for ATRA-induced developmental toxicity in humans](image)

Figure 8. Predicted in vivo developmental toxicity of ATRA in humans. 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 $5^{th}$ (A), the $50^{th}$ (B) and the $95^{th}$ (C) percentile of this virtual population are shown.

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\textsubscript{10} value for ATRA-induced developmental toxicity in rats amounted to 6 mg \cdot kg bw\textsuperscript{-1}. For humans, BMDL\textsubscript{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\textsubscript{10} values for ATRA-induced developmental toxicity for the 5\textsuperscript{th}, 50\textsuperscript{th} and 95\textsuperscript{th} percentile of this virtual population amounted to 0.02, 0.05 and 0.1 mg \cdot kg bw\textsuperscript{-1}, indicating a higher sensitivity of humans than of rats to ATRA-induced developmental toxicity.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure9.png}
\caption{BMD\textsubscript{10} values in rats for ATRA-induced developmental toxicity. BMD\textsubscript{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\textsubscript{10} values for different ATRA-induced malformations ranging from the BMDL\textsubscript{10} (lower limit of the 95\% confidence interval on the BMD\textsubscript{10}) to the BMDU\textsubscript{10} (upper limit of the 95\% confidence interval on the BMD\textsubscript{10}). The grey area represents the predicted BMD\textsubscript{10} (ranging from the BMDL\textsubscript{10} to the BMDU\textsubscript{10}).

* BMD\textsubscript{10} value could not be determined for this endpoint.}
\end{figure}

\section*{Discussion}

The aim of the present study was to assess the feasibility of predicting \textit{in vivo} developmental toxicity dose levels of the retinoid all-\textit{trans}-retinoic acid (ATRA) with reverse dosimetry, using a PBK model that is based on kinetic parameter values derived using \textit{in vitro} techniques only. To this end, \textit{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 BMD10 value for ATRA-induced developmental toxicity in rats is within one order of magnitude compared with BMD10 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 1<sup>st</sup> and the 99<sup>th</sup> 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 99<sup>th</sup> 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<sub>max</sub>. 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<sub>max</sub> (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 -D<sub>3</sub> 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 5th, the 50th and the 95th) of the dose-response curves for this virtual population reveals the distribution of sensitivity to ATRA-induced developmental toxicity, based on interindividual differences 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 5th 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 5th percentile of the population
Chapter 7

is 2.5-fold lower than the BMDL_{00} for the 50^{th} percentile of the population. This indicates that the default uncertainty factor for intraspecies differences in kinetics (amounting to \sqrt{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 \textit{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 \textit{et al}., 1995). This is probably due to increased expression levels of biotransformation enzymes (Collins \textit{et al}., 1995; Tay \textit{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 \textit{in vitro} concentration-response data as input in the blood compartment of the PBK model, as was done before for the glycol ethers (Louisse \textit{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 \textit{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 \textit{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 \textit{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 BMD10 value for ATRA-induced developmental toxicity and the BMD10 values determined from reported in vivo developmental toxicity studies described in the literature. This difference is small considering the experimental differences in BMD10 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 √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.
Chapter 7

References


Reverse dosimetry all-trans-retinoic acid


Nau, H. (2001). Teratogenicity of isotretinoin revisited: species variation and the role of all-trans-
Chapter 7


Reverse dosimetry all-trans-retinoic acid

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.


CHAPTER 8

Summary, general discussion, future perspectives and conclusions
Chapter 8

Summary

The aim of the present thesis is to provide the proof-of-principle that \textit{in vitro} concentration-response curves can be converted into \textit{in vivo} dose-response curves, using physiologically based kinetic (PBK) modeling with a reverse dosimetry approach, in order to provide alternative methods for \textit{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 \textit{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 \textit{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 \textit{in vitro} concentration-response curves into predicted \textit{in vivo} dose-response curves for developmental toxicity, using PBK modeling with a reverse dosimetry approach. Briefly, these factors include 1) an \textit{in vitro} model to obtain \textit{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 \textit{in vitro} model used was the embryonic stem cell (ES-D3 cell) differentiation assay of the validated embryonic stem cell test (EST) (Genschow \textit{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 \textit{in vitro} developmental toxicity endpoint assessed. The ES-D3 cell differentiation assay is the only validated \textit{in vitro} developmental toxicity assay for which no primary animal tissues are required (Genschow \textit{et al.}, 2002; Genschow \textit{et al.}, 2004). Therefore, alternatives for \textit{in vivo} developmental toxicity studies using this \textit{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
**Chapter 2**

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 (pHᵢ) 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 pHᵢ 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 pHᵢ 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 pHᵢ decrease at concentrations that induce inhibition of ES-D₃ cell differentiation, indicating that the pHᵢ 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 pHᵢ 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 pHᵢ on the mechanism of developmental toxicity of both chemicals. The results presented in chapter 3 indicate that a decrease of the pHᵢ 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-RAl 9-cis-retinoic acid (9-cis-RA), etretinate, acitretin and TTNPB ((E)-4-(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 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 EC50 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 \textit{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 \textit{in vitro} assay for developmental toxicity, 2) developments in the field of PBK modeling, 3) additional factors to be taken into account when translating \textit{in vitro} concentrations into \textit{in vivo} doses and 4) future perspectives and conclusions for the use of \textit{in vitro-in vivo} extrapolations using PBK modeling with a reverse dosimetry approach for risk assessment.

\textit{The choice of an \textit{in vitro} toxicity assay for developmental toxicity}

In the present thesis, the ES-D3 cell differentiation assay of the EST was used as \textit{in vitro} assay to determine \textit{in vitro} toxic effect concentrations for the endpoint developmental toxicity (Genschow \textit{et al.}, 2002; Genschow \textit{et al.}, 2004). The EST is the only validated \textit{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 \textit{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 \textit{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 \textit{et al.}, 2007; Kuegler \textit{et al.}, 2010; Theunissen \textit{et al.}, 2010; Theunissen \textit{et al.}, 2011; Zimmer \textit{et al.}, 2011a; Zimmer \textit{et al.}, 2011b), osteoblast differentiation (Zur Nieden \textit{et al.}, 2010a; Zur Nieden \textit{et al.}, 2010b) or endothelial differentiation (Fesstag \textit{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 \textit{et al.}, 2011). However, with the interpretation of toxicogenomic data, one must be aware that gene expression changes \textit{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 \textit{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 \textit{et al.}, 1985; Toumadje \textit{et al.}, 2003).

To assess whether the three ECVAM-validated \textit{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 \textit{et al.}, 2004; Piersma \textit{et al.}, 2004; Spielmann \textit{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 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

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.
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-D$_3$ 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-D$_3$ cells and ES-D$_3$ 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-D$_3$ 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 EC50 values (± SD) were calculated from values reported in the ECVAM validation studies.

<table>
<thead>
<tr>
<th>Assay</th>
<th>Readout parameter</th>
<th>Reported EC50 in assay</th>
<th>Serum conditions</th>
<th>Calculated free EC50 in assay</th>
<th>Exposure duration</th>
<th>Calculated free EAUC50 in assay</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MAA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ES-D3 cell differentiation assay</td>
<td>inhibition of ES-D3 cell differentiation</td>
<td>2.0 (±0.87) mM</td>
<td>20% FCS</td>
<td>2.0 (±0.87) mM</td>
<td>10 days</td>
<td>480 (±210) mM h</td>
</tr>
<tr>
<td>Rat postimplantation VEC assay</td>
<td>induction of malformations</td>
<td>2.3 (±1.5) mM</td>
<td>100% rat serum</td>
<td>2.3 (±1.5) mM</td>
<td>2 days</td>
<td>110 (±72) mM h</td>
</tr>
<tr>
<td>Limb bud micromass test</td>
<td>inhibition of limb bud cell differentiation</td>
<td>7.9 (±2.3) mM</td>
<td>5% FCS</td>
<td>7.9 (±2.2) mM</td>
<td>5 days</td>
<td>960 (±260) mM h</td>
</tr>
<tr>
<td><strong>ATRA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ES-D3 cell differentiation assay</td>
<td>inhibition of ES-D3 cell differentiation</td>
<td>b. b. (±0.8) nM</td>
<td>20% FCS</td>
<td>0.059 (±0.078) nM</td>
<td>10 days</td>
<td>14 (±18) nM h</td>
</tr>
<tr>
<td>Rat postimplantation VEC assay</td>
<td>induction of malformations</td>
<td>540 (±290) nM</td>
<td>100% rat serum</td>
<td>1.1 (±0.59) nM</td>
<td>2 days</td>
<td>53 (±18) nM h</td>
</tr>
<tr>
<td>Limb bud micromass test</td>
<td>inhibition of limb bud cell differentiation</td>
<td>17 (±9.4) nM</td>
<td>5% FCS</td>
<td>0.58 (±0.32) nM</td>
<td>5 days</td>
<td>70 (±39) nM h</td>
</tr>
</tbody>
</table>

* Averages calculated from reported data of ECVAM validation study (Genschow *et al.*, 2004).
* Averages calculated from reported data of ECVAM validation study (Piersma *et al.*, 2004).
* Averages calculated from reported data of ECVAM validation study (Spielmann *et al.*, 2004).
* Calculated by correcting the nominal EC50 values for the free fraction (Figure 1).
* Calculated by multiplying the free EC50 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\text{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)?

---

**Summary, general discussion, future perspectives and conclusions**

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\text{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 $C_{\text{max}}$, 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 $C_{\text{max}}$ 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 $C_{\text{max}}$ (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 $C_{\text{max}}$ (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.
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 $C_{\text{max}}$ 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 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.](image-url)
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 \( C_{\text{max}} \), 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 \( C_{\text{max}} \) 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
Summary, general discussion, future perspectives and conclusions

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](image)

**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 $C_{\text{max}}$ 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).
Chapter 8

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


Chapter 8


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.


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.


Summary, general discussion, future perspectives and conclusions


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 ontwikkelingstoxicologie 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 ontwikkelingstoxiciteitsstudies zou daarom een significante bijdrage leveren aan de vervanging, vermindering en verfijning (3Vs) 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 ontwikkelingstoxiciteit. 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-anpak, 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 kinetic 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 ontwikkelingstoxiciteit data 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 verschaf. Twee series modelstoffen werden geselecteerd, welke zijn 1) een groep glycol ethers, die gebioactiveerd moeten worden om ontwikkelingstoxiciteit te veroorzaken, die hydroofiel zijn, een lage mate van eiwitbinding vertonen en een niet-specifiek werkingsmechanisme hebben, en 2) een groep retinoiden, 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 conceptueel 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 celllijnen, en 2) hele-embryo-kweek (WEC)-testen. Er wordt een overzicht gegeven van de verschillende 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 alkoxyazijnzuurmetabolieten 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 acetzalamide (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, in het ontwikkelingstoxiciteitsmechanisme van beide chemicaliën bevestigt. De resultaten die gepresenteerd zijn in hoofdstuk 3 geven aan dat een afname van de pH, waarschijnlijk de wijze is waarop de alkoxyazijnzuurmetaboliënten van de glycolethers ontwikkelingstoxiciteit veroorzaken, en bevestigen de rol van de alkoxyazijnzuurmetaboliënten 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 azijnzuurmetaboliënten in plaats van met de moeder-glycolethers, in ogenschouw nemende dat latere PBK-modellering (hoofdstuk 5) rekening zou houden met de bioactivatie van de glycolethers tot hun alkoxyazijnzuurmetaboliënten.

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 alkoxyazijnzuurmetaboliënten 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 metaboliënten de ontwikkelingstoxische stoffen in vivo zijn. De potenties van de alkoxyazijnzuurmetaboliënten 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.
Samenvatting, toekomstperspectieven en conclusies

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 verschillende 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 nanomolare tot micromolare 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-signaltransductie-pathways (Collins en Mao, 1999) dan het waarschijnlijk niet-specifieke ontwikkelings-toxiciteitsmechanisme van de glycolethers (inductie van pH-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 vergelijk 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 verschillende 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 modelparameetwaarden voor het PBK-model van ATRA. Deze studie verschaf 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 gepparameteriseerd 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 gebruikt omgekeerde dosimetry 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\textsubscript{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\textsubscript{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
gprioriteerd voor verdere risicobehoordeling.

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

Samenvatting, toekomstperspectieven en conclusies

243
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
Samenvatting, toekomstperspectieven en conclusies


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.


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, altij d 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
Dankwoord / Acknowledgements


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, Suleyman, 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 paranymphen 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!

This, 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!
ABOUT THE AUTHOR
Curriculum Vitae


About the author
List of publications


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
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

Jochem Louisse, 2012