

The PrecisionTox chemical library: creation of a chemical collection to discover evolutionary conserved biomolecular signatures of toxicity

Rubén Martínez ¹, Juan Carlos González-Sánchez², Stavroula I. Sampani³, Stefan Scholz⁴, Beate I. Escher⁵, Luise Henneberger⁵, Julia Huchthausen⁵, Maurice Whelan³, Thomas Dickmeis⁶, Carsten Weiss⁶, John K. Colbourne^{7,*}, Jonathan H. Freedman ^{7,8,*}

¹Leitat Technological Center, Ecotoxicology & Biodegradability Unit, Terrassa, 08225, Spain

²Biochemistry Center (BZH), Heidelberg University, Heidelberg, 69120, Germany

³European Commission, Joint Research Centre (JRC), Ispra (VA), 21027, Italy

⁴Department of Ecotoxicology, Helmholtz Centre for Environmental Research—UFZ, Leipzig, 04318, Germany

⁵Department of Cell Toxicology, Helmholtz Centre for Environmental Research—UFZ, Leipzig, 04318, Germany

⁶Institute of Biological and Chemical Systems—Biological Information Processing, Karlsruhe Institute of Technology, Eggenstein-Leopoldshafen, 76344, Germany

⁷Centre for Environmental Research and Justice and School of Biosciences, University of Birmingham, Birmingham, B15 2TT, UK

⁸Gillings School of Global Public Health, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, United States

*Corresponding authors: Jonathan H. Freedman, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, United States. Email:

j.freedman.1@bham.ac.uk; John K. Colbourne, School of Biosciences, University of Birmingham, Birmingham, B15 2TT, UK. Email:

J.K.Colbourne@bham.ac.uk

Abstract

The large number and diversity of chemicals currently in use present significant challenges in assessing their human and environmental health risks due to a paucity of toxicological data. To address this shortage, high-throughput screening technologies are used to rapidly evaluate the toxicity of these chemicals. Suitable chemical libraries are crucial to evaluate the performance of these technologies and generate the cognate toxicity data. Unlike traditional chemical libraries designed for specific disease targets or receptor interactions, the PrecisionTox collection prioritizes diversity in targets and mechanisms of toxicity to ensure broad applicability in toxicity predictions to test the concept of phylotoxology. Phylotoxology proposes that mechanisms of toxicity are evolutionarily conserved among distantly related species. Furthermore, the application of phylotoxology can contribute to the reduction of mammalian species in toxicity testing. Here, an approach for generating a chemical library based on chemical properties—physicochemical, biomolecular, and toxicological—as well as practical considerations, including compound availability, cost, purity, and shipping regulations, is reported. From an initial pool of over 1,500 nominees, a set of 200 chemicals was selected based on multiple criteria, including organ toxicity, environmental exposure, structure, modes of action, and toxicological relevance. Additionally, information on baseline toxicity, Absorption, Distribution, Metabolism, and Excretion properties and utility for in vitro testing was collected. This work underscores the necessity of thoughtful chemical selection to refine toxicological models, improve hazard identification, and support regulatory efforts to protect human and environmental health.

Keywords: chemical libraries; new approach methodologies; adverse outcome pathways; physicochemical properties; baseline-toxicity prediction model

An estimated 350,000 chemicals are being used both commercially and globally (Wang et al. 2020). Humans and other species are exposed to many of these chemicals, for which frequently little to no toxicological information is available (Wang et al. 2020; Persson et al. 2022). To address this gap in chemical risk and hazard information, methods have been developed to measure a variety of toxicological endpoints using in vivo and in vitro high-throughput screening (HTS) technologies (Richard et al. 2016). At the same time, New Approach Methodologies (NAMs) offer a sustainable alternative to animal testing (Colbourne et al. 2025). NAMs enhance human relevance and can reduce costs and resources, facilitating the rapid screening of large numbers of chemicals and the identification of potential hazards. In vitro HTS platforms acquire data at

rates of millions of chemicals per week (Wunder et al. 2008), whereas in vivo platforms can examine tens to a few thousand chemicals per week (Murphey and Zon 2006).

The ability to rapidly measure toxicological effects of large numbers of chemicals using different endpoints and analyze the cognate data continues to develop. Chemical libraries have been created using an array of criteria to test toxicological endpoints. For instance, compounds have been selected to target specific organs (liver, kidney, heart), disease endpoints (cancer, developmental neurotoxicity, fibrosis), receptors or enzymes (serotonin transporters, G-protein coupled receptors, acetylcholinesterase), biological pathways (fatty acid oxidation, protein kinase, endocrine/androgen), chemical classes (metal/metalloids, pesticides),

or chemical structures (bisphenols, perfluorooctanoic acids, polycyclic aromatic hydrocarbons) (Richard et al. 2021).

A variety of approaches have been used to collect compounds for testing with known or anticipated toxicities into training sets or reference libraries. The simplest approach is to use chemicals contained in predefined lists. Agencies throughout the world, such as the Joint Research Centre (JRC), the Organization for Economic Co-operation and Development (OECD), and ToxCast and the Partnership for the Assessment of Risks from Chemicals (PARC) (Snapkow et al. 2024), are constantly assembling lists of chemicals that meet different criteria (e.g. JRC validation sets, OECD Test Guidelines or Guidance documents). Other methods include literature mining to identify compounds with specific toxicological characteristics (Canham et al. 2020), generating and subsequently synthesizing compounds based on Quantitative structure–activity relationship (QSAR) models, and synthesizing congeners of known drugs (Ashenden 2018). The number of compounds in these libraries can range from 10's for small projects to >10,000 for larger projects such as Tox21 (Richard et al. 2021).

PrecisionTox is an international consortium formed under the European Commission Horizon 2020 program with the goal of identifying toxicity pathways and cognate biomarkers associated with chemical exposure using NAMs (PrecisionTox Consortium 2023). To achieve this goal, PrecisionTox is testing the concept of phylo-toxicology, which proposes that many biomolecular and toxicological responses to chemical exposure are shared among distantly related species by evolutionarily conserved processes and pathways. Thus, screening chemicals in a suite of phylogenetically diverse species, including biomedical and ecologically relevant models, can be used to assess chemical hazards to wildlife and humans without the use of traditional mammalian species (Colbourne et al. 2022). To evaluate this concept, a collection of 200 compounds was assembled and tested in a human cell line, *Caenorhabditis elegans*, *Daphnia magna*, *Drosophila melanogaster*, and embryos of *Danio rerio* and *Xenopus laevis*. Unlike other libraries and smaller test sets, the defining characteristic of the PrecisionTox collection is its diversity in target organs, receptor binding, affected biological pathways, and chemical class. Consequently, there are hundreds of thousands of potential candidates.

To choose 200 chemicals for this collection, a chemical selection strategy was established to select compounds that target specific organs, have environmental relevance, particular chemical structures, and/or modes of action (MoA). This collection was designed to be a community resource that contains potential “MOA anchor chemicals” (OECD 2017). Here, the rationale for choosing chemicals for the PrecisionTox collection is presented. To define the characteristics of this collection, experimental and theoretical data on physicochemical, toxicological, and Absorption, Distribution, Metabolism, and Excretion (ADME) characteristics were collected. For easier access to this information, the data has been incorporated into the PrecisionTox Data Visualization Tool. Ultimately, this chemical collection and cognate data will be combined with transcriptomic- and metabolomic-based toxicological data from the 6 alternative test models to test the phylo-toxicology hypothesis with the goal of advancing the use of NAMs (PrecisionTox Consortium 2023).

Materials and methods

Selection of chemicals for the PrecisionTox collection

Several overall considerations were maintained during the chemical selection process. This included identifying reference

compounds, well-studied chemicals with large amounts of associated toxicological and mechanistic data. Similarly, chemicals with limited or poorly understood mechanistic data were selected to test future toxicity predictions. General and taxon-specific toxicants, which targeted taxa from various parts of the phylogenetic tree, were selected. These were used to test the phylo-toxicology hypothesis. Assembling the chemical collection also considered identifying and addressing issues in the logistics of obtaining and shipping chemicals to international partners. Additionally, it enabled the various groups in PrecisionTox to develop Standard Operating Procedures for toxicity testing, database development, and analysis processes.

Chemical nominations

The first step in building the chemical collection was to define its chemical space. From discussions within PrecisionTox, stakeholders, and partner EU projects, including Animal-free Safety Assessment of Chemicals: Project Cluster for Implementation of Novel Strategies (ASPIs), priority was given to several chemical groups (Table 1). ASPIs has a strong focus on hepatotoxicants, nephrotoxicants, and developmental neurotoxicants (DNTs). Thus, the phylogenetic approach of PrecisionTox is synergistic and mutually beneficial to the project cluster by including these chemical groups. Chemical structure-specific toxicants were selected, including conazoles, acrylamides, and imidazoles. This group of chemicals was selected as part of collaborative case-study projects investigating the grouping of substances based on bioactivity data from multiple testing platforms between PrecisionTox, chemical regulatory agencies (acrylamides and imidazoles), and PARC (conazoles). Furthermore, environmentally relevant chemicals were prioritized (Finckh et al. 2022; Scholz et al. 2022; Finckh et al. 2024). To maintain diversity in the PrecisionTox collection, a target number of 20 to 30 chemicals per group was chosen.

After defining the toxicological space in Table 1, the next step was to compile chemical nominees for inclusion into the collection. Nominees were obtained from consortium members and by manually mining available datasets from the US National Toxicology Program, the JRC (Sund and Deceuninck 2021), and the literature. Additional nominees were obtained from other international projects involved in developing NAMs and non-animal testing paradigms (Westmoreland et al. 2022; De Castelbajac et al. 2023). These were included to maximize collaborative efforts among these projects (Fig. 1). Some chemicals were excluded from nomination because of the biological limitations of the PrecisionTox test organisms. For example, pulmonary and dermal toxicants were not considered because none of the test species have systems analogous to lungs or mammalian skin.

Table 1. Groups used in chemical selections.

Organ-specific toxicants	Reference
Hepatotoxicants	(Albrecht et al. 2019)
Nephrotoxicants	(Su et al. 2016)
Developmental neurotoxicants	(Aschner et al. 2017; Blum et al. 2023)
Cardiotoxicants	(Sirenko et al. 2017)
Chemical class-specific toxicants	
Conazoles	(Peyton et al. 2015)
Acrylamides	(Koszucka et al. 2020)
Imidazoles	(Sharma et al. 2021)
Environmentally relevant chemicals	(Finckh et al. 2022; Scholz et al. 2022; Finckh et al. 2024)

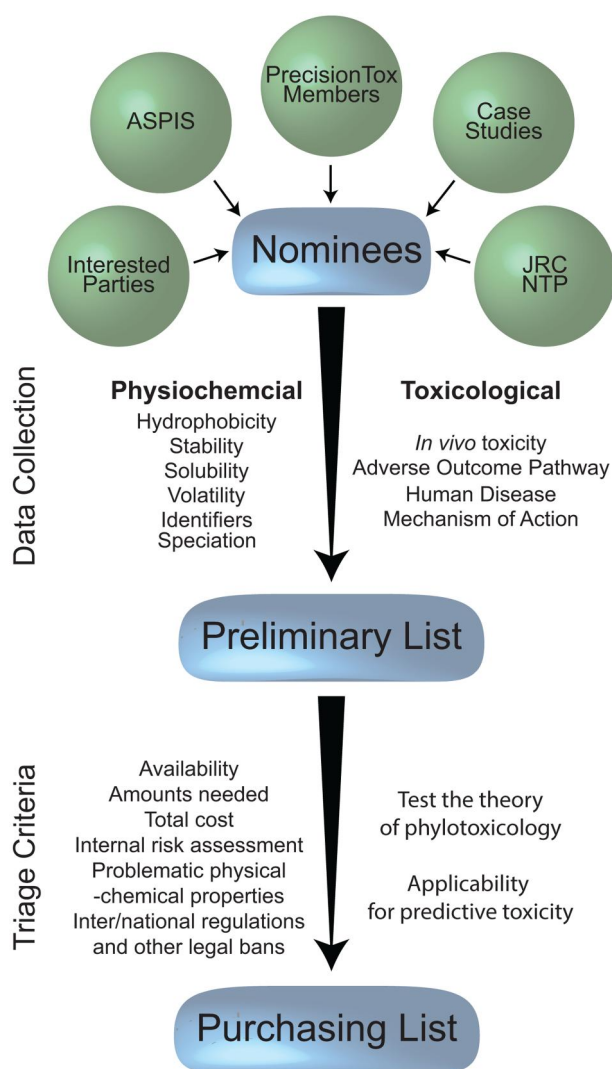


Fig. 1. Chemical selection process. Diagram of the PrecisionTox chemical collection selection process. Briefly, chemical nominees are obtained from multiple sources from within and without the consortium. After collecting physicochemical and toxicological data, nominees are selected for purchase based on the triage criteria.

Triaging process

Preliminary lists of several hundred chemical nominees within each group of toxicants were subjected to a tiered selection approach. First, following the compilation of physicochemical properties of the nominees, those with physicochemical properties outside the applicability domain or resulting in reduced bioavailability in the bioassays were removed. Those properties included volatility and hydrophobicity, as well as stability and solubility in aqueous media and/or DMSO. Volatile chemicals were removed based on an empirical threshold of ionization-corrected air–water partition constant $D_{aw} > 10^{-4}$ (Escher et al. 2019a). Hydrophobic chemicals with $\log D_{lip/w}$ (pH 7.4) > 4 were avoided to meet the criteria of in vitro assays, as they are challenging to dose in the small organism bioassays (see below). These criteria avoided potential declines in chemical test concentrations during exposure. Exceptions were made, however, for the inclusion of some compounds with a specific MoA if representatives within the defined range of physicochemical properties were unavailable. The first triaging step excluded approximately 75% of the original nominees.

The second step of the triaging process focused on pragmatic logistic aspects of obtaining and shipping of the compounds: Availability, amounts, cost, purity, inter/national regulations, and human safety. After compiling logistic data including potential manufacturers/providers, price per gram, purity, UN numbers, transport toxicity and packaging groups, International Air Transport Association classification, and exempted amounts allowed for shipping, the following restrictions were implemented. Generally, only chemicals with purities $>98\%$ were considered. Based on the available experimental and theoretical toxicity information and the experimental design for each model species, calculations were performed to determine the amount of each chemical required. Depending on amounts needed and cost, economically non-affordable chemicals were omitted, and a limit of 1,000€ per chemical was imposed for practical reasons. Additionally, some chemicals were not commercially available nor in the required amounts or formats. Nominees were also discarded if they were classified as a prohibited/restricted chemical (e.g. potential for misuse or addiction, chemical warfare agents, explosives). Chemicals considered too toxic to be used in the laboratory without special training and precautions (e.g. gases, potent inhalable carcinogens, potent toxicants, etc.) were also discarded. A summary of the exclusion list criteria is presented in Fig. 1. The second level of triaging excluded an additional ~45% of the candidates from the first step.

Final selection and considerations

For several of the chemical groups, the first 2 triaging processes reduced the nominees to 20 to 30 compounds. In these cases, no further selection was necessary. In situations where a larger number of nominees remained, a third triaging step was introduced. Chemicals were selected based on additional toxicological parameters and their applicability in the identification of conserved toxicity pathways and biomarkers. These parameters included relevance to human exposure, in vivo and in vitro toxicities in multiple species, chemical transformation via metabolism, and molecular mechanism of toxicity. Information on these endpoints was collected by literature mining and from multiple databases (see Supplementary Material 2). Chemical identifiers including chemical name, CAS number, DSSTox substance identifier (DTXSID), or InChIKey were used in database searches and literature mining.

To maintain diversity in chemical space, only 1 or 2 chemicals that affect a specific mechanism of toxicity, Key Event, receptor, or Molecular Initiating Event per group were included. Finally, higher priorities for inclusion were given to chemicals identified in the human exposome (Barupal and Fiehn 2019; Neveu et al. 2020).

It was deemed crucial to minimize possible bias during the preparation of sample material for toxicity testing and omic analyses (PrecisionTox Consortium 2023). Therefore, each chemical used in the project was purchased from a single vendor and the same manufacturer's lot. To maintain consistency across the laboratories, chemicals were shipped from the manufacturer to a single location, where they were stored under controlled conditions. Based on the solubility and stability of the test compound, chemical aliquots were distributed as dry powder or as a DMSO stock solution (see Supplementary Material 1) to the toxicity testing groups in Europe, United Kingdom, and the United States to allow them to expose each corresponding model organism to the same chemical batch, including its possible impurities. As noted above, local and international shipping regulations/restrictions were also considered part of the triaging process and during the

aliquoting process (e.g. maximum amount per vial allowed by regulation) (Fig. 1).

We included 3 known baseline toxicity compounds—N-methylaniline, diphenylamine, and butoxyethanol—in the compound selection (Kluver et al. 2016). They served as negative controls for specific, reactive, or uncoupling interactions. Furthermore, compounds selected for specific organ toxicity can serve as negative controls for other organ-specific interactions.

Physicochemical properties

Physicochemical properties of each nominee were collected from public databases or calculated de novo using the methods and models described in [Supplementary Material 3](#) using the cognate data listed in [Supplementary Material 4](#). Physicochemical data were derived from various experimental and prediction sources potentially associated with uncertainty. Hence, for the interpretation of toxicity data derived from the nominees, the uncertainty and potentially a re-measurement of physicochemical data to avoid any bias in data interpretation have to be considered.

Henry coefficients (K_H) were estimated using OPERA version 2.6 or obtained from the CompTox Chemicals Dashboard (Williams et al. 2021; US Environmental Protection Agency 2023). They were then converted to air–water partition coefficients (K_{aw}) and corrected for ionization at a given pH to obtain the ionization-corrected air–water distribution ratio D_{aw} (pH).

Octanol–water partition constants (K_{ow}) were retrieved from experimental data in original publications or CompTox. If experimental data were not available, the K_{ow} was predicted using a poly-parameter linear solvation energy relationship (LSER) using the UFZ-LSER database (Ulrich et al. 2021). Alternatively, the mean K_{ow} was determined from the predicted values using KOWWIN v1.67 and ACD/Labs consensus data obtained from CompTox.

Experimental acidity constants (pK_a) were obtained from the literature or experimentally measured using a Sirius T3 titrator (Pion, Inc.) (Niu et al. 2022; Huchthausen et al. 2024). If experimental data were not available or the prediction indicated that speciation did not change in the pH range relevant for the bioassays, pK_a values were predicted with ACD pK_a /GALAS (ACD/Percepta 2015). Fractions of all neutral or zwitterionic ($=f_{neutral}$), negative, double negative, positive, or double positive species were calculated using the Henderson–Hasselbalch equation (Escher et al. 2020).

Prediction of in vitro and in vivo properties

Membrane–water partition constants

The uptake of chemicals into biological membranes is important in understanding bioavailability and baseline toxicity, the minimal toxicity that every chemical exhibits (Huchthausen et al. 2024; Qin et al. 2024). The distribution ratio between liposomes, which are typically used as proxies of biological membranes, and water ($D_{lip/w}$) is composed of the $K_{lip/w}$ of all species, including neutral and multiple charged species that can be measured or predicted with LSER or from the K_{ow} (Lee et al. 2021; Ulrich et al. 2021).

Protein–water partition constants

Bovine serum albumin (BSA) serves as surrogate for protein binding in medium. If experimental data were available, the distribution ratio between BSA and water, $D_{BSA/w}$ at pH 7 to 7.4, served as a proxy for protein binding. If experimental data were not available, the $K_{BSA/w}$ of the neutral species was calculated with a LSER or was predicted from $\log K_{ow}$ (Endo and Goss 2011). The $K_{BSA/w}$ of

cations was assumed to be the same as for neutral chemicals. Because anionic chemicals have a higher affinity to BSA, the prediction model for the non-specific portion of the sorption isotherm of anionic polyfluoroalkyl substances was used (Qin et al. 2024). The $K_{BSA/w}$ was fixed at 1.31 for hydrophilic chemicals ($\log K_{ow} < 2$) (deBruyn and Gobas 2007).

For chemicals where experimental data or LSER parameters were unavailable for cellular proteins, the $K_{SP/w}$ was predicted from K_{ow} (Endo et al. 2012). Anion binding to structural proteins (SPs) is weaker than to BSA, and equations analogous to those for $D_{BSA/w}$ (pH) were used to predict $D_{SP/w}$ (pH) (Henneberger et al. 2016).

In vitro distribution modeling to estimate in vitro chemical availability

Various in vitro chemical distribution models for cell-based bioassays exist (Proença et al. 2021). They typically describe processes of binding to medium components, plastic, and cells, as well as volatility. Volatile chemicals were not included in the PrecisionTox chemical collection. Media of the cellular bioassays are typically supplemented with 2% to 10% fetal bovine serum (FBS). Under such circumstances, binding to the plastic is negligible (Fischer et al. 2018). This is different for water-based media where dosing solutions may have to be renewed (depending on physicochemical properties) or concentrations measured (Fischer et al. 2018). The remaining processes are binding to medium components (proteins and lipids) and cells. The freely dissolved fraction f_{free} can be calculated with Equation 1

$$f_{free} = \frac{1}{1 + D_{medium/w} \times \frac{V_{protein+lipid, medium}}{V_{medium}} + D_{cell/w} \times \frac{V_{cell}}{V_{medium}}} \quad (1)$$

The partition constant between medium components and water $D_{medium/w}$ can be predicted by mass balance models assuming that proteins and lipids are the dominant sorptive phases. Their volume fractions of proteins ($VF_{protein, medium}$) and lipids ($VF_{lipid, medium}$) in the medium are considered in the mass balance equation (Equation 2). BSA is typically used as a surrogate for medium proteins.

$$D_{medium/w} = D_{BSA/w} \times VF_{protein, medium} + D_{lip/w} \times VF_{lipid, medium} \quad (2)$$

The $VF_{protein, medium}$ is the ratio of volume of proteins $V_{protein, medium}$ to the sum of the volume of both sorptive phases, proteins, and lipids, $V_{protein+lipid, medium}$ (Equation 3), and analogously for $VF_{lipid, medium}$ (Equation 4)

$$VF_{protein, medium} = \frac{V_{protein, medium}}{V_{protein+lipid, medium}} \quad (3)$$

$$VF_{lipid, medium} = \frac{V_{lipid, medium}}{V_{protein+lipid, medium}} \quad (4)$$

Analogously, $D_{cell/w}$ can be predicted by Equation 5. The most abundant proteins in cells are SPs, for which muscle proteins are better surrogates than BSA (Henneberger et al. 2016). In cells, the volume of water $V_{water, cell}$ is also considered.

$$D_{cell/w} = D_{SP/w} \times \frac{V_{protein, cell}}{V_{cell}} + D_{lip/w} \times \frac{V_{lipid, cell}}{V_{cell}} + \frac{V_{water, cell}}{V_{cell}} \quad (5)$$

Baseline toxicity

The nominal baseline cytotoxicity concentration for mammalian cell lines can be calculated from the mass balance model from the critical baseline-toxic membrane concentration $IC_{10, \text{membrane}}$ (Equation 6).

$$IC_{10, \text{nom, baseline}} = \frac{IC_{10, \text{membrane}}}{D_{\text{lip/w}}} \times \left(1 + D_{\text{BSA/w}} \times VF_{\text{protein, medium}} + D_{\text{lip/w}} \times VF_{\text{lipid, medium}} + D_{\text{cell/w}} \times \frac{V_{\text{cell}}}{V_{\text{bioassay}}} \right) \quad (6)$$

$IC_{10, \text{membrane}}$ varies little across species and cell types, but is dependent on the measured toxicity endpoint. The $IC_{10, \text{membrane}}$ for cell lines used in PrecisionTox, for a 10% reduction of cell proliferation as endpoint, was estimated to be approximately 26 mmol/L_{lip} (Huchthausen et al. 2024). For a medium supplemented with 10% FBS, the volume fraction of proteins in medium, $VF_{\text{protein, medium}}$, was 0.3% and of lipids, $VF_{\text{lipid, medium}}$ was 0.007%. The volume fraction of proteins in cells $VF_{\text{protein, cells}}$ was 3% and of lipids $VF_{\text{lipid, cell}}$ was 0.5%, which are means of typical experimental data (Huchthausen et al. 2024).

The freely dissolved chemical concentration $IC_{10, \text{free}}$ (Equation 7) can be used to compare the in vitro cytotoxicity with lethal effects in vivo.

$$IC_{10, \text{free}} = \frac{IC_{10, \text{membrane}}}{D_{\text{lip/w}}} \quad (7)$$

The baseline toxicity lethal concentration for 50% of the test organisms in vivo, LC_{50} , was predicted using linear regression QSAR models given in Equation 8.

$$-\log LC_{50} = a \log D_{\text{lip/w}} + b \quad (8)$$

The use of the $D_{\text{lip/w}}$ as a QSAR descriptor, rather than K_{ow} , allows one to also apply QSAR to ionizable compounds.

The baseline toxicity prediction model for *C. elegans* was rescaled from a K_{ow} -based model using of Equation 4 in Supplementary Material 3 and the descriptors are $a=0.81$ and $b=1.15$. The model for *D. magna* ($a=0.82$, $b=1.48$) was derived from the K_{ow} -based model using 48-h LC_{50} (Zhao et al. 1998). The QSAR for *D. rerio* was directly developed with $D_{\text{lip/w}}$ and has $a=0.99$ and $b=0.78$ (Kluver et al. 2019). There are no baseline toxicity prediction models for *X. laevis* and *Drosophila*. Therefore, experimental LC_{50} data for 22 chemicals in the PrecisionTox collection that were identified as baseline toxicants in both cell lines and zebrafish were used to derive new regressions. For both species, the baseline toxicants correlated linearly with the $\log D_{\text{lip/w}}$, yielding $a=0.61$ and $b=2.12$ ($r^2 = 0.690$) for *X. laevis* ($n=22$). For *Drosophila*, only the BMD_{50} (mol/L_{feed}) of 11 baseline toxicants were causing effects resulting in a regression with $a=0.83$ and $b=0.52$ ($r^2 = 0.724$). These QSARs are considered preliminary, but they serve to define the dosing concentrations in the PrecisionTox experiments.

ADME information

The values of in vitro measured intrinsic hepatic clearance (CL_{int}) and unbound fraction available in plasma (F_u) were retrieved from the Integrated Chemical Environment database version 4.1 (Bell et al. 2020). Human values have been reported where available. In the case of CL_{int} when human values were unavailable, rodent values were reported. Finally, OPERA-calculated values

were reported in the absence of experimental data (see Supplementary Material 5). This information enables further downstream biokinetic analyses to allow for species extrapolation and human-relevant estimations by employing physiologically based toxicokinetic modeling for reverse and forward dosimetry.

Data visualization

The PrecisionTox Data Visualization Tool (PDVT) is a web-based tool developed as a Flash application using Cytoscape.js (Franz et al. 2016) for network visualization (<https://dex.precisio tox.org/chem/>). The source code is freely available at: <https://github.com/precisio tox/chem-viz>.

Results

Selection of chemicals for the PrecisionTox collection

Utilizing the protocols for chemical nomination, triaging, and purchasing presented in Fig. 1, a collection of 200 chemicals was selected from >1,500 candidates for subsequent toxicological and omics analyses (see Supplementary Material 6). As an example of the process, a list of 188 unique DNTs was assembled from information collected by the NTP, ECVAM, and ASPIS. Physicochemical and toxicological data were collected for nominees from public databases (see Supplementary Material 2). This was a semi-automated process; therefore, all types of data were collected for all nominees, independent of their use in the final chemical selection.

Some nominees were deselected based on physicochemical properties. For example, trichloroethylene has an air–water partition constant of $D_{aw} = 0.40$, which is greater than the 10^{-4} cut-off; PBDE 47 with a $K_{ow} = 6.8$ is too hydrophobic for testing in our systems ($K_{ow} > 4$). Information on availability, cost, and other shipping restrictions for the remaining candidates were then collected and used to deselect from the remaining candidates. Several of the nominees are restricted substances (e.g. cocaine, heroin) or too toxic to use under normal laboratory conditions (e.g. methylmercury). Additional nominees were removed based on the calculated amounts needed for testing in the PrecisionTox model systems and associated costs. For example, thapsigargin, domoic acid, and thalidomide would have cost €78,000, €19,000, and €1,400, respectively, making their incorporation into the collection problematic. For instance, the triaging protocol outline in Fig. 1 reduced the original 188 DNT nominees to 30 chemicals, which were included in the collection.

Characteristics of the PrecisionTox chemical collection

The chemicals in this collection, including identifiers (CAS number, DSSTox substance ID, SMILES and InChIKey) and Toxicity Groups are presented in Supplementary Material 6. Additionally, Chemical and Products Database (CPDat) use categories are included (Dionisio et al. 2018; US Environmental Protection Agency 2025). CPDat is a database maintained by the US EPA that catalogues commercially sourced chemicals, using standard product categories, based on how they are used. The value of providing use categories is to avoid a bias by a focus on a certain group of chemicals. For example, industrial chemicals are typically not designed for biological activity. Therefore, a focus on such chemicals could limit the applicability domain of NAMs developed with a focus on such chemicals.

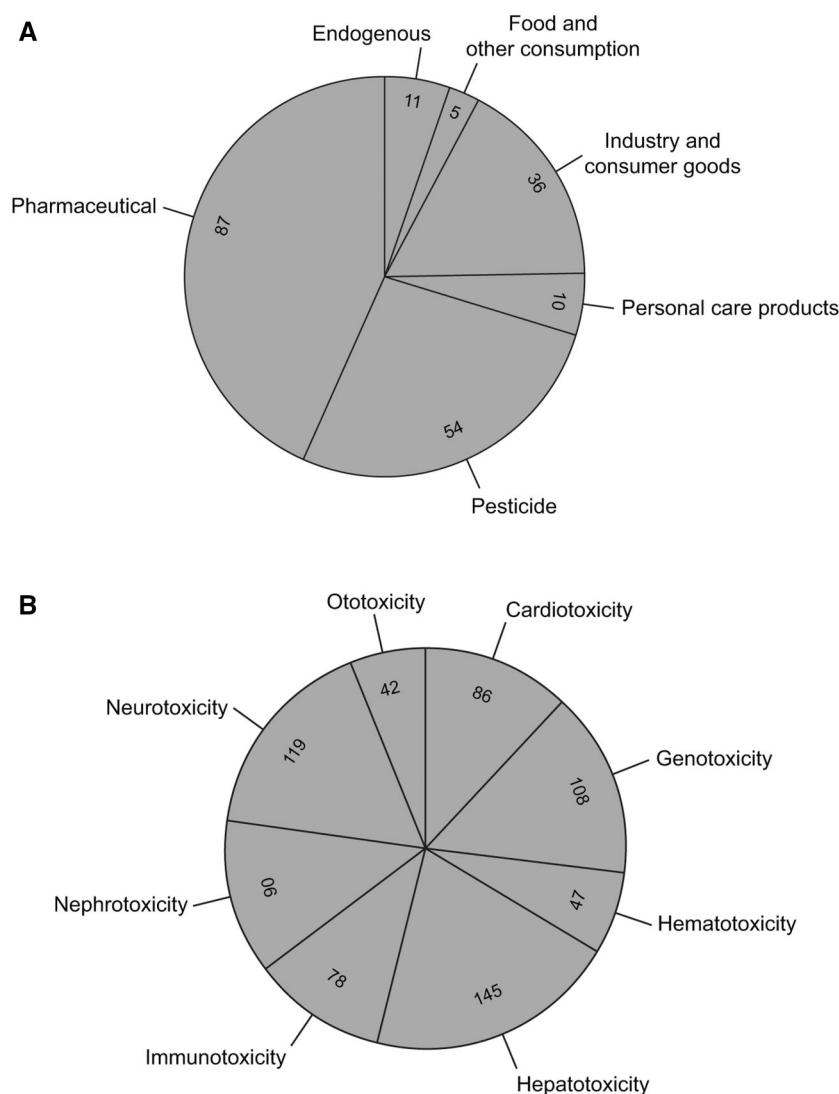


Fig. 2. Distribution of chemicals in toxicity and use categories. Chemicals were assigned to toxicity A) or use B) categories. Values in the pie sections are the number of chemicals in that category. Use and toxicity values for each chemical can be found in [Supplementary Materials 6 and 7](#).

Chemicals in the PrecisionTox collection are associated with 300 CPDat use categories, with the majority identified as drugs, industrial manufacturing compounds, pesticides, and food additives. Additionally, they mapped to 1 of 6 use categories that are associated with chemicals found in human plasma ([Braun et al. 2024](#)) ([Fig. 2A](#)).

Toxicology properties

Toxicological properties, including toxicity category and baseline toxicity predictions for all test animal species and Adverse Outcome Pathways (AOPs) of the chemical collection, are presented in [Supplementary Material 7](#). AOPs, Key Events, and Molecular Initiating Events associated with each chemical were identified by searching the EPA's Adverse Outcome Pathway Database (AOP-DB) ([Mortensen et al. 2021](#)). The AOP-DB was searched with the DTXSIDs of the PrecisionTox chemicals. Of the 200 chemicals in the collection, all except 24 are assigned to 1 or more of 8 toxicity categories ([Fig. 2B](#)). The majority of the chemicals are assigned to multiple categories, with hepatotoxicants and neurotoxicants being the most populated. Of the 200 chemicals, 87 were associated with 1 or more unique AOPs ([Fig. 3](#)) (see [Supplementary Material 7](#)).

Baseline toxicity predictions

The predicted baseline toxicity represents the minimal toxicity associated with any chemical ([Qin et al. 2024](#)). It can be used for planning of the dosing concentrations as well as the interpretation of the data. The specificity ratio, the ratio of the predicted IC_{10} or LC_{50} to the corresponding experimental data, indicates which chemicals are toxic because they are hydrophobic and have a high baseline toxicity or if they are reactive or have specific targets. QSAR predictions for the PrecisionTox in vivo models varied within a factor of 100 per chemical ([Fig. 4](#); see [Supplementary Material 7](#)). The predicted in vitro $IC_{10(free)}$ was similar to the predicted in vivo LC_{50} . In general, hydrophobic chemicals were less sensitive in vivo, and hydrophilic chemicals have a similar sensitivity.

ADME information

To provide additional information on the chemical collection, potential hepatic clearance rates were calculated. Results show intrinsic hepatic clearance values above $10 \mu\text{l}/\text{min}/10^6$ cells for 75 compounds. Four substances have rapid clearance values above $100 \mu\text{l}/\text{min}/10^6$ cells, as determined by human in vitro assays: Propofol (CAS No. 2078-54-8), methylparaben (CAS No. 99-76-3),

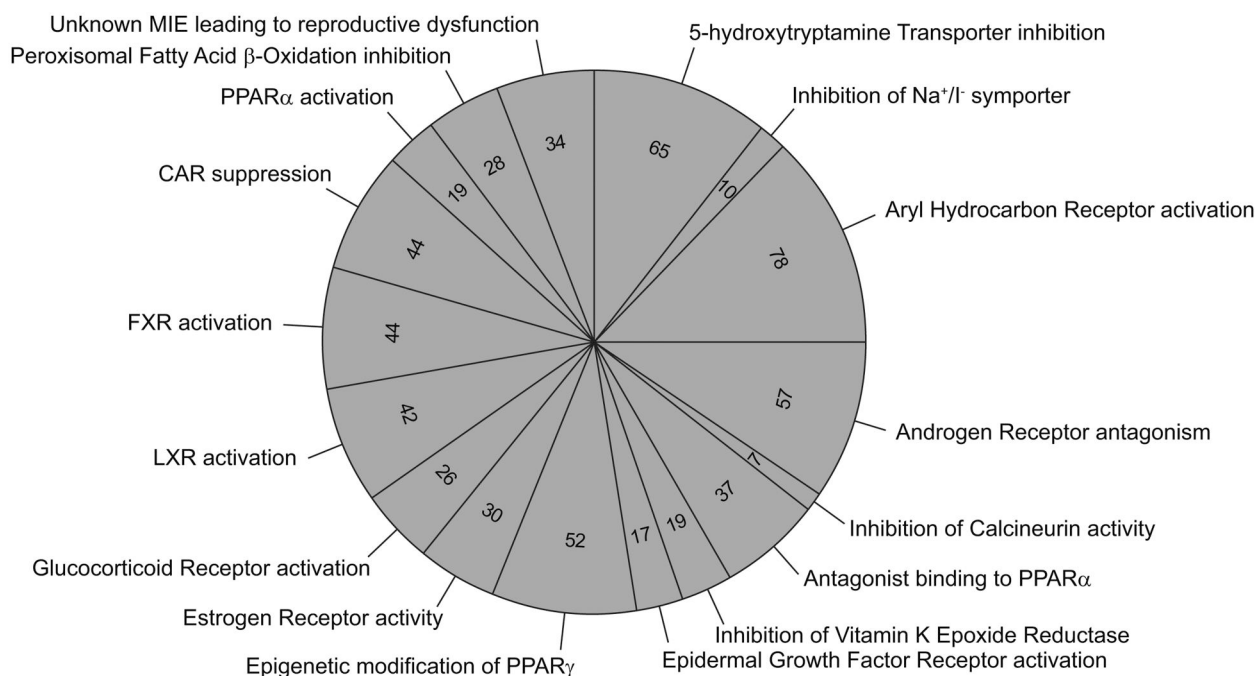


Fig. 3. Distribution of chemicals in adverse outcome pathways. Chemicals were mapped to AOPs using the EPA AOP-DB and the number of chemicals in each AOP presented. To simplify the figure and minimize redundancy, several AOPs were combined into a single pathway. For example, 5-hydroxytryptamine transporter (5-HTT) inhibition leading to population increase (AOP ID195); 5-HTT inhibition leading to decreased shelter seeking and increased predation (AOP ID98); 5-HTT inhibition leading to population decline (AOP ID97); 5-HTT inhibition leading to decreased reproductive success and population decline (AOP ID203); and 5-HTT inhibition leading to increased reproductive success and population increase (AOP ID204) were combined to the group labeled “5-hydroxytryptamine transporter inhibition.” Values in the sections are the number of chemicals in that category. AOP information can be found in [Supplementary Material 7](#).

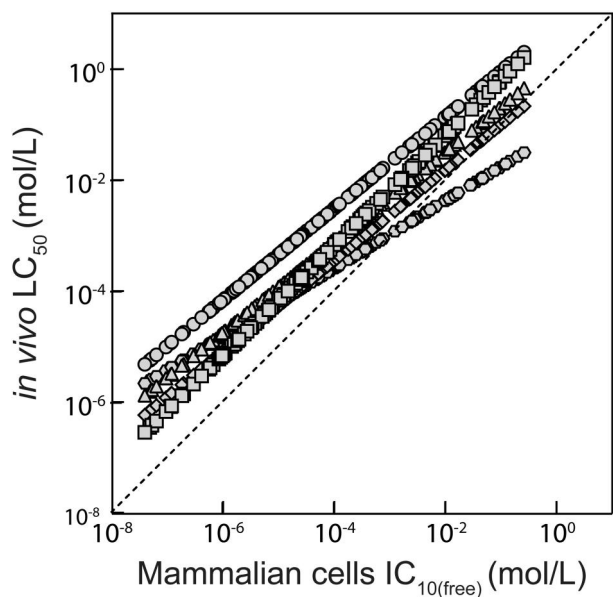


Fig. 4. In vitro–in vivo comparison between the baseline toxicity predictions for mammalian cell cytotoxicity $IC_{10(free)}$ normalized to freely dissolved concentrations (in vitro) and in vivo LC_{50} values for *D. melanogaster* (circle), *D. rerio* (square), *C. elegans* (triangle), *D. magna* (diamond), and *X. laevis* (hexagon).

bisphenol E (CAS No. 2081-08-5), and chlorpromazine (CAS No. 69-09-0), with clearance rates of 107, 107.5, 119, and 195.1 $\mu\text{L}/\text{min}/10^6$ cells, respectively. Based on experimental in vitro human and rodent data, approximately one-third of the chemicals in the PrecisionTox collection have clearance values below

2.5 $\mu\text{L}/\text{min}/10^6$ cells. This observation indicates a very slow rate of metabolism in mammalian test systems. An IA clearance value of 0 indicates little to no metabolism, a potential for bioaccumulation, or a lack of data. This highlights the need for a thorough review of existing data and potentially additional testing to better understand the metabolic fate of these chemicals (Gouliarmou et al. 2018). Human plasma protein-unbound fraction values for 133 compounds were $\geq 50\%$, indicating a high availability of the free chemical in plasma. This could lead to an increased potential for interaction with molecular targets, including enzymes or receptors. Additionally, metals and inorganic compounds are out of the model's applicability domain, and data for tunicamycin, cadmium chloride, sodium arsenite, tributyltin, triethyltin bromide, sodium perchlorate, and lithium (cation) were unavailable (see [Supplementary Material 5](#)).

Physicochemical properties

The K_{ow} of the neutral species range from -4.63 to 8.50 , covering 12 orders of magnitude with a mean of 2.31 (Fig. 5A). Over 80 of the chemicals are 75% to 100% in neutral form, almost 60 are more than 80% charged, with the remainder having multiple charges at pH 7.4 (Fig. 5B). The $D_{lip/w}$ (pH 7.4) of the collection ranges from -1 to 8.5 , with more than 80% of the compounds between -1 and 4 (Fig. 5C). The lowest value was set to -1 , because according to Gobas et al. (1988), that is the empirical lowest distribution ratio. Additionally, the QSAR is not valid if the affinity to membrane bilayers is too low. We also predicted binding to plasma (albumin) and SPs. As most were derived from the log K_{ow} or $D_{lip/w}$, the distributions look similar to Fig. 5C (see [Supplementary Material 4](#)).

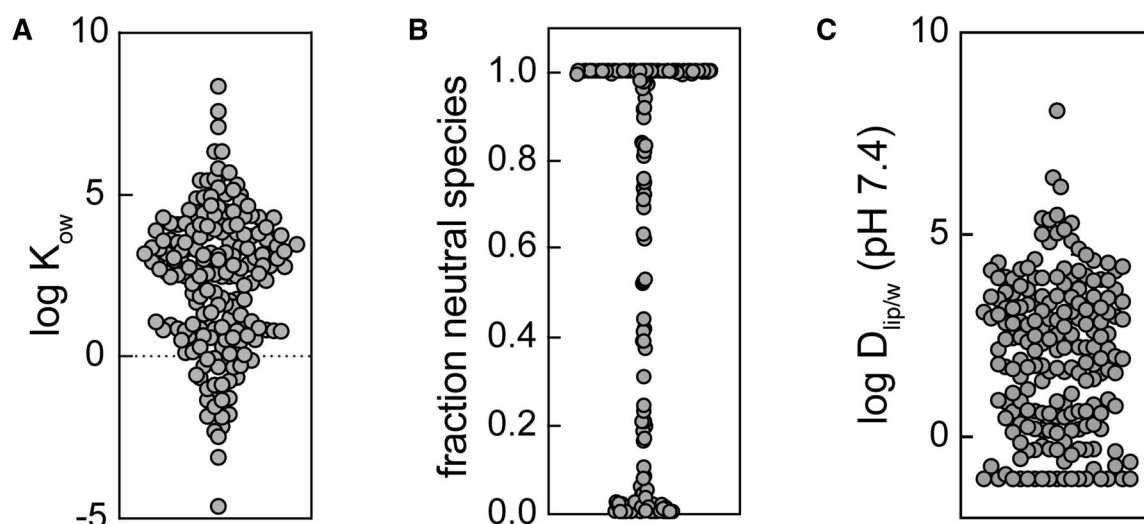


Fig. 5. Physicochemical properties of the PrecisionTox chemical collection. A) Distribution of octanol–water partition constants K_{ow} of the neutral species of all test chemicals. B) Distribution of fractions of neutral species of all test chemicals derived from the acidity constants pK_a . C) Distribution of the ionization-corrected liposome water distribution ratio ($D_{lip/w}$) of all species at a pH 7.4. Detailed data can be found in [Supplementary Material 4](#).

In vitro distribution modeling to estimate availability

The fraction of freely dissolved chemical in the bioassay medium (f_{free}) decreases with increasing hydrophobicity (Fig. 6). Chemicals with $\log D_{lip/w} < 3$ are typically fully dissolved, and those with $\log D_{lip/w} > 5$ are almost completely bound. Fulvestrant (CAS No. 129453-61-8) is extremely hydrophobic, $\log D_{lip/w} > 8$, which is outside the domain of the selection criteria, so the prediction should be treated with caution.

As the serum protein binding dominates the reduction of f_{free} , the relationship between f_{free} and $D_{BSA/w}$ is much more uniform than the relationship between f_{free} and $D_{lip/w}$ (Fig. 6). This is because $\log D_{lip/w}$ and $D_{BSA/w}$ are only proportional for neutral and positively charged chemicals, but anions bind stronger to proteins than to lipids.

In a medium that is supplemented with 10% FBS, the fraction taken up into the cells is negligible, and as the freely dissolved concentrations are low for the very hydrophobic chemicals, the binding to the plastic plates is also negligible. In addition, many test chemicals are partially or fully charged, and charged chemicals do not diffuse into the polymers. However, many cell culture plates are treated to make cells better adherent, and the tissue culture treatment makes the plastic surfaces anionic, allowing cations to bind. Unfortunately, there are no binding constants available yet for treated plate materials.

Uptake into cells and zebrafish embryos can be predicted reliably using established models (Grasse et al. 2024; Proença et al. 2021), but for all other assays, such prediction models do not exist because baseline toxicity QSARs are based only on empirical correlations with $D_{lip/w}$ and not on a mechanistic disposition model.

PrecisionTox data visualization tool

The PDVT is a web-based application developed to facilitate the exploration and analysis of the PrecisionTox chemical collection: <https://dex.precisiontox.org/chem>. Currently, the database contains all of the physicochemical and toxicological data reported in this manuscript and [supplemental materials](#). The database and PDVT are open-ended, and in the future will include novel experimental data generated by PrecisionTox.

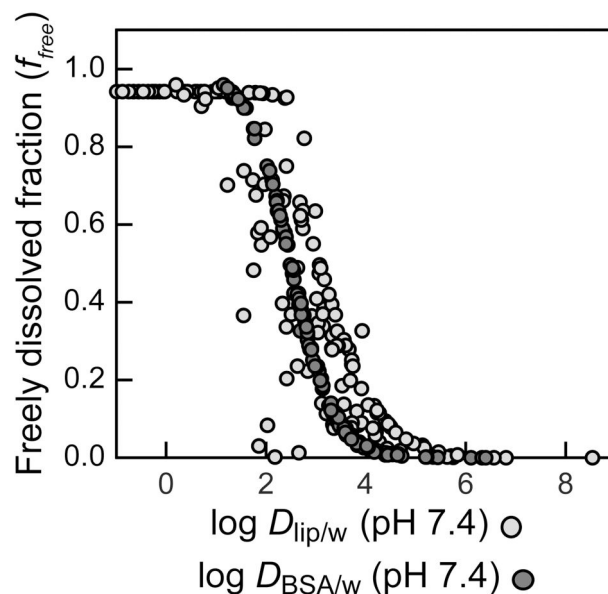


Fig. 6. Freely dissolved fraction f_{free} in the bioassay medium (Equation 1) as a function of liposome–water distribution ratio ($D_{lip/w}$ [pH 7.4]) (light gray dots) or protein–water distribution ratio ($D_{BSA/w}$ [pH 7.4]) (dark gray dots).

The PDVT features an interactive network that allows users to visualize how chemicals are classified according to 2 main categories: Use and Toxicity Endpoint (Fig. 2). PDVT has the ability to toggle between these classification systems. The differently colored nodes represent specific subcategories within each classification. For example, under the Use classification, nodes include *pesticide*, *pharmaceutical*, *industry*, *consumer goods*, and others (Fig. 7A). Each node can be clicked to reveal all chemicals associated with that particular subcategory, which will be shown connected to it via edges (Fig. 7B). These chemicals are simultaneously displayed in a table below the graph, providing a clear and organized view of the corresponding entries (Fig. 7C).

The web application has a search functionality that allows users to focus on individual chemicals. The single chemical view

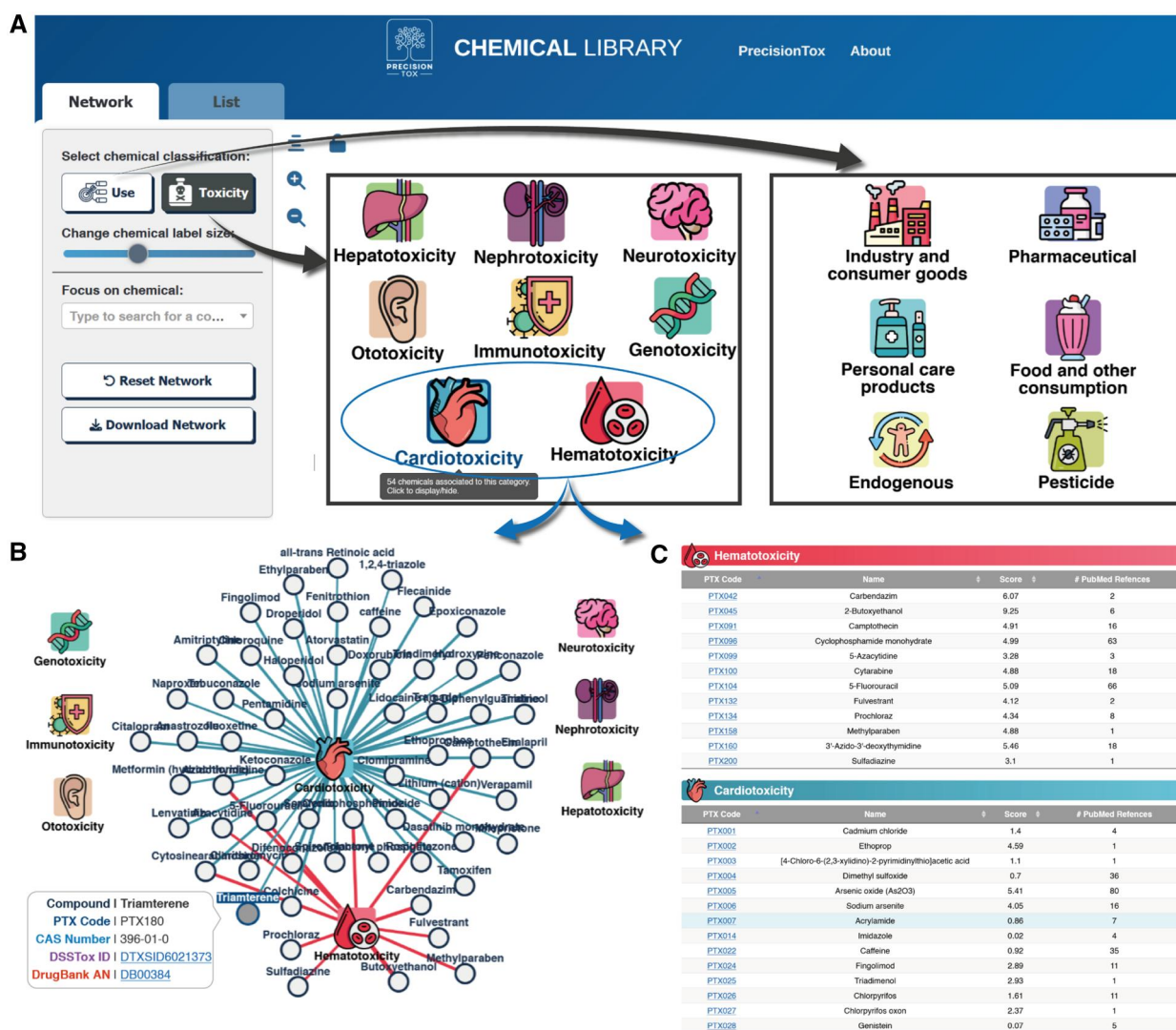


Fig. 7. General view on PDVT. A) An interactive diagram shows the different categories for chemical classification according to 2 different classification systems: “use” and “toxicity endpoint,” which the user can toggle between. B) Clicking on 1 or more of these nodes will display all the chemicals included in those categories as additional child nodes. Chemical can be associated to more than 1 category, which is represented by differently colored edges. C) In addition to the network, dynamically changing table will show the full list of chemicals associated to the last clicked category (<https://dex.precisioxtox.org/chem>).

page comprises a new network graph that displays all collected information for that compound. This includes its use and toxic endpoint categories, physicochemical properties, mechanisms of action, baseline toxicity data, molecular targets, and adverse outcome pathways. Each of these main attributes is represented as a node that can be clicked to display all specific data as additional child nodes (Fig. 8A). For example, physicochemical properties comprise up to 14 different attributes, including molecular weight, density, Henry coefficient, etc., whereas mechanisms of action contain relevant available annotations obtained from DrugBank and T3DB. For further convenience, chemical information is also displayed in a structured table format beneath the graph (Fig. 8B). A snapshot of the current network view can be obtained at any time in image or vector graphics formats. Finally, the web application includes a comprehensive table listing the entire chemical collection. This table consolidates all available data for each chemical into clearly defined columns, encompassing physicochemical properties, use, toxicity classifications, mechanisms of action, and other relevant information. Users can

filter and sort the table as needed to support targeted analysis, as well as download it in different formats.

Discussion

Over the past several decades, the ability to screen large numbers of chemicals for their biological effects has become an integral part of the discovery of new drugs, pesticides, and other commercial products, as well as toxicity testing. Additionally, HTS and associated data analysis workflows have contributed to the development of novel test methods. This has been accomplished by assessing their biological relevance via the interrogation of large numbers of chemicals to identify applicability domains and potential limitations, such as chemical space coverage and/or the range of expected responses. High-throughput screens also assess the reliability of a test method to increase reproducibility and minimize variability.

The starting point for the utilization of any HTS is the creation of a chemical library. There are 2 basic components in the

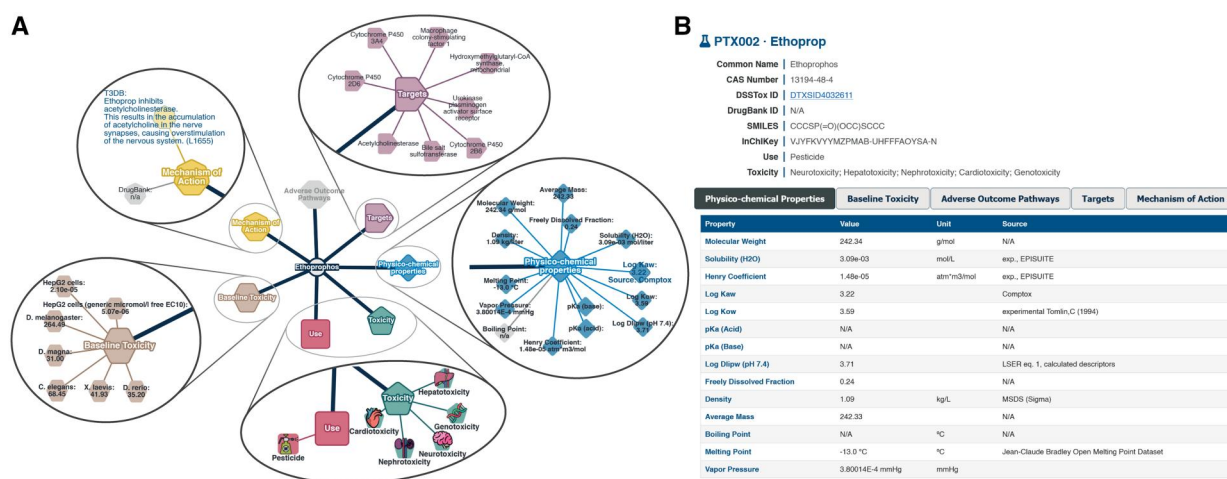


Fig. 8. Single chemical view on PDVT. A) The network shows the selected chemical surrounded by nodes that represent the collected chemical information categories. Each can be expanded individually by clicking to show additional nodes displaying the associated data. B) All this information is also displayed below in a more typical knowledge base format including the diverse chemical identifiers and a table where users can toggle between the different types of annotations.

creation of a chemical library. First is to define the parameters or chemical space of the collection: Targets for a specific protein/receptor, pathway/mechanism of response, disease/health endpoint, or to test a variety of responses (Richard et al. 2021). Multiple approaches have been used to create chemical libraries, which include synthesizing congeners of biologically active compounds (Sahoo et al. 2022), designing virtual libraries based on QSAR (Nowak et al. 2024), collecting chemicals with known responses (Sund and Deceuninck 2021), and pooling compounds of interest selected by various parties (Attene-Ramos et al. 2013). The second component in the creation of a chemical library are logistic and practical considerations: Can a compound be purchased or synthesized in the required amount, at the necessary purity, and at an acceptable cost? Also, can the chemical be tested in the selected biological system under available laboratory conditions?

Here, we report on the creation of a unique, fit-for-purpose chemical library designed to test the phylotoxology hypotheses (Colbourne et al. 2022; PrecisionTox Consortium 2023). In contrast to other libraries, the PrecisionTox chemical collection was not created to target a specific disease, protein, or pathway. Rather, diversity was at the center of its design. To select and subsequently test this chemical collection, a robust chemical identification and logistics protocol was developed (Fig. 1). From >1,500 nominees, 200 compounds were selected, purchased, and distributed to PrecisionTox testing laboratories in the EU, the United Kingdom, and the United States. Nominations were collected to reflect compounds with a variety of potential MoA. Problematic nominees were removed based on physicochemical properties obtained from public databases or calculated de novo. Practical considerations (cost, purity, availability, and the ability to send the compounds to the international PrecisionTox testing laboratories) were also important deciding factors. Ultimately, diverse representation throughout chemical space determined a nominee's inclusion in the collection.

With a few exceptions, the protocol functioned as anticipated. One problem occurred due to an unforeseen chemical property. N-(hydroxymethyl)acrylamide (CAS No. 924-42-5) unexpectedly oligomerized in solution and therefore was removed from testing. Logistical challenges related to regulatory restrictions affected the inclusion of pentobarbital (CAS No. 76-74-4), which is subject

to export prohibitions due to its potential use in certain jurisdictions. As a result, it was also removed from the collection to comply with applicable laws and regulations.

Based on the selection criteria, the final PrecisionTox collection is diverse, but due to the limited number of chemicals that could be tested, it is not inclusive. Chemicals in the collection target several of the organs typically associated with human toxicity, including the liver, kidney, heart, and the nervous and immune systems (Table 1, Fig. 2B). Chemicals that target the lung, skin, and the visual apparatus were excluded during selection. As each of these organ targets are not present in several or all of the PrecisionTox model species, they were not suitable for testing the phylotoxology hypothesis.

This chemical collection is biased toward pharmaceuticals and pesticides (Fig. 2A). This is consistent with the nomination process and selection protocol, as large amounts of physicochemical, toxicological, and exposure data are available for chemicals in these categories. These use categories typically comprise compounds with a specific MoA, which make them suitable as reference compounds. These reference compounds can be used to define MoA, mechanism of toxicity, biomarkers, and ethology for chemicals. In contrast, non-specifically acting compounds (e.g. with toxicity driven by hydrophobic interaction with cellular membranes) would result in effect concentration close to QSARs predictions (e.g. Huchthausen et al. 2023, 2024) and similar effect concentrations across phylogenetically distant groups. Hence, for the development of alternative approaches for the screening of chemicals based on NAMs and reduced reliance on animal tests (Sewell et al. 2024), particularly methods capturing specific modes of interactions (e.g. binding to specific target molecules) are needed. This is also a prerequisite to use the biological effect patterns in a read-across approach beyond the similarity of chemical structure (Escher et al. 2019b).

The retrieval of information on the physicochemical properties and baseline toxicity predictions helped the testing laboratories handle such a diverse set of chemicals and dose them appropriately. A large number of the chemicals have a specific MoA. Three compounds that were expected to show no specific MoA were intentionally included as reference compounds to demonstrate that experimental effect concentrations for these compounds would be close to QSAR predictions. It can be

expected that, specific to the *in vivo* model, more compounds will exhibit effect concentrations close to QSAR prediction, e.g. when the specific target may not be available in the corresponding model. Additionally, compounds included due to their environmental occurrence and to obtain missing toxicity information also may represent non-specifically acting compounds. By comparing the experimentally observed effects (i.e. cytotoxicity, mortality, mechanism-specific endpoints) to the predicted baseline toxicity, it is possible to quantify and estimate the reliability to detect specific effects of NAMs and benchmark them with respect to traditional animal-test-based approaches.

It should be noted that compounds with difficult physicochemical properties (volatility, hydrophobicity) were excluded for practical reasons. Such compounds require special exposure regimes. Excluding such compounds increases the throughput for benchmarking the approach(es) targeted with our library. Given the availability of special testing routines for difficult compounds, however, this does not limit *per se* the application domain for these compounds. Testing of volatile or hydrophobic chemicals is a regulatory requirement (see OECD 2019). Testing these compound in the context of PrecisionTox, however, would require additional efforts that would reduce the number of compounds included in the chemical collection to develop NAMs. It is therefore advisable to avoid these compounds during NAM development. Using appropriate guidance on testing difficult compounds, they can, however, be tested using developed NAMs.

One issue that needs to be taken into account when building a chemical library is the amount of available data for any individual compound. Experimental and theoretical data are readily available on the physicochemical properties through PubChem, CompTox, and other databases (see Supplementary Material 2). Where data are missing from public databases, *de novo* calculations can be used to fill these gaps. All of the chemical nominees for this collection had sufficient physicochemical data for triaging. Additionally, the selected compounds are also readily available at high purities, low cost, and international shipping regulations are well defined.

In vivo and *in vitro* toxicity, AOP, exposure, and toxicity mechanism data are generally not as complete. For example, of the 200 chemicals in the PrecisionTox collection, only 87 have been assigned to an AOP; less than half of the collection are found in DrugBank, and <50 have human exposure information. This paucity of toxicological information reflects the need for additional data and justifies its inclusion in this multipurpose library for toxicity testing. The ADME *in vitro* parameters, CL_{int} and F_u (human and rat), were extracted from the Integrated Chemical Environment database only after the chemical selection process had been completed, with the purpose of enabling further downstream analysis, such as PBK modeling to estimate effective dose-extrapolation assumptions to humans.

To summarize, from over 1,500 nominees, a collection of 200 chemicals was obtained for the PrecisionTox Consortium with diversity in chemical space as its core principle. This collection is a well-characterized set of compounds that will be used to test the phylotoxicology hypothesis. Selections are based on physicochemical and toxicological properties, as well as practical considerations. This report highlights the need for strategic selection criteria for any chemical library development, ensuring the appropriate level of coverage within chemical space while balancing practical constraints.

Acknowledgments

We want to acknowledge Patricia Solórzano (Leitat) for her involvement in chemical logistics. Some of the chemicals for this work were provided by the Division of Translational Toxicology, National Institute of Environmental Health Sciences, National Institutes of Health, U.S. Department of Health and Human Services via contract HHSN273201400020C (MRI Global, Kansas City, MO).

Author contributions

John K. Colbourne (Conceptualization, Funding Acquisition), Beate I. Escher (Methodology, Resources, Formal Analysis, Data Curation, Visualization, Writing—Review & Editing), Thomas Dickmeis (Conceptualization, Validation, Investigation, Data Curation, Writing—Review & Editing), Jonathan H. Freedman (Conceptualization, Resources, Methodology, Writing—Original Draft, Writing—Review & Editing, Supervision, Project Administration), Juan Carlos González-Sánchez (Software, Visualization, Writing—Original Draft, Writing—Review & Editing), Luise Henneberger (Formal Analysis, Writing—Review & Editing), Julia Huchthausen (Formal Analysis, Writing—Review & Editing), Rubén Martínez (Methodology, Resources, Data Curation, Writing—Review & Editing, Visualization), Stavroula I. Sampani (Conceptualization, Formal Analysis, Investigation, Data Curation, Writing—Original Draft, Writing—Review & Editing), Stefan Scholz (Software, Formal Analysis, Data Curation, Writing, Review & Editing), Carsten Weiss (Conceptualization, Writing—Review & Editing), and Maurice Whelan (Conceptualization, Writing—Review & Editing)

Supplementary material

Supplementary material is available at Toxicological Sciences online.

Funding

This project has received funding from the European Union's Horizon 2020 Research and Innovation Program under Grant Agreement No 965406. The work presented in this publication was performed as part of ASPIS. This output reflects only the authors' view, and the European Union cannot be held responsible for any use that may be made of the information contained therein.

Conflicts of interest. None declared.

References

- ACD/Percepta. 2015. Build 2726. Toronto (ON): Advanced Chemistry Development, Inc. [accessed 2024 Jun]. <http://www.acdlabs.com>.
- Albrecht W, Kappenberg F, Brecklinghaus T, Stoeber R, Marchan R, Zhang M, Ebbert K, Kirschner H, Grinberg M, Leist M, et al. 2019. Prediction of human drug-induced liver injury (DILI) in relation to oral doses and blood concentrations. *Arch Toxicol*. 93:1609–1637.
- Aschner M, Ceccatelli S, Daneshian M, Fritsche E, Hasiwa N, Hartung T, Hogberg HT, Leist M, Li A, Mundi WR, et al. 2017. Reference compounds for alternative test methods to indicate developmental neurotoxicity (DNT) potential of chemicals: example lists and criteria for their selection and use. *ALTEX*. 34:49–74.

- Ashenden SK. 2018. Screening library design. *Methods Enzymol.* 610:73–96.
- Attene-Ramos MS, Miller N, Huang R, Michael S, Itkin M, Kavlock RJ, Austin CP, Shinn P, Simeonov A, Tice RR, et al. 2013. The Tox21 robotic platform for the assessment of environmental chemicals—from vision to reality. *Drug Discov Today.* 18:716–723.
- Barupal DK, Fiehn O. 2019. Generating the blood exposome database using a comprehensive text mining and database fusion approach. *Environ Health Perspect.* 127:97008.
- Bell S, Abedini J, Ceger P, Chang X, Cook B, Karmaus AL, Lea I, Mansouri K, Phillips J, McAfee E, et al. 2020. An integrated chemical environment with tools for chemical safety testing. *Toxicol In Vitro.* 67:104916.
- Blum J, Masjosthusmann S, Bartmann K, Bendt F, Dolde X, Dönmez A, Förster N, Holzer A-K, Hübenthal U, Keßel HE, et al. 2023. Establishment of a human cell-based in vitro battery to assess developmental neurotoxicity hazard of chemicals. *Chemosphere.* 311:137035.
- Braun G, Herberth G, Krauss M, König M, Wojtysiak N, Zenclussen AC, Escher BI. 2024. Neurotoxic mixture effects of chemicals extracted from blood of pregnant women. *Science.* 386:301–309.
- Canham SM, Wang Y, Cornett A, Auld DS, Baeschlin DK, Patoor M, Skaanderup PR, Honda A, Llamas L, Wendel G, et al. 2020. Systematic chemogenetic library assembly. *Cell Chem Biol.* 27:1124–1129.
- Colbourne JK, Escher SE, Lee R, Vinken M, van de Water B, Freedman JH. 2025. Animal-free Safety Assessment of Chemicals: Project Cluster for Implementation of Novel Strategies (ASPIS) definition of new approach methodologies. *Environ Toxicol Chem.* 44:2365–2397. <https://doi.org/10.1093/etojnl/vgae093>
- Colbourne JK, Shaw JR, Sostare E, Rivetti C, Derelle R, Barnett R, Campos B, LaLone C, Viant MR, Hodges G. 2022. Toxicity by descent: a comparative approach for chemical hazard assessment. *Environ Adv.* 9:100287.
- De Castelbajac T, Aiello K, Arenas CG, Svingen T, Ramhøj L, Zalko D, Barouki R, Vanhaecke T, Rogiers V, Audebert M, et al. 2023. Innovative tools and methods for toxicity testing within PARC work package 5 on hazard assessment. *Front Toxicol.* 5:1216369.
- deBruyn AMH, Gobas FAPC. 2007. The sorptive capacity of animal protein. *Environ Sci Technol.* 26:1803–1808.
- Dionisio KL, Phillips K, Price PS, Grulke CM, Williams A, Biryol D, Hong T, Isaacs KK. 2018. The chemical and products database, a resource for exposure-relevant data on chemicals in consumer products. *Sci Data.* 5:180125.
- Endo S, Bauerfeind J, Goss K-U. 2012. Partitioning of neutral organic compounds to structural proteins. *Environ Sci Technol.* 46:12697–12703.
- Endo S, Goss KU. 2011. Serum albumin binding of structurally diverse neutral organic compounds: data and models. *Chem Res Toxicol.* 24:2293–2301.
- Escher B, Abagyan R, Embry M, Klüver N, Redman A, Zarfl C, Parkerton T. 2020. Recommendations for improving methods and models for aquatic hazard assessment of ionizable organic chemicals. *Environ Toxicol Chem.* 39:269–286.
- Escher BI, Glauch L, König M, Mayer P, Schlichting R. 2019a. Baseline toxicity and volatility cutoff in reporter gene assays used for high-throughput screening. *Chem Res Toxicol.* 32:1646–1655.
- Escher SE, Kamp H, Bennekou SH, Bitsch A, Fisher C, Graepel R, Hengstler JG, Herzler M, Knight D, Leist M, et al. 2019b. Towards grouping concepts based on new approach methodologies in chemical hazard assessment: the read-across approach of the EU-ToxRisk project. *Arch Toxicol.* 93:3643–3667.
- Finckh S, Beckers LM, Busch W, Carmona E, Dulio V, Kramer L, Krauss M, Posthuma L, Schulze T, Slootweg J, et al. 2022. A risk based assessment approach for chemical mixtures from wastewater treatment plant effluents. *Environ Int.* 164:107234.
- Finckh S, Carmona E, Borchardt D, Büttner O, Krauss M, Schulze T, Yang S, Brack W. 2024. Mapping chemical footprints of organic micropollutants in European streams. *Environ Int.* 183:108371.
- Fischer FC, Cirpka O, Goss KU, Henneberger L, Escher BI. 2018. Application of experimental polystyrene partition constants and diffusion coefficients to predict the sorption of organic chemicals to well plates in in vitro bioassays. *Environ Sci Technol.* 52:13511–13522.
- Franz M, Lopes CT, Huck G, Dong Y, Sumer O, Bader GD. 2016. Cytoscape.js: a graph theory library for visualisation and analysis. *Bioinformatics.* 32:309–311.
- Gobas FA, Lahittete JM, Garofalo G, Shiu WY, Mackay D. 1988. A novel method for measuring membrane-water partition coefficients of hydrophobic organic chemicals: comparison with 1-octanol-water partitioning. *J Pharm Sci.* 77:265–272.
- Gouliarmou V, Lostia AM, Coecke S, Bernasconi C, Bessems J, Dorne JL, Ferguson S, Testai E, Remy UG, Brian Houston J, et al. 2018. Establishing a systematic framework to characterise in vitro methods for human hepatic metabolic clearance. *Toxicol In Vitro.* 53:233–244.
- Grasse N, Massei R, Seiwert B, Scholz S, Escher BI, Reemtsma T, Fu Q. 2024. Impact of biotransformation on internal concentrations and specificity classification of organic chemicals in the zebrafish embryo (*Danio rerio*). *Environ Sci Technol.* 58:17898–17907.
- Henneberger L, Goss KU, Endo S. 2016. Partitioning of organic ions to muscle protein: experimental data, modeling, and implications for in vivo distribution of organic ions. *Environ Sci Technol.* 50:7029–7036.
- Huchthausen J, Braasch J, Escher BI, König M, Henneberger L. 2024. Effects of chemicals in reporter gene bioassays with different metabolic activities compared to baseline toxicity. *Chem Res Toxicol.* 37:744–756.
- Huchthausen J, Escher BI, Grasse N, König M, Beil S, Henneberger L. 2023. Reactivity of acrylamides causes cytotoxicity and activates oxidative stress response. *Chem Res Toxicol.* 36:1374–1385.
- Klüber N, Bittermann K, Escher BI. 2019. QSAR for baseline toxicity and classification of specific modes of action of ionizable organic chemicals in the zebrafish embryo toxicity test. *Aquat Toxicol.* 207:110–119.
- Klüber N, Vogs C, Altenburger R, Escher BI, Scholz S. 2016. Development of a general baseline toxicity QSAR model for the fish embryo acute toxicity test. *Chemosphere.* 164:164–173.
- Koszucka A, Nowak A, Nowak I, Motyl I. 2020. Acrylamide in human diet, its metabolism, toxicity, inactivation and the associated European Union legal regulations in food industry. *Crit Rev Food Sci Nutr.* 60:1677–1692.
- Lee J, Braun G, Henneberger L, König M, Schlichting R, Scholz S, Escher BI. 2021. Critical membrane concentration and mass-balance model to identify baseline cytotoxicity of hydrophobic and ionizable organic chemicals in mammalian cell lines. *Chem Res Toxicol.* 34:2100–2109.
- Mortensen HM, Senn J, Levey T, Langley P, Williams AJ. 2021. The 2021 update of the EPA's adverse outcome pathway database. *Sci Data.* 8:169.
- Murphey RD, Zon LI. 2006. Small molecule screening in the zebrafish. *Methods.* 39:255–261.
- Neveu V, Nicolas G, Salek RM, Wishart DS, Scalbert A. 2020. Exposome-Explorer 2.0: an update incorporating candidate

- dietary biomarkers and dietary associations with cancer risk. *Nucleic Acids Res.* 48:D908–D912.
- Niu L, Henneberger L, Huchthausen J, Krauss M, Ogefere A, Escher BI. 2022. pH-Dependent partitioning of ionizable organic chemicals between the silicone polymer polydimethylsiloxane (PDMS) and water. *ACS Environ Au.* 2:253–262.
- Nowak D, Huczyński A, Bachorz RA, Hoffmann M. 2024. Machine learning application for medicinal chemistry: colchicine case, new structures, and activity prediction. *Pharmaceuticals (Basel).* 17:173.
- OECD. 2017. Guidance on grouping of chemicals. 2nd ed. OECD Series on Testing and Assessment.
- OECD. 2019. Guidance document on aquatic toxicity testing of difficult substances and mixtures. OECD series on testing and assessment, Paris, France.
- Persson L, Carney Almroth BM, Collins CD, Cornell S, de Wit CA, Diamond ML, Fantke P, Hasselov M, MacLeod M, Ryberg MW, et al. 2022. Outside the safe operating space of the planetary boundary for novel entities. *Environ Sci Technol.* 56:1510–1521.
- Peyton LR, Gallagher S, Hashemzadeh M. 2015. Triazole antifungals: a review. *Drugs Today (Barc).* 51:705–718.
- PrecisionTox Consortium. 2023. The precision toxicology initiative. *Toxicol Lett.* 383:33–42.
- Proença S, Escher BI, Fischer FC, Fisher C, Gregoire S, Hewitt N, Nicol B, Paini A, Kramer NI. 2021. Effective exposure of chemicals in in vitro cell systems: a review of chemical distribution models. *Toxicol In Vitro.* 73:105133.
- Qin W, Henneberger L, Glüge J, König M, Escher BI. 2024. Baseline toxicity model to identify the specific and nonspecific effects of per- and polyfluoroalkyl substances in cell-based bioassays. *Environ Sci Technol.* 58:5727–5738.
- Richard AM, Huang R, Waidyanatha S, Shinn P, Collins BJ, Thillainadarajah I, Grulke CM, Williams AJ, Lougee RR, Judson RS, et al. 2021. The Tox21 10K compound library: collaborative chemistry advancing toxicology. *Chem Res Toxicol.* 34:189–216.
- Richard AM, Judson RS, Houck KA, Grulke CM, Volarath P, Thillainadarajah I, Yang C, Rathman J, Martin MT, Wambaugh JF, et al. 2016. ToxCast chemical landscape: paving the road to 21st century toxicology. *Chem Res Toxicol.* 29:1225–1251.
- Sahoo SK, Ahmad MN, Kaul G, Nanduri S, Dasgupta A, Chopra S, Yaddanapudi VM. 2022. Synthesis and evaluation of triazole congeners of nitro-benzothiazinones potentially active against drug resistant *Mycobacterium tuberculosis* demonstrating bactericidal efficacy. *RSC Med Chem.* 13:585–593.
- Scholz S, Nichols JW, Escher BI, Ankley GT, Altenburger R, Blackwell B, Brack W, Burkhard L, Collette TW, Doering JA, et al. 2022. The eco-exposome concept: supporting an integrated assessment of mixtures of environmental chemicals. *Environ Toxicol Chem.* 41:30–45.
- Sewell F, Alexander-White C, Brescia S, Currie RA, Roberts R, Roper C, Vickers C, Westmoreland C, Kimber I. 2024. New approach methodologies (NAMs): identifying and overcoming hurdles to accelerated adoption. *Toxicol Res (Camb).* 13:tfae044.
- Sharma P, LaRosa C, Antwi J, Govindarajan R, Werbovetz KA. 2021. Imidazoles as potential anticancer agents: an update on recent studies. *Molecules.* 26:4213.
- Sirenko O, Grimm FA, Ryan KR, Iwata Y, Chiu WA, Parham F, Wignall JA, Anson B, Cromwell EF, Behl M, et al. 2017. In vitro cardiotoxicity assessment of environmental chemicals using an organotypic human induced pluripotent stem cell-derived model. *Toxicol Appl Pharmacol.* 322:60–74.
- Snappow I, Smith NM, Arnesdotter E, Beekmann K, Blanc EB, Braeuning A, Corsini E, Sollner Dolenc M, Duivenvoorde LPM, Sundstøl Eriksen G, et al. 2024. New approach methodologies to enhance human health risk assessment of immunotoxic properties of chemicals—a PARC (Partnership for the Assessment of Risk from Chemicals) project. *Front Toxicol.* 6:1339104.
- Su R, Xiong S, Zink D, Loo LH. 2016. High-throughput imaging-based nephrotoxicity prediction for xenobiotics with diverse chemical structures. *Arch Toxicol.* 90:2793–2808.
- Sund J, Deceuninck P. 2021. EURL ECVAM library of reference chemicals. European Commission, Joint Research Centre (JRC). <https://data.europa.eu/data/datasets/92614229-d020-4d96-941c-c9604e525c9e?locale=en>
- Ulrich N, Endo S, Brown TN, Watanabe N, Bronner G, Abraham MH, Goss K-U. 2021. UFZ-LSER database v 3.2.1 [Internet]. Leipzig, Germany: Helmholtz Centre for Environmental Research-UFZ. [accessed 2021 March 21]. <http://www.ufz.de/lserd>.
- US Environmental Protection Agency. 2023. CompTox Chemicals Dashboard [accessed 2023 Dec 20]. <https://comptox.epa.gov/dashboard/>.
- US Environmental Protection Agency. 2025. Chemical and products database (CPDat). [accessed 2024 Jan]. <https://www.epa.gov/chemical-research/chemical-and-products-database-cpdatt>.
- Wang Z, Walker GW, Muir DCG, Nagatani-Yoshida K. 2020. Toward a global understanding of chemical pollution: a first comprehensive analysis of national and regional chemical inventories. *Environ Sci Technol.* 54:2575–2584.
- Westmoreland C, Bender HJ, Doe JE, Jacobs MN, Kass GEN, Madia F, Mahony C, Manou I, Maxwell G, Prieto P, et al. 2022. Use of new approach methodologies (NAMs) in regulatory decisions for chemical safety: report from an EPA deep dive workshop. *Regul Toxicol Pharmacol.* 135:105261.
- Williams AJ, Lambert JC, Thayer K, Dorne JCM. 2021. Sourcing data on chemical properties and hazard data from the US-EPA CompTox Chemicals Dashboard: a practical guide for human risk assessment. *Environ Int.* 154:106566.
- Wunder F, Kalthof B, Müller T, Hüser J. 2008. Functional cell-based assays in microliter volumes for ultra-high throughput screening. *Comb Chem High Throughput Screen.* 11:495–504.
- Zhao YH, Ji GD, Cronin MT, Dearden JC. 1998. QSAR study of the toxicity of benzoic acids to *Vibrio fischeri*, *Daphnia magna* and carp. *Sci Total Environ.* 216:205–215.