Development of a generic zebrafish embryo PBPK model and application to the developmental toxicity assessment of valproic acid analogs

Ségolène Siméon[†], Katharina Brotzmann[‡], Ciaran Fisher[§], Iain Gardner[§], Steve Silvester[‖,1], Richard Maclennan[‖,2], Paul Walker[‖], Thomas Braunbeck[‡] and Frederic Y. Bois*[,§].

[†] INERIS, METO unit, Parc ALATA BP2, Verneuil en Halatte, France.

[‡] University of Heidelberg, Aquatic Ecology and Toxicology, Centre for Organismal Studies (COS), Im Neuenheimer Feld 504, D-69120 Heidelberg, Germany.

[§] CERTARA UK Limited, Simcyp Division, Level 2-Acero, 1 Concourse Way, Sheffield, S1 2BJ, United Kingdom.

[‖] Cyprotex Discovery Ltd., No. 24 Mereside, Alderley Park, Macclesfield, Cheshire, SK10 4TG, United Kingdom.

[1] Steve Silvester is currently employed at Alderley Analytical Ltd. Alderley Park, Macclesfield, Cheshire, SK10 4TG, United Kingdom.
[2] Richard Maclennan is currently employed at Appleyard Lees IP LLP, The Lexicon Mount Street, Manchester, Greater Manchester, M2 5NT, United Kingdom.

Highlights

- A zebrafish embryo PBPK model for neutral or ionizable chemicals is proposed.
- Organs' growth and changes in metabolic clearance with time are considered.
- Bayesian calibration can be used to improve predictions.
- We apply the model to valproic acid analogues' developmental toxicity.

Abstract

In order to better explain, predict, or extrapolate to humans the developmental toxicity effects of chemicals to zebrafish (*Danio rerio*) embryos, we developed a physiologically-based pharmacokinetic (PBPK) model designed to predict organ concentrations of neutral or ionizable chemicals, up to 120 hours post-fertilization. Chemicals' distribution is modeled in the cells, lysosomes, and mitochondria of ten organs of the embryo. The model's partition coefficients are calculated with sub-models using physicochemical properties of the chemicals of interest. The model accounts for organ growth and changes in metabolic clearance with time. We compared *ab initio* model predictions to data obtained on culture medium and embryo concentrations of valproic acid (VPA) and nine analogs during continuous dosing under the OECD test guideline 236. We further improved the predictions by estimating metabolic clearance and partition coefficients from the data by Bayesian calibration. We also assessed the performance of the model at reproducing data published by Brox *et al*. (2016) on VPA and 16 other chemicals. We finally compared dose-response relationships calculated for mortality and malformations on the basis of predicted whole embryo concentrations *versus* those based on nominal water concentrations. The use of target organ concentrations substantially shifted the magnitude of dose-response parameters and the relative toxicity ranking of chemicals studied.

Keywords

PBPK model, *Danio rerio*, Zebrafish embryo, Bayesian, Development, Effect concentration, Valproic acid, Internal concentration, Toxicity.

1 Introduction

Prediction of chemicals' developmental and reproductive toxicity is a complex challenge. Toxicity assays are in majority conducted in mammals, due to their recognized efficacy at predicting toxicity in humans. However, mammalian assays are expensive, strictly regulated by law, and time-consuming. Given the morphological and developmental similarities among vertebrates, fish are a relevant test alternative to mammals [1,2]. For various reasons, the zebrafish embryo is particularly attractive in toxicology and pharmacology [3]. First, its transparency allows the visual detection of malformations, without interrupting development or invasive interventions [4]. Second, the number of eggs laid is high, and its development time is short [5]. Third, it is easy to maintain in the laboratory [6]. Finally, there are considerable gene homologies and neurophysiological similarities between zebrafish, mammals and humans [7,8]. Therefore, evidence obtained with zebrafish embryos should at least partially be translatable to humans [9]. From a regulatory point of view, until the age of 120 h, zebrafish embryos are an alternative to experiments with adult vertebrate species, because they are not protected by European animal welfare legislation until five days post-fertilization (hpf) [10,11].

Yet, extrapolation of toxicity from zebrafish to humans requires, at least, accounting for differences in pharmacokinetics between the two species [12]. Such differences might translate into differences in target organ concentrations for the same systemic exposure dose. In addition, knowing chemical concentrations in organs is fundamental to understand dose-response relationships [13,14]. Internal concentrations are often difficult to measure, but physiologically-based pharmacokinetic (PBPK) models can estimate them [15,16]. PBPK models connect anatomy, physiology, and biochemical processes to understand and compute a chemical's fate in the body. They allow approximate predictions of chemicals' concentration-time profiles in experimentally inaccessible organs from minimal data. Thereby, they provide mechanistic insight into toxicity and help reduce time, cost and need for animal experiments [17].

PBPK models have been published for adult zebrafish, mostly for ecotoxicological risk assessment [18–20]. Recently, Brox *et al.* [21] used a one-compartment two-parameter model developed by Gobas and Zhang [22] to explore the impact of physicochemical properties of polar compounds and of biological processes on embryo concentrations. This model, however, cannot describe decreases in concentrations due to metabolism [23] or dilution by organ growth, and Brox *et al.* concluded the necessity of more sophisticated toxicokinetic models for the zebrafish embryo.

In order to better explain, predict, and extrapolate developmental toxicity observed in zebrafish embryos, we developed a generic PBPK model integrating organ growth and hepatic metabolism. The model assumes quasi steady-state distribution between zebrafish cells, lysosomes, and mitochondria in different tissues (yolk, liver, gut, muscle, skeleton, eye, brain, heart, skin, and lumped other tissues). The model is generic in that it can simulate the distribution of many chemicals in zebrafish embryos on the basis of their physicochemical properties: chemicals' partition coefficients between cells or organelles and culture medium are calculated with the Simcyp® virtual *in vitro* intracellular distribution (VIVD) model [24]. The model can therefore be used for high-throughput predictions of internal concentrations in zebrafish.

We applied our model to the analysis of developmental toxicity data on valproic acid (VPA) and nine of its analogs: 2,2-dimethylvaleric acid, 2-ethylbutyric acid, 2-ethylhexanoic acid, 2-methylhexanoic acid, 2-methylpentanoic acid, 2-propylheptanoic acid, 4-eneVPA, 4-pentenoic acid and hexanoic acid. VPA is a notorious teratogenic antiepileptic and thymoregulator, inducing neural tube defects in mammalian embryos, probably by inhibition of histone deacetylase, interference with folate metabolism and inducing oxidative stress. However, the mechanism of action remains not well known [25–27]. VPA exerts its pharmacological effects mainly in the central nervous system by inhibition of the citric acid cycle and elevation of γ-aminobutyric acid (GABA) level [26]. Chemicals with similar structure can have similar properties, but for toxicological properties this should be backed-up by *in silico* and *in vitro* evidence. We demonstrate how the model can be used to base the toxicity ranking of VPA and the above analogs on internal concentration estimates rather than on nominal water medium concentrations, thereby helping transferability of zebrafish embryo test results to human risk assessment. In addition,

we use the data published by Brox *et al*. [21] on 16 other chemicals to discuss the model performance for a larger class of chemicals.

2 Materials and methods

2.1 EU-ToxRisk zebrafish experiments

2.1.1 Test chemicals

Except for 4-eneVPA (Santa Cruz Biotechnology, Dallas, Texas, USA; 98 % purity), all other chemicals (valproic acid, 2,2-dimethylvaleric acid, 2-ethylbutyric acid, 2-ethylhexanoic acid, 2-methylhexanoic acid, 2-methylpentanoic acid, 2-propylheptanoic acid, 4-pentenoic acid and hexanoic acid) were purchased at the highest purity available from Sigma (Deisenhofen, Germany). After initial range-finding tests, the ten chemicals were tested at three to eight different concentrations prepared from a 100 % DMSO stock solutions (n = 3). Nominal and analytically measured water concentrations as well as measured total embryo concentrations are summarized in Table S1 of the Supporting Information. The highest concentration tested led sometimes to 100% mortality of the embryos. In such cases, no concentration measurement was made.

2.1.2 Fish embryo toxicity testing

Adult wild-type zebrafish (*Danio rerio*) of the 'Westaquarium strain' were kept at the fish facilities of the Aquatic Ecology and Toxicology Group at the University of Heidelberg (licensed under no. 35-9185.64/BH). Based on OECD test guideline 236 as well as on complementary published work [28–31], embryos were raised and exposed until an age of 120 hours post-fertilization (hpf). According to the current EU animal welfare legislation, exposure of zebrafish may be extended to 120 hpf in cases of inconclusive observations until 96 hpf [11].

For initiation of the tests, embryos were immersed in the test solutions at the 16 cell-stage at the latest (≤ 90 min; before cleavage of blastodisc). To start exposures with minimum delay, twice the number of eggs eventually needed per treatment group were picked from the same batch of eggs and transferred into 100 ml crystallization dishes with the test concentrations or negative (artificial water according to

ISO 7346-3) and positive controls (24.7 µM of 3,4-dichloroaniline) [33] . At 3 hpf at the latest, viable eggs were selected for normal development under the stereomicroscope (≥ 30-fold magnification) and transferred to a final volume of 1 ml into 24 well plates (one embryo per well), which had been pre-exposed to the test solutions for 24 h to account for potential adsorption of the test solutions to the plastic walls of the wells. Test solutions were replaced at 24, 48, 72 and 96 hpf (without changing the well plates). Embryos were not dechorionated and hatched on their own at approximately 72 hpf.

Prior to replacement of the test solutions as well as at completion of the test (120 h), embryos were analyzed for macroscopically discernable alterations including the four morphological core endpoints listed by OECD TG 236 (coagulation, non-detachment of the tail, non-formation of somites and lack of heartbeat) [33] as well as any additional sublethal observation such as scoliosis/lordosis, eye deformation, loss of pigmentation, various types of edemata and general skeletal deformations [6,34,35].

For documentation, morphological alterations were recorded with a Zeiss Axio Cam ICc1 camera mounted on a Zeiss Olympus CKX41 microscope (Carl Zeiss, Oberkochen, Germany) and analyzed using the Zeiss imaging program Zen lite 2011. After termination of the exposure, embryos were anesthetized in 2 ml Eppendorf cups immersed into crushed ice for 30 min, washed three times with artificial water to remove superficial chemicals from the embryo bodies and shock-frozen with liquid nitrogen for subsequent chemical analysis of internal doses.

2.1.3 Toxicity data analysis

Effective concentrations (EC) leading to 10, 20 and 50% of mortality or malformations (grouped together as “total effect”) were calculated using ToxRat Prof. Vers. 2.10 (ToxRat Solutions, Alsdorf, Germany).

2.1.4 Chemical analysis of actual chemicals´ concentration

Concentrations in the embryo and in the culture medium at 72 or 120 hpf for the ten VPA analogs studied were measured by liquid chromatography and mass spectrometry. 100 µL ultrapure water were added to each vial containing 2 to 10 zebrafish embryos and mixed thoroughly. To 100 µL of embryo, incubated water of standard curve sample, 10 µL of rosuvastatin (internal standard, IS, Sigma-Aldrich)

in 50% methanol (VWR) were added, 300 µL or acetonitrile (Fisher Scientific) were then added to each sample. The samples in tube were centrifuged at 3000 rpm for 10 minutes, those in plates were centrifuged at 13,000 rpm for 5 minutes. 300 µL of supernatant were transferred to a fresh plate and subjected to a dry down under a stream of nitrogen at 50^{o} C for approximately 30 minutes (until approximately 50 µL sample remained). To each well or sample, 100 µL ultrapure water were added, the plate was sealed and mixed thoroughly. The sample plates were then placed into the autosampler attached to a Sciex TripleTOF 6600 Quadrupole Time-Of-Flight (QTOF) mass analyzer (AB Sciex, Singapore). The conditions used when running the samples are shown in Table S2 of the Supporting Information.

2.2 Brox *et al*. zebrafish experiments

The 17 chemicals studied by Brox *et al*. [21] were 2,4-diclorophenoxyacetic acid, atropine, benzocaine, caffeine, chloramphenicol, cimetidine, clofibric acid, colchicine, cyclophosphamide, metoprolol, metribuzin, phenacetin, phenytoin, sulfamethoxazole, theophylline, thiacloprid and valproic acid. For each compound, Brox *et al*. performed two independent experiments, except for three compounds for which a third one was done to reduce experimental uncertainty. The experiments were performed on zebrafish embryos from 4 hpf (four-cell stage) to 96 hpf, exposed to external concentrations comprised between 10 to 250 mg/L, without medium change. The approximate stability of external concentrations during the experiments was checked. The quantity of chemical in each embryo was measured at 24, 48, 72 and 96 hpf. For metribuzin, phenacetin and benzocaine, measurements at 6 hpf were also performed. Three vials of 20 mL containing nine dechorionated embryos in 18 mL exposure solution, and three control vials, were used for each sampling time. Replicates were processed separately. The experiment was performed in a closed system at 26 ± 1 °C.

2.3 PBPK model structure

Figure 1 presents the schematic structure of our zebrafish embryo model. Ten compartments are considered because of interest of compound effect on their development: yolk, liver, skeleton, gut, eye, brain, heart, skin, muscles, other organs and tissues. Mitochondria and lysosomes in tissues (except the

yolk) form two additional compartments. These organelles have a specific critical pH of 4.5 for lysosomes and 8 for mitochondria, leading to potential 'ion trapping' phenomena (sequestration of compounds because of their differential ionization between organelles and cellular embryo). The model also considers air and plastic to medium partition. Protein binding and ionization in the medium are also accounted for. In the experiments reported in this article the culture medium contained no proteins; therefore, protein binding was turned off in the model. Instantaneous diffusion across the various compartments is assumed. Metabolic clearance is modeled as a dynamic process in the liver. Organ volume growth with time is also modeled. Therefore, chemical's concentrations and quantities in the model change with time, even though (for notation simplicity) time indexing is not always explicit in the following equations.

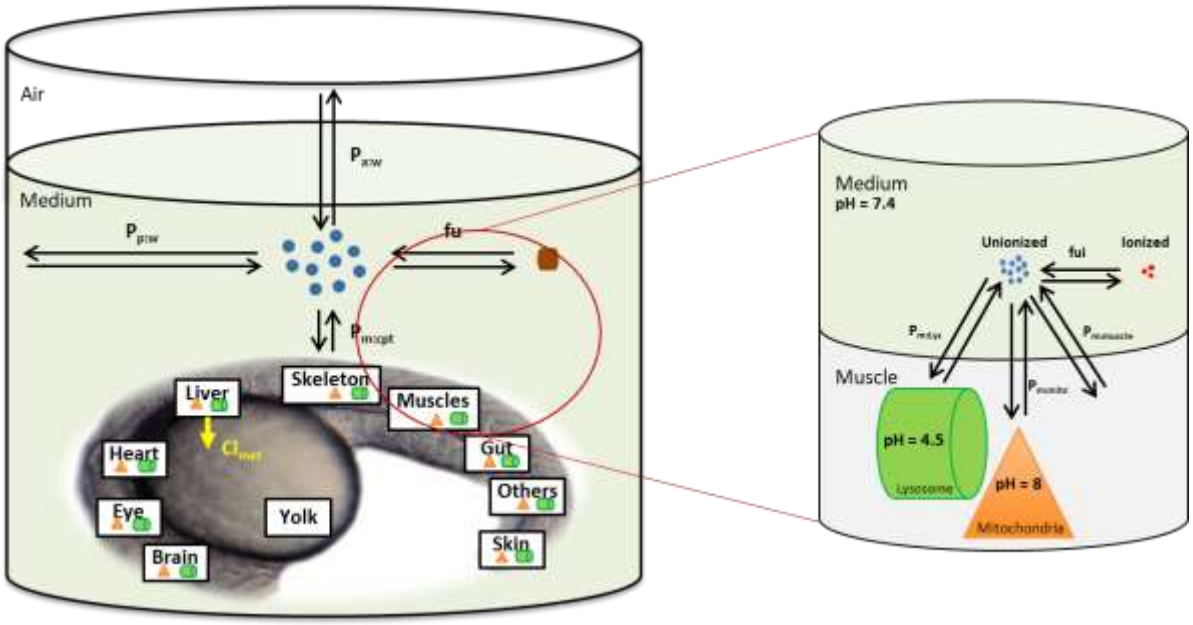

Figure 1: Structure of the zebrafish embryo model. The chemical of interest partitions between the various compartments and can be metabolized in the liver. It can also partition to the air and bind to the plastic walls and culture medium proteins. Legend: Cl_{met}: Metabolic clearance; $P_{p:w}$: Plastic to water partition coefficient; $P_{a:w}$: Air to water partition coefficient; $P_{m:cpt}$: Compartment to medium partition coefficient; *fu*: Fraction unbound in medium; *fui*: Fraction unionized in medium.

Concentrations in organs, yolk, lysosomes and mitochondria (C_i), are assumed to be at any time proportional to the concentration unbound in medium ($C_{medium,u}$). The proportionality factors are the

medium unbound over organs, yolk, lysosomes or mitochondria partition coefficients ($P_{mu:i}$) eventually corrected by a scaling factor (f_{pc}):

$$C_i = \frac{C_{medium,u}}{f_{pc} \times P_{mu:i}} \quad (1)$$

$C_{medium,u}$ is computed according to:

$$C_{medium,u} = \frac{Q_{parent}}{\frac{V_{medium}}{fu} + P_{a:w} \times fu_u \times V_{air}(t) + P_{p:w} \times S_{medium}(t) + \sum(P_{mu:j} \times V_j(t))} \quad (2)$$

where V_{medium} is the volume of culture medium, *fu* the fraction unbound in medium, *fui* the fraction unionized in medium, $P_{a:w}$ and $P_{p:w}$ are respectively the air and plastic to water partition coefficient, $V_{air}(t)$ the volume of air in head space at time *t*, $S_{medium}(t)$ the surface area of medium in contact with plastic. $V_j(t)$ are the volumes of yolk, liver, gut, muscle, skeleton, eye, brain, heart, skin, and other tissues at time *t*. The total concentration in medium (C_{medium}) is:

$$C_{medium} = \frac{C_{medium,u}}{fu} \quad (3)$$

Since we model the developing embryo, it is necessary to consider organ growth. The embryo volume without yolk at time *t*, $V_{embryo}(t)$, is computed as:

$$V_{embryo}(t) = V_{embryo}(120) \times \sum sc_k(t) + V_{embryo}(0) \quad (4)$$

where $V_{embryo}(120)$ is the volume of embryo at 120 hpf, $sc_k(t)$ represents the fraction of $V_{embryo\ (120hpf)}$ taken up by organ *k* (liver, gut, muscle, skeleton, eye, brain, heart, skin, or other tissues) at time *t*, and $V_{embryo}(0)$ is the embryo volume at the start of the experiment.

$V_{embryo}(0)$ is computed by assuming that the fertilized egg is a half-sphere of radius $r_{embryo,0}$ equal to 0.13 mm [36]:

$$V_{embryo}(0) = \frac{2}{3} \times \pi r_{embryo,0}^3 \quad (5)$$

The yolk volume at time *t*, $V_{yolk}(t)$, is given by:

$$V_{yolk}(t) = V_{yolk}(0) \times \exp(-K_{d,yolk} \times t) \quad (6)$$

where $K_{d,yolk}$ is the yolk consumption rate constant, and $V_{yolk}(0)$ is the yolk volume at start of experiment, computed according to the hypothesis that it is a sphere of radius $r_{yolk,0}$ equal to 0.4 mm [36].

$$V_{yolk}(0) = \frac{4}{3} \times \pi r_{yolk,0}^3 \quad (7)$$

The total volume of the embryo, $V_{embryototal}(t)$, is the sum of $V_{yolk}(t)$ and $V_{embryo}(t)$.

$$V_{embryo_{total}}(t) = V_{yolk}(t) + V_{embryo}(t) \quad (8)$$

The organ volumes at time t, $V_k(t)$, for liver, gut, muscle, skeleton, eye, brain, heart, skin, and other tissues are computed as:

$$V_k(t) = V_{embryo}(120) \times sc_k(t) \quad (9)$$

The fractional volumes $sc_k(t)$ were computed for each organ according to the time-dependent equations:

$$sc_k(t) = \exp\big(K_{g,k} \times (t - \tau_k)\big) - 1 \quad (10)$$

where τ_k is the time of growth initiation for organ k. Before τ_k, organ k volume is null. The organ growth rates ($K_{g,k}$) were calibrated used published information and our own data on embryos (see next section).

The surface area of medium in contact with plastic, S_{medium}(t), is:

$$S_{medium}(t) = 4 \times \frac{V_{content}(t)}{D_{well}} + \pi \left(\frac{D_{well}}{2}\right)^2 \quad (11)$$

where D_{well} is the well diameter, and $V_{content}(t)$ is the total volume of content per well, computed as:

$$V_{content}(t) = V_{medium} + V_{embryo_{total}}(t) \quad (12)$$

$V_{air}(t)$ is computed as the difference between well volume and medium, embryo and yolk volumes:

$$V_{air}(t) = V_{well} - V_{content}(t) \quad (13)$$

The air concentration depends on the air to water partition coefficient, eventually corrected by the scaling factor (f_{pc}), and on the fraction unbound and unionized in medium:

$$C_{air} = f_{pc} \times P_{a:w} \times C_{medium,u} \times fu_u \quad (14)$$

Similarly, the quantity of chemical bound to the culture walls per unit surface area depends on the plastic to water partition coefficient (which has the dimension of a length):

$$C_{plastic} = P_{p:w} \times C_{medium,u} \quad (15)$$

If metabolism is assumed to be linear in the embryo, the total quantity of metabolites formed per unit time in system is proportional to the number of liver cells of the embryo (N_{cell}), metabolic clearance per liver cell (Cl_{met}) and parent chemical concentration in liver cells (C_{liver}):

$$\frac{Q_{met}}{dt} = N_{cells} \times Cl_{met} \times C_{liver} \quad (16)$$

If metabolism is assumed to be saturable, a Michaelis-Menten term is used and Eq. 16 becomes:

$$\frac{Q_{met}}{dt} = N_{cells} \times V_{max} \times C_{liver}/(K_m + C_{liver}) \quad (17)$$

where V_{max} is the maximum rate of metabolism and K_m the Michaelis-Menten constant.

The total quantity of parent molecules in the system (Q_{parent}) depends of quantities in medium, air, total embryo and plastic. The model can account for the saturation of plastic binding by pre-incubation with the test substance prior to embryo exposure. The model splits Q_{parent} into two quantities, both function of time. A labile quantity present in medium and air (Q_{labile}), reset at 0 at each medium change; and a fixed quantity in embryo and bound on plastic walls (Q_{fixed}), impervious to medium renewals.

$$Q_{parent} = Q_{labile} + Q_{fixed} \quad (18)$$

$$\frac{Q_{labile}}{dt} = -f_{labile} \times \frac{Q_{met}}{dt} \quad (19)$$

$$\frac{Q_{fixed}}{dt} = -(1 - f_{labile}) \times \frac{Q_{met}}{dt} \quad (20)$$

where f_{labile} is the fraction of Q_{parent} in medium and air:

$$f_{labile} = \frac{\frac{C_{medium,u}}{fu} \times V_{medium} + C_{air} \times V_{air}(t)}{Q_{parent}} \quad (21)$$

N_{cell} changes with time and is calculated from liver volume at time t, $V_{liver}(t)$, and hepatocyte volume (V_{hep}):

$$N_{cells} = \frac{V_{liver}(t)}{V_{hep}} \quad (22)$$

The quantity in medium (Q_{medium}) depends on concentration unbound in medium, adjusted by the fraction unbound:

$$Q_{medium} = \frac{C_{medium,u} \times V_{medium}}{fu_u} \quad (23)$$

The quantities in organs, yolk and air (Q_i) are computed as:

$$Q_i = C_i \times V_i(t) \quad (24)$$

The quantity in embryo, excepting yolk, Q_{embryo}, is the sum of organ quantities, and the concentration in embryo (C_{embryo}) is computed as:

$$Q_{embryo} = \sum Q_{organ} \quad (25)$$

$$C_{embryo} = \frac{Q_{embryo}}{V_{embryo}} \quad (26)$$

The total quantity of compound in the embryo, $Q_{embryo,total}$, is the sum of $Q_{yolk}(t)$ and Q_{embryo}:

$$Q_{embryo_{total}} = Q_{yolk} + Q_{embryo} \quad (27)$$

The quantity bound to plastic ($Q_{plastic}$) is given by:

$$Q_{plastic} = C_{plastic} \times S_{medium} \quad (28)$$

The quantity in lysosomes (Q_{lyso}) and mitochondria (Q_{mito}) depend on the volume of the embryo and the fractions of lysosome (f_{lyso}) and mitochondria (f_{mito}) in cells, respectively:

$$Q_{lyso} = C_{lyso} \times V_{embryo}(t) \times f_{lyso} \quad (29)$$

$$Q_{mito} = C_{mito} \times V_{embryo}(t) \times f_{mito} \quad (30)$$

The length of the embryo was calibrated by ourselves with data from Kimmel *et al.* [36], using the following empirical equation:

$$L_{embryo} = A \times \frac{t^B}{C^B + t^B} + D \quad (31)$$

where A = 0.0260 dm, B = 4.397, C = 1617 minutes and D = 0.00755 dm.

At the start of a simulation, time zero, the quantity of parent compound (Q_{parent}) in the system is known. It includes was has been introduced in the medium plus what is already attached to the plastic walls in case of pre-incubation (Eq. 15 applies then). Therefore, the faction f_{labile} can be computed (Eq. 21) and used to set the initial values for the state variable Q_{labile} and Q_{fixed}, as corresponding fractions of Q_{parent}.

Medium renewals introduce a discontinuity in the solution of the two differential equations of the model (for Q_{labile} and Q_{fixed}). That case is handled classically: Integration is stopped just at the renewal time; Q_{fixed} is kept at its current value and Q_{labile} is made equal to the new quantity introduced in the system; Simulation is restarted from the renewal time.

2.4 PBPK model parameters

2.4.1 Partition coefficients

To estimate the fraction of chemical unbound in medium (*fu*), the fraction unionized in medium (*fui*), the plastic to water partition coefficient ($P_{p:w}$), the air to water partition coefficient ($P_{a:w}$), and the organs (liver, gut, muscle, skeleton, eye, brain, heart, skin, others, yolk, lysosomes and mitochondria) to medium unbound partition coefficient ($P_{mu:i}$), we use the Simcyp® VIVD model [24]. It computes parameter values on the basis of physicochemical properties of the substance considered (*logP*; Henry's constant; *pKa*; compound's character: mono or dibasic, mono or diacid, neutral, ampholyte; molecular weight; blood to plasma ratio and fraction unbound in bovine serum, *pH* and membrane potential) and embryo organ properties, obtained from the literature on embryos (for the yolk) [37,38] or adult fish [18]. As there was no bovine serum or other proteins in the zebrafish culture medium, *fu* was set equal to 1 for all compounds.

2.4.2 Physiological parameters

The model's physiological parameters are given in Table S3. Organ growth rates were estimated from our own data on total embryo volume and volume without yolk at different times (see Figure S1 in Supporting Information) using the following procedure: The yolk consumption rate constant was estimated using a simple exponential decay equation. The embryo's volume at 120 hpf ($V_{embryo\ (120hpf)}$),

the fractional volume of muscle at 120 hpf ($sc_{muscle\ (120hpf)}$) and the "other organs" fractional volume at 120 hpf ($sc_{others\ (120hpf)}$) were estimated by fitting the organ growth part of the model to the data. Calibration was performed by MCMC simulations in a Bayesian statistical framework [39–41]. The data was assumed to be log-normally distributed around the model predictions with a geometric standard deviation σ (estimate of residual uncertainty). Non-informative uniform priors were used for the three parameters to calibrate, so as to "let the data speak". Two MCMC chains of 10000 iterations were simulated and one of every two random samples produced were recorded. Convergence of the two chains was assessed using Gelman and Rubin's Rhat convergence criterion [42].

The fractional volumes, $sc_{k\ (120hpf)}$, for the rest of the organs were then computed by rescaling their literature values [18,43]:

$$sc_{k\ (120hpf)} = sc_{k\ litterature\ (120hpf)} \times \frac{(1-sc_{muscle\ fitted\ (120hpf)}-sc_{others\ fitted\ (120hpf)})}{(1-sc_{muscle\ litterature\ (120hpf)}-sc_{others\ litterature\ (120hpf)})} \quad (32)$$

Organ growth rates were finally obtained by:

$$K_{g,k} = \frac{\ln(sc_{k\ (120hpf)}+1)}{(t_{final}-\tau_k)} \quad (33)$$

t_{final} being equal to 120 hours.

2.4.3 Estimation of metabolic clearance and partition coefficient scaling factor

To improve the model fit to the data beyond that obtained with *ab initio* predictions, metabolic clearance (Cl_{met}) and the scaling factor (f_{pc}) were calibrated for individual chemicals on the basis of chemical concentration data in the embryo and medium. Fits were performed with data obtained at 120 hpf (and at 72 hpf for VPA). Those parameters were calibrated jointly, using MCMC simulations. The data on C_{medium} and C_{embryo} was assumed to be log-normally distributed around the model predictions (taken as geometric mean) with geometric standard deviation σ. The two non-detectable concentration data points were excluded from the analysis. The SD σ was also calibrated by sampling and assumed to be *a priori* distributed normally around 1.5 ± 1.5 SD with a truncation from 1.5 to 10 (that is, between 50% error and a 10-fold error at most). Two Markov chains of 10000 iterations were simulated for each chemical, and one in two random parameter samples were recorded. The last half of each recorded set of samples

was kept and convergence of the two chains was assessed using Gelman and Rubin's convergence criterion.[42] The prior distribution of metabolic clearance was assumed to be uniform (*i.e.*, uninformative) from 0 to either 10^{-11}, 10^{-10}, or 10^{-9} L/min, depending on compound (see Table 1). Different upper bounds were used to speedup convergence of the chains toward the posterior distribution, but they do not affect the posterior estimates as they were not reached during sampling at convergence. The prior distribution of f_{pc} was assumed to be uniform from 0 to 5 (*i.e.*, uninformative).

2.5 Software

The static model equations of the VIVD model were coded in R version 3.4.3 [44]. The corresponding R script was used as a preprocessor to obtain chemical-specific parameter values for input to the embryo model. All the dynamic model simulations and MCMC calibrations were performed with GNU MCSIM version 5.6.6 (https://www.gnu.org/software/mcsim/) [45]. The model code is given as supplemental material and on the web at https://sites.googles.com/site/modelecotoxtox/Software.

3 Results and discussion

3.1 Physiological parameters' calibration

The model accounts for the embryo's organ growth over time. We estimated the yolk consumption rate constant, the embryo's volume at 120 hpf, the fractional volume of muscle at 120 hpf and the "other organs" fractional volume at 120 hpf on the basis of our experimental embryo volume data. Figure S1 (in Supporting Information) shows the observed and predicted organ growth over time. Predicted total embryo volume and embryo volume without yolk fit the data rather well: the median relative error for the embryo volume estimate without yolk is equal to 1.09, and for the total embryo volume it is equal to 0.94. The Figure S2 also shows the contribution of each organ to the total embryo volume. The estimated parameter values are given in Table S3 of the Supporting Information.

3.2 *Ab initio* predictions of embryo concentrations

The parameter values obtained with the VIVD model for VPA and its analogs are given in Table S4 (Supporting Information). The primary physicochemical properties input to the VIVD model for

calculation of those values are given in Supplemental Table S5 (Supporting Information). 2-propyl-heptanoic acid had the highest octanol over water partition coefficient of all the analogs studied here. All analogs were predicted to be mainly present in ionized form in the medium. The partition coefficient for plastic binding is difficult to interpret directly because of its dimension (a length). Yet, it is useful to assess the impact of the materials used on the kinetics of the test chemicals. This is best done by comparing the predictions of the quantities bound to plastic and present in water. In our case, for all substances, the quantity bound to plastic was predicted to be about 0.2% of what is in water and therefore negligible.

The compounds were predicted to partition preferentially to the yolk, except for VPA and 2-propylheptanoic acid, which partition to water rather than yolk. They also all have higher affinity for the other embryo tissues than for medium. Affinity for lysosomes was predicted to be ten to a hundred times higher than for the other compartments for all compounds, except again for 2-propylheptanoic acid.

Figure S2 (Supporting Information) shows the difference of organ concentrations as a function of time for VPA. We can see that VPA concentrates preferably in mitochondria and weakly in lysosomes.

We compared the observed analogs' concentrations to those obtained when using VIVD-predicted parameters. For that case, metabolic clearance (Cl_{met}) was set to 0 L/min, because we did not have a way to predict it) and f_{pc} to 1. Figure 2 shows that those *ab initio* predictions tend to be underestimates of embryo concentrations (median relative error 0.11). The medium concentration data were well predicted (median relative error 0.90).

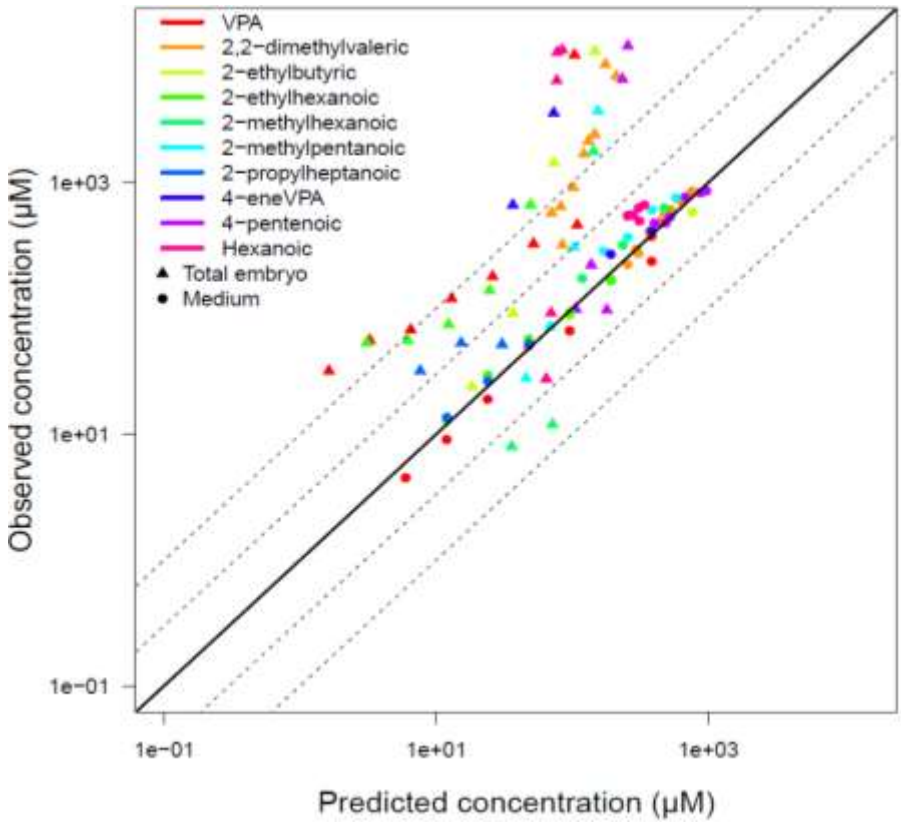

Figure 2: Observed *versus* model-predicted concentrations in the zebrafish embryo and in culture medium for valproic acid and nine analogs. The VIVD-computed parameter values were used without further adjustment. Metabolic clerance was set to zero and the correction factor of partition coefficient to one. The black line corresponds to perfect predictions. Dashed lines delineate the three- and ten-fold error bounds.

The VIVD model estimates of partition coefficients are certainly not perfect, for example due to imprecision affecting some input parameters, such as Henry's law constant. For low volatility compounds, the assumption of instantaneous partitioning between culture medium and air in the head-space may also over-predict distribution to the head-space. Conversely, particularly for high volatility compounds, the model will likely under-predict loss to the head-space if the experimental system is not hermetically sealed (as assumed by the model). Modelling distribution into the head-space air as a dynamic process, as we did for metabolism, might improve predictions [46]. However, that would complicate the model and add many parameters. The physicochemical tissue properties values (except for the yolk) were obtained in adult fish because we did not have embryo fish specific data: This should be improved with specific measurements. It should also be noted that the VIVD model has been

developed for relatively well-behaved small molecules and does not have a universal domain of applicability. It does not predict metabolic clearance either. Actually, there are no published QSAR or other *in silico* methods to predict metabolic clearances in zebrafish embryo. QSAR methods have been developed to predict the biotransformation half-lives or rates in adult fish [47,48]. However, since there are important physiological differences between adult fish and embryos, embryos fall out of the applicability domain of these QSARs. It would be interesting to understand whether these QSARs nevertheless provide useful upper bounds or lower bounds on the embryo biotransformation parameters.

Note also that the VIVD model is quite general and considers a number of parameters that may be irrelevant for the modelling of zebrafish embryo internal concentrations in specific cases. For instance, under normal conditions, there should be no proteins in the exposure medium. We chose to keep those features for compatibility with the VIVD model and greater generality of our embryo model, so that it can be used for *ab initio* predictions for a large number of chemicals.

3.3 Improving predictions by fitting VPA and analogs EU-ToxRisk data

For the *ab initio* predictions, metabolism was neglected. This may be acceptable for VPA analogs, since neglecting metabolism should lead to overestimated internal concentrations, while in fact the predictions were underestimates. To check the assumption of negligible metabolism, we estimated Cl_{met} by calibration with the data. Observed *versus* best fit (maximum posterior probability) concentrations are presented on Supporting Information's Figure S3. Cl_{met} value. Predictions of total embryo concentrations were out the ten-fold interval error and under-estimated, while predictions of medium concentration were included in the ten-fold error interval.

Supplemental Figure S5 shows the observed and predicted concentrations of VPA and its analogs in medium and in the total embryo (including yolk) as a function of time, after estimating Cl_{met} only. Table S6 (Supporting Information) summarizes the posteriors distributions of Cl_{met} and σ. The maximum posterior estimates are the most likely, best fitting, values. For the various compounds, the metabolic clearance best estimates were of the order of 10^{-15} to 10^{-13} L/min, which for an embryo volume of about 3×10^{-7} L correspond to half-lives well above 3500 hours (about 150 days). This implies negligible

metabolism of VPA and analogs in our zebrafish embryos and validates our *ab initio* choice of null metabolic clearances.

A better way to improve the internal concentration estimates of VPA analogs was to adjust the VIVD predicted partition coefficients using the concentration measurements. MCMC simulations were used to calibrate Cl_{met} and f_{pc} together on the basis of the data. Figure 3 plots observed *versus* predicted concentrations in that case. Here, the model predictions were also obtained using the best fitting parameter values. Overall, the points were better aligned with the perfect fit line and mostly comprised within the ten-fold error interval. Hexanoic acid, 4-pentenoic, 2-methylhexanoic acid and 2-methylpentanoic acid were the four worst predicted chemicals and did not fall in the three-fold error interval. The medium concentrations were under-predicted by the model with a median relative error of 0.72. The embryo concentrations were also slightly under-predicted with a median relative error of 0.83. The best estimate of σ corresponded on average to a factor 2.5.

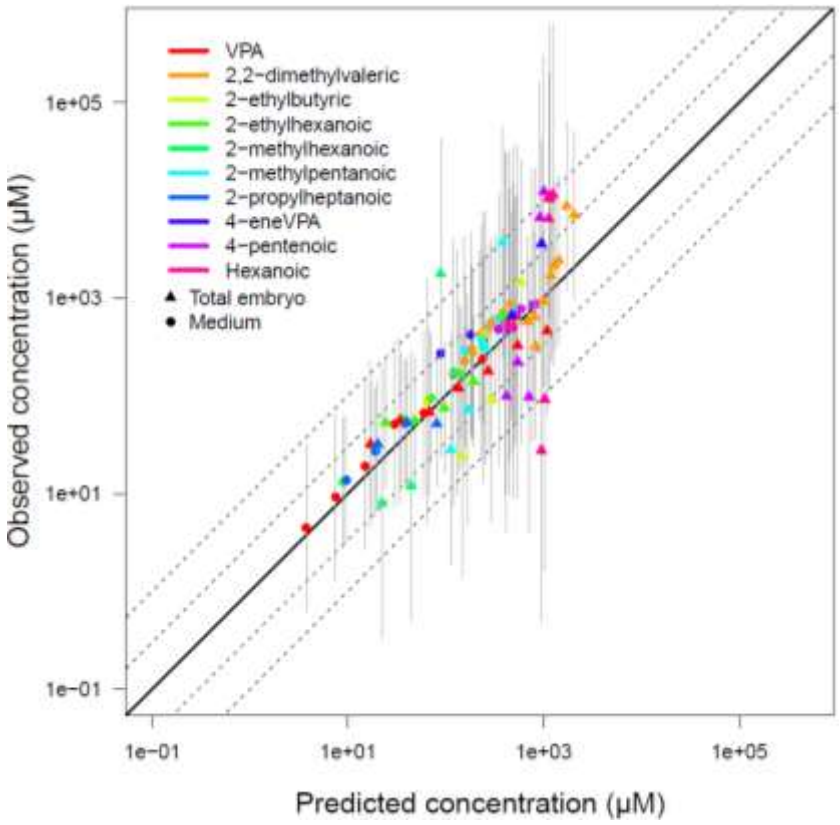

Figure 3: Observed *versus* predicted concentrations of valproic acid (VPA) and nine analogs, in the zebrafish embryo and culture medium. This is the best fit obtained after Cl_{met} and f_{pc} were simultaneously

estimated. The solid black line corresponds to perfect fit. Dashed lines correspond to the three- and ten-fold error intervals. The grey bars correspond to +/- one residual SD (σ).

Figure 4 presents the observed and predicted concentrations of VPA in medium and in the total embryo (including yolk) as a function of time, following exposure to various nominal concentrations, after simultaneous estimation of Cl_{met} and f_{pc}. Since the model was linear with dose, all concentrations were normalized to a nominal concentration of 1 mM to simplify the Figures (one model estimate only is needed for all doses). The discontinuities of the concentration-time curve are due to the daily changes of culture medium. Similar results are shown for the VPA analogs in Supplemental Figure S4. Almost all predicted medium and embryo time-course concentrations are in the 95% confidence interval of the model predictions (but those intervals can be large, showing a large uncertainty in measurements and consequently in model predictions).

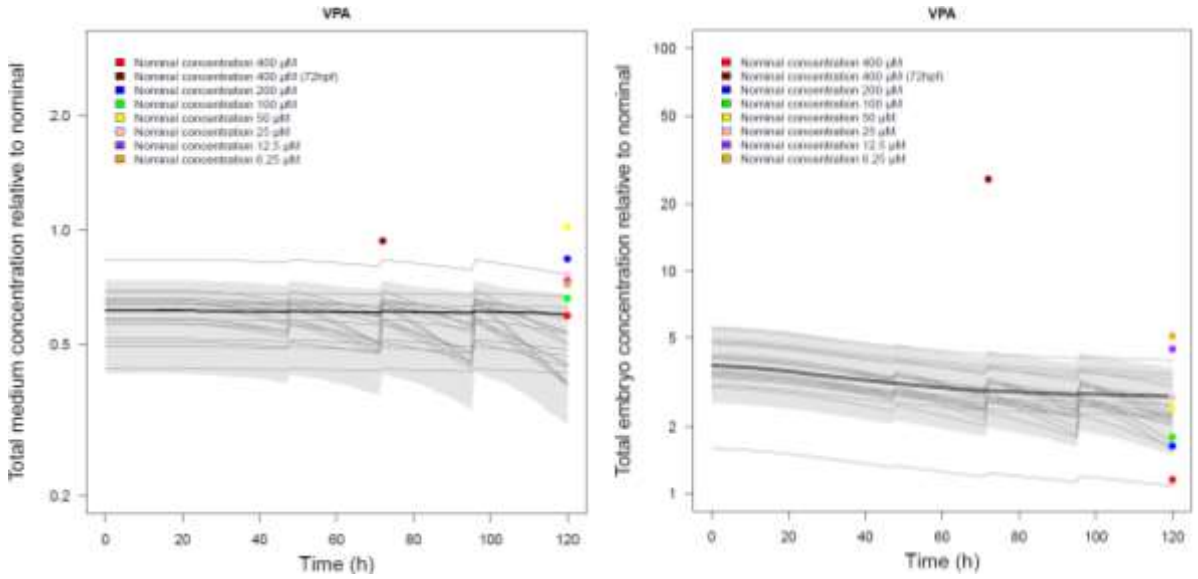

Figure 4: Predicted (lines) and observed (dots) VPA concentrations in medium (left) and in total embryo (right) as a function of time, after estimating Cl_{met} and f_{pc}. All concentrations were normalized to the nomimal concentration for plotting. Culture medium was replaced every day. The grey area defines the 95% confidence interval. The thick black line is the maximum posterior predicted concentration time-course. The thin lines are 20 predictions obtained using random parameter vectors drawn from their posterior distribution.

Table 1 summarizes, for each chemical, the posterior distributions of Cl_{met}, f_{pc} and uncertainty SD σ obtained by MCMC calibration with the concentration data. For the various compounds, the metabolic clearance best estimates were between 10^{-12} to 10^{-16} L/min. The highest upper confidence limit on a clearance was found for 2-Methylhexanoic acid, which for an embryo volume of about 3×10^{-7} L corresponds to a half-life of about 11 hours. The half-lives corresponding to the best estimates would be considerably higher. Therefore, the conclusion of negligible metabolism for those compounds in the zebrafish embryo appears to be coherent. The best estimates of the partition coefficient scaling factor, f_{pc}, are in the range of 0.608 (for 2-methylhexanoic acid) to 27.9 for 4-eneVPA. According to their 95% confidence intervals, the f_{pc} values for all compounds are significantly different from 1. It appears that the medium over tissue partition coefficient predicted by the VIVD model had to be increased to improve data fit for he compounds of interest. The values of σ, corresponding on average to a factor 2.5, shows large uncertainties in measurements and modeling. 2-methylhexanoic acid, 4-pentenoic acid and hexanoic acid are particularly affected, with uncertainties higher than a factor 3. Causes for such large uncertainties are multiple and cumulate their effects: the precision of the measurement method is limited, there is variability between embryos, initial concentration in the medium may be different from the nominal concentration, there may be some loss of the substance in the air, loss by degradation other than metabolic, loss or amplification during sample preparation, *etc.*

Table 1: Estimation of means, standard deviation (SD), 95% confidence intervals (IC95%) and maximum posterior (MP) value for the metabolic clearance, the partition coefficient correction factor, and residual uncertainty SD σ, for VPA and its analogs.

Compound	Cl_{met} (L/min)			f_{pc}			σ		
	Mean ± SD	IC 95%	MP	Mean ± SD	IC 95%	MP	Mean ± SD	IC 95%	MP
Valproic acid**	$3.55\times10^{-12} \pm 3.10\times10^{-12}$	$[1.38\times10^{-13}; 1.14\times10^{-11}]$	3.49×10^{-13}	20.0 ± 7.04	[9.46;36.9]	16.8	2.27 ± 0.397	[1.74;3.26]	1.97
2,2- Dimethylvaleric acid***	$1.67\times10^{-12} \pm 1.81\times10^{-12}$	$[5.25\times10^{-14}; 7.08\times10^{-12}]$	8.13×10^{-16}	2.28 ± 0.377	[8.61;30.4]	15.9	2.28 ± 0.377	[1.77;3.21]	2.00
2-Ethylbutyric acid**	$1.23\times10^{-11} \pm 1.20\times10^{-11}$	$[3.29\times10^{-13}; 4.32\times10^{-11}]$	5.59×10^{-13}	16.7 ± 10.4	[3.43;43.5]	11.9	3.44 ± 0.739	[2.32;5.24]	2.89
2-Ethylhexanoic acid***	$2.49\times10^{-12} \pm 2.20\times10^{-12}$	$[6.32\times10^{-14}; 8.36\times10^{-12}]$	4.16×10^{-14}	11.7 ± 3.24	[6.19;18.9]	10.6	1.87 ± 3.24	[1.52;2.73]	1.54
2-Methylhexanoic acid*	$8.67\times10^{-11} \pm 8.89\times10^{-11}$	$[1.80\times10^{-12}; 3.17\times10^{-10}]$	1.34×10^{-12}	1.76 ± 1.66	[0.257;6.86]	0.608	3.79 ± 0.803	[2.58;5.70]	3.18
2-Methylpentanoic acid**	$2.06\times10^{-11} \pm 1.71\times10^{-11}$	$[8.42\times10^{-13}; 6.65\times10^{-11}]$	1.76×10^{-12}	3.90 ± 2.12	[1.04 ; 8.96]	2.71	3.23 ± 0.693	[2.22;4.91]	2.71
2-Propylheptanoic acid***	$3.67\times10^{-12} \pm 2.62\times10^{-12}$	$[1.24\times10^{-13}; 9.41\times10^{-12}]$	3.43×10^{-13}	4.17 ± 1.78	[1.67;8.64]	3.13	2.02 ± 0.486	[1.51;3.33]	1.51
4-eneVPA***	$8.15\times10^{-12} \pm 8.15\times10^{-12}$	$[1.70\times10^{-13}; 3.09\times10^{-11}]$	2.67×10^{-14}	35.4 ± 22.7	[6.29;92.5]	27.9	3.33 ± 0.799	[2.07;5.11]	2.43
4-Pentenoic acid**	$1.51\times10^{-11} \pm 1.41\times10^{-11}$	$[6.73\times10^{-13}; 5.12\times10^{-11}]$	3.09×10^{-12}	6.93 ± 4.10	[1.60;17.5]	4.55	3.59 ± 0.669	[16.5;19.2]	3.20
Hexanoic acid***	$9.23\times10^{-12} \pm 8.65\times10^{-12}$	$[2.94\times10^{-13}; 3.46\times10^{-11}]$	3.59×10^{-13}	16.1 ± 11.3	[3.05;47.6]	9.10	4.36 ± 0.769	[3.15;6.19]	4.04

* Prior on CL_{met}: Uniform $(0, 10^{-9})$; prior on f_{pc}: Uniform (0, 5).

** Prior on CL_{met}: Uniform $(0, 10^{-10})$; prior on f_{pc}: Uniform (0, 5).

*** Prior on CL_{met}: Uniform $(0, 10^{-11})$; prior on f_{pc}: Uniform (0, 5).

A limitation of our model is that it does not consider the chorion nor active transport: The diffusion of chemicals within the embryo is assumed to be instantaneous. This assumption is reasonable given the small size of the embryo, but we do not yet have data to test its validity. We are currently working on improving the model to account for rate-limiting diffusions. Note that dechorionation could remove a significant fraction of chemical bound to the chorion. This would not affect the concentration measured in the embryo, but would prevent an assessment of mass balance and product loss (for example by metabolism) in the experiment if the chorionic concentration were not measured. It would be interesting to measure concentrations with and without the chorion to better understand its effects on the pharmacokinetics in the embryo. For transporters, there appear to be similarities between mammals and zebrafish embryo efflux transporters [49]. This may imply that active transport can play a role in modulating zebrafish embryos' exposure to xenobiotics, but we do not have sufficient to data to verify it for the VPA analogs investigated here. Likewise, our modeling of organ growth was based on limited data on organogenesis and embryo volumes. There is room for improvement with specific measurements of embryo volumes as a function of time. Note also that the Monte-Carlo simulated confidence intervals we estimated and presented on the Figures only reflect parametric uncertainty, and not structural model uncertainty.

Figure S6 shows the observed concentration**s** of VPA and analogs in the total embryo as a function of nominal concentration. Despite the small number of data points to infer on saturation, we also considered Michaelis-Menten kinetics as an alternative to linear ones, except for VPA, 2-propylheptanoic acid, and 4-eneVPA, for which linearity seemed to apply. In the case of saturable metabolism, parameters V_{max} and K_m were calibrated together with f_{pc}.

Figure S7 presents the kinetic profile for 2-ethylbutyric acid when saturable metabolism is assumed. The 95% confidence intervals were large and data fits for medium and total embryo were not improved compared to linear kinetics. Metabolism remained negligible.

3.4 Prediction of the VPA data of Brox *et al.*

Without any additional adjustment, we performed simulations of Brox *et al*. data on VPA with our best parameter estimates for Cl_{met} and f_{pc} (Table 1). Figure 5 shows an over-prediction with a median relative error of a factor 2. A difference of this order is expected because our above parameterization is quite uncertain and the laboratories which produced the data used somewhat different methods (dechorionation in Brox *et al.* experiments, different analytical methods *etc*.) This is a limited validation of the model, but at least for VPA, it predicts reasonably well a very different data set. Actually, some of the random prediction curves shown on Figure 5 are closer to the data. We examined the Cl_{met} and f_{pc} values leading to the four prediction curves closest to the data. Their means were 7.9 ± 2.6 picoL/min, and 13 ± 6, respectively, while the means in Table 1 are 3.55 picoL/min for Cl_{met}, and 20 for f_{pc}. So, Brox data point to a somewhat higher metabolism of VPA and lesser correction of the VIVD partition coefficient estimates than the EU-ToxRisk data.

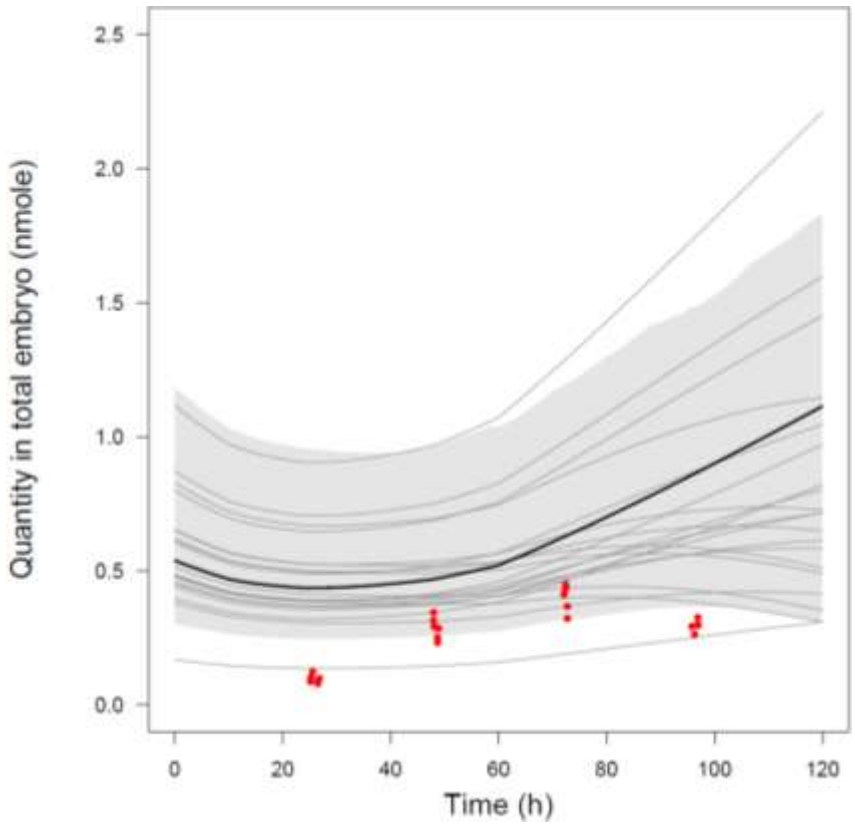

Figure 5: Predicted (lines) and data by Brox *et al*. (dots). VPA quantities in total embryo as a function of time, after estimating Cl_{met} and f_{pc} from our data. The culture medium was not replaced. The grey area

defines the 95% confidence interval. The thick black line is the maximum posterior prediction. The thin lines are 20 predictions obtained using random parameter vectors drawn from their posterior distribution.

3.5 Fit of the model to Brox *et al*. data

For all chemicals assayed by Brox *et al*. we also calibrated the model. Figure 6 shows the observed *versus* predicted quantities per embryo. Model predictions were obtained using the best fitting parameter values. Most points fall within the three-fold error interval, showing that the model can describe the data reasonably well. All the kinetic profiles (Figure S8) and fitted parameter values (Table S7) are shown in Supplementary information. The model captures most of the time courses correctly for those compounds. There are some misfits (Metribuzin, Phenytoin, Thiacloprid), which could be due to chemicals actively transported, or not penetrating the chorion, *etc*., but additional model complexity or fitting would be needed to improve this. Note that the metabolic clearance values obtained (Table S7) are again very low, and the ups and downs of the concentration time-courses are sufficiently explained by changes in organ sizes and yolk consumption. It would be very interesting to confirm this poor metabolic capacity of the embryo by measuring the expected metabolites, but this is challenging, because minute amounts of metabolites are expected to be formed. It might be worth designing experiments, in which the volume of the water medium would be very small (and the experiment time relatively short) to avoid low concentrations of the soluble metabolites in the medium.

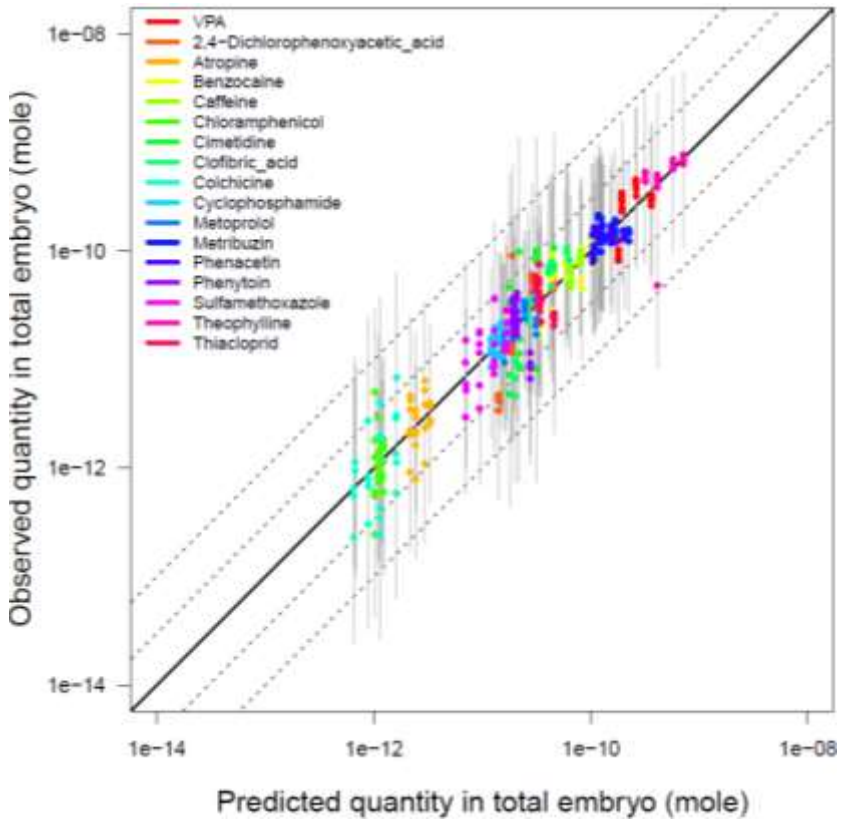

Figure 6: Observed *versus* predicted quantities in the embryo, for the 17 Brox *et al.* studied compounds. This is the best fit obtained with Cl_{met}, f_{pc} and $P_{a:w}$ simultaneous estimated. The black line corresponds to perfect fit. Dashed lines describe the three- and ten-fold error. The grey bars are the uncertainty σ.

3.6 Using the model to correct effective concentrations for pharmacokinetics

The main aim of our modeling effort was to estimate internal embryo concentrations in order to base effective concentration (EC) estimates on them – rather than on nominal concentration – for better mechanistic interpretations and improved risk assessments. Relating effects to internal concentrations should correct for pharmacokinetic differences between species. To show the impact of a proper accounting of cellular concentrations, we calculated two sets of ECs in the zebrafish embryo: For the first, we use nominal medium concentrations as a measure of dose. For the second, we used the model-predicted embryo concentration at 120 hpf (see Figure 4) as the dose. Because the model is linear with respect to dose, it is possible to obtain and apply a model-computed pharmacokinetic correction factor (f_k) to correct nominal dose ECs. This factor is specific to each compound and can be computed as:

$$f_k = \frac{Nominal\ concentration}{Model\text{-}predicted\ embryo\ concentration} \quad (35)$$

The corrected ECs are simply obtained as:

$$Corrected\ EC = \frac{Nominal\ dose\ EC}{Pharmacokinetic\ factor} \quad (36)$$

Table 2 shows the concentrations causing 10, 20 and 50 percent effects (embryo death or malformations, cumulated), uncorrected and corrected for pharmacokinetic at 120 hpf. The data-calibrated model (Figure 3, Table1) was used. Some compounds have high f_k values (almost a factor 5 for 2-methylhexanoic acid and hexanoic acid), but for 4-pentenoic acid, we estimated that no pharmacokinetic correction was necessary. Figure 7 compares the concentrations causing 10% of effects, before and after correction by f_k. We can observe a different ranking between the chemicals. The difference between the minimum and the maximum EC_{10} value is also wider after pharmacokinetic correction.

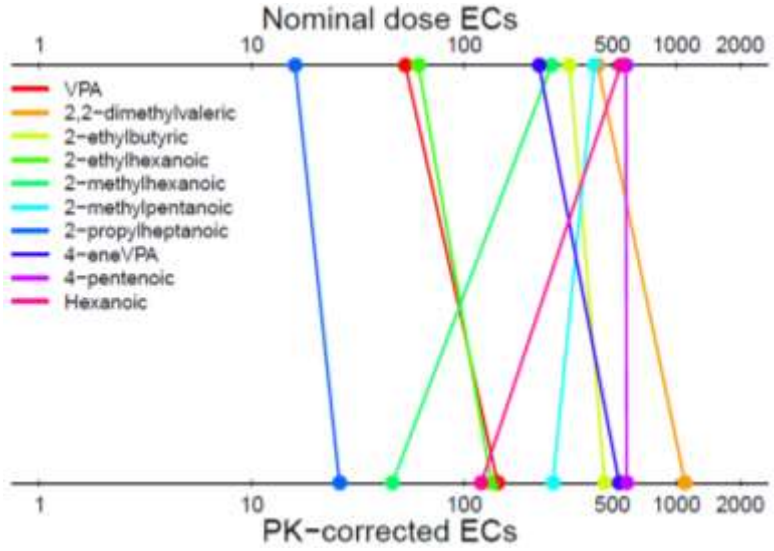

Figure 7: Illustration of the differences between VPA and analogs concentrations (in µM) inducing 10% effects (mortality or malformations, cumulated) when calculated from nominal dose or after correction for pharmacokinetics.

Despite its limitations, the model should be useful for risk assessment. It is simple to use and runs very quickly on a personal computer. For specific malformations or organ toxicity, the *ab initio* predicted tissue/organ concentrations should be used instead of the total embryo concentration. Obviously, if concentration measurements are available to improve the model, they should be used, even if that entails

some model fitting. The f_k values we obtained are also a useful summary of the bioaccumulation of VPA and its analogs in the zebrafish embryo. Our results indicate that all the VPA analogs tested accumulate in the embryo and that metabolic clearance is insignificant. For example, for the same medium concentrations, VPA and 2,2-Dimethylvaleric acid lead to the highest concentrations in the embryo (2.5 times the medium concentration), and 2-Methyl hexanoic acid the lowest (21% of the medium concentration) (Table 2). The pharmacokinetic correction should make *EC* more predictive in inter-species extrapolation (however, we did not find suitable animal or human data to confirm this). More importantly, it changes the ranking of the analogs and the order of priorities in a risk assessment context. For example, VPA becoming the 5th most potent (coming from the 2nd rank) and 4-eneVPA going to the 8th rank (from the 4th).

Table 2: Estimated pharmacokinetic correction factor (f_k) and of the 10, 20 and 50% effect concentrations for VPA and its analogs (at 120 hpf) in the zebrafish embryo, uncorrected and corrected for pharmacokinetics (after simultaneous Bayesian calibration Cl_{met} and f_{pc}).

Compound	**f_k**		**EC_{10} (µM)**		**EC_{20} (µM)**		**EC_{50} (µM)**	
	MPV	IC 95%	Based on nominal dose	PK corrected	Based on nominal dose	PK corrected	Based on nominal dose	PK corrected
Valproic acid	0.40	[0.34;0.47]	53	133	65	163	96	240
2,2-Dimethylvaleric acid	0.40	[0.36;0.46]	427	1068	445	1113	483	1208
2-Ethylbutyric acid	0.83	[0.62;1.1]	314	378	369	457	510	614
2-Ethylhexanoic acid	1.3	[1.2;1.3]	61	47	71	55	103	79
2-Methylhexanoic acid	4.8	[2.8;7.7]	258	54	280	58	333	69
2-Methylpentanoic acid	1.7	[1.2;2.4]	412	242	425	250	482	284
2-Propylheptanoic acid	0.60	[0.48;0.73]	16	27	17	28	20	33
4-eneVPA	0.54	[0.42;0.74]	226	419	235	435	253	469
4-Pentenoic acid	1.0	[0.72;1.5]	579	579	600	600	645	645
Hexanoic acid	4.5	[4.4;4.8]	548	122	557	124	575	128

4 Conclusions

We developed the structure and equations of the first PBPK model for the zebrafish embryo. Our model integrates previously developed predictive models to account for the impact of physicochemical properties of the chemicals of interest on partition between test system components, cells of different types, and sub-cellular organelles. It also describes metabolism and anatomical volume changes of the embryo during growth. Its structure and parameter values integrate a large amount of biological information gathered from the scientific literature (on organ volumes *etc*.). However, adult fish physicochemical tissue properties were used, except for the yolk, and future measurements of those parameters in embryos should improve our approximations.

The model can receive several applications. At a first level, using only physicochemical data and QSAR predictions of metabolism, it can be used as a high-throughput *ab initio* tool to make approximate predictions of chemical distribution in different embryonic tissues, as a function of time and exposure levels. At a second-level, when experimental data are available on medium and whole fish or tissue concentrations, the model parameters can be calibrated to estimate *in vivo* metabolism, and make finer predictions. Our model can also be used to guide further research and help design future experiments (for example on the impact of the chorion, effects of transporters on diffusion, *etc*.)

For the chemicals studied here, we found that metabolism was quite low and that the observed kinetics could be reasonably explained simply by volume changes. If this were true on average for most chemicals (which should be checked), the assumption of negligible metabolism could be made, alleviating the need for specific data. That being said, we wish that validated QSAR models were available for the zebrafish embryo.

Yet, our model has been checked with only two sets of concentration time-course data, even if at different exposure levels for the same chemical. Even though it performs rather well at predicting those data, it is certainly not “validated”, and we should not fully trust its predictions: The predictions are afflicted by large uncertainties, which can be estimated by Monte Carlo simulations. More data should be collected to better check and improve the model predictions. We showed by how much data fitting

can improve the model predictions and construe that as an incentive to develop more data. One of the virtues of PBPK models is actually to beg for more data and to direct research questions. The fact that animal PBPK models often have only plasma concentration data to “validate” them does not prevent their extensive use in the pharmaceutical industry, including in regulatory contexts. In any case, our model can be used to relate zebrafish embryo effects to cellular exposures, as demonstrated for VPA analogs. Its use should improve the extrapolation of zebrafish embryo data to human for safety assessment.

Associated content

Supporting Information

Table S1, Test and measured concentrations used in the FET; Table S2, Mass spectrometric and chromatography conditions; Table S3, Zebrafish embryo physiological parameter; Table S4, Partition coefficient values predicted by the VIVD model; Table S5, Compound's physicochemical properties; Table S6, Clearance estimation values if this is the only estimated parameter; Table S6, Parameter estimation values from Brox *et al*. data. Figure S1, Observed and modeled organ growth; Figure S2, Organ concentrations (µM) as a function of time (h), after estimating Cl_{met} and f_{pc} for VPA; Figure S3, Observed *versus* predicted concentrations with Cl_{met} estimate only; Figure S4, Predicted and observed VPA analogs' concentrations after estimating Cl_{met} and f_{pc}; Figure S5, Predicted and observed VPA analogs' concentrations after estimating Cl_{met} only; Figure S6, VPA and analogs observed concentrations in total embryo as a function of nominal concentration; Figure S7, Predicted and observed 2-ethylbutyric acid concentrations after estimating V_{max}, K_m and f_{pc}, according to Michaelis-Menten kinetic assumption; (PDF); Figure S8, Predicted and observed Brox *et al*. compound concentrations after estimating Cl_{met} and f_{pc} and $P_{a:w}$; Model Code.

Corresponding author

Email: *frederic.bois@certara.com

Present Addresses

Steve Silvester is currently employed at Alderley Analytical Ltd. Alderley Park, Macclesfield, Cheshire, SK10 4TG, United Kingdom.

Richard Maclennan is currently employed at Appleyard Lees IP LLP, The Lexicon Mount Street, Manchester, Greater Manchester, M2 5NT, United Kingdom.

Notes

The authors declare no competing financial interest.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 681002 (*Eu-ToxRisk*).

References

[1] A.J. Hill, H. Teraoka, W. Heideman, R.E. Peterson, Zebrafish as a Model Vertebrate for Investigating Chemical Toxicity, Toxicological Sciences. 86 (2005) 6–19. https://doi.org/10.1093/toxsci/kfi110.

[2] W. Larisch, K.-U. Goss, Modelling oral up-take of hydrophobic and super-hydrophobic chemicals in fish, Environmental Science: Processes & Impacts. (2018). https://doi.org/10.1039/C7EM00495H.

[3] W. Larisch, T.N. Brown, K.-U. Goss, A toxicokinetic model for fish including multiphase sorption features: A high-detailed, physiologically based toxicokinetic model, Environmental Toxicology and Chemistry. 36 (2017) 1538–1546. https://doi.org/10.1002/etc.3677.

[4] A. Bernut, G. Lutfalla, L. Kremer, Regard à travers le danio pour mieux comprendre les interactions hôte/pathogène, Médecine/Sciences. 31 (2015) 638–646. https://doi.org/10.1051/medsci/20153106017.

[5] J.R. Goldsmith, C. Jobin, Think Small: Zebrafish as a Model System of Human Pathology, Journal of Biomedicine and Biotechnology. 2012 (2012) 1–12. https://doi.org/10.1155/2012/817341.

[6] R. Nagel, DarT: The embryo test with the Zebrafish Danio rerio--a general model in ecotoxicology and toxicology, ALTEX. 19 Suppl 1 (2002) 38–48.

[7] K. Howe, M.D. Clark, C.F. Torroja, J. Torrance, C. Berthelot, M. Muffato, J.E. Collins, S. Humphray, K. McLaren, L. Matthews, S. McLaren, I. Sealy, M. Caccamo, C. Churcher, C. Scott, J.C. Barrett, R. Koch, G.-J. Rauch, S. White, W. Chow, B. Kilian, L.T. Quintais, J.A. Guerra-Assunção, Y. Zhou, Y. Gu, J. Yen, J.-H. Vogel, T. Eyre, S. Redmond, R. Banerjee, J. Chi, B. Fu, E. Langley, S.F. Maguire, G.K. Laird, D. Lloyd, E. Kenyon, S. Donaldson, H. Sehra, J. Almeida-King, J. Loveland, S. Trevanion, M. Jones, M. Quail, D. Willey, A. Hunt, J. Burton, S. Sims, K. McLay, B. Plumb, J. Davis, C. Clee, K. Oliver, R. Clark, C. Riddle, D. Eliott, G. Threadgold, G. Harden, D. Ware, B. Mortimer, G. Kerry, P. Heath, B. Phillimore, A. Tracey, N. Corby, M. Dunn, C. Johnson, J. Wood, S. Clark, S. Pelan, G. Griffiths, M. Smith, R. Glithero, P. Howden, N. Barker, C. Stevens, J. Harley, K. Holt, G. Panagiotidis, J. Lovell, H. Beasley, C. Henderson, D. Gordon, K. Auger, D. Wright, J. Collins, C. Raisen, L. Dyer, K. Leung, L. Robertson, K. Ambridge, D. Leongamornlert, S. McGuire, R. Gilderthorp, C.

Griffiths, D. Manthravadi, S. Nichol, G. Barker, S. Whitehead, M. Kay, J. Brown, C. Murnane, E. Gray, M. Humphries, N. Sycamore, D. Barker, D. Saunders, J. Wallis, A. Babbage, S. Hammond, M. Mashreghi-Mohammadi, L. Barr, S. Martin, P. Wray, A. Ellington, N. Matthews, M. Ellwood, R. Woodmansey, G. Clark, J. Cooper, A. Tromans, D. Grafham, C. Skuce, R. Pandian, R. Andrews, E. Harrison, A. Kimberley, J. Garnett, N. Fosker, R. Hall, P. Garner, D. Kelly, C. Bird, S. Palmer, I. Gehring, A. Berger, C.M. Dooley, Z. Ersan-Ürün, C. Eser, H. Geiger, M. Geisler, L. Karotki, A. Kirn, J. Konantz, M. Konantz, M. Oberländer, S. Rudolph-Geiger, M. Teucke, K. Osoegawa, B. Zhu, A. Rapp, S. Widaa, C. Langford, F. Yang, N.P. Carter, J. Harrow, Z. Ning, J. Herrero, S.M.J. Searle, A. Enright, R. Geisler, R.H.A. Plasterk, C. Lee, M. Westerfield, P.J. de Jong, L.I. Zon, J.H. Postlethwait, C. Nüsslein-Volhard, T.J.P. Hubbard, H.R. Crollius, J. Rogers, D.L. Stemple, The zebrafish reference genome sequence and its relationship to the human genome, Nature. 496 (2013) 498–503. https://doi.org/10.1038/nature12111.

[8] C.S. Martinez, D.A. Feas, M. Siri, D.E. Igartúa, N.S. Chiaramoni, S. del V. Alonso, M.J. Prieto, In vivo study of teratogenic and anticonvulsant effects of antiepileptics drugs in zebrafish embryo and larvae, Neurotoxicology and Teratology. 66 (2018) 17–24. https://doi.org/10.1016/j.ntt.2018.01.008.

[9] M. Ekker, M.-A. Akimento, Le poisson zèbre (danio rerio) un modèle en biologie du développement, Médecine/Sciences. 7 (1991) 553–560.

[10] B. Kais, R. Ottermanns, F. Scheller, T. Braunbeck, Modification and quantification of in vivo EROD live-imaging with zebrafish (Danio rerio) embryos to detect both induction and inhibition of CYP1A, Science of The Total Environment. 615 (2018) 330–347. https://doi.org/10.1016/j.scitotenv.2017.09.257.

[11] U. Strähle, S. Scholz, R. Geisler, P. Greiner, H. Hollert, S. Rastegar, A. Schumacher, I. Selderslaghs, C. Weiss, H. Witters, T. Braunbeck, Zebrafish embryos as an alternative to animal experiments—A commentary on the definition of the onset of protected life stages in animal welfare regulations, Reproductive Toxicology. 33 (2012) 128–132. https://doi.org/10.1016/j.reprotox.2011.06.121.

[12] N. Quignot, J. Hamon, F.Y. Bois, Extrapolating <i>in vitro<\i> results to predict human toxicity, in: A. Bal-Price, P. Jennings (Eds.), In Vitro Toxicology Systems, Springer Science, New-York, 2014: pp. 531–550.

[13] P.F. Landrum, M.J. Lydy, H. Lee, Toxicokinetics in aquatic systems: Model comparisons and use in hazard assessment, Environmental Toxicology and Chemistry. 11 (1992) 1709–1725. https://doi.org/10.1002/etc.5620111205.

[14] E.S. Salmina, D. Wondrousch, R. Kühne, V.A. Potemkin, G. Schüürmann, Variation in predicted internal concentrations in relation to PBPK model complexity for rainbow trout, Science of The Total Environment. 550 (2016) 586–597. https://doi.org/10.1016/j.scitotenv.2016.01.107.

[15] K. Krishnan, T. Peyret, Physiologically Based Toxicokinetic (PBTK) Modeling in Ecotoxicology, in: J. Devillers (Ed.), Ecotoxicology Modeling, Springer US, Boston, MA, 2009: pp. 145–175. https://doi.org/10.1007/978-1-4419-0197-2_6.

[16] J. Feng, Y. Gao, M. Chen, X. Xu, M. Huang, T. Yang, N. Chen, L. Zhu, Predicting cadmium and lead toxicities in zebrafish (Danio rerio) larvae by using a toxicokinetic–toxicodynamic model that considers the effects of cations, Science of The Total Environment. 625 (2018) 1584–1595. https://doi.org/10.1016/j.scitotenv.2018.01.068.

[17] M. Khazaee, C.A. Ng, Evaluating parameter availability for physiologically based pharmacokinetic (PBPK) modeling of perfluorooctanoic acid (PFOA) in zebrafish, Environmental Science: Processes & Impacts. 20 (2018) 105–119. https://doi.org/10.1039/C7EM00474E.

[18] A.R.R. Péry, J. Devillers, C. Brochot, E. Mombelli, O. Palluel, B. Piccini, F. Brion, R. Beaudouin, A Physiologically Based Toxicokinetic Model for the Zebrafish *Danio rerio*, Environmental Science & Technology. 48 (2014) 781–790. https://doi.org/10.1021/es404301q.
[19] A. Grech, C. Tebby, C. Brochot, F.Y. Bois, A. Bado-Nilles, J.-L. Dorne, N. Quignot, R. Beaudouin, Generic physiologically-based toxicokinetic modelling for fish: Integration of environmental factors and species variability, Science of The Total Environment. 651 (2019) 516–531. https://doi.org/10.1016/j.scitotenv.2018.09.163.
[20] M. Brinkmann, C. Schlechtriem, M. Reininghaus, K. Eichbaum, S. Buchinger, G. Reifferscheid, H. Hollert, T.G. Preuss, Cross-species extrapolation of uptake and disposition of neutral organic chemicals in fish using a multispecies physiologically-based toxicokinetic model framework, Environmental Science and Technology. 50 (2016) 1914–1923. https://doi.org/10.1021/acs.est.5b06158.
[21] S. Brox, B. Seiwert, E. Küster, T. Reemtsma, Toxicokinetics of Polar Chemicals in Zebrafish Embryo (*Danio rerio*): Influence of Physicochemical Properties and of Biological Processes, Environmental Science & Technology. 50 (2016) 10264–10272. https://doi.org/10.1021/acs.est.6b04325.
[22] F.A.P.C. Gobas, X. Zhang, Measuring bioconcentration factors and rate constants of chemicals in aquatic organisms under conditions of variable water concentrations and short exposure time, Chemosphere. 25 (1992) 1961–1971. https://doi.org/10.1016/0045-6535(92)90035-P.
[23] J.C. Otte, B. Schultz, D. Fruth, E. Fabian, B. van Ravenzwaay, B. Hidding, E.R. Salinas, Intrinsic xenobiotic metabolizing enzyme activities in early life stages of zebrafish (*Danio rerio*), Toxicological Sciences. 159 (2017) 86–93. https://doi.org/10.1093/toxsci/kfx116.
[24] C. Fisher, S. Siméon, M. Jamei, I. Gardner, Y.F. Bois, VIVD: Virtual in vitro distribution model for the mechanistic prediction of intracellular concentrations of chemicals in in vitro toxicity assays, Toxicology in Vitro. 58 (2019) 42–50. https://doi.org/10.1016/j.tiv.2018.12.017.
[25] K. Fathe, A. Palacios, R.H. Finnell, Brief report novel mechanism for valproate-induced teratogenicity: Novel Mechanism for Valproate-Induced Teratogenicity, Birth Defects Research Part A: Clinical and Molecular Teratology. 100 (2014) 592–597. https://doi.org/10.1002/bdra.23277.
[26] C.-M. Chuang, C.-H. Chang, H.-E. Wang, K.-C. Chen, C.-C. Peng, C.-L. Hsieh, R.Y. Peng, Valproic Acid Downregulates RBP4 and Elicits Hypervitaminosis A-Teratogenesis—A Kinetic Analysis on Retinol/Retinoic Acid Homeostatic System, PLoS ONE. 7 (2012) e43692. https://doi.org/10.1371/journal.pone.0043692.
[27] C.J. Phiel, F. Zhang, E.Y. Huang, M.G. Guenther, M.A. Lazar, P.S. Klein, Histone Deacetylase Is a Direct Target of Valproic Acid, a Potent Anticonvulsant, Mood Stabilizer, and Teratogen, Journal of Biological Chemistry. 276 (2001) 36734–36741. https://doi.org/10.1074/jbc.M101287200.
[28] T. Braunbeck, E. Lammer, Fish Embryo Toxicity Assays (UBA contract number 20385422), (2006).
[29] E. Lammer, G.J. Carr, K. Wendler, J.M. Rawlings, S.E. Belanger, Th. Braunbeck, Is the fish embryo toxicity test (FET) with the zebrafish (Danio rerio) a potential alternative for the fish acute toxicity test?, Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology. 149 (2009) 196–209. https://doi.org/10.1016/j.cbpc.2008.11.006.
[30] T. Braunbeck, B. Kais, E. Lammer, J. Otte, K. Schneider, D. Stengel, R. Strecker, The fish embryo test (FET): origin, applications, and future, Environmental Science and Pollution Research. 22 (2015) 16247–16261. https://doi.org/10.1007/s11356-014-3814-7.
[31] M.R. Embry, S.E. Belanger, T.A. Braunbeck, M. Galay-Burgos, M. Halder, D.E. Hinton, M.A. Léonard, A. Lillicrap, T. Norberg-King, G. Whale, The fish embryo toxicity test as an animal

alternative method in hazard and risk assessment and scientific research, Aquatic Toxicology. 97 (2010) 79–87. https://doi.org/10.1016/j.aquatox.2009.12.008.
[32] ISO, International Organization for Standardization. Water quality - Determination of the 28 acute lethal toxicity of substances to a freshwater fish [Brachydanio rerio Hamilton-Buchanan 29 (Teleostei, Cyprinidae)]. ISO 7346-3: Flow-through method. Available: [http://www.iso.org]., (1996).
[33] OECD, OECD guidelines for the testing of chemicals. Fish Embryo Acute Toxicity (FET) Test, (2013). https://www.oecd-ilibrary.org/environment/test-no-236-fish-embryo-acute-toxicity-fet-test_9789264203709-en.
[34] T. Braunbeck, M. Boettcher, H. Hollert, T. Kosmehl, E. Lammer, E. Leist, M. Rudolf, N. Seitz, Towards an alternative for the acute fish LC(50) test in chemical assessment: the fish embryo toxicity test goes multi-species -- an update, ALTEX. 22 (2005) 87–102.
[35] H. Hollert, S. Keiter, N. König, M. Rudolf, M. Ulrich, T. Braunbeck, A new sediment contact assay to assess particle-bound pollutants using zebrafish (danio rerio) embryos, Journal of Soils and Sediments. 3 (2003) 197–207. https://doi.org/10.1065/jss2003.09.085.
[36] C.B. Kimmel, W.W. Ballard, S.R. Kimmel, B. Ullmann, T.F. Schilling, Stages of embryonic development of the zebrafish, Developmental Dynamics. 203 (1995) 253–310. https://doi.org/10.1002/aja.1002030302.
[37] N. Hachicho, S. Reithel, A. Miltner, H.J. Heipieper, E. Küster, T. Luckenbach, Body mass parameters, lipid profiles and protein contents of zebrafish embryos and effects of 2,4-dinitrophenol exposure, PLoS One. 10 (2015) e0134755. https://doi.org/10.1371/journal.pone.0134755.
[38] M. Hagedorn, F.W. Kleinhans, R. Freitas, J. Liu, E.W. Hsu, D.E. Wildt, W.F. Rall, Water distribution and permeability of zebrafish embryos, Brachydanio rerio, The Journal of Experimental Zoology. 278 (1997) 356–371. https://doi.org/10.1002/(sici)1097-010x(19970815)278:6<356::aid-jez3>3.0.co;2-n.
[39] P. Bernillon, F.Y. Bois, Statistical issues in toxicokinetic modeling: a Bayesian perspective, Environmental Health Perspectives. 108 (suppl. 5) (2000) 883–893.
[40] F.Y. Bois, Bayesian inference, in: B. Reisfeld, A.N. Mayeno (Eds.), Computational Toxicology Vol. II, Humana Press, New-York, 2012: pp. 597–636.
[41] A.F.M. Smith, G.O. Roberts, Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods, Journal of the Royal Statistical Society Series B. 55 (1993) 3–23.
[42] A. Gelman, D.B. Rubin, Inference from iterative simulation using multiple sequences (with discussion), Statistical Science. 7 (1992) 457–511.
[43] J.W. Nichols, J.M. McKim, M.E. Andersen, M.L. Gargas, H.J. Clewell, R.J. Erickson, A physiologically based toxicokinetic model for the uptake and disposition of waterborne organic chemicals in fish, Toxicol. Appl. Pharmacol. 106 (1990) 433–447.
[44] R Development Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria., 2013. http://www.R-project.org.
[45] F.Y. Bois, GNU MCSim: Bayesian statistical inference for SBML-coded systems biology models, Bioinformatics. 25 (2009) 1453–1454. https://doi.org/10.1093/bioinformatics/btp162.
[46] J.M.Z. Comenges, E. Joossens, J.V.S. Benito, A. Worth, A. Paini, Theoretical and mathematical foundation of the Virtual Cell Based Assay – A review, Toxicology in Vitro. 45 (2017) 209–221. https://doi.org/10.1016/j.tiv.2016.07.013.
[47] E. Papa, L. van der Wal, J.A. Arnot, P. Gramatica, Metabolic biotransformation half-lives in fish: QSAR modeling and consensus analysis, Science of The Total Environment. 470–471 (2014) 1040–1046. https://doi.org/10.1016/j.scitotenv.2013.10.068.
[48] J.A. Arnot, W. Meylan, J. Tunkel, P.H. Howard, D. Mackay, M. Bonnell, R.S. Boethling, A QUANTITATIVE STRUCTURE–ACTIVITY RELATIONSHIP FOR PREDICTING

METABOLIC BIOTRANSFORMATION RATES FOR ORGANIC CHEMICALS IN FISH, Environ Toxicol Chem. 28 (2009) 1168. https://doi.org/10.1897/08-289.1.

[49] S. Fischer, N. Klüver, K. Burkhardt-Medicke, M. Pietsch, A.-M. Schmidt, P. Wellner, K. Schirmer, T. Luckenbach, Abcb4 acts as multixenobiotic transporter and active barrier against chemical uptake in zebrafish (Danio rerio) embryos, BMC Biology. 11 (2013) 69. https://doi.org/10.1186/1741-7007-11-69.