NANYANG TECHNOLOGICAL UNIVERSITY
SINGAPORE

XENOBIOTIC BIOTRANSFORMATION IN THE ERA OF EXPOSOME: DATABASE FUSION, COMPUTER ASSISTED PREDICTION, AND CHEMICAL-CHEMICAL INTERACTION

PENG BO

SCHOOL OF CIVIL AND ENVIRONMENTAL ENGINEERING

2022

XENOBIOTIC BIOTRANSFORMATION IN THE ERA OF EXPOSOME: DATABASE FUSION, COMPUTER ASSISTED PREDICTION, AND CHEMICAL-CHEMICAL INTERACTION

PENG BO

School of Civil and Environmental Engineering

A thesis submitted to the Nanyang Technological University in partial fulfilment of the requirement for the degree of Doctor of Philosophy

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

2021-SEP-28

Peng Bo

. . . . . .
Date

. . . . . .
PENG BO

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

2021-SEP-28

. . . . . .
Date

. . . . . .
FANG MINGLIANG

Authorship Attribution Statement

This thesis contains material from [2] paper(s) published in the following peer-reviewed journal(s) / from papers accepted at conferences in which I am listed as an author.

Chapter 5 is published as Peng, B., Liu, M., Han, Y., Wanjaya, E. R., & Fang, M. (2019). Competitive Biotransformation Among Phenolic Xenobiotic Mixtures: Underestimated Risks for Toxicity Assessment. *Environmental Science & Technology*, *53*(20), 12081-12090. doi:10.1021/acs.est.9b04968

The contributions of the co-authors are as follows:

- Asst/Prof Fang provided the initial project direction and edited the manuscript drafts.
- I prepared the manuscript drafts.
- I co-designed the study with Asst/Prof Fang and did all of the lab work at the School of Civil and Environmental Engineering and the Nanyang Environment and Water Research Institute. I also went through the data and analyzed it.
- Dr. Liu assisted in the sample preparation.
- Dr. Han and Miss. Wanjaya assisted in the development of the instrument method.

Chapter 6 is published as Peng, B.; Zhao, H.; Keerthisinghe, T. P.; Yu, Y.; Chen, D.; Huang, Y.; Fang, M., Gut microbial metabolite p-cresol alters biotransformation of bisphenol A: Enzyme competition or gene induction? Journal of Hazardous Materials 2022, 426, 128093.

The contributions of the co-authors are as follows:

- Asst/Prof Fang provided the initial project direction and edited the manuscript drafts.
- I prepared the manuscript drafts and completed the *in vitro* experiments and data analysis.
- Dr. Huang, Miss. Yu conducted *in vivo* experiments and data analysis.
- Mr. Zhao conducted the docking analysis
- Dr. Keerthisinghe, Dr. Chen, and Dr. Huang contributed to suggestions

and revisions for the manuscript

2021-SEP-28

. . . . . .

Date

. . . . . .

PENG BO

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to those who have provided invaluable assistance and support in the completion of my thesis. First and foremost is my supervisor, Professor Fang Mingliang, who encourages and supports me unconditionally throughout my doctoral study. This thesis would not have been possible without his kind mentorship, insightful advice, and great patience. His academic enthusiasm and wisdom have inspired me and helped me get through a hard time.

Special thanks go to other members of my thesis advisory committee: Professor Cao Bin and Professor Zhou Yan. I am very fortunate to receive their constructive comments and helpful feedback on my thesis. I am very grateful for their generous help and for making the time in their busy schedules to serve on my committee.

I am also grateful for the various opportunities that the School and the University have given me over the years. I have benefitted greatly from the financial support provided by the NTU research scholarship. I also enjoyed the interactions with my supportive colleagues, research fellows, and labmates, which have made this memorable and fruitful intellectual journey. I thank the research and Ph.D office for wonderful administrative support.

Finally, I would like to dedicate this thesis to my beloved parents and dearest boyfriend. I could not have finished my doctoral study without their continuous support and love.

TABLE OF CONTENTS

SUMMARY

Human and biota are chronically exposed to thousands of chemicals from numerous environmental sources through various pathways. Pesticides, plasticizers, flame retardants, pharmaceuticals, and other synthetic chemicals, for example, can enter the environment and food chain, triggering adverse effects and illness. Absorption, distribution, metabolism, and elimination (ADME) are the general disposition processes when living organisms like fish, birds, humans, rodents and some other mammals, are exposed to a foreign chemical (i.e., xenobiotic). Fundamentally, these processes regulate xenobiotics' concentration at the action site, as well as the occurrence of adverse events, thereby becoming major contributors to the toxicity potential of compounds. Among all ADME processes, biotransformation (i.e., metabolism) plays a crucial role because it has the ability to directly transform the chemical structures of such compounds, therefore altering their lifetimes, bioavailability, and biological effects. To date, the prediction of xenobiotic biotransformation largely depends on experience and most prediction models have been challenged with low accuracy. In addition, living organisms are exposed to hundreds to thousands of chemicals at the same time in the era of exposome. However, there is very little information on the chemical-chemical interaction on the xenobiotic transformation.

This thesis begins with the development of a comprehensive biotransformation database, which contains multi-species data on common environmental contaminants, pharmaceuticals, and human endogenous metabolites. Text mining and database fusion were used to collect biotransformation information from multiple species, comprising 7,800 biotransformation reactions from 1,599 substrates. Unlike earlier biotransformation-related databases, the species gathered in this collection include birds, fish, rodents, and other mammals in addition to humans. This database covers substrate structures and their associated products, reaction types, the names or kinds of enzymes catalyzing biotransformation, biosystem types, and references. For the first time, environmental contaminants and various species are being addressed as a focus for data curation in biotransformation studies.

Following that, this thesis has developed a novel computational approach for biotransformation prediction based on the database. The prediction approach began

with automatically retrieved reaction templates, which were obtained by comparing the SMARTS pattern, atomic number, total number of hydrogens, formal charge, degree, number of radical electrons, aromaticity, bond orders, and atomic number of neighbouring atoms of the mapped atom-to-atom reaction of reactants and products. Subsequentially, chemical similarity metrics were used to rank prediction candidates. The results revealed that chemical similarity was a useful metric for predicting metabolite structure. Extensive evaluations showed that the current model outcompeted previous methods including Biotransformer, GLORYx, and SyGMa. These findings suggest that the proposed method can give a simple and automated solution that does not require any expert knowledge. We also combined the first two parts as an open-access data portal, EcoBioTrans (https://www.ecobiotrans.asia)[1]. Users are allowed to query by compound name and simplified molecular input line entry specification (SMILES) for retrieving the database or submit compound structure for retrieving prediction results.

The third part of the thesis explored the biotransformation behavior of environmental phenolic compounds under single exposure and concurrent exposure. Bisphenol A (BPA) was chosen as a model molecule to explore its biotransformation in liver microsome and cell models after single/co-exposure to additional phenolic xenobiotics (Triclosan (TCS), Tetrabromobisphenol A (TBBPA), Bisphenol S (BPS)) and a combination of 22 phenolic compounds. The results suggested that biotransformation of phenolic xenobiotics can be significantly altered by co-exposure, which provides referential evidence on underestimated risks of simultaneous exposure to environmental toxicants.

This thesis further tentatively discovered the possible interplay between the gut microbiome and biotransformation of environmental contaminants via both *in vivo* and *in vitro* experiments. Bisphenol A (BPA) and *p*-cresol were utilized as model compounds. The results indicated that biotransformation behaviors differed between the *in vitro* and *in vivo* results. *In vivo* mouse examination using *p*-cresol injection exhibited enhancing effect on BPA metabolism, which is rarely found in mixture studies. However, in both *in vitro* models of liver S9 fractions and HepG2 cell line, *p*-cresol is found as a strong inhibitor in a non-competitive pattern for BPA

[1] Please use the server address (http://1.117.57.232/) to view the website as our server is being audited and certified.

biotransformation. The subsequent close investigation revealed that the expression of biotransformation enzyme genes including *Ugt1a1*, *Ugt2b1*, *or Sult1a1* were dynamically induced after the *p*-cresol treatment.

In sum, this thesis work has built a comprehensive biotransformation database, provided a novel xenobiotic biotransformation prediction platform, and further extended our insight into the biotransformation interaction between chemicals. The results can be applied as some fundamental tools and knowledge in xenobiotic biotransformation.

LIST OF PUBLICATIONS

- **Peng, B.**, Liu, M., Han, Y., Wanjaya, E. R., & Fang, M. (2019). Competitive Biotransformation Among Phenolic Xenobiotic Mixtures: Underestimated Risks for Toxicity Assessment. *Environmental Science & Technology, 53*(20), 12081-12090. doi:10.1021/acs.est.9b04968
- **Peng, B.**; Zhao, H.; Keerthisinghe, T. P.; Yu, Y.; Chen, D.; Huang, Y.; Fang, M., Gut microbial metabolite p-cresol alters biotransformation of bisphenol A: Enzyme competition or gene induction? *Journal of Hazardous Materials 2022*, 426, 128093.
- **Peng, B.**, & Fang, M. EcoBioTrans: A Comprehensive Database Utilizing Multi-Species Biotransformation Data. ***To Be Submitted.***
- **Peng, B.**, & Fang, M. Computer-Assisted Biotransformation Prediction Based on Molecular Similarity. ***To Be Submitted.***
- Yang, J. J., Han, Y., Mah, C. H., Wanjaya, E., **Peng, B.**, Xu, T. F., . . . Fang, M. L. (2020). Streamlined MRM method transfer between instruments assisted with HRMS matching and retention-time prediction. *Analytica Chimica Acta, 1100*, 88-96. doi:https://doi.org/10.1016/j.aca.2019.12.002
- Zhang, Y., Liu, M., **Peng, B.**, Jia, S., Koh, D., Wang, Y., . . . Fang, M. (2020). Impact of Mixture Effects between Emerging Organic Contaminants on Cytotoxicity: A Systems Biological Understanding of Synergism between Tris(1,3-dichloro-2-propyl)phosphate and Triphenyl Phosphate. Environmental Science & Technology, 54(17), 10722-10734. doi:10.1021/acs.est.0c02188
- Liu, M., **Peng, B.**, Su, G., Fang, M. L. Reactive Flame Retardants: Are They Safer Replacements? *Environmental Science & Technology*, ***Under Review.***

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Description
ADME	Absorption, distribution, metabolism, and elimination
ADEs	Adverse drug events
BDAGE	Bisphenol A Diglycidyl Ether
BFDGE	Bisphenol F Diglycidyl Ether
BPA	Bisphenol A
CAPD	Continuous ambulatory peritoneal dialysis
CLA	Conjugated linoleic acid
DCA	Deoxycholic acid
DDI	Drug-drug interactions
DHB	Dihydroxybenzophenone
EH	Epoxide hydrolases
EGCG	Epigallocatechin gallate
EU	European Union
FDA	U.S. Food and Drug Administration
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
HMDB	Human Metabolome Database
HPV	High Production Volume
HR-MS	High-resolution mass spectrometry
IUCN	International Union for Conservation of Nature
LCA	Lithocholic acid
MDR	Multidrug resistance protein
UDCA	Ursodeoxycholic acid
GI	Gastrointestinal
TCS	Triclosan
TBBPA	Tetrabromobisphenol A
BPS	Bisphenol S
PAH	Polycyclic aromatic hydrocarbons
PBS	Phosphate buffered saline
RT-PCR	Real-time polymerase chain reaction
PCBs	Polychlorinated biphenyls
PDB	Protein Data Bank
PKKB	PharmacoKinetics Knowledge Base
PUG	Power User Gateway
QED	Quantitative estimation of drug-likeness
ESI	Electrospray ionization
DMEM	Dulbecco’s modified Eagle’s medium
FBS	Fetal bovine serum
DMSO	Dimethyl sulfoxide
QA	Quality assurance
QC	Quality control
SEM	Standard error of the mean

SMILES	Simplified molecular input line entry specification
SMARTS	SMILES arbitrary target specification
MBP	4-Methyl-2,4-bis(4-hydroxyphenyl)pent-1-ene
MnBP	Mono-n-butyl phthalate
BzBP	Benzylbutyl phthalate
MBzP	Mono-benzyl phthalate
MEHP	Mono-2-ethylhexyl phthalate
DiNP	Di-isononyl phthalate
DiDP	Di-isodecyl phthalate
PFOA	Perfluorooctanoic acid
mRNA	Messenger RNA
MW	Molecular weight
NADPH	Nicotinamide adenine dinucleotide phosphate
NQO	NADPH-quinone oxidoreductase
DTT	DL-Dithiothreitol
NAPQI	N-acetyl-p-benzoquinonimine
NSAIDs	Nonsteroidal anti-inflammatory drugs
logP	Partition coefficient
PFOS	Perfluorooctanesulfonic acid
PFOA	Perfluorooctanoic acid
UGT	UDP-glucuronosyltransferases
VOCs	Volatile organic compounds
RLM	Rat liver microsomes
SULT	Sulfotransferases
GST	Glutathione-S-transferases
SOM	Site of metabolism
HR-MS	High-resolution mass spectrometry
CYP	Cytochromes P450
TDCPP	Tris(1,3-dichloroisopropyl) phosphate
TMA	Trimethylamine
PXR	Pregnane X receptor
GLP	Glucagon-like peptide
CaSR	Ca^{2+} sensing receptor
PPAR	Proliferator-activated receptor
GLP	Glucagon-like peptide
NK	natural killer
NMDAR	N-methyl-D-aspartate receptor
a7nAChR	A7 nicotinic acetylcholine receptor
DRD2	Dopamine receptor D2
DDIs	Drug-drug interactions
ADEs	Adverse drug events
TPhP	Triphenyl phosphate

Chapter 1. Introduction

1.1 Background

Human and biota are chronically exposed to thousands of chemicals from numerous environmental sources through various pathways (Jiang et al. 2018). For example, pesticides, plasticizers, flame retardants, pharmaceuticals, and other synthetic chemicals can enter the environment and the food chain, causing unwanted effects and disease (Escher et al. 2020). When living organisms such as fish, birds, humans, rodents are exposed to these chemicals, the general disposition processes are known as absorption, distribution, metabolism, and elimination (ADME) (Klaassen et al. 2013). Among all ADME processes, biotransformation plays a central role because it can directly alter the chemical structures of such compounds, thus modifying their lifetime, bioavailability, and biological effects (Koppel et al. 2017). In particular, biotransformation could affect foreign chemical toxicity in two ways: it could make it less toxic (detoxication), or it could make it more toxic (activation) by converting it to a more reactive form, such as the oxidation of ethanol to acetaldehyde.

Hence, extensive studies have been done in recent decades to identify the biotransformation pathways of a wide range of medications, toxins, and industrial substances. To begin, in drug discovery, the basis for determining the biological effects of putative medicines is a thorough understanding of their metabolic pathways. There are a large number of *in vivo* and *in vitro* studies on the pharmacokinetics or metabolism of drugs, including ziprasidone (Prakash et al. 1997), ambroxol (Hammer et al. 1978), acetaminophen (McGill and Jaeschke 2013), and ftorafur (Benvenuto et al. 1978). Moreover, understanding the metabolism of all types of chemical substance is important for risk assessment in the field of toxicology. In some cases, even if the compounds themselves are not toxic, their metabolites cause the toxic effects. For instance, acetaminophen hepatotoxicity is mediated by the reactive metabolite N-acetyl-p-benzoquinonimine (NAPQI) (Yoon et al. 2016); MBP (4-Methyl-2,4-bis(4-hydroxyphenyl)pent-1-ene), an active metabolite of bisphenol A (BPA), has significantly more potent estrogenic toxicity than BPA itself (Okuda et al. 2010). Lastly, in the field of environmental monitoring, the

characterization or identification of a target compound is commonly performed by determining the level of its metabolites (Ekman et al. 2015). Most chemicals are transformed into metabolites within several hours after entering the body systems, and they are commonly found in living species as the form of metabolite (Kammann et al. 2013).

Meanwhile, the curation of biotransformation data and the creation of *in silico* models for predicting biotransformation are active fields of research. The growth, evolution, and accessibility of high-resolution mass spectrometry (HR-MS) promoted advanced detection in biological and environmental samples and increased the amount of available data. Previous research has therefore developed biotransformation-relevant databases, including commercial ones like the Accelrys Metabolite Database (*Accelrys Metabolite Database* 2011) and the Fujitsu ADME Database ("ADME database (Fujitsu Kyushu Systems Ltd 2014),"), and publicly available databases such as DrugBank (Wishart 2006), Human Metabolome Database (HMDB) (Wishart 2007), EAWAG Biocatalysis/Biodegradation Database (Gao et al. 2010) and Transformer (Hoffmann et al. 2014). Based on these biotransformation data, a variety of *in silico* computational algorithms and integrated approaches have been developed and are available as online services or commercial applications. The three main research fields of biotransformation prediction approaches are prediction of metabolic sites (SOMs), prediction of metabolite structures, and prediction of metabolizing enzyme binding, all of which provide critical assistance and guidance for experimenters.

However, there are a few key knowledge gaps related to biotransformation that could be addressed for a more comprehensive understanding. First, there is a scarcity of biotransformation relevant data. On the one hand, some of the crucial data required to build prediction models are locked away in private and commercial databases like the Accelrys Metabolite database. On the other hand, the coverage of data is limited on the biotransformations of pharmaceuticals in the human body system. Nonetheless, many environmental chemicals, such as flame retardants, plasticizers, and so on, harm not just human health but also the entire ecosystem. Second, few prediction methods has been developed with enough metabolic information of

environmental chemicals. Furthermore, most prediction approaches for metabolite structure necessitate manually extracting the template from massive reactions, which is time-consuming and demands a significant amount of time and effort from specialists. Third, though simultaneous exposure to multiple compounds increases the risk of interaction, most biotransformation studies were conducted using individual chemicals, and whether co-exposure of multiple environmental chemicals will affect each other's fate in the human body needs further research. Lastly, despite the fact that some gut microbial metabolites are processed by the mammalian system in a manner analogous to xenobiotics, the possible interplay between the gut microbiome and biotransformation of environmental chemicals has never been investigated yet.

1.2 Purpose and scope

The objectives of this PhD work are to

(i) Establish a comprehensive database incorporating multi-species biotransformation information as well as physicochemical properties of common environmental pollutants, pharmaceuticals, gut microbiota derived metabolites, and develop an open-access data portal that can conduct searches for content associated with chemicals in this database.

(ii) Develop a novel *in silico* computational method for biotransformation prediction that identifies probable metabolite structure without expert hand-coding templates and robustly ranks the predicted metabolites by similarity metrics; integrate the prediction function into the database website.

(iii) Investigate biotransformation behavior of environmental phenolic compounds under single exposure and concurrent exposure; tentatively explore the mixture effect on biotransformation after exposure to 22 common organic contaminants of human-relevant level.

(iv) Characterize the possible interplay between the gut microbiome and biotransformation of environmental contaminants via both *in vivo* and *in vitro* experiments.

1.3 Dissertation overview

This dissertation includes seven chapters. Chapter 1 and 2 are the introduction and literature review. Chapter 7 is the conclusions and recommendations.

Chapter 3 curated biotransformation information among multiple species and set up a comprehensive database, which harbors biotransformation data of common environmental pollutants, along with pharmaceuticals and human endogenous metabolites. To establish this data repository, we collected 7,800 biotransformation reactions of 1,599 substrates by text mining and database fusion. The collected species are not restricted to human, but also extended to birds, fish, rodents, and other mammals. This database contains structures of substrates and their corresponding products, types of reactions, the names or types of enzymes catalyzing biotransformation, types of biosystems, and references. Users are also allowed to search biotransformation results based on structural similarity. Overall, for the first time, environmental pollutants and multiple species are targeted as a focus for biotransformation data curation. This integrative platform makes a great contribution to the scientific understanding of the biotransformation path in the whole ecosystem and serves as a guidance for chemical management.

Chapter 4 presented a novel computational method with an automatically extracted reaction template, which further ranks candidates by molecular similarity metric. Our work demonstrates molecular similarity to be an effective metric. Extensive evaluation reveals that the current model achieves 45.25% recall and 26.37% precision at the top 3 candidates and 65.67% recall and 7.96 % precision at the top 20 candidates, respectively. Our results indicate that the proposed approach can provide a simple and automatic solution that does not restrict to any expert knowledge.

Chapter 5 selected bisphenol A (BPA) as a model compound. Its biotransformation was investigated under single/co-exposure to other phenolic xenobiotics (Triclosan (TCS), Tetrabromobisphenol A (TBBPA), and Bisphenol S (BPS)) in liver microsome and cell models. The result shows that binary exposures exhibit significant inhibitory effects on BPA metabolism, especially the sulfate conjugation.

In combination of analysis on inhibition models and enzyme activity, the inhibition effect was suggested to be primarily driven by competition for metabolizing enzymes. A mixture with 22 phenolic chemicals was further examined to disrupt BPA at various human-relevant levels. Again, the result demonstrates significant inhibition on BPA metabolism, indicating the possible natural existence of our finding. Overall, our results show that biotransformation of phenolic xenobiotics can be significantly altered by co-exposure, which provides referential evidence on underestimated risks of simultaneous exposure to environmental toxicants.

Chapter 6 utilized bisphenol A (BPA) and *p*-cresol as model compounds to explore whether gut microbial metabolites could affect environmental phenol metabolism on both *in vitro* and *in vivo* models. The results indicated that biotransformation behaviors differed between the *in vitro* and *in vivo* results. *In vivo* mouse examination using *p*-cresol injection exhibited enhancing effect on BPA metabolism, which is rarely found in mixture studies. The serum level of BPA was significantly lower in the *p*-cresol group with a clear dose-response relationship. However, in both *in vitro* models of liver S9 fractions and HepG2 cell line, *p*-cresol is found as a strong inhibitor in a non-competitive pattern for BPA biotransformation. *In silico* docking approach was utilized to explore the non-competitive inhibition mechanism by estimating the binding affinity of *p*-cresol, BPA and two allosteric binding sites of key enzyme SULT 1A1. Next, a close investigation revealed that the expression of biotransformation enzyme genes including *Ugt1a1, Ugt2b1, or Sult1a1* were dynamically induced after the *p*-cresol treatment. Overall, our results provided a novel insight into the biotransformation interaction between gut microbiome and environmental contaminants. The discrepancy between *in vitro* and *in vivo* models suggested that both enzyme competition and dynamic gene regulation should be considered when evaluating the interaction between gut microbiome and environmental xenobiotics.

Chapter 2. Review of Theory and Previous Work

2.1. Introduction of biotransformation

2.1.1 Absorption, distribution, metabolism, and elimination (ADME)

Absorption, distribution, metabolism, and elimination (ADME) are the general disposition processes when living organisms like fish, birds, humans, rodents, and some other mammals, are exposed to a foreign chemical (i.e., xenobiotic), such as pharmaceutical, dietary component, and environmental pollutant. Fundamentally, these processes regulate xenobiotics' concentration at the action site, as well as the occurrence of adverse events, thereby becoming major contributors to the toxicity potential of compounds (Klaassen et al. 2013; Vrbanac and Slauter 2017).

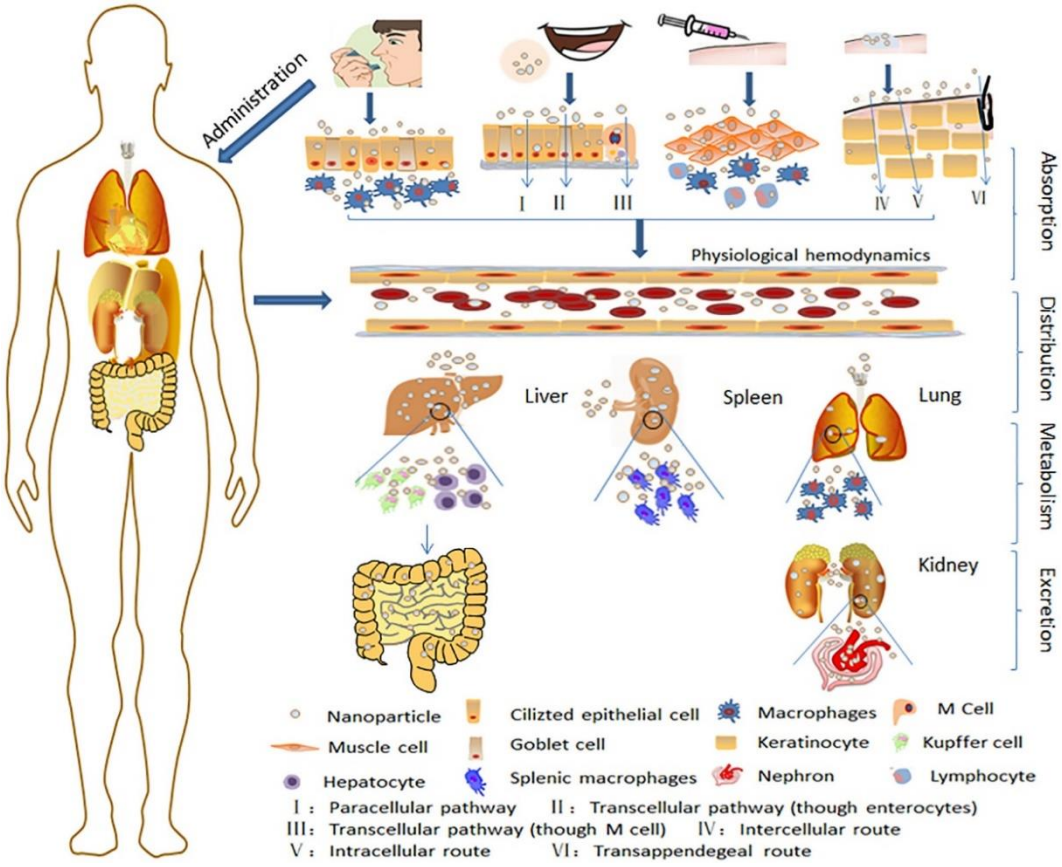


Figure 2.1 A schematic overview of absorption, distribution, metabolism (biotransformation), and elimination (ADME) process, which is adapted from previous study (Zhang et al. 2020).

Specifically, absorption is the process by which xenobiotics penetrate body

membranes to enter the systematic circulation. For most animals, the main sites of absorption include the gastrointestinal tract (GI tract), lungs and skin. Thereinto, the GI tract is the most important site for xenobiotics that are absorbed via the oral route, including most drugs and numerous environmental toxicants. The subsequent process is distribution, where a xenobiotic is distributed to tissues throughout the body after entering the bloodstream. That is followed by the central process once xenobiotics are transported to the liver through the hepatic portal vein, which is metabolism, also named biotransformation. Biotransformation is a series of enzyme-catalyzed processes which modify the physiochemical properties of foreign chemicals from one absorbed (lipophilic) to one that can be readily eliminated from the body (hydrophilic). Subsequent to biotransformation, toxicants can be excreted from the body in several routes, consisting of urinary excretion, fecal excretion, exhalation, and other uncommon routes (Caldwell et al. 1995; Klaassen et al. 2013).

Scientific efforts on ADME studies have promoted a comprehensive understanding of the fate of a xenobiotic and been of great significance to the field of toxicology. In particular, the toxicologic profile of a substance unequivocally depends on whether a xenobiotic and/or its metabolites are present in appropriate concentrations at its sites of action. Poor absorption of a toxicant may limit or even prevent toxicity because it may never reach a sufficiently high concentration at a potential site of action to trigger toxic effects. Likewise, rapid biotransformation prevents a toxicant from reaching a sufficiently high concentration at a potential site of action, and thus toxic effects will not be induced. For example, ochratoxin A is toxic, carcinogenic, and organ-specific due to poor biotransformation and slow elimination of metabolites; perfluorinated compounds such as perfluorooctanesulfonic acid (PFOS) and perfluorooctanoic acid (PFOA) have been attracting a lot of attention as they are barely metabolized in mammals, which exacerbates their toxicological effects including neurotoxicity, carcinogenesis, genotoxicity and epigenetic effects, immunotoxicity, reproductive and developmental toxicities, and effects on the endocrine system (Klaassen et al. 2013; Stahl et al. 2011; Tran et al. 2020). However, if a xenobiotic can permeate cells at a high rate of passive diffusion but cannot be biotransformed at an appreciable rate, the xenobiotic, such as Dioxin (2,3,7,8-tetrachol-orodibenzo-*p*-dioxin or TCDD), can have a long residence time in the body thus resulting in exacerbated toxicity

(Klaassen et al. 2013). Under certain conditions, nevertheless, the biotransformed metabolites are more toxic than the parent chemical. For example, MBP (4-Methyl-2,4-bis(4-hydroxyphenyl)pent-1-ene), an active metabolite of bisphenol A (BPA), has much greater estrogenic toxicity than BPA itself (Okuda et al. 2010).

2.1.2 Biotransformation

Among all ADME processes, biotransformation (i.e., metabolism) plays a central role because it changes the chemical structures of xenobiotics from those that favor absorption across biological membranes (namely, lipophilicity) to those that favor elimination in urine or bile (namely, hydrophilicity). This process is critical because many ingested xenobiotics are commonly lipophilic and would accumulate to toxic levels without metabolism and subsequent elimination. However, biotransformation may also convert the xenobiotics into toxic substances, comprising mutagenic, carcinogenic, or teratogenic molecules, like the above-mentioned acetaminophen. It influences parameters such as absorption, passive membrane permeation, transport and binding to macromolecules as well (Braga and Andrade 2012). In general, biotransformation can change the biological properties of a xenobiotic, which could make the xenobiotic less toxic (detoxication), or more toxic(activation) (Klaassen et al. 2013).

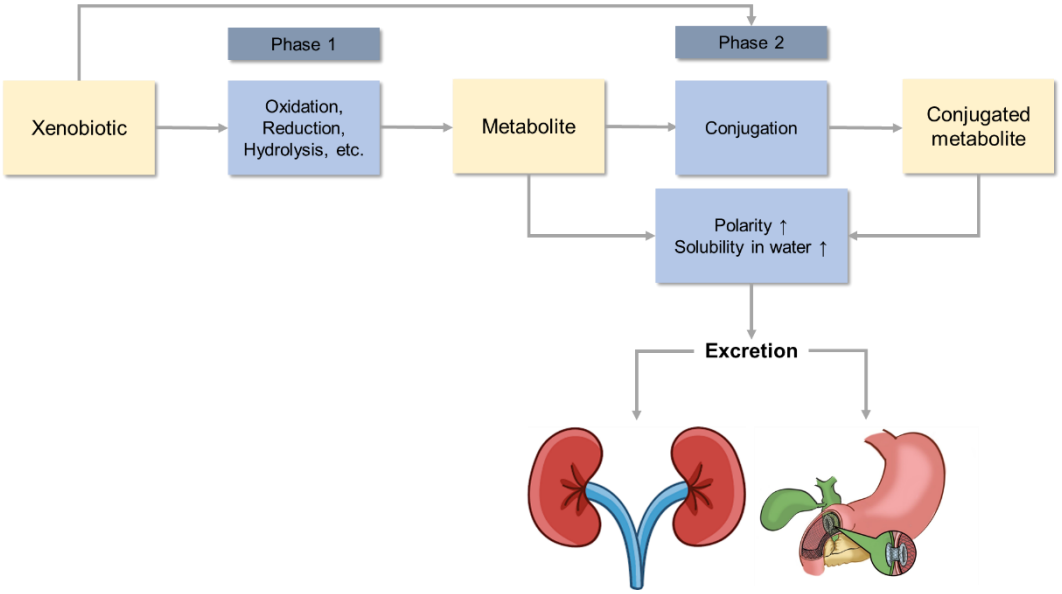


Figure 2.2 A schematic overview of phase I and phase II biotransformation reactions.

Biotransforming reactions are commonly classified into two groups, phase I (Functionalization) and phase II (Conjugation). Phase I reactions include hydrolysis, reduction, and oxidation whereas Phase II biotransformation reactions involve conjugation reactions such as glucuronidation, sulfonation (or sulfation), glutathione conjugation, methylation, acetylation, conjugation with amino acids, and other

unusual conjugations. Apart from methylation and acetylation, conjugation reactions lead to a great rise in their hydrophilicity, so they greatly promote the excretion of xenobiotics (Hayes and Kruger 2014; Klaassen et al. 2013).

The enzymes that catalyze xenobiotic biotransformation are usually called drug-metabolizing enzymes, which are summarized in Table 2.1. The major biotransformation reactions involved in the oxidation, reduction, hydrolysis, and conjugation are shown in Table 2.2 (Klaassen et al. 2013; Vrbanac and Slauter 2017). However, some biotransformation reactions are not catalyzed by mammalian enzymes that listed in Table 2.1. Chances are that gut microbes modified the xenobiotics' chemical structures as well (Koppel et al. 2017).

The majority of phase I reactions are catalyzed by the CYPs (Cytochrome P450s), a superfamily of monooxygenases presenting in eukaryotes as well as in prokaryotes (Koenig et al. 2012). Xenobiotics that bind to receptor proteins can induce the expression of CYP genes, which in turn enhances protein synthesis and associated enzyme activity (Honkakoski and Negishi 1998). Several enzyme superfamilies are involved in phase II conjugating reactions, including uridine diphosphate-glucuronosyltransferase (UGT), sulfotransferase (SULT), and glutathione S-transferase (GST) (Xu et al. 2005).

Table 2.1 Table of enzymes, localizations, and reactions for xenobiotic metabolizing enzymes (adapted from previous study) (Vrbanac and Slauter 2017)

Enzyme Types	Localizations	Reactions
PHASE I		
Cytochrome P450s (CYP)	Endoplasmic reticulum	C and O oxidation, dealkylation, and other oxidations
Flavin-containing monooxygenases (FMO)	Endoplasmic reticulum	N, S, and P oxidation
Epoxide hydrolases (EH)	Endoplasmic reticulum, cytosol	Hydration of epoxides
Reductases	Cytosol	Reduction reactions
Monoamine oxidase		Oxidative deamination of monoamines
HYDROLASES		
Esterases	Endoplasmic reticulum, cytosol	Hydrolyzes ester bonds
Amidases	Endoplasmic reticulum, cytosol	Hydrolyzes amide bonds
PHASE II		
UDP-glucuronosyltransferases (UGT)	Endoplasmic reticulum	Addition of glucuronic acid
Sulfotransferases (SULT)	Cytosol	Addition of sulfate
Glutathione-S-transferases (GST)	Endoplasmic reticulum	Addition of glutathione
Amino acid transferases	Endoplasmic reticulum	Addition of amino acid
Methyltransferases	Endoplasmic reticulum	Addition of methyl group
N-acetyltransferases	Endoplasmic reticulum	Addition of acetyl group
Miscellaneous transferases		
MISCELLANEOUS ENZYMES		
Alcohol dehydrogenase	Cytosol	Oxidation of alcohols
Aldehyde dehydrogenase	Cytosol	Oxidation of aldehydes
NADPH-quinone oxidoreductase (NQO)	Endoplasmic reticulum	Reduction of quinones

Table 2.2 General pathways of xenobiotic biotransformation (adapted from previous study) (Klaassen et al. 2013).

Reaction Class	Specific reaction
Conjugation	Methylation
	Acetylation
	Glucuronidation
	Sulfonation
	Formation of Acyl-CoA Thioesters
	Amino Acid Conjugation
	Glutathione Conjugation
	Thiosulfate Formation
	Unusual Conjugation Reactions
	Phosphorylation
Hydrolysis	Demethylation
	Deethylation
	Deacetylation
	Decarboxylation
	Dephosphorylation
	Hydrolysis
Reduction	Azo- and Nitro-Reduction
	Carbonyl Reduction
	Disulfide Reduction
	Sulfoxide and N-Oxide Reduction
	Quinone Reduction
	Dihydropyrimidine Dehydrogenation
	Dehydroxylation
	Aldehyde Oxidase—Reductive Reactions
	Reductive Dehalogenation
Oxidation	Alcohol, Aldehyde, Ketone Oxidation–Reduction
	Alcohol Dehydrogenation
	Aldehyde Dehydrogenation
	Dimeric Dihydrodiol Dehydrogenation
	Molybdenum Hydroxylysis
	Xanthine Oxidoreduction
	Aldehyde Oxidation
	Amine Oxidation
	Aromatization
	Peroxidase-Dependent Cooxidation
	Oxidative Dehalogenation
	Flavin Monooxygenases Related Oxidation

2.1.3 Species differences in ADME processes

Prior research has intensively studied species differences in regard to ADME processes. In terms of elimination, species variant urinary excretion of weak organic acids and bases is commonly identified because the pH of urine varies greatly between species (Hutchinson et al. 2014; Klaassen et al. 2013; Martignoni 2006).

However, it is certain that the basis for species differences in both the qualitative and quantitative aspects of xenobiotic biotransformation and toxicity is the species-specific metabolic enzyme activities. Their expression patterns and substrate specificities can vary greatly between species, implying that when extrapolating from in vitro and animal testing results, there is a risk of missing toxic metabolites formed in humans (Kirchmair et al. 2015). For example, species difference in coumarin metabolism and toxicity (activation by epoxidation in rats, detoxication by aromatic hydroxylation in humans) reflects a species difference in hepatic microsomal cytochrome P450, whereas glutathione transferase activity accounts for the difference in susceptibility to aflatoxin-induced liver toxicity between rats and mice. In addition, previous studies on the biotransformation of benfuracarb by hepatic microsomes taken from seven mammalian species revealed that the clearance rates for the seven species ranged from 1.4 to 3.5 as a result of different cytochrome P450 enzyme expression. Besides, when compared on a weight-normalized basis, small mammals have a tendency to eliminate drugs faster than humans due to more active hepatic monooxygenase activities (Hutchinson et al. 2014; Klaassen et al. 2013). Likewise, body weight and hepatic microsomal monooxygenase activity are also highly correlated across birds (Walker and Ronis 1989). Nevertheless, unlike the aforementioned species, fish tend to preserve much lower Phase I enzyme activities devoid of a strong correlation with body weight. This is probably because of their peculiar excretion approach, diffusion across gills into vast ambient water, and reduced pressure for the evolution of highly active detoxification enzymes (Du et al. 2014). Moreover, entire pathways of xenobiotic biotransformation are low or absent in some species. Binu Shrestha *et al.* found that phenolic compounds including acetaminophen, aspirin, and structurally-related toxicants like phenolic plasticizers found in the diet and environment, are toxic to the domestic cat at a greater extent,

because pseudogenization of the gene is in charge of phenol detoxification enzyme (Shrestha et al. 2011).

Specifically, UDP-glucuronosyltransferase (UGT), one of the most common enzymes in living organisms which expresses in various tissues and catalyzes glucuronide conjugation, has been discovered to have variable expression patterns among species. As an example, the UGT1E family is critical to phase II metabolism in birds, whereas mammalian UGT1A isoforms are proved to be essential for humans. Figure 2.3's phylogenetic tree shows that the avian UGT1E isoforms might evolve independently of the mammalian UGT1A isoforms as a result of separate locations (Kawai et al. 2019). This might be due to the fact that the origin of birds is closer to reptiles than mammals in biological classification, meaning they may have UGT isoforms in common with reptile species instead of mammals. Previous research has also found that zebra finches and lizards have UGT isoforms derived from the same ancestral variable exon that mammals do not.(Huang and Wu 2010).

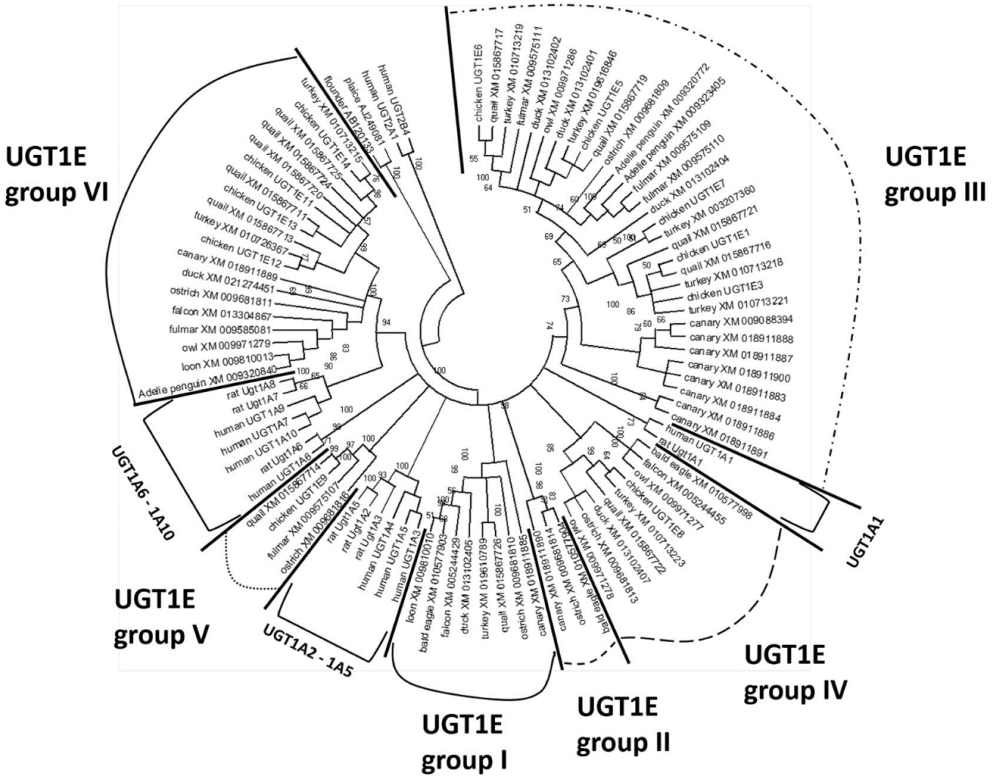


Figure 2.3 Phylogenetic trees of the UGT isoforms using mammalian UGT 1A and avian UGT 1E amino acid sequences, which is adapted from literature (Kawai et al. 2019).

Besides, Cytochrome (CYP) enzymes that catalyze metabolic processes of both endogenous and foreign substances have also been revealed to have different expression patterns among species. Except for plants, the enzyme superfamily may be found in a wide variety of species despite type (Ahmad et al. 2021). A total of eighteen CYP gene families have been discovered in fish, with members including CYP1, CYP2, CYP3, CYP4, CYP5, CYP7, CYP8, CYP11, CYP17, CYP19, CYP20, CYP21, CYP24, CYP26, CYP27, CYP39, CYP46 and CYP51 (Uno et al. 2012). A comparative analysis of human and Japanese pufferfish genomes revealed that six CYP families (1, 3, 4, 5, 17, and 21) could be found in both species, although several subfamilies were unique to each one, such as 1C and 3B in pufferfish and 4A, 4B, 4X, and 4Z in humans (Nelson 2003). Another comparative analysis of human and birds also reported the absence of clear orthology between avian and human CYP2C or CYP3A genes. For instance, certain avian features were discovered in the members of the CYP2 gene family, including the presence of several genes in the avian CYP2J, CYP2AB, CYP2AC, and CYP2W subfamilies, while humans were only endowed with a single gene. Besides, according to the phylogeny presented in Figure 2.4, the CYP3 family was split into two subgroups: the avian CYP3A subgroup and the human CYP3A subgroup. Birds have two CYP3A genes, CYP3A37 and CYP3A80, while humans have four CYP3A genes (Watanabe et al. 2013).

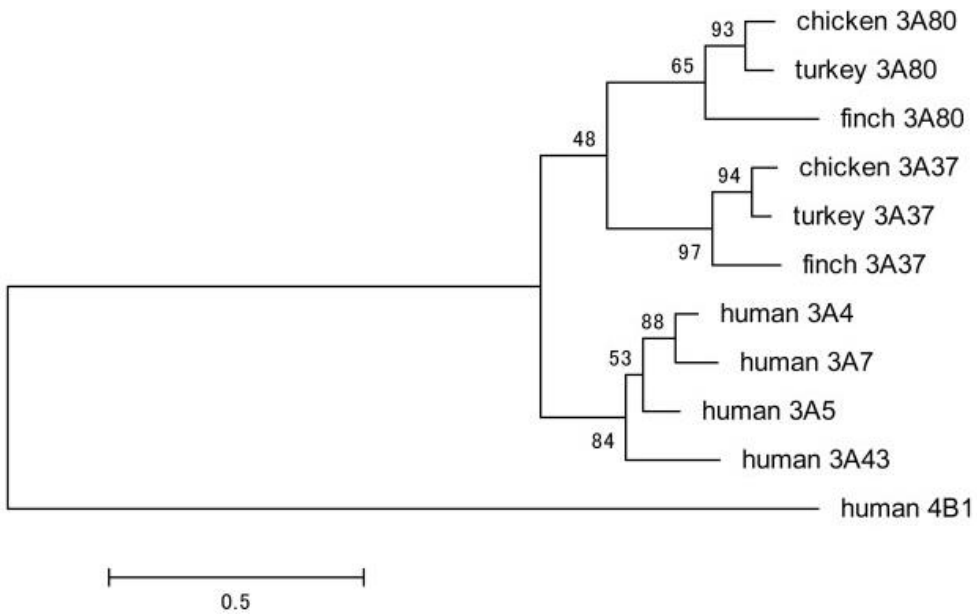


Figure 2.4 CYP3 amino acid sequence phylogeny from chicken, zebra finch, turkey,

and human (Watanabe et al. 2013). The maximum likelihood tree was created using MEGA5 software. The numbers on the branches indicate the number of times per 100 bootstrap replicates that the branch appeared in the trees, estimated by a random resampling of the data. The scale bar represents 50 substitutions per 100 residues.

Therefore, owing to these discrepancies in ADME processes, exposures to the same chemicals might have distinct effects on different species. Even if a compound is designed for human benefit and harmless to human, it may have deleterious impacts on other species. Karen A. Kidd et al. discovered that chronic exposure of fathead minnows to low concentrations of the potent 17-ethynylestradiol resulted in feminization of males via the production of vitellogenin mRNA and protein. Such effects on gonadal development are demonstrated by intersex in males and altered oogenesis in females, and, ultimately, the species' near extinction (Kidd et al. 2007). Similarly, as a result of large and rapid vulture population declines on the Indian subcontinent, these species have been listed as critically endangered by the IUCN. Diclofenac-induced poisoning of vultures through ingestion of tissues from livestock is consistent with the evidence of population decline, cause of death, and toxicity (Cuthbert et al. 2006). Thus, more detailed risk assessments should be taken to characterize the biological fates of environmental pollutants of the wild species in the ecosystem.

2.2. Biotransformation databases

To date, the curation of biotransformation data and prediction of biotransformation products is an active field of research due to its significant role. For the prediction of xenobiotic metabolism, a wide range of computational methods and integrated approaches have been developed and present as web services or commercial software. Indeed, almost all computational models are based on experimental data. The amount of data available, as well as its quality, determines the model's coverage and performance. This work summarized recent progress on the literature of xenobiotic metabolism using databases in later sections.

2.2.1. Introduction of biotransformation databases

Recently, there has been an increase in research efforts to enhance public access to experimental data on small organic molecules, but data on xenobiotic metabolism are still scarce. There are only a few databases in the public domain that are relevant to this topic (see TABLE 2.3), though data collections on metabolism have been built up over decades. On the one hand, relevant data used to construct these models are either locked away in proprietary and commercial databases, such as the Accelrys Metabolite database and the Fujitsu ADME database (*Accelrys Metabolite Database*, 2011; "ADME database (Fujitsu Kyushu Systems Ltd 2014),"). On the other hand, the coverage of data is limited on the biotransformations of pharmaceuticals in the human body system or in the microbial system. For example, the EAWAG Biocatalysis/Biodegradation Database has made excellent progress on collecting microbial biocatalytic reactions and biodegradation pathways for primarily xenobiotic, chemical compounds, but the data is limited to environmental biodegradation (Gao et al. 2010). In addition, the other well-known database with biotransformation information primarily focused on pharmacokinetics data of drugs, such as DrugBank, BindingDB (Liu et al. 2007; Wishart et al. 2018). Nevertheless, a large number of environmental chemicals, such as flame retardants and plasticizers, endanger not only human health, but also the entire ecosystem. Moreover, published data on xenobiotic metabolism of environmental chemicals, which may serve as an information source for a new database, are currently fragmented over many journals and publications. Thus far, there are only a few existing databases containing biotransformation information.

2.2.2. Previous database summary

Accelrys Metabolite Database, also called MDL Metabolite Database, is a commercial database containing biotransformation reaction information with the longest history and the largest dataset (*Accelrys Metabolite Database* 2011). It was established by MDL Information Systems in 1994, via collecting data from literature (Pfeifer et al. 1983). After several rounds of merging and updating, the most recent version contains ~20,000 parent compounds and ~100,000 biotransformation reactions, over 98% of which are pharmaceuticals. It is basically a database of biotransformations found in mammals, 17,289 of which are human (Borodina et al.,

2003).

MetaBase, is another commercially available database of biotransformations, developed by GeneGo (Peach et al. 2012). It is a large database that includes literature data on small molecule–protein, protein–protein and protein–DNA interactions, signaling pathways, regulatory networks and diseases. The primary focus of MetaBase is the biological properties and effects of drugs and drug-like molecules. Now it has been merged into MetaCore (https://clarivate.com/products/metacore/, Clarivate Analytics, London, UK).

In terms of freely available biotransformation metabolite database, a popular and relatively long-lived, open-source tool is DrugBank. It is a free, comprehensive online resource that provides detailed information on FDA-approved drugs as well as experimental medications that are undergoing FDA approval. For the biotransformation aspect, it reports information of substrates, products as well as enzyme if possible. However, the coverage of DrugBank is limited to pharmaceuticals (Wishart et al. 2018). Another freely available database for biotransformation is MetXBioDB, which is a database that consists of a manually curated collection of > 2000 experimentally confirmed biotransformations derived from the literature. Most of the substrates are overlapped with DrugBank while the coverage extends to pesticide and other industrial chemicals a little bit. Each biotransformation in MetXBioDB includes a starting reactant, a reaction product, the name or type of the enzyme catalysing the biotransformation, the type of reaction, and one or more citations. It also collected over 30 biotransformation reactions in human gut system (Djoumbou-Feunang et al. 2019). Transformer provides comprehensive information on the transformation and transport of xenobiotics in the human body. It contains the interactions of phase I and II enzymes and drug transporters with drugs, prodrugs, alimentary and Traditional Chinese Medicine compounds, but does not have information of the products (Hoffmann et al. 2014).

There are some other databases that predominantly focused on the primary metabolism along with endogenous metabolites, like METLIN, and Human Metabolome Database (HMDB). METLIN actually commenced as a database for characterising known metabolites and now contains a large amount of high-resolution tandem mass spectrometry (MS/MS) data. The Human Metabolites Database

(HMDB) is another freely accessible electronic database that contains detailed information about small molecule metabolites found in the human body (Smith 2005; Wishart 2007).

In addition, some databases spend more effort on collecting data on pharmacokinetics parameters, including clearance rates, and half-lives. PharmacoKinetics Knowledge Base (PKKB) developed by Soochow University is an example. PKKB collected 10,000 experimental ADMET measurements of 1685 drugs from rat and mouse, including Caco-2 permeability, octanol/water partition coefficient, human bioavailability, solubility, dissociation constant, intestinal absorption (Cao 2012). Company Cerep also created a commercial database for pharmacokinetic information, which is BioPrint (Krejsa et al. 2003).

Table 2.3 Databases and datasets with metabolism-related content.

Name	Company/institution	Availability	Compound Coverage	Species Coverage	Key feature	Metabolizing enzymes	Biotransformation data	PK parameters	Drug transporters	Bioactivities	Ref
ADME DB	Fujitsu	Commercial	Drug	Human	Pharmacokinetic information	√		√	√	√	("ADME database (Fujitsu Kyushu Systems Ltd, 2014),")
AurSCOPE ADME	Aureus Sciences	Commercial	Drug	Human	ADMET properties	√					(Peach et al., 2012)
BioPrint	Cerep	Commercial	Drug	Human	Pharmacokinetic information					√	(Krejsa et al., 2003)
ChEMBL	European Bioinformatics Institute	Free	Bioactive molecules	Human	Bioassay data	√			√	√	(Gaulton, 2011)
DrugBank	University of Alberta	Free	Drug	Human	Pharmaceutical information	√	√	√	√	√	(David S. Wishart et al., 2018)
MetaBase	GeneGo	Commercial	Drug	Human	Biological properties and effects of drugs and drug-like molecules	CYP only	√				("Metabase (Thomson Reuters, 2013),")
Metabolite	Accelrys	Commercial	Drug	Mammal	Biotransformation data		√				(*Accelrys*

											Metabolite Database, 2011)
Microsomal Stability	Evolvus	Commercial	Drug	Human	Microsomal stability			√			("ADME database (Fujitsu Kyushu Systems Ltd, 2014),")
PubChem	NCBI, NIH	Free	Full coverage	Human	Comprehensive chemical information	√	√	√	√	√	(Kim et al., 2019)
QSAR World	Strand Life Sciences	Free	Drug	Human	Clearance, half-life of drugs			√			("ADME database (Fujitsu Kyushu Systems Ltd, 2014),")
WOMBAT-PK	Sunset Molecular Discovery	Commercial	Drug	Human	Pharmacokinetic data			√			(Olah et al., 2007)
BindingDB	University of California San Diego	Free	Drug	Human	Bioactivity focused on drug targets	√	√	√	√	√	(T. Liu et al., 2007)
Drug Database (GOSTAR)	Excelra	Commercial	Drug	Human	Comprehensive resource for metabolites of approved drugs	√	√	√		√	("GOSTAR drug

											database (GVK Biosciences Private Limited, 2013),")
EAWAG-BBD	The University of Minnesota	Free	Environmental chemicals	Microbe	Biodegradation of xenobiotics	√	√				(Gao et al. 2011)
Human Metabolome Database (HMDB)	The Metabolomics Innovation Centre, CA	Free	Endogenous Metabolite	Human	Human endogenous metabolites	√	√	√	√	√	(Wishart et al. 2007)
KEGG	Kyoto University	Free	Drug	Human	Large database comprising ~20 different sub-collections of data on metabolism	√	√		√	√	(Kanehisa et al. 2012)
Metabolism and Transport Drug Interaction Database (DIDB)	The University of Washington	Free	Drug	Human	Drug–drug interactions in humans			√	√		(Hachad et al. 2010)
METLIN	The Scripps Research Institute	Free	Endogenous Metabolite	Human	Metabolite Mass Spectra		√				(Smith 2005)
PharmGKB	Stanford University	Free	Drug	Human	Encyclopaedia focused on drugs	√	√	√	√	√	(Hewett 2002)
PKKB	Soochow University	Free	Drug	Rodent	Pharmacokinetic data	√	√	√	√	√	(Cao, 2012)

Protein Data Bank (PDB)	US governmental agencies	Free	Full coverage	Human	Structures of biomolecules	√	√	√	√	√	(J. Kirchmair et al., 2008)
SuperCyp	Charité—University Medicine Berlin	Free	Drug	Human	Drug–CYP interactions	CYP only					(Preissner, 2010)
MetXBioDB	University of Alberta	Free	Drug	Human	Biotransformation data	√	√				(Djoumbou-Feunang et al., 2019)

2.3. Biotransformation prediction models

Here, this study classifies current computational approaches to predict xenobiotic metabolism into three major categories: (1) prediction of sites of metabolism (SOMs), (2) prediction of metabolites structures, and (3) prediction of metabolizing enzymes binding (Kirchmair et al. 2012). As the first two methods have a broader range of applications, this study will go over them in greater depth subsequently.

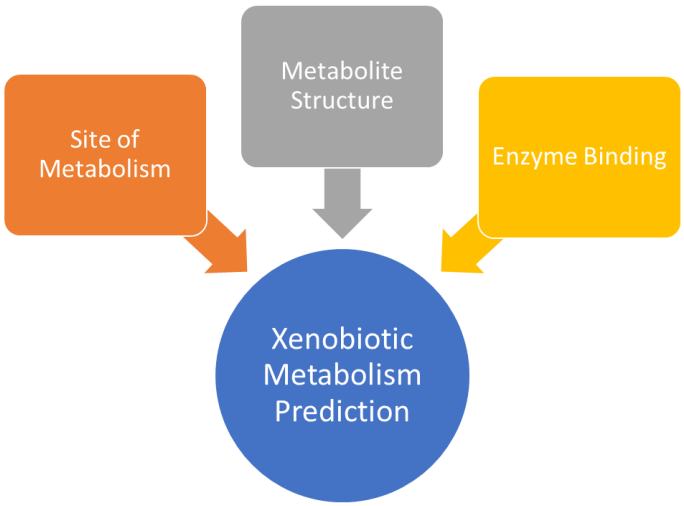


Figure 2.5 Overview on categories of computational approaches to predict xenobiotic metabolism.

2.3.1 Sites of metabolism

Identifying a molecule's SOMs may provide crucial hints in the development of a medicinal chemistry strategy to optimize metabolic properties and, as a result, essential parameters such as toxicity, bioavailability, and bioactivity. SOMs are located primarily based on the chemical reactivity and orientation of the ligand bound to the enzyme's catalytic site. Nowadays, a great number of computational approaches are available, which attempt to identify the most likely SOMs. The details of SOM prediction model are summarized in Table 2.4. Methods used to determine reactivity include approaches that use fingerprint-based data mining, shape-focused methods and molecular interaction fields, protein–ligand docking, and combinations of these. SOMs or potential metabolites are typically predicted using techniques that focus on a single aspect of metabolic reactions, such as the reaction

energy barrier, geometrical properties, or pharmacokinetic properties (Kirchmair et al. 2012).

Table 2.4 Computational methods for predicting sites of metabolism.

Software	Key Component	Coverage	Licensing	Ref
QMBO	Reactivity-based method	CYP3A4, CYP2C9	Free	(Afzelius et al. 2007)
CypScore	Surface electrostatics+ semi-empirical method	Individual CYP reactions	Commercial	(Hennemann et al. 2009)
Metaprint2D	Fingerprint-based data mining method	Phase I & II	Free	(Carlsson et al. 2010)
ROCS	Shaped-focused method	CYP2C9	Free	(Sykes et al. 2008)
MetaSite	Molecular interaction fields+ reactivity estimator	Variety of CYPs	Commercial	(Cruciani et al. 2005)
SMARTCyp	DFT-derived reaction energies	Variety of CYPs	Free	(Rydberg et al. 2010)
StarDrop	Semiempirical method	CYP2C9, CYP2D6, CYP3A4	Commercial	(*StarDrop* 2011)
RS-Predictor	Descriptors + SVM-like ranking	Variety of CYPs	Free	(Zaretzki et al. 2011)
MLite	Docking+ reactivity estimator	CYP3A4	Free	(Oh et al. 2008)
IDSite	Induced-fit docking+ reactivity estimator	CYP2C9, CYP2D6, CYP3A4	Commercial	(Li et al. 2011)
MEXAler	Reaction rules	Phase II	Commercial	("MEXAlert (CompuDrug International Inc., 2001),")
QikProp	Reaction rules	~20 phase I reactions	Commercial	("QikProp (Schrödinger LLC 2014),")
ADMET Predictor Metabolism module	ANN ensemble	Variety of CYPs	Commercial	(*ADMET Predictor, Metabolite Software Module* 2011)
Percepta P450 Regioselectivity module	PLS (GALAS)	Human liver microsomal metabolism and CYP1A2, CYP2C9, CYP2C19, CYP2D6, CYP3A4	Commercial	("Percepta Platform (ACD/Labs Inc. 2014),")

RSWebPredictor	MIRank (SVM)	Variety of CYPs	Free	(Zaretzki 2013)
FAME	Random forest	Phase I & II	Free for academic use	(Kirchmair 2013)
XenoSite	Neural Network	Variety of CYPs	Free	(Zaretzki et al. 2013)

2.3.2 Prediction of metabolites

Computational approaches to predict metabolites of small organic molecules are also important and these approaches rely heavily on experimental data. Many computational methods have been developed to predict SOMs, but few have been developed to predict xenobiotic metabolic products. Current prediction methods are dominated by knowledge-based systems, which integrate common biotransformation reaction templates in SMIRKS format. Almost all mainstream prediction models utilized reaction rules, such as SyGMa, Biotransformer, EAWAG-BBD(Djoumbou-Feunang et al. 2019; Ridder and Wagener 2008). The reaction rules were developed based on the general understanding of metabolic processes or a visual assessment of commonly occurring reactions in the experimental dataset (Ridder and Wagener 2008). For example, Figure 2.6 illustrated that the reaction of oxidation of primary alcohol was encoded into a SMIRKS reaction as "[*:1][CH2][OH1]>>[*:1][C:2](=O)O", where * represents for any chemical group, [CH2] represents for carbon with two hydrogen atoms, [OH1] represents for oxygen with one hydrogen atoms, =O represents for oxygen with double bond.

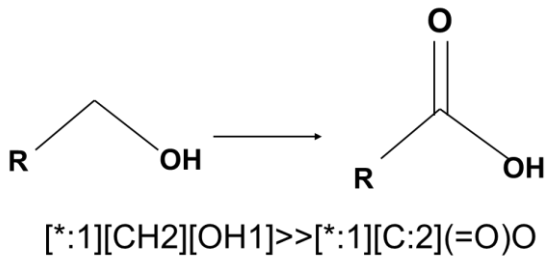


Figure 2.6 A general rule for oxidation of a primary alcohol (Ridder and Wagener 2008).

All these reaction rules were hand-encoded by experts based on their experience and knowledge. SyGMa utilized 118 phase 1 knowledge-based reaction rules and 27 phase 2 knowledge-based reaction rules (Ridder and Wagener 2008). Djoumbou-Feunang *et al* developed another freely available tool for metabolite prediction, which is BioTransformer with 163 CYP rules and 74 phase 2 rules (Djoumbou-Feunang et

al. 2019). In addition, another freely available tool for metabolite prediction called GLORYx was established with 261 reaction rules.

Recently, some new approaches utilizing machine learning techniques have been built up for prediction of drug metabolites. MetaTrans developed by Litsa et al. is the first prediction model that utilized neural machine translation, which was adapted from a transformer model for organic synthesis (Litsa et al. 2020; Schwaller et al. 2019). Unlike traditional rule-based approaches, MetaTrans treated metabolite prediction as a sequence translation problem with chemical compounds represented using the SMILES notation thus bypassing the reaction rules. Besides, CypReact employed machine learning technique in another way, which is the classification model. It took molecular structure as input and trained the classification model into 9 categories that include nine kinds of important human cytochrome P450 (Tian et al. 2018).

Table 2.5 Computational methods for predicting metabolite structures.

Software/Service	Key Component	Coverage	Licensing	Ref
MetabolExpert	Knowledge based system	Phase I & Phase II	Commercial	(Darvas and Kaiser 1987)
Meteor	Knowledge based system	Phase I & Phase II	Commercial	(Marchant et al. 2008)
SyGMa	Knowledge based system	Phase I & Phase II	Free to academia	(Ridder and Wagener 2008)
TIMES	Knowledgebased system	Phase I & Phase II	Commercial	(Mekenyan et al. 2004)
JChem Metabolizer module	Knowledge based system	Phase I & Phase II	Commercial	(*JChem, Metabolizer Software Module* 2011)
Metaprint2D-React	Fingerprint-based data mining approach	Phase I	Commercial	(Carlsson et al. 2010)
MetaSite	Molecular interaction fields & reactivity estimator	Variety of CYPs	Commercial	(Cruciani et al. 2005)

Meta-PC	Knowledge based system	Phase I & Phase II	Commercial	(Klopman et al. 1994)
MetaDrug	Knowledge based system	Phase I & Phase II	Commercial	(Ekins 2006)
EAWAG-BBD Pathway Prediction System	Knowledge based system	Phase I & Phase II	Free	(Gao et al. 2010)
Biotransformer	Knowledge based system	Phase I & Phase II	Free	(Djoumbou-Feunang et al. 2019)
GLORYx	Knowledge based system	Phase I & Phase II	Free	(Bruyn Kops et al. 2020)
MetaTrans	Neural machine translation	Phase I & Phase II	Free	(Litsa et al. 2020)
MetaTox	Knowledge based system	nine classes of reactions	Free	(Rudik et al. 2017)
CypReact	Machine learning approach	Phase I	Free	(Tian et al. 2018)

2.4. Environmental Exposome

2.4.1 Introduction of exposome

Humans are inevitably exposed to a multitude of small molecules that are foreign to the body (xenobiotics), including dietary components, environmental chemicals, and pharmaceuticals (Koppel et al. 2017). In this context, the term "environmental exposome" has been used to refer to all environmental exposures (internal and specific/general external) that individuals encounter from conception until the end of their lives. In our previous work, more than 20,000 chemicals were collected to compile a human exposome database (Zhao et al. 2020). Environmental and food chain contaminated by these chemicals can result in adverse health outcomes for humans and other animals. More recent industrial and agricultural chemicals, pharmaceuticals, and personal care products are not generally persistent, but they are ubiquitous due to their widespread use and global sources. The abiotic and biotic transformation products may be more persistent and occur at higher concentrations than their parent compounds (Fenner et al. 2013). Simultaneous exposure to multiple compounds increases the risk of interaction, which we will discuss in detail in subsequent section 2.4. Additionally, part of this study focused on phenolic xenobiotics and gut microbiota-derived metabolites, which details will also be reviewed in the later sections 2.4.2 and 2.4.3, respectively.

2.4.2 Phenolic xenobiotics

Phenolic chemicals are a class of chemical compounds consisting of a hydroxyl group (-OH) bonded directly to an aromatic hydrocarbon group. Phenolic xenobiotics, such as Bisphenol A (BPA), Bisphenol S (BPS), Triclosan (TCS), and Tetrabromobisphenol A (TBBPA) are an important group of ubiquitous environmental contaminants with wide detection in environment and human samples (Chen et al. 2016; Harrad et al. 2009; Yueh and Tukey 2016). Most phenolic xenobiotics are generally thought to be rapidly detoxified to glucuronide and sulfate conjugates by phase II metabolic enzymes (Ashrap et al. 2017; Klaassen et al. 2013). Due to the similar structure, previous studies show that BPA, BPS, TCS and TBBPA share common metabolic pathways, with major metabolites of sulfate and/or

glucuronide conjugates detected in human serum and urine samples (Gerona et al. 2013; Ho et al. 2017; Oh et al. 2018; Provencher et al. 2014).

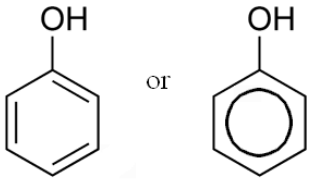


Figure 2.7 Structure of phenol

Bisphenol A (BPA) is the common name for 2,2-(4,4′-dihydroxydiphenyl) propane, 4,4′-isopropyllidenediphenol, alternatively, 2,2′-bis(4-hydroxyphenyl) propane, an organic compound with two phenol moieties. BPA is used in the synthesis of polycarbonates, epoxy resins and thermal paper. The global demand for BPA is predicted to grow from 3.9 million tons in 2006 to about 5 million tons in 2010 (Huang et al. 2012). As a consequence, its occurrence has been reported in nearly all tested samples including urine of adults and children, serum of pregnant women, breast milk, follicular and amniotic fluid, cord blood and placental tissue, and human foetal livers (Rochester 2013; Vandenberg et al. 2010). Besides, many adverse health outcomes of BPA have been reported, including endocrine disruption, obesity, diabetes, hepatotoxicity, neurotoxicity, and immunotoxicity (Michalowicz 2014; Prins et al. 2018). Moreover, one of BPA metabolites, 4-methyl-2,4-bis(4-hydroxyphenyl)pent-1-ene is reported to be up to 1,000 times potent as an estrogen than BPA.(Moreman et al. 2018)

The increasing health concerns over BPA have facilitated the replacement of BPA from consumer products with substitutes like Bisphenol S (BPS). BPS is used for a variety of industrial applications, for example, as a wash fastening agent in cleaning products, an electro plating solvent, and a constituent of phenolic resin. BPS is also used as a developer in thermal paper, including products marketed as “BPA-free paper” (Rochester and Bolden 2015). Its detected concentrations and frequencies in urine samples have been comparable to those of BPA (Zhou et al. 2014). The *in vitro* literature indicates that BPS has actions and potencies similar to those of BPA and supports the biological plausibility of their hormonal activity *in vivo*. This is not surprising because BPS is structural analogue of BPA and thus mechanisms of action

would be expected to be similar. BPS exhibits multidirectional toxic effects on reproductive systems, endocrine systems, and nervous systems in mammals, and may also be an incentive of oxidative stress (Rochester and Bolden 2015).

Triclosan (TCS), which is 5-chloro-2-(2,4-dichlorophenoxy) phenol, was first launched into health care industry around 1970. TCS is a synthetic lipid-soluble antimicrobial agent that has been used globally as a disinfectant, antiseptic or preservative in clinical settings, personal care products, household item, and medical devices. It was recently phased out in US due to the concerns of hepatotoxic, mutagenic and carcinogenic effects (Administration 2016; Yueh and Tukey 2016).

Tetrabromobisphenol A (TBBPA) is one of the most extensively used flame retardant, with global use over 170,000 metric tons/year. TBBPA as a persistent compound has been detected in soil, sediment, sludge, surface water, dust and air. Major TBBPA exposure pathways for humans include breast milk, diet and indoor air and dust (Malkoske et al. 2016). The current data suggests that TBBPA may be an endocrine disruptor and immunotoxicant and is currently under serious investigation in EU ("Scientific Opinion on Tetrabromobisphenol A (TBBPA) and its derivatives in food," 2011).

In addition, there are some other phenolic compounds with various applications are frequently detected in human and environmental samples. The information is summarized at Table 2.6.

Table 2.6 Summary of basic information of 22 compounds in the phenolic xenobiotic mixtures and their maximum (MAX), geometric mean (GM) concentrations, sample type and size.

No.	Compound Name	Cas-No.	Uses	Concentration in Urine (ng/mL) or Blood (ng/mL)		Sample Type	Sample Size (n)	Ref
		131-56-6	Ultraviolet (UV) light absorber	Maximum (MAX)	Geometric Mean (GM)			
1	2,4-dihydroxybenzophenone(DHB)	80-09-1	Plasticizer	1911.7	21.9	Urine	15593	(Prevention 2018)
2	Bipshenol S (BPS)	101-20-2	Antiseptic in detergents and plastics.	4.7	0.427	Urine	2682	(Prevention 2018)
3	Triclocarban	3380-34-5	Antiseptic	24.5	0.200[c]	Urine	2686	(Prevention 2018)
4	Triclosan (TCS)	94-26-8	Antiseptic and additives in cosmetics	599.7	13.8	Urine	15593	(Prevention 2018)
5	Butyl Paraben	120-47-8	Antifungal preservative and food addictive.	18.28	1.16[c]	Urine	13076	(Prevention 2018)
6	Ethyl Paraben	99-76-3	Anti-fungal agent and food preservative.	99.62	5.44[c]	Urine	13076	(Prevention 2018)
7	Methyl Paraben	13674-87-8	Flame retardants, pesticides, plasticizers, and nerve gases.	1056.6	52.4	Urine	13076	(Prevention 2018)
8	Tris(1,3-dichloroisopropyl)phosphate (TDCPP)[a]	115-86-6	Flame retardant and plasticizer.	8.60	0.856	Urine	2646	(Prevention 2018)
9	Triphenyl phosphate (TPhP)	84-74-2	Plasticizer.	16.57	2.71	Blood	69	(Zhao et al. 2017; Li et al. 2017; Li et al. 2015)
10	Di-n-butyl phthalate (DBP)	131-70-4	Metabolite of di(n-butyl) phthalate	13.87	3.45	Blood	184	(Frederiksen et al. 2010; Högberg et al. 2007; Li et al.

								2018)
11	Mono-n-butyl phthalate (MnBP)	85-68-7	Plasticizer	110.2	16.8	Urine	21003	(Prevention 2018)
12	Benzylbutyl phthalate (BzBP)	4376-20-9	Urinary phthalate metabolite.	232.8	16.7	Urine	18462	(Prevention 2018)
13	Mono-2-ethylhexyl phthalate (MEHP)	2528-16-7	Used as plasticizers	31.3	2.70	Urine	21003	(Prevention 2018)
14	Mono-benzyl phthalate (MBzP)	28553-12-0	Used to make plastics more flexible	70.0	8.30	Urine	21003	(Prevention 2018)
15	Di-isononyl phthalate (DiNP)	26761-40-0	Plasticizer	13.8	2.00[c]	Urine	21003	(Prevention 2018)
16	Di-isodecyl phthalate (DiDP)	446-72-0	An isoflavone	20.9	2.67	Urine	13075	(Prevention 2018)
17	Genistein	335-67-1	Used in fire-fighting applications, cosmetics, greases and lubricants, paints, polishes and adhesives	781.3	29.7	Urine	15689	(Prevention 2018)
18	Perfluorooctanoic acid (PFOA)	1675-54-3	Used as constituent of epoxy resins	10.7	3.50	Blood	14178	(Prevention 2018)
19	Bisphenol A Diglycidyl Ether (BADGE)	5581-32-8	A component of commercial liquid epoxy resins commonly used in the food-packing industry and in dental sealants	2.50	0.500	Blood	223	(Kim et al. 2015)
20	Bisphenol A bis(2,3-dihydroxypropyl) ether (BADGE-2H_2O)	2095-03-6	Used as a material of interior coating for food cans	9.50	2.30	Blood	243	(Kim et al. 2015; Wang et al. 2015)
21	Bisphenol-F-diglycidyl ether (BFDGE)	79-94-7	Flame retardant in electronics, paper, plastics, textiles, carpeting and furniture.	180	26.2	Blood	20	(Wang et al. 2015)
22	Tetrabromobisphenol A (TBBPA)			7.58	0.68	Blood	265	(NAGAYAMA et al. 2000; Thomsen et al. 2002; Xiao et al. 2011)

2.4.3 Gut microbiota derived metabolites

The human gut microbiota is a diverse and complex microbial community that interacts with the host and is closely related to human biology. The estimated 10^{13} microbes that inhabit the human gastrointestinal (GI) tract make key contributions to many processes, including training intestinal epithelial barrier, modulating immuno-function, digesting host indigestible nutrients, producing vitamins and hormones and preventing pathogenic bacterium colonization (Koppel et al. 2017). A considerable number of xenobiotic responsive genes have been identified among multiple gut bacterial phyla, including those which encode antibiotic resistance, drug metabolism, and stress responses. It has been revealed that the metabolites produced by the gut microbiome play an important role in the microbiome-host interactions (Wikoff et al. 2009).

Besides, GI tract is one of the most important sites where toxicants are absorbed (Klaassen et al. 2013). The gut microbiome works as the "first-pass" of external stressors (e.g., diet and environmental chemicals) before entry into the internal circulation. The first pass effect is a phenomenon of drug metabolism whereby the concentration of a drug is greatly reduced before it reaches the systemic circulation. It is the fraction of drug lost during the process of absorption which is generally related to the liver and gut wall.

In addition, many metabolites from the gut microbiome have increasingly been identified as important modulators of human biological processes, and any dysregulation on those metabolites could induce some adverse effects (Zhang et al. 2018). A recent research reported that gut microbiome-encoded enzymes can directly and significantly impact intestinal and systemic drug metabolism (Zimmermann et al. 2019). Additionally, a human gut metabolite, *p*-cresol, was reported to reduce the effective systematic capacity of drug acetaminophen by competitive O-sulfonation (Clayton et al. 2009). Moreover, previous study reported that a large numbers of gut microbiota derived metabolites undergo phase I and phase II metabolism, including indole derivatives, phenyl derivatives, flavones and others like dihydroxyquinoline and 3-Carboxy-4-methyl-5-pentyl2-furanpropionic acid (Wikoff et al. 2009). These metabolites after xenobiotic biotransformation have been identified in human blood samples and the detected concentrations are quite noticeable.

Table 2.7 Gut microbial metabolites and host health implications (Rajapaksha Pathiranage Tharushi Prabha, 2021).

Category	Subcategory	Metabolite	Molecular functions	Host health implications
Amino acid-related	Tryptophan and indole derivatives	Indolepropionic acid	Pregnane X receptor (PXR) Agonist (Liu et al., 2020) AHR agonist (Liu et al. 2020)	Immunomodulatory (Donia and Fischbach 2016) Promotes intestinal epithelial barrier function ((Liu et al. 2020) Neuroprotective (Liu et al. 2020) Antioxidant (Liu et al. 2020)
		Indole	AHR antagonist (Jin et al. 2014) Modulates Glucagon-like peptide (GLP)-1 secretion (Zhang and Davies 2016)	Maintains host-microbe homeostasis at mucosal surface (Zhang and Davies 2016)
		Indole-3-aldehyde	AHR agonist (Zhang and Davies 2016)	Maintains host-microbe homeostasis at mucosal surface (Zhang and Davies 2016)
		Skatole	AHR agonist (Liu et al. 2020)	Na
		Tryptamine	AHR agonist (Hubbard et al. 2015)	Neurotransmitter (Donia and Fischbach 2016)
		Indole-3-acetic acid	AHR agonist (Liu et al. 2020) Inhibits microglial activation (Liu et al. 2020)	Anti-inflammatory effects in nonalcoholic fatty liver disease (Liu et al. 2020) Renal toxicity (Liu et al. 2020)
		Tryptophan	Regulates Ca2+-sensing receptor (CaSR) signaling pathway (Tan et al. 2017)	Inflammatory bowel diseases (Tan et al. 2017) Major depression, chronic brain injury, amyotrophic lateral sclerosis (Dehhaghi et al. 2019)
		Serotonin /5-hydroxytryptamine	Regulation of adipose tissue energy storage and expenditure (Cervenka et al. 2017) Brain-mediated appetite control (Cervenka et al. 2017)	Diabetes (Cervenka et al. 2017) Fat accumulation in adipose tissues (Cervenka et al. 2017) Depression, fatigue, and impaired cognitive function (Kaur et al. 2019) Neurotransmitter (Kaur et al. 2019)
		5-Hydroxy-L-tryptophan	AHR agonist (Bittinger et al. 2003)	NA

	Xanthurenic acid	AHR agonist (Hubbard et al. 2015)	NA
	Melatonin	Stimulation of IL-2, IL-6, and IL-12 production (Srinivasan et al. 2005) Decrease the activity of natural killer (NK) cells, granulocytes, and macrophages (Srinivasan et al. 2005)	Immunomodulatory (Srinivasan et al. 2005)
	Formylkynurenine	NA	NA
	Kynurenine	AHR agonist (Cervenka et al. 2017)	Alzheimer disease, anxiety, amyotrophic lateral sclerosis, Schizophrenia (Dehhaghi et al. 2019)
	Kynurenic acid	N-methyl-D-aspartate receptor (NMDAR) (Cervenka et al. 2017) a7 nicotinic acetylcholine receptor (a7nAChR) antagonist (Cervenka et al. 2017) GPR35 agonist (Cervenka et al. 2017) AHR agonist (Cervenka et al. 2017)	Modulate local inflammation (Cervenka et al. 2017) Alzheimer disease, anxiety, amyotrophic lateral sclerosis, Schizophrenia (Dehhaghi et al. 2019)
	3-Hydroxy-L-kynurenine	NA	NA
	Quinolinic acid/ Pyridinedicarboxylic acid	NMDAR agonist (Cervenka et al. 2017) TH1 target cells' apoptosis (Dehhaghi et al. 2019) Selectively suppress the proliferation of killer cells, and CD4+ and CD8+ T lymphocytes (Dehhaghi et al. 2019)	Alzheimer disease, dementia complex, Huntington disease, and multiple sclerosis (Dehhaghi et al. 2019)
	Indole-3-acetaldehyde	NA	NA
	Indole-3-acetaldoxime	NA	NA
	Indole-3-acetonitrile	NA	NA
	Indole-3-pyruvic acid	NA	NA
	Indolelactic acid	NA	Ameliorates salt-sensitive hypertension (Liu et al. 2020)
	Indoleacetyl glutamine	NA	NA

Tyrosine and phenyl derivatives	Phenyllactic acid	Highly potent activation of hominids contain a third member (HCA3) (Peters et al., 2019)	NA
	Phenethylamine	Agonist of dopamine receptor D2 (DRD2)/4 (Liu et al. 2020)	Neurotransmitter (Donia and Fischbach 2016)
	Phenylacetic acid	NA	Toxic to colonocytes (Liu et al. 2020) Triggers steatosis (Liu et al. 2020)
	4-ethylphenyl sulfate	NA	Autistic spectrum disorder (Zhang and Davies 2016)
	P-cresol	NA	Induce genotoxic effects on colonocytes (Saito et al. 2018)
	P-cresol sulfate	Activates nicotinamide adenine dinucleotide phosphate (NADPH) oxidase (Zhang and Davies 2016) Th1-type cellular immune responses (Shiba et al. 2014)	Damages cell membranes (Zhang and Davies 2016) Induces apoptosis (Zhang and Davies 2016)
	Hippuric acid	NA	Metabolic syndrome (Pallister et al. 2017)
	3-phenylpropionate /Hydrocinnamate	Inhibitor of branched-chain α -keto acid dehydrogenase (Tso et al. 2013)	Positively correlate with gut microbial diversity (Pallister et al. 2017)
	Phenylacetylglutamine	NA	Cardiovascular disease (Wilmanski et al. 2019)
	Cinnamoylglycine	NA	Obesity (Wilmanski et al. 2019)
	Phenylpropionylglycine	NA	
	Phenylacetylglycine	NA	
	4-Hydroxyphenylacetic acid	NA	Genotoxic effects on human enterocytes (Liu et al. 2020)
	Tyramine	Agonist of trace amine-associated receptors (TAAR)1 (Liu et al. 2020)	Neuromodulators (Liu et al. 2020) Vasodilation (Liu et al. 2020)
	Trans-Cinnamate/Cinnamic acid		Regulates glucose transport (Liu et al. 2020) Antitumor (Liu et al. 2020)
	Tyrosine	Regulates CaSR signaling pathway (Tan et al. 2017)	Inflammatory bowel diseases (Tan et al. 2017)

		Coumaric acid/4-Hydroxycinnamate	NA	NA
		Phenylalanine	Regulates CaSR signaling pathway (Tan et al., 2017)	Inflammatory bowel diseases (Tan et al. 2017)
	BCAAs	L-leucine	NA	Insulin resistance (Pedersen et al. 2016) Type ii diabetes (Pedersen et al. 2016)
		Isoleusine	NA	Insulin resistance (Pedersen et al. 2016) Type ii diabetes (Pedersen et al. 2016)
		L-valine	NA	Insulin resistance (Pedersen et al. 2016) Type ii diabetes (Pedersen et al. 2016)
	Others	5-aminovaleric acid	NA	
		γ-aminobutyric acid (GABA)	NA	Neurotransmitter (Bhat et al. 2010)
		3-aminoisobutyric acid	NA	
		Imidazole propionate	NA	Insulin resistance (Koh et al. 2018)
Lipids and fatty acids (FAs)	SCFAs	Acetate	Activates GPCR41 and GPCR43 (Zhang and Davies 2016)	Anti-inflammatory (Besten et al. 2013) Improves insulin sensitivity (Besten et al. 2013)
		Propionate	Activates GPCR41 and GPCR43 (Zhang and Davies 2016) Upregulates GLP-1, Peptide YY (PYY), leptin (Zhang and Davies 2016)	Anti-lipogenic (Besten et al. 2013) Increases energy expenditure (Besten et al. 2013)
		Butyrate	Activates GPCR41 and GPCR43 (Zhang and Davies 2016) Activates GPCR109A (Zhang and Davies 2016) Modulates peroxisome proliferator-activated receptor (PPAR)-γ (Zhang and Davies 2016)	
	Trimethylamine N-oxide	Trimethylamine (TMA)	NA	NA

(TMAO) related	TMAO	NA	Cardiovascular disease (Wang et al. 2011)
FAs	Conjugated linoleic acid (CLA)	NA	Body weight gain (Mensink 2005) Cardiovascular disease (Mensink 2005)
	Conjugate linolenic acids (CLnAs)	NA	Insulin resistance (Risérus et al. 2004)
	10-hydroxy-cis-12-octadecenoate	NA	
Bile acids	Deoxycholic acid (DCA)	TGR5 agonist (Wahlström et al. 2016) FXR agonist (Jia et al. 2018)	Cholesterol gallstone disease (Ridlon et al. 2014) Colon cancer (Ridlon et al. 2014) Obesity-associated hepatocellular carcinoma (Ridlon et al. 2014)
	Lithocholic acid (LCA)	FXR agonist (Jia et al. 2018)	Colon cancer (Mcgarr et al. 2005)
	Ursodeoxycholic acid (UDCA)	Farnesoid X receptor (FXR) antagonist (Wahlström et al. 2016)	NA
	Glyco-cholic acid	FXR agonist (Wahlström et al. 2016)	NA
	Tauro-cholic acid	FXR agonist (Wahlström et al. 2016)	NA
	Glyco-chenodeoxycholic acid	FXR agonist (Wahlström et al. 2016)	NA

2.5 Chemical-chemical interactions

Chemicals in a complex mixture can contribute to toxicity even if their concentration is lower than the effect threshold and/or analytical detection limit. Chemicals with similar modes of toxic action tend to adhere to the mixture concept of "concentration addition," whereas those with different modes of action adhere to the concept of "independent action" (Escher et al. 2020). Among all the chemical-chemical interactions, drug-drug interactions (DDIs) have been attracting the most attention as it is critical to minimize unexpected adverse drug events (ADEs) and to maximize synergistic benefits when treating a disease. Prior research revealed that thirty percent of all reported ADEs were associated with DDIs. Moreover, ADEs caused by DDIs are one of the major reasons for drug withdrawal from the market (Ryu et al. 2018).

The mechanisms of DDIs are commonly classified into three classes: pharmacokinetic, pharmacodynamic, and pharmaceutical incompatibility. The relationship between drug concentration and drug response is described as pharmacodynamics whereas the relationship between the rates of change of drug concentrations in different parts of the body is described as pharmacokinetics. Pharmacodynamic interactions occur when drugs with similar or opposing pharmacological effects interact with one another. Competition at molecular or cellular sites of action is one of the underlying mechanisms. The third category includes pharmaceutical interactions caused by chemical or physical incompatibility between the drug preparations used, with calcium chloride and sodium bicarbonate preparations being an example of such an interaction (Corrie and Hardman 2017).

Pharmacokinetic interactions that occur when one drug alters the absorption, distribution, metabolism or excretion of the other drug, are the primary causes of unexpected drug interactions. This type of response varies between patients without any particular drug pattern and can be difficult to predict. Among several enzyme families involved in drug metabolism, the most important one is the cytochrome P450 (CYP) enzyme family. Inhibition of a cytochrome P450 enzyme improves the concentration of some drugs by impeding their biotransformation. For instance, clarithromycin is a strong inhibitor of CYP3A-catalysed simvastatin metabolism,

thus enhancing the risk of myopathy. Meanwhile, inhibition of cytochrome P450 enzymes can also be used medicinally. For instance, another strong inhibitor of CYP3A that named ritonavir obstructs metabolism of other protease inhibitors thus rising their effectiveness in treating HIV. Conversely, the concentration of some drugs could be declined by enhancing their metabolism if the cytochrome P450 enzyme is inducted. For instance, carbamazepine is a strong inducer of CYP3A that enhances the metabolism of the combined oral contraceptive, thus increasing the risk of unwanted pregnancy (Juurlink et al. 2003).

From an analysis on pharmacokinetic-based drug-drug interaction (DDI) data for drugs approved by the U.S. Food and Drug Administration in 2017, researchers found that the observed medication interactions could be explained by CYP3A inhibition and induction for the most part. Two-thirds of the DDIs were metabolized by CYP3A, which is followed by CYP2D6, CYP1A2, and the CYP2C families. For drugs as inhibitors of enzymes, inhibitors of CYP3A accounted for the greatest number of medications (n = 16), followed by CYP2C9, CYP1A2, CYP2C8, CYP2D6, CYP2C19, and CYP2B6. For drugs as inducers of enzymes, there were similar amounts of drugs as inducers of CYP3A (n = 10), CYP2B6 (n = 10), and CYP1A2 (n = 9) (Yu et al. 2019).

Additionally, it has been reported that a human gut metabolite, *p*-cresol, reduced the effective systematic capacity of drug acetaminophen by competitive O-sulfonation (Clayton et al. 2009). Besides, food-drug, and herb-drug interactions are also big concerns. For example, grapefruit juice possesses high interaction with almost all types of drugs, which occurs via inhibition of CYP3A enzymes. Furthermore, a number of studies have been documented on the interaction of warfarin and cranberry juice. Tamoxifen is a successful anti-tumor agent. But co-administration with sesame seeds has negative effect on inducing regression of established MCF-7 tumor size (Bushra et al. 2011).

Chapter 3. EcoBioTrans: A Comprehensive Database Utilizing Multi-Species Biotransformation Data

3.1 Overview

Through modifying their lifetimes, bioavailabilities, and biological effects, xenobiotic biotransformation can play a principal role in the activation, detoxification, and elimination of foreign chemicals including environmental pollutants, pharmaceuticals, and food additives. Over the past several decades, scientific efforts have been mainly invested into monitoring studies and prediction studies on this field; however, a comprehensive understanding on biotransformation has not been achieved. Extant studies focus on metabolites identification due to the maturity of analytic method and researchers' easier access to analytic tools. To broaden our research scope, informatics approaches for systems-level data are needed. To date, researchers do not have an openly available high-quality biotransformation database that incorporates substance with an adequate coverage from multiple species. In this work, we set up an open-access data portal, EcoBioTrans (https://www.ecobiotrans.asia)[2], which is an open access database harboring biotransformation data of common environmental pollutants, along with pharmaceuticals and human endogenous metabolites. To establish this data repository, we collected 7,800 biotransformation reactions of 1,599 substrates by text mining and database fusion. The collected species are not restricted to human, but also extended to bird, fish, rodent, and other mammals. This database contains structures of substrates and their corresponding products, types of reactions, the names or types of enzymes catalyzing biotransformation, types of biosystems, and references. Users are also allowed to search biotransformation results based on structural similarity. Overall, for the first time, environmental pollutants and multiple species are targeted as focus for biotransformation data curation. This integrative platform makes a great contribution to the scientific understanding of biotransformation path in the whole ecosystem and serves as a guide for chemical

[2] Please use the server address (http://1.117.57.232/) to view the website as our server is being audited and certified.

management.

3.2 Introduction

Biotransformation is crucial for the activation, detoxification, and elimination of xenobiotics, such as environmental pollutants, pharmaceuticals, food additives, and biologically foreign substances. Through biotransformation, xenobiotics will be converted into metabolites with different physicochemical and pharmacological properties. Specific enzymes like cytochrome P450s and UDP glucuronosyltransferases are commonly employed in biotransformation processes to make xenobiotic compounds more soluble and easily excreted. Without xenobiotic biotransformation, the numerous foreign chemicals to which we are exposed would eventually accumulate to toxic levels, either unintentionally or intentionally (Klaassen et al. 2013). Recently, awareness is increasing that human is inevitably exposed to a cocktail of massive xenobiotics due to ubiquitous and continuous use of them, especially some industrial chemicals and pesticides. In our previous work, more than 20,000 chemicals were collected to compile a human exposome database (Zhao et al. 2020). These chemicals can enter the environmental and the food chain, causing unwanted effects and disease to human and other species. In fact, in biomonitoring studies, a large portion of persistent organic pollutants exist as their transformation products (Escher et al. 2020). Therefore, understanding xenobiotic biotransformation is essential in the following aspects:

First, in drug discovery, the biological effects of potential drug cannot be determined if metabolites are not well identified. In some cases, some metabolites are more potent than parent compound. For instance, cyclophosphamide is a prodrug that is implicated to be activated by P450s (Clayton et al. 2009). Secondly, in the field of toxicology, understanding the metabolism of all types of chemicals is important, even if the compounds themselves are not toxic, and it is their metabolites that cause the toxic effects. For instance, acetaminophen hepatotoxicity is mediated by the reactive metabolite N-acetyl-p-benzoquinonimine (NAPQI) (Yoon et al. 2016). In addition, in monitoring field, the characterization or identification of xenobiotic metabolites

from biological or environmental samples is quite difficult and is not unlike natural product identification or dereplication. The biotransformation pattern and ecological toxicity could be vastly different between human and other species in the ecosystem, such as fish and bird (Patterson et al. 2014). A comprehensive understanding of biotransformation pathways of specific species would facilitate a rapid identification process.

To date, the curation of biotransformation data and prediction of biotransformation product is an active field of research. However, the available of suitable data is limited. On the one hand, relevant data used to construct these models are either locked away in proprietary and commercial databases, such as the Accelrys Metabolite database. On the other hand, the coverage of data is limited on the biotransformations of pharmaceuticals in human body system or in microbial system. For example, the EAWAG Biocatalysis/Biodegradation Database has made excellent progress on collecting microbial biocatalytic reactions and biodegradation pathways for primarily xenobiotic, chemical compounds (Gao et al. 2010). Nevertheless, a large number of environmental chemicals, such as flame retardants, plasticizers, endanger not only human health, but also the entire ecosystem.

There are only a few existing databases containing xenobiotic metabolism information. In terms of freely available biotransformation metabolite database, a popular and relatively long-lived, open-source tool is DrugBank. DrugBank is a comprehensive, freely available web resource containing detailed drug, drug-target, drug action and drug interaction information about FDA-approved drugs as well as experimental drugs going through the FDA approval process. For biotransformation aspect, it reports information of substrates, products as well as enzyme if possible. However, the coverage of DrugBank is limited to pharmaceuticals (David S. Wishart et al., 2018). Another freely available database for biotransformation is MetXBioDB, which is a database that consists of a manually curated collection of > 2000 experimentally confirmed biotransformations derived from the literature. Most of substrates are overlapped with DrugBank while the coverage extends to pesticide and other industrial chemicals a little bit. Each biotransformation in MetXBioDB includes a starting reactant (structure and identifiers), a reaction product (structure and

identifiers), the name or type of the enzyme catalyzing the biotransformation, the type of reaction, and one or more citations. It also collected over 30 biotransformation reactions in human gut system (Djoumbou-Feunang et al. 2019). Transformer provides comprehensive information on the transformation and transport of xenobiotics in the human body. It contains the interactions of phase I and II enzymes and drug transporters with drugs, prodrugs, alimentary and Traditional Chinese Medicine compounds, but not what the product is (Hoffmann et al. 2014).

Despite the scientific significance in the understanding of xenobiotic biotransformation, there remains a major gap between the vast knowledge on it. First, though growth, evolution, and accessibility of high-resolution mass spectrometry (HR-MS) has promoted advanced detection in biological and environmental samples, significantly less effort has been devoted to the characterization or understanding of xenobiotic biotransformation among multiple species (i.e., not only human), partly due to the great number of species, and their structural complexity. Given the diversity of biological components that build our environment, it is vital to compile the informatics of chemical and biological interaction between xenobiotic biotransformation and the environment. Second, the biotransformation of environmental pollutants including flame retardants, plasticizers, polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs) and other industrial chemicals lack adequate scientific attention. Though these chemicals are not generally persistent, the continuous global usage make them constantly circulating in the ecosystem. Even some problematic industrial chemicals are regulated, the replaced chemicals tend to attain similar biological properties, including comparable biotransformation pathways. Notable examples contain bisphenol A, bisphenol S, bisphenol AF, and bisphenol F (Schmidt et al. 2013).

To address these needs, we have developed EcoBioTrans, including a web server, as well as a database repository that shall facilitate the curation and standardization of the experimental biotransformation information of multiple species. The web portal allows users to search compounds by simplified molecular input line entry specification (SMILES), and compound name. Along with detailed biotransformation data of 578 environmental chemicals, 978 pharmaceuticals and 43 human

metabolites, their molecular structures, descriptors, physiochemical properties are presented as well. Such a data integration pipeline can be further developed into a protocol that may be adopted by a wider community. Furthermore, we developed an algorithm for querying based on similarity. Even if no biotransformation results of search compound are available, users can acquire results from compounds with similar structures. In addition, we accomplished a systematic analysis on curated biotransformation data, especially on the diverse distribution of drugs and environmental compounds, as well as various species. Lastly, we reported BDE 99 as a case study, and analyzed the biotransformation difference among different species. Future aspects of development were also discussed in Conclusions. In the whole, in the present work, we aimed to establish a comprehensive biotransformation database that include possible data from all categories of chemicals (environmental chemicals, pharmaceuticals and endogenous human metabolites) derived from multiple species (human, rodent, fish, bird and other mammals). These efforts are intended to enhance the understanding of chemical impacts on the whole ecosystem and help to guide the chemical management.

3.3 Methodology

3.3.1 Data collection, processing, and storage

For the data collection, we conducted two rounds of literature searches. The first phase concentrated primarily on some highly concerned environmental pollutants based on our previous work on human indoor exposome (Dong et al. 2019). After thorough search on 385 environmental chemicals, we have found 3,475 literatures related to biotransformation of 213 compounds. For the second phase, we selected 6 keywords, which are 'metabolite', 'metabolism', 'biotransformation', 'microsome', 'toxicokinetics', 'pharmacokinetics', coupled with compound name from U.S. EPA High Production Volume (HPV) List to obtain a comprehensive collection of literature reports on PubMed. Thereafter, we have found 485,400 literatures from second round of search. Then we manually curated the biotransformation information from these literatures and assign them into specific categories. After compiling the list of compounds, their chemical structures and identifiers were retrieved from PUG REST, the REST-style version of PUG (Power User Gateway) for accessing PubChem data and services (Kim et al. 2018). Subsequently, we conducted an expert review to parse these biotransformation data and classified them into specific reaction class for further analysis. Lastly, we combined our database with previous publicly available database DrugBank and MetXBioDB to achieve better coverage. The physicochemical properties were calculated by rdkit (https://www.rdkit.org) based on their chemical structures, including molecular weight, number of aromatic rings, topological polar surface area, rotatable bond count, topological polar surface area, partition coefficient (LogP), and quantitative estimation of drug-likeness (QED).

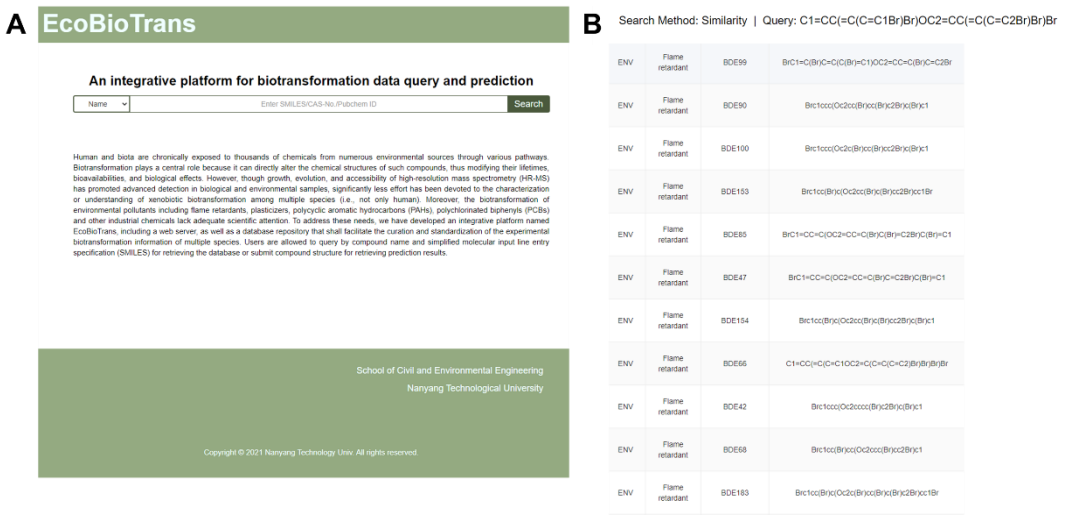


Figure 3.1 EcoBioTrans interface. (A) Homepage; (B) query results of "C1=CC(=C(C=C1Br)Br)OC2=CC(=C(C=C2Br)Br)Br" (SMILES of BDE 99) by the similarity search.

3.3.2 Database architecture and search algorithm

EcoBioTrans was assembled on the Python-based web development framework Django 1.11, with information stored in the MySQL relational database. We provide a variety of query methods to facilitate the acquisition of toxins with similar structures or characteristics, such as structure, similarity, and text-based queries. EcoBioTrans provides a convenient and easy-to-use interface for users. Two services, normal query and similarity query, are designed to support data retrieval, whose input parameters and output information will be elaborated respectively. In the normal query pattern, two molecular submission approaches are provided by pasting the SMILES string or submitting the compound name. Once a user submits the job, the webserver will automatically standardize the input SMILES strings and deliver the result. In the similarity query pattern, one molecular submission approaches are provided by entering SMILES strings. In order to assess molecular similarities based on two-dimensional structure, a fingerprinting technique and a similarity metric are used. The similarity query function is based on the Tanimoto coefficient and the Morgan fingerprint as descriptors, both of which was employed by the open-source cheminformatics module, rdkit. The metric is calculated based on the equation:

$$S_{AB} = \frac{C}{A + B - C}$$

Where S_{AB} is the similarity score, A is the features of molecule A, B is the features of molecule B, C is the common features in molecule A and B.

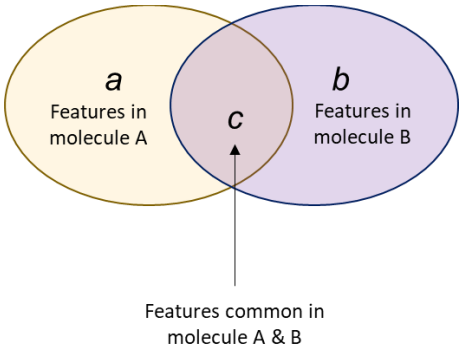


Figure 3.2 Illustration of calculation of similarity metrics.

3.4 Results and discussion

3.4.1 Database overview

EcoBioTrans portal consists of two major components, a database for harboring the most recent biotransformation datasets as well as a web server to visualize these datasets. Using the searching terms stated in the Method part, we received > 48k studies from PubMed and manual curated 7.8k biochemical reactions of multiple species including human, bird, rodent, fish, and other mammals. Afterwards, we attempted to merge more relevant biotransformation data from publicly available databases for a broad coverage. In total, EcoBioTrans consists of 1,599 chemicals and 7,800 biotransformation reactions. For retrieving the database, users can query by compound name, and simplified molecular input line entry specification (SMILES).

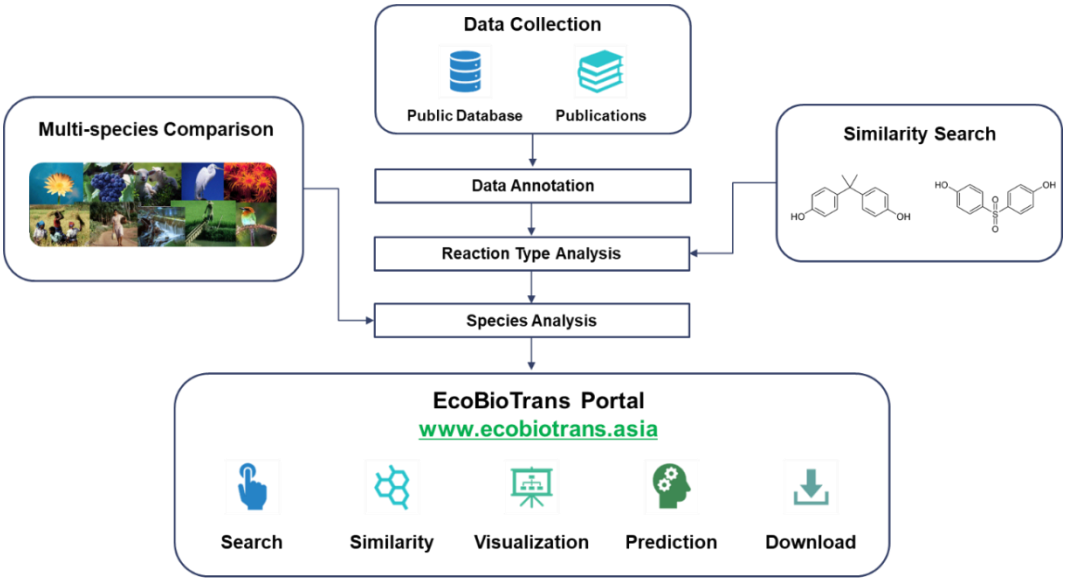


Figure 3.3 A schematic overview of the EcoBioTrans database and web server pipeline.

Biotransformation data are curated from public databases and publications. After quality control and pre-processing, the biotransformation information is annotated, categorized into different reaction class and then compared among multiple species.

The present version of EcoBioTrans consists of 1,599 chemicals from 3 different data sources, including 2 publicly available database, DrugBank and MetXBioDB as well as published publications. The detailed distribution of data source and compound category are presented in Figure 3.4 A and B. MetXBioDB contributes 790 compounds as the primary data source while 471 compounds (mostly drugs) are supplemented from DrugBank. In addition to publicly available database, we manually curated extra 378 compounds, mostly environmental pollutants from published literature. For the first time, we have focused the biotransformation data curation on environmental chemicals, which subcategory distributed are listed in Figure 3.4C.

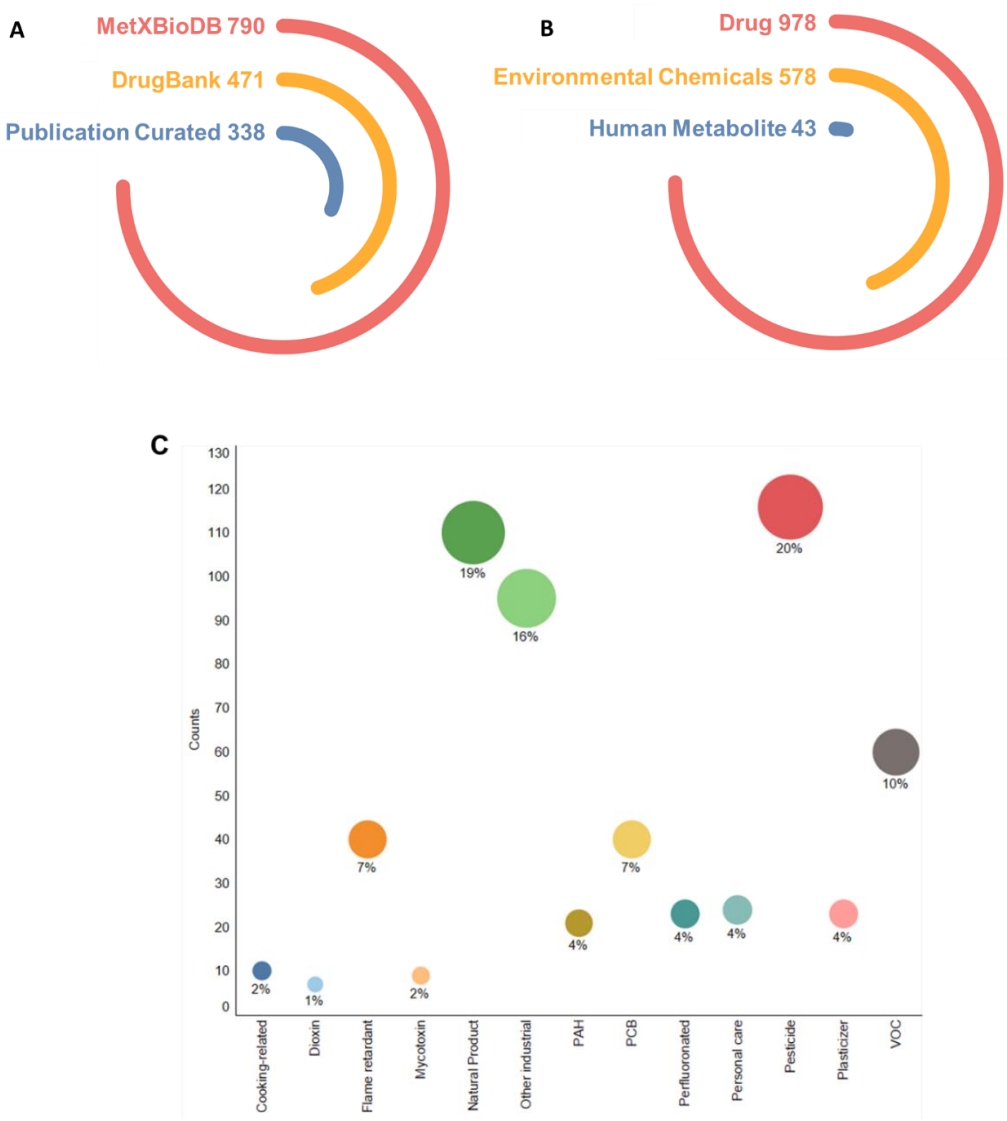


Figure 3.4 Database overview. Distribution of (A) data source, (B) compound category, and (C) compound subcategory of environmental chemicals.

Among the 578 environmental chemicals, pesticide holds the largest proportion at 20% of total, implying a large amount of research effort on it over the past decades. Following closely behind are the natural products, such as phytochemicals, flavonoids, which account for 19% of the total. The biotransformation pathways of a considerable amount of volatile organic compounds (VOCs) are also under scientific attention that 58 chemicals are curated in the database. Flame retardants and PCBs share roughly the same amount of curated data, each accounting for 7% of the whole. PAH, perfluorinated compounds, plasticizers, and personal care products comprise 4% of all environmental chemicals, respectively. Besides, the number of dioxin

category and cooking-related category are less than others, only at around 20. In general, Overall, our collection of biotransformation data on environmental pollutants is extensive and representative, covering all main categories.

3.4.2 The comparative analysis of pharmaceuticals and environmental chemicals

It is known that the physicochemical properties of small molecules, such as lipophilicity, size, hydrogen bonding, and ionization state, broadly influence their biological activities and toxicity profiles (Gleeson, Hersey, Montanari, & Overington, 2011). To gain insights into the possible difference between the pharmaceuticals and environmental chemicals in terms of their biotransformation behavior, we calculated several key physicochemical properties related to drug-likeness parameters of full sets of 978 drugs, 578 environmental chemicals and 47 human metabolites (Bickerton et al. 2012).

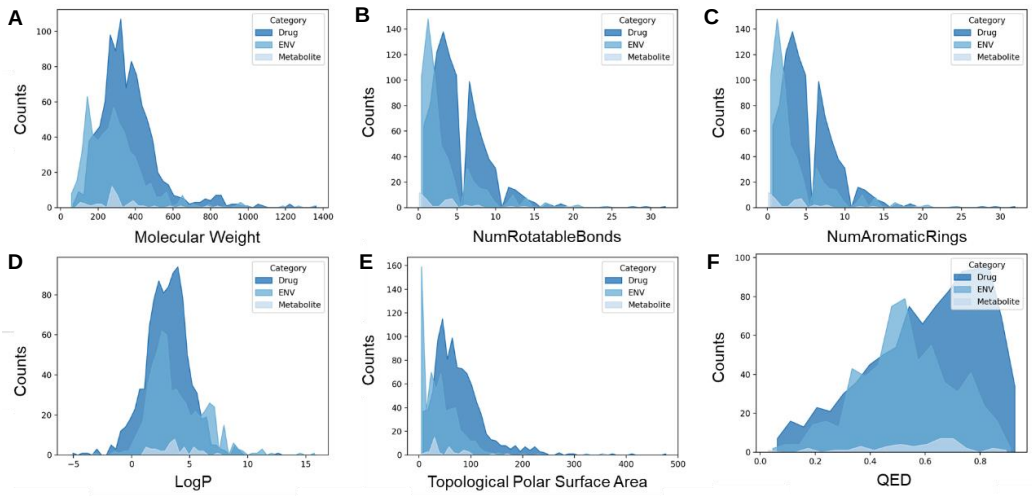


Figure 3.5 Distribution of physicochemical properties representing chemical diversity of pharmaceuticals (Drug), environmental chemicals (ENV), and human metabolite (Metabolite). (A) The distribution of molecular weight; (B) the distribution of number of rotatable bonds; (C) the distribution of aromatic rings; (D) the distribution of LogP; (E) the distribution of topological surface area; (F) the distribution of QED socre.

Density plots depict the distribution of six physicochemical properties of

EcoBioTrans, including molecular weight (MW), partition coefficient (logP), number of rotatable bonds, number of aromatic rings, topological polar surface area, and quantitative estimation of drug-likeness (QED). As presented in Figure 3.5, the molecular weights of chemicals in both categories are not randomly distributed, as pharmaceuticals are mainly found at 200 – 600 Da while a vast majority of environmental chemicals locate at 50 – 400 Da. We also observe similar pattern on number of rotatable bonds, number of aromatic rings, and topological polar surface area, suggesting pharmaceuticals tend to preserve more complex molecular structures. Specifically, environmental chemicals largely present at the range of 0-5 of rotatable bonds and aromatic rings. Meanwhile, a major portion of pharmaceuticals are subject to 0-10 rotatable bonds and aromatic rings. Increasing aromatic ring count are reported to exert effects on several developability parameters that are lipophilicity- and size-independent and have a detrimental effect on mammal bioavailability parameters (Ritchie et al. 2011). This is probably a reason of different biotransformation behavior between the two chemical groups. Topological polar surface area is a metric for the optimization of a compound's ability to permeate cells (Pajouhesh and Lenz 2005). Environmental chemicals mainly fall in the range below 100, whereas drugs' range are slightly wider to 130. It should be noted that molecules with less than 90 angstroms squared are possibly able to penetrate the blood–brain barrier (Hitchcock and Pennington 2006). Thus, environmental chemicals with higher bioavailability might pose a much more damaging effect than drugs. In terms of partition coefficient (logP) values, most compounds of both drugs, environmental chemicals and human metabolites fell in the range of 0-8. Thereafter, we calculated an overall parameter called quantitative estimation of drug-likeness (QED) to compare the overall difference among these three groups, and the result shows that environmental chemicals behave much diversely while most pharmaceuticals locate at range over 0.5, demonstrating that the biological activity of environmental chemicals is possibly in a manner different to drugs. In summary, compounds on pharmaceutical category are peculiarly prone to more rotatable bonds and aromatic bonds, greater molecular weight, coupled with broader topological surface area, all of which indicate more complicated molecular structures, higher bioavailabilities and more active biological activities.

3.4.3 Reaction distribution analysis

After parsing and curating all reactions in the EcoBioTrans database, we classified them into 4 main categories: hydrolysis, reduction, oxidation, and conjugation. In Figure 3.7, each row represents for a reaction subclass affiliated with one main reaction class, such as oxidative deamination, oxidative desulfuration, hydroxylation under main class of oxidation. To identify possible discrepancy of chemical categories, we highlighted them into different colors. Overall, the oxidation reaction is revealed as the most common reaction type with 3,404 reactions in total while the reduction reactions is relatively scarce at 254 reactions. Besides, we summarized 2,002 reactions under hydrolysis category and 1,643 reactions under conjugation category. For dehalogenation reactions, we particularly listed it out for its diverse mechanisms and distinguished frequency on environmental chemicals. There are three major mechanisms for removing halogens, known as reductive dehalogenation involving replacement of a halogen with hydrogens, oxidative dehalogenation that a halogen and hydrogen are replaced with oxygen, as well as elimination. On all these reaction subclasses of dehalogenation, we discovered more occurrence of environmental chemicals, implying that dehalogenation might be a special reaction for environmental category. This is a probable result of larger structural distribution of environmental chemicals on halogens. However, the number of demethylation reactions from the drug category exceeds environmental chemicals by 344. Likewise, in terms of N-oxidation category and deacetylation category, 74 and 66 pharmaceutical reactions are presented respectively, but rare environmental reactions are shown. Besides, we found that hydroxylation is the most frequently encountered reaction type on both sets, comprising 1,002 counts for drugs and 1,251 counts for environmental chemicals, respectively. Following that is the glucuronidation reaction from the conjugation class, which contributes 404 reactions on drug and 504 reactions on environmental chemicals. Sulfonation is another conjugation reaction that occurs frequently, accounting for 209 environmental reactions and 106 pharmaceutical reactions. In general, the overall trends for the reaction distribution of drugs and environmental pollutants are consistent, but the subcategories behave differently in a variety of ways. Furthermore, the EcoBioTrans database is the first to focus on environmental chemicals and collect 3,978 biotransformation reactions from them.

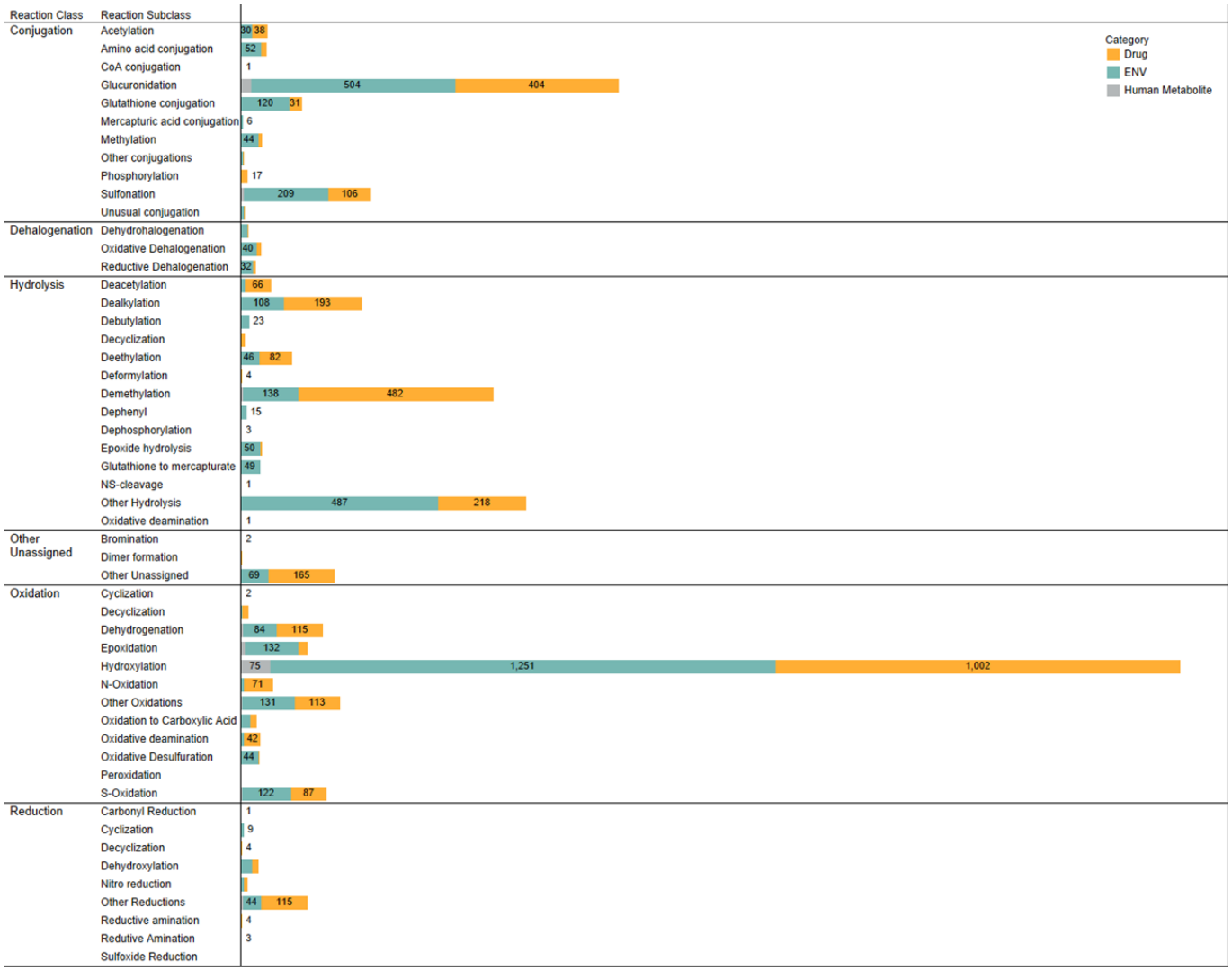

Reaction Class
Reaction Subclass
Conjugation
Acetylation 30 38
Amino acid conjugation 52
CoA conjugation 1
Glucuronidation 504 404
Glutathione conjugation 120 31
Mercapturic acid conjugation 6
Methylation 44
Other conjugations
Phosphorylation 17
Sulfonation 209 106
Unusual conjugation
Dehalogenation
Dehydrohalogenation
Oxidative Dehalogenation 40
Reductive Dehalogenation 32
Hydrolysis
Deacetylation 66
Dealkylation 108 193
Debutylation 23
Decyclization
Deethylation 46 82
Deformylation 4
Demethylation 138 482
Dephenyl 15
Dephosphorylation 3
Epoxide hydrolysis 50
Glutathione to mercapturate 49
NS-cleavage 1
Other Hydrolysis 487 218
Oxidative deamination 1
Other Unassigned
Bromination 2
Dimer formation
Other Unassigned 69 165
Oxidation
Cyclization 2
Decyclization
Dehydrogenation 84 115
Epoxidation 132
Hydroxylation 75 1,251 1,002
N-Oxidation 71
Other Oxidations 131 113
Oxidation to Carboxylic Acid
Oxidative deamination 42
Oxidative Desulfuration 44
Peroxidation
S-Oxidation 122 87
Reduction
Carbonyl Reduction 1
Cyclization 9
Decyclization 4
Dehydroxylation
Nitro reduction
Other Reductions 44 115
Reductive amination 4
Redutive Amination 3
Sulfoxide Reduction
Category
Drug
ENV
Human Metabolite

Figure 3.6 Distribution of the various types of biotransformation reactions on environmental chemicals, pharmaceuticals, and human metabolites. Colors of green, yellow, grey represent for categories of environmental chemicals, pharmaceuticals, human metabolites, respectively.

3.4.4 Species distribution and difference

Conspicuously, human being is not the only species under the risk of chemical pollution. Countless animals and insects in the ecosystem are exposed to these manmade chemicals like environmental pollutants as well. To gain the comprehensive understanding of biotransformation processes in the whole ecosystem, for the first time, we collected biotransformation information from several species, particularly a variety of species belong to bird and fish. We curated biotransformation reactions from 578 environmental compounds in total, some of which contain information from multiple species, while 429 compounds contain information from only one. Due to its prime biological significance, the majority of substrates contain experimental or monitoring information of humans. That is followed by the rodent, the most prevalent and easily accessible experimental species. The data from fish and birds are both relatively scarce, totaling around seventy.

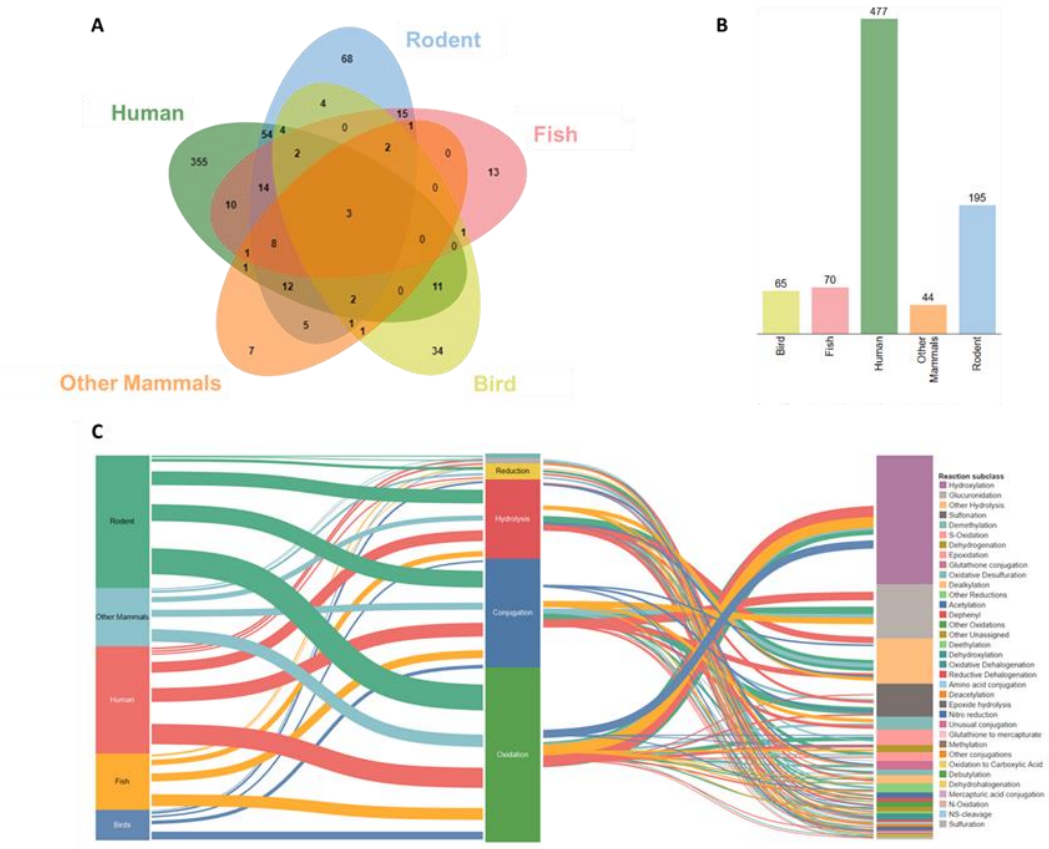


Figure 3.7 (A) Overlap analysis of 578 environmental chemicals of 5 major species categories mapping in EcoBioTrans database; (B) distribution of biotransformation

reactions among five species categories including bird, rodent, human, fish, and other mammals; (C) Sankey diagram that illustrates the distribution of the biotransformation reactions identified in terms of biosystems, assigned their occurrences in reaction classes and reaction subclasses.

Consequently, to characterize the distribution of these compounds from difference species on reaction classes, we presented a Sankey diagram that illustrate how substrates flow to reaction classes on each biosystem category. Consistent with the overall distribution trend, the oxidation class and hydroxylation subclass are the two regions that all species gravitate toward the most. We found 139, 108, 62, 40 oxidation reactions on rodent domain, human domain, fish domain and bird domain, respectively. For each of domain, over 70 percent of oxidation reactions are hydroxylation reactions on average. Glucuronidation from the conjugation class is shown as the following one, with obvious flow from all five species domains. Sulfonation is another conjugation reaction that is important for all species, containing over 20 reactions from human and rodent, respectively. In general, we did not observe significantly diverse distribution of reactions among species domains. However, there are two exceptional subclasses, dehydrohalogenation and methylation. Specially, fish displayed considerably inclination to dehydrohalogenation belong to dehalogenation reaction class comparing with other domains. Besides, birds tend to perform methylation at a greater frequency than other species. In particular, 43 methylation reactions were curated from bird species while only 11 from human and 1 from rodent.


Conjugation
Dehalogenation
Hydrolysis
Oxidation
Reduction
Acetylation
Amino acid conjugation
Glucuronidation
Glutathione conjugation
Mercapturic acid conjugation
Methylation
Other conjugations
Sulfonation
Unusual conjugation
Dehydrohalogenation
Oxidative Dehalogenation
Reductive Dehalogenation
Deacetylation
Dealkylation
Debutylation
Deethylation
Demethylation
Dephenyl
Epoxide hydrolysis
Glutathione to mercapturate
NS-cleavage
Other Hydrolysis
Dehydrogenation
Epoxidation
Hydroxylation
N-Oxidation
Other Oxidations
Oxidation to Carboxylic Acid
Oxidative Desulfuration
S-Oxidation
Dehydroxylation
Nitro reduction
Other Reductions
Sulfuration
Category
Sub
CID
Drug
Drug
5323
404291
ENV
Dioxin
36614
Flame retardant
8289
14716
26177
36159
93311
95170
Mycotoxin
4580
40024
186907
Natural Product
10168
Other industrial
980
8679
PAH
2336
9154
9171
31423
Personal care
4632
7456
7547
8571
456201
Biosystem
Bird
Fish
Human
Other Mammals
Rodent

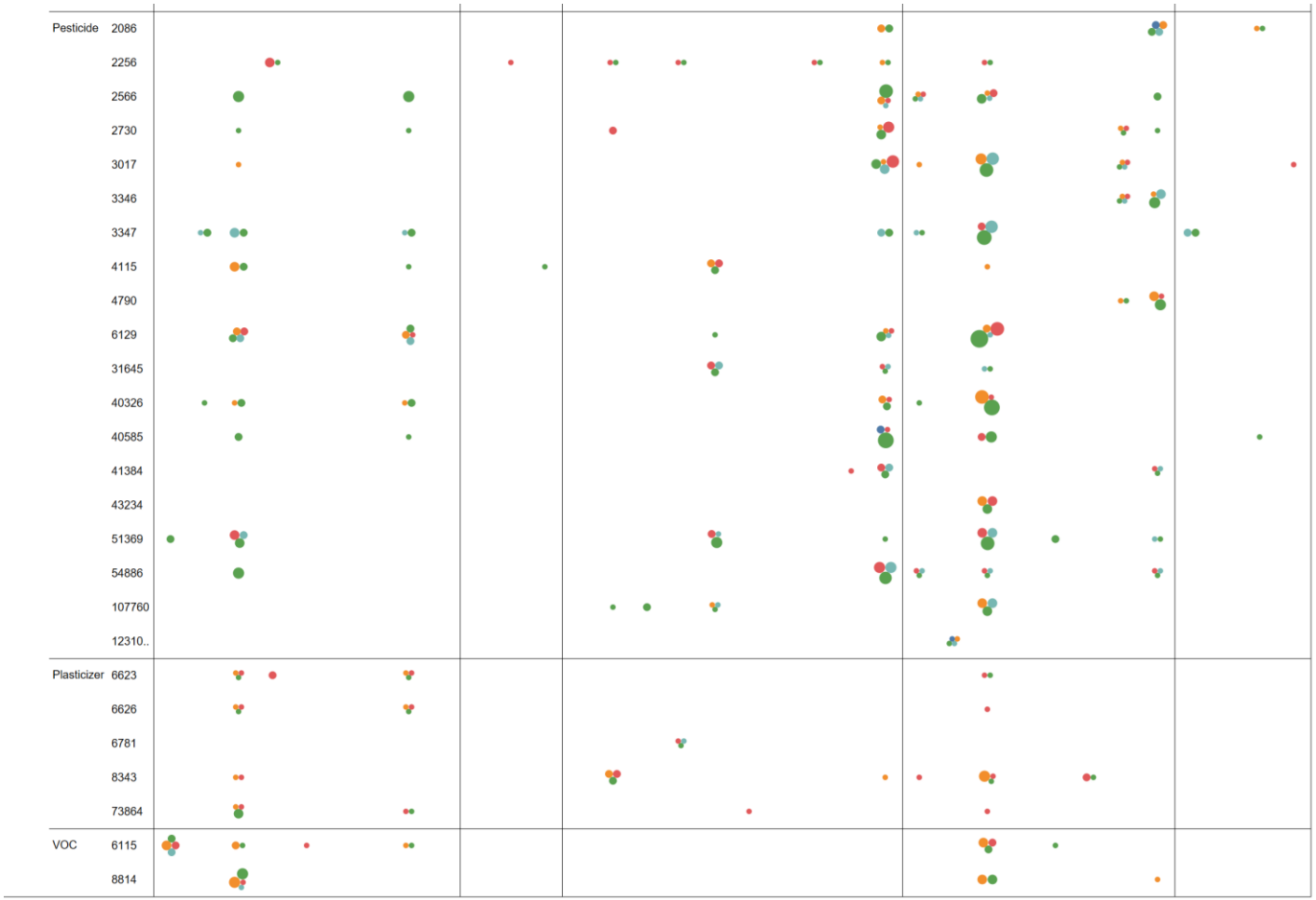


Figure 3.8 Overview of the distribution of reaction classes and reaction subclasses of the summarized compounds that have biotransformation

information on 3 or more species domains. Each of color set represent for one species domain while the size of circle shows occurrence of reactions.

To further explore the different biotransformation behaviour on diverse biosystems, we selected substrates with biotransfromation information on three or more species categories. As can been seen from Figure 3.9, the category contains most substances is pesticide, followed by flame retardant and plastizer. On the whole, we found clustering tendency of most compounds among species. Little variability between species has been disclosed. Corresponding to one specific compound category under one reaction type, multiple species domains exhibited similar occurrence results. Consistent with previous findings of reaction distribution part, we discovered most circles under the columns of hydroxylation reaction and glucuronidation reaction. However, for the dehalogenation reaction, most occurrences are from human and fish. Likewise, in terms of methylation reaction, only two compounds under the flame retardant category were displayed from bird domain.

3.4.5 Case study of EcoBioTrans database

Here, we utilize BDE 99 as an example to elucidate the differences of xenobiotic biotransformation behavior among different species. BDE 99 (2,2',4,4',5-Pentabromodiphenyl ether), well-known as one of the most frequently detected brominated flame retardants in environmental samples and biological systems, has been demonstrated to attribute endocrine disruption effects, developmental effects, and carcinogenic effects (Wu et al. 2020). Furthermore, biotransformation information for BDE 99 is unavailable from other databases such as PubChem, DrugBank, and HMDB (Kim et al., 2019; Wishart et al. 2007; Wishart et al. 2018).

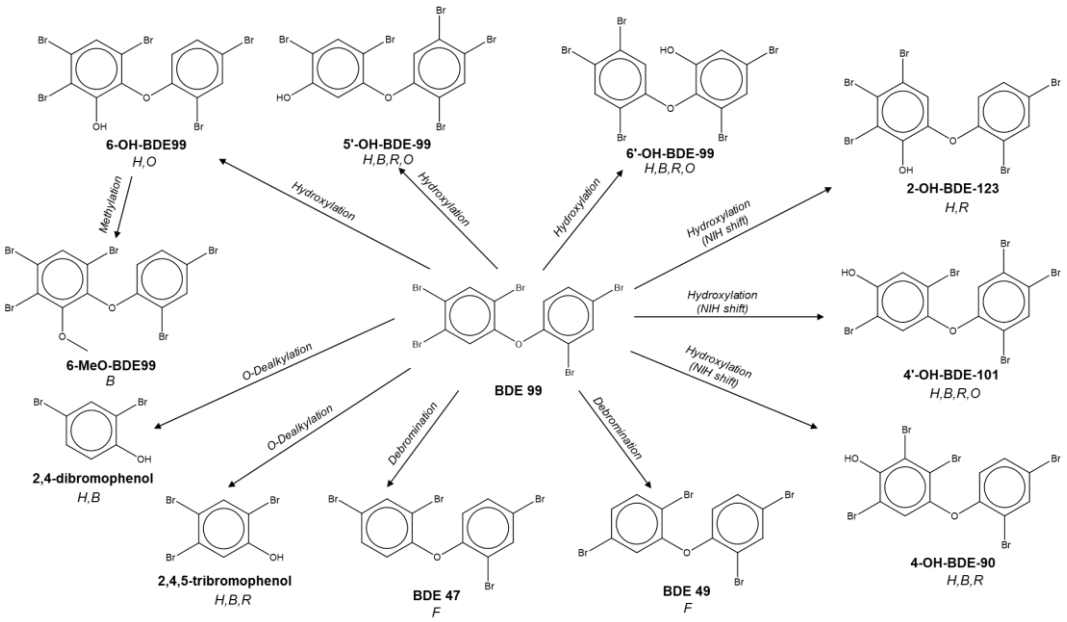


Figure 3.9 Curated biotransformation pathways of BDE 99 in 5 biosystems: Human (H), Rodent (R), Bird (B), Fish (F) and other mammals (O).

The EcoBioTrans has curated BDE 99 biotransformation information from *in vitro* or *in vivo* studies of human, rodent, fish, bird and cat, which is assigned as other mammal. As can been seen from Figure 3.10 A, hydroxylation compensates the largest portion of all biotransformation pathways and leads to six hydroxylated products, that are 6'-hydroxy-BDE 99, 6'-hydroxy-BDE 99, 6-hydroxy-BDE 99, 4-hydroxy-BDE 90, 4'-hydroxy-BDE 101 and 2-hydroxy-BDE 123. Among the total 6 hydroxylated products, half is normal hydroxylation while the other half is product from hydroxylation involved with NIH shift. No significant difference has been observed among human, rodent, bird, and other mammal. However, from *in vivo* and *in vitro* studies of fishes, we have not found any hydroxylated product of BDE 99. Instead, two debrominated products, BDE 47 and BDE 49, are reported only on fish studies, implying that reductive dehalogenation might be a relatively inclining pathway for species in fish category. Likewise, one unique biotransformation product, 6-MeO-BDE 99, is identified in bird category. 6-MeO-BDE 99 is a product of methylation reaction subsequent to 6-OH-BDE 99. Among all the biotransformation pathways in EcoBioTrans, methylation is a unique reaction type for environmental chemicals, that are mainly found in polychlorinated biphenyls and

polybrominated diphenyl ethers from bird category. Nevertheless, some drugs undergo methylation reaction in human, such as ciprofloxacin, abametapir and milnacipran. Besides, we collected 2,4-dibromophenol and 2,4,5-tribromophenol, which are hydrolysis products from O-dealkylation in biosystems of human, bird and rodent. Overall, the illustrative biotransformation pathways of BDE 99 validated that the biotransformation performances diverge a lot amongst different species. This is a probable result of their distinct metabolic enzyme systems, disparate bioavailability efficient, and varying elimination rate.

3.5 Short Summary

In this work, we set up an open-access data portal, EcoBioTrans (https://www.ecobiotrans.asia) [3], which is an open access database harboring biotransformation data of common environmental pollutants, along with pharmaceuticals and human endogenous metabolites. To establish this data repository, we collected 7,800 biotransformation reactions of 1,599 substrates by text mining and database fusion.

For the first time, data from various species and compound classes are collected thoroughly. The collected species are not restricted to human, but also extended to bird, fish, rodent, and other mammals. This database contains structures of substrates and their corresponding products, types of reactions, the names or types of enzymes catalyzing biotransformation, types of biosystems, and references. Users are also allowed to search biotransformation results based on structural similarity. Overall, environmental contaminants and various species are being addressed as a focus for biotransformation data curation for the first time. This integrative platform contributes significantly to the scientific understanding of the biotransformation pathway in the entire ecosystem and acts as a guide for chemical risk management.

[3] Please use the server address (http://1.117.57.232/) to view the website as our server is being audited and certified.

Chapter 4. Computer-Assisted Biotransformation Prediction Based on Molecular Similarity

4.1 Overview

Biotransformation is crucial for the activation, detoxification, and elimination of foreign chemicals, including pharmaceuticals as well as environmental pollutants. To date, prediction of biotransformation product is an active field of research due to its significant role. However, there is a knowledge gap in that almost all prediction tools were developed for pharmaceuticals, despite the fact that countless environmental substances also undergo biotransformation, and their metabolic pathways may differ slightly from pharmaceuticals. Furthermore, current methodologies are based on hand-encoding metabolic transformation rules, which require a significant amount of effort from experts. Here we presented a novel computational method with automatically extracted reaction template, which further ranked candidates by molecular similarity metric. Our work demonstrated molecular similarity to be an effective metric. Extensive evaluation revealed that the current model achieved 45.25% recall and 26.37% precision at the top 3 candidates and 65.67% recall and 7.96 % precision at the top 20 candidates, respectively. Our results indicate that the proposed approach can provide a simple and automatic solution that does not restrict to any expert knowledge. Lastly, BioSim is freely available as a web server at www.ecobiotrans.asia and is also provided as a Python package upon request.

4.2 Introduction

Xenobiotic metabolism plays an indispensable role in the activation, detoxification, and elimination of foreign chemicals, such as environmental pollutants, pharmaceuticals, food additives, and biologically foreign substances. It converts xenobiotics into metabolites with different physicochemical and pharmacological properties through a series of specific enzyme catalyzed reactions (Klaassen et al. 2013). Biotransforming reactions are commonly classified into two groups, phase I (Functionalization) and phase II (Conjugation). Phase I reactions include hydrolysis, reduction, and oxidation whereas Phase II biotransformation reactions involve conjugation reactions such as glucuronidation, sulfonation (or sulfation), glutathione

conjugation, methylation, acetylation, conjugation with amino acids, and other unusual conjugations. This process is critical because many ingested xenobiotics are commonly lipophilic and would accumulate to toxic levels without metabolism and subsequent elimination. For example, ochratoxin A is toxic, carcinogenic, and organ-specific due to poor biotransformation and slow elimination of metabolites (Tran et al. 2020). However, biotransformation may also convert the xenobiotics to toxic substances, comprising mutagenic, carcinogenic, or teratogenic molecules. The hepatotoxicity of acetaminophen, for example, is mediated by the reactive metabolite N-acetyl-*p*-benzoquinonimine (NAPQI) (Yoon et al. 2016). It also influences parameters such as absorption, passive membrane permeation, transport and binding to macromolecules (Braga and Andrade 2012). Overall, biotransformation plays an essential part in the field of toxicology, especially environmental toxicology. In particular, the toxicologic profile of a substance unequivocally depends on whether a xenobiotic and/or its metabolites are present in appropriate concentrations at its sites of action.

Therefore, the curation of biotransformation data and prediction of biotransformation product is an active field of research due to its significant role. Over the past 100 years, continuous research effort has been put into computational approaches for biotransformation prediction. The three main study fields of computer-aided metabolic prediction approaches are prediction of metabolic sites (SOMs), prediction of metabolites structures, and prediction of metabolizing enzymes binding, which provide crucial support and guidance for experimenters. Thereinto, the first two methods have a broader range of applications (Kirchmair et al. 2012). SOM prediction methods typically have a high level of prediction accuracy. However, most of these methods, such as CypReact, are limited to CYP450 catalytic reactions and can only predict unstable sites rather than metabolite structures (Tian et al. 2018). SOM predictions are not the same as identifying the correct bioinformations that will occur at an atomic position, and they do not provide information about the type of reaction that will take place. This makes it difficult to draw quantitative inferences regarding a molecule's metabolic potential because of these restrictions.

From the very first attempt of computer-assistance in the prediction of xenobiotic

metabolism, a large number of automated programs have relied on encoding reaction templates. The second way is to predict the structure of metabolite without preliminary prediction of SOMs. Primarily it is realized in expert systems, which predict products of metabolism using dictionaries of biotransformation. MetabolExpert, META, Meteor, UM-PPS, SyGMa, TIMES, the JChem Metabolizer module, GLORYx, Biotransformer are the most well-known expert systems for predicting xenobiotic metabolites (Darvas and Kaiser 1987; Bruyn et al. 2020; Djoumbou-Feunang et al. 2019; Gao et al. 2010; Klopman et al. 1994; Marchant et al. 2008; Mekenyan et al. 2004; Ridder and Wagener 2008). These systems contain rules of transformation of parent compounds into their metabolites. The main disadvantage of this approach is that extracting the template is time-consuming and requires a significant amount of time and effort from experts. The first open-source predictor of the comprehensive metabolite structure is Sygma, which utilizes 118 phase 1 knowledge-based reaction rules and 27 phase 2 knowledge-based reaction rules (Ridder and Wagener 2008). Recently, Djoumbou-Feunang et al. (2019) developed another freely available tool for metabolite prediction, which is BioTransformer with 163 CYP rules and 74 phase 2 rules (Djoumbou-Feunang et al. 2019). However, Biotransformer does not rank the candidate metabolite. In addition, another freely available tool for metabolite prediction called GLORYx was established with 261 reaction rules.

However, though biotransformation is crucial for environmental toxicological studies, none of any prediction method was developed with enough metabolic information of environmental chemicals. Based on our previous observations on the comparison of biotransformation behavior between drug and environmental chemicals, though the overall trends for the reaction distribution of drugs and environmental pollutants are consistent, the subcategories behave differently in a variety of ways. For instance, dehalogenation might be a special reaction for environmental category, which is a probable result of larger structural distribution of environmental chemicals on halogens. Here in this work, we presented a novel computational method with automatically extracted reaction template, which significantly increased the reaction templates to 1,800. Moreover, with more biotransformation data from environmental chemicals, this model is proved to behave

better on prediction of environmental chemicals than previous methods. Subsequent ranking strategy was adopted from retrosynthesis strategy, which is calculated by similarity metrics. Prior research on retrosynthesis field reveals that molecular similarity is a surprisingly effective metric for proposing and ranking one-step retrosynthetic reaction based on analogy to precedent reactions, which is consistent with our results that achieves 45.25% recall and 26.37% precision at the top 3 candidates and 65.67% recall and 7.96 % precision at the top 20 candidates, respectively. Overall, BioSim is proved to be a powerful prediction method for one-step biotransformation reaction and its web service is also freely accessible at www.ecobiotrans.asia.

4.3 Methodology

4.3.1 Dataset collection

This work collected metabolic reaction data from a database that we created previously called EcoBioTrans. The EcoBioTrans database is an open access database harboring biotransformation data of common environmental pollutants, along with pharmaceuticals and human endogenous metabolites. It was created to aid in the summarization of biotransformation rules, as well as the training and validation of machine learning biotransformation prediction models. Every biotransformation reaction in EcoBioTrans includes the information of reaction substrate, the information of reaction product, the name or type of the enzyme catalyzing the biotransformation, the type of reaction, the source citation, and the biosystem. All abnormal structures such as free radicals, metal chelating and structural flaws were removed from the data so that the different reaction sites could be distinguished. The pre-treatment resulted in 5,600 reaction records.

4.3.2 Automatic extraction of biotransformation reactions template

The pipeline for automatic template extraction was adapted from previous prediction models in retrosynthesis field, based on the cheminformatics framework RDKit.(Coley et al. 2019) As illustrated in Figure 4.1, the following steps were involved in completing the automatic extraction in general. To begin, the

biotransformation reactions were converted into SMILES strings of the form reactants>>product. Following that, the reaction SMILES strings were mapped atom-to-atom using an unsupervised attention-guided atom-mapping method called Rxnmapper (Schwaller et al. 2021). The local properties of each atom in the product are then evaluated to see if they are identical to those of its reactants. If two atoms have the same SMARTS pattern, atomic number, total number of hydrogens, formal charge, degree, number of radical electrons, aromaticity, bond order, and atomic number of neighboring atoms, they are considered identical. Non-identical atoms, as well as any mapped reactant atoms that do not appear in the product, would be included in the list of changed atoms. Reaction centers are substructures composed of atoms and bonds that change during the reaction. Next, SMARTS patterns are generated for reactants and products, respectively. A strict SMARTS pattern is defined for each atom that changes between reactants and products, including unmapped atoms belonging to leaving groups. When reacting atoms are part of a specific substructural motif, such as, glucuronide conjugate, glutathione conjugate, sulfate conjugate, the entire motif would be included in the template. A SMARTS atom pattern is defined for each atom that is a neighbor of a reacting atom (or belongs to a special functional group). Radius = 1 was used to extract reaction rules because it has been shown in previous retrosynthesis studies to have high sensitivity and specificity (Segler et al. 2018). A SMILES string is formed by converting a molecular fragment containing only the atoms surrounding the reaction center. To generate an overall SMARTS pattern for the reactants, the SMILES tokens of the atoms are replaced by their custom SMARTS patterns. The SMARTS patterns are then sorted alphabetically in an attempt to canonicalize the order of disconnected reactant/product fragments. The final reaction template is created as a concatenation reactant_smarts >> product_smarts.

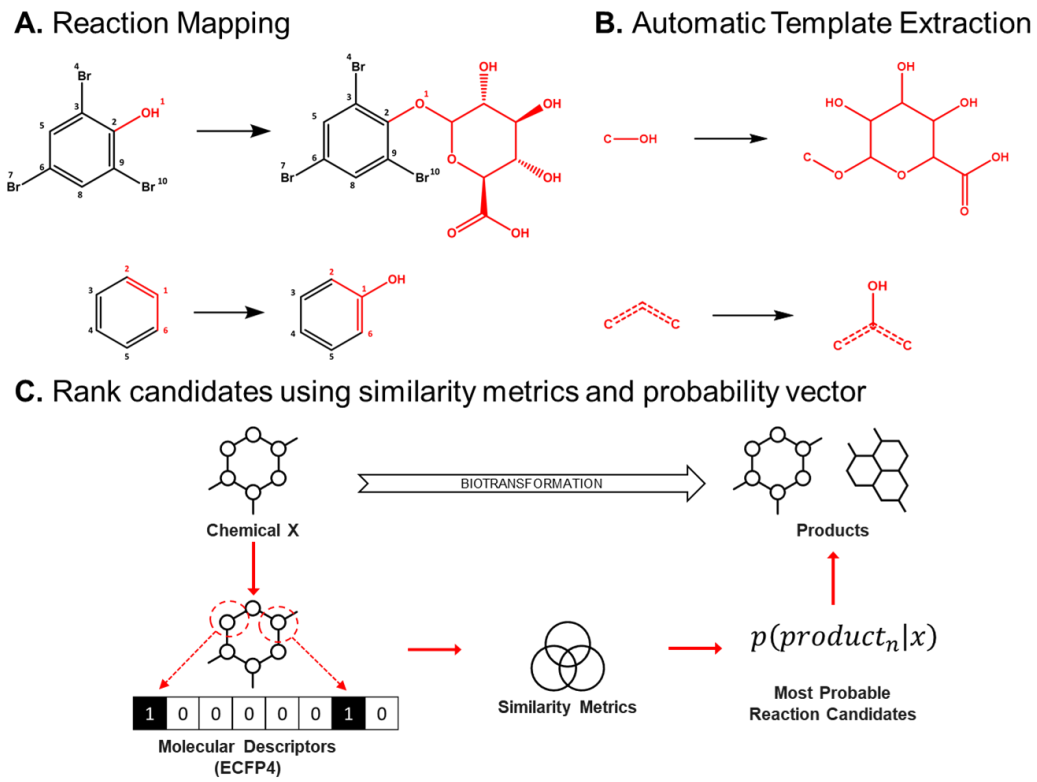


Figure 4.1 Overall workflow of BioSim prediction.

4.3.3 Reaction templates application based on similarity calculation

Following the acquisition of a comprehensive set of reaction templates, they were applied to the prediction compounds using a similarity score. In order to assess molecular similarities based on two-dimensional structure, a fingerprinting technique and a similarity metric are used. As a result of this, we employed Morgan circular fingerprints with varied settings from RDKit in this study. Essentially, the Morgan fingerprint is an expanded connectivity fingerprint reimplemented in a new way (ECFP). As part of this research, we used four fingerprinting approaches, including Morgan fingerprints of radius 2 without features, Morgan fingerprints of radius 3 without features, Morgan fingerprints of radius 2 with features, and Morgan fingerprints of radius 3 with features.

There are a number of similarity metrics that we attempted in this study as well. The similarity metrics evaluate how similar two molecules are to each other in different ways. The first is the Dice similarity, which quantifies the similarity between two

fingerprint vectors A and B by calculating the ratio of overlapping substructure prevalence, along with the Tanimoto metric and Tversky similarity. The detailed equations for similarity metrics calculation are listed in Table 4.1.

Table 4.1 Equations of similarity metrics

Similarity Metric	Equation
Dice similarity	$S_{AB} = \frac{2C}{A+B}$
Tanimoto similarity	$S_{AB} = \frac{C}{A+B-C}$
Tversky similarity_1	$S_{AB} = \frac{C}{\alpha A+\beta B-C}$
Tversky similarity_2	$S_{AB} = \frac{C}{\alpha A+\beta B-C}$

Where S_{AB} is the similarity score, A is the features of molecule A, B is the features of molecule B, C is the common features in molecule A and B, α in Tversky similarity_1 is 1.0 and α in Tversky similarity_1 is 0.5 while β in Tversky similarity_1 is 0.5 and α in Tversky similarity_1 is 1.0.

4.3.4 Evaluation procedures

In order to evaluate the performance of this model, we performed a comparative analysis with three popular *in silico* metabolism prediction tools, namely Biotransformer, GLORYx and Sygma.(Bruyn et al. 2020; Djoumbou-Feunang et al 2019; Ridder and Wagener 2008)

In terms of test datasets, we referred to several previous studies as follows:

Biotransformer first conducted a comparative assessment of the performance of BioTransformer and Meteor in predicting single-step human metabolism of 40 pharmaceuticals and pesticides, *randomly selected* from DrugBank and T3DB. Then it conducted test on the CYP450-catalyzed single-step metabolism of the 60 molecules. GLORYx manually curated test data set from literatures based on the top 100 best-selling drugs from 2018. The final test data set contains 37 parent molecules with 136 first-generation metabolites in total. The transformer model, MetaTrans,

collected the test set for the comparison consisted of 65 drugs with a total of 179 metabolites, which cover a wide range of enzymes. It might be observed that the performance of same prediction tool varies a lot on different works, which is a result of different test datasets.

However, the selection of previous test dataset sizes is limited to dozens of compounds, which could easily be biased. Hence in this work, we utilized all one-step biotransformation reactions involved in environmental category of our database and conducted a cross validation. For each compound to test, we deleted reaction templates related to this compound and test its accuracy.

4.3.5 Calculation of precision and recall

Precision and recall are important evaluation metrics and they are calculated as:

$$Precision = tp/(tp + fp)$$

$$Recall = tp/(tp + fn)$$

Where tp is true positive (equivalent with hit), fp is false positive (equivalent with hit), fn is false negative (equivalent with hit).

Thus, the precision metric is calculated as total identified metabolites are divided by total prediction output metabolites. The recall metric is calculated as total identified metabolites are divided by total metabolites.

4.4 Results and discussion

4.4.1 Introduction of overall scheme

The overall scheme of this work is illustrated in Figure 4.2. To begin, reaction precedents are retrieved from the knowledge base based on reactant similarity, S_{react},

which is assigned a value between 0 and 1. The calculation of molecular similarity is discussed at method section. After that, a highly local transform containing fully specified leaving groups is extracted from each precedent reaction and applied to the target compound. In contrast to traditional template extraction approaches that attempt to include neighboring atoms as necessary context, these templates contain only the atoms that are immediately involved in the reaction (specified by atomic identity, aromaticity, number of hydrogen atoms). Using the example of Figure 4.1, the template for a hydroxylation reaction would consist only of the two aromatic carbons that are bridged in the product and the additional O. Besides, the template for a glucuronidation reaction would consist only of the one aromatic carbon and one hydroxy group with the special glucuronide group conserved. Next, the valid products were compared with template products, and a product similarity score, S_{pro}, were obtained by same similarity calculation method. Lastly, the candidates were ranked by the result of multiplying S_{react} and S_{pro}.

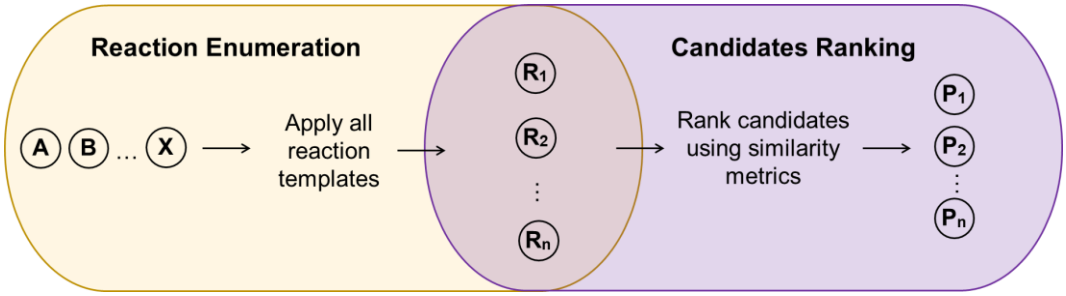


Figure 4.2 Overall scheme of BioSim prediction.

4.4.2 Top10 templates in dataset

After carefully parsing and extracting procedures, we have extracted 1,800 reaction templates from all biotransformation reactions. In Figure 4.3, the top five most popular biotransformation templates extracted from environmental chemicals and drugs are presented.

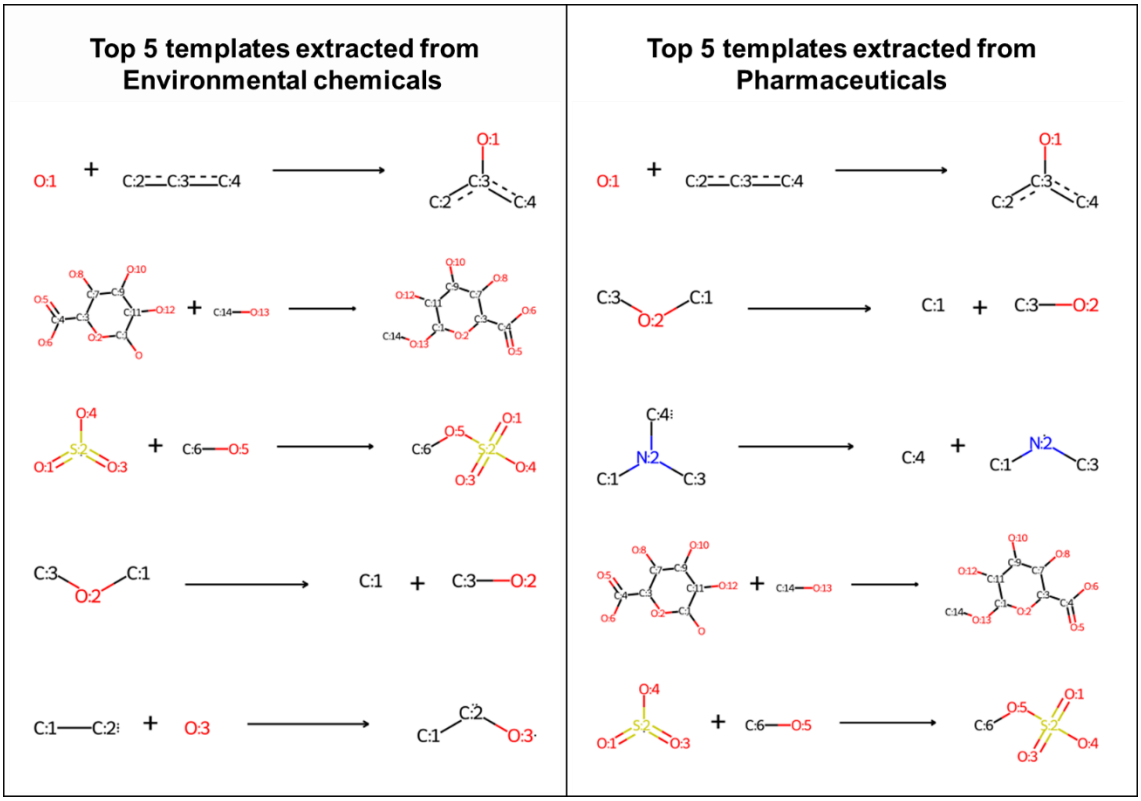


Figure 4.3 Top 5 extracted templates from environmental chemicals and drugs in the dataset.

Specifically, the top five most popular reaction templates from environmental chemicals are: hydroxylation reaction of benzene group, glucuronide conjugation reaction, sulfonation reaction, demethylation reaction and hydroxylation reaction of alkyl group. In contrast, the top five most popular reaction templates from pharmaceuticals are: hydroxylation reaction of benzene group, O-demethylation reaction, N-demethylation reaction, glucuronide conjugation reaction and sulfonation reaction. In general, the overall trends for templates from these two categories are consistent that hydroxylation, glucuronidation, sulfonation, demethylation comprised of the most popular templates. However, their ranking orders are not that consistent. In environmental chemicals, the second and third position reaction types are denoted by glucuronidation and sulfonation, respectively. Whereas the O-demethylation reaction and the N-demethylation reaction, are the second and third position reactions in pharmaceutical category. This is a probable result of different structural distribution of environmental chemicals and drugs.

4.4.3 Similarity calculation comparison

As illustrated earlier, in this work, we attempted ranking strategy with four fingerprinting techniques and four similarity metric calculation methods. The fingerprinting techniques include Morgan fingerprint at radius 2 and radius 3 with or without features. The similarity metric calculation methods are well discussed in method section, which are Dice similarity coefficient, Tanimoto similarity coefficient and two variant Tversky similarity coefficients. Overall, their performance is quite similar across all sixteen combinations.

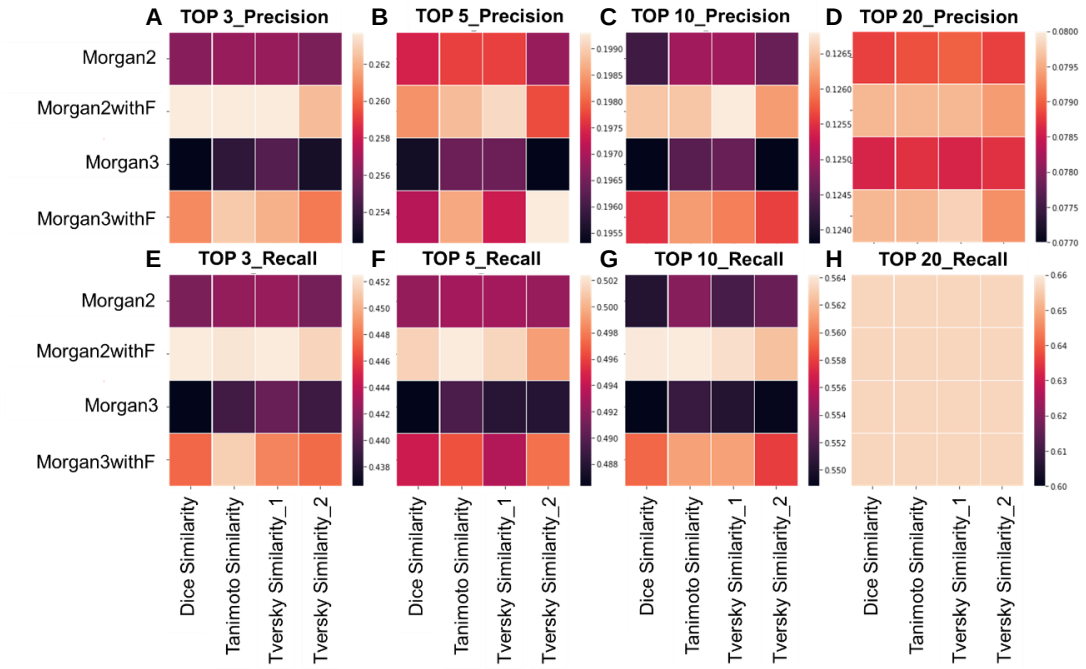


Figure 4.4 Precision and recall results when predicting biotransformation at different fingerprinting and similarity schemes. Each square shows the top-n fractional precision and recall results for that fingerprinting and similarity combination, colored between purple (low accuracy) and nude (high accuracy).

Specifically, Morgan fingerprints of radius 2 with features performed best for the top three predicted candidates, achieving the highest precisions and recalls. Meanwhile, Tversky similarity_1 has a higher accuracy than others when calculating the

similarity metric for top-3 results. Therefore, Morgan fingerprints of radius 2 combined with Tversky similarity_1 calculation is the best combined method for top-3 fractional precision and recall results, attaining the highest precision of 26.37 percent and the highest recall of 45.25 percent. Likewise, Morgan fingerprints of radius 2 with features shows best ranking capability in top 5 candidates and top 10 candidates. However, the precision results from Morgan fingerprints of radius 3 are slightly higher than radius 2 on the top 20 candidates. Tanimoto similarity shows marginally better recall results for the top-5 and top-10 fractional recall results, whereas Tversky similarity 1 calculation performs better on precision. As the ability to rank candidates becomes less important as the number of candidates increases, the Morgan radius 2 fingerprint with features was selected in combination with the first Tversky similarity metric for use on the test set for this study.

4.4.4 Accuracy comparison

Table 4.2 Comparative assessment of BioSim's, GLORYx's, SyGMa's and BioTransformer's predictions of single-step biotransformation for 970 environmental chemicals.

	Method	Precision (%)	Recall (%)
Top3	BioSim	26.37	45.25
	GLORYx	12.81	21.62
	SyGMa	19.63	36.54
	Biotransformer	-	-
Top5	BioSim	19.91	50.15
	GLORYx	11.01	27.79
	SyGMa	17.05	48.81
	Biotransformer	-	-
Top10	BioSim	12.68	56.37
	GLORYx	8.86	37.88
	SyGMa	11.15	56.23

	Biotransformer	-	-
Top20	BioSim	7.96	65.67
	GLORYx	8.86	37.85
	SyGMa	7.28	60.31
	Biotransformer	12.52	40.34

Table 4.2 illustrates BioSim's, GLORYx's, SyGMa's and BioTransformer's predictions of single-step biotransformation for 970 environmental chemicals. Prior research summarized that the ranking of metabolites is the key challenge of knowledge-based approaches.(Kirchmair et al. 2015) To have a more detailed examination of performances of these models, we compare the precision rate and recall rate of these models based on different prediction rankings. The assessment was performed by comparing the precision (i.e. the fraction of true metabolites among the predicted ones) and recall (i.e. the fraction of true metabolites that were predicted over the total number of true metabolites) for each model. The precision indicates the validity of the results, while the recall indicates the completeness of the results.

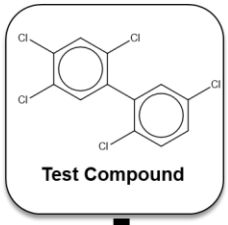
In particular, BioSim achieves a significantly higher precision rate (26.37%) for the top three predicted candidates than GLORYx (12.81%) and SyGMa (19.63%). Furthermore, BioSim achieves a significantly higher recall rate (45.25%) than GLORYx (21.62%) and SyGMa (36.54%). In addition to this higher precision and recall, the Top 5 sample results show that BioSim's precision rate (19.91%) is significantly higher than GLORYx's precision rate (11.01%) and SyGMa's precision rate (17.05%), and BioSim's recall rate (50.15%) is higher than GLORYx's recall rate (27.79%) and SyGMa's recall rate (48.81%). Such pattern of results is consistent when the candidates range gets larger. Results of Top 10 prediction also suggest that BioSim has the highest precision rate (12.68%) and the highest recall rate (56.37%), compared to GLORYx (precision rate is 8.86% and recall rate is 37.88%) and SyGMa (precision rate is 11.15% and recall rate is 56.23%). Lastly, we compare prediction performance of these models in the Top 20 sample. Results show that BioSim has a lower precision rate (7.96%) than GLORYx (8.86%) and Biotransformer (12.52%),

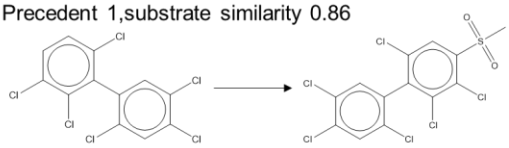
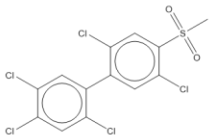
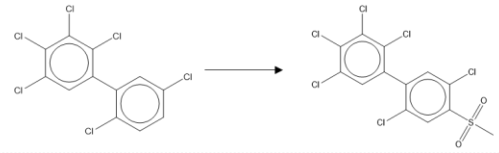
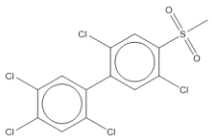
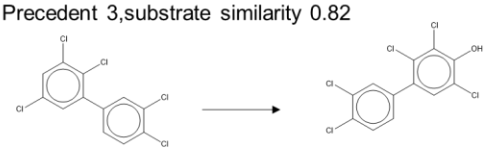
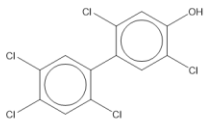
but a higher precision rate than SyGMa (7.28%). Consistent with results of previous samples, we find that BioSim's recall rate is the highest (65.67%), compared to GLORYx (37.85%), SyGMa (60.31%), and Biotransformer (40.34%). In summary, there is a significant difference in performance between the four tools, with BioSim outperforming other methods in both aspects.

4.4.5 Case study of BioSim prediction

Figure 4.5A shows an example of a biotransformation product predicted for PCB 101. After recalling up to 20 reaction precedents in order of decreasing product similarity, the precedent reaction site is extracted and matched against the target compound. Precedent reactions 1 and 2 both lead to methyl sulfone formation, which templates are from chemicals with similar structures. The Precedent reaction 3 provides another reaction product from hydroxylation reaction. The attempt to perform a response may fail for other preceding reactions. Alternatively, the similarity scores could be lower. At last, BioSim succeed to perform prediction to generate two products of PCB 101.

A



Reaction Precedent (sorted by substrate similarity)	Reacted Product	Final Score
Precedent 1,substrate similarity 0.86 		0.74
Precedent 2,substrate similarity 0.86 		0.74
Precedent 3,substrate similarity 0.82 		0.72
Precedent x…		

B

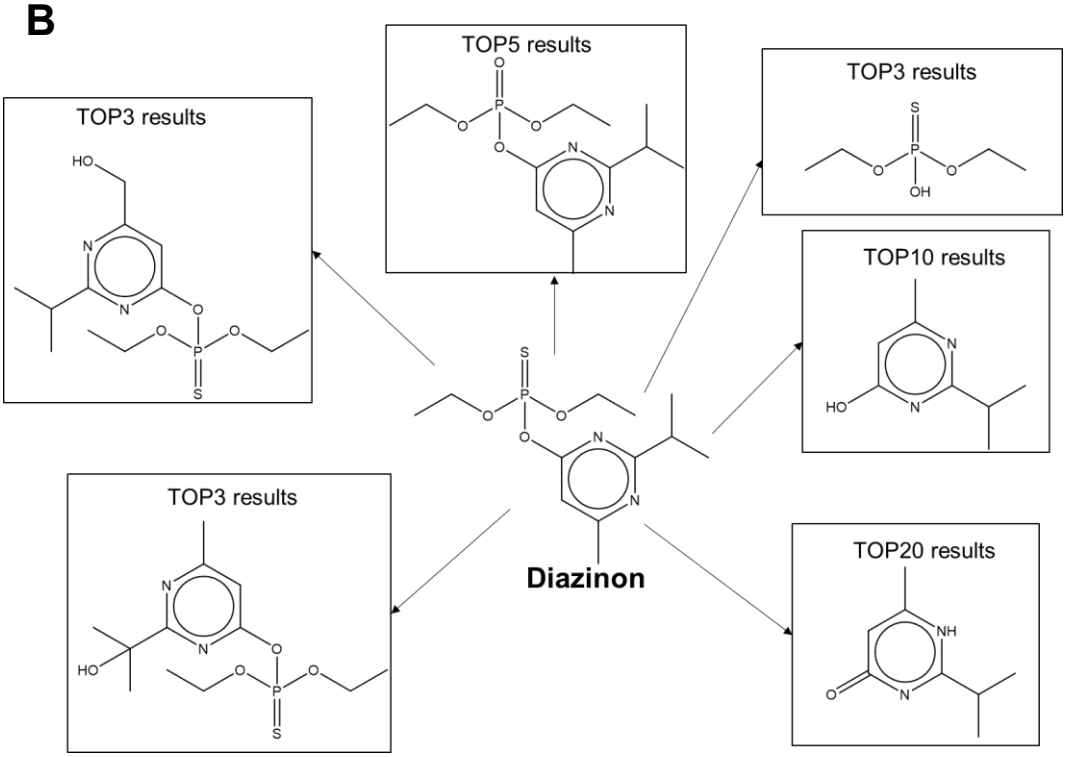


Figure 4.5 Examples of metabolites predicted by BioSim. (A) the workflow of how BioSim predict environmental chemical PCB 101. Similarity scores are shown using the Morgan radius 2 fingerprint with features. (B) the predicted results of Diazinon.

As can been seen from figure 4.5 B that presented a prediction example of Diazinon, 3 of 6 results can be accurately predicted by BioSim top 3 results. One of the remaining three can be properly predicted by the top five, one by the top ten, and one by the top twenty. Overall, BioSim exhibited outstanding and precise predicting accuracy on environmental chemicals.

4.4.6 Conclusions

In this work, we present BioSim, a freely available, open access software tool that enables the rapid, accurate, and comprehensive prediction of xenobiotic biotransformation, particularly for environmental chemicals. BioSim can also be accessed as a free web service at www.ecobiotrans.asia. The reactions predicted by BioSim cover both Phase I and Phase II metabolism. Overall, BioSim was shown to perform better than GLORYx, SyGMa and BioTransformer, which are three major prediction tools. For a long time, sorting and manually extracting templates has been key challenges for biotransformation prediction. Instead of having experts manually encode reaction templates, this study presents a method that can automatically extract reaction templates. In addition, for the first time, a ranking technique based on similarity measures has been proposed for biotransformation prediction. The simple and fast technique outperforms other strategies when it comes to candidate ranking. More importantly, BioSim's performance will improve with more reaction information input. In the future, BioSim will address many of the predictive metabolic needs of environmental scientists.

However, there are a few limitations of BioSim. First, more data, especially on drug metabolism, will improve the effectiveness of this technique, supporting the use of such tools in drug development for expediting and strengthening safety studies. Second, the site prediction is inaccurate at this point. It is possible to improve the accuracy of the procedure by integrating the SOM technique in the future.

4.5 Short Summary

Biotransformation is crucial for the activation, detoxification, and elimination of foreign chemicals, including pharmaceuticals as well as environmental pollutants. To date, prediction of biotransformation product is an active field of research due to its significant role. However, there is a knowledge gap in that almost all prediction tools were developed for pharmaceuticals, despite the fact that countless environmental substances also undergo biotransformation, and their metabolic pathways may differ slightly from pharmaceuticals. Furthermore, current methodologies are based on hand-encoding metabolic transformation rules, which require a significant amount of effort from experts. Here we presented a novel computational method with automatically extracted reaction template, which further ranked candidates by molecular similarity metric. Our work demonstrated molecular similarity to be an effective metric. Extensive evaluation revealed that the current model achieved 45.25% recall and 26.37% precision at the top 3 candidates and 65.67% recall and 7.96 % precision at the top 20 candidates, respectively. Our results indicate that the proposed approach can provide a simple and automatic solution that does not restrict to any expert knowledge. Lastly, BioSim is freely available as a web server at www.ecobiotrans.asia and is also provided as a Python package upon request.

Chapter 5. Competitive Biotransformation among Phenolic Xenobiotic Mixtures: Underestimated Risks for Toxicity Assessment

5.1 Overview

Humans are inevitably exposed to a complex mixture of organic contaminants (i.e., xenobiotics) through diet, environment and behavior. Biotransformation makes key contributions to the elimination of xenobiotics and greatly mediates the toxicity. However, most biotransformation studies were conducted using individual chemical, and whether co-exposure of multiple environmental chemicals will affect each other's fate in human body is still in its infancy. In this study, bisphenol A (BPA) was selected as a model compound. Its biotransformation was investigated under single/co-exposure to other phenolic xenobiotics (Triclosan (TCS), Tetrabromobisphenol A (TBBPA), and Bisphenol S (BPS)) in liver microsome and cell models. The result shows that binary exposures exhibit significant inhibitory effects on BPA metabolism, especially the sulfate conjugation. In combination of analysis on inhibition models and enzyme activity, the inhibition effect was suggested to be primarily driven by competition for metabolizing enzymes. A mixture with 22 phenolic chemicals was further examined to disrupt BPA at various human-relevant levels. Again, the results demonstrate significant inhibition on BPA metabolism, indicating the possible natural existence of our finding. Overall, our results show that biotransformation of phenolic xenobiotics can be significantly altered by co-exposure, which provides referential evidence on underestimated risks of simultaneous exposure to environmental toxicants.

5.2 Introduction

Humans are frequently exposed to a multitude of small molecules that are foreign to the body (xenobiotics), including dietary components, environmental chemicals, and pharmaceuticals (Koppel et al., 2017). A human blood exposome review reported occurrence of more than 1,000 small molecules with wide ranges of concentrations (Rappaport, Barupal, Wishart, Vineis, & Scalbert, 2014). Under current chemical risk assessment regime, parent compounds were mostly considered in the toxicity

characterization (Rodricks 2007). However, biotransformation is a crucial step to mediate the xenobiotics' toxicity and eliminate them by transforming them into metabolites with altered activities, lifetimes, bioavailabilities, and biological effects within the body. Meanwhile, biotransformation is accomplished by a limited number of enzymes with broad substrate specifies; therefore, many xenobiotics share common enzymes during metabolism in human (Klaassen et al. 2013). Drug-drug interactions by disrupting absorption, distribution, metabolism, or excretion (ADME) have been well documented over the past years and undesirable side effects could be involved in medication (Baxter 2010). Additionally, it has been reported in a more recent study that a human gut metabolite, *p*-cresol, reduced the effective systematic capacity of drug acetaminophen by competitive O-sulfonation (Clayton et al. 2009). Nonetheless, to date, there is a vast knowledge gap in understanding interaction among environmental xenobiotics in human metabolism.

Phenolic chemicals such as Bisphenol A (BPA), Bisphenol S (BPS), Triclosan (TCS), and Tetrabromobisphenol A (TBBPA) are an important group of ubiquitous environmental contaminants with wide detection in environment and human samples (Chen et al. 2016; Harrad et al. 2009; Yueh and Tukey 2016). BPA is used in the synthesis of polycarbonates, epoxy resins and thermal paper with more than 3.8 million tons annual production and a wide range of adverse health outcomes have been reported, including endocrine disruption, obesity, diabetes, hepatotoxicity, neurotoxicity, and immunotoxicity (Michalowicz 2014). With respect to BPS, a BPA replacement, its detected concentrations and frequencies in urine samples have been comparable to those of BPA (Zhou et al. 2014). Triclosan (TCS), as a broad-spectrum antimicrobial agent, has been extensively used in personal care products and have been recently phased out in US due to the concerns of hepatotoxic, mutagenic and carcinogenic effects (Administration 2016; Yueh and Tukey 2016). Tetrabromobisphenol A (TBBPA) is one of most extensively used flame retardant, with global use over 170,000 metric tons/year. The current data suggests that TBBPA may be also an endocrine disruptor and immunotoxicant and is currently under being heavily investigated in EU.("Scientific Opinion on Tetrabromobisphenol A (TBBPA) and its derivatives in food" 2011) Most phenolic xenobiotics are generally recognized to be rapidly detoxified to more water-soluble glucuronide and sulfate conjugates by

phase II metabolic enzymes (Ashrap et al. 2017). Due to the similar structure, previous studies show that BPA, BPS, TCS and TBBPA share common metabolic pathways, with major metabolites of sulfate and/or glucuronide conjugates detected in human serum and urine samples (Gerona et al. 2013; Ho et al. 2017; Liao and Kannan 2012; Oh et al. 2018; Provencher et al. 2014). Therefore, it is of great interest to investigate if co-exposure to above chemicals affect each other's biotransformation fate in human metabolism.

Hence, this work aims to investigate the mutual effects on biotransformation between phenolic xenobiotics (i.e., BPA, TCS, TBBPA, and BPS) in two *in vitro* models of liver microsome and HepG2 cell line due to the key role of liver in metabolism. To further examine its human relevance, a mixture of over 20 environmental-related phenolic chemicals was prepared at detected concentrations in human biological fluids samples and then applied to BPA xeno-biotransformation. All possible BPA metabolites in intra-cellular and extra-cellular medium under different treatments were measured and compared. Competitive and non-competitive drug-drug interaction model and key enzyme expression were used to further identify the mechanism of interaction. Our study provides detailed information on underestimated risks of simultaneous exposure of multiple toxicants, which should be considered in the toxicity assessment and toxicokinetics prediction.

5.3 Methodology

5.3.1 Chemicals and materials

All the solvents and reagents used in this study are HPLC grade or higher. Bisphenol A (Ring-$^{13}C_{12}$, 99%) was purchased from Cambridge Isotope Laboratories (MA, U.S.A). Bisphenol A monosulfate sodium salt (native and D6 labeled form), and Bisphenol A β-D glucuronide (native and ^{13}C labeled form) were purchased from Toronto Research Chemicals (North York, Canada). BPA, TCS, TBBPA, BPS and all other chemicals in prepared phenolic xenobiotic mixtures (see Table B8-9) were purchased from Sigma Aldrich (Singapore) with a purity > 97% unless stated otherwise.

5.3.2 Liver microsomal incubation

The incubation with rat liver microsome (RLM, Invitrogen Life Technologies, SG) was modified from previous studies (Fang et al. 2015; Karrer et al. 2018). For each substrate, we prepared at least 7 serial dilution from high concentration stock in DMSO and further diluted them from the highest concentration (8 μM or 50 μM or 250 μM) to the lowest concentration with DMSO content is <0.1% in the reaction mixture to a final volume of 200 μL, containing 50 mM phosphate buffer (PBS) at pH 7.4, 10 mM magnesium chloride, 10 mM DL-Dithiothreitol (DTT) and 0.05 mg/mL rat liver microsome protein. Due to the low solubility of tetrabromobisphenol A (TBBPA) in PBS buffer, 8 μM is the highest dose in microsomal incubation. For triclosan, 0-250 μM was employed in microsomal incubation but the measured concentrations were employed to replace the nominal concentration. For all the models and calculations, measured concentration was applied throughout this study if there is any over-saturation.

We first placed the reaction mixtures on ice during sample preparation and then pre-warmed them for 5 min at 37°C. The addition of mixture containing uridine 5'-diphosphate-glucuronic acid (final concentration: 5 mM), 3'-Phosphoadenosine-5'-phosphosulfate (final concentration: 50 μM), reduced glutathione (final concentration: 10 μM), nicotinamide adenine dinucleotide phosphate (final concentration: 5 mM) started the reaction. After a 30-min incubation, reactions were terminated via an addition of 200 μL ice-cold methanol. Samples were then centrifuged at 13,500 rounds per minute (rpm) for 30 min, and the supernatant was collected for further analysis.

5.3.3 Cell culture and exposure experiment

HepG2 human liver cancer cells (ATCC® HB-8065™) were grown in Dulbecco's Modified Eagle's Medium (DMEM, Gibco, SG) containing 10% heat-inactivated foetal bovine serum (FBS, Gibco Invitrogen), and maintained in a humidified 37 °C incubator with 5% CO_2. For the experiment, cells were seeded in 6-well plates containing a final volume of 3 mL medium. We seeded exactly the same number of cells in the beginning, which is approximately 0.3×10^6 cells. After 12 hours'

incubation, they were treated with DMSO control or substrates (DMSO: 0.1%) for another 48 hours, reaching ~1.2 $\times 10^6$ cells. All the selected concentrations or combination of any compounds for cell culture experiment (e.g., TCS< 10μM; BPA<20μM) after 48 hours exhibit no significant changes in cell numbers and viability, meaning the cell counts roughly maintained comparable as the control group.

Cell viability tests were conducted using Alamar Blue assay. When the exposure experiments were finished, 20 μL of Resazurin (0.74 mg/mL in PBS) was added directly to cell medium. Cells were covered with tinfoil and incubated at 37 °C, 5% CO_2 for 18 hours. The absorbance was read at florescence 580-610 nm, absorbance 600 nm. The cell viability was calculated by setting the control cells as 100%.

5.3.4 Solubility measurement of hydrophobic compounds in culture medium and PBS buffer

We measured the concentrations of all tested compounds in both PBS and cell culture medium to address the concerns of the over-saturation during the incubation. All the details can be found in Appendix B. Briefly, we have found TCS and TBBPA were over-saturated at high concentration in PBS buffer medium. However, their solubility in cell culture medium was much higher and all the measured concentration matched the nominal values. TCS and TBBPA at the high doses of 10 μM and 50 μM did not reach saturation in cell culture medium with 10% FBS.

5.3.5 Metabolite extraction

The metabolites extraction was modified from our previous study (Beyer et al. 2017). Briefly, 6-well plates were placed on ice and the medium was transferred into Eppendorf vial to further analyze the chemical or their metabolite level in the extra-cellular medium. To extract the metabolites in the intra-cellular medium, cells were rapidly washed with 1 mL ice-cold PBS and then harvested with a cell scraper in 2 mL ice-cold quenching mixture solution of methanol, acetonitrile and Milli-Q water (2:2:1, v/v/v) pre-spiked with isotope internal standards (^{13}C-BPA, BPAS-d6, and ^{13}C-BPAG) and finally transferred into a 2 mL Eppendorf vial. After that, samples

were vortexed followed by three cycles of shock-freezing-thawing and sonication at 4 °C for 10 min. Samples were subsequently incubated for 1 hour at -20 °C and later centrifuged at 13,000 rpm and 4 °C for 15 min to precipitate out proteins. The supernatant was dried in a vacuum concentrator (Labconco, Inc. Singapore) at 4 °C and reconstituted in ACN: H_2O (1:1, v/v). BPA metabolites were subsequently collected into LC vials and stored under -80 °C till analysis.

5.3.6 Instrumental Analysis and QA/AC

Analyses of all samples were performed using an Agilent 1260 HPLC system coupled with a 6550 Q-TOF mass spectrometry (Q-TOF MS) (Agilent Technologies, Singapore). Samples were analyzed with an Atlantis T3 column (3 mm, 2.1 × 100 mm) for MS analysis in electro-spray ionization (ESI) in both positive and negative mode. The mobile phase comprised of LC-MS grade Water with 2 mM of ammonium acetate (A) and acetonitrile (B). A gradient elution at a flow rate of 0.2 mL/min was used for the separation: 0-3 min: 95% of A; increase to 95% B at 14 min and maintain to 17 min; 17-18 min: decrease from 95% to 5% B and 1 min was used as a post-run equilibrium. The compounds and their metabolites were identified using either authentic standards or accurate mass and MS/MS fragmentations. All samples were in triplicates in this study. All samples in either microsome or cell study were prepared in triplicates in this study. The single or dual chemical dosing experiment was repeated at least twice and the result agreed with each other. For each batch, one procedural and laboratory blank was run alongside to correct for the background level during sample preparation and analysis. We also ran a pooled sample for every six samples to check the stability of the instrument during sample analysis, which was very important for the relative quantification of metabolites without standards or internal isotope standards. The calibration standards were also run before and after the working list with wide dynamic linear range to cover the concentration variation in samples. Any peak with single/noise ratio (S/N) < 10 was considered as noise. Since there was no cleanup step in the sample preparation, we assumed the recovery was 100% and matrix effect could be a big concern, especially for those without internal standard. However, we found the matrix effect for the metabolites were highly consistent between replicates (with RSD < 5%), which is probably due to

homogeneity of the cell or microsomal samples.

5.3.7 Reverse transcription polymerase chain reaction (RT-PCR)

Cells were first stabilized by RNA cell protection reagent (Qiagen, SG) and stored at -20 °C. Total RNA from HepG2 cells was isolated using TRIzol reagent according to the manufacturer's instructions (Invitrogen Life Technologies, SG). RNA was purified by DNase I (Thermo Scientific, SG) and reverse transcribed to cDNA using a High Capacity cDNA Reverse Transcription Kit (Thermo Scientific, SG) using the primers listed in Supplemental Table B6 for Sulfotransferase Family 1A Member 1 (SULT1A1), and membrane transporter, multidrug resistance protein 1 (MDR 1) gene expression. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as the house-keeping gene for the expression normalization. RT-PCR analysis was performed using SsoAdvanced^{TM} SYBR Green Supermix (Bio-Rad Laboratories, SG). Reactions were completed on a LightCycler 480 (Roche, SG).

5.3.8 Data processing

All samples were performed in triplicate for each substrate concentration. Michaelis-Menten constants of K_m and V_{max} were obtained using GraphPad Prism 8.0 (GraphPad Software Inc., San Diego, CA). The parameters were calculated based on the eq. 1

$$V = V_{max}[S]/(K_m + [S]) \quad (1)$$

where V_{max} is the maximum enzyme velocity and K_m is the substrate concentration at which the reaction rate is half of V_{max}; [S] is the substrate concentration at the enzyme active site.

ANOVA and Dunnett's post-hoc tests were also performed using GraphPad Prism 8.0 to determine statistical significance.

5.3.9 Prediction of inhibitory effect from in vitro data

The magnitude of the inhibition was estimated as competitive inhibition derived from

the Michaelis-Menten model. The calculation was based on the eq. 2

$$V = V_{max}[S]/(K_m\left(1 + \frac{[I]}{K_i}\right) + [S]) \quad (2)$$

Non-competitive inhibition model was applied using the eq. 3

$$V = V_{max}[S]/(K_m + [S])\left(1 + \frac{[I]}{K_i}\right) \quad (3)$$

where K_i is inhibition constant from the *in vitro* inhibition experiment; [I] is the inhibitor concentration at the enzyme active site; V_{max} and K_m are Michaelis-Menten constants obtained from *in vitro* experiment without inhibitors, V_{max} is the maximum enzyme velocity without inhibitors and K_m is the substrate concentration at which the reaction rate is half of V_{max}; [S] is the substrate concentration at the enzyme active site.

5.4 Results and discussion

5.4.1 Cell viability exposed to single compound and binary mixtures

The HepG2 cells were first incubated with different concentrations of BPA, TCS, TBBPA and BPS (0-100 μM); respectively. As shown in Figure 5.1A, TCS produced a relatively obvious impairment on cell viability at high concentrations. After 48-h exposure, 50 μM TCS resulted in statistically significant cell lethality and it was ~50% of the DMSO control (**$p < 0.01$). More reduction in cell viability was found when cells were exposed to 100 μM TCS. On the contrary, cell viabilities were not significantly altered by BPA, BPS and TBBPA even at maximum concentration (100 μM). Though, in the binary chemical exposure, cells co-exposed to 100 μM BPA with 100 μM BPS or 100 μM TBBPA concurrently exhibited significant cytotoxicity, with reduction of 79% and 74% in cell viability, respectively (Figure 5.1B). The binary exposure of 50 μM TCS and 50 μM BPA produced more cell lethality compared to cells exposed to 50 μM TCS individually. Overall, our results demonstrate that co-exposure to binary compounds resulted in amplified cytotoxicity.

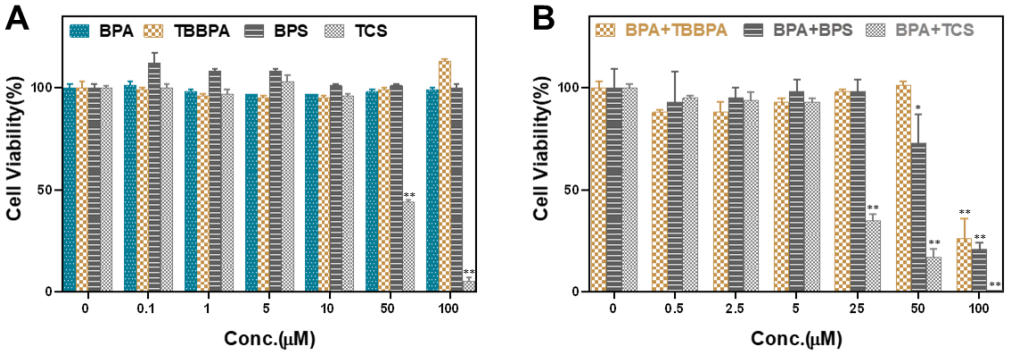


Figure 5.1 Viability of HepG2 cells exposed to single and binary compounds. (A) Exposure to BPA, TBBPA, BPS and TCS individually with different concentrations (μM); (B) Binary exposure to BPA with TBBPA, BPS or TCS at different concentrations. In Figure 5.1B, cells were exposed to BPA with either TBBPA, BPS or TCS at the concentration ratio of 1:1 at every point. Significant differences were indicated in comparison of the control. $*p < 0.05$, $**p < 0.01$ versus control group. The control group is exposed to 0.1% DMSO only. Data shown are the mean ± standard deviation of 4 replicates.

5.4.2 BPS, TCS and TBBPA inhibited biotransformation of BPA in incubation with rat liver microsomes (RLM)

We first evaluated and identified all the possible bio-transformation metabolites from BPA and other three phenol compounds. The predicted metabolic pathways of BPA, TCS, TBBPA, and BPS all comprised sulfate conjugates and glucuronide conjugates, according to the prediction model we presented forward in Chapter 4. Once the expected metabolite structures have been obtained, all the parent compounds and their metabolites were then identified using either authentic standards or accurate mass and MS/MS fragmentations. The result can be found in Table B1-4 and Figure B1-8. Briefly, with incubation of RLM, BPA was metabolized into the glucuronide conjugate, sulfate conjugate, and several other metabolites produced by cytochrome P450 dependent pathways, including hydroxy products of BPA, BPA sulfate, and BPA glucuronide. For TBBPA, BPS, and TCS, sulfate conjugate and glucuronide conjugate were the major metabolites as well. Besides, hydroxy forms of substrates, their sulfate conjugates and their glucuronide conjugates have also been identified.

In this study, we evaluated inhibition of BPA biotransformation under co-exposure to BPS, TCS, and TBBPA, by incubation with RLM. A substrate concentration range from 0 to 250/50/8 μM with or without 20 μM BPA was applied. Figure 5.2 (A.C.E) presents alterations of four typical BPA metabolites with different concentrations of binary compounds and Figure 5.2 (B.D.F) shows the Lineweaver-Burk plots for glucuronide conjugates of BPS, TCS and TBBPA, respectively. Overall, BPA metabolites went down with rising doses of BPS, TCS or TBBPA. The Lineweaver-Burk plots for BPS and TCS suggests inhibitions of competitive type.

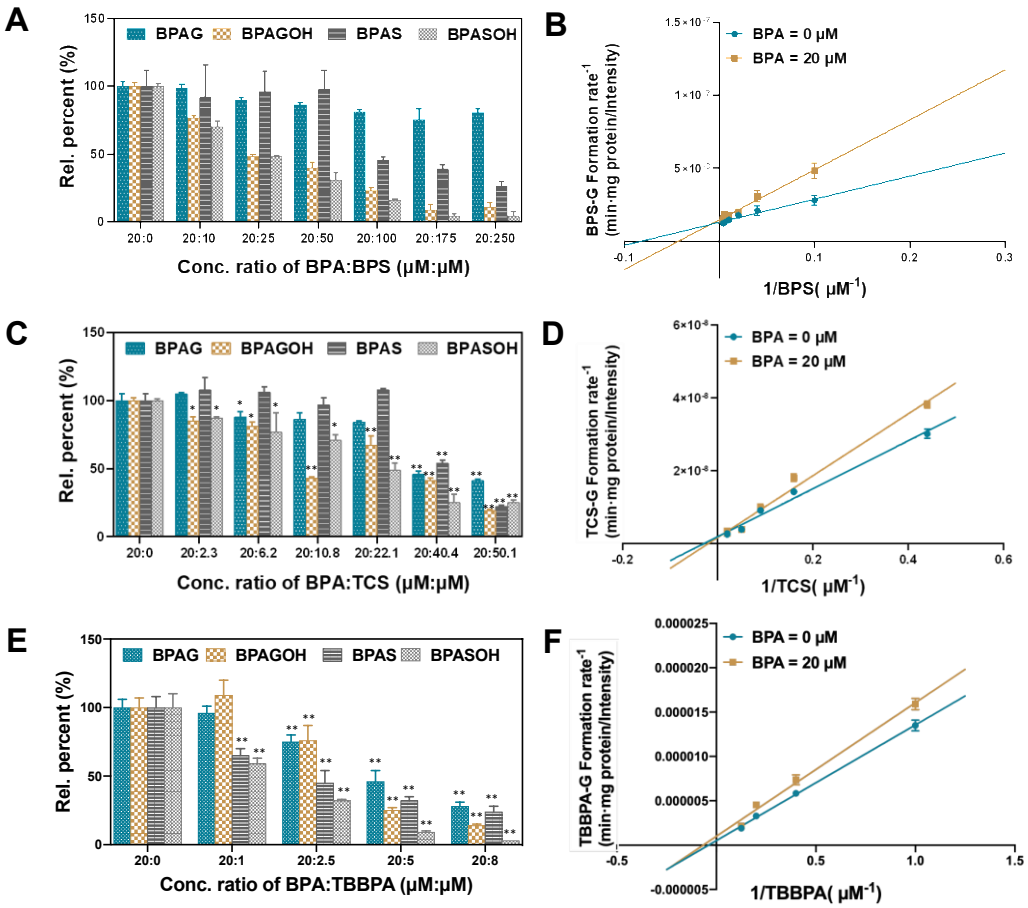


Figure 5.2 BPS, TCS and TBBPA inhibited biotransformation of BPA during incubation with rat liver microsomes. (A.C.E) BPA metabolites measured after incubation with rat liver microsomes by binary exposure to BPS (A), TCS (C), and TBBPA (E). (B.D.F) Lineweaver-Burk plots for Bisphenol A inhibition of BPS, TCS and TBBPA with incubation of rat liver microsomes (RLM). The control group is exposed to 20 μM BPA only and "Rel. Percent" represents the relative change

comparing with the control. Data shown are the mean ± standard deviation of triplicate experiment. *$p < 0.05$ **$p < 0.01$ versus control group.

Glucuronide conjugates were the major metabolites in incubation of all other three phenolic compounds due to the high abundance of uridine glucuronide transferase (UGT) in RLM (Klaassen et al., 2013). The V_{max} for BPA glucuronidation was determined to be 3.30 ± 0.16 nmol/ (min·mg), and the K_m was 71.28 ± 9.20 μM. Among all the BPA metabolites, BPASOH (hydroxy BPA sulfate) was the one that can be most significantly altered by exposure to other phenolic compounds. For example, BPASOH dropped to less than 50% with 2.5 μM TBBPA, 25 μM BPS or 22.1 μM TCS, while half reduction of BPAG was observed with 40 and 50 μM TCS as well as 5 μM and 8 μM TBBPA. The highest doses, 8 μM TBBPA and 50 μM TCS, produced ~ 70% inhibitory effects on BPAG. Under co-exposure with TBBPA, BPAS (BPA sulfate) and BPASOH (hydroxy BPA sulfate) declined more noticeably, as fell to ~30% at 5 μM and 2.5 μM further declined to 24% and 3% with 8 μM TBBPA, respectively. For BPAG (BPA glucuronide) and BPAGOH (hydroxy BPA glucuronide), significant inhibition starts to be observed from 2.5 μM TBBPA. The highest 8 μM TBBPA resulted in approximately 70% diminution of BPAG and BPAGOH. On binary exposure of BPS and BPA, BPAGOH and BPASOH deceased relatively dramatically, starting from 10 μM BPS and falling to 11% and 4% at 250 μM; respectively. Significant reduction of BPAS was found from 100 μM BPS and further decreased to 26% at 250 μM. Similarly, BPAGOH and BPASOH dropped dramatically with co-exposure to TCS. Specifically, 2.3 μM TCS significantly reduced 15% formation of BPAGOH and BPASOH whereas BPAS was significantly inhibited to 55% at 40 μM. At 50 μM, BPAGOH and BPASOH decreased to 19% and 25%, respectively. Besides, we also identified that the biotransformation of BPS, TCS, or TBBPA was inhibited by BPA. At varying concentrations of BPS, TCS, or TBBPA, their major glucuronic conjugate was significantly lower when co-exposed to 20 μM BPA (Figure 5.2 B.D.F). The Lineweaver-Burk plots suggest that the inhibition types were likely to be competitive.

5.4.3 BPS, TCS and TBBPA inhibited biotransformation of BPA in HepG2 cells

We further conducted BPA biotransformation under similar treatment of TBBPA,

BPS and TCS in HepG2 cell line, as results illustrated in Figure 5.3. Specifically, HepG2 cells were exposed to 0-50 μM BPS, 0-10 μM TCS, and 0-50 μM TBBPA with 20 μM BPA or without BPA. All the selected concentrations are not supposed to induce any cytotoxicity (as shown in Figure 5.1). Consistent with previous studies on HepG2 cells and biomonitoring data, sulfate conjugates were observed to be the predominant metabolites(Bursztyka et al. 2008; Liu et al. 2017). The V_{max} for BPA sulfation was determined to be 1.41 ± 0.13 ng/ (h·10^6 cells), and the K_m was 27.88 ± 5.20 μM.

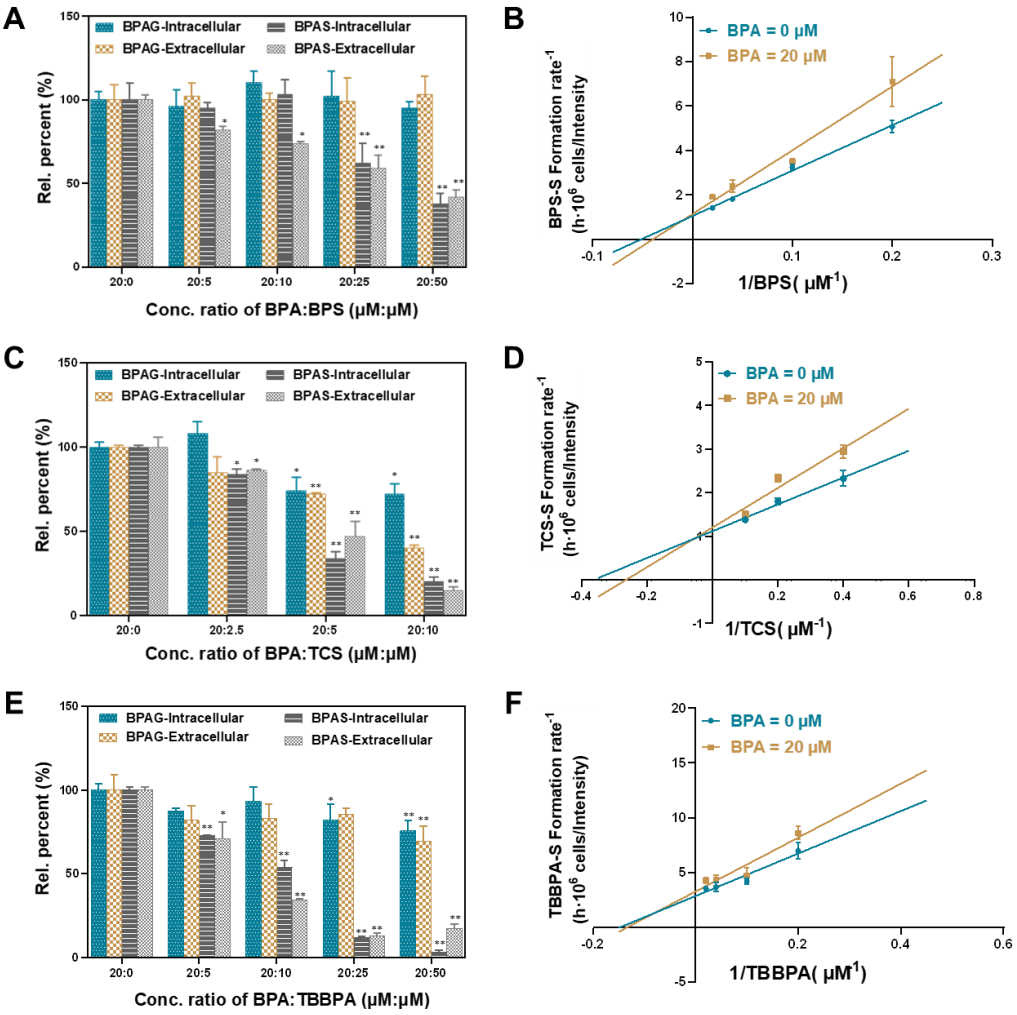


Figure 5.3 BPS, TCS and TBBPA inhibited biotransformation of BPA on HepG2 cells. (A.C.E) BPA metabolites were measured after incubation with HepG2 cells treated with binary compounds at various concentration ratios. The control group is exposed to 20 μM BPA only. (B.D.F) Lineweaver-Burk plots for inhibition of BPS-S (i.e.,

sulfate conjugates), TCS-S and TBBPA-S formation by Bisphenol A in HepG2 cells. Data shown are the mean ± standard deviation of triplicate experiment. *$p < 0.05$, **$p < 0.01$ versus control group.

An analysis of BPA metabolites revealed significant inhibition on biotransformation between binary exposure of BPA with other three phenolic compounds. As shown in Figure 5.3, both intracellular and extracellular BPAS (BPA sulfate) declined in presence with either BPS, TCS, or TBBPA, which further reinforced the fact of competition among BPA and other phenolic compounds. The reducing pattern is in good agreement with microsomal data, though significant decreases start from much lower concentrations: 5 μM TBBPA, 5 μM BPS and 2.5 μM TCS. With co-exposure with 25 μM TBBPA, 50 μM BPS or 5 μM TCS, BPAS decreased to less than half of the control. Correspondingly, reduction was also found in conjugation of TBBPA, TCS and BPS. With the presence of 20 μM BPA, the representative Lineweaver-Burk plots for the inhibition of BPS-S, TCS-S, and TBBPA-S suggest the mutual inhibition effect of BPA on other three phenolic compounds (Figure 5.3 B.D.F). Besides, the extracellular data are highly correlated with the intracellular data, indicating that the reduction of metabolites formation was not possibly due to the accelerated extracellular transportation.

Unlike sulfate conjugates, glucuronide conjugates were less affected by binary exposure. During incubation with TCS, a significant decrease of BPAG begins from 5 μM with an inhibition effect approximately 30%. However, there was no significant reduction observed with BPS at varying concentrations for BPAG. In addition, the biotransformation of TCS, BPS or TBBPA was also found to be inhibited by BPA.

5.4.4 The inhibition of BPA transformation was due to competitive enzyme binding rather than enzyme activity

The inhibition effect of other phenolic compounds on BPA in the cell model could be explained by either enzyme competition or down-regulation of enzyme expression. To figure out this inhibition mechanism, we first examined the enzyme expression using RT-PCR. One major metabolizing enzyme participated in metabolism of phenolic compounds in HepG2 cells is sulfotransferases (SULTs) and

SULT 1A1 has been reported to play an important role in conjugating sulfonic acid (Wu et al. 2011). Sulfate conjugated compounds are also the metabolite that have been most sensitive in biotransformation in the binary exposure. For the detoxification process, after hydrophilic part conjugated to toxicants, excretion would be facilitated via membrane transporter protein in phase III of their metabolism, with the anionic groups acting as affinity tags for a variety of membrane transporters (Klaassen et al. 2013). Therefore, we investigated the expression levels of SULT 1A1 and multidrug resistance protein 1(MDR 1) in cells under single exposure to 20 μM BPA, 25 μM BPS, and 25 μM TBBPA as well as binary exposure to 20 μM BPA with 25 μM BPS, and 20 μM BPA with 25 μM TBBPA. As shown in Figure 5.4A, the expression of MDR 1 was not significantly altered by exposure to any single or binary compounds. In contrast, the enzyme expression of SULT1A1 was enhanced by binary exposure. Specifically, the treatment of 20 μM BPA increased the SULT 1A1 expression by 1.8 folds and combination with 25 μM TBBPA and 25 μM BPS further increased it to ~3 folds. These results show that exposure to phenolic xenobiotics induce the expression level of some metabolizing enzymes. Thus, inhibitory effect would not be likely a result of down regulation of metabolizing enzymes.

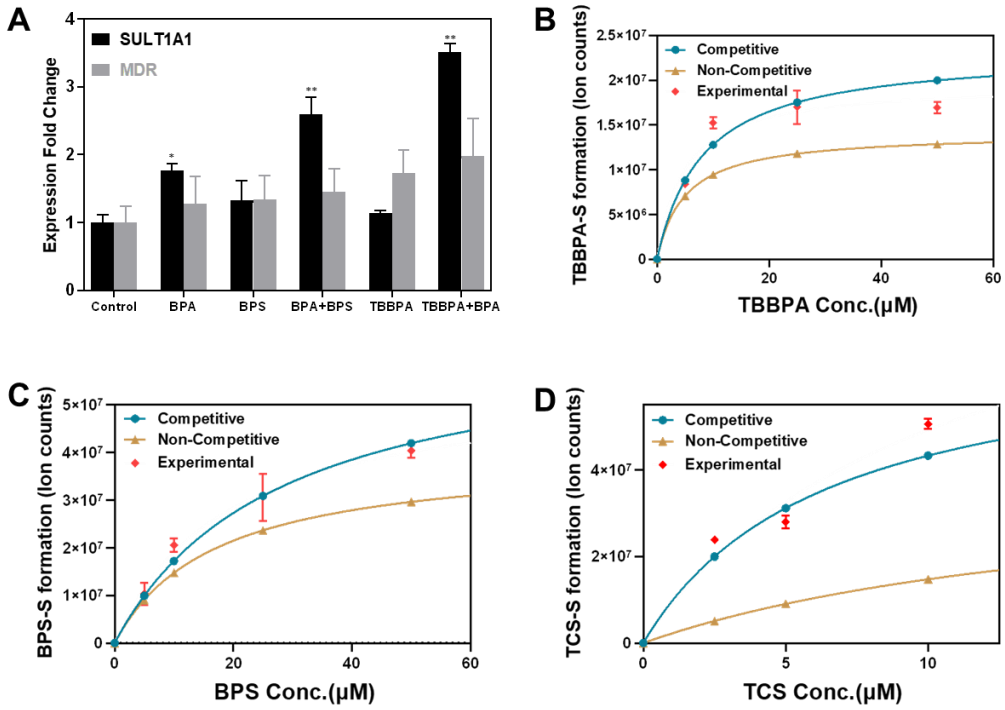


Figure 5.4 (A) Phenolic xenobiotics regulate gene expression of xenobiotic metabolizing enzyme SULT 1A1 and MDR 1. mRNA expression of SULT 1A1 and MDR 1 was measured after single or binary chemical treatment in HepG2 cell line. Data are the mean ± standard error; n = 6. *$p < 0.05$, **$p < 0.01$ versus control group. (B-D) Quantitative prediction of inhibition potential by two inhibition models. (B) Quantitative prediction of TBBPA metabolite formation and comparison with experimental data from HepG2 cell. Competitive inhibition model $R^2 = 0.94$; non-competitive inhibition $R^2 = 0.58$. (C) Quantitative prediction of BPS metabolite formation. Competitive inhibition model $R^2 = 0.82$; non-competitive inhibition: no correlation. (D) Quantitative prediction of TCS metabolite formation. Competitive inhibition model $R^2 = 0.92$; non-competitive inhibition $R^2 = 0.47$.

To further investigate whether the inhibition is competitive or non-competitive, we have used different inhibitive models derived from the Michaelis-Menten equation (See methodology) to predict the formation of sulfate conjugates of those phenolic compounds. The results are further compared with the experimental data obtained from HepG2 cell line. As evident from Figure 5.4 (B-D), the competitive inhibition models provide a more robust prediction comparing with non-competitive inhibition

for all three tested inhibitors. For BPS metabolites formation, the correlation coefficient R^2 between experimental data and competitive inhibition model is 0.94 whereas the R^2 of non-competitive inhibition is 0.58. For TBBPA, competitive R^2 is 0.92 and non-competitive R^2 is 0.47. The difference of TCS metabolites formation is even more distinct. All these results indicate that BPA, as well as other phenolic compounds in this study underwent competitive inhibition and the metabolite formation under binary/multiple exposure could be predicted by parameters obtained in *in vitro* experiment of single chemical.

5.4.5 Phenolic xenobiotic mixtures at environmentally relevant exposure levels alter BPA biotransformation

In the incubation of BPA and other three phenolic compounds, highly competitive binding to xeno-biotransformation enzymes was observed at moderate μM levels. This level is good to conduct mechanism study; however, it is unlikely to be environmentally relevant. As mentioned above, humans are simultaneously exposed to hundreds of environmental chemicals and countless xenobiotics share the metabolic pathways like sulfonation and glucuronidation. In this study, we further explore whether the phenolic xenobiotics at environmentally relevant exposure levels would affect the xenobiotic transformation. To better mimic human exposure, we prepared phenolic and benzoic acid compound mixtures containing 22 phenolic compounds at environmentally relevant exposure levels, which were obtained from extensive bio-monitoring data of published literature (See Table B8 and Text B5). All these xenobiotic compounds are commonly found in our daily life, including plasticizers (e.g., BPS and Phthalates), personal care products (e.g., parabens, triclosan, triclocarban) and others (e.g., benzoic acid). Geometric mean (GM), 0.5-fold of GM, maximum (MAX), and 5-fold of maximum (5MAX) levels of mixtures in the biological fluids were adopted at incubation with microsome while GM, 0.5-fold, 2-fold, 4-fold, 8-fold of GM, maximum (MAX) were applied to HepG2 cells. None of those tested concentration showed any obvious cell toxicity. For BPA, 20 μM was used in microsomal incubation and 20, 1, 0.5 and 0.1 μM were selected for HepG2 cells. Previous studies reported human detected BPA values with a range between non-detected to 0.5 μM and BPA with 0.5 and 0.1 μM is within the human

relevance levels (Schönfelder et al., 2002). The impacts of phenolic xenobiotic mixtures on BPA biotransformation products were demonstrated in Figure 5.5. In general, increasing exposure concentration of the phenolic xenobiotic mixtures impeded the biotransformation of BPA.

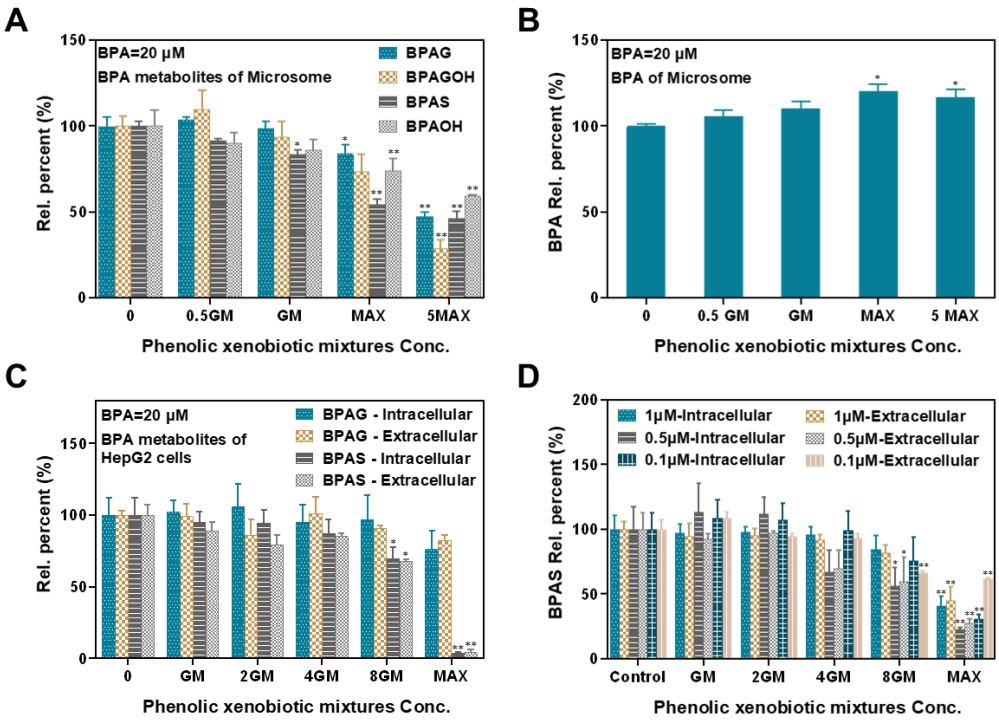


Figure 5.5 Phenolic xenobiotic mixtures at environmentally relevant exposure levels inhibited biotransformation of BPA. (A-B) BPA metabolites and BPA were measured after incubation with rat liver microsomes (RLM). Cells were exposed to 20 μM BPA with or without phenolic xenobiotic mixtures. (C-D) BPA metabolites in cell extracts and extracellular medium were measured after incubation with HepG2 cells. Cells were exposed to 20 μM (C), 1 μM, 0.5 μM and 0.1 μM BPA (D) with or without phenolic xenobiotic mixtures. Geometric mean (GM), 0.5-fold (0.5GM), maximum (MAX), and 5-fold maximum (5MAX) levels of mixtures from previous studies on the environmentally relevant exposure levels were adopted in microsomal experiment and GM, 2-fold (2 GM), 4-fold (4 GM), 8-fold of GM (8 GM), and MAX were used in the cell study. Data are the mean ± standard error; n = 3. The control group is exposed to 20, 1, 0.5 and 0.1 μM BPA only. *$p < 0.05$, **$p < 0.01$ versus control

group.

After incubation of BPA and phenolic xenobiotic mixtures with RLM, as the most major metabolites, a significant reduction of BPAG (BPA glucuronide) was observed from MAX concentration and became more dramatic at 5 MAX. Compared to the control group, BPAG dropped to approximately 80% and 50% of the control for MAX and 5MAX; respectively. Similar pattern can be observed in hydroxy BPA glucuronide (BPAGOH) and the decline was more dramatic, with reductions of 27 % and 71% at MAX and 5-fold MAX; respectively. Unlike above, we observed that the reduction of BPA sulfate (BPAS) starts from GM (83%), and further declined at MAX (54%) and 5-fold MAX (50%). Significant decrease of BPAOH (hydroxy BPA) was also detected with mixtures at MAX and 5-fold MAX. Meanwhile, the unconjugated BPA went up to ~120% of the control group with phenolic xenobiotic mixtures at MAX and 5-fold MAX, which suggested a saturated conjugation pathway inhibited the biotransformation of BPA.

For HepG2 cells, overall, a declining tendency of BPAS with increasing dose of phenolic xenobiotic mixtures was observed as well. After incubation of 20 μM BPA and phenolic xenobiotic mixtures, a significant reduction of BPAS was observed at 8GM and MAX. Specifically, the relative abundance of BPAS was ~70 % of the control group at 8 GM exposure level and further decreased to less than 10% at MAX point. A similar observation was found in the binary exposure of phenolic xenobiotic mixtures and BPA at lower concentrations of 1, 0.5 and 0.1 μM. Significant inhibitory effect was also observed at high GM and MAX concentration and the amount of BPAS dropped to less than half of the control group. Overall, significant inhibition effect of BPA can be observed by co-exposure to phenol mixtures at environmentally relevant exposure levels. Changes in cell extracts and extra-cellular medium could reflect the changes in the liver organ and plasma *in vivo*; respectively.

5.4.5 Implications and Limitations

Biotransformation of xenobiotics plays a vital role in mediating the toxicity. The decreasing BPAS and increasing BPA due to co-exposure implies that the addition of multiple xenobiotics sharing same pathways can inhibit the biotransformation

process of other compounds. Through comparing the expression of a few metabolizing enzymes and different inhibitive models derived from Michaelis-Menten equation, it suggests the altered BPA metabolism was likely to be competitive inhibition-driven. Thus, this competitive enzyme activities may lead to the prolonged effect of contaminants themselves. The enhanced toxicity due to poor biotransformation was reported in many drugs such as acetaminophen and rifampin (Clayton et al., 2009; Corsini & Bortolini, 2013). In terms of BPA, our cytotoxicity result suggested there was increased toxicity upon dual chemical treatment. To date, it is still unclear whether the toxicity of BPA is due to itself or its metabolites. In this study, among all BPA metabolites identified, 4,4'-(4-methylpent-2-ene-2,4-diyl) diphenol (MBP) is the most toxic metabolite, which was reported to be up to 1,000 times more potent as an estrogen than BPA (Moreman et al., 2018; Yoshihara et al., 2004). However, the MBP amount stayed relatively stable upon dual treatment, comparing with other deceasing metabolites like sulfate conjugates. The toxicity associated with altered biotransformation of BPA for various toxicological end points should be further investigated in the future.

The example of BPA can be extrapolated to other environmental chemicals which shares overlapping biotransformation enzymes. Recent scientific evidence on human exposome shows that we are constantly exposing to ever-increasing chemicals, both endogenous and exogenous ones, from food, consumer products and the environment (Athersuch and Keun 2015). In a review of human blood exposome, more than 1,000 small molecules are found in human blood with wide ranges of concentrations (Rappaport et al., 2014). In another study, >600 possible putative exposure biomarkers have been found in a human pooled urine (Jia et al. 2019). It is of great interest to investigate how all the biotransformation of those compounds can be affected by the resting super complex mixtures. In this work, we found that phenolic xenobiotic mixtures at environmentally relevant exposure levels could alter BPA biotransformation; however, the inhibition mechanism still demands to be settled by further studies. Overall, these results suggested the disrupted bio-transformation due to simultaneous exposure to xenobiotic mixtures should be considered in future risk assessment.

However, it should be noted that there are several limitations of this study. One of them is the relevance between *in vitro* and *in vivo* models. For example, the enzyme activity in *in vivo* model can be more effective than the *in vitro* model, which might make the enzyme competition less intensive (Zotter et al. 2017). For another, the enzyme activity of liver cancer cell line, HepG2 cells, might not be exactly identical to *in vivo* human liver cells. Besides, the rat liver microsomes might not fully simulate real biotransformation in human body. Though by combination of data from cells and microsomes, we have a whole picture of competitive inhibition of these phenolic compounds; however, further investigation using *in vivo* model should be considered in future studies. In addition, we have observed some over-saturation of dosed chemicals at high concentration, especially in the microsome study. To generate slow and reliable release of hydrophobic chemicals like TBBPA, future research can explore passive dosing scheme on these chemicals (Schmidt 2013; Smith et al. 2012). Furthermore, the increased toxicity might be related to many other factors, not only reduced biotransformation of BPA. Future studies should include more research into other parts of the influence mechanism.

5.5 Short Summary

This study employed bisphenol A (BPA) as a model compound to investigate its biotransformation under single/co-exposure to other phenolic xenobiotics (Triclosan (TCS), Tetrabromobisphenol A (TBBPA), Bisphenol S (BPS)) and a mixture with 22 phenolic chemicals in liver microsome and cell models. The results suggest that binary exposures exhibit significant inhibitory effects on BPA metabolism, especially the sulfate conjugation. In combination of further analysis on inhibition models and enzyme activity, the inhibition effect was suggested to be primarily driven by competition for metabolizing enzymes. The mixture with 22 phenolic chemicals was further examined to disrupt BPA at various human-relevant levels. Again, the result demonstrates significant inhibition on BPA metabolism, indicating the possible natural existence of our finding. Overall, our results show that biotransformation of phenolic xenobiotics can be significantly altered by co-exposure, which provides referential evidence on underestimated risks of simultaneous exposure to environmental toxicants.

Chapter 6. Gut Microbial Metabolite *p*-Cresol Alters Biotransformation of Bisphenol A: Enzyme Competition or Gene Induction?

6.1 Overview

Recent studies on pharmaceuticals have revealed the direct and indirect mechanisms that link the human gut microbiome to xenobiotic biotransformation. Though environmental contaminants compose a vital portion of xenobiotics and share overlapping biotransformation pathways with gut microbial metabolites, the possible interplay between gut microbiome and biotransformation of environmental contaminants remains obscure. This study utilized bisphenol A (BPA) and *p*-cresol as model compounds to explore whether gut microbial metabolites could affect environmental phenol metabolism on both *in vitro* and *in vivo* models. Some distinct biotransformation behavior has been observed between the *in vitro* and *in vivo* results. *In vivo* mouse examination using *p*-cresol injection exhibited enhancing effect on BPA metabolism, which is rarely found in mixture studies. Serum level of BPA was significantly lower in the *p*-cresol group with a clear dose-response relationship. However, in both *in vitro* models of liver S9 fractions and HepG2 cell line, *p*-cresol is found as a strong inhibitor in a non-competitive pattern for BPA biotransformation. *In silico* docking approach was utilized to explore the non-competitive inhibition mechanism by estimating the binding affinity of *p*-cresol, BPA and two allosteric binding site of key enzyme SULT 1A1. Next, a close investigation revealed that the expression of biotransformation enzyme genes including *Ugt1a1, Ugt2b1, or Sult1a1* were dynamically induced after the *p*-cresol treatment. Overall, our results provided a novel insight into the biotransformation interaction between gut microbiome and environmental contaminants. The discrepancy between *in vitro* and *in vivo* models suggested that both enzyme competition and dynamic gene regulation should be considered when evaluating the interaction between gut microbiome and environmental xenobiotics.

6.2 Introduction

Xenobiotic metabolism is key to the activation, detoxification, and elimination of

foreign chemicals like environmental pollutants, pharmaceuticals, food additives and biologically foreign substances, as it modifies these compounds into metabolites with different physicochemical and pharmacological properties (Klaassen et al. 2013). Recent studies have revealed that the gut microbiota can influence the metabolism of drugs, which may change their efficacy and/or side effect profiles. Direct metabolisms including the transformations of xenobiotics by gut microbes (e.g., reduction, hydrolysis, and acylation) have been intensively documented over the past years. Nevertheless, compounds that resist microbial metabolism can still be influenced by the gut microbiota through several indirect mechanisms (Spanogiannopoulos et al. 2016). Indirect mechanisms involve more complex host–microbial interactions that modulate host pathways for xenobiotic metabolism or transport. For example, comparisons between germ-free and colonized mice have revealed that the microbiota affects the expression of several host genes involved in biotransformation (Hooper et al. 2001). Moreover, some gut microbial metabolites are processed by mammalian system in a manner analogous to xenobiotics. This might lead to increased toxicity or half-life of xenobiotics due to competition between the drug and microbial metabolites for the same enzymes or transporters that are involved in drug detoxification or elimination. An example of this type of interaction can be observed with acetaminophen (one of the most widely used drugs worldwide) and *p*-cresol (an end product of tyrosine and phenylalanine metabolism) (Wikoff et al. 2009). A previous pharmacometabonomic study shows that the competition between *p*-cresol and acetaminophen impedes the ability of the host to detoxify acetaminophen, which might lead to severe hepatocellular damage (Clayton et al. 2009). *p*-Cresol is a frequently detected compound in human plasma that the serum *p*-cresol level of healthy people is 8.6 ± 3.0 μM (mean ± SD), whereas the serum *p*-cresol levels of continuous ambulatory peritoneal dialysis (CAPD), uremic, and hemodialysis patients are 7–10 times higher, which could reach 62.0 ± 19.5, 87.8 ± 31.7 and 88.7 ± 49.3 μM, respectively (De Smet et al. 1998).

Along with pharmaceuticals and dietary compounds, environmental chemicals compose an important portion of xenobiotics. Moreover, their metabolic pathways are often overlapping. For instance, *p*-cresol, acetaminophen and bisphenol A are all substrates of the human cytosolic sulfotransferase 1A1 (SULT1A1) (Gamage et al.

2006). Based on established knowledge in drug-drug interaction, chemicals that undergo overlapping metabolic pathway are more likely to interfere with each other (Ryu et al. 2018). Our previous work investigated such interaction between environmental phenolic xenobiotics that share similar metabolic pathway, and demonstrated that a mixture of 22 phenolic chemicals at human-relevant levels significantly disrupts BPA biotransformation (Peng et al. 2019). Tyler Pollock et al. also proved that concurrent exposure to tetrabromobisphenol A (a flame retardant) and triclosan (an antimicrobial agent) elevated concentrations of dietary BPA in mouse serum and reproductive tissues (Pollock et al. 2017).

However, the possible interplay between the gut microbiome and biotransformation of environmental contaminants still remains obscure. Hence, in this study, BPA was selected as a model environmental compound as a widely detected contaminant with extensively investigated metabolic pathways involving multiple Phase I & II enzymes (cytochromes P450, UGTs and SULTs) while *p*-cresol was chosen as a model gut microbial metabolite due to its frequent occurrence and high detection (Michalowicz 2014). First, we examined the effects of *p*-cresol on BPA metabolism on *in vivo* mouse model. Then, the *in vitro* investigation was conducted on liver S9 fractions and HepG2 cell line. To explore the mechanism underlying *p*-cresol's high affinity to sulfotransferase (SULT) and its non-competitive inhibition pattern, we utilized an *in silico* molecular docking approach and examined cresol isomers. Lastly, we explored for some key metabolism gene expression of mouse livers. To our knowledge, this is the first study to report the biotransformation interaction between environmental chemicals and gut microbial metabolites.

6.3 Methodology

6.3.1 Chemicals and Materials

All chemicals and solvents were purchased from Sigma Aldrich (Singapore) with a purity of >97%, unless stated otherwise. The isotopic standard, bisphenol A (Ring-$^{13}C_{12}$, 99%) was purchased from Cambridge Isotope Laboratories (Andover, MA, United States). Bisphenol A monosulfate sodium salt (native and D6 labeled form), and Bisphenol A β-D glucuronide (native and ^{13}C labeled form) were purchased from Toronto Research Chemicals (North York, Canada). HPLC grade water, methanol, and acetonitrile (ACN) were purchased from Thermo Fisher Scientific (Singapore).

6.3.2 Incubation of chemicals with liver subcellular fractions

Rat (Sprague-Dawley) S9 fractions was purchased from Invitrogen Life Technologies (Singapore). The experimental procedures were modified from our previous studies.(Fang et al. 2015; Peng et al. 2019) The single-compound stock solution of each chemical was prepared in DMSO and further diluted to <0.5 % in the reaction mixture to a final volume of 150 μL. The reaction mixtures containing 50 mM phosphate buffer at pH 7.4, 10 mM magnesium chloride, 10 mM dithiothreitol and rat liver subcellular fractions (S9 fractions, final concentration: 0.1 mg/mL) were kept on ice during sample preparation. Subsequent to 5-min pre-warming step, the reaction was initiated by addition of enzyme cofactor mixture comprised of uridine 5'-diphosphate-glucuronic acid (final concentration: 2 mM), 3'-phosphoadenosine-5'-phosphosulfate (final concentration: 0.5 mM), reduced glutathione (final concentration: 1 mM), nicotinamide adenine dinucleotide phosphate (final concentration: 5 mM). The incubation was carried out at 37 °C for 30 min and terminated by addition of 150 μL ice-cold methanol. The samples were then centrifuged at 13,500 rpm for 30 min, and the supernatant was collected for further analysis. A full dose-response experiment was conducted for *p*-cresol by incubation of 10 μM BPA with 0, 15, 25, 50 and 175 μM *p*-cresol. The experimental procedures of cresol isomers (*m*-cresol and *o*-cresol) were conducted in the same fashion. Conversely, BPA at concentrations ranging from 0, 5, 10, 25, 50 and 100 μM in combination with 10 μM cresol were also incubated in liver S9 fractions.

6.3.3 Cell culture, exposure experiment and metabolite extraction

HepG2 human liver cancer cell (ATCC® HB-8065™) was obtained from the American Type Culture Collection (ATCC, Rockville, MD, USA). Cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM, Gibco, SG) containing 10% heat-inactivated fetal bovine serum and maintained in humidified 37 °C atmosphere containing 5% CO_2. The cells were seeded in 6-well plates with a density of 300,000 cells/well for 12 h and then treated with DMSO control or substrates (final DMSO: 0.1%) for another 48 h. Substrates were 10 μM BPA with 0-100 μM *p*-cresol, *m*-cresol or *o*-cresol, and 0-50 μM BPA with/without 5/25 μM *p*-cresol. Metabolites were subsequently extracted following our prior studies using ice-cold quenching mixture solution of methanol, acetonitrile and Milli-Q water (2:2:1, v/v/v) pre-spiked with isotope internal standards (^{13}C-BPA, BPAS-d6, and ^{13}C-BPAG) (Beyer et al. 2017; Peng et al. 2019). BPA at concentrations ranging from 0, 2.5, 5, 10, 25 and 50 μM with 10 μM cresol were also applied in HepG2 cells due to cytotoxicity concern.

6.3.4 Quantification of BPA and its metabolites in in vitro models

Analyses of all samples from *in vitro* experiment were performed using an Agilent 1260 HPLC system coupled with a 6550 Q-TOF mass spectrometry (Q-TOF MS) (Agilent Technologies, Singapore). Samples were analyzed with an Atlantis T3 column (3 μm, 2.1 × 100 mm, Waters) for MS analysis in electro-spray ionization (ESI) in negative mode. The LC mobile phases were water containing 2 mM of ammonium acetate (A) and 100% acetonitrile (B). The mobile phase flow rate was 0.2 mL/min and the gradient was as follows: 0-3 min: 95% of A; increased to 95% B at 14 min and maintained to 17 min; 17-18 min: decreased from 95% to 5% B and 1 min was used as a post-run equilibrium. The compounds and their metabolites were identified using either authentic standards or accurate mass and MS/MS fragmentations. The detailed qualification and quantification of those BPA metabolites and *p*-cresol sulfate can refer to our previous study and Appendix B (Peng et al. 2019). BPA, BPAS, BPAG were identified using authentic standards and absolutely quantified using the isotope-spiking internal standard method. For other identified metabolites without authentic standards, we confirmed their identity with

exact m/z, MS/MS fragmentations, and our in-house retention time prediction models (Peng et al. 2019; Yang et al. 2020). Semi-quantification of those compounds were conducted using its response during LC-MS analysis.

6.3.5 Animal treatment and BPA quantification

Animal experiments in this study were approved by the Institutional Animal Care and Use Committee of Jinan University (Ethical Approval Number: IACUC-20191125-05, Guangzhou, China). Sixty CD-1 female mice (8 weeks old) were purchased from Beijing Charles River Laboratories (Vital River, China). Mice were acclimated for 10 days prior to experiments and were maintained under a 12-h light/dark cycle at a room temperature of 25 ± 2 °C and relative humidity of 55 ± 10%, with and fed on a standard chow. Each mouse was randomly allocated to control or treatment group via tail intravenous injection and sampled at different time points post injection (i.e., 30, 60, 120 and 240 min, Figure 6.1A). The dosage selection for each treatment group was selected according to human relevant exposure levels. Briefly, one dose of BPA at 50 μg/kg bw and two doses of *p*-cresol at 171 and 1,972 μg/kg bw were used in this study. Blood samples were collected from each mouse via cardiac puncture and plasma samples were obtained for BPA analysis. A portion of liver samples were collected and stored at -80 °C for further gene expression analysis. Each time point per treatment group was conducted in n = 5 replicates. To quantify the trace level of BPA concentration in the mouse blood, we used a sensitive multiple reaction monitoring method of LC-MS in this study (detailed in Appendix B).

6.3.6 Reverse transcription polymerase chain reaction (RT-PCR)

The gene expression of UDP-Glucuronosyltransferases 1A1 (*Ugt1a1*), UDP-Glucuronosyltransferases 2B1 (*Ugt2b1*) and Sulfotransferase (*Sult1a1*), which code the major enzymes of BPA biotransformation, were quantified using real-time PCR method (detailed in Appendix B).

6.3.7 Data Processing

ANOVA and Dunnett's post-hoc tests were performed using GraphPad Prism 8.4

(GraphPad Software Inc., San Diego, USA) to determine statistical significance. IC_{50} (half maximal inhibitory concentration) values with 95% confidence intervals (CI) were calculated using variable slope model (four parameters model) as *eq*.(1):

$$Y = 100/\left(1 + \frac{IC_{50}}{X}\right)^{HillSlope} \quad (1)$$

where X is dose and Y is normalized response, 100% down to 0%, decreasing as X increases; IC_{50} is at same concentration units as X; Hillslope is slope factor or Hill slope, unitless.

The K_m values and V_{max} values were obtained using Michaelis-Menten model. The parameters were calculated based on *eq*.(2):

$$v = V_{max}[S]/K_m + [S] \quad (2)$$

where V_{max} and K_m are Michaelis-Menten constants obtained from the *in vitro* experiment; V_{max} is the maximum enzyme velocity and K_m is the substrate concentration at which the reaction rate is half of V_{max}; [S] is the substrate concentration at the enzyme active site; v is enzyme velocity.

6.3.8 Molecular Docking

To compare the SULT binding affinity difference among exogenous xenobiotics, molecular docking was further adopted to illustrate their binding capacity. The binding energy (E) was calculated by the *eq*.(3):

$$E = -R \times T \times (\log K_{eq}) \quad (3)$$

where, *R, T* and K_{eq} stand for ideal gas constant, Kelvin temperature and equilibrium constant in the binding process; respectively. The structures of SULT1A1 was obtained from RCSB PDB database (https://www.rcsb.org/), while the information of its two allosteric sites (the catechin-binding site and the NSAID-binding site) was adopted from previous studies (Cook, Wang, Girvin, & Leyh, 2016; T. Wang, Cook, & Leyh, 2017). The binding affinity analysis was processed by Autodock Tools 1.5.6.(Morris et al., 2009) The 3D structures of the SULT1A1 protein and its

allosteric sites were presented by 3D Structure Viewer (https://servicesn.mbi.ucla.edu/viewer/), and further ligand interaction analysis was carried out with Ligand Plot (Laskowski & Swindells, 2011). The poses and vacuum electrostatic surfaces of the ligands or proteins were presented by Pymol (https://pymol.org/2/).

6.4 Results and discussion

6.4.1 p-Cresol enhanced BPA metabolism in in vivo animal model

To determine the *in vivo* impact of *p*-cresol on the biotransformation of BPA, we conducted the experiment on sixty 8-week-old female CD-1 mice via tail intravenous injection. Three treatment groups included 50 μg/kg bw BPA (BPA Control), 50 μg/kg bw BPA with 171 μg/kg bw *p*-cresol (BPA + Low *p*-cresol) and 50 μg/kg bw BPA with 1972 μg/kg bw *p*-cresol (BPA + High *p*-cresol). The plasma BPA was quantified and presented in Figure 6.1B. Overall, co-treatment with *p*-cresol induced a dose-dependent decrease in concentrations of BPA in plasma at all time points. At 30 min, the plasma BPA of control group was significantly higher ($p < 0.01$) than the other two groups, which was demonstrated to be 0.7, 0.4 and 0.2 ng/mL for BPA Control, BPA + Low *p*-cresol, and BPA + High *p*-cresol, respectively. At 60 and 120 min, the control was faintly higher than the *p*-cresol treatment group, though the latter reached equilibrium possibly due to the faster metabolism of BPA in the group. At 240 min, the BPA concentration of control group remained higher ($p < 0.05$) than the rest groups. The dose-dependent decrease of BPA concentrations implies that BPA biotransformation is significantly accelerated by the co-administration of *p*-cresol. With the presence of *p*-cresol, more proportion of injected BPA was metabolized and eliminated from mice, which might moderate the half-life of BPA.

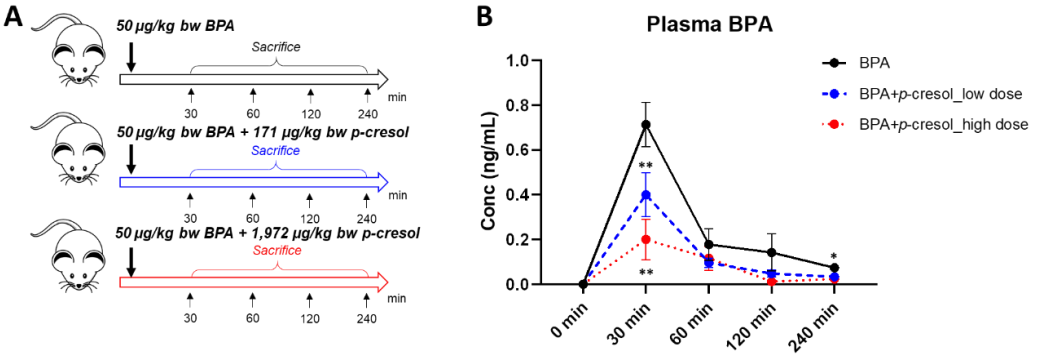


Figure 6.1 (A) Treatment regime of BPA and *p*-cresol in 8 weeks old female mice. (B) Plasma time-concentration profile for BPA from adult female CD-1 mice following tail intravenous injection of a single dose of 50 µg/kg bw BPA with or without 171 µg/kg or 1,972 µg/kg bw *p*-cresol. Data shown are the mean ± standard error of five replicates experiment. *$p < 0.05$, **$p < 0.01$ vs control group.

Previous research suggests that the enhanced biotransformation is quite unusual and unexpected, since most studies reported reduced metabolic capacity with co-administration owing to competition for host enzymes or transporters. It has been demonstrated that concurrent exposure to environmental phenolic chemicals elevated concentrations of BPA in mouse serum and reproductive tissues (Pollock et al. 2017). T. Andrew Clayton et al. also revealed that competitive o-sulfonation of *p*-cresol reduced systemic metabolic capacity of acetaminophen. Nonetheless, they only gave mice the acetaminophen dose and compared the urine metabolite profiles before and after the dose (Clayton et al. 2009). Therefore, such observed discrepancy may stem from the additional injections of *p*-cresol in our study. This unusual pattern drew our interest for further *in vitro* evaluation.

6.4.2 p-Cresol inhibited BPA biotransformation in in vitro models

To explore the mechanism of interaction between *p*-cresol and BPA, we further examined BPA biotransformation under similar treatment in two *in vitro* models of liver S9 fractions and HepG2 cell line. As illustrated in Figure 6.2 A and B, BPA biotransformation were inhibited in both liver S9 fraction and HepG2 models. In liver S9 models, with the presence of 10 µM *p*-cresol, conversion of BPA into BPAS

declined to ~50% compared to the control group. The highest dose, 175 μM *p*-cresol caused ~80% inhibitory effect on the sulfate conjugate. On the contrary, even though glucuronide conjugates were found to be the major metabolites in the incubation of liver S9 fractions due to the high abundance of uridine glucuronide transferase, the glucuronide conjugates were not observed to be affected much by *p*-cresol. 100 μM and 175 μM *p*-cresol resulted in slight decrease of BPAG formation. We observed similar trend in HepG2 model that both intracellular and extracellular BPAS were reduced in the presence of *p*-cresol. The significant decrease started at 5 μM *p*-cresol. Notably, under co-exposure to *p*-cresol, BPAS formation declined to ~70% and ~50% with 5 and 10 μM *p*-cresol, respectively. Moreover, 25 μM *p*-cresol resulted in more than half diminution of BPAS and 50 μM/100 μM *p*-cresol inhibited nearly all BPAS formation. Even at 5 μM (lower than the geometric mean concentration of human serum detection), co-exposure to *p*-cresol resulted in significant reduction of BPAS formation both in intracellular and extracellular content (De Smet et al., 1998). Nevertheless, unlike sulfate conjugates, we did not observe any alteration of BPA glucuronide in the presence of *p*-cresol across tested concentrations.

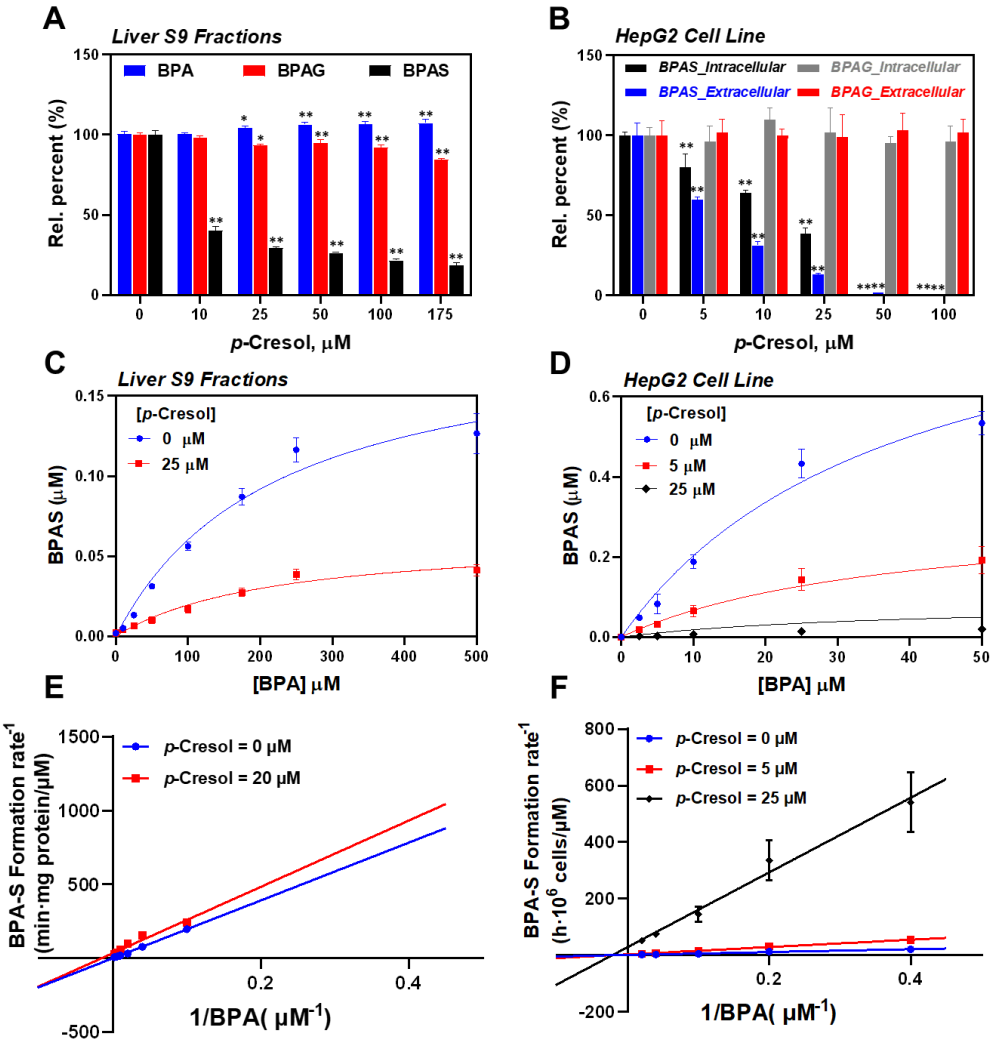


Figure 6.2 p-Cresol inhibited BPA biotransformation in in vitro models.
(A) Change of BPA metabolites of BPAS and BPAG under co-exposure to 0-175 μM *p*-cresol with liver S9 fractions; (B) Change of BPA metabolites of BPAS and BPAG under co-exposure to 0-175 μM *p*-cresol with HepG2 cells; (C) Michaelis-Menten plot and (E) Lineweaver-Burk plot for BPA inhibition of *p*-cresol (25 μM) with incubation of rat liver S9 fractions; (D) Michaelis-Menten plot and (F) Lineweaver-Burk plot for BPA inhibition of *p*-cresol (5 and 25 μM) with incubation of HepG2 cell line. Data shown are the mean ± standard deviation of triplicate experiment. **p* < 0.05, ***p* < 0.01 vs control group.

As *p*-cresol exhibited remarkable inhibitory effect on BPA metabolism *in vitro*, we further utilized the Michaelis-Menten model and the Lineweaver-Burk

approximation to determine the strength and type of inhibition of *p*-cresol against sulfate conjugation. In brief, 0-500 μM BPA with or without 25 μM *p*-cresol was incubated with liver S9 fractions whereas 0-50 μM BPA with or without 5/25 μM *p*-cresol was employed on HepG2 cells. In the incubation of liver S9 fractions without *p*-cresol, the V_{max} for BPA sulfation was determined to be 0.19 ± 0.02 μmol/(min·mg) and the K_m was 220.50 ± 40.04 μM. When co-incubated with 25 μM *p*-cresol, the V_{max}' for BPA sulfation decreased to 0.06 ± 0.008 μmol/(min·mg), while the K_m' stayed relatively unchanged at 218.00 ± 64.60 μM (Michaelis-Menten plot presented as Figure 6.2C). As shown in Figure 6.2D, the V_{max} for BPA sulfation with HepG2 cells was determined to be 0.98 ± 0.11 μmol/(h·10^6 cells), and the K_m was 38.7 ± 11.24 μM. When BPA was added with 5/25 μM *p*-cresol, the V_{max} for BPA sulfation reduced to 0.36 ± 0.09 μmol/(h·10^6 cells) and 0.04 ± 0.07 μmol/(h·10^6 cells), respectively; whereas the K_m was relatively stable at 42.31 ± 14.44 μM and 45.25 ± 14.34 μM, respectively. In contrast, the K_m values of *p*-cresol sulfation are remarkably lower, which were determined to be 8.38 ± 3.43 μM with liver S9 fractions and 3.48 ± 1.32 μM on HepG2 cells. The low K_m values of *p*-cresol indicated its much higher affinity to SULT, which could explain its potent inhibition on BPA sulfation. As illustrated in Lineweaver-Burk plots (Figure 6.2E & 6.2F), the linear relationship in the double-reciprocal plots confirmed a non-competitive inhibition profile in incubation of both liver S9 fractions and HepG2 cells. This mode of inhibition is consistent with the Michaelis-Menten plots (Figure 6.2C & 6.2D), as V_{max} was dramatically changed but K_m remained unchanged as non-competitive inhibitor was increasingly added.

6.4.3 Cresol isomers inhibited BPA sulfate formation with distinct specificities

Next, we examined the inhibition effect of cresol isomers on BPAS formation as previous studies on the crystal structures of SULT indicated that the position of the acceptor -OH group(s) amenable to sulfation is crucial for enzymatic activities (B. Wu et al., 2011). In this study, we performed dose-response inhibition examination of two additional cresol isomers *m*-cresol and *o*-cresol on BPA transformation, though *m*-cresol and *o*-cresol are mainly used in industrial production and were rarely detected in human samples (Yeung et al. 2002). Figure 6.3 shows dose-response

curves of *p, m, o*-cresol with incubation of both liver S9 fractions (A) and HepG2 cells (B and C) and Table B10 listed the IC_{50} values of *p,m,o*-cresol to BPA sulfate formation.

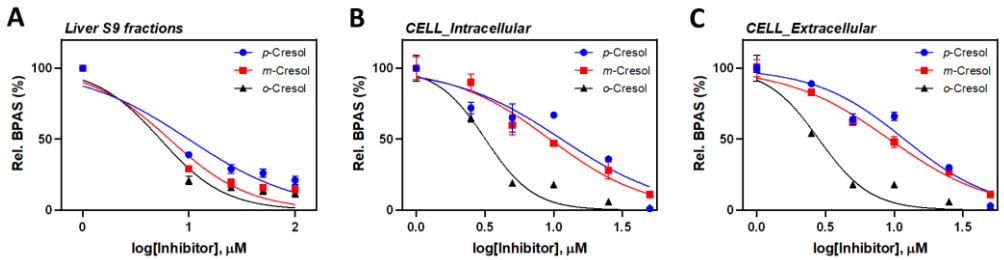


Figure 6.3 Dose-response curves for p,m,o-cresol to BPA (10 µM) sulfate formation in incubation of HepG2 cell line and rat liver S9 fractions.Dose-response curves for BPAS inhibition of incubation with (A) rat liver S9 fractions; (B) extracellular HepG2 cell medium; and (C) intracellular part of HepG2 cells. Data shown are the mean ± standard deviation of triplicate experiment.

BPA metabolites showed a consistent trend with liver S9 fractions in both intracellular and extracellular segments of cells. We observed that *o*-cresol was the strongest inhibitor of BPAS formation with incubation of both HepG2 cell and liver S9 fractions with IC_{50} values of 2.82 µM and 0.23 µM, respectively. It was followed by *m*-cresol that reduced half amount of BPAS formation when it reached 9.13 µM and 0.94 µM with HepG2 cell and liver S9 fractions, respectively. In contrast, the IC_{50} values of *p*-cresol were the highest, which were 2.81 µM with S9 fractions and 11.42 µM with HepG2 cells. The descending order of cresol isomers might be a result of the different hydroxy group position. For example, the orientation of the -OH group determines the affinity and K_m values of 3-OH form of resveratrol is much lower than the 4-OH form with incubation of liver cytosol and SULT 1E1 (Wu et al. 2011).

6.4.4 Binding affinity evaluation of chemicals and SULT1A1 protein by docking

Prior studies demonstrated that the predominant human cytosolic sulfotransferase in the liver is SULT1A1 and it harbors two noninteracting allosteric sites (presented in Figure 6.4A), each of which binds to a different molecular family: the catechins

(naturally occurring flavanols) and nonsteroidal anti-inflammatory drugs (NSAIDs) (Cook et al., 2016). From the above *in vitro* results, we observed a non-competitive inhibition pattern from the concurrent exposure to *p*-cresol and bisphenol A (BPA), which might involve reversible binding of allosteric sites. Hence, we further evaluated the binding affinity between the two allosteric sites of sulfotransferase (SULT) 1A1 protein and several chemicals, including the model compound BPA and cresol isomers using the molecular docking analysis. Figure 6.4B presents the ligand binding of BPA with the catechin-binding site (upper panel) and the NSAID-binding site (lower panel), while Figure 6.4C presents for the binding of *p*-cresol. The details of other compounds including *m*-cresol, *o*-cresol, acetaminophen and epigallocatechin gallate (EGCG) are presented in Figure B10. In general, all complexes have hydrogen bond formation, and the main forces include electrostatic repulsion and hydrophobic interactions (as illustrated by red scattered lines).

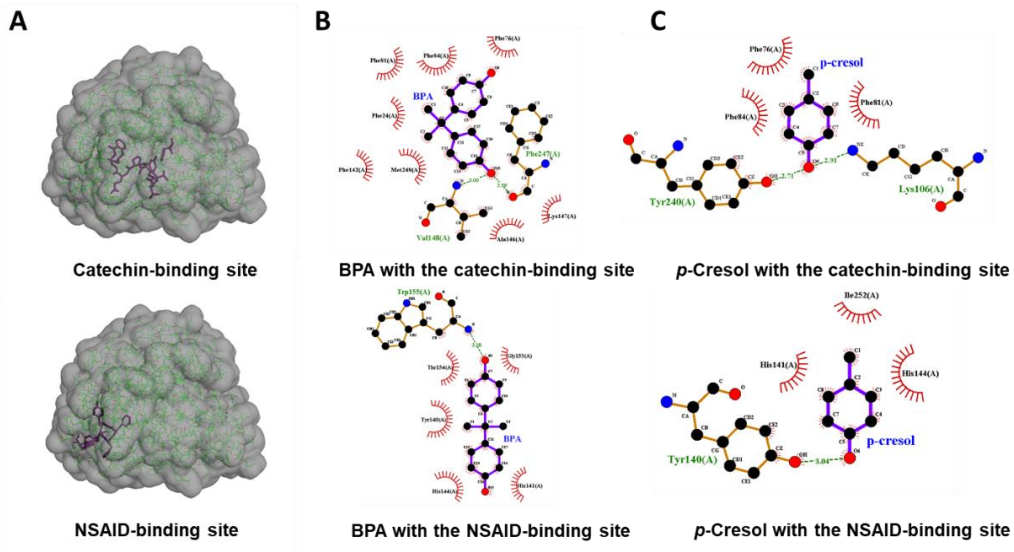


Figure 6.4 The docking analysis of BPA and p-cresol with two allosteric sites of SULT1A1. (A) The structures of SULT1A1 protein, where the upper is catechin-binding site and the bottom is the NSAID-binding site; (B) the binding analysis of BPA with the catechin-binding site (upper) and the NSAID-binding site (lower); (C) the binding analysis of *p*-cresol with the catechin-binding site (upper) and the NSAID-binding site (lower); The docking pockets of catechin-binding site and the NSAID-binding site are (0.375, 40, 40, 44, 22.191, 15.201, 16.633) and (0.375, 30,

20, 38, 5.894, 8.23, 8.532); respectively. The red scattered lines present hydrophobic interactions; the green dashed lines show hydrogen bond formation; the red feather-like shapes illustrate the adjacent residues of SULT1A1.

The summary of binding energy to two allosteric sites are listed in Table B11. Among all the tested compounds, BPA and cresol showed the lowest binding energy and suggested their high affinity to SULT1A1, which was even stronger than the two pharmaceutical drugs of acetaminophen and catechin EGCG. It can be seen that the overall binding energy to NSAID-binding site is lower, which is primarily due to one more developed hydrogen bond in the catechin-binding site. For example, the binding energy of BPA with the catechin-binding site was calculated to be -6.31 kcal/mol whereas it is -2.54 kcal/mol for NSAID-binding site. The huge difference implies that the catechin-binding site presents a greater binding affinity than the NSAID-binding site and BPA might prefer to bind to the catechin-binding site only. We identified a similar pattern for the positive catechin compound epigallocatechin gallate (EGCG), suggesting the validity of our modeling. For the cresol isomers, the binding energy to catechin-binding site is also higher than NSAID-binding site (i.e., -4.69 kcal/mol V.S. -2.31 kcal/mol). However, the difference was not as extreme as BPA or EGCG and it could be speculated that the cresol isomers might be able to bind two allosteric sites with increased substrate concentration. This result is consistent with one previous crystal structure study on SULT 1A1 that can accommodate two small molecules like *p*-nitrophenol (*p*-NP) at both active sites at the same time (Gamage et al. 2003). These results further validated our findings that *p*-cresol inhibited BPAS formation in a non-competitive pattern.

6.4.5 p-Cresol enhanced BPA metabolism and xenobiotic transformation gene expression in in vivo animal model

To investigate the potential mechanism to elucidate this difference between *in vivo* and *in vitro*, we hypothesized that metabolic enzymes might be altered upon exposure and examined the gene expression of a few key biotransformation enzymes (e.g., UDP-glucuronosyltransferases and sulfotransferase) in the livers of the mouse. We

found a dynamic change of those genes with various doses and exposure time (Figure 6.5).

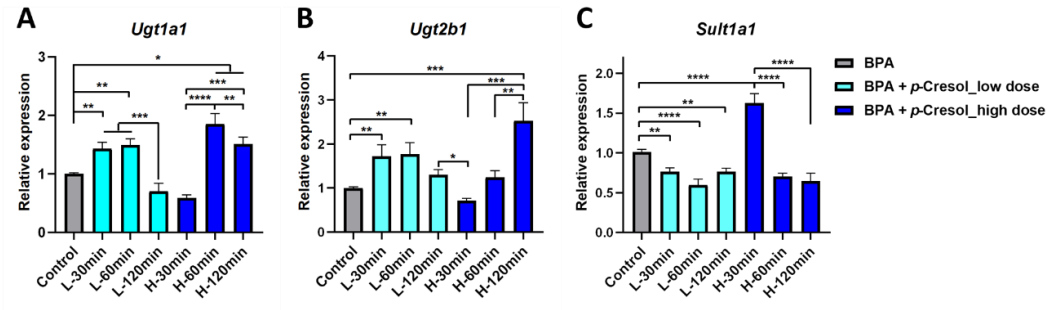


Figure 6.5 The mRNA expression levels of key biotransformation enzymes in mouse livers after single treatment of 50 μg/kg bw BPA (control group), 50 μg/kg bw BPA with 171 μg/kg bw *p*-cresol (L group) and 50 μg/kg bw BPA with 1972 μg/kg bw *p*-cresol (H group) at 30, 60 and 120 min. (A) UDP-glucuronosyltransferases 1A1 gene (*Ugt1a1*); (B) UDP-glucuronosyltransferases 2B1 gene (*Ugt2b1*); (C) sulfotransferase 1A1 gene (*Sult1a1*). Data are the mean ± standard error; n = 9. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

Specifically, in the lower *p*-cresol treatment, both UDP-glucuronosyltransferases genes (*Ugt1a1* and *Ugt2b1*) were upregulated to ~1.5 fold at 30 min and 60 min, before returning to normal at 120 min. In the higher *p*-cresol treatment group, there was a slight inhibition of *Ugt2b1* at 30 min but a significant upregulation at 60 min and 120 min. Similar pattern was observed on *Ugt1a1* at 60 min and 120 min. In addition, *Sult1a1* has been highly induced by the treatment at the early stage of 30 min, though there is no change for other treatments. Prior research reported that phase II metabolism by UDP-glucuronosyltransferases (UGTs) and sulfotransferases (SULTs) is the predominant metabolic pathway during the first-pass metabolism of BPA and *p*-cresol (Michalowicz 2014; Morinaga et al. 2004). Therefore, it is likely that co-exposure of *p*-cresol and BPA has upregulated those enzymes which promotes the biotransformation of BPA. Such change of gene expression was to our great surprise given the short exposure period. It is known that stable upregulation of xenobiotic transformation enzymes usually takes days or even a couple of weeks to achieve. For examples, phenobarbital induced the mRNA expression of the

cytochrome P-450 (CYP 450) enzyme CYP1A2 and CYP2B10 in multiple mice strains after receiving a once-daily injection for 3 days (Sakuma et al. 1999). However, the transient expression change of those genes can be crucial for the xenobiotics with short half-lives such as BPA. Prior research has reported the role of gut microbiome derived metabolites in the regulation of xenobiotic transformation enzymes. For instance, gene *sult3a1* (Sulfotransferase family 3A, member 1) and gene *Cyp2b10* (Cytochrome P450, family 2, subfamily b, polypeptide 10) in liver tissues were less expressed in the germ-free mouse compared with the normal animal (Björkholm et al. 2009).

6.4.6 Implications of interaction between gut microbiome derived metabolites and environmental chemicals

Although biotransformation is a crucial step to mediate the xenobiotics' toxicity and regulate their terminal half-lives, few toxicity studies have considered interactions of toxicants with gut microbial metabolites during biotransformation process. Most previous studies have been focusing on the interactions among environmental chemicals (Oh et al. 2018; Peng et al. 2019; Huang et al. 2020). In this study, we have further explored the biotransformation interaction between BPA and gut microbial metabolites. We demonstrated that *p*-cresol at 5 μM (lower than geometric median of human serum detections: 8.6 μM) could significantly inhibit BPA biotransformation *in vitro*, especially its sulfate conjugation. This result is consistent with a previous metabolomics study that identified sulfation as a prevalent host response to gut bacterial metabolites (Wikoff et al. 2009). Notably, the saturation of conventional metabolic pathways might elicit unexpected enhanced toxicity. For example, once an overdose of acetaminophen saturates the glucuronidation and sulfation pathways, the reactive intermediate N-acetyl-p-benzoquinone imine (NAPQI) is massively accumulated, which induces oxidative stress and ultimately leads to hepatocellular damage (Johnson et al. 2014). The mechanism underlying these susceptible sulfotransferase activities still requires to be further investigated by more studies. However, the interaction between gut microbiome metabolites and xenobiotics can be quite complex. On the one hand, multiple xenobiotics might compete for common biotransformation enzymes; on the other hand, the induction or inhibition of gene

expression resulted from the gut microbial metabolite exposure is another important factor to consider. Environmental chemicals at their relevant exposure doses are unlikely to change these enzyme activities; however, the gut microbial metabolites are capable of impacting on the host xenobiotic transformation gene expression due to its much higher concentration. Hence, simple enzymatic competition models should be studied in conjunction with other approaches to predict the interaction between gut microbiome derived metabolites and environmental toxicants. Future research efforts on risk assessment can further explore this complex interaction among the chemicals in the biotransformation.

However, it should be noted that there are several limitations of this study. One of them is the possible gender difference in response to *p*-cresol and BPA treatment due to the potentially different transcriptomic responses between genders. For instance, the gene transcription might be more sensitive in the female due to the weak estrogen receptor activation effect of BPA (Acconcia et al. 2015). Second, gut microbes can interact with biotransformation process via a number of distinct and complex mechanisms. BPA, for example, may be able to alter the structure of the gut microbiome, which in turn may affect cresol levels (Lai et al. 2016). Although we have gained a sufficient understanding of the biotransformation interaction between environmental phenols and gut microbial metabolites through a combination of *in vivo* and *in vitro* examination, further investigation into other aspects of the influence mechanism should be considered in future studies.

6.5 Short Summary

This study utilized bisphenol A (BPA) and *p*-cresol as model compounds to explore whether gut microbial metabolites could affect environmental phenol metabolism on both *in vitro* and *in vivo* models. The results indicated that biotransformation behaviors differed between the *in vitro* and *in vivo* results. *In vivo* mouse examination using *p*-cresol injection exhibited enhancing effect on BPA metabolism, which is rarely found in mixture studies. Serum level of BPA was significantly lower in the *p*-cresol group with a clear dose-response relationship. However, in both *in vitro* models of liver S9 fractions and HepG2 cell line, *p*-cresol is found as a strong inhibitor in a non-competitive pattern for BPA biotransformation. *In silico* docking approach was utilized to explore the non-competitive inhibition mechanism by estimating the binding affinity of *p*-cresol, BPA and two allosteric binding site of key enzyme SULT 1A1. Next, a close investigation revealed that the expression of biotransformation enzyme genes including *Ugt1a1*, *Ugt2b1, or Sult1a1* were dynamically induced after the *p*-cresol treatment. Overall, our results provided a novel insight into the biotransformation interaction between gut microbiome and environmental contaminants. The discrepancy between *in vitro* and *in vivo* models suggested that both enzyme competition and dynamic gene regulation should be considered when evaluating the interaction between gut microbiome and environmental xenobiotics.

Chapter 7. Conclusions and Recommendations

7.1 Conclusions

The first section of this thesis curated biotransformation information among multiple species and set up a comprehensive database, which harbors biotransformation data of common environmental pollutants, along with pharmaceuticals and human endogenous metabolites. To establish this data repository, we collected 7,800 biotransformation reactions of 1,599 substrates by text mining and database fusion. The collected species are not restricted to human, but also extended to birds, fish, rodents, and other mammals. This database contains structures of substrates and their corresponding products, types of reactions, the names or types of enzymes catalyzing biotransformation, types of biosystems, and references. Users are also allowed to search biotransformation results based on structural similarity. Overall, for the first time, environmental pollutants and multiple species are targeted as focus for biotransformation data curation. This integrative platform makes a great contribution to the scientific understanding of the biotransformation path in the whole ecosystem and serves as a guidance for chemical management.

Furthermore, a novel computational method with an automatically extracted reaction template was presented, which further ranks candidates by molecular similarity metric. This work demonstrates molecular similarity to be an effective metric. Extensive evaluation reveals that the current model achieves 45.25% recall and 26.37% precision at the top 3 candidates and 65.67% recall and 7.96 % precision at the top 20 candidates, respectively. On the whole, the results indicate that the proposed approach can provide a simple and automatic solution that does not restrict to any expert knowledge.

The third section of this study selected bisphenol A (BPA) as a model compound. Its biotransformation was investigated under single/co-exposure to other phenolic xenobiotics (Triclosan (TCS), Tetrabromobisphenol A (TBBPA), and Bisphenol S (BPS)) in liver microsome and cell models. The result shows that binary exposures exhibit significant inhibitory effects on BPA metabolism, especially the sulfate conjugation. In combination of analysis on inhibition models and enzyme activity,

the inhibition effect was suggested to be primarily driven by competition for metabolizing enzymes. A mixture with 22 phenolic chemicals was further examined to disrupt BPA at various human-relevant levels. Again, the result demonstrates significant inhibition on BPA metabolism, indicating the possible natural existence of our finding. Overall, our results show that biotransformation of phenolic xenobiotics can be significantly altered by co-exposure, which provides referential evidence on underestimated risks of simultaneous exposure to environmental toxicants.

Lastly, this study utilized bisphenol A (BPA) and *p*-cresol as model compounds to explore whether gut microbial metabolites could affect environmental phenol metabolism on both *in vitro* and *in vivo* models. The results indicated that biotransformation behaviors differed between the *in vitro* and *in vivo* results. *In vivo* mouse examination using *p*-cresol injection exhibited enhancing effect on BPA metabolism, which is rarely found in mixture studies. The serum level of BPA was significantly lower in the *p*-cresol group with a clear dose-response relationship. However, in both *in vitro* models of liver S9 fractions and HepG2 cell line, *p*-cresol is found as a strong inhibitor in a non-competitive pattern for BPA biotransformation. *In silico* docking approach was utilized to explore the non-competitive inhibition mechanism by estimating the binding affinity of *p*-cresol, BPA and two allosteric binding sites of key enzyme SULT 1A1. Next, a close investigation revealed that the expression of biotransformation enzyme genes including *Ugt1a1*, *Ugt2b1, or Sult1a1* was dynamically induced after the *p*-cresol treatment. Overall, our results provided a novel insight into the biotransformation interaction between gut microbiome and environmental contaminants. The discrepancy between *in vitro* and *in vivo* models suggested that both enzyme competition and dynamic gene regulation should be considered when evaluating the interaction between gut microbiome and environmental xenobiotics.

In sum, the major conclusions and contributions of this study were:

(i) Establishing a comprehensive biotransformation database, which harbors multi-species biotransformation data of common environmental pollutants,

along with pharmaceuticals and human endogenous metabolites.

(ii) Developing a novel computational method that robustly predicted biotransformation metabolite. The method started with an automatically extracted reaction template was presented, which further ranks candidates by molecular similarity metric.

(iii) Presenting a novel xenobiotic biotransformation prediction platform integrating multiple query function and prediction function.

(iv) Revealing that biotransformation of phenolic xenobiotics can be significantly altered by co-exposure, which provides referential evidence on underestimated risks of simultaneous exposure to environmental toxicants.

(v) Demonstrating that gut microbial metabolites could affect environmental phenol metabolism on both *in vitro* and *in vivo* models, but in distinct ways.

(vi) Providing a novel insight into the biotransformation interaction between gut microbiome and environmental contaminants.

7.2 Recommendations

Several knowledge gaps related to biotransformation of environmental chemicals are addressed by our findings and would inform further research:

First, a comprehensive biotransformation database, which harbors data of common environmental pollutants, along with pharmaceuticals and human endogenous metabolites, has been established. However, the coverage of species and compounds could be further expanded. Second, a novel computational method was developed for robust metabolite structure prediction. But the specific prediction for each species could be improved in the future. Subsequently, this work revealed that biotransformation of phenolic xenobiotics can be significantly altered by co-exposure, which provides referential evidence on underestimated risks of simultaneous exposure to environmental toxicants. Third, the complicated interplay between gut microbiome and biotransformation of environmental contaminants were revealed via *in vivo* and *in vitro* experiments.

However, there are some limitations as well. First, additional information such as gene expression, organ distribution, and gender variance might be added to the biotransformation database in the future. Second, by adding new approaches, the model's accuracy might be enhanced even more. The knowledge-based strategy, for example, might be coupled with a classification prediction model based on machine learning techniques. Another limitation is the relevance between *in vitro* and *in vivo* models. In addition, we have observed some over-saturation of dosed chemicals at high concentrations, especially in the microsome study. To generate slow and reliable release of hydrophobic chemicals, future research can explore passive dosing schemes on these chemicals. Lastly, this thesis demonstrated that biotransformation behaviors were a discrepancy between the *in vitro* and *in vivo* results when gut metabolites and environmental chemicals were co-exposed. Based on this intriguing discovery, more in-depth research should be carried out.

References

Accelrys Metabolite Database. (2011).

ADME database (Fujitsu Kyushu Systems Ltd, 2014).

ADMET Predictor, Metabolite Software Module. (2011).

Administration, F. a. D. (2016). Safety and Effectiveness of Consumer Antiseptics; Topical Antimicrobial Drug Products for Over-the-Counter Human Use. In (pp. 61106-61130). Washington, DC.

Afzelius, L., Arnby, C. H., Broo, A., Carlsson, L., Isaksson, C., Jurva, U., . . . Weidolf, L. (2007). State-of-the-art tools for computational site of metabolism predictions: Comparative analysis, mechanistical insights, and future applications. *Drug Metab. Rev., 39*, 61.

Ahmad, H. I., Afzal, G., Jamal, A., Kiran, S., Khan, M. A., Mehmood, K., . . . Almas, S. (2021). In Silico Structural, Functional, and Phylogenetic Analysis of Cytochrome (CYPD) Protein Family. *Biomed Res Int, 2021*, 5574789. doi:10.1155/2021/5574789

Allen, F., Pon, A., Wilson, M., Greiner, R., & Wishart, D. (2014). CFM-ID: a web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra. *Nucleic Acids Research, 42*(W1), W94-W99. doi:10.1093/nar/gku436

Ashrap, P., Zheng, G., Wan, Y., Li, T., Hu, W., Li, W., . . . Hu, J. (2017). Discovery of a widespread metabolic pathway within and among phenolic xenobiotics. *Proc Natl Acad Sci U S A, 114*(23), 6062-6067. doi:10.1073/pnas.1700558114

Athersuch, T. J., & Keun, H. C. (2015). Metabolic profiling in human exposome studies. *Mutagenesis, 30*(6), 755-762.

Battal, D., Cok, I., Unlusayin, I., & Tunctan, B. (2014). Development and validation of an LC-MS/MS method for simultaneous quantitative analysis of free and conjugated bisphenol A in human urine. *Biomed Chromatogr, 28*(5), 686-693. doi:10.1002/bmc.3090

Baxter, K. (2010). Stockley's Drug Interactions Pocket Companion 2011. [electronic resource]: London and Chicago : Pharmaceutical Press, 2010.

Benvenuto, J. A., Lu, K., Hall, S. W., Benjamin, R. S., & Loo, T. L. (1978). Disposition and metabolism of 1-(tetrahydro-2-furanyl)-5-fluorouracil (ftorafur) in humans. *Cancer Res, 38*(11 Pt 1), 3867-3870.

Beyer, B. A., Fang, M., Sadrian, B., Montenegro-Burke, J. R., Plaisted, W. C., Kok, B. P. C., . . . Lairson, L. L. (2017). Metabolomics-based discovery of a metabolite that enhances oligodendrocyte maturation. *Nature Chemical Biology, 14*(1), 22-28. doi:10.1038/nchembio.2517

Bickerton, G. R., Paolini, G. V., Besnard, J., Muresan, S., & Hopkins, A. L. (2012). Quantifying the chemical beauty of drugs. *Nature Chemistry, 4*(2), 90-98. doi:10.1038/nchem.1243

Björkholm, B., Bok, C. M., Lundin, A., Rafter, J., Hibberd, M. L., & Pettersson, S. (2009). Intestinal Microbiota Regulate Xenobiotic Metabolism in the Liver. *PLoS One, 4*(9), e6958. doi:10.1371/journal.pone.0006958

Borodina, Y., Sadym, A., Filimonov, D., Blinova, V., Dmitriev, A., & Poroikov, V. (2003). Predicting Biotransformation Potential from Molecular Structure. *Journal of Chemical Information and Computer Sciences, 43*(5), 1636-1646.

doi:10.1021/ci034078l
Bursztyka, J., Perdu, E., Pettersson, K., Pongratz, I., Fernández-Cabrera, M., Olea, N., . . . Cravedi, J. P. (2008). Biotransformation of genistein and bisphenol A in cell lines used for screening endocrine disruptors. *Toxicology in Vitro, 22*(6), 1595-1604. doi:10.1016/j.tiv.2008.06.013
Bushra, R., Aslam, N., & Khan, A. Y. (2011). Food-drug interactions. *Oman medical journal, 26*(2), 77-83. doi:10.5001/omj.2011.21
C Braga, R., & H Andrade, C. (2012). QSAR and QM/MM approaches applied to drug metabolism prediction. *Mini reviews in medicinal chemistry, 12*(6), 573-582.
Caldwell, J., Gardner, I., & Swales, N. (1995). An Introduction to Drug Disposition: The Basic Principles of Absorption, Distribution, Metabolism, and Excretion. *Toxicologic Pathology, 23*(2), 102-114. doi:10.1177/019262339502300202
Cao, D. (2012). ADMET evaluation in drug discovery. 11. PharmacoKinetics Knowledge Base (PKKB): a comprehensive database of pharmacokinetic and toxic properties for drugs. *J. Chem. Inf. Model., 52*.
Carlsson, L., Spjuth, O., Adams, S., Glen, R. C., & Boyer, S. (2010). Use of historic metabolic biotransformation data as a means of anticipating metabolic sites using MetaPrint2D and Bioclipse. *BMC Bioinformatics, 11*, 362.
Chen, D., Kannan, K., Tan, H., Zheng, Z., Feng, Y. L., Wu, Y., & Widelka, M. (2016). Bisphenol Analogues Other Than BPA: Environmental Occurrence, Human Exposure, and Toxicity-A Review. *Environ Sci Technol, 50*(11), 5438-5453. doi:10.1021/acs.est.5b05387
Clayton, T. A., Baker, D., Lindon, J. C., Everett, J. R., & Nicholson, J. K. (2009). Pharmacometabonomic identification of a significant host-microbiome metabolic interaction affecting human drug metabolism. *Proc Natl Acad Sci U S A, 106*(34), 14728. doi:10.1073/pnas.0904489106
Coley, C. W., Green, W. H., & Jensen, K. F. (2019). RDChiral: An RDKit Wrapper for Handling Stereochemistry in Retrosynthetic Template Extraction and Application. *Journal of chemical information and modeling, 59*(6), 2529-2537. doi:10.1021/acs.jcim.9b00286
Cook, I., Wang, T., Girvin, M., & Leyh, T. S. (2016). The structure of the catechin-binding site of human sulfotransferase 1A1. *Proceedings of the National Academy of Sciences, 113*(50), 14312. doi:10.1073/pnas.1613913113
Corrie, K., & Hardman, J. G. (2017). Mechanisms of drug interactions: pharmacodynamics and pharmacokinetics. *Anaesthesia & Intensive Care Medicine, 18*(7), 331-334. doi:https://doi.org/10.1016/j.mpaic.2017.04.006
Corsini, A., & Bortolini, M. (2013). Drug-induced liver injury: the role of drug metabolism and transport. *J Clin Pharmacol, 53*(5), 463-474. doi:10.1002/jcph.23
Cruciani, G., Carosati, E., De Boeck, B., Ethirajulu, K., Mackie, C., Howe, T., & Vianello, R. (2005). MetaSite: Understanding metabolism in human cytochromes from the perspective of the chemist. *J. Med. Chem., 48*, 6970.
Cuthbert, R., Green, R. E., Ranade, S., Saravanan, S., Pain, D. J., Prakash, V., & Cunningham, A. A. (2006). Rapid population declines of Egyptian vulture (Neophron percnopterus) and red-headed vulture (Sarcogyps calvus) in India. *Animal Conservation, 9*(3), 349-354.
Darvas, F., & Kaiser, K. L. E. (1987). MetabolExpert: An Expert System for

Predicting Metabolism of Substances.
de Bruyn Kops, C., Šícho, M., Mazzolari, A., & Kirchmair, J. (2020). Gloryx: Prediction of the metabolites resulting from phase 1 and phase 2 biotransformations of xenobiotics. *Chemical research in toxicology, 34*(2), 286-299.
De Smet, R., David, F., Sandra, P., Van Kaer, J., Lesaffer, G., Dhondt, A., . . . Vanholder, R. (1998). A sensitive HPLC method for the quantification of free and total p-cresol in patients with chronic renal failure. *Clinica Chimica Acta, 278*(1), 1-21. doi:https://doi.org/10.1016/S0009-8981(98)00124-7
Djoumbou-Feunang, Y., Fiamoncini, J., Gil-de-la-Fuente, A., Greiner, R., Manach, C., & Wishart, D. S. (2019). BioTransformer: a comprehensive computational tool for small molecule metabolism prediction and metabolite identification. *Journal of Cheminformatics, 11*(1), 2. doi:10.1186/s13321-018-0324-5
Dong, T., Zhang, Y., Jia, S., Shang, H., Fang, W., Chen, D., & Fang, M. (2019). Human Indoor Exposome of Chemicals in Dust and Risk Prioritization Using EPA's ToxCast Database. *Environmental Science & Technology, 53*(12), 7045-7054. doi:10.1021/acs.est.9b00280
Du, B., Haddad, S. P., Luek, A., Scott, W. C., Saari, G. N., Kristofco, L. A., . . . Chambliss, C. K. (2014). Bioaccumulation and trophic dilution of human pharmaceuticals across trophic positions of an effluent-dependent wadeable stream. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1656), 20140058.
Ekins, S. (2006). Algorithms for network analysis in systems-ADME/tox using the MetaCore and MetaDrug platforms. *Xenobiotica, 36*.
Ekman, D. R., Skelton, D. M., Davis, J. M., Villeneuve, D. L., Cavallin, J. E., Schroeder, A., . . . Collette, T. W. (2015). Metabolite Profiling of Fish Skin Mucus: A Novel Approach for Minimally-Invasive Environmental Exposure Monitoring and Surveillance. *Environmental Science & Technology, 49*(5), 3091-3100. doi:10.1021/es505054f
Escher, B. I., Stapleton, H. M., & Schymanski, E. L. (2020). Tracking complex mixtures of chemicals in our changing environment. *Science, 367*(6476), 388-392.
Fang, M., Webster, T. F., Ferguson, P. L., & Stapleton, H. M. (2015). Characterizing the peroxisome proliferator-activated receptor (PPARgamma) ligand binding potential of several major flame retardants, their metabolites, and chemical mixtures in house dust. *Environ Health Perspect, 123*(2), 166-172. doi:10.1289/ehp.1408522
Fanrong Zhao, Mo Chen, Fumei Gao, Huan Shen, & Jianying Hu. (2017). Organophosphorus Flame Retardants in Pregnant Women and Their Transfer to Chorionic Villi. *Environ. Sci. Technol., 51*(11), 6489-6497.
Fenner, K., Canonica, S., Wackett, L. P., & Elsner, M. (2013). Evaluating pesticide degradation in the environment: blind spots and emerging opportunities. *Science, 341*(6147), 752-758.
Frederiksen, H., Jørgensen, N., & Andersson, A. M. (2010). Correlations between phthalate metabolites in urine, serum, and seminal plasma from young Danish men determined by isotope dilution liquid chromatography tandem mass spectrometry. *Journal of analytical toxicology, 34*(7), 400-410.
Gamage, N., Barnett, A., Hempel, N., Duggleby, R. G., Windmill, K. F., Martin, J.

L., & McManus, M. E. (2006). Human sulfotransferases and their role in chemical metabolism. *Toxicological Sciences, 90*(1), 5-22.

Gamage, N. U., Duggleby, R. G., Barnett, A. C., Tresillian, M., Latham, C. F., Liyou, N. E., . . . Martin, J. L. (2003). Structure of a human carcinogen-converting enzyme, SULT1A1. Structural and kinetic implications of substrate inhibition. *J Biol Chem, 278*(9), 7655-7662. doi:10.1074/jbc.M207246200

Gao, J., Ellis, L. B., & Wackett, L. P. (2011). The University of Minnesota Pathway Prediction System: multi-level prediction and visualization. *Nucleic Acids Res., 39*.

Gao, J., Ellis, L. B. M., & Wackett, L. P. (2010). The University of Minnesota Biocatalysis/Biodegradation Database: improving public access. *Nucleic Acids Research, 38*(suppl_1), D488-D491. doi:10.1093/nar/gkp771

Gaulton, A. (2011). ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res., 40*.

Gerona, R. R., Woodruff, T. J., Dickenson, C. A., Pan, J., Schwartz, J. M., Sen, S., . . . Hunt, P. A. (2013). Bisphenol-A (BPA), BPA glucuronide, and BPA sulfate in midgestation umbilical cord serum in a northern and central California population. *Environ Sci Technol, 47*(21), 12477-12485. doi:10.1021/es402764d

Gleeson, M. P., Hersey, A., Montanari, D., & Overington, J. (2011). Probing the links between in vitro potency, ADMET and physicochemical parameters. *Nature reviews Drug discovery, 10*(3), 197-208.

GOSTAR drug database (GVK Biosciences Private Limited, 2013).

Hachad, H., Ragueneau-Majlessi, I., & Levy, R. H. (2010). A useful tool for drug interaction evaluation: The University of Washington Metabolism and Transport Drug Interaction Database. *Hum. Genom., 5*.

Halden, R. U. (2014). On the need and speed of regulating triclosan and triclocarban in the United States. *Environmental Science & Technology, 48*(7), 3603-3611. doi:10.1021/es500495p

Halden, R. U., & Paull, D. H. (2005). Co-Occurrence of Triclocarban and Triclosan in U.S. Water Resources. *Environmental Science & Technology, 39*(6), 1420-1426. doi:10.1021/es049071e

Hammer, R., Bozler, G., Jauch, R., Koss, F. W., & Hadamovsky, H. (1978). [Ambroxol, comparative studies of pharmacokinetics and biotransformation in rat, rabbit, dog and man (author's transl)]. *Arzneimittelforschung, 28*(5a), 899-903.

Hanet, N., Lancon, A., Delmas, D., Jannin, B., Chagnon, M. C., Cherkaoui-Malki, M., . . . Heydel, J. M. (2008). Effects of endocrine disruptors on genes associated with 17beta-estradiol metabolism and excretion. *Steroids, 73*(12), 1242-1251. doi:10.1016/j.steroids.2008.06.005

Harrad, S., Abdallah, M. A.-E., Rose, N. L., Turner, S. D., & Davidson, T. A. (2009). Current-Use Brominated Flame Retardants in Water, Sediment, and Fish from English Lakes. *Environ Sci Technol, 43*(24), 9077-9083. doi:10.1021/es902185u

Hayes, A. W., & Kruger, C. L. (2014). *Hayes' principles and methods of toxicology*: Crc Press.

Hennemann, M., Friedl, A., Lobell, M., Keldenich, J., Hillisch, A., Clark, T., & Göller, A. H. (2009). CypScore: Quantitative prediction of reactivity toward

cytochromes P450 based on semi-empirical molecular orbital theory. *ChemMedChem, 4*, 657.
Hewett, M. (2002). PharmGKB: the Pharmacogenetics Knowledge Base. *Nucleic Acids Res., 30*.
Hitchcock, S. A., & Pennington, L. D. (2006). Structure−Brain Exposure Relationships. *Journal of Medicinal Chemistry, 49*(26), 7559-7583. doi:10.1021/jm060642i
Ho, K. L., Yuen, K. K., Yau, M. S., Murphy, M. B., Wan, Y., Fong, B. M., . . . Lam, M. H. (2017). Glucuronide and sulfate conjugates of tetrabromobisphenol A (TBBPA): Chemical synthesis and correlation between their urinary levels and plasma TBBPA content in voluntary human donors. *Environ Int, 98*, 46-53. doi:10.1016/j.envint.2016.09.018
Hoffmann, M. F., Preissner, S. C., Nickel, J., Dunkel, M., Preissner, R., & Preissner, S. (2014). The Transformer database: biotransformation of xenobiotics. *Nucleic Acids Research, 42*(D1), D1113-D1117. doi:10.1093/nar/gkt1246
Högberg, J., Hanberg, A., Berglund, M., Skerfving, S., Remberger, M., Calafat, A. M., & Håkansson, H. (2007). Phthalate diesters and their metabolites in human breast milk, blood or serum, and urine as biomarkers of exposure in vulnerable populations. *Environmental health perspectives, 116*(3), 334-339.
Honkakoski, P., & Negishi, M. (1998). Regulation of cytochrome P450 genes by nuclear receptors. *Mol. Cell. Biol., 18*, 5652-5658.
Hooper, L. V., Wong, M. H., Thelin, A., Hansson, L., Falk, P. G., & Gordon, J. I. (2001). Molecular analysis of commensal host-microbial relationships in the intestine. *Science, 291*(5505), 881-884.
Huang, H., & Wu, Q. (2010). Cloning and comparative analyses of the zebrafish Ugt repertoire reveal its evolutionary diversity. *PLoS One, 5*(2), e9144.
Huang, Y. Q., Wong, C. K. C., Zheng, J. S., Bouwman, H., Barra, R., Wahlström, B., . . . Wong, M. H. (2012). Bisphenol A (BPA) in China: A review of sources, environmental levels, and potential human health impacts. *Environment International, 42*, 91-99. doi:https://doi.org/10.1016/j.envint.2011.04.010
Hutchinson, T. H., Madden, J. C., Naidoo, V., & Walker, C. H. (2014). Comparative metabolism as a key driver of wildlife species sensitivity to human and veterinary pharmaceuticals. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 369*(1656), 20130583. doi:10.1098/rstb.2013.0583
JChem, Metabolizer Software Module. (2011).
Jia, S., Xu, T., Huan, T., Chong, M., Liu, M., Fang, W., & Fang, M. (2019). Chemical Isotope Labeling Exposome (CIL-EXPOSOME): One High-Throughput Platform for Human Urinary Global Exposome Characterization. *Environ Sci Technol, 53*(9), 5445-5453. doi:10.1021/acs.est.9b00285
Jiang, C., Wang, X., Li, X., Inlora, J., Wang, T., Liu, Q., & Snyder, M. (2018). Dynamic Human Environmental Exposome Revealed by Longitudinal Personal Monitoring. *Cell, 175*(1), 277-291 e231. doi:10.1016/j.cell.2018.08.060
Johnson, B. P., Walisser, J. A., Liu, Y., Shen, A. L., McDearmon, E. L., Moran, S. M., . . . Bradfield, C. A. (2014). Hepatocyte circadian clock controls acetaminophen bioactivation through NADPH-cytochrome P450 oxidoreductase. *Proc Natl Acad Sci U S A, 111*(52), 18757-18762.

doi:10.1073/pnas.1421708111
Juurlink, D. N., Mamdani, M., Kopp, A., Laupacis, A., & Redelmeier, D. A. (2003). Drug-Drug Interactions Among Elderly Patients Hospitalized for Drug Toxicity. *JAMA, 289*(13), 1652-1658. doi:10.1001/jama.289.13.1652
Kammann, U., Askem, C., Dabrowska, H., Grung, M., Kirby, M. F., Koivisto, P., . . . Strand, J. (2013). Interlaboratory Proficiency Testing for Measurement of the Polycyclic Aromatic Hydrocarbon Metabolite 1-Hydroxypyrene in Fish Bile for Marine Environmental Monitoring. *Journal of AOAC INTERNATIONAL, 96*(3), 635-641. doi:10.5740/jaoacint.12-080
Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., & Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res., 40*, D109.
Karrer, C., Roiss, T., von Goetz, N., Gramec Skledar, D., Peterlin Masic, L., & Hungerbuhler, K. (2018). Physiologically Based Pharmacokinetic (PBPK) Modeling of the Bisphenols BPA, BPS, BPF, and BPAF with New Experimental Metabolic Parameters: Comparing the Pharmacokinetic Behavior of BPA with Its Substitutes. *Environ Health Perspect, 126*(7), 077002. doi:10.1289/EHP2739
Kawai, Y. K., Shinya, S., Ikenaka, Y., Saengtienchai, A., Kondo, T., Darwish, W. S., . . . Ishizuka, M. (2019). Characterization of function and genetic feature of UDP-glucuronosyltransferase in avian species. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology, 217*, 5-14. doi:https://doi.org/10.1016/j.cbpc.2018.11.001
Kidd, K. A., Blanchfield, P. J., Mills, K. H., Palace, V. P., Evans, R. E., Lazorchak, J. M., & Flick, R. W. (2007). Collapse of a fish population after exposure to a synthetic estrogen. *Proceedings of the National Academy of Sciences, 104*(21), 8897-8901.
Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., . . . Yu, B. (2019). PubChem 2019 update: improved access to chemical data. *Nucleic Acids Research, 47*(D1), D1102-D1109.
Kim, S., Thiessen, P. A., Cheng, T., Yu, B., & Bolton, E. E. (2018). An update on PUG-REST: RESTful interface for programmatic access to PubChem. *Nucleic Acids Research, 46*(W1), W563-W570. doi:10.1093/nar/gky294
Kim, S. I., Yang, Y. J., Hong, Y. P., Myung, S. C., & Kim, S. C. (2015). Distribution of serum bisphenol A diglycidyl ether and its metabolite in Korean adult men and its association with reproductive hormone levels. *Molecular & Cellular Toxicology, 11*(1), 71-78.
Kirchmair, J. (2013). FAst MEtabolizer (FAME): a rapid and accurate predictor of sites of metabolism in multiple species by endogenous enzymes. *J. Chem. Inf. Model., 53*.
Kirchmair, J., Goller, A. H., Lang, D., Kunze, J., Testa, B., Wilson, I. D., . . . Schneider, G. (2015). Predicting drug metabolism: experiment and/or computation? *Nat Rev Drug Discov, 14*(6), 387-404. doi:10.1038/nrd4581
Kirchmair, J., Markt, P., Distinto, S., Schuster, D., Spitzer, G. M., Liedl, K. R., . . . Wolber, G. (2008). The Protein Data Bank (PDB), its related services and software tools as key components for in silico guided drug discovery. *J. Med. Chem., 51*, 7021.
Kirchmair, J., Williamson, M. J., Tyzack, J. D., Tan, L., Bond, P. J., Bender, A., &

Glen, R. C. (2012). Computational prediction of metabolism: sites, products, SAR, P450 enzyme dynamics, and mechanisms. *Journal of chemical information and modeling, 52*(3), 617-648.

Klaassen, C. D., Casarett, L. J., & Doull, J. (2013). *Casarett and Doull's toxicology : the basic science of poisons*: New York : McGraw-Hill Education / Medical, c2013.

8th ed.

Klopman, G., Dimayuga, M., & Talafous, J. (1994). META. 1. A program for the evaluation of metabolic transformation of chemicals. *J. Chem. Inf. Model., 34*, 1320.

Koenig, S., Fernández, P., & Solé, M. (2012). Differences in cytochrome P450 enzyme activities between fish and crustacea: Relationship with the bioaccumulation patterns of polychlorobiphenyls (PCBs). *Aquatic Toxicology, 108*, 11-17. doi:https://doi.org/10.1016/j.aquatox.2011.10.016

Koppel, N., Maini Rekdal, V., & Balskus, E. P. (2017). Chemical transformation of xenobiotics by the human gut microbiota. *Science, 356*(6344). doi:10.1126/science.aag2770

Krejsa, C. M., Horvath, D., Rogalski, S. L., Penzotti, J. E., Mao, B., Barbosa, F., & Migeon, J. C. (2003). Predicting ADME properties and side effects: the BioPrint approach. *Current opinion in drug discovery and development, 6*(4), 470-480.

Laskowski, R. A., & Swindells, M. B. (2011). LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. In: ACS Publications.

Li, J., Schneebeli, S. T., Bylund, J., Farid, R., & Friesner, R. A. (2011). IDSite: An accurate approach to predict P450-mediated drug metabolism. *J. Chem. Theory Comput., 7*, 3829.

Li, X., Sun, H., Yao, Y., Zhao, Z., Qin, X., Duan, Y., & Wang, L. . (2018). Distribution of Phthalate Metabolites between Paired Maternal–Fetal Samples. *Environmental science & technology, 52*(11), 6626-6635.

Liao, C., & Kannan, K. (2012). Determination of free and conjugated forms of bisphenol A in human urine and serum by liquid chromatography–tandem mass spectrometry. *Environmental Science & Technology, 46*(9), 5003-5009.

Litsa, E. E., Das, P., & Kavraki, L. E. (2020). Prediction of drug metabolites using neural machine translation. *Chemical science, 11*(47), 12777-12788.

Liu, J., Li, J., Wu, Y., Zhao, Y., Luo, F., Li, S., . . . Martin, J. W. (2017). Bisphenol A Metabolites and Bisphenol S in Paired Maternal and Cord Serum. *Environ Sci Technol, 51*(4), 2456-2463. doi:10.1021/acs.est.6b05718

Liu, T., Lin, Y., Wen, X., Jorissen, R. N., & Gilson, M. K. (2007). BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res., 35*.

Malkoske, T., Tang, Y., Xu, W., Yu, S., & Wang, H. (2016). A review of the environmental distribution, fate, and control of tetrabromobisphenol A released from sources. *Sci Total Environ, 569-570*, 1608-1617. doi:10.1016/j.scitotenv.2016.06.062

Marchant, C. A., Briggs, K. A., & Long, A. (2008). In silico tools for sharing data and knowledge on toxicity and metabolism: DEREK for Windows, METEOR, and VITIC. *Toxicol. Mech. Methods, 18*, 177.

Martignoni, M. (2006). Species and strain differences in drug metabolism in liver and

intestine: University Library Groningen][Host].
McGill, M. R., & Jaeschke, H. (2013). Metabolism and disposition of acetaminophen: recent advances in relation to hepatotoxicity and diagnosis. *Pharm Res, 30*(9), 2174-2187. doi:10.1007/s11095-013-1007-6
Mekenyan, O. G., Dimitrov, S. D., Pavlov, T. S., & Veith, G. D. (2004). A systematic approach to simulating metabolism in computational toxicology. I. The TIMES heuristic modelling framework. *Curr. Pharm. Des., 10*, 1273.
Metabase (Thomson Reuters, 2013).
MEXAlert (CompuDrug International Inc., 2001).
Michalowicz, J. (2014). Bisphenol A--sources, toxicity and biotransformation. *Environ Toxicol Pharmacol, 37*(2), 738-758. doi:10.1016/j.etap.2014.02.003
Moreman, J., Takesono, A., Trznadel, M., Winter, M. J., Perry, A., Wood, M. E., . . . Tyler, C. R. (2018). Estrogenic Mechanisms and Cardiac Responses Following Early Life Exposure to Bisphenol A (BPA) and Its Metabolite 4-Methyl-2,4-bis(p-hydroxyphenyl)pent-1-ene (MBP) in Zebrafish. *Environ Sci Technol, 52*(11), 6656-6665. doi:10.1021/acs.est.8b01095
Morinaga, Y., Fuke, C., Arao, T., & Miyazaki, T. (2004). Quantitative analysis of cresol and its metabolites in biological materials and distribution in rats after oral administration. *Legal Medicine, 6*(1), 32-40. doi:https://doi.org/10.1016/j.legalmed.2003.08.005
Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry, 30*(16), 2785-2791.
NAGAYAMA, J., TSUJI, H., & TAKASUGA, T. (2000). Comparison between brominated flame retardants and dioxins or organochlorine compounds in blood levels of Japanese adults. *Organohalogen compounds, 48*, 27-30.
Nelson, D. R. (2003). Comparison of P450s from human and fugu: 420 million years of vertebrate P450 evolution. *Arch Biochem Biophys, 409*(1), 18-24. doi:10.1016/s0003-9861(02)00553-2
Oh, J., Choi, J. W., Ahn, Y. A., & Kim, S. (2018). Pharmacokinetics of bisphenol S in humans after single oral administration. *Environ Int, 112*, 127-133. doi:10.1016/j.envint.2017.11.020
Oh, W. S., Kim, D. N., Jung, J., Cho, K. H., & No, K. T. (2008). New combined model for the prediction of regioselectivity in cytochrome P450/3A4 mediated metabolism. *J. Chem. Inf. Model., 48*, 591.
Okuda, K., Fukuuchi, T., Takiguchi, M., & Yoshihara, S. (2011). Novel pathway of metabolic activation of bisphenol A-related compounds for estrogenic activity. *Drug Metab Dispos, 39*(9), 1696-1703. doi:10.1124/dmd.111.040121
Okuda, K., Takiguchi, M., & Yoshihara, S. i. (2010). In vivo estrogenic potential of 4-methyl-2,4-bis(4-hydroxyphenyl)pent-1-ene, an active metabolite of bisphenol A, in uterus of ovariectomized rat. *Toxicology Letters, 197*(1), 7-11. doi:https://doi.org/10.1016/j.toxlet.2010.04.017
Olah, M., Rad, R., Ostopovici, L., Bora, A., Hadaruga, N., Hadaruga, D., . . . Oprea, T. I. (2007). WOMBAT and WOMBAT-PK: Bioactivity Databases for Lead and Drug Discovery. *Chemical Biology*, 760-786. doi:https://doi.org/10.1002/9783527619375.ch13b

Pajouhesh, H., & Lenz, G. R. (2005). Medicinal chemical properties of successful central nervous system drugs. *NeuroRX, 2*(4), 541-553. doi:10.1602/neurorx.2.4.541

Patterson, D. G., Aylward, L. L., Covaci, A., & Franzblau, A. (2014). Chapter 5 - Biomonitoring for POPs. In G. O'Sullivan & C. Sandau (Eds.), *Environmental Forensics for Persistent Organic Pollutants* (pp. 163-197). Amsterdam: Elsevier.

Peach, M. L., Zakharov, A. V., Liu, R., Pugliese, A., Tawa, G., Wallqvist, A., & Nicklaus, M. C. (2012). Computational tools and resources for metabolism-related property predictions. 1. Overview of publicly available (free and commercial) databases and software. *Future Medicinal Chemistry, 4*(15), 1907-1932. doi:10.4155/fmc.12.150

Peng, B., Liu, M., Han, Y., Wanjaya, E. R., & Fang, M. (2019). Competitive Biotransformation Among Phenolic Xenobiotic Mixtures: Underestimated Risks for Toxicity Assessment. *Environmental Science & Technology, 53*(20), 12081-12090. doi:10.1021/acs.est.9b04968

Peng, B.; Zhao, H.; Keerthisinghe, T. P.; Yu, Y.; Chen, D.; Huang, Y.; Fang, M., Gut microbial metabolite p-cresol alters biotransformation of bisphenol A: Enzyme competition or gene induction? Journal of Hazardous Materials 2022, 426, 128093.

Peng Li, Jun Jin, Ying Wang, Jicheng Hu, Meng Xu, Yiming Sun, & YulongMa. (2017). Concentrations of organophosphorus, polybromobenzene, and polybrominated diphenyl ether flame retardants in human serum, and relationships between concentrations and donor ages. *Chemosphere, 171*, 654-660.

Peng Li, Qiu-Xu Li, Yu-long Ma, Jun Jin, Ying Wang, & Yang Tian. (2015). Determination of Organophosphate Esters in Human Serum Using Gel Permeation Chromatograph and Solid Phase Extraction Coupled with Gas Chromatography-Mass Spectrometry. *Chinese Journal of Analytical Chemistry, 43*(7), 1033-1039.

Percepta Platform (ACD/Labs Inc., 2014).

Pfeifer, S., Gildemeister, B., & Pöhlmann, H. (1983). *Biotransformation von Arzneimitteln* (Vol. 5): Verlag Volk und Gesundheit, VEB.

Pollock, T., Mantella, L., Reali, V., & deCatanzaro, D. (2017). Influence of Tetrabromobisphenol A, with or without Concurrent Triclosan, upon Bisphenol A and Estradiol Concentrations in Mice. *Environ Health Perspect, 125*(8), 087014. doi:10.1289/EHP1329

Prakash, C., Kamel, A., Gummerus, J., & Wilner, K. (1997). Metabolism and excretion of a new antipsychotic drug, ziprasidone, in humans. *Drug Metab Dispos, 25*(7), 863-872.

Preissner, S. (2010). SuperCYP: a comprehensive database on cytochrome P450 enzymes including a tool for analysis of CYP-drug interactions. *Nucleic Acids Res., 38*.

Prevention, C. f. D. C. a. (2018). Fourth National Report on Human Exposure to Environmental Chemicals. U.S.

Prins, G. S., Hu, W. Y., Xie, L., Shi, G. B., Hu, D. P., Birch, L., & Bosland, M. C. (2018). Evaluation of Bisphenol A (BPA) Exposures on Prostate Stem Cell Homeostasis and Prostate Cancer Risk in the NCTR-Sprague-Dawley Rat: An

NIEHS/FDA CLARITY-BPA Consortium Study. *Environ Health Perspect, 126*(11), 117001. doi:10.1289/EHP3953

Provencher, G., Berube, R., Dumas, P., Bienvenu, J. F., Gaudreau, E., Belanger, P., & Ayotte, P. (2014). Determination of bisphenol A, triclosan and their metabolites in human urine using isotope-dilution liquid chromatography-tandem mass spectrometry. *J Chromatogr A, 1348*, 97-104. doi:10.1016/j.chroma.2014.04.072

QikProp (Schrödinger LLC, 2014).

Rappaport, S. M., Barupal, D. K., Wishart, D., Vineis, P., & Scalbert, A. (2014). The blood exposome and its role in discovering causes of disease. *Environ Health Perspect, 122*(8), 769-774. doi:10.1289/ehp.1308015

Ridder, L., & Wagener, M. (2008). SyGMa: Combining Expert Knowledge and Empirical Scoring in the Prediction of Metabolites. *ChemMedChem, 3*(5), 821-832. doi:10.1002/cmdc.200700312

Ritchie, T. J., Macdonald, S. J. F., Young, R. J., & Pickett, S. D. (2011). The impact of aromatic ring count on compound developability: further insights by examining carbo- and hetero-aromatic and -aliphatic ring types. *Drug Discovery Today, 16*(3), 164-171. doi:https://doi.org/10.1016/j.drudis.2010.11.014

Rochester, J. R. (2013). Bisphenol A and human health: a review of the literature. *Reprod Toxicol, 42*, 132-155. doi:10.1016/j.reprotox.2013.08.008

Rochester, J. R., & Bolden, A. L. (2015). Bisphenol S and F: A Systematic Review and Comparison of the Hormonal Activity of Bisphenol A Substitutes. *Environ Health Perspect, 123*(7), 643-650. doi:10.1289/ehp.1408989

Rodricks, J. V. (2007). Calculated risks : the toxicity and human health risks of chemicals in our environment: Cambridge ; New York : Cambridge University Press, 2007.

2nd ed., with major revisions and additions.

Rudik, A. V., Bezhentsev, V. M., Dmitriev, A. V., Druzhilovskiy, D. S., Lagunin, A. A., Filimonov, D. A., & Poroikov, V. V. (2017). MetaTox: web application for predicting structure and toxicity of xenobiotics' metabolites. *Journal of chemical information and modeling, 57*(4), 638-642.

Rydberg, P., Gloriam, D. E., Zaretzki, J., Breneman, C., & Olsen, L. (2010). SMARTCyp: A 2D method for prediction of cytochrome P450-mediated drug metabolism. *ACS Med. Chem. Lett., 1*, 96.

Ryu, J. Y., Kim, H. U., & Lee, S. Y. (2018). Deep learning improves prediction of drug–drug and drug–food interactions. *Proceedings of the National Academy of Sciences, 115*, E4304-E4311.

Sakuma, T., Ohtake, M., Katsurayama, Y., Jarukamjorn, K., & Nemoto, N. (1999). Induction of CYP1A2 by Phenobarbital in the Livers of Aryl Hydrocarbon-Responsive and -Nonresponsive Mice. *Drug Metabolism and Disposition, 27*(3), 379. Retrieved from http://dmd.aspetjournals.org/content/27/3/379.abstract

Schauer, U. M. D., Völkel, W., & Dekant, W. (2006). Toxicokinetics of Tetrabromobisphenol A in Humans and Rats after Oral Administration. *Toxicological Sciences, 91*(1), 49-58. doi:10.1093/toxsci/kfj132

Schmidt, J., Kotnik, P., Trontelj, J., Knez, Ž., & Mašič, L. P. (2013). Bioactivation of bisphenol A and its analogs (BPF, BPAF, BPZ and DMBPA) in human

liver microsomes. *Toxicology in Vitro, 27*(4), 1267-1276. doi:https://doi.org/10.1016/j.tiv.2013.02.016

Schmidt, S. N., Holmstrup, M., Smith, K. E. C., & Mayer, P. (2013). Passive Dosing of Polycyclic Aromatic Hydrocarbon (PAH) Mixtures to Terrestrial Springtails: Linking Mixture Toxicity to Chemical Activities, Equilibrium Lipid Concentrations, and Toxic Units. *Environmental Science & Technology, 47*(13), 7020-7027. doi:10.1021/es3047813

Schönfelder, G., Wittfoht, W., Hopp, H., Talsness, C. E., Paul, M., & Chahoud, I. (2002). Parent bisphenol A accumulation in the human maternal-fetal-placental unit. *Environmental Health Perspectives, 110*(11), A703-A707. doi:10.1289/ehp.110-1241091

Schwaller, P., Hoover, B., Reymond, J.-L., Strobelt, H., & Laino, T. (2021). Extraction of organic chemistry grammar from unsupervised learning of chemical reactions. *Science Advances, 7*(15), eabe4166. doi:10.1126/sciadv.abe4166

Schwaller, P., Laino, T., Gaudin, T., Bolgar, P., Hunter, C. A., Bekas, C., & Lee, A. A. (2019). Molecular Transformer: A Model for Uncertainty-Calibrated Chemical Reaction Prediction. *ACS Central Science, 5*(9), 1572-1583. doi:10.1021/acscentsci.9b00576

Scientific Opinion on Tetrabromobisphenol A (TBBPA) and its derivatives in food. (2011). *EFSA Journal, 9*(12). doi:10.2903/j.efsa.2011.2477

Segler, M. H. S., Preuss, M., & Waller, M. P. (2018). Planning chemical syntheses with deep neural networks and symbolic AI. *Nature, 555*(7698), 604-610.

Shrestha, B., Reed, J. M., Starks, P. T., Kaufman, G. E., Goldstone, J. V., Roelke, M. E., . . . Court, M. H. (2011). Evolution of a major drug metabolizing enzyme defect in the domestic cat and other felidae: phylogenetic timing and the role of hypercarnivory. *PLoS One, 6*(3), e18046.

Smith, C. A. (2005). Metlin: a metabolite mass spectral database. *Ther. Drug Monit., 27*.

Smith, K. E. C., Rein, A., Trapp, S., Mayer, P., & Karlson, U. G. (2012). Dynamic Passive Dosing for Studying the Biotransformation of Hydrophobic Organic Chemicals: Microbial Degradation as an Example. *Environmental Science & Technology, 46*(9), 4852-4860. doi:10.1021/es204050u

Song, Y., Xie, P., & Cai, Z. (2018). Metabolism of bisphenol S in mice after oral administration. *Rapid Communications in Mass Spectrometry, 32*(6), 495-502.

Spanogiannopoulos, P., Bess, E. N., Carmody, R. N., & Turnbaugh, P. J. (2016). The microbial pharmacists within us: a metagenomic view of xenobiotic metabolism. *Nature Reviews Microbiology, 14*(5), 273-287. doi:10.1038/nrmicro.2016.17

Stahl, T., Mattern, D., & Brunn, H. (2011). Toxicology of perfluorinated compounds. *Environmental Sciences Europe, 23*(1), 38. doi:10.1186/2190-4715-23-38

StarDrop. (2011).

Sykes, M. J., McKinnon, R. A., & Miners, J. O. (2008). Prediction of metabolism by cytochrome P450 2C9: Alignment and docking studies of a validated database of substrates. *J. Med. Chem., 51*, 780.

Thomsen, C., Lundanes, E., & Becher, G. (2002). Brominated flame retardants in archived serum samples from Norway: a study on temporal trends and the role

of age. . *Environmental science & technology, 36*(7), 1414-1418.
Tian, S., Djoumbou-Feunang, Y., Greiner, R., & Wishart, D. S. (2018). CypReact: A Software Tool for in Silico Reactant Prediction for Human Cytochrome P450 Enzymes. *Journal of chemical information and modeling, 58*(6), 1282-1291. doi:10.1021/acs.jcim.8b00035
Tran, V. N., Viktorová, J., & Ruml, T. (2020). Mycotoxins: Biotransformation and bioavailability assessment using caco-2 cell monolayer. *Toxins, 12*(10), 628.
Uno, T., Ishizuka, M., & Itakura, T. (2012). Cytochrome P450 (CYP) in fish. *Environmental Toxicology and Pharmacology, 34*(1), 1-13. doi:https://doi.org/10.1016/j.etap.2012.02.004
Vandenberg, L. N., Chahoud, I., Heindel, J. J., Padmanabhan, V., Paumgartten, F. J. R., & Schoenfelder, G. (2010). Urinary, circulating, and tissue biomonitoring studies indicate widespread exposure to bisphenol A. *Environmental Health Perspectives, 118*(8), 1055-1070. doi:10.1289/ehp.0901716
Vrbanac, J., & Slauter, R. (2017). Chapter 3 - ADME in Drug Discovery. In A. S. Faqi (Ed.), *A Comprehensive Guide to Toxicology in Nonclinical Drug Development (Second Edition)* (pp. 39-67). Boston: Academic Press.
Walker, C. H., & Ronis, M. J. J. (1989). The monooxygenases of birds, reptiles and amphibians. *Xenobiotica, 19*(10), 1111-1121.
Wang, L., Xue, J., & Kannan, K. (2015). Widespread occurrence and accumulation of bisphenol A diglycidyl ether (BADGE), bisphenol F diglycidyl ether (BFDGE) and their derivatives in human blood and adipose fat. *Environmental science & technology, 49*(5), 3150-3157.
Wang, T., Cook, I., & Leyh, T. S. (2017). The NSAID allosteric site of human cytosolic sulfotransferases. *Journal of Biological Chemistry, 292*(49), 20305-20312.
Watanabe, K. P., Kawai, Y. K., Ikenaka, Y., Kawata, M., Ikushiro, S.-I., Sakaki, T., & Ishizuka, M. (2013). Avian cytochrome P450 (CYP) 1-3 family genes: isoforms, evolutionary relationships, and mRNA expression in chicken liver. *PLoS One, 8*(9), e75689-e75689. doi:10.1371/journal.pone.0075689
Wikoff, W. R., Anfora, A. T., Liu, J., Schultz, P. G., Lesley, S. A., Peters, E. C., & Siuzdak, G. (2009). Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proceedings of the National Academy of Sciences, 106*(10), 3698. doi:10.1073/pnas.0812874106
Wishart, D. S. (2006). DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res., 34*.
Wishart, D. S. (2007). HMDB: the human metabolome database. *Nucleic Acids Res., 35*.
Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., . . . Wilson, M. (2018). DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Research, 46*(D1), D1074-D1082. doi:10.1093/nar/gkx1037
Wu, B., Basu, S., Meng, S., Wang, X., & Hu, M. (2011). Regioselective sulfation and glucuronidation of phenolics: insights into the structural basis. *Current drug metabolism, 12*(9), 900-916. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/21933112
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3426368/
Wu, J., Yue, H., & Cai, Z. (2009). Investigation on metabolism and pharmacokinetics

of triclosan in rat plasma by using UPLC-triple quadrupole MS. *Se pu= Chinese journal of chromatography*, *27*(5), 724-730.

Wu, Z., He, C., Han, W., Song, J., Li, H., Zhang, Y., . . . Wu, W. (2020). Exposure pathways, levels and toxicity of polybrominated diphenyl ethers in humans: A review. *Environmental Research, 187*, 109531. doi:https://doi.org/10.1016/j.envres.2020.109531

Xiao, Z., Feng, J., Shi, Z., Li, J., Zhao, Y., & Wu, Y. (2011). Determination of three brominated flame retardants in human serum using solid-phase extraction coupled with ultra-performance liquid chromatography-tandem mass spectrometry and gas chromatography-mass spectrometry. *Chinese journal of chromatography*, *29*(12), 1165-1172.

Xu, C., Li, C. Y., & Kong, A. N. (2005). Induction of phase I, II and III drug metabolism/transport by xenobiotics. *Arch Pharm Res, 28*(3), 249-268. doi:10.1007/bf02977789

Yang, J. J., Han, Y., Mah, C. H., Wanjaya, E., Peng, B., Xu, T. F., . . . Fang, M. L. (2020). Streamlined MRM method transfer between instruments assisted with HRMS matching and retention-time prediction. *Analytica Chimica Acta, 1100*, 88-96. doi:https://doi.org/10.1016/j.aca.2019.12.002

Yeung, S., Lan, W., Huang, C., Lin, C., Chan, C., Chang, M., & Jeng, J. (2002). Scavenging property of three cresol isomers against H2O2, hypochlorite, superoxide and hydroxyl radicals. *Food and Chemical Toxicology, 40*(10), 1403-1413.

Yoon, E., Babar, A., Choudhary, M., Kutner, M., & Pyrsopoulos, N. (2016). Acetaminophen-Induced Hepatotoxicity: a Comprehensive Update. *Journal of clinical and translational hepatology, 4*(2), 131-142. doi:10.14218/JCTH.2015.00052

Yoshihara, S. i., Makishima, M., Suzuki, N., & Ohta, S. (2001). Metabolic Activation of Bisphenol A by Rat Liver S9 Fraction. *Toxicological Sciences, 62*(2), 221-227. doi:10.1093/toxsci/62.2.221

Yoshihara, S. i., Mizutare, T., Makishima, M., Suzuki, N., Fujimoto, N., Igarashi, K., & Ohta, S. (2004). Potent Estrogenic Metabolites of Bisphenol A and Bisphenol B Formed by Rat Liver S9 Fraction: Their Structures and Estrogenic Potency. *Toxicological Sciences, 78*(1), 50-59. doi:10.1093/toxsci/kfh047

Yu, J., Petrie, I. D., Levy, R. H., & Ragueneau-Majlessi, I. (2019). Mechanisms and Clinical Significance of Pharmacokinetic-Based Drug-Drug Interactions with Drugs Approved by the U.S. Food and Drug Administration in 2017. *Drug Metabolism and Disposition, 47*(2), 135. doi:10.1124/dmd.118.084905

Yueh, M. F., & Tukey, R. H. (2016). Triclosan: A Widespread Environmental Toxicant with Many Biological Effects. *Annu Rev Pharmacol Toxicol, 56*, 251-272. doi:10.1146/annurev-pharmtox-010715-103417

Zalko, D., Prouillac, C., Riu, A., Perdu, E., Dolo, L., Jouanin, I., . . . Cravedi, J.-P. (2006). Biotransformation of the flame retardant tetrabromo-bisphenol A by human and rat sub-cellular liver fractions. *Chemosphere, 64*(2), 318-327. doi:https://doi.org/10.1016/j.chemosphere.2005.12.053

Zaretzki, J. (2013). RS-WebPredictor: a server for predicting CYP-mediated sites of metabolism on drug-like molecules. *Bioinformatics, 29*.

Zaretzki, J., Bergeron, C., Rydberg, P., Huang, T. W., Bennett, K. P., & Breneman,

C. M. (2011). RS-Predictor: A new tool for predicting sites of cytochrome P450-mediated metabolism applied to CYP 3A4. *J. Chem. Inf. Model., 51*, 1667.

Zaretzki, J., Matlock, M., & Swamidass, S. J. (2013). XenoSite: Accurately Predicting CYP-Mediated Sites of Metabolism with Neural Networks. *Journal of chemical information and modeling, 53*(12), 3373-3383. doi:10.1021/ci400518g

Zhang, Y., Keerthisinghe, T. P., Han, Y., Liu, M., Wanjaya, E. R., & Fang, M. (2018). "Cocktail" of Xenobiotics at Human Relevant Levels Reshapes the Gut Bacterial Metabolome in a Species-Specific Manner. *Environ Sci Technol, 52*(19), 11402-11410. doi:10.1021/acs.est.8b02629

Zhao, F., Li, L., Chen, Y., Huang, Y., Keerthisinghe Tharushi, P., Chow, A., . . . Fang, M. Risk-Based Chemical Ranking and Generating a Prioritized Human Exposome Database. *Environmental Health Perspectives, 129*(4), 047014. doi:10.1289/EHP7722

Zhou, X., Kramer, J. P., Calafat, A. M., & Ye, X. (2014). Automated on-line column-switching high performance liquid chromatography isotope dilution tandem mass spectrometry method for the quantification of bisphenol A, bisphenol F, bisphenol S, and 11 other phenols in urine. *J Chromatogr B Analyt Technol Biomed Life Sci, 944*, 152-156. doi:10.1016/j.jchromb.2013.11.009

Zimmermann, M., Zimmermann-Kogadeeva, M., Wegmann, R., & Goodman, A. L. (2019). Mapping human microbiome drug metabolism by gut bacteria and their genes. *Nature*. doi:10.1038/s41586-019-1291-3

Zotter, A., Bäuerle, F., Dey, D., Kiss, V., & Schreiber, G. (2017). Quantifying enzyme activity in living cells. *Journal of Biological Chemistry, 292*(38), 15838-15848.

Appendix A

Table A1. The compound list of biotransformation database

Category	Source substrate	Substrate smiles
ENV	N-nitrosodimethylamine	CN(C)N=O
ENV	n-nitrosodibutylamine	CCCCN(CCCC)N=O
ENV	Meiqx	Cc1cnc2ccc3c([nH]c(=N)n3C)c2n1
ENV	Homo-Tienilic Alcohol	O=C(c1cccs1)c1ccc(OCCCO)c(Cl)c1Cl
ENV	furfuryl alcohol	OCc1ccco1
ENV	Dulcin	CCOc1ccc(NC(=N)O)cc1
ENV	Acrylamide	C=CC(N)=O
ENV	4-Propylanisole	CCCc1ccc(OC)cc1
ENV	1-Nitrosopyrrolidine	O=NN1CCCC1
ENV	2,3-dichloro-dibenzo-p-dioxin	Clc1cc2c(cc1Cl)Oc1ccccc1O2
ENV	2,3,7-TRICHLORODIBENZO-P-DIOXIN (2,3,7-triCDD)	Clc1ccc2c(c1)Oc1cc(Cl)c(Cl)cc1O2
ENV	2,3,7,8-tetrachlorodibenzo-p-dioxin	Clc1cc2c(cc1Cl)Oc1cc(Cl)c(Cl)cc1O2
ENV	2,3,3',4',5-Pentachlorobiphenyl	Clc1cc(Cl)c(Cl)c(-c2ccc(Cl)c(Cl)c2)c1
ENV	1-monochloro-dibenzo-p-dioxin	Clc1cccc2c1Oc1ccccc1O2
ENV	1,3,6,8-Tetrachlorodibenzo-P-dioxin	Clc1cc(Cl)c2c(c1)Oc1c(Cl)cc(Cl)cc1O2
ENV	1,2,7,8-TETRACHLORODIBENZO-P-DIOXIN	Clc1cc2c(cc1Cl)Oc1c(ccc(Cl)c1Cl)O2
Drug	sdz_imm_125	C/C=C/C[C@@H](C)[C@@H](O)[C@H]1C(O)=N[C@@H](CC)C(=O)N(C)CC(=O)N(C)[C@@H](CC(C)C)C(O)=N[C@@H](C(C)C)N(C)C(=O)[C@@H](CC(C)C)C(=O)N(C)[C@@H](C(C)C)C(=O)N(C)[C@@H](CC(C)C)C(O)=N[C@@H](C)C(O)=N[C@H](COCCO)C(=O)N(C)[C@@H](CC(C)C)C(=O)N(C)[C@@H](CC(C)C)C(=O)N(C)[C@@H](C(C)C)C(=O)N1C
Drug	Psoralen	O=c1ccc2cc3ccoc3cc2o1
Drug	Zotepine	CN(C)CCOC1=Cc2ccccc2Sc2ccc(Cl)cc21
Drug	Zoplicone	CN1CCN(C(=O)O[C@@H]2c3nccnc3C(=O)N2c2ccc(Cl)cn2)CC1
Drug	Zonisamide	NS(=O)(=O)Cc1noc2ccccc12
Drug	Zomepirac	Cc1cc(CC(=O)O)n(C)c1C(=O)c1ccc(Cl)cc1
Drug	zolpidem	Cc1ccc(-c2nc3ccc(C)cn3c2CC(=O)N(C)C)cc1
Drug	Zolmitriptan	CN(C)CCc1c[nH]c2ccc(C[C@H]3COC(=O)N3)cc12

Drug	Ziprasidone	O=C1Cc2cc(CCN3CCN(c4nsc5ccccc45)CC3)c(Cl)cc2N1
Drug	Zileuton	CC(c1cc2ccccc2s1)N(O)C(N)=O
Drug	Zidovudine	Cc1cn([C@H]2C[C@H](N=[N+]=[N-])[C@@H](CO)O2)c(=O)[nH]c1=O
Drug	zatosetron	CN1[C@H]2CC[C@@H]1CC(N=C(O)c1cc(Cl)cc3c1OC(C)(C)C3)C2
Drug	Zaltoprofen	CC(C(=O)O)c1ccc2c(c1)CC(=O)c1ccccc1S2
Drug	Zaleplon	CCN(C(C)=O)c1cccc(-c2ccnc3c(C#N)cnn23)c1
Drug	zafirlukast	CC1=CC=CC=C1S(=O)(=O)NC(=O)C2=CC(=C(C=C2)CC3=CN(C4=C3C=C(C=C4)NC(=O)OC5CCCC5)C)OC
Drug	YM-64227	CCc1ccc2c(C3CCCCC3)nc(=O)n(CC)c2n1
Drug	warfarin-s	CC(=O)C[C@@H](c1ccccc1)c1c(O)oc2ccccc2c1=O
Drug	warfarin-r	CC(=O)C[C@H](c1ccccc1)c1c(O)oc2ccccc2c1=O
Drug	Vortioxetine	Cc1ccc(Sc2ccccc2N2CCNCC2)c(C)c1
Drug	Vorinostat	O=C(CCCCCCC(=O)Nc1ccccc1)NO
Drug	voriconazole	C[C@@H](c1ncncc1F)[C@](O)(Cn1cncn1)c1ccc(F)cc1F
Drug	Vorapaxar	CCOC(=O)N[C@@H]1CC[C@@H]2[C@@H](C1)C[C@H]1C(=O)O[C@H](C)[C@H]1[C@H]2/C=C/c1ccc(-c2cccc(F)c2)cn1
Drug	Vinorelbine	CCC1=C[C@@H]2CN(C1)Cc1c([nH]c3ccccc13)[C@@](C(=O)OC)(c1cc3c(cc1OC)N(C)[C@H]1[C@@](O)(C(=O)OC)[C@H](OC(C)=O)[C@]4(CC)C=CCN5CC[C@]31[C@@H]54)C2
Drug	Vinflunine	CC[C@]12C=CCN3CC[C@@]4(c5cc([C@@]6(C(=O)OC)C[C@@H]7C[C@@H](C(C)(F)F)CN(Cc8c6[nH]c6ccccc86)C7)c(OC)cc5N(C)[C@H]4[C@@](O)(C(=O)OC)[C@@H]1OC(C)=O)[C@@H]32
Drug	Vinblastine	CC[C@]1(O)C[C@@H]2CN(CCc3c([nH]c4ccccc34)[C@@](C(=O)OC)(c3cc4c(cc3OC)N(C)[C@H]3[C@@](O)(C(=O)OC)[C@H](OC(C)=O)[C@]5(CC)C=CCN6CC[C@]43[C@@H]65)C2)C1
Drug	Vilanterol	OCc1cc([C@@H](O)CNCCCCCCOCCOCc2c(Cl)cccc2Cl)ccc1O
Drug	Vicriviroc	COC[C@@H](c1ccc(C(F)(F)F)cc1)N1CCN(C2(C)CCN(C(=O)c3c(C)ncnc3C)CC2)C[C@@H]1C
Drug	Veratric acid	COc1ccc(C(=O)O)cc1OC
Drug	Verapamil	COc1ccc(CCN(C)CCCC(C#N)(c2ccc(

		OC)c(OC)c2)C(C)C)cc1OC
Drug	Venlafaxine	COc1ccc(C(CN(C)C)C2(O)CCCCC2)cc1
Drug	vanoxerine	Fc1ccc(C(OCCN2CCN(CCCc3ccccc3)CC2)c2ccc(F)cc2)cc1
Drug	Vandetanib	COc1cc2c(Nc3ccc(Br)cc3F)ncnc2cc1OCC1CCN(C)CC1
Drug	valspodar	C/C=C/C[C@@H](C)C(=O)[C@H]1C(O)=N[C@@H](C(C)C)C(=O)N(C)CC(=O)N(C)[C@@H](CC(C)C)C(O)=N[C@@H](C(C)C)C(=O)N(C)[C@@H](CC(C)C)C(O)=N[C@@H](C)C(O)=N[C@H](C)C(=O)N(C)[C@@H](CC(C)C)C(=O)N(C)[C@@H](CC(C)C)C(=O)N(C)[C@@H](C(C)C)C(=O)N1C
Drug	Valsartan	CCCCC(=O)N(Cc1ccc(-c2ccccc2-c2nn[nH]n2)cc1)[C@H](C(=O)O)C(C)C
Drug	Valrubicin	CCCCC(=O)OCC(=O)[C@]1(O)Cc2c(O)c3c(c(O)c2[C@@H](O[C@H]2C[C@H](NC(=O)C(F)(F)F)[C@H](O)[C@H](C)O2)C1)C(=O)c1c(OC)cccc1C3=O
Drug	valproic acid	CCCC(CCC)C(=O)O
Drug	valdecoxib	Cc1onc(-c2ccccc2)c1-c1ccc(S(N)(=O)=O)cc1
Drug	Valbenazine	COc1cc2c(cc1OC)[C@H]1C[C@@H](OC(=O)[C@@H](N)C(C)C)[C@H](CC(C)C)CN1CC2
Drug	v11294	CCNc1ncn(Cc2ccc(OC)c(OC3CCCC3)c2)c2nc(C(C)C)nc1-2
Drug	Uridine triacetate	CC(=O)OC[C@H]1O[C@@H](n2ccc(=O)[nH]c2=O)[C@H](OC(C)=O)[C@@H]1OC(C)=O
Drug	Urethane	CCOC(=N)O
Drug	Udenafil	CCCOc1ccc(S(=O)(=O)NCCC2CCCN2C)cc1-c1nc(=O)c2c([nH]1)c(CCC)nn2C
Drug	Tyrphostin B42	N#C/C(=C\c1ccc(O)c(O)c1)C(O)=NCc1ccccc1
Drug	trovafloxacin	N[C@H]1[C@@H]2CN(c3nc4c(cc3F)c(=O)c(C(=O)O)cn4-c3ccc(F)cc3F)C[C@H]12
Drug	tropisetron	CN1[C@H]2CC[C@@H]1CC(OC(=O)c1c[nH]c3ccccc13)C2
Drug	Troleandomycin	CO[C@H]1C[C@H](O[C@@H]2[C@@H](C)C(=O)O[C@H](C)[C@H](C)[C@H](OC(C)=O)[C@@H](C)C(=O)[C@]3(CO3)C[C@H](C)[C@H](O[C@@H]3O[C@H](C)C[C@H](N(C)C)[C@H]3OC(C)=O)[C@H]2C)O[C

		@@H](C)[C@@H]1OC(C)=O
Drug	Trolamine	OCCN(CCO)CCO
Drug	Troglitazone	Cc1c(C)c2c(c(C)c1O)CCC(C)(COc1ccc(CC3SC(=O)N=C3O)cc1)O2
Drug	Trofosfamide	O=P1(N(CCCl)CCCl)OCCCN1CCCl
Drug	Tripelennamine	CN(C)CCN(Cc1ccccc1)c1ccccn1
Drug	Trimethoprim	COc1cc(Cc2cnc(N)nc2N)cc(OC)c1OC
Drug	trimethadione	CN1C(=O)OC(C)(C)C1=O
Drug	Trimebutine	CCC(COC(=O)c1cc(OC)c(OC)c(OC)c1)(c1ccccc1)N(C)C
Drug	Triflusal	CC(=O)Oc1cc(C(F)(F)F)ccc1C(=O)O
Drug	Trifluridine	O=c1[nH]c(=O)n([C@H]2C[C@H](O)[C@@H](CO)O2)cc1C(F)(F)F
Drug	Triethylenetetramine	NCCNCCNCCN
Drug	triazolam	Cc1nnc2n1-c1ccc(Cl)cc1C(c1ccccc1Cl)=NC2
Drug	Triamterene	Nc1nc(N)c2nc(-c3ccccc3)c(N)nc2n1
Drug	Triamcinolone	C[C@]12C[C@H](O)[C@@]3(F)[C@@H](CCC4=CC(=O)C=C[C@@]43C)[C@@H]1C[C@@H](O)[C@]2(O)C(=O)CO
Drug	trazodone	O=c1n(CCCN2CCN(c3cccc(Cl)c3)CC2)nc2ccccn12
Drug	traxoprodil	C[C@@H]([C@@H](O)c1ccc(O)cc1)N1CCC(O)(c2ccccc2)CC1
Drug	Travoprost	CC(C)OC(=O)CCC/C=C\C[C@H]1[C@@H](O)C[C@@H](O)[C@@H]1/C=C/[C@@H](O)COc1cccc(C(F)(F)F)c1
Drug	Trandolapril	CCOC(=O)[C@H](CCc1ccccc1)N[C@@H](C)C(=O)N1[C@H](C(=O)O)C[C@H]2CCCC[C@@H]21
Drug	Tramadol	COc1cccc([C@@]2(O)CCCC[C@@H]2CN(C)C)c1
Drug	Trabectedin	COc1cc2c(cc1O)CCN[C@]21CS[C@@H]2c3c(OC(C)=O)c(C)c4c(c3[C@H](COC1=O)N1[C@@H](O)[C@@H]3Cc5cc(C)c(OC)c(O)c5[C@H]([C@H]21)N3C)OCO4
Drug	tpa-023	CCn1ncnc1COc1nn2c(-c3ccccc3F)nnc2cc1C(C)(C)C
Drug	tozasertib	Cc1cc(=Nc2cc(N3CCN(C)CC3)nc(Sc3ccc(NC(=O)C4CC4)cc3)n2)[nH][nH]1
Drug	Toremifene	CN(C)CCOc1ccc(/C(=C(/CCCl)c2ccccc2)c2ccccc2)cc1
Drug	Torasemide	Cc1cccc(Nc2ccncc2S(=O)(=O)NC(=O)NC(C)C)c1
Drug	Topiramate	CC1(C)O[C@@H]2CO[C@@]3(COS(N)(=O)=O)OC(C)(C)O[C@H]3[C@@H]2O1

Drug	Toluene	Cc1ccccc1
Drug	Tolterodine	Cc1ccc(O)c([C@H](CCN(C(C)C)C(C)C)c2ccccc2)c1
Drug	Tolmetin	Cc1ccc(C(=O)c2ccc(CC(=O)O)n2C)cc1
Drug	Tolfenamic acid	Cc1c(Cl)cccc1Nc1ccccc1C(=O)O
Drug	Tolcapone	Cc1ccc(C(=O)c2cc(O)c(O)c([N+](=O)[O-])c2)cc1
Drug	tolbutamide	CCCCN=C(O)NS(=O)(=O)c1ccc(C)cc1
Drug	Tizanidine	Clc1ccc2nsnc2c1NC1=NCCN1
Drug	tirilazad	C[C@@H]1C[C@H]2[C@@H]3CCC4=CC(=O)C=C[C@]4(C)C3=CC[C@]2(C)[C@H]1C(=O)CN1CCN(c2cc(N3CCCC3)nc(N3CCCC3)n2)CC1
Drug	Tipiracil	N=C1CCCN1Cc1[nH]c(=O)[nH]c(=O)c1Cl
Drug	Tiopronin	CC(S)C(=O)NCC(=O)O
Drug	Tioguanine	Nc1nc(=S)c2[nH]cnc2[nH]1
Drug	timolol	CC(C)(C)NC[C@H](O)COc1nsnc1N1CCOCC1
Drug	ticrynafen	O=C(O)COc1ccc(C(=O)c2cccs2)c(Cl)c1Cl
Drug	ticlopidine	Clc1ccccc1CN1CCc2sccc2C1
Drug	Ticagrelor	CCCSc1nc(N[C@@H]2C[C@H]2c2ccc(F)c(F)c2)c2nnn([C@@H]3C[C@H](OCCO)[C@@H](O)[C@H]3O)c2n1
Drug	Thymol	Cc1ccc(C(C)C)c(O)c1
Drug	Thioridazine	CSc1ccc2c(c1)N(CCC1CCCCN1C)c1ccccc1S2
Drug	Thiocolchicoside	O=C(N[C@@H]4C1=C\C(=O)C(/SC)=C\C=C1c3c(cc(O[C@@H]2O[C@@H]([C@@H](O)[C@H](O)[C@H]2O)CO)c(OC)c3OC)CC4)C
Drug	Thienorphine	CO[C@]12CC[C@]3(CC1[C@](C)(O)CCc1cccs1)[C@H]1Cc4ccc(O)c5c4[C@@]3(CCN1CC1CC1)C2O5
Drug	thiabendazole	c1ccc2[nH]c(-c3cscn3)nc2c1
Drug	Theophylline	Cn1c(=O)c2[nH]cnc2n(C)c1=O
Drug	Thalidomide	O=C1CCC(N2C(=O)c3ccccc3C2=O)C(=O)N1
Drug	Tezacaftor	CC(C)(CO)c1cc2cc(NC(=O)C3(c4ccc5c(c4)OC(F)(F)O5)CC3)c(F)cc2n1C[C@@H](O)CO
Drug	tetranor-(+)-s-145	O=C(O)CC[C@@H]1[C@@H](NS(=O)(=O)c2ccccc2)[C@H]2CC[C@@H]1C2
Drug	Tetrahydrofolic acid	Nc1nc(=O)c2c([nH]1)NCC(CNc1ccc(C(=O)N[C@@H](CCC(=O)O)C(=O)O)cc1)N2
Drug	Tetracaine	CCCCNc1ccc(C(=O)OCCN(C)C)cc1

Drug	Tetrabenazine	COc1cc2c(cc1OC)C1CC(=O)C(CC(C)C)CN1CC2
Drug	Terfenadine	CC(C)(C)c1ccc(C(O)CCCN2CCC(C(O)(c3ccccc3)c3ccccc3)CC2)cc1
Drug	Terbinafine	CN(C/C=C/C#CC(C)(C)C)Cc1cccc2ccccc12
Drug	Terazosin	COc1cc2nc(N3CCN(C(=O)C4CCCO4)CC3)nc(N)c2cc1OC
Drug	Tenoxicam	CN1C(C(=O)Nc2ccccn2)=C(O)c2sccc2S1(=O)=O
Drug	Tenofovir alafenamide	CC(C)OC(=O)[C@H](C)N[P@](=O)(CO[C@H](C)Cn1cnc2c(N)ncnc21)Oc1ccccc1
Drug	Tenofovir	C[C@H](Cn1cnc2c(N)ncnc21)OCP(=O)(O)O
Drug	Teniposide	COc1cc([C@@H]2c3cc4c(cc3[C@@H](O[C@@H]3O[C@@H]5CO[C@@H](c6cccs6)O[C@H]5[C@H](O)[C@H]3O)[C@H]3COC(=O)[C@H]23)OCO4)cc(OC)c1O
Drug	Temsirolimus	CO[C@H]1C[C@@H]2CC[C@@H](C)[C@@](O)(O2)C(=O)C(=O)N2CCCC[C@H]2C(=O)O[C@H]([C@H](C)C[C@@H]2CC[C@@H](OC(=O)C(C)(CO)CO)[C@H](OC)C2)CC(=O)[C@H](C)/C=C(\C)[C@@H](O)[C@@H](OC)C(=O)[C@H](C)C[C@H](C)/C=C/C=C/C=C/1C
Drug	Temazepam	CN1C(=O)C(O)N=C(c2ccccc2)c2cc(Cl)ccc21
Drug	tegaserod	CCCCCNC(=N)N/N=C/c1c[nH]c2ccc(OC)cc12
Drug	Tegafur	O=c1nc(O)c(F)cn1C1CCCO1
Drug	Tedizolid phosphate	Cn1nnc(-c2ccc(-c3ccc(N4C[C@H](COP(=O)(O)O)OC4=O)cc3F)cn2)n1
Drug	Tedizolid	Cn1nnc(-c2ccc(-c3ccc(N4C[C@H](CO)OC4=O)cc3F)cn2)n1
Drug	tecalcet	COc1cccc([C@@H](C)NCCCc2ccccc2Cl)c1
Drug	Tazemetostat	CCN(c1cc(-c2ccc(CN3CCOCC3)cc2)cc(C(=O)Nc2c(C)cc(C)[nH]c2=O)c1C)C1CCOCC1
Drug	Tazarotene	CCOC(=O)c1ccc(C#Cc2ccc3c(c2)C(C)(C)CCS3)nc1
Drug	taxol	CC(=O)O[C@H]1C(=O)[C@]2(C)[C@@H](O)C[C@H]3OC[C@@]3(OC(C)=O)[C@@H]2[C@H](OC(=O)c2ccccc2)[C@]2(O)C[C@H](OC(=O)[C@H](O)[C@@H](N=C(O)c3ccccc3)c3c

		cccc3)C(C)=C1C2(C)C
Drug	tauromustine	CN(C)S(=O)(=O)CCN=C(O)N(CCCl)N=O
Drug	Tasosartan	Cc1nc(C)c2c(n1)N(Cc1ccc(-c3ccccc3-c3nn[nH]n3)cc1)C(=O)CC2
Drug	Tapentadol	CC[C@@H](c1cccc(O)c1)[C@@H](C)CN(C)C
Drug	tangeretin	COc1ccc(-c2cc(=O)c3c(OC)c(OC)c(OC)c(OC)c3o2)cc1
Drug	tamsulosin	CCOc1ccccc1OCCN[C@H](C)Cc1ccc(OC)c(S(N)(=O)=O)c1
Drug	tamoxifen	CC/C(=C(\c1ccccc1)c1ccc(OCCN(C)C)cc1)c1ccccc1
Drug	tamarixetin	COc1ccc(-c2oc3cc(O)cc(O)c3c(=O)c2O)cc1O
Drug	Tafenoquine	COc1cc(C)c2c(Oc3cccc(C(F)(F)F)c3)c(OC)cc(NC(C)CCCN)c2n1
Drug	Tacrolimus	C=CC[C@@H]1/C=C(\C)C[C@H](C)C[C@H](OC)[C@H]2O[C@@](O)(C(=O)C(=O)N3CCCC[C@H]3C(=O)O[C@H](/C(C)=C/[C@@H]3CC[C@@H](O)[C@H](OC)C3)[C@H](C)[C@@H](O)CC1=O)[C@H](C)C[C@@H]2OC
Drug	tacrine	N=c1c2c([nH]c3ccccc13)CCCC2
Drug	Syringetin	COc1cc(-c2oc3cc(O)cc(O)c3c(=O)c2O)cc(OC)c1O
Drug	suprofen	CC(C(=O)O)c1ccc(C(=O)c2cccs2)cc1
Drug	Sumatriptan	CNS(=O)(=O)Cc1ccc2[nH]cc(CCN(C)C)c2c1
Drug	Sulphadimethoxine	COc1cc(NS(=O)(=O)c2ccc(N)cc2)nc(OC)n1
Drug	Sulindac	CC1=C(CC(=O)O)c2cc(F)ccc2/C1=C\c1ccc(S(C)=O)cc1
Drug	Sulfinpyrazone	O=C1C(CCS(=O)c2ccccc2)C(=O)N(c2ccccc2)N1c1ccccc1
Drug	Sulfapyridine	Nc1ccc(S(=O)(=O)Nc2ccccn2)cc1
Drug	Sulfamethoxazole	Cc1cc(NS(=O)(=O)c2ccc(N)cc2)no1
Drug	sulfamerazine	Cc1ccnc(NS(=O)(=O)c2ccc(N)cc2)n1
Drug	Stanolone	C[C@]12CC[C@H]3[C@@H](CC[C@H]4CC(=O)CC[C@@]43C)[C@@H]1CC[C@@H]2O
Drug	Spironolactone	CC(=O)S[C@@H]1CC2=CC(=O)CC[C@]2(C)[C@H]2CC[C@@]3(C)[C@@H](CC[C@@]34CCC(=O)O4)[C@@H]21
Drug	Spirapril	CCOC(=O)[C@H](CCc1ccccc1)N[C@@H](C)C(=O)N1CC2(C[C@H]1C(=O)O)SCCS2
Drug	sparteine	C1CCN2C[C@@H]3C[C@@H](CN4

		CCCC[C@H]34)[C@@H]2C1
Drug	Sorafenib	CNC(=O)c1cc(Oc2ccc(NC(O)=Nc3ccc(Cl)c(C(F)(F)F)c3)cc2)ccn1
Drug	Solifenacin	O=C(O[C@H]1CN2CCC1CC2)N1CCc2ccccc2[C@@H]1c1ccccc1
Drug	Sofosbuvir	CC(C)OC(=O)[C@H](C)N[P@](=O)(OC[C@H]1O[C@@H](n2ccc(=O)[nH]c2=O)[C@](C)(F)[C@@H]1O)Oc1ccccc1
Drug	Sodium oxybate	O=C([O-])CCCO.[Na+]
Drug	S-MTPPA	Cc1ccsc1-c1ccc(C(C)(C)C(=O)O)cc1
Drug	s-methyl diethylthiocarbamate	CCN(CC)C(=O)SC
Drug	S-Mephenytoin	CC[C@@]1(c2ccccc2)N=C(O)N(C)C1=O
Drug	Sitagliptin	N[C@@H](CC(=O)N1CCn2c(nnc2C(F)(F)F)C1)Cc1cc(F)c(F)cc1F
Drug	Sirolimus	CO[C@H]1C[C@@H]2CC[C@@H](C)[C@@](O)(O2)C(=O)C(=O)N2CCCC[C@H]2C(=O)O[C@H]([C@H](C)C[C@@H]2CC[C@@H](O)[C@H](OC)C2)CC(=O)[C@H](C)/C=C(\C)[C@@H](O)[C@@H](OC)C(=O)[C@H](C)C[C@H](C)/C=C/C=C/C=C/1C
Drug	Simvastatin	CCC(C)(C)C(=O)O[C@H]1C[C@@H](C)C=C2C=C[C@H](C)[C@H](CC[C@@H]3C[C@@H](O)CC(=O)O3)[C@H]21
Drug	Sildenafil	CCCc1nn(C)c2c(=O)[nH]c(-c3cc(S(=O)(=O)N4CCN(C)CC4)ccc3OCC)nc12
Drug	SFZ-47 (prodrug with a methylphenyl moiety)	O=C1C2=CC=CN2CC1CNC3=CC=C(C)C=C3
Drug	Sevoflurane	FCOC(C(F)(F)F)C(F)(F)F
Drug	sertraline	CN[C@H]1CC[C@@H](c2ccc(Cl)c(Cl)c2)c2ccccc21
Drug	sertindole	OC1=NCCN1CCN1CCC(c2cn(-c3ccc(F)cc3)c3ccc(Cl)cc23)CC1
Drug	Seratrodast	CC1=C(C)C(=O)C(C(CCCCCC(=O)O)c2ccccc2)=C(C)C1=O
Drug	Sennosides	O=C(O)c1cc(O)c2c(c1)C(C1c3cc(C(=O)O)cc(O)c3C(=O)c3c(OC4OC(CO)C(O)C(O)C4O)cccc31)c1cccc(OC3OC(CO)C(O)C(O)C3O)c1C2=O
Drug	Selumetinib	Cn1cnc2c(F)c(Nc3ccc(Br)cc3Cl)c(C(=O)NOCCO)cc21
Drug	Selexipag	CC(C)N(CCCCOCC(=O)NS(C)(=O)=O)c1cnc(-c2ccccc2)c(-c2ccccc2)n1
Drug	selegiline	C#CCN(C)[C@H](C)Cc1ccccc1
Drug	Segesterone acetate	C=C1C[C@H]2[C@@H]3CCC4=CC(=O)CC[C@@H]4[C@H]3CC[C@]2(C)[C@@]1(OC(C)=O)C(C)=O
Drug	schembl9652240	C[C@H]1S[C@H](c2cccnc2)N(C)C1

		=O
Drug	schembl896413	COc1cc(C[C@@H]2SC(=O)N=C2O)ccc1C(O)=NCc1ccc(C(F)(F)F)cc1
Drug	schembl895799	C[C@@H](C#Cc1ccc(Cc2ccccc2)s1)NC(=N)O
Drug	schembl7750160	CN(C)CCCCCN1c2ccccc2Sc2ccc(C(F)(F)F)cc21
Drug	schembl7562446	CC(C)(C)c1cc(C[C@@H]2SCN=C2O)cc(C(C)(C)C)c1O
Drug	schembl69463	CCCOc1ccc(S(=O)(=O)NCC[C@@H]2CCCN2C)cc1-c1nc2c(CCC)nn(C)c2c(=O)[nH]1
Drug	schembl3207291	CN[C@@H]1CN(c2ccc3c(=O)c(C(=O)O)cn(-c4nccs4)c3n2)C[C@H]1OC
Drug	schembl237672	CC[N+](C)(CC)CCc1c(C)c2ccc(OC)cc2oc1=O
Drug	schembl17043265	CC[C@@H](C(=O)c1ccc(C)cc1)N1CCCC1
Drug	schembl14329530	CCCCCCOc1ccc2ccc(=O)oc2c1
Drug	schembl14141926	COC(=O)[C@H](c1ccccc1Cl)N1CCC2=C(CC(=O)S2)C1
Drug	schembl12998564	COc1ccccc1N=C(O)/C(=N/Nc1ccc([N+](=O)[O-])cc1OC)C(C)=O
Drug	S-Bioallethrin	C=CCC1=C(C)[C@@H](OC(=O)[C@@H]2[C@@H](C=C(C)C)C2(C)C)CC1=O
Drug	Saxagliptin	N#C[C@@H]1C[C@@H]2C[C@@H]2N1C(=O)[C@@H](N)C12CC3CC(CC(O)(C3)C1)C2
Drug	Saquinavir	CC(C)(C)NC(=O)[C@@H]1C[C@@H]2CCCC[C@@H]2CN1C[C@@H](O)[C@H](Cc1ccccc1)NC(=O)[C@H](CC(N)=O)NC(=O)c1ccc2ccccc2n1
Drug	Salmeterol	OCc1cc(C(O)CNCCCCCCOCCCCc2ccccc2)ccc1O
Drug	Salbutamol	CC(C)(C)NCC(O)c1ccc(O)c(CO)c1
Drug	Sacubitril	CCOC(=O)[C@H](C)C[C@@H](Cc1ccc(-c2ccccc2)cc1)NC(=O)CCC(=O)O
Drug	S-2-[4-(3-methyl-2-thienyl)phenyl]propionic acid	Cc1ccsc1-c1ccc(C(C)C(=O)O)cc1
Drug	Rupatadine	Cc1cncc(CN2CCC(=C3c4ccc(Cl)cc4CCc4cccnc43)CC2)c1
Drug	Rufinamide	NC(=O)c1cn(Cc2c(F)cccc2F)nn1
Drug	ruboxistaurin	CN(C)C[C@@H]1CCn2cc(c3ccccc32)C2=C(C(=O)N=C2O)c2cn(c3ccccc23)CCO1
Drug	R-Tolperisone	Cc1ccc(C(=O)[C@H](C)CN2CCCCC2)cc1
Drug	RPR-106541	CCC[C@@H]1O[C@@H]2C[C@H]3[C@@H]4C[C@H](F)C5=CC(=O)CC[C@]5(C)[C@@]4(F)[C@@H](O)C[C@]3(C)[C@]2(SC)O1

Drug	RP73401	COc1ccc(C(=O)N=c2c(Cl)c[nH]cc2Cl)cc1OC1CCCC1
Drug	roxithromycin	CC[C@H]1OC(=O)[C@H](C)[C@@H](O[C@@H]2C[C@](C)(OC)[C@H](O)[C@@H](C)O2)[C@@H](C)[C@@H](O[C@@H]2O[C@H](C)C[C@@H](N(C)C)[C@H]2O)[C@](C)(O)C[C@@H](C)/C(=N\OCOCCOC)[C@H](C)[C@H](O)[C@]1(C)O
Drug	Rotigotine	CCCN(CCc1cccs1)[C@H]1CCc2c(O)cccc2C1
Drug	Rosuvastatin	CC(C)c1nc(N(C)S(C)(=O)=O)nc(-c2ccc(F)cc2)c1/C=C/[C@H](O)C[C@H](O)CC(=O)O
Drug	Rosiglitazone	CN(CCOc1ccc(CC2SC(=O)NC2=O)cc1)c1ccccn1
Drug	ropivacaine	CCCN1CCCC[C@H]1C(O)=Nc1c(C)cccc1C
Drug	ropinirole	CCCN(CCC)CCc1cccc2c1CC(O)=N2
Drug	Romidepsin	C/C=C1\NC(=O)[C@H]2CSSCC/C=C/[C@H](CC(=O)N[C@H](C(C)C)C(=O)N2)OC(=O)[C@H](C(C)C)NC1=O
Drug	Rolapitant	C[C@@H](OC[C@@]1(c2ccccc2)CC[C@]2(CCC(=O)N2)CN1)c1cc(C(F)(F)F)cc(C(F)(F)F)c1
Drug	Roflumilast	O=C(Nc1c(Cl)cncc1Cl)c1ccc(OC(F)F)c(OCC2CC2)c1
Drug	Rofecoxib	CS(=O)(=O)c1ccc(C2=C(c3ccccc3)C(=O)OC2)cc1
Drug	Rocuronium	C=CC[N+]1([C@H]2C[C@H]3[C@@H]4CC[C@H]5C[C@H](O)[C@@H](N6CCOCC6)C[C@]5(C)[C@H]4CC[C@]3(C)[C@H]2OC(C)=O)CCCC1
Drug	RL-3142	CN(C)c1ccc(/C=C/C(=O)c2ccc(-n3ccnc3)cc2)cc1
Drug	Rizatriptan	CN(C)CCc1c[nH]c2ccc(Cn3cncn3)cc12
Drug	Ritonavir	CC(C)c1nc(CN(C)C(=O)N[C@H](C(=O)N[C@@H](Cc2ccccc2)C[C@H](O)[C@H](Cc2ccccc2)NC(=O)OCc2cncs2)C(C)C)cs1
Drug	risperidone	Cc1nc2n(c(=O)c1CCN1CCC(c3noc4cc(F)ccc34)CC1)CCCC2
Drug	riluzole	N=c1[nH]c2ccc(OC(F)(F)F)cc2s1
Drug	Rifamycin	CO[C@H]1/C=C/O[C@@]2(C)Oc3c(C)c(O)c4c(O)c(cc(O)c4c3C2=O)NC(=O)/C(C)=C\C=C\[C@H](C)[C@H](O)[C@@H](C)[C@@H](O)[C@@H](C)[C@H](OC(C)=O)[C@@H]1C
Drug	rifalazil	CO[C@H]1/C=C/O[C@@]2(C)Oc3c(C)c(O)c4c(=O)c(c5oc6cc(N7CCN(CC(C)C)CC7)cc(O)c6nc-

		5c4c3C2=O)N=C(O)/C(C)=C\C=C\[C@H](C)[C@H](O)[C@@H](C)[C@H](O)[C@H](C)[C@H](OC(C)=O)[C@@H]1C
Drug	rifabutin	CO[C@H]1/C=C\O[C@@]2(C)Oc3c(C)c(O)c4c(c3C2=O)C2=NC3(CCN(CC(C)C)CC3)NC2=C(CC(=O)/C(C)=C\C=C\[C@H](C)[C@@H](O)[C@H](C)[C@@H](O)[C@H](C)[C@H](OC(C)=O)[C@H]1C)C4=O
Drug	R-Ibuprofen	CC(C)Cc1ccc([C@@H](C)C(=O)O)cc1
Drug	RG12525	c1ccc(-c2nn[nH]n2)c(COc2ccc(OCc3ccc4ccccc4n3)cc2)c1
Drug	Revefenacin	CN(CCN1CCC(OC(=O)Nc2ccccc2-c2ccccc2)CC1)C(=O)c1ccc(CN2CCC(C(N)=O)CC2)cc1
Drug	retigabine	CCOC(O)=Nc1ccc(NCc2ccc(F)cc2)cc1N
Drug	Resacetophenone	CC(=O)c1ccc(O)cc1O
Drug	repaglinide	CCOc1cc(CC(O)=N[C@@H](CC(C)C)c2ccccc2N2CCCCC2)ccc1C(=O)O
Drug	Remdesivir	CCC(CC)COC(=O)[C@H](C)N[P@](=O)(OC[C@H]1O[C@@](C#N)(c2ccc3c(N)ncnn23)[C@H](O)[C@@H]1O)Oc1ccccc1
Drug	reduced_haloperidol	O[C@@H](CCCN1CCC(O)(c2ccc(Cl)cc2)CC1)c1ccc(F)cc1
Drug	reboxetine	CCOc1ccccc1O[C@@H](c1ccccc1)[C@@H]1CNCCO1
Drug	Ranolazine	COc1ccccc1OCC(O)CN1CCN(CC(=O)Nc2c(C)cccc2C)CC1
Drug	Ranitidine	CNC(=C[N+](=O)[O-])NCCSCc1ccc(CN(C)C)o1
Drug	Ramipril	CCOC(=O)[C@H](CCc1ccccc1)N[C@@H](C)C(=O)N1[C@H](C(=O)O)C[C@@H]2CCC[C@@H]21
Drug	raloxifene	O=C(c1ccc(OCCN2CCCCC2)cc1)c1c(-c2ccc(O)cc2)sc2cc(O)ccc12
Drug	Rabeprazole (rabeprazole_thioether)	COCCCOc1ccnc(CSc2nc3ccccc3[nH]2)c1C
Drug	Rabeprazole	COCCCOc1ccnc(CS(=O)c2nc3ccccc3[nH]2)c1C
Drug	qun	CCN(CC)CCC[C@@H](C)N=c1c2ccc(Cl)cc2[nH]c2ccc(OC)cc12
Drug	Quinotolast	O=c1c(C(O)=Nc2nn[nH]n2)cc(Oc2ccccc2)c2ccccn12
Drug	quinine	C=C[C@H]1CN2CC[C@H]1C[C@H]2[C@H](O)c1ccnc2ccc(OC)cc12
Drug	quinidine	C=C[C@H]1CN2CC[C@H]1C[C@@H]2[C@@H](O)c1ccnc2ccc(OC)cc12

Drug	Quinestrol	C#C[C@]1(O)CC[C@H]2[C@@H]3CCc4cc(OC5CCCC5)ccc4[C@H]3CC[C@@]21C
Drug	Quinapril	CCOC(=O)[C@H](CCc1ccccc1)N[C@@H](C)C(=O)N1Cc2ccccc2C[C@H]1C(=O)O
Drug	Quinagolide	CCCN1C[C@@H](NS(=O)(=O)N(CC)CC)C[C@@H]2Cc3c(O)cccc3C[C@H]21
Drug	quetiapine	OCCOCCN1CCN(C2=Nc3ccccc3Sc3ccccc32)CC1
Drug	quazepam	Fc1ccccc1C1=NCC(=S)N(CC(F)(F)F)c2ccc(Cl)cc21
Drug	pyrazoloacridine	COc1ccc2c(c1)-c1nn(CCCN(C)C)c3ccc([N+](=O)[O-])c(c13)N2
Drug	Pyrazinamide	NC(=O)c1cnccn1
Drug	Psilocin	CN(C)CCc1c[nH]c2cccc(O)c12
Drug	Pseudoephedrine	CN[C@@H](C)[C@@H](O)c1ccccc1
Drug	Prucalopride	COCCCN1CCC(NC(=O)c2cc(Cl)c(N)c3c2OCC3)CC1
Drug	Propranolol	CC(C)NCC(O)COc1cccc2ccccc12
Drug	Propofol	CC(C)c1cccc(C(C)C)c1O
Drug	Propiolactone	O=C1CCO1
Drug	Propafenone	CCCNCC(O)COc1ccccc1C(=O)CCc1ccccc1
Drug	Propacetamol	CCN(CC)CC(=O)Oc1ccc(NC(C)=O)cc1
Drug	Promethazine	CC(CN1c2ccccc2Sc2ccccc21)N(C)C
Drug	promazine	CN(C)CCCN1c2ccccc2Sc2ccccc21
Drug	Proguanil	CC(C)NC(=N)NC(=N)Nc1ccc(Cl)cc1
Drug	Prochlorperazine	CN1CCN(CCCN2c3ccccc3Sc3ccc(Cl)cc32)CC1
Drug	Procarbazine	CNNCc1ccc(C(=O)NC(C)C)cc1
Drug	Procainamide	CCN(CC)CCNC(=O)c1ccc(N)cc1
Drug	Primidone	CCC1(c2ccccc2)C(=O)NCNC1=O
Drug	Prednisone	C[C@]12CC(=O)[C@H]3[C@@H](CCC4=CC(=O)C=C[C@@]43C)[C@@H]1CC[C@]2(O)C(=O)CO
Drug	Prednisolone acetate	CC(=O)OCC(=O)[C@@]1(O)CC[C@H]2[C@@H]3CCC4=CC(=O)C=C[C@]4(C)[C@H]3[C@@H](O)C[C@@]21C
Drug	Prednisolone	C[C@]12C[C@H](O)[C@H]3[C@@H](CCC4=CC(=O)C=C[C@@]43C)[C@@H]1CC[C@]2(O)C(=O)CO
Drug	Prazosin	COc1cc2nc(N3CCN(C(=O)c4ccco4)CC3)nc(N)c2cc1OC
Drug	Pravastatin	CC[C@H](C)C(=O)O[C@H]1C[C@H](O)C=C2C=C[C@H](C)[C@H](CC[C@@H](O)C[C@@H](O)CC(=O)O)[

		C@H]21
Drug	Prasugrel	CC(=O)Oc1cc2c(s1)CCN(C(C(=O)C1CC1)c1ccccc1F)C2
Drug	Pranidipine	COC(=O)C1=C(C)NC(C)=C(C(=O)OC/C=C/c2ccccc2)C1c1cccc([N+](=O)[O-])c1
Drug	pradefovir	Nc1ncnc2c1ncn2CCOC[P@]1(=O)OCC[C@@H](c2cccc(Cl)c2)O1
Drug	PNU-96391	CCCN1CCC[C@@H](c2cccc(S(C)(=O)=O)c2)C1
Drug	Pitolisant	Clc1ccc(CCCOCCCN2CCCCC2)cc1
Drug	piroxicam	CN1C(C(O)=Nc2ccccn2)=C(O)c2ccccc2S1(=O)=O
Drug	Piperazine	C1CNCCN1
Drug	Piperaquine	Clc1ccc2c(N3CCN(CCCN4CCN(c5ccnc6cc(Cl)ccc56)CC4)CC3)ccnc2c1
Drug	Pioglitazone	CCC1=CN=C(CCOC2=CC=C(CC3SC(=O)NC3=O)C=C2)C=C1
Drug	Pindolol	CC(C)NCC(O)COc1cccc2[nH]ccc12
Drug	pimozide	Oc1nc2ccccc2n1C1CCN(CCCC(c2ccc(F)cc2)c2ccc(F)cc2)CC1
Drug	pimobendan	COc1ccc(-c2nc3cc(C4=NN=C(O)C[C@@H]4C)ccc3[nH]2)cc1
Drug	pilocarpine	CC[C@@H]1C(=O)OC[C@@H]1Cc1cncn1C
Drug	Pholcodine	CN1CC[C@]23c4c5ccc(OCCN6CCOCC6)c4O[C@H]2[C@@H](O)C=C[C@H]3[C@H]1C5
Drug	phenytoin	O=C1NC(=O)C(N1)(C1=CC=CC=C1)C1=CC=CC=C1
Drug	Phenylephrine	CNC[C@H](O)c1cccc(O)c1
Drug	Phenylbutyric acid	O=C(O)CCCc1ccccc1
Drug	Phenylacetic acid	O=C(O)Cc1ccccc1
Drug	phentolamine	Cc1ccc(N(CC2=NCCN2)c2cccc(O)c2)cc1
Drug	Phentermine	CC(C)(N)Cc1ccccc1
Drug	Phenprocoumon	CCC(c1ccccc1)c1c(O)c2ccccc2oc1=O
Drug	Phenoxymethylpenicillin	CC1(C)S[C@@H]2[C@H](NC(=O)COc3ccccc3)C(=O)N2[C@H]1C(=O)O
Drug	phenol	Oc1ccccc1
Drug	Phenobarbital	CCC1(C(=O)NC(=O)NC1=O)C1=CC=CC=C1
Drug	Pheniramine	CN(C)CCC(c1ccccc1)c1ccccn1
Drug	phenformin	N=C(N)NC(=N)NCCc1ccccc1
Drug	Phenelzine	NNCCc1ccccc1
Drug	Phendimetrazine	C[C@H]1[C@H](c2ccccc2)OCCN1C
Drug	phencyclidine	c1ccc(C2(N3CCCCC3)CCCCC2)cc1
Drug	Phenazopyridine	Nc1ccc(/N=N/c2ccccc2)c(N)n1
Drug	phenacetin	CCOc1ccc(N=C(C)O)cc1

Drug	perphenazine	OCCN1CCN(CCCN2c3ccccc3Sc3ccc(Cl)cc32)CC1
Drug	Perospirone	O=C1[C@H]2CCCC[C@H]2C(=O)N1CCCCN1CCN(c2nsc3ccccc23)CC1
Drug	Perindopril	CCC[C@H](N[C@@H](C)C(=O)N1[C@H](C(=O)O)C[C@@H]2CCCC[C@@H]21)C(=O)OCC
Drug	Perhexiline	C1CCC(C(C[C@H]2CCCCN2)C2CCCCC2)CC1
Drug	perazine	CN1CCN(CCCN2c3ccccc3Sc3ccccc32)CC1
Drug	pentoxyresorufin	CCCCCOc1ccc2nc3ccc(=O)cc-3oc2c1
Drug	Pentoxifylline	CC(=O)CCCCn1c(=O)c2c(ncn2C)n(C)c1=O
Drug	Pentaerythritol tetranitrate	O=[N+]([O-])OCC(CO[N+](=O)[O-])(CO[N+](=O)[O-])CO[N+](=O)[O-]
Drug	pefloxacin	CCn1cc(C(=O)O)c(=O)c2cc(F)c(N3CCN(C)CC3)cc21
Drug	Paroxetine	Fc1ccc([C@@H]2CCNC[C@H]2COc2ccc3c(c2)OCO3)cc1
Drug	Paramethadione	CCC1(C)OC(=O)N(C)C1=O
Drug	Paraldehyde	CC1OC(C)OC(C)O1
Drug	pantoprazole	COc1ccnc(C[S@@](=O)c2nc3ccc(OC(F)F)cc3[nH]2)c1OC
Drug	Palbociclib	CC(=O)c1c(C)c2cnc(Nc3ccc(N4CCNCC4)cn3)nc2n(C2CCCC2)c1=O
Drug	Paclitaxel	CC(=O)O[C@H]1C(=O)[C@]2(C)[C@@H](O)C[C@H]3OC[C@@]3(OC(C)=O)[C@H]2[C@H](OC(=O)c2ccccc2)[C@]2(O)C[C@H](OC(=O)[C@H](O)[C@@H](NC(=O)c3ccccc3)c3ccccc3)C(C)=C1C2(C)C
Drug	Oxyquinoline	Oc1cccc2cccnc12
Drug	Oxymorphone	CN1CC[C@]23c4c5ccc(O)c4O[C@H]2C(=O)CC[C@@]3(O)[C@H]1C5
Drug	Oxycodone	COc1ccc2c3c1O[C@H]1C(=O)CC[C@@]4(O)[C@@H](C2)N(C)CC[C@]314
Drug	Oxybutynin	CCN(CC)CC#CCOC(=O)C(O)(c1ccccc1)C1CCCCC1
Drug	Oxoprenolol	C=CCOc1ccccc1OCC(O)CNC(C)C
Drug	OXODIPINE	CCOC(=O)C1=C(C)NC(C)=C(C(=O)OC)C1c1cccc2c1OCO2
Drug	Oxetacaine	CN(C(=O)CN(CCO)CC(=O)N(C)C(C)(C)Cc1ccccc1)C(C)(C)Cc1ccccc1
Drug	Oxcarbazepine	NC(=O)N1c2ccccc2CC(=O)c2ccccc21
Drug	Oxazepam	O=C1Nc2ccc(Cl)cc2C(c2ccccc2)=NC1O
Drug	Oxabolone cipionate	C[C@]12CC[C@@H]3[C@H]4CCC(=O)C(O)=C4CC[C@H]3[C@@H]1CC[C@@H]2OC(=O)CCC1CCCC1
Drug	OT-7100	CCCCc1cc(NC(=O)c2cc(OC)c(OC)c(

		OC)c2)n2nccc2n1
Drug	Orphenadrine	Cc1ccccc1C(OCCN(C)C)c1ccccc1
Drug	Ornidazole	Cc1ncc([N+](=O)[O-])n1CC(O)CCl
Drug	org_30659	C#C[C@]1(O)C=C[C@H]2[C@@H]3CCC4=CC(=O)CC[C@H]4[C@H]3C(=C)C[C@@]21C
Drug	Orciprenaline	CC(C)NCC(O)c1cc(O)cc(O)c1
Drug	Orantinib	Cc1[nH]c(/C=C2\C(O)=Nc3ccccc32)c(C)c1CCC(=O)O
Drug	Ondansetron	Cc1nccn1CC1CCc2c(c3ccccc3n2C)C1=O
Drug	onapristone	CN(C)c1ccc([C@H]2C[C@]3(C)[C@H](CC[C@]3(O)CCCO)[C@@H]3CCC4=CC(=O)CCC4=C32)cc1
Drug	omeprazole	COc1ccc2nc(S(=O)Cc3ncc(C)c(OC)c3C)[nH]c2c1
Drug	Olsalazine	O=C(O)c1cc(/N=N/c2ccc(O)c(C(=O)O)c2)ccc1O
Drug	olopatadine	CN(C)CC/C=C1/c2ccccc2COc2ccc(CC(=O)O)cc21
Drug	Olodaterol	COc1ccc(CC(C)(C)NC[C@H](O)c2cc(O)cc3c2OCC(=O)N3)cc1
Drug	olanzapine	Cc1cc2c(s1)Nc1ccccc1N=C2N1CCN(C)CC1
Drug	O-isopropyl acetaminophen	CC(O)=Nc1ccc(OC(C)C)cc1
Drug	Odanacatib	CC(C)(F)C[C@H](N[C@@H](c1ccc(-c2ccc(S(C)(=O)=O)cc2)cc1)C(F)(F)F)C(=O)NC1(C#N)CC1
Drug	Norverapamil	COc1ccc(CCNCCC[C@](C#N)(c2ccc(OC)c(OC)c2)C(C)C)cc1OC
Drug	nortriptyline	CNCCC=C1c2ccccc2CCc2ccccc21
Drug	Norgestimate	C#C[C@]1(OC(C)=O)CC[C@H]2[C@@H]3CCC4=C/C(=N/O)CC[C@@H]4[C@H]3CC[C@@]21CC
Drug	nordiazepam	OC1=Nc2ccc(Cl)cc2C(c2ccccc2)=NC1
Drug	n-nitroso-n-methylaniline (NMPhA)	CN(N=O)c1ccccc1
Drug	N-methyl-3,4methylenedioxyamphetamine	CNC(C)Cc1ccc2c(c1)OCO2
Drug	Nitrofurantoin	O=C1CN(/N=C/c2ccc([N+](=O)[O-])o2)C(=O)N1
Drug	Nitazoxanide	CC(=O)Oc1ccccc1C(=O)Nc1ncc([N+](=O)[O-])s1
Drug	nisoldipine	COC(=O)C1=C(C)NC(C)=C(C(=O)OCC(C)C)C1c1ccccc1[N+](=O)[O-]
Drug	Nirvanol	CCC1(c2ccccc2)N=C(O)N=C1O
Drug	Nintedanib	COC(=O)c1ccc2c(c1)NC(=O)/C2=C(\Nc1ccc(N(C)C(=O)CN2CCN(C)CC2)cc1)c1ccccc1
Drug	Nimodipine	COCCOC(=O)C1=C(C)NC(C)=C(C(=

		O)OC(C)C)C1c1cccc([N+](=O)[O-])c1
Drug	nifedipine	COC(=O)C1=C(C)NC(C)=C(C(=O)OC)C1c1ccccc1[N+](=O)[O-]
Drug	Nicoboxil	CCCCOCCOC(=O)c1cccnc1
Drug	Nicergoline	CO[C@]12C[C@@H](COC(=O)c3cncc(Br)c3)CN(C)[C@@H]1Cc1cn(C)c3cccc2c13
Drug	Nicardipine	COC(=O)C1=C(C)NC(C)=C(C(=O)OCCN(C)Cc2ccccc2)C1c1cccc([N+](=O)[O-])c1
Drug	Niacin	O=C(O)c1cccnc1
Drug	nevirapine	Cc1ccnc2c1N=C(O)c1cccnc1N2C1CC1
Drug	Netupitant	Cc1ccccc1-c1cc(N2CCN(C)CC2)ncc1N(C)C(=O)C(C)(C)c1cc(C(F)(F)F)cc(C(F)(F)F)c1
Drug	Netarsudil	Cc1ccc(C(=O)OCc2ccc([C@@H](CN)C(=O)Nc3ccc4cnccc4c3)cc2)c(C)c1
Drug	Nelfinavir	Cc1c(O)cccc1C(O)=N[C@@H](CSc1ccccc1)[C@H](O)CN1C[C@H]2CCCC[C@H]2C[C@H]1C(O)=NC(C)(C)C
Drug	Nelarabine	COc1nc(N)nc2c1ncn2[C@@H]1O[C@H](CO)[C@@H](O)[C@@H]1O
Drug	nefazodone	CCc1nn(CCCN2CCN(c3cccc(Cl)c3)CC2)c(=O)n1CCOc1ccccc1
Drug	Nebivolol	OC(CNCC(O)C1CCc2cc(F)ccc2O1)C1CCc2cc(F)ccc2O1
Drug	NE-100	CCCN(CCC)CCc1ccc(OC)c(OCCc2ccccc2)c1
Drug	Nateglinide	CC(C)[C@H]1CC[C@@H](CC1)C(=O)N[C@H](CC1=CC=CC=C1)C(O)=O
Drug	Naproxen	COc1ccc2cc([C@H](C)C(=O)O)ccc2c1
Drug	Naltrexone	O=C1CC[C@@]2(O)[C@H]3Cc4ccc(O)c5c4[C@@]2(CCN3CC2CC2)[C@H]1O5
Drug	Nalmefene	C=C1CC[C@@]2(O)[C@H]3Cc4ccc(O)c5c4[C@@]2(CCN3CC2CC2)[C@H]1O5
Drug	Nalidixic acid	CCn1cc(C(=O)O)c(=O)c2ccc(C)nc21
Drug	nabumetone	COc1ccc2cc(CCC(C)=O)ccc2c1
Drug	n_ac_pcbc	N[C@@H](CS/C(Cl)=C(/Cl)C(Cl)=C(Cl)Cl)C(=O)O
Drug	mycophenolic acid	COc1c(C)c2c(c(O)c1C/C=C(\C)CCC(=O)O)C(=O)OC2
Drug	Mycophenolate mofetil	COc1c(C)c2c(c(O)c1C/C=C(\C)CCC(=O)OCCN1CCOCC1)C(=O)OC2
Drug	muraglitazar	COc1ccc(OC(=O)N(CC(=O)O)Cc2ccc(OCCc3nc(-c4ccccc4)oc3C)cc2)cc1
Drug	Mupirocin	C/C(=C\C(=O)OCCCCCCCCC(=O)O

		)C[C@@H]1OC[C@H](C[C@@H]2O[C@H]2[C@@H](C)[C@H](C)O)[C@@H](O)[C@H]1O
Drug	morphine	CN1CC[C@]23c4c5ccc(O)c4OC2[C@@H](O)C=CC3C1C5
Drug	montelukast	CC(C)(O)c1ccccc1CC[C@@H](SCC1(CC(=O)O)CC1)c1cccc(/C=C/c2ccc3ccc(Cl)cc3n2)c1
Drug	Mometasone furoate	C[C@@H]1C[C@H]2[C@@H]3CCC4=CC(=O)C=C[C@]4(C)[C@@]3(Cl)[C@@H](O)C[C@]2(C)[C@@]1(OC(=O)c1ccco1)C(=O)CCl
Drug	Molsidomine	CCOC(=O)[N-]c1c[n+](N2CCOCC2)no1
Drug	Moexipril	CCOC(=O)[C@H](CCc1ccccc1)N[C@@H](C)C(=O)N1Cc2cc(OC)c(OC)cc2C[C@H]1C(=O)O
Drug	Modafinil	NC(=O)CS(=O)C(c1ccccc1)c1ccccc1
Drug	Moclobemide	O=C(NCCN1CCOCC1)c1ccc(Cl)cc1
Drug	MK-429	COc1ccc([C@H](CC(=O)O)N2CCN(CCCc3ccc4c(n3)NCCC4)C2=O)cn1
Drug	Misoprostol	CCCCC(C)(O)C/C=C/[C@H]1[C@H](O)CC(=O)[C@@H]1CCCCCCC(=O)OC
Drug	Misonidazole	COCC(O)Cn1ccnc1[N+](=O)[O-]
Drug	Mirtazapine	CN1CCN2c3ncccc3Cc3ccccc3C2C1
Drug	miocamycin	CCC(=O)O[C@H]1[C@@H](C)O[C@@H](O[C@H]2[C@@H](N(C)C)[C@@H](O)[C@@H](O[C@H]3[C@@H](CC=O)C[C@@H](C)[C@@H](OC(C)=O)/C=C/C=C/C[C@@H](C)OC(=O)C[C@@H](OC(=O)CC)[C@H]3OC)O[C@@H]2C)C[C@@]1(C)OC(C)=O
Drug	Minoxidil	N=c1nc(N2CCCCC2)cc(N)n1O
Drug	Minocycline	CN(C)c1ccc(O)c2c1C[C@H]1C[C@H]3[C@H](N(C)C)C(O)=C(C(N)=O)C(=O)[C@@]3(O)C(O)=C1C2=O
Drug	Minaprine	Cc1cc(-c2ccccc2)nnc1NCCN1CCOCC1
Drug	Milnacipran	CCN(CC)C(=O)C1(c2ccccc2)CC1CN
Drug	mifepristone	CC#C[C@]1(O)CC[C@H]2[C@@H]3CCC4=CC(=O)CCC4=C3[C@@H](c3ccc(N(C)C)cc3)C[C@@]21C
Drug	Midostaurin	CO[C@@H]1[C@H](N(C)C(=O)c2ccccc2)C[C@H]2O[C@]1(C)n1c3ccccc3c3c4c(c5c6ccccc6n2c5c31)C(=O)NC4
Drug	Midodrine	COc1ccc(OC)c(C(O)CNC(=O)CN)c1
Drug	Midazolam	Cc1ncc2n1-c1ccc(Cl)cc1C(c1ccccc1F)=NC2
Drug	Mianserine	CN1CCN2c3ccccc3Cc3ccccc3C2C1

Drug	Mexiletine	Cc1cccc(C)c1OCC(C)N
Drug	Metyrapone	CC(C)(C(=O)c1cccnc1)c1cccnc1
Drug	Metronidazole	Cc1ncc([N+](=O)[O-])n1CCO
Drug	Metoprolol	COCCc1ccc(OCC(O)CNC(C)C)cc1
Drug	Metoclopramide	CCN(CC)CCNC(=O)c1cc(Cl)c(N)cc1OC
Drug	Methylprednisolone	C[C@H]1C[C@H]2[C@@H]3CC[C@](O)(C(=O)CO)[C@@]3(C)C[C@H](O)[C@@H]2[C@@]2(C)C=CC(=O)C=C12
Drug	Methylphenobarbital	CCC1(c2ccccc2)C(=O)NC(=O)N(C)C1=O
Drug	Methylphenidate	COC(=O)C(c1ccccc1)C1CCCCN1
Drug	Methyldopa	C[C@](N)(Cc1ccc(O)c(O)c1)C(=O)O
Drug	Methyl salicylate	COC(=O)c1ccccc1O
Drug	Methyl nicotinate	COC(=O)c1cccnc1
Drug	methoxsalen	COc1c2occc2cc2ccc(=O)oc12
Drug	Methotrexate	CN(Cc1cnc2nc(N)nc(N)c2n1)c1ccc(C(=O)N[C@@H](CCC(=O)O)C(=O)O)cc1
Drug	Methocarbamol	COc1ccccc1OCC(O)COC(N)=O
Drug	Methimazole	Cn1cc[nH]c1=S
Drug	Methenamine	C1N2CN3CN1CN(C2)C3
Drug	methaqualone	Cc1ccccc1-n1c(C)nc2ccccc2c1=O
Drug	Methamphetamine	CNC(C)Cc1ccccc1
Drug	methadone	CCC(=O)C(CC(C)N(C)C)(c1ccccc1)c1ccccc1
Drug	Metamfetamine	CN[C@@H](C)Cc1ccccc1
Drug	Metamizol	c1ccccc1N2N(C)C(C)=C(C2=O)N(C)CS(=O)(=O)O
Drug	Mestranol	C#C[C@]1(O)CC[C@H]2[C@@H]3CCc4cc(OC)ccc4[C@H]3CC[C@@]21C
Drug	mesalazine	Nc1ccc(O)c(C(=O)O)c1
Drug	Mequitazine	c1ccc2c(c1)Sc1ccccc1N2CC12CCN(CC1)CC2
Drug	Mequinol	COc1ccc(O)cc1
Drug	Mequindox	CC1=C([N+](=O)C2=CC=CC=C2N1[O-])C(=O)C
Drug	Mephenytoin	CCC1(c2ccccc2)NC(=O)N(C)C1=O
Drug	meperidine	CCOC(=O)C1(c2ccccc2)CCN(C)CC1
Drug	Meloxicam	Cc1cnc(N=C(O)C2=C(O)c3ccccc3S(=O)(=O)N2C)s1
Drug	Mefloquine	OC(c1cc(C(F)(F)F)nc2c(C(F)(F)F)cccc12)C1CCCCN1
Drug	mefenamic acid	Cc1cccc(Nc2ccccc2C(=O)O)c1C
Drug	Medroxyprogesterone Acetate	CC(=O)O[C@]1(C(C)=O)CC[C@H]2[C@@H]3C[C@H](C)C4=CC(=O)CC[C@]4(C)[C@H]3CC[C@@]21C
Drug	Meclizine	Cc1cccc(CN2CCN(C(c3ccccc3)c3ccc(

		Cl)cc3)CC2)c1
Drug	Mebendazole	COC(=O)Nc1nc2ccc(C(=O)c3ccccc3)cc2[nH]1
Drug	m-Chlorophenylpiperazine	Clc1cccc(N2CCNCC2)c1
Drug	MaxiPost	COc1ccc(Cl)cc1C1(F)C(O)=Nc2cc(C(F)(F)F)ccc21
Drug	maprotiline	CNCCCC12CCC(c3ccccc31)c1ccccc12
Drug	M23OH	CCCCc1cc(N)n2ncc(O)c2n1
Drug	Lurasidone	O=C1[C@H]2[C@@H]3CC[C@@H](C3)[C@H]2C(=O)N1C[C@@H]1CCCC[C@H]1CN1CCN(c2nsc3ccccc23)CC1
Drug	Lumiracoxib	Cc1ccc(Nc2c(F)cccc2Cl)c(CC(=O)O)c1
Drug	Lumefantrine	CCCCN(CCCC)CC(O)c1cc(Cl)cc2c1-c1ccc(Cl)cc1/C2=C/c1ccc(Cl)cc1
Drug	l-trofosfamide	O=[P@@]1(N(CCCl)CCCl)OCCCN1CCCl
Drug	Loxoprofen	CC(C(=O)O)c1ccc(CC2CCCC2=O)cc1
Drug	Loxepine	CN1CCN(C2=Nc3ccccc3Oc3ccc(Cl)cc32)CC1
Drug	lovastatin	CC[C@H](C)C(=O)O[C@H]1C[C@@H](C)C=C2C=C[C@H](C)[C@H](CC[C@@H]3C[C@@H](O)CC(=O)O3)[C@H]21
Drug	LOSIGAMONE	COC1=CC(=O)OC1C(O)c1ccccc1Cl
Drug	losartan	CCCCc1nc(Cl)c(CO)n1Cc1ccc(-c2ccccc2-c2nn[nH]n2)cc1
Drug	lornoxicam	CN1C(C(O)=Nc2ccccn2)=C(O)c2sc(Cl)cc2S1(=O)=O
Drug	Lormetazepam	CN1C(=O)C(O)N=C(c2ccccc2Cl)c2cc(Cl)ccc21
Drug	Lorcaserin	C[C@H]1CNCCc2ccc(Cl)cc21
Drug	Lorazepam	OC1=Nc2ccc(Cl)cc2C(c2ccccc2Cl)=NC1O
Drug	loratadine	CCOC(=O)N1CCC(=C2c3ccc(Cl)cc3CCc3cccnc32)CC1
Drug	Lopinavir, D-valine diastereomer-	Cc1cccc(C)c1OCC(O)=N[C@@H](Cc1ccccc1)[C@@H](O)C[C@H](Cc1ccccc1)N=C(O)[C@@H](C(C)C)N1CCCN=C1O
Drug	loperamide	CN(C)C(=O)C(CCN1CCC(O)(c2ccc(Cl)cc2)CC1)(c1ccccc1)c1ccccc1
Drug	Lonidamine	O=C(O)c1nn(Cc2ccc(Cl)cc2Cl)c2ccccc12
Drug	Lofexidine	CC(Oc1c(Cl)cccc1Cl)C1=NCCN1
Drug	Lobeglitazone	COc1ccc(Oc2cc(N(C)CCOc3ccc(CC4SC(=O)NC4=O)cc3)ncn2)cc1
Drug	lisofylline	C[C@@H](O)CCCCn1c(=O)c2c(ncn2C)n(C)c1=O

Drug	Lisdexamfetamine	C[C@@H](Cc1ccccc1)NC(=O)[C@@H](N)CCCCN
Drug	linomide	CN(C(=O)c1c(O)c2ccccc2n(C)c1=O)c1ccccc1
Drug	Linezolid	CC(=O)NC[C@H]1CN(c2ccc(N3CCOCC3)c(F)c2)C(=O)O1
Drug	Linagliptin	CC#CCn1c(N2CCC[C@@H](N)C2)nc2c1c(=O)n(Cc1nc(C)c3ccccc3n1)c(=O)n2C
Drug	l-ifosfamide	O=[P@]1(NCCCl)OCCCN1CCCl
Drug	lidocaine	CCN(CC)CC(=O)Nc1c(C)cccc1C
Drug	Licofelone	CC1(C)Cc2c(-c3ccccc3)c(-c3ccc(Cl)cc3)c(CC(=O)O)n2C1
Drug	Levothyroxine	N[C@@H](Cc1cc(I)c(Oc2cc(I)c(O)c(I)c2)c(I)c1)C(=O)O
Drug	Levosimendan	C[C@@H]1CC(=O)NN=C1c1ccc(NN=C(C#N)C#N)cc1
Drug	Levoleucovorin	Nc1nc(=O)c2c([nH]1)NC[C@H](CNc1ccc(C(=O)N[C@@H](CCC(=O)O)C(=O)O)cc1)N2C=O
Drug	Levofloxacin	C[C@H]1COc2c(N3CCN(C)CC3)c(F)cc3c(=O)c(C(=O)O)cn1c23
Drug	Levodopa	N[C@@H](Cc1ccc(O)c(O)c1)C(=O)O
Drug	Levocetirizine	O=C(O)COCCN1CCN([C@H](c2ccccc2)c2ccc(Cl)cc2)CC1
Drug	Levocarnitine	C[N+](C)(C)C[C@H](O)CC(=O)[O-]
Drug	levobupivacaine	CCCCN1CCCC[C@H]1C(O)=Nc1c(C)cccc1C
Drug	Levetiracetam	CC[C@@H](C(N)=O)N1CCCC1=O
Drug	Levamlodipine	CCOC(=O)C1=C(COCCN)NC(C)=C(C(=O)OC)[C@@H]1c1ccccc1Cl
Drug	Levacetylmethadol	CC[C@H](OC(C)=O)C(C[C@H](C)N(C)C)(c1ccccc1)c1ccccc1
Drug	Leucovorin	Nc1nc2c(c(=O)[nH]1)N(C=O)C(CNc1ccc(C(=O)N[C@@H](CCC(=O)O)C(=O)O)cc1)CN2
Drug	Letrozole	N#Cc1ccc(C(c2ccc(C#N)cc2)n2cncn2)cc1
Drug	Letermovir	COc1cccc(N2CCN(C3=Nc4c(F)cccc4[C@H](CC(=O)O)N3c3cc(C(F)(F)F)ccc3OC)CC2)c1
Drug	leflunomide	Cc1oncc1C(=O)Nc1ccc(C(F)(F)F)cc1
Drug	l-cyclophosphamide	O=[P@]1(N(CCCl)CCCl)NCCCO1
Drug	lauric acid	CCCCCCCCCCCC(=O)O
Drug	Latanoprostene bunod	O=C(CCC/C=C\C[C@H]1[C@@H](O)C[C@@H](O)[C@@H]1CC[C@@H](O)CCc1ccccc1)OCCCCO[N+](=O)[O-]
Drug	Latanoprost	CC(C)OC(=O)CCC/C=C\C[C@H]1[C@@H](O)C[C@@H](O)[C@@H]1CC[C@@H](O)CCc1ccccc1

Drug	Lasofoxifene	Oc1ccc2c(c1)CC[C@H](c1ccccc1)[C@@H]2c1ccc(OCCN2CCCC2)cc1
Drug	laquinimod	CCN(C(=O)c1c(O)c2c(Cl)cccc2n(C)c1=O)c1ccccc1
Drug	Lansoprazole	Cc1cc(OCC(F)(F)F)c(C)c(C[S@@](=O)c2nc3ccccc3[nH]2)n1
Drug	Lamotrigine	Nc1nnc(-c2cccc(Cl)c2Cl)c(N)n1
Drug	Lamivudine	Nc1ccn([C@@H]2CS[C@H](CO)O2)c(=O)n1
Drug	Lactulose	OC[C@H]1O[C@](O)(CO)[C@@H](O)[C@@H]1O[C@@H]1O[C@H](CO)[C@H](O)[C@H](O)[C@H]1O
Drug	Lacosamide	COC[C@@H](NC(C)=O)C(=O)NCc1ccccc1
Drug	Labetalol	CC(CCc1ccccc1)NCC(O)c1ccc(O)c(C(N)=O)c1
Drug	L-775606	Fc1cccc(CCN2CCN(CCCc3c[nH]c4ccc(-n5cnnc5)cc34)CC2)c1
Drug	l_775606	Fc1cccc(CCN2CCN(CCCc3cc4cc(-n5cnnc5)ccc4[nH]3)CC2)c1
Drug	kr-60436	COc1ccc(N2CCc3c2c2cccc(OC(F)(F)F)c2[nH]c3=NCCO)c(C)c1
Drug	KR-33028	N#Cc1cccc2sc(C(=O)NC(=N)N)cc12
Drug	Ketotifen	CN1CCC(=C2c3ccccc3CC(=O)c3sccc32)CC1
Drug	Ketorolac	O=C(c1ccccc1)c1ccc2n1CCC2C(=O)O
Drug	Ketoprofen	CC(C(=O)O)c1cccc(C(=O)c2ccccc2)c1
Drug	ketobemidone	CCC(=O)C1(c2cccc(O)c2)CCN(C)CC1
Drug	Ketamine	CNC1(c2ccccc2Cl)CCCCC1=O
Drug	k-11777	CN1CCN(C(O)=N[C@@H](Cc2ccccc2)C(O)=N[C@H](/C=C/S(=O)(=O)c2ccccc2)CCc2ccccc2)CC1
Drug	JNJ-10198409	COc1cc2c(cc1OC)-c1[nH][nH]c(=Nc3cccc(F)c3)c1C2
Drug	ivermectin	CO[C@H]1C[C@H](O[C@H]2[C@H](C)O[C@@H](O[C@@H]3/C(C)=C\C[C@@H]4CC[C@@H](C[C@]5(CC[C@H](C(C)C)CO5)O4)OC(=O)[C@@H]4C=C(C)[C@@H](O)[C@@H]5CC/C(=C\C=C/[C@@H]3C)[C@@]54O)C[C@@H]2OC)O[C@@H](C)[C@@H]1O
Drug	Ivacaftor	CC(C)(C)c1cc(C(C)(C)C)c(NC(=O)c2c[nH]c3ccccc3c2=O)cc1O
Drug	Itraconazole	CCC(C)n1ncn(-c2ccc(N3CCN(c4ccc(OC[C@H]5CO[C@](Cn6cncn6)(c6ccc(Cl)cc6Cl)O5)cc4)CC3)cc2)c1=O
Drug	Istradefylline	CCn1c(=O)c2c(nc(/C=C/c3ccc(OC)c(

		OC)c3)n2C)n(CC)c1=O
Drug	isotretinoin	CC1=C(/C=C/C(C)=C/C=C/C(C)=C\C(=O)O)C(C)(C)CCC1
Drug	Isosorbide mononitrate	O=[N+]([O-])O[C@@H]1CO[C@@H]2[C@@H](O)CO[C@H]12
Drug	isoquinoline	c1ccc2cnccc2c1
Drug	Isopropyl myristate	CCCCCCCCCCCCCC(=O)OC(C)C
Drug	isoniazid	NNC(=O)c1ccncc1
Drug	Isocarboxazid	Cc1cc(C(=O)NNCc2ccccc2)no1
Drug	irinotecan	CCc1c2c(nc3ccc(OC(=O)N4CCC(N5CCCCC5)CC4)cc13)-c1cc3c(c(=O)n1C2)COC(=O)[C@]3(O)CC
Drug	Irbesartan	CCCCC1=NC2(CCCC2)C(=O)N1Cc1ccc(-c2ccccc2-c2nn[nH]n2)cc1
Drug	ipriflavone	CC(C)Oc1ccc2c(=O)c(-c3ccccc3)coc2c1
Drug	Ipratropium	CC(C)[N+]1(C)[C@H]2CC[C@@H]1C[C@H](OC(=O)C(CO)c1ccccc1)C2
Drug	Inositol	O[C@H]1[C@H](O)[C@@H](O)[C@H](O)[C@@H](O)[C@H]1O
Drug	Indomethacin	COc1ccc2c(c1)c(CC(=O)O)c(C)n2C(=O)c1ccc(Cl)cc1
Drug	indiplon	CC(=O)N(C)c1cccc(-c2ccnc3c(C(=O)c4cccs4)cnn23)c1
Drug	indinavir	CC(C)(C)N=C(O)[C@@H]1CN(Cc2ccccnc2)CCN1C[C@@H](O)C[C@@H](Cc1ccccc1)C(O)=N[C@H]1c2ccccc2C[C@H]1O
Drug	Indapamide	CC1Cc2ccccc2N1NC(=O)c1ccc(Cl)c(S(N)(=O)=O)c1
Drug	Indacaterol	CCc1cc2c(cc1CC)CC(NC[C@H](O)c1ccc(O)c3nc(O)ccc13)C2
Drug	imipramine	CN(C)CCCN1c2ccccc2CCc2ccccc21
Drug	imatinib	Cc1ccc(NC(=O)c2ccc(CN3CCN(C)CC3)cc2)cc1Nc1nccc(-c2cccnc2)n1
Drug	iloperidone	COc1cc(C(C)=O)ccc1OCCCN1CCC(c2noc3cc(F)ccc23)CC1
Drug	Ifosfamide	O=P1(NCCCl)OCCCN1CCCl
Drug	Idelalisib	CC[C@H](Nc1ncnc2[nH]cnc12)c1nc2ccccc(F)c2c(=O)n1-c1ccccc1
Drug	Idebenone	COC1=C(OC)C(=O)C(CCCCCCCCCCCCCO)=C(C)C1=O
Drug	Icosapent ethyl	CC/C=C\C/C=C\C/C=C\C/C=C\C/C=C\CCCC(=O)OCC
Drug	Ibrutinib	C=CC(=O)N1CCC[C@@H](n2nc(-c3ccc(Oc4ccccc4)cc3)c3c(N)ncnc32)C1
Drug	Hydroxyzine	OCCOCCN1CCN(C(c2ccccc2)c2ccc(Cl)cc2)CC1
Drug	Hydroxyoxid acid	CC1(C)[C@@H]2CC[C@@]3(C)[C

		@@H](C2)[C@@H](C(=O)O)CC[C@@]13O
Drug	Hydroxycitronellal	CC(CC=O)CCCC(C)(C)O
Drug	Hydroxychloroquine	CCN(CCO)CCCC(C)Nc1ccnc2cc(Cl)ccc12
Drug	Hydromorphone	CN1CC[C@]23c4c5ccc(O)c4O[C@H]2C(=O)CC[C@H]3[C@H]1C5
Drug	hydrocodone	COc1ccc2c3c1O[C@H]1C(=O)CC[C@H]4[C@@H](C2)N(C)CC[C@]314
Drug	Hydralazine	NNc1nncc2ccccc12
Drug	Hexobarbital	CN1C(=O)NC(=O)C(C)(C2=CCCCC2)C1=O
Drug	Hesperetin	COc1ccc([C@@H]2CC(=O)c3c(O)cc(O)cc3O2)cc1O
Drug	harmine	COc1ccc2c(c1)[nH]c1c(C)nccc12
Drug	harmaline	COc1ccc2c3c([nH]c2c1)C(C)=NCC3
Drug	Halothane	FC(F)(F)C(Cl)Br
Drug	haloperidol	O=C(CCCN1CCC(O)(c2ccc(Cl)cc2)CC1)c1ccc(F)cc1
Drug	Halofantrine	CCCCN(CCCC)CCC(O)c1cc2c(Cl)cc(Cl)cc2c2cc(C(F)(F)F)ccc12
Drug	guanabenz	N=C(N)N/N=C/c1c(Cl)cccc1Cl
Drug	Guaifenesin	COc1ccccc1OCC(O)CO
Drug	gts-21	COc1ccc(/C=C2\CCCN=C2c2cccnc2)c(OC)c1
Drug	Griseofulvin	COC1=CC(=O)C[C@@H](C)[C@]12Oc1c(Cl)c(OC)cc(OC)c1C2=O
Drug	granisetron	CN1[C@H]2CCC[C@@H]1CC(NC(=O)c1nn(C)c3ccccc13)C2
Drug	Glycyrrhizic acid	CC1(C)[C@@H](O[C@H]2O[C@H](C(=O)O)[C@@H](O)[C@H](O)[C@H]2O[C@@H]2O[C@H](C(=O)O)[C@@H](O)[C@H](O)[C@H]2O)CC[C@@]2(C)[C@H]1CC[C@]1(C)[C@@H]2C(=O)C=C2[C@@H]3C[C@@](C)(C(=O)O)CC[C@]3(C)CC[C@]21C
Drug	Glyburide	COc1ccc(Cl)cc1C(O)=NCCc1ccc(S(=O)(=O)NC(O)=NC2CCCCC2)cc1
Drug	Glipizide	Cc1cnc(C(O)=NCCc2ccc(S(=O)(=O)NC(O)=NC3CCCCC3)cc2)cn1
Drug	Glimepiride	CCC1=C(C)CN(C(O)=NCCc2ccc(S(=O)(=O)NC(O)=NC3CCC(C)CC3)cc2)C1=O
Drug	Gliclazide	Cc1ccc(S(=O)(=O)NC(O)=NN2CC3CCCC3C2)cc1
Drug	Glasdegib	CN1CC[C@@H](NC(=O)Nc2ccc(C#N)cc2)C[C@@H]1c1nc2ccccc2[nH]1
Drug	gepirone	CC1(C)CC(=O)N(CCCCN2CCN(c3nccccn3)CC2)C(=O)C1
Drug	Gemifloxacin	CO/N=C1\CN(c2nc3c(cc2F)c(=O)c(C(=O)O)cn3C2CC2)CC1CN
Drug	Gemfibrozil	Cc1ccc(C)c(OCCCC(C)(C)C(=O)O)c1

Drug	Gemcitabine Monophosphate	Nc1ccn([C@@H]2O[C@H](COP(=O)(O)O)[C@@H](O)C2(F)F)c(=O)n1
Drug	Gemcitabine diphosphate	Nc1ccn([C@@H]2O[C@H](COP(=O)(O)OP(=O)(O)O)[C@@H](O)C2(F)F)c(=O)n1
Drug	Gemcitabine	Nc1ccn([C@@H]2O[C@H](CO)[C@@H](O)C2(F)F)c(=O)n1
Drug	gefitinib	COc1cc2ncnc(Nc3ccc(F)c(Cl)c3)c2cc1OCCCN1CCOCC1
Drug	Gavestinel	O=C(O)c1[nH]c2cc(Cl)cc(Cl)c2c1/C=C/C(O)=Nc1ccccc1
Drug	Gatifloxacin	COc1c(N2CCNC(C)C2)c(F)cc2c(=O)c(C(=O)O)cn(C3CC3)c12
Drug	gamma-Hydroxybutyric acid	O=C(O)CCCO
Drug	Galantamine	COc1ccc2c3c1O[C@H]1C[C@@H](O)C=C[C@@]31CCN(C)C2
Drug	galangin	O=c1c(O)c(-c2ccccc2)oc2cc(O)cc(O)c12
Drug	FYX-051	N#Cc1cc(-c2nc(-c3ccncc3)n[nH]2)ccn1
Drug	Fusidic acid	CC(=O)O[C@H]1C[C@@]2(C)[C@@H](C[C@@H](O)[C@H]3[C@@]4(C)CC[C@@H](O)[C@@H](C)[C@@H]4CC[C@@]32C)/C1=C(\CCC=C(C)C)C(=O)O
Drug	Furosemide	NS(=O)(=O)c1cc(C(=O)O)c(NCc2ccco2)cc1Cl
Drug	Furazolidone	O=C1OCCN1N=Cc1ccc([N+](=O)[O-])o1
Drug	furafylline	Cc1nc2c(=O)n(C)c(=O)n(Cc3ccco3)c2[nH]1
Drug	Frovatriptan	CN[C@@H]1CCc2[nH]c3ccc(C(N)=O)cc3c2C1
Drug	Fostamatinib	COc1cc(Nc2ncc(F)c(Nc3ccc4c(n3)N(COP(=O)(O)O)C(=O)C(C)(C)O4)n2)cc(OC)c1OC
Drug	Fospropofol	CC(C)c1cccc(C(C)C)c1OCOP(=O)(O)O
Drug	Fosphenytoin	O=C1NC(c2ccccc2)(c2ccccc2)C(=O)N1COP(=O)(O)O
Drug	Fosinopril	CCC(=O)O[C@@H](O[P@](=O)(CCCCc1ccccc1)CC(=O)N1C[C@H](C2CCCCC2)C[C@H]1C(=O)O)C(C)C
Drug	Formoterol	COc1ccc(CC(C)NCC(O)c2ccc(O)c(NC=O)c2)cc1
Drug	Formestane	C[C@]12CC[C@H]3[C@@H](CCC4=C(O)C(=O)CC[C@@]43C)[C@@H]1CCC2=O
Drug	Fomepizole	Cc1cn[nH]c1
Drug	fluvoxamine	COCCCC/C(=N\OCCN)c1ccc(C(F)(F)F)cc1
Drug	fluvastatin	CC(C)n1c(/C=C/[C@H](O)C[C@H](

		O)CC(=O)O)c(-c2ccc(F)cc2)c2ccccc21
Drug	Fluticasone propionate	CCC(=O)O[C@]1(C(=O)SCF)[C@H](C)C[C@H]2[C@@H]3C[C@H](F)C4=CC(=O)C=C[C@]4(C)[C@@]3(F)[C@@H](O)C[C@@]21C
Drug	Fluticasone furoate	C[C@@H]1C[C@H]2[C@@H]3C[C@H](F)C4=CC(=O)C=C[C@]4(C)[C@@]3(F)[C@@H](O)C[C@]2(C)[C@@]1(OC(=O)c1ccco1)C(=O)SCF
Drug	flutamide	CC(C)C(O)=Nc1ccc([N+](=O)[O-])c(C(F)(F)F)c1
Drug	Flurbiprofen	CC(C(=O)O)c1ccc(-c2ccccc2)c(F)c1
Drug	Flurazepam	CCN(CC)CCN1C(=O)CN=C(c2ccccc2F)c2cc(Cl)ccc21
Drug	fluphenazine	OCCN1CCN(CCCN2c3ccccc3Sc3ccc(C(F)(F)F)cc32)CC1
Drug	Fluoxetine	CNCCC(Oc1ccc(C(F)(F)F)cc1)c1ccccc1
Drug	Fluorodopa (18F)	N[C@@H](Cc1cc(O)c(O)cc1[18F])C(=O)O
Drug	flunitrazepam	CN1C(=O)CN=C(c2ccccc2F)c2cc([N+](=O)[O-])ccc21
Drug	flunarizine	Fc1ccc(C(c2ccc(F)cc2)N2CCN(C/C=C/c3ccccc3)CC2)cc1
Drug	Fludrocortisone	C[C@]12C[C@H](O)[C@@]3(F)[C@@H](CCC4=CC(=O)CC[C@@]43C)[C@@H]1CC[C@]2(O)C(=O)CO
Drug	flucloxacillin	Cc1onc(-c2c(F)cccc2Cl)c1C(O)=N[C@@H]1C(=O)N2[C@@H](C(=O)O)C(C)(C)S[C@H]12
Drug	Floxuridine	O=c1[nH]c(=O)n([C@H]2C[C@H](O)[C@@H](CO)O2)cc1F
Drug	flosequinan	Cn1cc(S(C)=O)c(=O)c2ccc(F)cc21
Drug	Florbetapir (18F)	CNc1ccc(/C=C/c2ccc(OCCOCCOCC[18F])nc2)cc1
Drug	Flecainide	OC(=NC[C@@H]1CCCCN1)c1cc(OCC(F)(F)F)ccc1OCC(F)(F)F
Drug	fk_506	C=CC[C@@H]1/C=C(/C)C[C@H](C)C[C@H](OC)[C@H]2O[C@@](O)(C(=O)C(=O)N3CCCC[C@H]3C(=O)O[C@H](/C(C)=C/[C@@H]3CC[C@@H](O)[C@H](OC)C3)[C@H](C)[C@H](O)CC1=O)[C@H](C)C[C@@H]2OC
Drug	Finasteride	CC(C)(C)NC(=O)[C@H]1CC[C@H]2[C@@H]3CC[C@H]4NC(=O)C=C[C@]4(C)[C@H]3CC[C@]12C
Drug	Fidaxomicin	CCc1c(Cl)c(O)c(Cl)c(O)c1C(=O)O[C@H]1[C@H](O)[C@H](OC)[C@H](OC/C2=C\C=C\C[C@H](O)/C(C)=C/[

		C@H](CC)[C@@H](O[C@@H]3OC(C)(C)[C@@H](OC(=O)C(C)C)[C@H](O)[C@@H]3O)/C(C)=C/C(C)=C/C[C@@H]([C@@H](C)O)OC2=O)O[C@@H]1C
Drug	Fexofenadine	CC(C)(C(=O)O)c1ccc(C(O)CCCN2CCC(C(O)(c3ccccc3)c3ccccc3)CC2)cc1
Drug	fentanyl	CCC(=O)N(c1ccccc1)C1CCN(CCc2ccccc2)CC1
Drug	Fenretinide	CC1=C(/C=C/C(C)=C/C=C/C(C)=C/C(=O)Nc2ccc(O)cc2)C(C)(C)CCC1
Drug	Fenproporex	C[C@@H](Cc1ccccc1)NCCC#N
Drug	Fenoprofen	CC(C(=O)O)c1cccc(Oc2ccccc2)c1
Drug	Fenofibrate	CC(C)OC(=O)C(C)(C)Oc1ccc(C(=O)c2ccc(Cl)cc2)cc1
Drug	Fenfluramine	CCNC(C)Cc1cccc(C(F)(F)F)c1
Drug	Felodipine	CCOC(=O)C1=C(C)NC(C)=C(C(=O)OC)C1c1cccc(Cl)c1Cl
Drug	Felbamate	NC(=O)OCC(COC(N)=O)c1ccccc1
Drug	Febuxostat	Cc1nc(-c2ccc(OCC(C)C)c(C#N)c2)sc1C(=O)O
Drug	Favipiravir	NC(=O)c1nc(F)cnc1O
Drug	Famotidine	NC(N)=Nc1nc(CSCCC(N)=NS(N)(=O)=O)cs1
Drug	ezlopitant	COc1ccc(C(C)C)cc1CN[C@H]1C2CCN(CC2)[C@H]1C(c1ccccc1)c1ccccc1
Drug	Ezetimibe	O=C1[C@H](CC[C@H](O)c2ccc(F)cc2)[C@@H](c2ccc(O)cc2)N1c1ccc(F)cc1
Drug	everolimus	CO[C@@H]1C[C@@H]2CC[C@@H](C)[C@@](O)(O2)C(=O)C(=O)N2CCCC[C@H]2C(=O)O[C@H]([C@H](C)C[C@@H]2CC[C@@H](OCCO)[C@H](OC)C2)CC(=O)[C@H](C)/C=C(/C)[C@@H](O)[C@@H](OC)C(=O)[C@H](C)C[C@H](C)/C=C\C=C\C=C\1C
Drug	Eucalyptol	CC1(C)O[C@]2(C)CC[C@H]1CC2
Drug	etoricoxib	Cc1ccc(-c2ncc(Cl)cc2-c2ccc(S(C)(=O)=O)cc2)cn1
Drug	etoposide	COc1cc([C@@H]2c3cc4c(cc3[C@H](O[C@H]3O[C@H]5CO[C@@H](C)O[C@@H]5[C@@H](O)[C@H]3O)[C@H]3COC(=O)[C@H]23)OCO4)cc(OC)c1O
Drug	etoperidone	CCc1nn(CCCN2CCN(c3cccc(Cl)c3)CC2)c(=O)n1CC
Drug	Etodolac	CCc1cccc2c3c([nH]c12)C(CC)(CC(=O)O)OCC3
Drug	etizolam	CCc1cc2c(s1)-n1c(C)nnc1CN=C2c1ccccc1Cl

Drug	ethylmorphine	CCOc1ccc2c3c1O[C@H]1[C@@H](O)C=C[C@H]4[C@@H](C2)N(C)CC[C@]314
Drug	Ethotoin	CCN1C(=O)NC(c2ccccc2)C1=O
Drug	Ethionamide	CCc1cc(C(N)=S)ccn1
Drug	ethinylestradiol	C#C[C@]1(O)CC[C@H]2[C@@H]3CCc4cc(O)ccc4[C@H]3CC[C@@]21C
Drug	ethanol	CCO
Drug	eszopiclone	CN1CCN(C(=O)O[C@H]2c3nccnc3C(=O)N2c2ccc(Cl)cn2)CC1
Drug	estazolam	Clc1ccc2c(c1)C(c1ccccc1)=NCc1nncn1-2
Drug	Esomeprazole	COc1ccc2nc([S@@](=O)Cc3ncc(C)c(OC)c3C)[nH]c2c1
Drug	Esmolol	COC(=O)CCc1ccc(OCC(O)CNC(C)C)cc1
Drug	Eslicarbazepine acetate	CC(=O)O[C@H]1Cc2ccccc2N(C(N)=O)c2ccccc21
Drug	erythromycin	CC[C@H]1OC[C@H](C)[C@@H](O[C@H]2C[C@@](C)(OC)[C@@H](O)[C@H](C)O2)[C@@H](C)[C@@H](O[C@@H]2O[C@H](C)C[C@H](N(C)C)[C@H]2O)[C@](C)(O)C[C@@H](C)C[C@H](C)[C@H](O)[C@]1(C)O
Drug	Ertugliflozin	CCOc1ccc(Cc2cc([C@]34OC[C@](CO)(O3)[C@@H](O)[C@H](O)[C@H]4O)ccc2Cl)cc1
Drug	Erlotinib	C#Cc1cccc(Nc2ncnc3cc(OCCOC)c(OCCOC)cc23)c1
Drug	Epoprostenol	CCCCC[C@H](O)/C=C/[C@H]1[C@H](O)C[C@@H]2O/C(=C\CCCC(=O)O)C[C@@H]21
Drug	eplerenone	COC(=O)[C@@H]1CC2=CC(=O)CC[C@]2(C)[C@@]23O[C@@H]2C[C@@]2(C)[C@@H](CC[C@@]24CCC(=O)O4)[C@H]13
Drug	Epitestosterone	C[C@]12CC[C@H]3[C@@H](CCC4=CC(=O)CC[C@@]43C)[C@@H]1CC[C@H]2O
Drug	Epinephrine	CNC[C@H](O)c1ccc(O)c(O)c1
Drug	Ephedrine	CN[C@@H](C)[C@H](O)c1ccccc1
Drug	Enzacamene	Cc1ccc(/C=C2/C(=O)C3(C)CCC2C3(C)C)cc1
Drug	Entrectinib	CN1CCN(c2ccc(C(=O)Nc3n[nH]c4ccc(Cc5cc(F)cc(F)c5)cc34)c(NC3CCOCC3)c2)CC1
Drug	Entacapone	CCN(CC)C(=O)/C(C#N)=C/c1cc(O)c(O)c([N+](=O)[O-])c1
Drug	Enflurane	FC(F)OC(F)(F)C(F)Cl
Drug	Endoxifen	CC/C(=C(\c1ccc(O)cc1)c1ccc(OCCN

		C)cc1)c1ccccc1
Drug	Encainide	COc1ccc(C(=O)Nc2ccccc2CCC2CCCCN2C)cc1
Drug	Enalapril	CCOC(=O)[C@H](CCc1ccccc1)N[C@@H](C)C(=O)N1CCC[C@H]1C(=O)O
Drug	Emtricitabine	Nc1nc(=O)n([C@@H]2CS[C@H](CO)O2)cc1F
Drug	Empagliflozin	OC[C@H]1O[C@@H](c2ccc(Cl)c(Cc3ccc(O[C@H]4CCOC4)cc3)c2)[C@H](O)[C@@H](O)[C@@H]1O
Drug	Emedastine	CCOCCn1c(N2CCCN(C)CC2)nc2ccccc21
Drug	eletriptan	CN1CCC[C@@H]1C[C@@H]1CNc2ccc(CCS(=O)(=O)c3ccccc3)cc21
Drug	EIDD-2801	CC(C)C(=O)OC[C@H]1O[C@@H](n2cc/c(=N/O)[nH]c2=O)[C@H](O)[C@@H]1O
Drug	efavirenz	OC1=Nc2ccc(Cl)cc2[C@@](C#CC2CC2)(C(F)(F)F)O1
Drug	Edaravone	Cc1cc(O)n(-c2ccccc2)n1
Drug	ecabapide	CN=C(O)c1cccc(NCC(O)=NCCc2ccc(OC)c(OC)c2)c1
Drug	ebastine	CC(C)(C)c1ccc(C(=O)CCCN2CCC(OC(c3ccccc3)c3ccccc3)CC2)cc1
Drug	Duloxetine	CNCC[C@H](Oc1cccc2ccccc12)c1cccs1
Drug	Droxicam	CN1c2c(oc(=O)n(-c3ccccn3)c2=O)-c2ccccc2S1(=O)=O
Drug	Drotaverine	CCOc1ccc(/C=C2\NCCc3cc(OCC)c(OCC)cc32)cc1OCC
Drug	Drospirenone	C[C@]12CC[C@H]3[C@H]([C@@H]1[C@@H]1C[C@@H]1[C@@]21CCC(=O)O1)[C@H]1C[C@H]1C1=CC(=O)CC[C@@]13C
Drug	Dronedarone	CCCCc1oc2ccc(NS(C)(=O)=O)cc2c1C(=O)c1ccc(OCCCN(CCCC)CCCC)cc1
Drug	Droloxifene	CC/C(=C(/c1ccc(OCCN(C)C)cc1)c1ccccc(O)c1)c1ccccc1
Drug	DRF_4367	COc1ccc(-c2cc(C(F)(F)F)nn2-c2ccc(S(N)(=O)=O)c(CO)c2)cc1
Drug	Doxorubicin	COc1cccc2c1C(=O)c1c(O)c3c(c(O)c1C2=O)C[C@@](O)(C(=O)CO)C[C@@H]3O[C@H]1C[C@H](N)[C@H](O)[C@H](C)O1
Drug	Doxepin	CN(C)CCC=C1c2ccccc2COc2ccccc21
Drug	Dosulepin	CN(C)CC/C=C1\c2ccccc2CSc2ccccc21
Drug	Dorzolamide	CCN[C@H]1C[C@H](C)S(=O)(=O)c2sc(S(N)(=O)=O)cc21
Drug	Dopamine	NCCc1ccc(O)c(O)c1

Drug	donepezil	COc1cc2c(cc1OC)C(=O)C(CC1CCN(Cc3ccccc3)CC1)C2
Drug	domperidone	Oc1nc2ccccc2n1CCCN1CCC(n2c(O)nc3cc(Cl)ccc32)CC1
Drug	domitroban	O=C(O)CCC/C=C\C[C@@H]1[C@@H](NS(=O)(=O)c2ccccc2)[C@H]2CC[C@@H]1C2
Drug	Dolutegravir	C[C@@H]1CCO[C@H]2Cn3cc(C(=O)NCc4ccc(F)cc4F)c(=O)c(O)c3C(=O)N21
Drug	Dolasetron	O=C(O[C@H]1C[C@@H]2C[C@@H]3C[C@H](C1)N2CC3=O)c1c[nH]c2ccccc12
Drug	Doconexent	CC/C=C\C/C=C\C/C=C\C/C=C\C/C=C\C/C=C\CCC(=O)O
Drug	docetaxel	CC(=O)O[C@@]12CO[C@@H]1C[C@H](O)[C@@]1(C)C(=O)[C@H](O)C3=C(C)[C@@H](OC(=O)[C@H](O)[C@@H](N=C(O)OC(C)(C)C)c4ccccc4)C[C@@](O)([C@@H](OC(=O)c4ccccc4)[C@@H]12)C3(C)C
Drug	DMXAA	Cc1ccc2c(=O)c3cccc(CC(=O)O)c3oc2c1C
Drug	Disopyramide	CC(C)N(CCC(C(=N)O)(c1ccccc1)c1ccccn1)C(C)C
Drug	disopiramide	CC(C)N(CC[C@](C(=N)O)(c1ccccc1)c1ccccn1)C(C)C
Drug	Diroximel fumarate	COC(=O)/C=C/C(=O)OCCN1C(=O)CCC1=O
Drug	diphenhydramine	CN(C)CCOC(c1ccccc1)c1ccccc1
Drug	Dimethylamphetamine	CC(CC1=CC=CC=C1)N(C)C
Drug	Dimethyl sulfoxide	CS(C)=O
Drug	diltiazem	COc1ccc([C@@H]2Sc3ccccc3N(CCN(C)C)C(=O)[C@@H]2OC(C)=O)cc1
Drug	Dihydromorphine	CN1CC[C@]23c4c5ccc(O)c4O[C@H]2[C@@H](O)CC[C@H]3[C@H]1C5
Drug	dihydrocodeine	COc1ccc2c3c1O[C@H]1[C@@H](O)CC[C@H]4[C@@H](C2)N(C)CC[C@]314
Drug	Digoxin	C[C@H]1O[C@@H](O[C@H]2[C@@H](O)C[C@H](O[C@H]3[C@@H](O)C[C@H](O[C@H]4CC[C@]5(C)[C@H]6C[C@@H](O)[C@]7(C)[C@@H](C8=CC(=O)OC8)CC[C@]7(O)[C@@H]6CC[C@@H]5C4)O[C@@H]3C)O[C@@H]2C)C[C@H](O)[C@@H]1O
Drug	digitoxin	C[C@H]1O[C@H](O[C@H]2[C@@H](O)C[C@H](O[C@H]3[C@@H](O)C[C@H](O[C@H]4CC[C@]5(C)[C@H]6CC[C@]7(C)[C@@H](C8=CC(=O)OC8)CC[C@]7(O)[C@@H]6CC[

		C@@H]5C4)O[C@@H]3C)O[C@@H]2C)C[C@H](O)[C@@H]1O
Drug	Difluprednate	CCCC(=O)O[C@]1(C(=O)COC(C)=O)CC[C@H]2[C@@H]3C[C@H](F)C4=CC(=O)C=C[C@]4(C)[C@@]3(F)[C@@H](O)C[C@@]21C
Drug	Didemethylimipramine	NCCCN1c2ccccc2CCc2ccccc21
Drug	Didanosine	O=c1nc[nH]c2c1ncn2[C@H]1CC[C@@H](CO)O1
Drug	Diclofenac	O=C(O)Cc1ccccc1Nc1c(Cl)cccc1Cl
Drug	Dichlorobenzyl alcohol	OCc1ccc(Cl)cc1Cl
Drug	diazepam	CN1C(=O)CN=C(c2ccccc2)c2cc(Cl)ccc21
Drug	Diamorphine	CC(=O)Oc1ccc2c3c1O[C@H]1[C@@H](OC(C)=O)C=C[C@H]4[C@@H](C2)N(C)CC[C@]314
Drug	Diacerein	CC(=O)Oc1cccc2c1C(=O)c1c(OC(C)=O)cc(C(=O)O)cc1C2=O
Drug	Dezocine	C[C@@]12CCCCC[C@@H](Cc3ccc(O)cc31)[C@@H]2N
Drug	dexverapamil	COc1ccc(CCN(C)CCC[C@](C#N)(c2ccc(OC)c(OC)c2)C(C)C)cc1OC
Drug	Dextrothyroxine	N[C@H](Cc1cc(I)c(Oc2cc(I)c(O)c(I)c2)c(I)c1)C(=O)O
Drug	Dextrorphan	CN1CC[C@@]23CCCC[C@@H]2[C@@H]1Cc1ccc(O)cc13
Drug	dextropropoxyphene	CCC(=O)O[C@](Cc1ccccc1)(c1ccccc1)[C@H](C)CN(C)C
Drug	dextromethorphan	COc1ccc2c(c1)[C@]13CCCC[C@@H]1[C@H](C2)N(C)CC3
Drug	Dextroamphetamine	C[C@H](N)Cc1ccccc1
Drug	dexpropranolol (propranolol)	CC(C)NC[C@@H](O)COc1cccc2ccccc12
Drug	Dexpanthenol	CC(C)(CO)[C@@H](O)C(=O)NCCCO
Drug	Dexmethylphenidate	COC(=O)[C@H](c1ccccc1)[C@H]1CCCCN1
Drug	Dexmedetomidine	Cc1cccc([C@H](C)c2cnc[nH]2)c1C
Drug	dexloxiglumide	CCCCCN(CCCOC)C(=O)[C@@H](CCC(=O)O)N=C(O)c1ccc(Cl)c(Cl)c1
Drug	Dexlansoprazole	Cc1c(OCC(F)(F)F)ccnc1C[S@@](=O)c1nc2ccccc2[nH]1
Drug	Dexfenfluramine	CCN[C@@H](C)Cc1cccc(C(F)(F)F)c1
Drug	Dexamethasone	C[C@@H]1C[C@H]2[C@@H]3CCC4=CC(=O)C=C[C@]4(C)[C@@]3(F)[C@@H](O)C[C@]2(C)[C@@]1(O)C(=O)CO
Drug	Desvenlafaxine	CN(C)CC(c1ccc(O)cc1)C1(O)CCCCC1
Drug	Desogestrel	C#C[C@]1(O)CC[C@H]2[C@@H]3CCC4=CCCC[C@@H]4[C@H]3C(=

		C)C[C@@]21CC
Drug	Desmethylsibutramine, (S)-	CN[C@@H](CC(C)C)C1(c2ccc(Cl)cc2)CCC1
Drug	Desimipramine	CNCCCN1c2ccccc2CCc2ccccc21
Drug	Deramciclane	CN(C)CCO[C@]1(c2ccccc2)C[C@H]2CC[C@]1(C)C2(C)C
Drug	Deoxypodophyllotoxin	COc1cc([C@@H]2c3cc4c(cc3C[C@H]3COC(=O)[C@@H]32)OCO4)cc(OC)c1OC
Drug	Delavirdine	CC(C)Nc1cccnc1N1CCN(C(=O)c2cc3cc(NS(C)(=O)=O)ccc3[nH]2)CC1
Drug	Delamanid	C[C@]1(COc2ccc(N3CCC(Oc4ccc(OC(F)(F)F)cc4)CC3)cc2)Cn2cc([N+](=O)[O-])nc2O1
Drug	Deflazacort	CC(=O)OCC(=O)[C@@]12N=C(C)O[C@@H]1C[C@H]1[C@@H]3CCC4=CC(=O)C=C[C@]4(C)[C@H]3[C@@H](O)C[C@@]12C
Drug	Debrisoquine	N=C(N)N1CCc2ccccc2C1
Drug	Darolutamide	CC(O)c1cc(C(=O)N[C@@H](C)Cn2ccc(-c3ccc(C#N)c(Cl)c3)n2)n[nH]1
Drug	Dapsone	Nc1ccc(S(=O)(=O)c2ccc(N)cc2)cc1
Drug	Dapagliflozin	CCOc1ccc(Cc2cc([C@@H]3O[C@H](CO)[C@@H](O)[C@H](O)[C@H]3O)ccc2Cl)cc1
Drug	Dantrolene	O=C1CN(/N=C/c2ccc(-c3ccc([N+](=O)[O-])cc3)o2)C(=O)N1
Drug	Danazol	C#C[C@]1(O)CC[C@H]2[C@@H]3CCC4=Cc5oncc5C[C@]4(C)[C@H]3CC[C@@]21C
Drug	D-alpha-Tocopherol acetate	CC(=O)Oc1c(C)c(C)c2c(c1C)CC[C@@](C)(CCC[C@H](C)CCC[C@H](C)CCCC(C)C)O2
Drug	Dalfampridine	Nc1ccncc1
Drug	Dacomitinib	COc1cc2ncnc(Nc3ccc(F)c(Cl)c3)c2cc1NC(=O)/C=C/CN1CCCCC1
Drug	Dacarbazine	CN(C)N/N=C1\N=CN=C1C(N)=O
Drug	Dabigatran etexilate	CCCCCCOC(=O)/N=C(/N)c1ccc(NCc2nc3cc(C(=O)N(CCC(=O)OCC)c4ccccn4)ccc3n2C)cc1
Drug	Dabigatran	Cn1c(CNc2ccc(C(=N)N)cc2)nc2cc(C(=O)N(CCC(=O)O)c3ccccn3)ccc21
Drug	Cyproterone acetate	CC(=O)O[C@]1(C(C)=O)CC[C@H]2[C@@H]3C=C(Cl)C4=CC(=O)[C@@H]5C[C@@H]5[C@]4(C)[C@H]3CC[C@@]21C
Drug	Cyproheptadine	CN1CCC(=C2c3ccccc3C=Cc3ccccc32)CC1
Drug	cyclosporine_a	C/C=C/C[C@@H](C)[C@@H](O)[C@@H]1C(O)=N[C@H](CC)C(=O)N(C)CC(=O)N(C)[C@H](CC(C)C)C(O)=N[C@@H](C(C)C)C(=O)N(C)[C@

		H](CC(C)C)C(O)=N[C@H](C)C(O)=N[C@@H](C)C(=O)N(C)[C@@H](CC(C)C)C(=O)N(C)[C@@H](CC(C)C)C(=O)N(C)[C@@H](C(C)C)C(=O)N1C
Drug	Cyclophosphamide	O=P1(N(CCCl)CCCl)NCCCO1
Drug	cyclobenzaprine	CN(C)CCC=C1c2ccccc2C=Cc2ccccc21
Drug	Cyanocobalamin	C/C1=C2/N=C(/C=C3\N=C(/C(C)=C4/[C@@H](CCC(N)=O)[C@](C)(CC(N)=O)[C@](C)([C@@H]5N=C1[C@](C)(CCC(=O)NC[C@@H](C)OP(=O)([O-])O[C@@H]1[C@@H](CO)O[C@H](n6cnc7cc(C)c(C)cc76)[C@@H]1O)[C@H]5CC(N)=O)N4[Co+]C#N)[C@@](C)(CC(N)=O)[C@@H]3CCC(N)=O)C(C)(C)[C@@H]2CCC(N)=O
Drug	Cyadox	N#CCC(=O)NN=Cc1c[n+]([O-])c2ccccc2[n+]1[O-]
Drug	CTK0G1203	C[C@@H](NC(=NC#N)N=c1cc[nH]cc1)C(C)(C)C
Drug	cp-199331	COc1ccc(NS(=O)(=O)C(F)(F)F)cc1C[C@@H]1COc2ccc(OCc3nc4cc(F)c(F)cc4s3)cc2[C@@H]1O
Drug	CP-195543	O=C(O)c1ccc(C(F)(F)F)cc1-c1ccc2c(c1)OC[C@H](Cc1ccccc1)[C@H]2O
Drug	cp_125_611	Oc1ccc(OC(F)(F)F)cc1CN[C@@H]1CCCN[C@H]1c1ccccc1
Drug	cp_122_721	COc1ccc(OC(F)(F)F)cc1CN[C@@H]1CCCN[C@H]1c1ccccc1
Drug	coumarin	O=c1ccc2ccccc2o1
Drug	Cortisone	C[C@]12CC(=O)[C@H]3[C@@H](CCC4=CC(=O)CC[C@@]43C)[C@@H]1CC[C@]2(O)C(=O)CO
Drug	Combretastatin A4	COc1ccc(/C=C\c2cc(OC)c(OC)c(OC)c2)cc1O
Drug	Colchicine	COc1cc2c(c(OC)c1OC)-c1ccc(OC)c(=O)cc1C(NC(C)=O)CC2
Drug	codeine	COc1ccc2c3c1O[C@H]1[C@@H](O)C=C[C@H]4[C@@H](C2)N(C)CC[C@]314
Drug	Cocaine	COC(=O)[C@H]1[C@@H](OC(=O)c2ccccc2)C[C@@H]2CC[C@H]1N2C
Drug	clozapine	CN1CCN(C2=Nc3cc(Cl)ccc3Nc3ccccc32)CC1
Drug	clotiazepam	CCc1cc2c(s1)N(C)C(=O)CN=C2c1ccccc1Cl
Drug	Clorazepic acid	O=C(O)C1N=C(c2ccccc2)c2cc(Cl)ccc2NC1=O
Drug	Clopidogrel	COC(=O)[C@@H](c1ccccc1Cl)N1CC

		c2sccc2C1
Drug	Clonazepam	O=C1CN=C(c2ccccc2Cl)c2cc([N+](=O)[O-])ccc2N1
Drug	clomipramine	CN(C)CCCN1c2ccccc2CCc2ccc(Cl)cc21
Drug	clomethiazole	Cc1ncsc1CCCl
Drug	Clofibric acid	CC(C)(Oc1ccc(Cl)cc1)C(=O)O
Drug	Clofarabine	Nc1nc(Cl)nc2c1ncn2[C@@H]1O[C@H](CO)[C@@H](O)[C@@H]1F
Drug	clobazam	CN1C(=O)CC(=O)N(c2ccccc2)c2cc(Cl)ccc21
Drug	Clioquinol	Oc1c(I)cc(Cl)c2cccnc12
Drug	clindamycin	CCC[C@@H]1C[C@@H](C(O)=N[C@H]([C@H](C)Cl)[C@H]2O[C@H](SC)[C@H](O)[C@@H](O)[C@H]2O)N(C)C1
Drug	Clarithromycin	CC[C@H]1OC(=O)[C@H](C)[C@@H](O[C@H]2C[C@@](C)(OC)[C@@H](O)[C@H](C)O2)[C@H](C)[C@@H](O[C@@H]2O[C@H](C)C[C@H](N(C)C)[C@H]2O)[C@](C)(OC)C[C@@H](C)C(=O)[C@H](C)[C@@H](O)[C@]1(C)O
Drug	Cladribine	Nc1nc(Cl)nc2c1ncn2[C@H]1C[C@H](O)[C@@H](CO)O1
Drug	Citalopram	CN(C)CCCC1(c2ccc(F)cc2)OCc2cc(C#N)ccc21
Drug	Cisatracurium	COc1ccc(C[C@@H]2c3cc(OC)c(OC)cc3CC[N@+]2(C)CCC(=O)OCCCCCOC(=O)CC[N@@+]2(C)CCc3cc(OC)c(OC)cc3[C@H]2Cc2ccc(OC)c(OC)c2)cc1OC
Drug	Cisapride	COc1cc(N)c(Cl)cc1C(O)=N[C@H]1CCN(CCCOc2ccc(F)cc2)C[C@H]1OC
Drug	cis-6-(2-acetylvinylthio)purine	CC(=O)/C=C\Sc1ncnc2[nH]cnc12
Drug	Ciprofloxacin	O=C(O)c1cn(C2CC2)c2cc(N3CCNCC3)c(F)cc2c1=O
Drug	cinnarizine	C(=C/c1ccccc1)\CN1CCN(C(c2ccccc2)c2ccccc2)CC1
Drug	Cinacalcet	C[C@@H](NCCCc1cccc(C(F)(F)F)c1)c1cccc2ccccc12
Drug	cilostazol	O=C1CCc2cc(OCCCCc3nnnn3C3CCCCC3)ccc2N1
Drug	Cilazapril	CCOC(=O)[C@H](CCc1ccccc1)N[C@H]1CCCN2CCC[C@@H](C(=O)O)N2C1=O
Drug	Choline C 11	C[N+](C)([11CH3])CCO.[Cl-]
Drug	chlorzoxazone	Oc1nc2cc(Cl)ccc2o1
Drug	Chlorquinaldol	Cc1ccc2c(Cl)cc(Cl)c(O)c2n1
Drug	chlorpropamide	CCCN=C(O)NS(=O)(=O)c1ccc(Cl)cc1

Drug	chlorpromazine	CN(C)CCCN1c2ccccc2Sc2ccc(Cl)cc21
Drug	Chloroxylenol	Cc1cc(O)cc(C)c1Cl
Drug	Chloroquine	CCN(CC)CCCC(C)Nc1ccnc2cc(Cl)ccc12
Drug	Chloramphenicol	O=C(N[C@H](CO)[C@H](O)c1ccc([N+](=O)[O-])cc1)C(Cl)Cl
Drug	Cevimeline	CC1O[C@@]2(CS1)CN1CCC2CC1
Drug	Cethromycin	CC[C@H]1OC(=O)[C@H](C)C(=O)[C@H](C)[C@@H](O[C@@H]2O[C@H](C)C[C@H](N(C)C)[C@H]2O)[C@](C)(OCC=Cc2cnc3ccccc3c2)C[C@@H](C)C(=O)[C@H](C)[C@H]2NC(=O)O[C@]12C
Drug	cerivastatin	COCc1c(C(C)C)nc(C(C)C)c(/C=C/[C@@H](O)C[C@@H](O)CC(=O)O)c1-c1ccc(F)cc1
Drug	celecoxib	Cc1ccc(-c2cc(C(F)(F)F)nn2-c2ccc(S(N)(=O)=O)cc2)cc1
Drug	Cefuroxime	CO/N=C(\C(=O)N[C@@H]1C(=O)N2C(C(=O)O)=C(COC(N)=O)CS[C@H]12)c1ccco1
Drug	Ceftaroline fosamil	CCO/N=C(\C(=O)N[C@@H]1C(=O)N2C(C(=O)[O-])=C(Sc3nc(-c4cc[n+](C)cc4)cs3)CS[C@H]12)c1nsc(NP(=O)(O)O)n1
Drug	Cefotaxime	CO/N=C(\C(=O)N[C@@H]1C(=O)N2C(C(=O)O)=C(COC(C)=O)CS[C@H]12)c1csc(N)n1
Drug	Cefepime	CO/N=C(\C(=O)N[C@@H]1C(=O)N2C(C(=O)[O-])=C(C[N+]3(C)CCCC3)CS[C@H]12)c1csc(N)n1
Drug	Cefapirin	CC(=O)OCC1=C(C(=O)O)N2C(=O)[C@@H](NC(=O)CSc3ccncc3)[C@H]2SC1
Drug	Carvedilol	COc1ccccc1OCCNCC(O)COc1cccc2[nH]c3ccccc3c12
Drug	Carteolol	CC(C)(C)NCC(O)COc1cccc2c1CCC(=O)N2
Drug	Carboxy celecoxib	NS(=O)(=O)c1ccc(-n2nc(C(F)(F)F)cc2-c2ccc(C(=O)O)cc2)cc1
Drug	Carbidopa	C[C@@](Cc1ccc(O)c(O)c1)(NN)C(=O)O
Drug	carbamazepine	NC(=O)N1c2ccccc2C=Cc2ccccc21
Drug	Captopril	C[C@H](CS)C(=O)N1CCC[C@H]1C(=O)O
Drug	Capecitabine	CCCCCOC(=O)Nc1nc(=O)n([C@@H]2O[C@H](C)[C@@H](O)[C@H]2O)cc1F
Drug	Candesartan cilexetil	CCOc1nc2cccc(C(=O)OC(C)OC(=O)

		OC3CCCCC3)c2n1Cc1ccc(-c2ccccc2-c2nnn[nH]2)cc1
Drug	Candesartan	CCOc1nc2cccc(C(=O)O)c2n1Cc1ccc(-c2ccccc2-c2nn[nH]n2)cc1
Drug	Camostat	CN(C)C(=O)COC(=O)Cc1ccc(OC(=O)c2ccc(NC(=N)N)cc2)cc1
Drug	Calcium carbimide	[Ca+2].[N-]=C=[N-]
Drug	Calcitriol	C=C1/C(=C\C=C2/CCC[C@@]3(C)[C@H]2CC[C@@H]3[C@H](C)CCCC(C)(C)O)C[C@@H](O)C[C@@H]1O
Drug	Cabergoline	C=CCN1C[C@H](C(=O)N(CCCN(C)C)C(=O)NCC)C[C@@H]2c3cccc4[nH]cc(c34)C[C@H]21
Drug	Cabazitaxel	CO[C@H]1C[C@H]2OC[C@@]2(OC(C)=O)[C@H]2[C@H](OC(=O)c3ccccc3)[C@]3(O)C[C@H](OC(=O)[C@H](O)[C@@H](NC(=O)OC(C)(C)C)c4ccccc4)C(C)=C([C@@H](OC)C(=O)[C@]12C)C3(C)C
Drug	Butyrfentanyl	CCCC(=O)N(c1ccccc1)C1CCN(CCc2ccccc2)CC1
Drug	Butylscopolamine	CCCC[N+]1(C)[C@H]2C[C@H](OC(=O)[C@H](CO)c3ccccc3)C[C@@H]1[C@H]1O[C@@H]21
Drug	Butalbital	C=CCC1(CC(C)C)C(=O)NC(=O)NC1=O
Drug	Buspirone	O=C1CC2(CCCC2)CC(=O)N1CCCCN1CCN(c2ncccn2)CC1
Drug	Bupropion	CC(NC(C)(C)C)C(=O)c1cccc(Cl)c1
Drug	Buprenorphine	CO[C@]12CC[C@@]3(C[C@@H]1[C@](C)(O)C(C)(C)C)[C@H]1Cc4ccc(O)c5c4[C@@]3(CCN1CC1CC1)[C@H]2O5
Drug	Bupranolol	Cc1ccc(Cl)c(OCC(O)CNC(C)(C)C)c1
Drug	Bufuralol	CCc1cccc2cc(C(O)CNC(C)(C)C)oc12
Drug	Budesonide	CCCC1O[C@@H]2C[C@H]3[C@@H]4CCC5=CC(=O)C=C[C@]5(C)[C@H]4[C@@H](O)C[C@]3(C)[C@]2(C(=O)CO)O1
Drug	brotizolam	Cc1nnc2n1-c1sc(Br)cc1C(c1ccccc1Cl)=NC2
Drug	brostallicin	C=C(Br)C(O)=Nc1cc(C(=O)Nc2cc(C(=O)Nc3cc(C(=O)Nc4cc(C(=O)NCCNC(=N)N)n(C)c4)n(C)c3)n(C)c2)n(C)c1
Drug	bropirimine	N=c1nc(O)c(Br)c(-c2ccccc2)[nH]1
Drug	Bronopol	O=[N+]([O-])C(Br)(CO)CO
Drug	bromperidol	O=C(CCCN1CCC(O)(c2ccc(Br)cc2)CC1)c1ccc(F)cc1
Drug	bromocriptine	CC(C)C[C@H]1C(=O)N2CCC[C@H]2[C@]2(O)O[C@](N=C(O)[C@@H]3

		C=C4c5cccc6[nH]c(Br)c(c56)C[C@H]4N(C)C3)(C(C)C)C(=O)N12
Drug	Bromazepam	OC1=Nc2ccc(Br)cc2C(c2ccccn2)=NC1
Drug	brofaromine	COc1cc(Br)c2oc(C3CCNCC3)cc2c1
Drug	Brodifacoum	O=c1oc2ccccc2c(O)c1C1CC(c2ccc(-c3ccc(Br)cc3)cc2)Cc2ccccc21
Drug	brn 4201400	Cc1nc2c(=O)n(C)c(=O)n(CC3CCCCC3)c2[nH]1
Drug	Brimonidine	Brc1c(NC2=NCCN2)ccc2nccnc12
Drug	Brexpiprazole	O=c1ccc2ccc(OCCCCN3CCN(c4cccc5sccc45)CC3)cc2[nH]1
Drug	bpr-0L075	COc1ccc2c(C(=O)c3cc(OC)c(OC)c(OC)c3)c[nH]c2c1
Drug	bosentan	COc1ccccc1Oc1c(NS(=O)(=O)c2ccc(C(C)(C)C)cc2)nc(-c2ncccn2)nc1OCCO
Drug	Bortezomib	CC(C)C[C@H](NC(=O)[C@H](Cc1ccccc1)NC(=O)c1cnccn1)B(O)O
Drug	bms_275183	COC(=O)O[C@@]12CO[C@@H]1C[C@H](O)[C@@]1(C)C(=O)[C@H](OC(C)=O)C3=C(C)[C@@H](OC(=O)[C@H](O)[C@@H](N=C(O)OC(C)(C)C)C(C)(C)C)C[C@@](O)([C@@H](OC(=O)c4ccccc4)[C@@H]12)C3(C)C
Drug	Blonanserin	CCN1CCN(c2cc(-c3ccc(F)cc3)c3c(n2)CCCCCC3)CC1
Drug	Bimatoprost	CCNC(=O)CCC/C=C\C[C@H]1[C@@H](O)C[C@@H](O)[C@@H]1/C=C/[C@@H](O)CCc1ccccc1
Drug	bifendate	COC(=O)c1cc(OC)c2c(c1-c1c(C(=O)OC)cc(OC)c3c1OCO3)OCO2
Drug	Bexarotene	C=C(c1ccc(C(=O)O)cc1)c1cc2c(cc1C)C(C)(C)CCC2(C)C
Drug	Betrixaban	COc1ccc(NC(=O)c2ccc(C(=N)N(C)C)cc2)c(C(=O)Nc2ccc(Cl)cn2)c1
Drug	Betamethasone phosphate	C[C@H]1C[C@H]2[C@@H]3CCC4=CC(=O)C=C[C@]4(C)[C@@]3(F)[C@@H](O)C[C@]2(C)[C@@]1(O)C(=O)COP(=O)(O)O
Drug	Betahistine	CNCCc1ccccn1
Drug	beta-artelinic acid	C[C@@H]1CC[C@H]2[C@@H](C)[C@@H](OCc3ccc(C(=O)O)cc3)O[C@@H]3O[C@@]4(C)CC[C@@H]1[C@@]23OO4
Drug	beta_arteether	CCO[C@H]1O[C@@H]2O[C@]3(C)CCC[C@H]4[C@H](C)CC[C@@H]([C@H]1C)[C@@]42OO3
Drug	Beta carotene	CC1=C(/C=C/C(C)=C/C=C/C(C)=C/C=C/C=C(C)/C=C/C=C(C)/C=C/C2=C(C)CCCC2(C)C)C(C)(C)CCC1

Drug	bergapten	COc1c2ccoc2cc2oc(=O)ccc12
Drug	Benzylpenicillin	CC1(C)S[C@@H]2[C@H](NC(=O)Cc3ccccc3)C(=O)N2[C@H]1C(=O)O
Drug	Benzydamine	CN(C)CCCOc1nn(Cc2ccccc2)c2ccccc12
Drug	Benzphetamine	C[C@@H](Cc1ccccc1)N(C)Cc1ccccc1
Drug	Benzonatate	CCCCNc1ccc(C(=O)OCCOCCOCCOCCOCCOCCOCCOCCOCCOC)cc1
Drug	benzodioxolyl-butanamine (BDB)	CC[C@H](N)Cc1ccc2c(c1)OCO2
Drug	Benzocaine	CCOC(=O)c1ccc(N)cc1
Drug	Benzhydrocodone	COc1ccc2c3c1O[C@H]1C(OC(=O)c4ccccc4)=CC[C@H]4[C@@H](C2)N(C)CC[C@]314
Drug	benzeneethanol, 2-[(2,6-dichlorophenyl)amino]-(reduced_diclofenac)	OCCc1ccccc1Nc1c(Cl)cccc1Cl
Drug	Benzatropine	CN1[C@H]2CC[C@@H]1C[C@H](OC(c1ccccc1)c1ccccc1)C2
Drug	Benidipine	COC(=O)C1=C(C)NC(C)=C(C(=O)O[C@@H]2CCCN(Cc3ccccc3)C2)[C@@H]1c1cccc([N+](=O)[O-])c1
Drug	Bendazac	O=C(O)COc1nn(Cc2ccccc2)c2ccccc12
Drug	Bendamustine	Cn1c(CCCC(=O)O)nc2cc(N(CCCl)CCCl)ccc21
Drug	Beclomethasone dipropionate	CCC(=O)OCC(=O)[C@@]1(OC(=O)CC)[C@@H](C)C[C@H]2[C@@H]3CCC4=CC(=O)C=C[C@]4(C)[C@@]3(Cl)[C@@H](O)C[C@@]21C
Drug	Bazedoxifene	Cc1c(-c2ccc(O)cc2)n(Cc2ccc(OCCN3CCCCCC3)cc2)c2ccc(O)cc12
Drug	barnidipine	COC(=O)C1=C(C)NC(C)=C(C(=O)O[C@H]2CCN(Cc3ccccc3)C2)[C@H]1c1cccc([N+](=O)[O-])c1
Drug	Bambuterol	CN(C)C(=O)Oc1cc(OC(=O)N(C)C)cc(C(O)CNC(C)(C)C)c1
Drug	Balsalazide	O=C(O)CCNC(=O)c1ccc(/N=N/c2ccc(O)c(C(=O)O)c2)cc1
Drug	Azimilide	CN1CCN(CCCCN2C(=O)CN(/N=C/c3ccc(-c4ccc(Cl)cc4)o3)C2=O)CC1
Drug	Azelastine	CN1CCCC(n2nc(Cc3ccc(Cl)cc3)c3ccccc3c2=O)CC1
Drug	Azathioprine	Cn1cnc([N+](=O)[O-])c1Sc1ncnc2nc[nH]c12
Drug	Avanafil	COc1ccc(CNc2nc(N3CCC[C@H]3CO)ncc2C(=O)NCc2ncccn2)cc1Cl
Drug	Atosiban	CCOc1ccc(C[C@H]2NC(=O)CCSSC[C@@H](C(=O)N3CCC[C@H]3C(=O)N[C@@H](CCCN)C(=O)NCC(N)=O

		)NC(=O)[C@H](CC(N)=O)NC(=O)[C@H]([C@@H](C)O)NC(=O)[C@H]([C@@H](C)CC)NC2=O)cc1
Drug	Atorvastatin	CC(C)c1c(C(O)=Nc2ccccc2)c(-c2ccccc2)c(-c2ccc(F)cc2)n1CC[C@@H](O)C[C@@H](O)CC(=O)O
Drug	Atomoxetine	CNCC[C@@H](Oc1ccccc1C)c1ccccc1
Drug	Atenolol	CC(C)NCC(O)COc1ccc(CC(N)=O)cc1
Drug	Ataluren	O=C(O)c1cccc(-c2noc(-c3ccccc3F)n2)c1
Drug	Asunaprevir	C=C[C@@H]1C[C@]1(NC(=O)[C@@H]1C[C@@H](Oc2ncc(OC)c3ccc(Cl)cc23)CN1C(=O)[C@@H](NC(=O)OC(C)(C)C)C(C)(C)C)C(=O)NS(=O)(=O)C1CC1
Drug	Astemizole	COc1ccc(CCN2CCC(Nc3nc4ccccc4n3Cc3ccc(F)cc3)CC2)cc1
Drug	Artesunate	C[C@@H]1CC[C@H]2[C@@H](C)[C@H](OC(=O)CCC(=O)O)O[C@@H]3O[C@@]4(C)CC[C@@H]1[C@@]23OO4
Drug	Artenimol	C[C@@H]1CC[C@H]2[C@@H](C)C(O)O[C@@H]3O[C@@]4(C)CC[C@@H]1[C@@]23OO4
Drug	Artemether	CO[C@H]1O[C@@H]2O[C@@]3(C)CC[C@H]4[C@H](C)CC[C@@H]([C@H]1C)[C@]42OO3
Drug	Arsenic trioxide	O=[As]O[As]=O
Drug	Aripiprazole lauroxil	CCCCCCCCCCCC(=O)OCN1C(=O)CCc2ccc(OCCCCN3CCN(c4cccc(Cl)c4Cl)CC3)cc21
Drug	Aripiprazole	ClC1=CC=CC(N2CCN(CCCCOC3=CC4=C(CCC(=O)N4)C=C3)CC2)=C1Cl
Drug	Arbidol	CCOC(=O)c1c(CSc2ccccc2)n(C)c2cc(Br)c(O)c(CN(C)C)c12
Drug	Aranidipine	COC(=O)C1=C(C)NC(C)=C(C(=O)OCC(C)=O)C1c1ccccc1[N+](=O)[O-]
Drug	Aprindine	CCN(CC)CCCN(c1ccccc1)C1Cc2ccccc2C1
Drug	aprepitant	C[C@@H](O[C@H]1OCCN(Cc2nc(O)n[nH]2)[C@H]1c1ccc(F)cc1)c1cc(C(F)(F)F)cc(C(F)(F)F)c1
Drug	Apremilast	CCOc1cc([C@@H](CS(C)(=O)=O)N2C(=O)c3cccc(NC(C)=O)c3C2=O)ccc1OC
Drug	Apomorphine	CN1CCc2cccc3c2[C@H]1Cc1ccc(O)c(O)c1-3
Drug	Apixaban	COc1ccc(-

		n2nc(C(N)=O)c3c2C(=O)N(c2ccc(N4CCCCC4=O)cc2)CC3)cc1
Drug	Apalutamide	CNC(=O)c1ccc(N2C(=S)N(c3cnc(C#N)c(C(F)(F)F)c3)C(=O)C23CCC3)cc1F
Drug	Antipyrine	O=C1C=C(C)N(C)N1c2ccccc2
Drug	Angiotensin II	CC[C@H](C)[C@H](NC(=O)[C@H](Cc1ccc(O)cc1)NC(=O)[C@@H](NC(=O)[C@H](CCCN=C(N)N)NC(=O)[C@@H](N)CC(=O)O)C(C)C)C(=O)N[C@@H](Cc1cnc[nH]1)C(=O)N1CCC[C@H]1C(=O)N[C@@H](Cc1ccccc1)C(=O)O
Drug	Anethole trithione	COc1ccc(-c2cc(=S)ss2)cc1
Drug	Anastrozole	CC(C)(C#N)c1cc(Cn2cncn2)cc(C(C)(C)C#N)c1
Drug	Anagrelide	O=C1CN2Cc3c(ccc(Cl)c3Cl)N=C2N1
Drug	Amrubicin	CC(=O)[C@]1(N)Cc2c(O)c3c(c(O)c2[C@@H](O[C@H]2C[C@H](O)[C@H](O)CO2)C1)C(=O)c1ccccc1C3=O
Drug	Amphetamine	CC(N)Cc1ccccc1
Drug	Amoxicillin	CC1(C)S[C@@H]2[C@H](NC(=O)[C@H](N)c3ccc(O)cc3)C(=O)N2[C@H]1C(=O)O
Drug	Amoxapine	Clc1ccc2c(c1)C(N1CCNCC1)=Nc1ccccc1O2
Drug	Amodiaquine	CCN(CC)Cc1cc(Nc2ccnc3cc(Cl)ccc23)ccc1O
Drug	amitriptyline	CN(C)CCC=C1c2ccccc2CCc2ccccc21
Drug	amiodarone	CCCCc1oc2ccccc2c1C(=O)c1cc(I)c(OCCN(CC)CC)c(I)c1
Drug	aminopyrine	CN(C)c1cn(C)n(-c2ccccc2)c1=O
Drug	Aminophenazone	Cc1c(N(C)C)c(=O)n(-c2ccccc2)n1C
Drug	Aminoflavone	Cc1c(F)c(N)c2c(=O)cc(-c3ccc(N)c(F)c3)oc2c1F
Drug	aminochrome	O=C1C=C2CCNC2=CC1=O
Drug	Amifampridine	Nc1ccncc1N
Drug	ambroxol	Nc1c(Br)cc(Br)cc1CN[C@H]1CC[C@@H](O)CC1
Drug	Ambrisentan	COC(c1ccccc1)(c1ccccc1)[C@H](Oc1nc(C)cc(C)n1)C(=O)O
Drug	Alvameline	CCn1nnc(C2=CCCN(C)C2)n1
Drug	Alprenolol	C=CCc1ccccc1OCC(O)CNC(C)C
Drug	alprazolam	Cc1nnc2n1-c1ccc(Cl)cc1C(c1ccccc1)=NC2
Drug	alpha-thujone	CC(C)[C@]12CC(=O)[C@H](C)[C@H]1C2
Drug	Alpelisib	Cc1nc(NC(=O)N2CCC[C@H]2C(N)=O)sc1-c1ccnc(C(C)(C)C(F)(F)F)c1
Drug	alosetron	Cc1nc[nH]c1CN1CCc2c(c3ccccc3n2C)C1=O

Drug	almotriptan	CN(C)CCc1c[nH]c2ccc(CS(=O)(=O)N3CCCC3)cc12
Drug	Allopurinol	Oc1ncnc2[nH]ncc12
Drug	alilusem	Cc1ccccc1C(=O)N1CC/C(=N\OS(=O)(=O)O)c2ccc(Cl)cc21
Drug	alfentanil	CCC(=O)N(c1ccccc1)C1(COC)CCN(CCn2nnn(CC)c2=O)CC1
Drug	Alectinib	CCc1cc2c(cc1N1CCC(N3CCOCC3)CC1)C(C)(C)c1[nH]c3cc(C#N)ccc3c1C2=O
Drug	albendazole	CCCSc1ccc2nc(N=C(O)OC)[nH]c2c1
Drug	Afloqualone	Cc1ccccc1-n1c(CF)nc2ccc(N)cc2c1=O
Drug	adinazolam	CN(C)Cc1nnc2n1-c1ccc(Cl)cc1C(c1ccccc1)=NC2
Drug	Ademetionine	C[S+](CC[C@H](N)C(=O)[O-])C[C@H]1O[C@@H](n2cnc3c(N)ncnc32)[C@H](O)[C@@H]1O
Drug	Adefovir	Nc1ncnc2c1ncn2CCOCP(=O)(O)O
Drug	Acetylsalicylic acid	CC(=O)Oc1ccccc1C(=O)O
Drug	Acetylcysteine	CC(=O)N[C@@H](CS)C(=O)O
Drug	Acetohexamide	CC(=O)c1ccc(S(=O)(=O)NC(=O)NC2CCCCC2)cc1
Drug	acetanilide	CC(O)=Nc1ccccc1
Drug	acetaminophen	CC(O)=Nc1ccc(O)cc1
Drug	Acemetacin	COc1ccc2c(c1)c(CC(=O)OCC(=O)O)c(C)n2C(=O)c1ccc(Cl)cc1
Drug	aceclofenac	O=C(O)COC(=O)Cc1ccccc1Nc1c(Cl)cccc1Cl
Drug	Acalabrutinib	CC#CC(=O)N1CCC[C@H]1c1nc(-c2ccc(C(=O)Nc3ccccn3)cc2)c2c(N)nccn12
Drug	ac1ocg1j	CN[C@@H](C)Cc1ccccc1OC
Drug	abiraterone	C[C@]12CC[C@H]3[C@@H](CC=C4C[C@@H](O)CC[C@@]43C)[C@@H]1CC=C2c1cccnc1
Drug	Abametapir	Cc1ccc(-c2ccc(C)cn2)nc1
Drug	7-Ethoxycoumarin	CCOC1=CC2=C(C=C1)C=CC(=O)O2
Drug	4-Methylumbelliferone	Cc1cc(=O)oc2cc(O)ccc12
Drug	2-chloro-3-pyridin-3-yl-5,6,7,8-tetrahydroindolizine-1-carboxamide	N=C(O)c1c(Cl)c(-c2cccnc2)n2c1CCCC2
Drug	1,4-Butanediol	OCCCCO
Drug	4-[5-(4-Fluoro-phenyl)-2-methylsulfanyl-3H-imidazol-4-yl]-pyridin-2-yl-(1-phenyl-ethyl)-amine	CSc1nc(-c2ccc(F)cc2)c(-c2ccnc(N[C@H](C)c3ccccc3)c2)[nH]1
Drug	[5-(Diethylamino)-2-methylpent-3-yn-2-yl] (2S)-2-cyclohexyl-2-hydroxy-2-phenylacetate	CCN(CC)CC#CC(C)(C)OC(=O)[C@@](O)(c1ccccc1)C1CCCCC1

Drug	[2-methyl-4-(o-tolylazo)phenyl]amine	Cc1cc(/N=N/c2ccccc2C)ccc1N
Drug	(s)-tomoxetine	CNCC[C@H](Oc1ccccc1C)c1ccccc1
Drug	(S)-Sibutramine	CC(C)C[C@H](N(C)C)C1(c2ccc(Cl)cc2)CCC1
Drug	(s)-propafenone	CCCNC[C@H](O)COc1ccccc1C(=O)CCc1ccccc1
Drug	(s)-oxybutynin	CCN(CC)CC#CCOC(=O)[C@@](O)(c1ccccc1)C1CCCCC1
Drug	(s)-norfenfluramine	C[C@H](N)Cc1cccc(C(F)(F)F)c1
Drug	(S)-mianserin	CN1CCN2c3ccccc3Cc3ccccc3[C@H]2C1
Drug	(S)-Methylenedioxyethylamphetamine	CCN[C@@H](C)Cc1ccc2c(c1)OCO2
Drug	(s)-mephobarbital	CC[C@@]1(c2ccccc2)C(=O)N(C)C(=O)N=C1O
Drug	(s)-hexobarbital	CN1C(=O)N=C(O)[C@](C)(C2=CCCCC2)C1=O
Drug	(s)-flurbiprofen	C[C@H](C(=O)O)c1ccc(-c2ccccc2)c(F)c1
Drug	(s)-fluoxetine	CNCC[C@H](Oc1ccc(C(F)(F)F)cc1)c1ccccc1
Drug	(s)-bupropion	C[C@H](NC(C)(C)C)C(=O)c1cccc(Cl)c1
Drug	(s)-azelastine	CN1CCC[C@H](n2nc(Cc3ccc(Cl)cc3)c3ccccc3c2=O)CC1
Drug	(S)-2-(2,6-Dichlorobenzamido)-3-(2',6'-dimethoxy-[1,1'-biphenyl]-4-yl)propanoic acid	COc1cccc(OC)c1-c1ccc(C[C@H](N=C(O)c2c(Cl)cccc2Cl)C(=O)O)cc1
Drug	(s)-(+)-ibuprofen	CC(C)Cc1ccc([C@H](C)C(=O)O)cc1
Drug	(S)-(-)-Mexiletine	Cc1cccc(C)c1OC[C@H](C)N
Drug	(s)-(-)-carvedilol	COc1ccccc1OCCNC[C@H](O)COc1cccc2[nH]c3ccccc3c12
Drug	(r)-trimipramine	C[C@H](CN(C)C)CN1c2ccccc2CCc2ccccc21
Drug	(R)-phenprocoumon	CC[C@H](c1ccccc1)c1c(O)c2ccccc2oc1=O
Drug	(R)-m-Methoxyamphetamine	COc1cccc(C[C@@H](C)N)c1
Drug	(R)-mianserin	CN1CCN2c3ccccc3Cc3ccccc3[C@@H]2C1
Drug	(r)-metoprolol	COCCc1ccc(OC[C@H](O)CNC(C)C)cc1
Drug	(r)-mda	C[C@@H](N)Cc1ccc2c(c1)OCO2
Drug	(r)-ketamine	CN[C@@]1(c2ccccc2Cl)CCCCC1=O
Drug	(R)-halothane	FC(F)(F)[C@H](Cl)Br
Drug	(r)-fluoxetine	CNCC[C@@H](Oc1ccc(C(F)(F)F)cc1)c1ccccc1
Drug	(r)-enflurane	FC(F)OC(F)(F)[C@H](F)Cl
Drug	(R)-colchicine	COc1cc2c(c(OC)c1OC)-

		c1ccc(OC)c(=O)cc1[C@H](N=C(C)O)CC2
Drug	(r)-chloroquine	CCN(CC)CCC[C@@H](C)Nc1ccnc2cc(Cl)ccc12
Drug	(r)-bupropion	C[C@@H](NC(C)(C)C)C(=O)c1cccc(Cl)c1
Drug	(R)-Bunitrolol	CC(C)(C)NC[C@@H](O)COc1ccccc1C#N
Drug	(r)-azelastine	CN1CCC[C@@H](n2nc(Cc3ccc(Cl)cc3)c3ccccc3c2=O)CC1
Drug	(R)-acenocoumarol	CC(=O)CC(c1ccc([N+](=O)[O-])cc1)c1c(O)c2ccccc2oc1=O
Drug	(R)-3,4-Methylenedioxymethamphetamine	CN[C@H](C)Cc1ccc2c(c1)OCO2
Drug	(R)-1-(4-Methoxyphenyl)propan-2-amine	COc1ccc(C[C@@H](C)N)cc1
Drug	(R)-(+)-Rebamipide	O=C(N[C@H](Cc1cc(O)nc2ccccc12)C(=O)O)c1ccc(Cl)cc1
Drug	(r)-(+)-felodipine	CCOC(=O)C1=C(C)NC(C)=C(C(=O)OC)[C@@H]1c1cccc(Cl)c1Cl
Drug	(R)-(-)-Mexiletine	Cc1cccc(C)c1OC[C@@H](C)N
Drug	(4R)-5-(methoxycarbonyl)-2,6-dimethyl-4-(3-nitrophenyl)-1,4-dihydropyridine-3-carboxylic acid	COC(=O)C1=C(C)NC(C)=C(C(=O)O)[C@H]1c1cccc([N+](=O)[O-])c1
Drug	(3z)-5-[(1alpha,4alpha)-3beta-[(phenylsulfonyl)amino]bicyclo[2.2.1]heptan-2alpha-yl]-3-pentenoic acid	O=C(O)C/C=C\C[C@@H]1[C@@H](NS(=O)(=O)c2ccccc2)[C@H]2CC[C@@H]1C2
Drug	(2s,5s)-troglitazone	Cc1c(C)c2c(c(C)c1O)CC[C@@](C)(COc1ccc(C[C@@H]3SC(=O)N=C3O)cc1)O2
Drug	(2S,3R)-3-(4-Hydroxy-phenyl)-2-[4-(2-piperidin-1-yl-ethoxy)-phenyl]-2,3-dihydro-benzo[1,4]oxathiin-6-ol	Oc1ccc([C@H]2Sc3cc(O)ccc3O[C@H]2c2ccc(OCCN3CCCCC3)cc2)cc1
Drug	(2S)-1-(1,3-Benzodioxol-5-yl)-N-methyl-2-butanamine	CC[C@@H](Cc1ccc2c(c1)OCO2)NC
Drug	(2R,3S)â€2â€(diphenylmethyl)â€Nâ€[2â€methoxyâ€5â€(propâ€1â€enâ€2â€yl)phenyl]methylâ€1â€azabicyclo[2.2.2]octanâ€3â€• amine	C=C(C)c1ccc(OC)c(CN[C@H]2C3CCN(CC3)[C@@H]2C(c2ccccc2)c2ccccc2)c1
Drug	(2R,3S)-2-((R)-1-(3,5-Bis(trifluoromethyl)phenyl)ethoxy)-3-(4-fluorophenyl)morpholine	C[C@@H](O[C@H]1OCCN[C@H]1c1ccc(F)cc1)c1cc(C(F)(F)F)cc(C(F)(F)F)c1
Drug	(2r)-2-[4-(3-methylthiophen-2-	Cc1ccsc1-

	yl)phenyl]propanoic acid	c1ccc([C@@H](C)C(=O)O)cc1
Drug	(2r)-1-(2-methoxyphenyl)propan-2-amine	COc1ccccc1C[C@@H](C)N
Drug	(1-methyl-5-piperazin-1-yl-3-propyl-1H-pyrazolo[4,3-d]pyrimidin-7-yl)-(5-methyl-pyridin-2-yl)-amine	CCCc1nn(C)c2c(Nc3ccc(C)cn3)nc(N3CCNCC3)nc12
Drug	(+)-halofantrine	CCCCN(CCCC)CC[C@@H](O)c1cc2c(Cl)cc(Cl)cc2c2cc(C(F)(F)F)ccc12
Drug	(+-)-Cisapride	COc1cc(N)c(Cl)cc1C(O)=NC1CCN(CCCOc2ccc(F)cc2)CC1OC
Drug	(+)-bisoprolol	CC(C)NC[C@@H](O)COc1ccc(COCCOC(C)C)cc1
Drug	(-)-zileuton	C[C@@H](c1cc2ccccc2s1)N(O)C(=N)O
Drug	(-)-thalidomide	O=C1N=C(O)CC[C@@H]1N1C(=O)c2ccccc2C1=O
Drug	(-)-pseudococaine	COC(=O)[C@@H]1[C@@H]2CC[C@H](C[C@H]1OC(=O)c1ccccc1)N2C
Drug	(-)-etodolac	CCc1cccc2c3c([nH]c12)[C@@](CC)(CC(=O)O)OCC3
Drug	(-)-ethosuximide	CC[C@]1(C)CC(O)=NC1=O
Drug	(-)-bisoprolol	CC(C)NC[C@H](O)COc1ccc(COCCOC(C)C)cc1
Human Metabolite	xanthurenic acid	O=C(O)c1cc(O)c2cccc(O)c2n1
Human Metabolite	Uroporphyrinogen I	O=C(O)CCc1c2[nH]c(c1CC(=O)O)Cc1[nH]c(c(CC(=O)O)c1CCC(=O)O)Cc1[nH]c(c(CC(=O)O)c1CCC(=O)O)Cc1[nH]c(c(CC(=O)O)c1CCC(=O)O)C2
Human Metabolite	Triiodothyronine	N[C@@H](Cc1cc(I)c(Oc2ccc(O)c(I)c2)c(I)c1)C(=O)O
Human Metabolite	trihydroxycoprostane	CC(C)CCC[C@@H](C)[C@H]1CC[C@H]2[C@H]3[C@H](C[C@H](O)[C@@]21C)[C@@]1(C)CC[C@@H](O)C[C@H]1C[C@H]3O
Human Metabolite	Thyroxine	NC(Cc1cc(I)c(Oc2cc(I)c(O)c(I)c2)c(I)c1)C(=O)O
Human Metabolite	testosterone	C[C@]12CC[C@H]3[C@@H](CCC4=CC(=O)CC[C@@]43C)[C@@H]1CC[C@@H]2O
Human Metabolite	Taurocholic acid	C[C@H](CCC(O)=NCCS(=O)(=O)O)[C@H]1CC[C@H]2[C@H]3[C@H](C[C@H](O)[C@@]21C)[C@@]1(C)CC[C@@H](O)C[C@H]1C[C@H]3O
Human Metabolite	serotonin	NCCc1c[nH]c2ccc(O)cc12

Human Metabolite	retinol	CC1=C(/C=C/C(C)=C/C=C/C(C)=C/CO)C(C)(C)CCC1
Human Metabolite	Retinoic acid, all-trans	CC1=C(/C=C/C(C)=C/C=C/C(C)=C/C(=O)O)C(C)(C)CCC1
Human Metabolite	retinal	CC1=C(/C=C/C(C)=C/C=C/C(C)=C/C=O)C(C)(C)CCC1
Human Metabolite	rac-4-hydroxynonenal	CCCCCC(O)C=CC=O
Human Metabolite	progesterone	CC(=O)[C@H]1CC[C@H]2[C@@H]3CCC4=CC(=O)CC[C@]4(C)[C@H]3CC[C@]12C
Human Metabolite	pregnenolone	CC(=O)[C@H]1CC[C@H]2[C@@H]3CC=C4C[C@@H](O)CC[C@]4(C)[C@H]3CC[C@]12C
Human Metabolite	pinoline	COc1ccc2[nH]c3c(c2c1)CCNC3
Human Metabolite	p-cresol	Cc1ccc(O)cc1
Human Metabolite	Norethisterone	C#C[C@]1(O)CC[C@H]2[C@@H]3CCC4=CC(=O)CC[C@@H]4[C@H]3CC[C@@]21C
Human Metabolite	N-Acetyl-5-HT	CC(O)=NCCc1c[nH]c2ccc(O)cc12
Human Metabolite	melatonin	COc1ccc2[nH]cc(CCN=C(C)O)c2c1
Human Metabolite	Lithocholic Acid	C[C@H](CCC(=O)O)[C@H]1CC[C@H]2[C@@H]3CC[C@@H]4C[C@H](O)CC[C@]4(C)[C@H]3CC[C@@]21C
Human Metabolite	isoursodeoxycholic acid	C[C@H](CCC(=O)O)[C@H]1CC[C@H]2[C@@H]3[C@@H](O)C[C@@H]4C[C@@H](O)CC[C@]4(C)[C@H]3CC[C@@]21C
Human Metabolite	ibrolipim	CCOP(=O)(Cc1ccc(C(O)=Nc2ccc(Br)cc2C#N)cc1)OCC
Human Metabolite	hyocholic acid	C[C@H](CCC(=O)O)[C@H]1CC[C@H]2[C@@H]3[C@H](O)[C@H](O)[C@@H]4C[C@H](O)CC[C@]4(C)[C@H]3CC[C@@]21C
Human Metabolite	Etiocholanolone	C[C@]12CC[C@H]3[C@@H](CC[C@@H]4C[C@H](O)CC[C@]34C)[C@@H]1CCC2=O
Human Metabolit	estrone	C[C@]12CC[C@@H]3c4ccc(O)cc4CC[C@H]3[C@@H]1CCC2=O

e		
Human Metabolite	Estradiol acetate	CC(=O)Oc1ccc2c(c1)CC[C@@H]1[C@@H]2CC[C@]2(C)[C@@H](O)CC[C@@H]12
Human Metabolite	estradiol	C[C@]12CC[C@@H]3c4ccc(O)cc4CC[C@H]3[C@@H]1CC[C@@H]2O
Human Metabolite	Dihyroxy-1H-indole	Oc1cc2cc[nH]c2cc1O
Human Metabolite	Dihydrocurcumin	COc1cc(/C=C/C(=O)/C=C(\O)CCc2ccc(O)c(OC)c2)ccc1O
Human Metabolite	dehydroepiandrosterone	C[C@]12CC[C@H]3[C@@H](CC=C4C[C@@H](O)CC[C@@]43C)[C@@H]1CCC2=O
Human Metabolite	Cholic acid	C[C@H](CCC(=O)O)[C@H]1CC[C@H]2[C@H]3[C@H](C[C@H](O)[C@@]21C)[C@@]1(C)CC[C@@H](O)C[C@H]1C[C@H]3O
Human Metabolite	Bilirubin	C=CC1=C(C)/C(=C/c2[nH]c(Cc3[nH]c(/C=C4\N=C(O)C(C)=C4C=C)c(C)c3CCC(=O)O)c(CCC(=O)O)c2C)N=C1O
Human Metabolite	arachidonic acid	CCCCC/C=C\C/C=C\C/C=C\C/C=C\CCCC(=O)O
Human Metabolite	Androstenedione	C[C@]12CC[C@H]3[C@@H](CCC4=CC(=O)CC[C@@]43C)[C@@H]1CCC2=O
Human Metabolite	2-hydroxyestrone	C[C@]12CC[C@@H]3c4cc(O)c(O)cc4CC[C@H]3[C@@H]1CCC2=O
Human Metabolite	2-hydroxyestradiol	C[C@]12CC[C@@H]3c4cc(O)c(O)cc4CC[C@H]3[C@@H]1CC[C@@H]2O
Human Metabolite	(s)-venlafaxine	COc1ccc([C@@H](CN(C)C)C2(O)CCCCC2)cc1
Human Metabolite	(13alpha)-3-methoxy-estra-1,3,5(10)-triene-17alpha-ol	COc1ccc2c(c1)CC[C@@H]1[C@@H]2CC[C@]2(C)[C@H]1CC[C@H]2O
Human Metabolite	Cholecalciferol	C=C1CC[C@H](O)C/C1=C/C=C1\CCC[C@@]2(C)[C@H]1CC[C@@H]2[C@H](C)CCCC(C)C
ENV	γ-HBCD	Br[C@H]1CC[C@H](Br)[C@H](Br)CC[C@@H](Br)[C@H](Br)CC[C@H]1Br
ENV	β-HBCD	Br[C@H]1CC[C@@H](Br)[C@H](Br)CC[C@@H](Br)[C@@H](Br)CC[C@H]1Br
ENV	α-HBCD	Br[C@H]1CC[C@@H](Br)[C@H](Br

		)CC[C@H](Br)[C@H](Br)CC[C@H]1Br
ENV	Tris(methylphenyl) phosphate	Cc1ccc(OP(=O)(Oc2ccc(C)cc2)Oc2ccc(C)cc2)cc1
ENV	Tris(2-ethylhexyl) phosphate	CCCCC(CC)COP(=O)(OCC(CC)CCCC)OCC(CC)CCCC
ENV	Tris(2-chloroisopropyl) phosphate	CC(CCl)OP(=O)(OC(C)CCl)OC(C)CCl
ENV	Tris(2-butoxyethyl) phosphate	CCCCOCCOP(=O)(OCCOCCCC)OCCOCCCC
ENV	Tris(2,3-dibromopropyl) phosphate	O=P(OCC(Br)CBr)(OCC(Br)CBr)OCC(Br)CBr
ENV	Tris(1,3-dichloro-2-propyl)phosphate	O=P(OC(CCl)CCl)(OC(CCl)CCl)OC(CCl)CCl
ENV	Tripropyl phosphate	CCCOP(=O)(OCCC)OCCC
ENV	Triphenyl Phosphate	O=P(Oc1ccccc1)(Oc1ccccc1)Oc1ccccc1
ENV	Tri-o-cresyl phosphate	Cc1ccccc1OP(=O)(Oc1ccccc1C)Oc1ccccc1C
ENV	Tetradecabromo-1,4-diphenoxybenzene	Brc1c(Br)c(Br)c(Oc2c(Br)c(Br)c(Oc3c(Br)c(Br)c(Br)c(Br)c3Br)c(Br)c2Br)c(Br)c1Br
ENV	Resorcinol Bis(diphenylphosphate)	O=P(Oc1ccccc1)(Oc1ccccc1)Oc1cccc(OP(=O)(Oc2ccccc2)Oc2ccccc2)c1
ENV	Pentaerythritol Dibromide	OCC(CO)(CBr)CBr
ENV	pentachlorobenzene	Clc1cc(Cl)c(Cl)c(Cl)c1Cl
ENV	Decabromodiphenyl Ethane	Brc1c(Br)c(Br)c(CCc2c(Br)c(Br)c(Br)c(Br)c2Br)c(Br)c1Br
ENV	Bis(2-ethylhexyl)-tetrabromophthalate	CCCCC(CC)COC(=O)c1c(Br)c(Br)c(Br)c(Br)c1C(=O)OCC(CC)CCCC
ENV	BDE 99	BrC1=C(Br)C=C(C(Br)=C1)OC2=CC=C(Br)C=C2Br
ENV	BDE 90	Brc1ccc(Oc2cc(Br)cc(Br)c2Br)c(Br)c1
ENV	BDE 85	BrC1=CC=C(OC2=CC=C(Br)C(Br)=C2Br)C(Br)=C1
ENV	BDE 66	C1=CC(=C(C=C1OC2=C(C=C(C=C2)Br)Br)Br)Br
ENV	BDE 49	Brc1ccc(Oc2cc(Br)ccc2Br)c(Br)c1
ENV	BDE 47	BrC1=CC=C(OC2=CC=C(Br)C=C2Br)C(Br)=C1
ENV	BDE 42	Brc1ccc(Oc2cccc(Br)c2Br)c(Br)c1
ENV	BDE 28	Brc1ccc(Oc2ccc(Br)cc2Br)cc1
ENV	BDE 25	C1=CC(=CC(=C1)Br)OC2=C(C=C(C=C2)Br)Br
ENV	BDE 209	BrC1=C(Br)C(Br)=C(Br)C(Br)=C1OC2=C(Br)C(Br)=C(Br)C(Br)=C2Br
ENV	BDE 137	C1=CC(=C(C=C1Br)Br)OC2=CC(=C(C(=C2Br)Br)Br)Br
ENV	BDE 123	C1=CC(=C(C=C1Br)Br)OC2=CC(=C(C(=C2)Br)Br)Br
ENV	BDE 68	Brc1cc(Br)cc(Oc2ccc(Br)cc2Br)c1

ENV	BDE 208	Brc1cc(Br)c(Br)c(Oc2c(Br)c(Br)c(Br)c(Br)c2Br)c1Br
ENV	BDE 203	Brc1cc(Br)c(Oc2c(Br)c(Br)c(Br)c(Br)c2Br)cc1Br
ENV	BDE 183	Brc1cc(Br)c(Oc2c(Br)cc(Br)c(Br)c2Br)cc1Br
ENV	BDE 154	Brc1cc(Br)c(Oc2cc(Br)c(Br)cc2Br)c(Br)c1
ENV	BDE 153	Brc1cc(Br)c(Oc2cc(Br)c(Br)cc2Br)cc1Br
ENV	BDE 15	Brc1ccc(Oc2ccc(Br)cc2)cc1
ENV	BDE 100	Brc1ccc(Oc2c(Br)cc(Br)cc2Br)c(Br)c1
ENV	2-ethylhexyldiphenyl phosphate	CCCCC(CC)COP(=O)(Oc1ccccc1)Oc1ccccc1
ENV	1,2-Dibromo-4-(1,2-dibromoethyl)cyclohexane	BrCC(Br)C1CCC(Br)C(Br)C1
Human Metabolite	indole	c1ccc2[nH]ccc2c1
Human Metabolite	Enterolactone	O=C1OC[C@H](Cc2cccc(O)c2)[C@H]1Cc1cccc(O)c1
Human Metabolite	benzoic acid	O=C(O)c1ccccc1
Human Metabolite	3-methylindole	Cc1c[nH]c2ccccc12
ENV	n_ac_dfec	CC(O)=N[C@@H](CSC(F)(F)[C@@H](OCF)C(F)F)C(=O)O
ENV	n_ac_ffvc	CC(O)=N[C@@H](CS/C(F)=C(/OCF)C(F)F)C(=O)O
ENV	schembl4081913	CN=C(O)Cc1ccc2ccn(C3CCN(CCc4ccccc4F)CC3)c2c1
ENV	schembl29464	CN1C(=O)[C@H](O)N=C(c2ccccc2)c2cc(Cl)ccc21
ENV	schembl149458	CN1CCCC[C@@H]1CCN1c2ccccc2Sc2ccc([S@@](C)=O)cc21
ENV	schembl4023988	COC(OC)[C@@]1(C)Oc2ccc(N)cc2[C@H](N(Cc2nnn(C)n2)c2ccc(Cl)cc2)[C@H]1O
ENV	Tentoxin	CC(C)C[C@@H]1N=C(O)[C@H](C)N(C)C(=O)CN=C(O)/C(=C\c2ccccc2)N(C)C1=O
ENV	Territrem a	COc1cc(-c2cc3c(c(=O)o2)C[C@]2(O)[C@@]4(C)C(=O)C=CC(C)(C)[C@]4(O)CC[C@@]2(C)O3)cc2c1OCO2
ENV	Ochratoxin A	CC1Cc2c(Cl)cc(C(=O)NC(Cc3ccccc3)C(=O)O)c(O)c2C(=O)O1
ENV	Deoxynivalenol	CC1=C[C@H]2O[C@@H]3[C@H](O)C[C@@](C)([C@]34CO4)[C@@]2(

		CO)[C@H](O)C1=O
ENV	Citrinin	CC1=C2C(=CO[C@H](C)[C@H]2C)C(=O)C(C(=O)O)=C1O
ENV	AFLATOXIN G2	COc1cc2c(c3oc(=O)c4c(c13)CCOC4=O)[C@@H]1CCO[C@@H]1O2
ENV	AFLATOXIN B2	COc1cc2c(c3oc(=O)c4c(c13)CCC4=O)[C@@H]1CCO[C@@H]1O2
ENV	AFLATOXIN B1	COc1cc2c(c3oc(=O)c4c(c13)CCC4=O)[C@@H]1C=CO[C@@H]1O2
ENV	AFLATOXIN G1	COc1cc2c(c3oc(=O)c4c(c13)CCOC4=O)C1C=COC1O2
ENV	α-pinene (αPN)	CC1=CCC2CC1C2(C)C
ENV	Wogonin	COc1c(O)cc(O)c2c(=O)cc(-c3ccccc3)oc12
ENV	trans-isoeugenol	C/C=C/c1ccc(O)c(OC)c1
ENV	Theobromine	Cn1cnc2c1c(O)nc(=O)n2C
ENV	Thalifendine	COC1=CC=C2C=C3c4cc5c(cc4CC[NH+]3C=C2C1=O)OCO5
ENV	Tetradecanoic acid	CCCCCCCCCCCCCC(=O)O
ENV	Stearic Acid	CCCCCCCCCCCCCCCCCC(=O)O
ENV	Sinapinic acid	COc1cc(/C=C/C(=O)O)cc(OC)c1O
ENV	Silybin	COc1cc([C@H]2Oc3cc([C@H]4Oc5cc(O)cc(O)c5C(=O)[C@@H]4O)ccc3O[C@@H]2CO)ccc1O
ENV	Senecionine	C/C=C1/C[C@H](C)[C@@](C)(O)C(=O)OCC2=CCN3CC[C@@H](OC1=O)[C@@H]23
ENV	Scutellarein	O=c1cc(-c2ccc(O)cc2)oc2cc(O)c(O)c(O)c12
ENV	Scopoletin	COc1cc2ccc(=O)oc2cc1O
ENV	Rutaecarpine	O=c1c2ccccc2nc2n1CCc1c-2[nH]c2ccccc12
ENV	Riddelliine	C=C1C/C(=C/C)C(=O)O[C@@H]2CCN3CC=C(COC(=O)[C@@]1(O)CO)[C@H]23
ENV	Rhein	O=C(O)c1cc(O)c2c(c1)C(=O)c1cccc(O)c1C2=O
ENV	Rhamnetin	COc1cc(O)c2c(=O)c(O)c(-c3ccc(O)c(O)c3)oc2c1
ENV	Resveratrol	Oc1ccc(/C=C/c2cc(O)cc(O)c2)cc1
ENV	Resokaempferol	O=c1c(O)c(-c2ccc(O)cc2)oc2cc(O)ccc12
ENV	Quercetin	O=c1c(O)c(-c2ccc(O)c(O)c2)oc2cc(O)cc(O)c12
ENV	Pterostilbene	COc1cc(/C=C/c2ccc(O)cc2)cc(OC)c1
ENV	Prunetin	COc1cc(O)c2c(=O)c(-c3ccc(O)cc3)coc2c1
ENV	Pristane	CC(C)CCCC(C)CCCC(C)CCCC(C)C
ENV	Prechrysophanol	Cc1cc(O)c2c(O)c3c(cc2c1)CC[C@@H](O)C3
ENV	Piperine	O=C(/C=C/C=C/c1ccc2c(c1)OCO2)N

		1CCCCC1
ENV	Physcion	COc1cc(O)c2c(c1)C(=O)c1cc(C)cc(O)c1C2=O
ENV	Phloretin	O=C(CCc1ccc(O)cc1)c1c(O)cc(O)cc1O
ENV	Phenethyl isothiocyanate	S=C=NCCc1ccccc1
ENV	Palmitic acid	CCCCCCCCCCCCCCCC(=O)O
ENV	Paeonol	COc1ccc(C(C)=O)c(O)c1
ENV	Oroxylin A	COc1c(O)cc2oc(-c3ccccc3)cc(=O)c2c1O
ENV	Oleic acid	CCCCCCCC/C=C\CCCCCCCC(=O)O
ENV	Norpatchoulenol	CC1(C)[C@@H]2CC[C@@]3(C)[C@H](C=CC[C@@]13O)C2
ENV	N'-nitrosonornicotine	O=NN1CCCC1c1cccnc1
ENV	N'-Nitrosoanabasine	O=NN1CCCCC1c1cccnc1
ENV	Nicotine	CN1CCC[C@H]1c1cccnc1
ENV	Narigenin	O=C1CC(c2ccc(O)cc2)Oc2cc(O)cc(O)c21
ENV	Myricetin	O=c1c(O)c(-c2cc(O)c(O)c(O)c2)oc2cc(O)cc(O)c12
ENV	Morin	O=c1c(O)c(-c2ccc(O)cc2O)oc2cc(O)cc(O)c12
ENV	Methyleugenol	C=CCc1ccc(OC)c(OC)c1
ENV	Magnolol	C=CCc1ccc(O)c(-c2cc(CC=C)ccc2O)c1
ENV	Maackiain	Oc1ccc2c(c1)OC[C@H]1c3cc4c(cc3O[C@@H]21)OCO4
ENV	Luteolin	O=c1cc(-c2ccc(O)c(O)c2)oc2cc(O)cc(O)c12
ENV	L-menthol	CC(C)[C@@H]1CC[C@@H](C)C[C@H]1O
ENV	Linoleic acid	CCCCC/C=C\C/C=C\CCCCCCCC(=O)O
ENV	kobophenol A	Oc1ccc(C2Oc3cc(O)cc(C4c5c(cc(O)cc5C5C(c6ccc(O)cc6)OC(c6ccc(O)cc6)C5c5cc(O)cc(O)c5)OC4c4ccc(O)cc4)c3C2c2cc(O)cc(O)c2)cc1
ENV	Kaempferol	O=c1c(O)c(-c2ccc(O)cc2)oc2cc(O)cc(O)c12
ENV	Jaceidin	COc1cc(-c2oc3cc(O)c(OC)c(O)c3c(=O)c2OC)ccc1O
ENV	Isorhamnetin	COc1cc(-c2oc3cc(O)cc(O)c3c(=O)c2O)ccc1O
ENV	Isoferulic Acid	COc1ccc(/C=C/C(=O)O)cc1O
ENV	Glycitein	COc1cc2c(=O)c(-c3ccc(O)cc3)coc2cc1O
ENV	Geraniol	CC(C)=CCC/C(C)=C/CO
ENV	Geraldol	COc1cc(-c2oc3cc(O)ccc3c(=O)c2O)ccc1O

ENV	Gentisic acid	C1=CC(=C(C=C1O)C(=O)O)O
ENV	Genistein	O=c1c(-c2ccc(O)cc2)coc2cc(O)cc(O)c12
ENV	Fraxetin	COc1cc2ccc(=O)oc2c(O)c1O
ENV	Formononetin	COc1ccc(-c2coc3cc(O)ccc3c2=O)cc1
ENV	Flavopiridol	CN1CC[C@H](c2c(O)cc(O)c3c(=O)cc(-c4ccccc4Cl)oc23)[C@H](O)C1
ENV	Fisetin	O=c1c(O)c(-c2ccc(O)c(O)c2)oc2cc(O)ccc12
ENV	Fenchone	CC12CCC(C1)C(C)(C)C2=O
ENV	Eupatilin	COc1ccc(-c2cc(=O)c3c(O)c(OC)c(O)cc3o2)cc1OC
ENV	Eugenol	C=CCc1ccc(O)c(OC)c1
ENV	Equol	Oc1ccc(C2COc3cc(O)ccc3C2)cc1
ENV	Epicatechin	Oc1cc(O)c2c(c1)OC(c1ccc(O)c(O)c1)C(O)C2
ENV	Emodin	Cc1cc(O)c2c(c1)C(=O)c1cc(O)cc(O)c1C2=O
ENV	eicosapentaenoic acid	CC/C=C\C/C=C\C/C=C\C/C=C\C/C=C\CCCC(=O)O
ENV	Diuron	CN(C)C(=O)Nc1ccc(Cl)c(Cl)c1
ENV	Diosmetin	COc1ccc(-c2cc(=O)c3c(O)cc(O)cc3o2)cc1O
ENV	Diallyl disulfide	C=CCSSCC=C
ENV	delta-9-Tetrahydrocannabinol	CCCCCc1cc(O)c2c(c1)OC(C)(C)[C@@H]1CCC(C)=C[C@@H]21
ENV	Decursinol angelate	C/C=C(/C)C(=O)O[C@H]1Cc2cc3ccc(=O)oc3cc2OC1(C)C
ENV	Datiscetin	O=c1c(O)c(-c2ccccc2O)oc2cc(O)cc(O)c12
ENV	Daidzein	O=c1c(-c2ccc(O)cc2)coc2cc(O)ccc12
ENV	Curcumin	COc1cc(/C=C/C(=O)CC(=O)/C=C/c2ccc(O)c(OC)c2)ccc1O
ENV	Cumene	CC(C)c1ccccc1
ENV	Cotinine	CN1C(=O)CC[C@H]1c1cccnc1
ENV	Cinnamaldehyde	O=C/C=C/c1ccccc1
ENV	Chrysophanol	Cc1cc(O)c2c(c1)C(=O)c1cccc(O)c1C2=O
ENV	Chrysoeriol	COc1cc(-c2cc(=O)c3c(O)cc(O)cc3o2)ccc1O
ENV	Chrysin	O=c1cc(-c2ccccc2)oc2cc(O)cc(O)c12
ENV	Chelidonine	CN1Cc2c(ccc3c2OCO3)[C@H]2[C@@H](O)Cc3cc4c(cc3[C@H]21)OCO4
ENV	Carvone	C=C(C)C1CC=C(C)C(=O)C1
ENV	Carvacrol	Cc1ccc(C(C)C)cc1O
ENV	Caffeine	Cn1c(=O)c2c(ncn2C)n(C)c1=O
ENV	Caffeic Acid	O=C(O)/C=C/c1ccc(O)c(O)c1
ENV	Bisdemethoxycurcumin	O=C(/C=C/c1ccc(O)cc1)CC(=O)/C=C/c1ccc(O)cc1

ENV	Biochanin a	COc1ccc(-c2coc3cc(O)cc(O)c3c2=O)cc1
ENV	Berberine	O1c2c(OC1)cc5c(c2)c4cc3ccc(OC)c(OC)c3c[n+]4CC5
ENV	Ascorbic acid	O=C1O[C@H]([C@@H](O)CO)C(O)=C1O
ENV	Arbutin	OC[C@H]1O[C@@H](Oc2ccc(O)cc2)[C@H](O)[C@@H](O)[C@@H]1O
ENV	Apigenin	O=c1cc(-c2ccc(O)cc2)oc2cc(O)cc(O)c12
ENV	Anethole	C/C=C/c1ccc(OC)cc1
ENV	Aloe emodin	O=C1c2cccc(O)c2C(=O)c2c(O)cc(CO)cc21
ENV	6-gingerol	CCCCCC(O)CC(=O)CCc1ccc(OC)c(O)c1
ENV	4-Cymene	CC(C)c1ccc(C(=O)O)cc1
ENV	2-Methylfuran	Cc1ccco1
ENV	1,8-Cineole	CC12CCC(CC1)C(C)(C)O2
ENV	1,4-cineole	CC(C)C12CCC(C)(CC1)O2
ENV	(S)-camphor	CC1(C)[C@H]2CC[C@]1(C)C(=O)C2
ENV	(1S,6R)-(+)-3-carene	CC1=CC[C@@H]2[C@H](C1)C2(C)C
ENV	(1S)-(-)-cis-Pinane	C[C@@H]1CC[C@H]2C[C@@H]1C2(C)C
ENV	(+)-Menthol	CC(C)[C@H]1CC[C@H](C)C[C@@H]1O
ENV	(+)-Limonene	C=C(C)[C@H]1CC=C(C)CC1
ENV	(+)-isothujone	CC(C)[C@]12CC(=O)[C@@H](C)[C@H]1C2
ENV	(+)-Camphor	CC1(C)[C@@H]2CC[C@@]1(C)C(=O)C2
ENV	(-)-menthone	CC(C)[C@@H]1CC[C@@H](C)CC1=O
ENV	(-)-limonene	C=C(C)[C@@H]1CC=C(C)CC1
ENV	(-)-Epigallocatechin gallate	O=C(O[C@@H]1Cc2c(O)cc(O)cc2O[C@@H]1c1cc(O)c(O)c(O)c1)c1cc(O)c(O)c(O)c1
ENV	(-)-Epigallocatechin	Oc1cc(O)c2c(c1)O[C@H](c1cc(O)c(O)c(O)c1)[C@H](O)C2
ENV	(-)-beta-Pinene	C=C1CC[C@H]2C[C@@H]1C2(C)C
Drug	Vitamin E	Cc1c(C)c2c(c(C)c1O)CC[C@@](C)(CCC[C@H](C)CCC[C@H](C)CCCC(C)C)O2
Drug	Vitamin D2	C=C1CC[C@H](O)C/C1=C/C=C1\CCC[C@@]2(C)[C@H]1CC[C@@H]2[C@H](C)/C=C/[C@H](C)C(C)C
Drug	tyramine	NCCc1ccc(O)cc1
ENV	zinc8660420	C=C(C)[C@H]1Cc2c(ccc3c2O[C@@H]2COc4cc(OC)c(OC)cc4[C@H]2C3=O)O1

ENV	zinc4556587	CC[C@@H]1C[C@H]2C[C@@H]3c4[nH]c5ccc(OC)cc5c4CCN(C2)[C@@H]31
ENV	zinc40442723	O=C1C=C2CN([C@@H](C(=O)C3CC3)c3ccccc3F)CC[C@H]2S1
ENV	zinc38309998	CCC(C)(C)C(=O)O[C@H]1C[C@@H](C)C=C2C=C[C@H](C)[C@H](CC[C@H](O)C[C@@H](O)CC(=O)O)[C@H]21
ENV	zinc36377943	CCC[C@@H]1O[C@@H]2C[C@H]3[C@@H]4CCC5=CC(=O)C=C[C@]5(C)[C@H]4[C@H](O)C[C@]3(C)[C@]2(C(=O)CO)O1
ENV	zinc33952517	C[C@H](C(=O)c1ccc2c(c1)OCO2)N1CCCC1
ENV	zinc33952402	Cc1ccc(C(=O)[C@@H](C)N2CCCC2)cc1
ENV	zinc21985223	CN(C)c1ccc([C@H]2C[C@@]3(C)[C@@H](CC[C@@]3(O)/C=C/CO)[C@@H]3CCC4=CC(=O)CCC4=C32)cc1
ENV	zinc188	C[C@H](CN(C)C)CN1c2ccccc2Sc2ccc(C#N)cc21
ENV	zinc1698316	COC(=O)[C@H]1[C@H]2C[C@H]3c4[nH]c5ccccc5c4CCN3C[C@@H]2CC[C@H]1O
ENV	zinc15020070	COc1ccc2c(c1)Oc1cc(O)ccc1[C@@]21OC(=O)c2ccccc21
ENV	zinc03978059	C[C@]12C[C@H](O)[C@@H]3[C@H](CCC4=CC(=O)CC[C@@]43C)[C@H]1CC[C@]2(O)C(=O)CO
ENV	Tetraphenyltin	c1ccc([Sn](c2ccccc2)(c2ccccc2)c2ccccc2)cc1
ENV	O-methoxycatechol	COOc1ccccc1O
ENV	N-Phenyl-2-naphthylamine	c1ccc(Nc2ccc3ccccc3c2)cc1
ENV	n-Hydroxy-3-aminobenzanthrone	O=C1c2ccccc2-c2ccc(NO)c3cccc1c23
ENV	N-acetyl-N-hydroxy-3-aminobenzanthrone	CC(=O)N(O)C1C=CC2=C3C(=CC=CC31)C(=O)c1ccccc12
ENV	N-Acetylbenzidine	CC(O)=Nc1ccc(-c2ccc(N)cc2)cc1
ENV	N-(Benzylideneamino)benzamide	O=C(NN=Cc1ccccc1)c1ccccc1
ENV	n-(2-phenylethyl)indomethacinamide	COc1ccc2c(c1)c(CC(O)=NCCc1ccccc1)c(C)n2C(=O)c1ccc(Cl)cc1
ENV	N-(1-Phenyl-1H-pyrazole-5-yl)-p-toluenesulfonamide	Cc1ccc(S(=O)(=O)Nc2ccnn2-c2ccccc2)cc1
ENV	Monochlorobimane	Cc1c(C)n2c(CCl)c(C)c(=O)n2c1=O
ENV	MNAN	CCCCCN(C)N=O
ENV	Hexafluoroisopropanol	OC(C(F)(F)F)C(F)(F)F
ENV	Diphenylamine	c1ccc(Nc2ccccc2)cc1
ENV	dibenzylfluorescein	O=C(OCc1ccccc1)c1ccccc1-c1c2ccc(=O)cc-2oc2cc(OCc3ccccc3)ccc12

ENV	chlorofluoroacetic acid	O=C(O)C(F)Cl
ENV	Catechol	Oc1ccccc1O
ENV	Biphenyl	c1ccc(-c2ccccc2)cc1
ENV	benzophenone	O=C(c1ccccc1)c1ccccc1
ENV	7-Butoxy-4-(trifluoromethyl)-2H-chromen-2-one	CCCCOc1ccc2c(C(F)(F)F)cc(=O)oc2c1
ENV	5-methychyrsene	Cc1cc2ccccc2c2ccc3ccccc3c12
ENV	4-Phenylphenol	Oc1ccc(-c2ccccc2)cc1
ENV	4-Phenylazophenol	Oc1ccc(/N=N/c2ccccc2)cc1
ENV	4-phenoxyaniline	Nc1ccc(Oc2ccccc2)cc1
ENV	4-nonylphenol	CCCCCCCCCc1ccc(O)cc1
ENV	4-n-nonylphenol diethoxylate	CCCCCCCCCc1ccc(OCCOCCO)cc1
ENV	4-nitropyrene	O=[N+]([O-])c1cc2cccc3ccc4cccc1c4c32
ENV	4-Nitrophenol	O=[N+]([O-])c1ccc(O)cc1
ENV	4-nitrophenethyl bromide	O=[N+]([O-])c1ccc(CCBr)cc1
ENV	4-nitrobenzyl chloride	O=[N+]([O-])c1ccc(CCl)cc1
ENV	4-nitroanisole	COc1ccc([N+](=O)[O-])cc1
ENV	4-nitro-3-(trifluoromethyl)aniline	Nc1ccc([N+](=O)[O-])c(C(F)(F)F)c1
ENV	4-n-Butylphenol	CCCCc1ccc(O)cc1
ENV	4-Methylbiphenyl	Cc1ccc(-c2ccccc2)cc1
ENV	4-methylaminomethyl-7-methoxycoumarin	CNCc1cc(=O)oc2cc(OC)ccc12
ENV	4'-Methoxy-alpha-pyrrolidinopropiophenone, (S)-	COc1ccc(C(=O)[C@H](C)N2CCCC2)cc1
ENV	4-Iodophenol	Oc1ccc(I)cc1
ENV	4-Iodobenzoic acid	O=C(O)c1ccc(I)cc1
ENV	4-Hydroxybenzophenone	O=C(c1ccccc1)c1ccc(O)cc1
ENV	4-Fluorophenol	Oc1ccc(F)cc1
ENV	4-Ethoxyphenol	CCOc1ccc(O)cc1
ENV	4-Cyclopentylphenol	Oc1ccc(C2CCCC2)cc1
ENV	4-Chlorophenol	Oc1ccc(Cl)cc1
ENV	4-Bromophenol	Oc1ccc(Br)cc1
ENV	4-Biphenylmethanol	OCc1ccc(-c2ccccc2)cc1
ENV	4-aminoveratrole	COc1ccc(N)cc1OC
ENV	4-Aminobiphenyl	Nc1ccc(-c2ccccc2)cc1
ENV	4,4'-Dichlorobiphenyl	Clc1ccc(-c2ccc(Cl)cc2)cc1
ENV	4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol	CN(CCCC(O)c1cccnc1)N=O
ENV	4-(ethylaminomethyl)-7-methoxycoumarin	CCNCc1cc(=O)oc2cc(OC)ccc12
ENV	4-(butylaminomethyl)-7-methoxycoumarin	CCCCNCc1cc(=O)oc2cc(OC)ccc12
ENV	3-Methylbiphenyl	Cc1cccc(-c2ccccc2)c1
ENV	2-propylquinoline	CCCc1ccc2ccccc2n1
ENV	2-nitropyrene	O=[N+]([O-

		])c1cc2ccc3cccc4ccc(c1)c2c34
ENV	2-naphthylamine	Nc1ccc2ccccc2c1
ENV	2-Methylbiphenyl	Cc1ccccc1-c1ccccc1
ENV	2-iodobenzyl bromide	BrCc1ccccc1I
ENV	2-acetamidofluorene	CC(O)=Nc1ccc2c(c1)Cc1ccccc1-2
ENV	2,8-Dihydroxyquinoline	Oc1ccc2cccc(O)c2n1
ENV	2,7-Dimethylnaphthalene	Cc1ccc2ccc(C)cc2c1
ENV	2,5-hexanedione	CC(=O)CCC(C)=O
ENV	2,4-dinitrobenzylalcohol	O=[N+]([O-])c1ccc(CO)c([N+](=O)[O-])c1
ENV	2,2-dichloro-2-fluoroethanol	OCC(F)(Cl)Cl
ENV	1-Naphthol	Oc1cccc2ccccc12
ENV	1-Naphtaldehyde	O=Cc1cccc2ccccc12
ENV	1-iodobutane	CCCCI
ENV	1-iodo-3-phenylpropane	ICCCc1ccccc1
ENV	1-chloro-2,4-dinitrobenzene	O=[N+]([O-])c1ccc(Cl)c([N+](=O)[O-])c1
ENV	1-chloro-2,3-dinitrobenzene	O=[N+]([O-])c1cccc(Cl)c1[N+](=O)[O-]
ENV	1-aminopyrene	Nc1ccc2ccc3cccc4ccc1c2c34
ENV	1,8-dinitropyrene	O=[N+]([O-])c1ccc2ccc3ccc([N+](=O)[O-])c4ccc1c2c34
ENV	1,6-dimethylnaphthalene	Cc1ccc2c(C)cccc2c1
ENV	1,3,5-trinitrobenzene	O=[N+]([O-])c1cc([N+](=O)[O-])cc([N+](=O)[O-])c1
ENV	(S)-4-Vinylcyclohexene	C=C[C@@H]1CC=CCC1
ENV	(R,R)-1-phenylpropylene oxide	CC1OC1c1ccccc1
ENV	(2R)-N-Benzyl-N-methyl-1-phenylpropan-2-amine	C[C@H](Cc1ccccc1)N(C)Cc1ccccc1
ENV	(+)-2-bromo-3-(4-nitrophenyl)propanoic acid	O=C(O)C(Br)Cc1ccc([N+](=O)[O-])cc1
ENV	Pyrene	c1cc2ccc3cccc4ccc(c1)c2c34
ENV	phenanthrene	c1ccc2c(c1)ccc1ccccc12
ENV	Naphthalene	c1ccc2ccccc2c1
ENV	Fluoranthene	c1ccc2c(c1)-c1cccc3cccc-2c13
ENV	dibenzo[a,l]pyrene	c1ccc2c(c1)cc1ccc3cccc4c5ccccc5c2c1c34
ENV	dibenzo[a,h]anthracene	c1ccc2c(c1)ccc1cc3c(ccc4ccccc43)cc12
ENV	Dibenz[a,j]acridine	c1ccc2c(c1)ccc1nc3ccc4ccccc4c3cc12
ENV	Chrysene	c1ccc2c(c1)ccc1c3ccccc3ccc21
ENV	Benzo[a]pyrene	c1ccc2c(c1)cc1ccc3cccc4ccc2c1c34
ENV	Benzo(k)fluoranthene	c1ccc2cc3c(cc2c1)-c1cccc2cccc-3c12
ENV	Benzo(j)fluoranthene	c1ccc2c3c(ccc2c1)-c1cccc2cccc-3c12
ENV	Benzo(b)fluoranthene	c1ccc2c(c1)-c1cccc3c1c-2cc1ccccc13
ENV	Acenaphthylene	C1=Cc2cccc3cccc1c23
ENV	Acenaphthene	c1cc2c3c(cccc3c1)CC2

ENV	2-methylnaphthalene	Cc1ccc2ccccc2c1
ENV	2-Ethylnaphthalene	CCc1ccc2ccccc2c1
ENV	2,2',3,3',4,5',6'-HEPTACHLOROBIPHENYL /PCB 177	Clc1ccc(-c2c(Cl)c(Cl)cc(Cl)c2Cl)c(Cl)c1Cl
ENV	1-nitropyrene	O=[N+]([O-])c1ccc2ccc3cccc4ccc1c2c34
ENV	1-n-hexylnaphthalene	CCCCCCc1cccc2ccccc12
ENV	1-ethylnaphthalene	CCc1cccc2ccccc12
ENV	1- methylnaphthalene	Cc1cccc2ccccc12
ENV	PCB132	Clc1ccc(-c2c(Cl)ccc(Cl)c2Cl)c(Cl)c1Cl
ENV	PCB126	Clc1ccc(-c2cc(Cl)c(Cl)c(Cl)c2)cc1Cl
ENV	PCB97	Clc1cc(Cl)c(-c2cccc(Cl)c2Cl)cc1Cl
ENV	PCB85	Clc1ccc(-c2ccc(Cl)c(Cl)c2Cl)c(Cl)c1
ENV	PCB79	Clc1cc(Cl)cc(-c2ccc(Cl)c(Cl)c2)c1
ENV	PCB77	C1=CC(=C(C=C1C2=CC(=C(C=C2)Cl)Cl)Cl)Cl
ENV	PCB66	Clc1ccc(-c2ccc(Cl)c(Cl)c2)c(Cl)c1
ENV	PCB64	Clc1ccc(-c2c(Cl)ccc(Cl)c2Cl)cc1
ENV	PCB208	Clc1cc(Cl)c(Cl)c(-c2c(Cl)c(Cl)c(Cl)c(Cl)c2Cl)c1Cl
ENV	PCB203	Clc1cc(Cl)c(-c2c(Cl)c(Cl)c(Cl)c(Cl)c2Cl)cc1Cl
ENV	PCB202	Clc1cc(Cl)c(Cl)c(-c2c(Cl)c(Cl)cc(Cl)c2Cl)c1Cl
ENV	PCB199	Clc1cc(-c2c(Cl)c(Cl)cc(Cl)c2Cl)c(Cl)c(Cl)c1Cl
ENV	PCB198	Clc1cc(Cl)c(Cl)c(-c2c(Cl)c(Cl)c(Cl)c(Cl)c2Cl)c1
ENV	PCB193	Clc1cc(-c2c(Cl)c(Cl)cc(Cl)c2Cl)cc(Cl)c1Cl
ENV	PCB187	Clc1cc(Cl)c(-c2c(Cl)c(Cl)cc(Cl)c2Cl)cc1Cl
ENV	PCB184	Clc1cc(Cl)c(-c2c(Cl)cc(Cl)c(Cl)c2Cl)c(Cl)c1
ENV	PCB183	Clc1cc(Cl)c(-c2c(Cl)cc(Cl)c(Cl)c2Cl)cc1Cl
ENV	PCB182	Clc1cc(Cl)c(-c2cc(Cl)c(Cl)c(Cl)c2Cl)c(Cl)c1
ENV	PCB180	Clc1cc(Cl)c(-c2cc(Cl)c(Cl)c(Cl)c2Cl)cc1Cl
ENV	PCB178	Clc1cc(Cl)c(Cl)c(-c2c(Cl)c(Cl)cc(Cl)c2Cl)c1
ENV	PCB175	Clc1cc(Cl)c(Cl)c(-c2c(Cl)cc(Cl)c(Cl)c2Cl)c1
ENV	PCB174	Clc1cc(-c2c(Cl)ccc(Cl)c2Cl)c(Cl)c(Cl)c1Cl
ENV	PCB172	Clc1cc(Cl)c(Cl)c(-c2cc(Cl)c(Cl)c(Cl)c2Cl)c1
ENV	PCB163	Clc1ccc(-

		c2c(Cl)c(Cl)cc(Cl)c2Cl)cc1Cl
ENV	PCB162	Clc1cc(Cl)c(Cl)c(-c2cc(Cl)c(Cl)c(Cl)c2)c1
ENV	PCB159	Clc1cc(Cl)cc(-c2cc(Cl)c(Cl)c(Cl)c2Cl)c1
ENV	PCB153	Clc1cc(Cl)c(-c2cc(Cl)c(Cl)cc2Cl)cc1Cl
ENV	PCB149	Clc1cc(Cl)c(-c2c(Cl)ccc(Cl)c2Cl)cc1Cl
ENV	PCB146	Clc1cc(Cl)c(Cl)c(-c2cc(Cl)c(Cl)cc2Cl)c1
ENV	PCB141	Clc1ccc(Cl)c(-c2cc(Cl)c(Cl)c(Cl)c2Cl)c1
ENV	PCB138	Clc1cc(Cl)c(-c2ccc(Cl)c(Cl)c2Cl)cc1Cl
ENV	PCB134	Clc1cccc(-c2c(Cl)c(Cl)cc(Cl)c2Cl)c1Cl
ENV	PCB130	Clc1cc(Cl)c(Cl)c(-c2ccc(Cl)c(Cl)c2Cl)c1
ENV	PCB127	Clc1cc(Cl)cc(-c2cc(Cl)c(Cl)c(Cl)c2)c1
ENV	PCB120	Clc1cc(Cl)cc(-c2cc(Cl)c(Cl)cc2Cl)c1
ENV	PCB118	Clc1ccc(-c2cc(Cl)c(Cl)cc2Cl)cc1Cl
ENV	PCB110	Clc1ccc(-c2c(Cl)ccc(Cl)c2Cl)cc1Cl
ENV	PCB108	Clc1cc(Cl)cc(-c2ccc(Cl)c(Cl)c2Cl)c1
ENV	PCB101	Clc1ccc(Cl)c(-c2cc(Cl)c(Cl)cc2Cl)c1
ENV	PCB3	Clc1ccc(-c2ccccc2)cc1
ENV	Perfluorooctane sulfonamide	NS(=O)(=O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	N-Ethyl-N-(2-hydroxyethyl)perfluorooctanesulfonamide (N-EtFOSE)	CCN(CCO)S(=O)(=O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	N-EtFOSA	CCNS(=O)(=O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	N-(2-hydroxyethyl)perfluorooctanesulfonamide (FOSE alcohol)	O=S(=O)(NCCO)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	8:2 PAP diester	O=P(O)(OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F)OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	8:2 mono-PAP	O=P(O)(O)OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	8:2 FTUCA	O=C(O)/C=C(\F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	8:2 FTOH Fluorotelomer alcohol	OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	8:2 FTCA fluorotelomer carboxylic acid	O=C(O)CC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	8:2 fluorotelomer aldehyde (8:2 FTAL)	O=CCC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F

ENV	8:2 Fluorotelomer Acrylate	C=CC(=O)OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	8:2 chlorinated polyfluoroalkyl ether sulfonate	O=S(=O)([O-])C(F)(F)C(F)(F)OC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)Cl
ENV	7:3-fluorotelomer acid (7:3-FTA)	O=C(O)/C=C/C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	6:2 PAP diester	O=P(O)(OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F)OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	6:2 mono-PAP	O=P(O)(O)OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	6:2 Fluorotelomer alcohol	OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	6:2 chlorinated polyfluoroalkyl ether sulfonate	O=S(=O)([O-])C(F)(F)C(F)(F)OC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)Cl
ENV	4:2 PAP diester	O=P(O)(OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)F)OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	4:2 mono-PAP	O=P(O)(O)OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	4:2 Fluorotelomer alcohol	OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	10:2 PAP diester	O=P(O)(OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F)OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	10:2 mono PAP	O=P(O)(O)OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	10:2 FTOH	OCCC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
ENV	Triclosan	Oc1cc(Cl)ccc1Oc1ccc(Cl)cc1Cl
ENV	Triclocarban	O=C(Nc1ccc(Cl)cc1)Nc1ccc(Cl)c(Cl)c1
ENV	Salicylic acid	O=C(O)c1ccccc1O
ENV	Propyl Paraben	CCCOC(=O)c1ccc(O)cc1
ENV	Octinoxate	CCCCC(CC)COC(=O)/C=C/c1ccc(OC)cc1
ENV	n-butyl paraben	CCCCOC(=O)c1ccc(O)cc1
ENV	Ketoconazole	CC(=O)N1CCN(c2ccc(OC[C@H]3CO[C@](Cn4ccnc4)(c4ccc(Cl)cc4Cl)O3)cc2)CC1
ENV	iso-butyl paraben	CC(C)COC(=O)c1ccc(O)cc1
Env	Hydroquinone	Oc1ccc(O)cc1
ENV	Heptyl paraben	CCCCCCCOC(=O)c1ccc(O)cc1
ENV	Gypsogenin	CC1(C)CCC2(C(=O)O)CCC3(C)C(=CCC4C5(C)CCC(O)C(C)(C=O)C5CCC

		43C)C2C1
ENV	Ferulic Acid	COc1cc(/C=C/C(=O)O)ccc1O
ENV	Farnesol	CC(C)=CCC/C(C)=C\CC/C(C)=C\CO
ENV	Ethyl Paraben	CCOC(=O)c1ccc(O)cc1
ENV	Cetyl alcohol	CCCCCCCCCCCCCCCCO
ENV	Benzyl Paraben	O=C(OCc1ccccc1)c1ccc(O)cc1
ENV	Benzyl alcohol	OCc1ccccc1
ENV	Benzoyl peroxide	O=C(OOC(=O)c1ccccc1)c1ccccc1
ENV	benzophenone-3	COc1ccc(C(=O)c2ccccc2)c(O)c1
ENV	benzophenone-2	O=C(c1ccc(O)cc1O)c1ccc(O)cc1O
ENV	Aminocaproic acid	NCCCCCC(=O)O
ENV	4-N-Octylphenol	CCCCCCCCc1ccc(O)cc1
ENV	1,2-dibromoethane	BrCCBr
ENV	Methyl paraben	COC(=O)C1=CC=C(C=C1)O
ENV	Warfarin	CC(=O)CC(c1ccccc1)c1c(O)c2ccccc2 oc1=O
ENV	Triphenyltin chloride	Cl[Sn](c1ccccc1)(c1ccccc1)c1ccccc1
ENV	Trifluralin	CCCN(CCC)c1c([N+](=O)[O-])cc(C(F)(F)F)cc1[N+](=O)[O-]
ENV	triflumuron	O=C(NC(=O)c1ccccc1Cl)Nc1ccc(OC(F)(F)F)cc1
ENV	Triadimefon	CC(C)(C)C(=O)C(Oc1ccc(Cl)cc1)n1cncn1
ENV	trans-Phenothrin	CC(C)=C[C@H]1[C@H](C(=O)OCc2cccc(Oc3ccccc3)c2)C1(C)C
ENV	TRANS-PERMETHRIN	CC1(C)[C@H](C=C(Cl)Cl)[C@H]1C(=O)OCc1cccc(Oc2ccccc2)c1
ENV	trans-Bromuconazole	Clc1ccc(C2(Cn3cncn3)CC(Br)CO2)c(Cl)c1
ENV	Tralomethrin	CC1(C)C(C(=O)OC(C#N)c2cccc(Oc3ccccc3)c2)C1C(Br)C(Br)(Br)Br
ENV	Tolclofos-methyl	COP(=S)(OC)Oc1c(Cl)cc(C)cc1Cl
ENV	Thiobencarb	CCN(CC)C(=O)SCc1ccc(Cl)cc1
ENV	Terbutryn	CCN=c1nc(NC(C)(C)C)[nH]c(SC)n1
ENV	terbuthylazine	CCN=c1nc(Cl)nc(NC(C)(C)C)[nH]1
ENV	Teflubenzuron	O=C(NC(=O)c1c(F)cccc1F)Nc1cc(Cl)c(F)c(Cl)c1F
ENV	Tebufenozide	CCc1ccc(C(=O)NN(C(=O)c2cc(C)cc(C)c2)C(C)(C)C)cc1
ENV	Sulprofos	CCCS[P@@](=S)(OCC)Oc1ccc(SC)cc1
ENV	Sulfuryl Fluoride	O=S(=O)(F)F
ENV	s-metolachlor	CCc1cccc(C)c1N(C(=O)CCl)[C@@H](C)COC
ENV	Simazine	CCN=c1nc(Cl)[nH]c(=NCC)[nH]1
ENV	Resmethrin	CC(C)=CC1C(C(=O)OCc2coc(Cc3ccccc3)c2)C1(C)C
ENV	Pyrithiobac	COC1=CC(=NC(=N1)SC2=C(C(=CC=C2)Cl)C(=O)O)OC
ENV	Pyridalyl	FC(F)(F)c1ccc(OCCCOc2c(Cl)cc(OC

		C=C(Cl)Cl)cc2Cl)nc1
ENV	Pyributicarb	COc1cccc(N(C)C(=S)Oc2cccc(C(C)(C)C)c2)n1
ENV	Pyribenzoxim	COc1cc(OC)nc(Oc2cccc(Oc3nc(OC)cc(OC)n3)c2C(=O)ON=C(c2ccccc2)c2ccccc2)n1
ENV	Propiconazole	CCCC1COC(Cn2cncn2)(c2ccc(Cl)cc2Cl)O1
ENV	Propazine	CC(C)Nc1nc(Cl)nc(NC(C)C)n1
ENV	Profenofos	CCCSP(=O)(OCC)Oc1ccc(Br)cc1Cl
ENV	Prochloraz	CCCN(CCOc1c(Cl)cc(Cl)cc1Cl)C(=O)n1ccnc1
ENV	Primicarb	Cc1nc(N(C)C)nc(OC(=O)N(C)C)c1C
ENV	Phorate	CCOP(=S)(OCC)SCSCC
ENV	Permethrin	CC1(C)C(C=C(Cl)Cl)C1C(=O)OCc1ccccc(Oc2ccccc2)c1
ENV	Pentachlorophenol	Oc1c(Cl)c(Cl)c(Cl)c(Cl)c1Cl
ENV	Parathion	CCOP(=S)(OCC)Oc1ccc([N+](=O)[O-])cc1
ENV	O,p'-DDD	Clc1ccc(C(c2ccccc2Cl)C(Cl)Cl)cc1
ENV	O,p′-DDT	Clc1ccc(C(c2ccccc2Cl)C(Cl)(Cl)Cl)cc1
ENV	N,N-diethyl-m-toluamide	CCN(CC)C(=O)c1cccc(C)c1
ENV	Myristicin	C=CCc1cc(OC)c2c(c1)OCO2
ENV	Molinate	CCSC(=O)N1CCCCCC1
ENV	Methyl parathion	COP(=S)(OC)Oc1ccc([N+](=O)[O-])cc1
ENV	Methyl bromide	CBr
ENV	Methoxychlor	COc1ccc(C(c2ccc(OC)cc2)C(Cl)(Cl)Cl)cc1
ENV	Methiozolin	Cc1ccsc1C1=NOC(C)(COCc2c(F)cccc2F)C1
ENV	metam-sodium	CNC(=S)[S-].[Na+]
ENV	METALAXYL	COCC(=O)N(c1c(C)cccc1C)C(C)C(=O)OC
ENV	Malathion	CCOC(=O)CC(SP(=S)(OC)OC)C(=O)OCC
ENV	linuron	CON(C)C(=O)Nc1ccc(Cl)c(Cl)c1
ENV	Kadethrin	CC1(C)[C@@H](C(=O)OCc2coc(Cc3ccccc3)c2)[C@H]1/C=C1\CCSC1=O
ENV	Imiprothrin	C#CCN1CC(=O)N(COC(=O)C2C(C=C(C)C)C2(C)C)C1=O
ENV	Imidacloprid	O=[N+]([O-])/N=C1\NCCN1Cc1ccc(Cl)nc1
ENV	Hexachlorobenzene	Clc1c(Cl)c(Cl)c(Cl)c(Cl)c1Cl
ENV	Heptachlor	ClC1=C(Cl)C2(Cl)C3C(Cl)C=CC3C1(Cl)C2(Cl)Cl
ENV	Furametpyr	Cc1nn(C)c(Cl)c1C(O)=Nc1cccc2c1C(C)OC2(C)C
ENV	Folpet	CC(Cl)(Cl)SN1C(=O)c2ccccc2C1=O

ENV	fipronil	N#Cc1nn(-c2c(Cl)cc(C(F)(F)F)cc2Cl)c(N)c1S(=O)C(F)(F)F
ENV	Fenvalerate	CC(C)C(C(=O)OC(C#N)c1cccc(Oc2ccccc2)c1)c1ccc(Cl)cc1
ENV	Fenthion	COP(=S)(OC)Oc1ccc(SC)c(C)c1
ENV	Fenpyroximate	Cc1nn(C)c(Oc2ccccc2)c1C=NOCc1ccc(C(=O)OC(C)(C)C)cc1
ENV	Etofenprox	CCOc1ccc(C(C)(C)COCc2cccc(Oc3ccccc3)c2)cc1
ENV	ethiprole	CCS(=O)c1c(C#N)nn(-c2c(Cl)cc(C(F)(F)F)cc2Cl)c1N
ENV	Esfenvalerate	CC(C)[C@H](C(=O)O[C@H](C#N)c1cccc(Oc2ccccc2)c1)c1ccc(Cl)cc1
ENV	EPTAM	CCCN(CCC)C(=O)SCC
ENV	Endosulfan-alpha	O=[S@]1OC[C@H]2[C@H](CO1)[C@@]1(Cl)C(Cl)=C(Cl)[C@]2(Cl)C1(Cl)Cl
ENV	D-trans-Tetramethrin	CC(C)=C[C@@H]1[C@@H](C(=O)OCN2C(=O)C3=C(CCCC3)C2=O)C1(C)C
ENV	Disulfoton	CCOP(=S)(OCC)SCCSCC
ENV	Dimethoate	CN=C(O)CSP(=S)(OC)OC
ENV	Diflubenzuron	O=C(NC(=O)c1c(F)cccc1F)Nc1ccc(Cl)cc1
ENV	Dieldrin	ClC1=C(Cl)[C@]2(Cl)[C@@H]3[C@@H]4C[C@H]([C@@H]3[C@@]1(Cl)C2(Cl)Cl)[C@H]1O[C@@H]41
ENV	Dichlorvos	COP(=O)(OC)OC=C(Cl)Cl
ENV	Dichlorophen	Oc1ccc(Cl)cc1Cc1cc(Cl)ccc1O
ENV	Dichlorodiphenyltrichloroethane	Clc1ccc(C(c2ccc(Cl)cc2)C(Cl)(Cl)Cl)cc1
ENV	Diazinon	CCOP(=S)(OCC)Oc1cc(C)nc(C(C)C)n1
ENV	Deltamethrin	CC1(C)[C@@H](C=C(Br)Br)[C@H]1C(=O)O[C@H](C#N)c1cccc(Oc2ccccc2)c1
ENV	Dehydroacetic acid	CC(=O)C1C(=O)C=C(C)OC1=O
ENV	Cyprodinil	Cc1cc(C2CC2)nc(Nc2ccccc2)n1
ENV	Cyhalothrin	CC1(C)C(/C=C(\Cl)C(F)(F)F)C1C(=O)OC(C#N)c1cccc(Oc2ccccc2)c1
ENV	Cyfluthrin	CC1(C)C(C=C(Cl)Cl)C1C(=O)OC(C#N)c1ccc(F)c(Oc2ccccc2)c1
ENV	Cycloprothrin	CCOc1ccc(C2(C(=O)OC(C#N)c3ccccc(Oc4ccccc4)c3)CC2(Cl)Cl)cc1
ENV	Clothianidin	CN/C(=N\[N+](=O)[O-])NCc1cnc(Cl)s1
ENV	cis-Dibromochrysanthemic acid	CC1(C)C(C=C(Br)Br)C1C(=O)O
ENV	Cinmethylin	Cc1ccccc1CO[C@@H]1C[C@@]2(C(C)C)CC[C@]1(C)O2

ENV	Chlorpyrifos	CCOP(=S)(OCC)Oc1nc(Cl)c(Cl)cc1Cl
ENV	Chlordane	ClC1=C(Cl)C2(Cl)C3C(Cl)C(Cl)CC3C1(Cl)C2(Cl)Cl
ENV	Chlorantraniliprole (DPX-E2Y45)	CNC(=O)c1cc(Cl)cc(C)c1NC(=O)c1cc(Br)nn1-c1ncccc1Cl
ENV	Carbosulfan	CCCCN(CCCC)SN(C)C(=O)Oc1cccc2c1OC(C)(C)C2
ENV	Carbofuran	CNC(=O)Oc1cccc2c1OC(C)(C)C2
ENV	Carbaryl	CNC(=O)Oc1cccc2ccccc12
ENV	Captan	O=C1C2CC=CCC2C(=O)N1SC(Cl)(Cl)Cl
ENV	Butachlor	CCCCOCN(C(=O)CCl)c1c(CC)cccc1CC
ENV	Bioresmethrin	CC(C)=C[C@@H]1[C@@H](C(=O)OCc2coc(Cc3ccccc3)c2)C1(C)C
ENV	Bifenthrin	Cc1c(COC(=O)[C@@H]2[C@H](/C=C(\Cl)C(F)(F)F)C2(C)C)cccc1-c1ccccc1
ENV	Benfuracarb	CCOC(=O)CCN(SN(C)C(=O)Oc1cccc2c1OC(C)(C)C2)C(C)C
ENV	Benalaxyl	COC(=O)C(C)N(C(=O)Cc1ccccc1)c1c(C)cccc1C
ENV	Azinphos methyl	COP(=S)(OC)SCn1nnc2ccccc2c1=O
ENV	Atrazine	CCNc1nc(Cl)nc(NC(C)C)n1
ENV	Ametryn	CCN=c1nc(SC)[nH]c(=NC(C)C)[nH]1
ENV	Allethrin	C=CCC1=C(C)C(OC(=O)C2C(C=C(C)C)C2(C)C)CC1=O
ENV	Aldrin	ClC1=C(Cl)[C@]2(Cl)[C@H]3[C@H]([C@@H]4C=C[C@H]3C4)[C@@]1(Cl)C2(Cl)Cl
ENV	Aldicarb	CNC(=O)ON=CC(C)(C)SC
ENV	Alachlor	CCc1cccc(CC)c1N(COC)C(=O)CCl
ENV	Acrolein	C=CC=O
ENV	Acetachlor	CCOCN(C(=O)CCl)c1c(C)cccc1CC
ENV	3-trifluoromethyl-4-nitrophenol	O=[N+]([O-])c1ccc(O)cc1C(F)(F)F
ENV	3-Phenoxybenozoic acid	O=C(O)c1cccc(Oc2ccccc2)c1
ENV	2-Phenylphenol	Oc1ccccc1-c1ccccc1
ENV	2,6-diisopropylnaphthalene	CC(C)c1ccc2cc(C(C)C)ccc2c1
ENV	2,6-dichlorobenzonitrile	N#Cc1c(Cl)cccc1Cl
ENV	2,4,6-tribromophenol	Oc1c(Br)cc(Br)cc1Br
ENV	1,4-dichlorobenzene	Clc1ccc(Cl)cc1
ENV	1,3-dichloropropene	Cl/C=C/CCl
ENV	(1S)-trans-Cyfluthrin	CC1(C)[C@@H](C(=O)OC(C#N)c2ccc(F)c(Oc3ccccc3)c2)[C@@H]1C=C(Cl)Cl
ENV	(1R,4S)-trans-prallethrin	C#CCC1=C(C)[C@@H](OC(=O)[C@@H]2[C@@H](C=C(C)C)C2(C)C)CC1=O
ENV	(1R,4S)-cis-prallethrin	C#CCC1=C(C)[C@@H](OC(=O)[C@

		@H]2[C@H](C=C(C)C)C2(C)C)CC1=O
ENV	(+)-fipronil	N#Cc1nn(-c2c(Cl)cc(C(F)(F)F)cc2Cl)c(N)c1[S@](=O)C(F)(F)F
ENV	(-)-cis-Phenothrin	CC(C)=C[C@@H]1[C@H](C(=O)OCc2cccc(Oc3ccccc3)c2)C1(C)C
ENV	4-aminobenzoate	Nc1ccc(C(=O)[O-])cc1
ENV	3-Ethylphenol	CCc1cccc(O)c1
ENV	Di-n-octyl phthalate	CCCCCCCCOC(=O)c1ccccc1C(=O)OCCCCCCCC
ENV	Dimethyl phthalate	COC(=O)c1ccccc1C(=O)OC
ENV	Diisononyl-cyclohexane-1,2-dicarboxylate	CC(C)CCCCCCOC(=O)C1CCCCC1C(=O)OCCCCCCC(C)C
ENV	Diisobutyl phthalate	CC(C)COC(=O)c1ccccc1C(=O)OCC(C)C
ENV	Dihexyl phthalate	CCCCCCOC(=O)c1ccccc1C(=O)OCCCCCC
ENV	Diethyl phthalate	CCOC(=O)c1ccccc1C(=O)OCC
ENV	Dibutyl phthalate	CCCCOC(=O)c1ccccc1C(=O)OCCCC
ENV	Di-2-ethylhexyl terephthalate	CCCCC(CC)COC(=O)c1ccc(C(=O)OCC(CC)CCCC)cc1
ENV	Di-2-ethylhexyl Adipate	CCCCC(CC)COC(=O)CCCCC(=O)OCC(CC)CCCC
ENV	Di-(2-ethylhexyl)phthalate	CCCCC(CC)COC(=O)c1ccccc1C(=O)OCC(CC)CCCC
ENV	Bisphenol Z	Oc1ccc(C2(c3ccc(O)cc3)CCCCC2)cc1
ENV	Bisphenol S	O=S(=O)(c1ccc(O)cc1)c1ccc(O)cc1
ENV	Bisphenol F diglycidyl ether	c1cc(OCC2CO2)ccc1Cc1ccc(OCC2CO2)cc1
ENV	Bisphenol F	Oc1ccc(Cc2ccc(O)cc2)cc1
ENV	Bisphenol B	CCC(C)(c1ccc(O)cc1)c1ccc(O)cc1
ENV	Bisphenol AF	Oc1ccc(C(c2ccc(O)cc2)(C(F)(F)F)C(F)(F)F)cc1
ENV	Bisphenol A glycidyl methacrylate	C=C(C)C(=O)OCC(O)COc1ccc(C(C)(C)c2ccc(OCC(O)COC(=O)C(=C)C)cc2)cc1
ENV	Bisphenol A	CC(C)(c1ccc(O)cc1)c1ccc(O)cc1
ENV	bis(2-propylheptyl) phthalate	CCCCCC(CCC)COC(=O)c1ccccc1C(=O)OCC(CCC)CCCCC
ENV	Benzyl butyl phthalate	CCCCOC(=O)c1ccccc1C(=O)OCc1ccccc1
ENV	Acrylonitrile	C=CC#N
ENV	3,3′-Dimethylbisphenol A	Cc1cc(C(C)(C)c2ccc(O)c(C)c2)ccc1O
ENV	2-Ethylhexyl-2,3,4,5-tetrabromobenzoate	CCCCC(CC)COC(=O)c1cc(Br)c(Br)c(Br)c1Br
ENV	VINYL CHLORIDE	CCCCOP(=O)(OCCCC)OCCCC
ENV	tri-N-butyltin	CCCC[SnH](CCCC)CCCC
ENV	Trichloroethylene	ClC=C(Cl)Cl
ENV	Trichloroethane	CC(Cl)(Cl)Cl

ENV	Tributyltin chloride	CCCC[Sn](Cl)(CCCC)CCCC
ENV	tert-butyl methyl ether	COC(C)(C)C
ENV	tert-butyl ethyl ether (ETBE)	CCOC(C)(C)C
ENV	tert-amyl methyl ether	CCC(C)(C)OC
ENV	styrene	C=Cc1ccccc1
ENV	quinoline	c1ccc2ncccc2c1
ENV	p-xylene	Cc1ccc(C)cc1
ENV	p-Toluidine	Cc1ccc(N)cc1
ENV	pentanol	CCCCCO
ENV	o-xylene	Cc1ccccc1C
ENV	Nonane	CCCCCCCCC
ENV	N-Nitrosodipropylamine	CCCN(CCC)N=O
ENV	n-methylformamide	CN=CO
ENV	Nitrosodiethylamine	CCN(CC)N=O
ENV	n,n-diethylformamide	CCN(C=O)CC
ENV	m-xylene	Cc1cccc(C)c1
ENV	Methylene chloride	ClCCl
ENV	isobutanol	CC(C)CO
ENV	hexane	CCCCCC
ENV	hexamethylphosphoramide	CN(C)P(=O)(N(C)C)N(C)C
ENV	Furan	c1ccoc1
ENV	ethylmethylnitrosamine	CCN(C)N=O
ENV	Dibutyltin dichloride	CCCC[Sn](Cl)(Cl)CCCC
ENV	cumene hydroperoxide	CC(C)(OO)c1ccccc1
ENV	butanol	CCCCO
ENV	butadiene	C=CC=C
ENV	beta-methyl styrene	C/C=C/c1ccccc1
ENV	benzene	c1ccccc1
ENV	Aniline	Nc1ccccc1
ENV	acetaldehyde	CC=O
ENV	4-vinylcyclohexene	C=CC1CC=CCC1
ENV	4-tert-Octylphenol	CC(C)(C)CC(C)(C)c1ccc(O)cc1
ENV	4-tert-Butylphenol	CC(C)(C)c1ccc(O)cc1
ENV	4-sec-Butylphenol	CCC(C)c1ccc(O)cc1
ENV	4-n-Propylphenol	CCCc1ccc(O)cc1
ENV	4-Isopropylphenol	CC(C)c1ccc(O)cc1
ENV	4-Ethylphenol	CCc1ccc(O)cc1
ENV	2-propanol	CC(C)O
ENV	2-nitrofluorene	O=[N+]([O-])c1ccc2c(c1)Cc1ccccc1-2
ENV	2-Isopropylnaphthalene	CC(C)c1ccc2ccccc2c1
ENV	2-ethylaniline	CCc1ccccc1N
ENV	2,6-Xylidine	Cc1cccc(C)c1N
ENV	2,5-dimethylaniline	Cc1ccc(C)c(N)c1
ENV	2,4-dimethylaniline	Cc1ccc(N)c(C)c1
ENV	2,4-Dichloroaniline	Nc1ccc(Cl)cc1Cl
ENV	2,3-dimethylaniline	Cc1cccc(N)c1C

ENV	2,2-dichloro-1,1,1-trifluoroethane	FC(F)(F)C(Cl)Cl
ENV	1-iodopropane	CCCI
ENV	1-iodopentane	CCCCCI
ENV	1-iodooctane	CCCCCCCCI
ENV	1-iodoheptane	CCCCCCCI
ENV	1-iodoethane	CCI
ENV	1,3-Dinitrobenzene	O=[N+]([O-])c1cccc([N+](=O)[O-])c1
ENV	1,2-dichlorobenzene	Clc1ccccc1Cl
ENV	1,2,4-trichlorobenzene	Clc1ccc(Cl)c(Cl)c1
ENV	tert-Butanol	CC(C)(C)O
Drug	Zopiclone	CN1CCN(CC1)C(=O)OC1N(C(=O)C2=NC=CN=C12)C1=NC=C(Cl)C=C1
Drug	Nordazepam	ClC1=CC=C2NC(=O)CN=C(C3=CC=CC=C3)C2=C1
Drug	Guanfacine	NC(=N)NC(=O)CC1=C(Cl)C=CC=C1Cl
Drug	Topiroxostat	N#CC1=NC=CC(=C1)C1=NNC(=N1)C1=CC=NC=C1
Drug	Tarenflurbil	C[C@@H](C(O)=O)C1=CC(F)=C(C=C1)C1=CC=CC=C1
Drug	sulfadiazine	Nc1ccc(S(=O)(=O)Nc2ncccn2)cc1
ENV	Pyridine N-oxide	O[n+]1ccccc1
Drug	Procaine	CCN(CC)CCOC(=O)C1=CC=C(N)C=C1
ENV	PhIP	CN1C2=C(N=CC(=C2)C3=CC=CC=C3)N=C1N
ENV	Demethoxycurcumin	COc1cc(/C=C/C(=O)CC(=O)/C=C/c2ccc(O)cc2)ccc1O

Appendix B

Text B1. Metabolite identification and their Quantification in *In Vitro* Samples.

We have identified BPA, BPAS, BPAG, TCS, BPS, and TBBPA using authentic standards and quantified these compounds using the isotope-spiking internal standard method. Specifically, for Bisphenol A (BPA), we purchased pure standard (Sigma, SG) and isotopic labeled standard Bisphenol A (Ring-$^{13}C_{12}$, 99%) from Cambridge Isotope Laboratories (MA, U.S.A) as internal standard. For BPA sulfate and BPA glucuronide, bisphenol A monosulfate sodium salt (native and D6 labeled form), and bisphenol A β-D glucuronide (native and ^{13}C labeled form) were purchased from Toronto Research Chemicals (North York, Canada). For TBBPA and TCS, we have used ^{13}C-TCS (Sigma, SG) as the internal standard. For other identified metabolites

without authentic standards, we confirmed their identity with exact m/z, MS/MS fragmentations, and our in-house retention time prediction models (Yang et al., unpublished data). The MS/MS fragmentation were also compared with in-silico prediction (CFM-ID method) or previously published data.(Allen, Pon, Wilson, Greiner, & Wishart, 2014) For the quantification of those metabolites without standards, we only used responses of ion count in the study and compare to the control. In order to avoid any error in the compound identification, we have only marked and quantified the metabolites with high confidence (i.e., matching accurate m/z; MS/MS fragmentation pattern; and retention time prediction or confirmed with authentic standards). For those with poor MS/MS quality due to the low abundance, we only left them as putative compounds.

For the biotransformation pathways, we first identified all possible metabolites using a prediction tool named SyGMa which is a python library for the Systematic Generation of potential Metabolites and then mapped them into the pathways reported in previous studies (Ho et al. 2017; Michalowicz 2014; Oh et al. 2018; Okuda et al. 2011; Ridder and Wagener 2008; Schauer et al. 2006; Song et al. 2018; Wu et al. 2009; Yoshihara et al. 2001; Yoshihara et al. 2004; Zalko et al. 2006). In this study, we investigated the metabolism of BPA *in vitro* in rat liver microsome and HepG2 cell line. With incubation of RLM, BPA was metabolized into the glucuronide conjugate, sulfate conjugate, and several other metabolites produced by cytochrome P450 dependent pathways, including hydroxy products of BPA, BPA sulfate, and BPA glucuronide. Figure B1 presents the proposed metabolic pathways of BPA composed of identified metabolites in this study, whereas all their instrumental parameters are illustrated in Table B1 and MS/MS spectrum in Figure B2. Briefly, BPA, BPA sulfate, and BPA glucuronide were confirmed with authentic standards. Hydroxy-BPA sulfate, (E)-4,4'-(4-methylpent-2-ene-2,4-diyl)diphenol, hydroxy-BPA, 4-methyl-2,4-bis(4-hydroxyphenyl)pent-1-ene were confirmed with MS/MS fragmentation. For those with poor MS/MS quality due to the low abundance, we only left them as putative compounds. The metabolites confirmed with high confidence are marked with asterisk. Consistent with previous findings(Battal et al. 2014; Gerona et al. 2013; Liu et al. 2017), glucuronidation and sulfonation were determined to be major pathways of BPA. Bisphenol A glucuronide

(BPAG) was the most abundant metabolite with incubation of RLM, as rat liver microsomes contain massive UDP-glucuronosyltransferase (UGT). The V_{max} for BPA glucuronidation was determined to be 3.30 ± 0.16 nmol/ (min·mg), and the K_m was 71.28 ± 9.20 μM. In the cell extracts, sulfate conjugate of BPA catalyzed by sulfotransferases (SULTs) was determined to be the most major metabolite, consistent with one previous study. (Bursztyka et al. 2008) The V_{max} for BPA sulfonation was determined to be 1.41 ± 0.13 ng/ (hr·10^6 cells), and the K_m was 27.88 ± 5.20 μM. In addition to hydroxy products like hydroxy BPAS, one additional metabolite catalyzed by enzyme CYP450 is isopropenylphenol, containing a pair of isomers (E)-4,4'-(4-methylpent-2-ene-2,4-diyl) diphenol (MBP) and 4-methyl-2,4-bis(p-hydroxyphenyl) pent-1-ene. MBP has been reported to be up to 1,000 times more potent as an estrogen than BPA (Moreman et al. 2018; Yoshihara et al. 2004).

For Bisphenol S (BPS), Triclosan (TCS) and Tetrabromobisphenol A (TBBPA), the identification follows the same method mentioned above. Overall, sulfate conjugate and glucuronide conjugate were the major metabolites as well for all of them. Besides, hydroxy forms of substrates, their sulfate conjugates and their glucuronide conjugates have also been identified in this work. Their proposed metabolic pathways are presented by Figure B3, Figure B5 and Figure B7, whereas all their instrumental parameters are illustrated in Table B2-4 and all their MS/MS spectrums in Figure B4, Figure B6 and Figure B8.

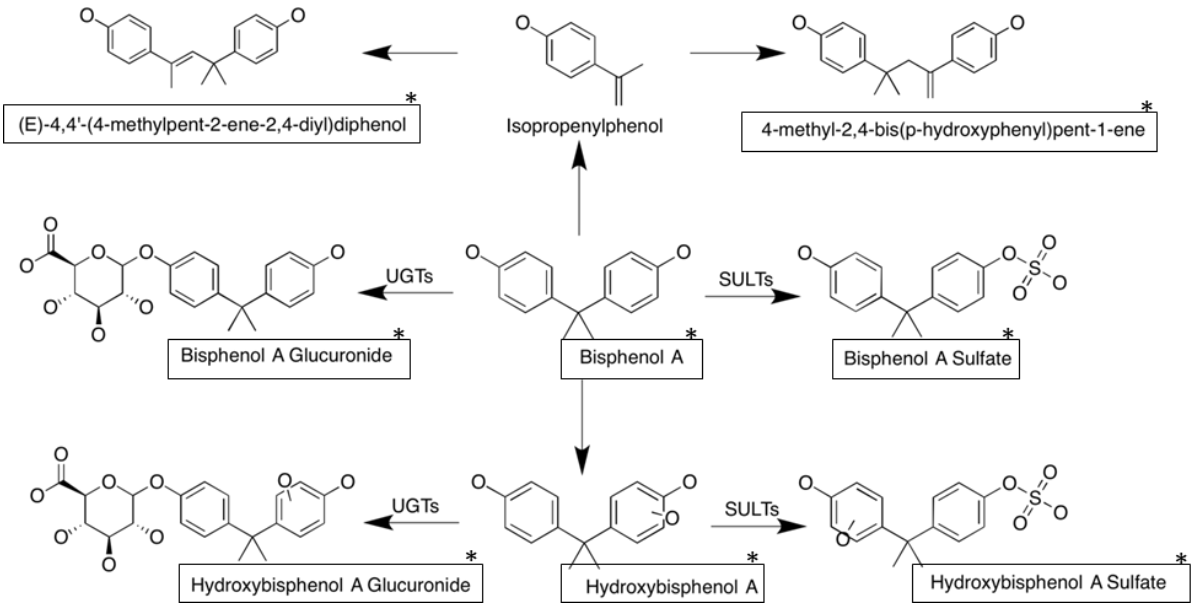


Figure B1. Proposed metabolic pathways of Bisphenol A *in vitro* and several relevant

enzymes. The metabolites confirmed with high confidence are marked with asterisk. For those with poor MS/MS quality due to the low abundance, we only left them as putative compounds.

Table B1. Retention time (RT) and molecular formulae of BPA and its metabolites

Compound	**Molecular formula**	**Molecular mass (M)a**	**$[M–H]^-$ Theoretical**	**$[M–H]^-$ Experimental**	**Error (ppm)**	**RT**
BPA	$C_{15}H_{16}O_2$	228.29	227.1078	227.1081	-1.32	9.217
BPA sulfate	$C_{15}H_{16}O_5S$	308.25	307.0646	307.0646	0.00	7.418
BPA glucuronide	$C_{21}H_{24}O_8$	404.32	403.1398	403.1406	-1.98	6.806
Hydroxy-BPA sulfate	$C_{15}H_{16}O_6S$	324.24	323.0595	323.0594	0.31	7.402
Hydroxy-BPA glucuronide	$C_{21}H_{25}O_8$	420.31	419.1398	419.1347	12.17	6.756
Hydroxy-BPA	$C_{15}H_{16}O_3$	244.29	243.1027	243.1033	-2.47	8.659
Isopropenylphenol	$C_9H_{10}O$	134.18	133.0659	133.0660	-0.75	9.206
(*E*)-4,4'-(4-methylpent-2-ene-2,4-diyl)diphenol	$C_{18}H_{20}O_2$	268.15	267.1391	267.1387	1.50	10.130
4-methyl-2,4-bis(4-hydroxyphenyl)pent-1-ene	$C_{18}H_{20}O_2$	268.15	267.1391	267.1386	1.87	10.841

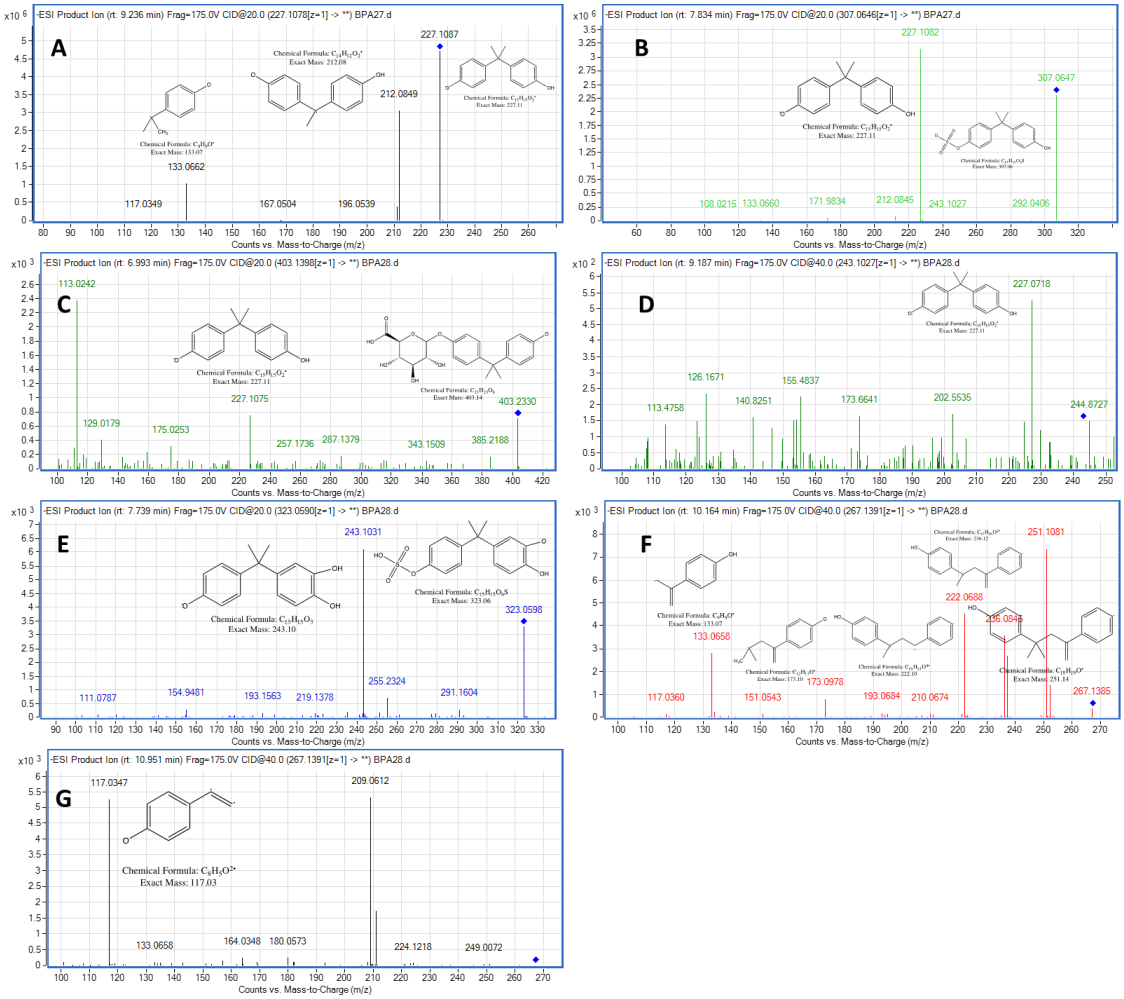


Figure B2. MS/MS spectrum examples of BPA and its metabolites (A) MS/MS spectrum of BPA; (B) MS/MS spectrum of BPAS; (C) MS/MS spectrum of BPAG; (D) MS/MS spectrum of BPAOH; (E) MS/MS spectrum of BPASOH; (F) MS/MS spectrum of (*E*)-4,4'-(4-methylpent-2-ene-2,4-diyl)diphenol; (G) MS/MS spectrum of 4-methyl-2,4-bis(4-hydroxyphenyl)pent-1-ene

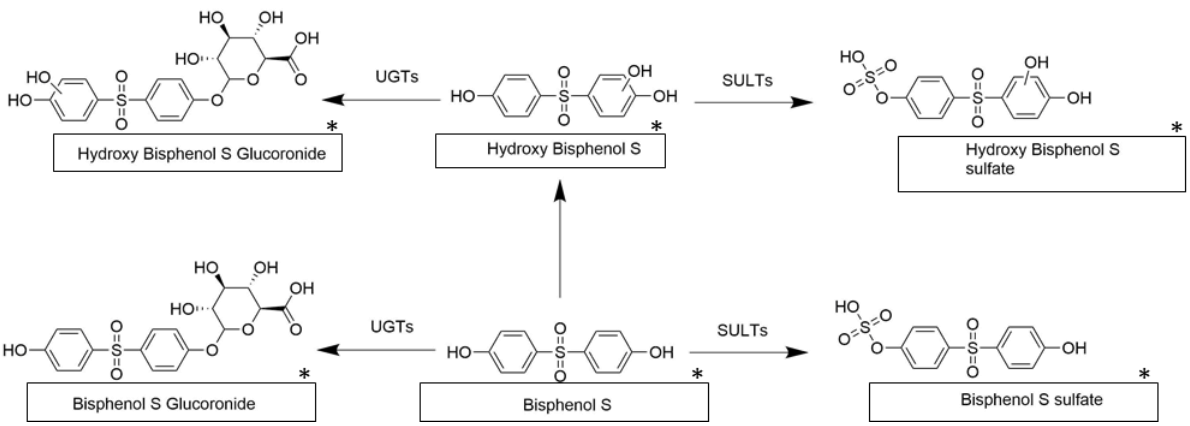


Figure B3. Proposed metabolic pathways of Bisphenol S *in vitro* and several relevant

enzymes. The metabolites confirmed with high confidence are marked with asterisk. For those with poor MS/MS quality due to the low abundance, we only left them as putative compounds.

Table B2. Retention time (RT) and molecular formulae of BPS and its metabolites

Compound	Molecular formula	Molecular mass (g/moL)	$[M–H]^-$ Theoretical	$[M–H]^-$ Experimental	Error (ppm)	RT
BPS	$C_{12}H_{10}O_4S$	250.28	249.0227	249.0223	1.61	7.809
BPS sulfate	$C_{12}H_{10}O_7S_2$	330.23	328.9795	328.9803	-2.43	6.700
BPS glucuronide	$C_{12}H_{18}O_{10}S$	426.31	425.0584	425.0565	4.47	6.187
Hydroxy-BPS sulfate	$C_{12}H_{11}O_7S_2$	346.22	344.9744	344.9723	6.09	7.758
Hydroxy-BPS glucuronide	$C_{12}H_{18}O_{11}S$	442.30	441.0497	441.0517	-4.53	8.205
Hydroxy-BPS	$C_{12}H_{10}O_5S$	266.27	265.0176	265.0154	8.30	7.983

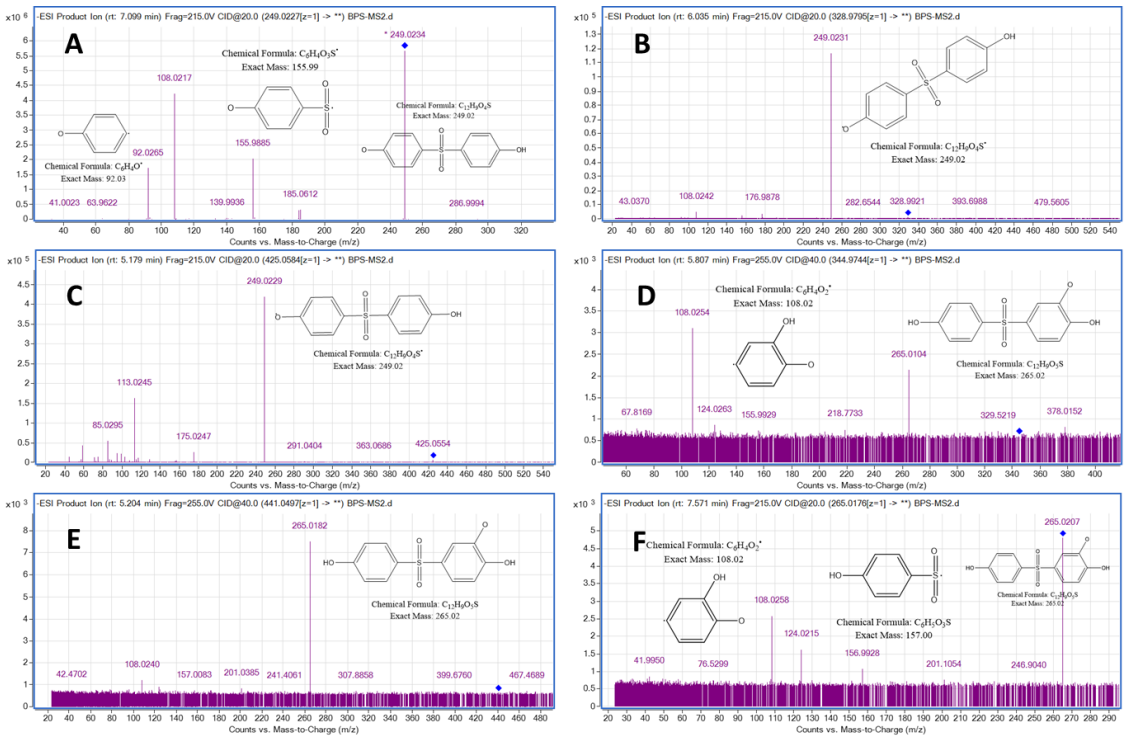


Figure B4. MS/MS spectrum examples of BPS and its metabolites. (A) MS/MS spectrum of BPS; (B) MS/MS spectrum of BPS-S; (C) MS/MS spectrum of BPS-G;

(D) MS/MS spectrum of BPS-SOH; (E) MS/MS spectrum of BPS-GOH; (F) MS/MS spectrum of BPS-OH.

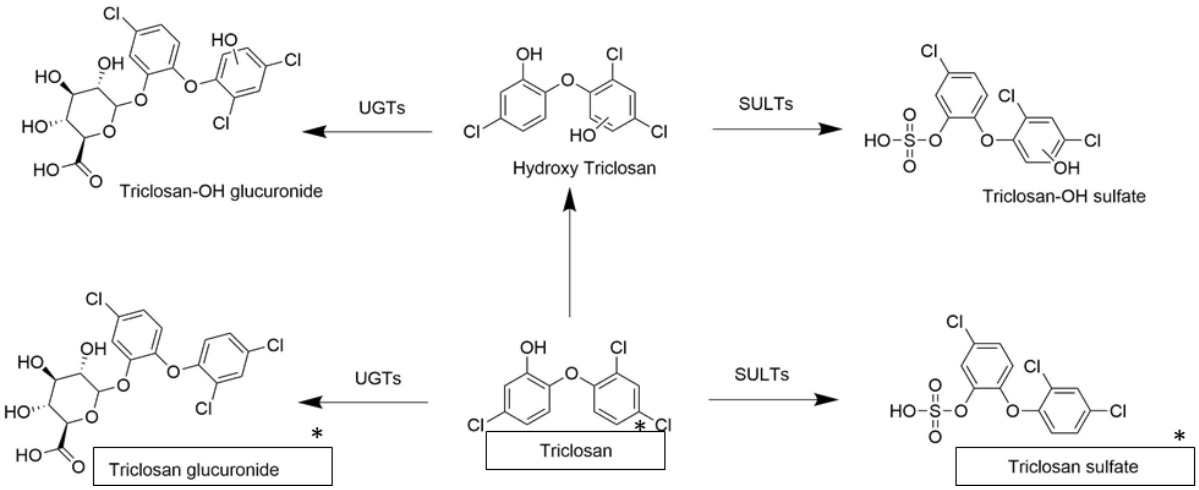


Figure B5. Proposed metabolic pathways of Triclosan *in vitro* and several relevant enzymes.

The metabolites confirmed with high confidence are marked with asterisk. For those with poor MS/MS quality due to the low abundance, we only left them as putative compounds.

Table B3. Retention time (RT) and molecular formulae of TCS and its metabolites.

Compound	**Molecular formula**	**Molecular mass (g/moL)**	**$[M–H]^-$ Theoretical**	**$[M–H]^-$ Experimental**	**Error (ppm)**	**RT**
TCS	$C_{12}H_7Cl_3O_2$	289.54	286.9439	286.9448	-3.14	11.316
TCS sulfate	$C_{12}H_7Cl_3O_5S$	369.49	368.8979	368.8979	0.00	8.649
TCS glucuronide	$C_{18}H_{15}Cl_3O_8$	465.57	462.9760	462.9768	-1.73	7.556
Hydroxy-TCS sulfate	$C_{12}H_7Cl_3O_6S$	385.49	384.8928	384.8931	-0.78	7.793
Hydroxy-TCS glucuronide	$C_{18}H_{15}Cl_3O_9$	481.56	478.9709	478.9705	0.84	6.976
Hydroxy-TCS	$C_{12}H_7Cl_3O_3$	305.53	302.9388	302.9388	0.00	9.908
2,4-Dichlorophenol	$C_6H_4Cl_2O$	163.00	160.9566	160.9567	-0.62	7.244

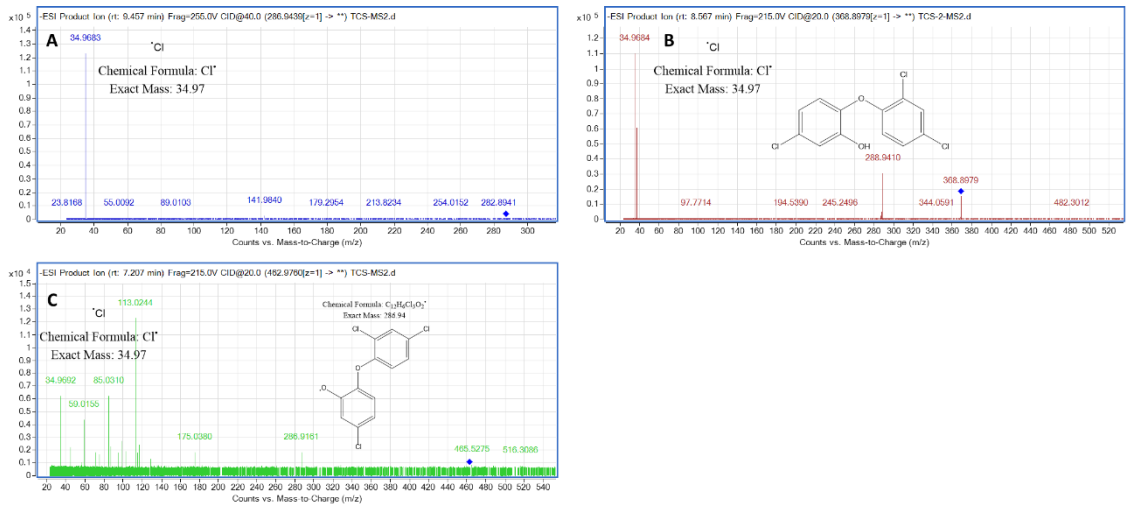


Figure B6. MS/MS spectrum examples of TCS and its metabolites. (A) MS/MS spectrum of TCS; (B) MS/MS spectrum of TCS-S; (C) MS/MS spectrum of TCS-G.

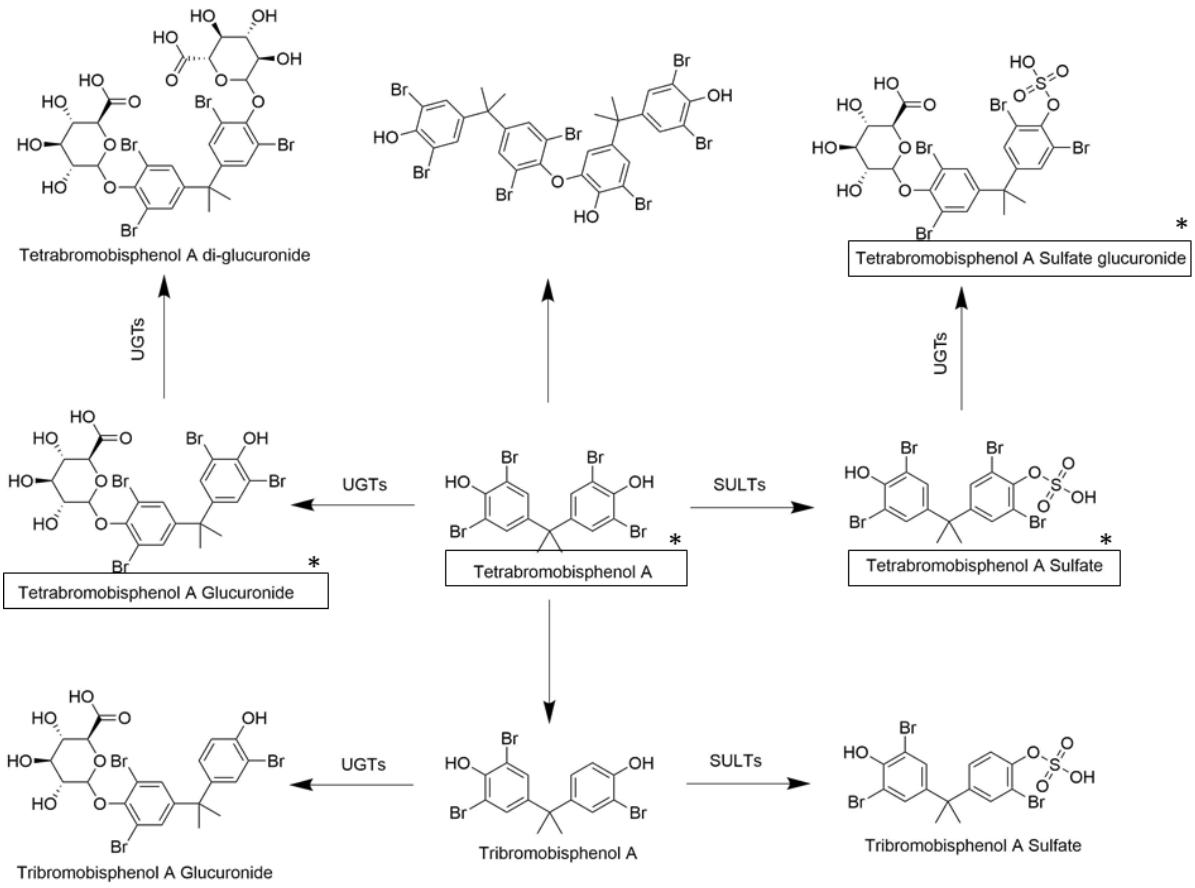


Figure B7. Proposed metabolic pathways of TBBPA *in vitro* and several relevant enzymes.

The metabolites confirmed with high confidence are marked with asterisk. For those with poor MS/MS quality due to the low abundance, we only left them as putative

compounds.

Table B4. Retention time (RT) and molecular formulae of TBBPA and its metabolites.

Compound	Molecular formula	Molecular mass (g/moL)	$[M–H]^-$ Theoretical	$[M–H]^-$ Experimental	Error (ppm)	RT
TBBPA	$C_{15}H_{12}Br_4O_2$	543.88	542.7458	542.7454	0.74	11.224
TBBPA sulfate	$C_{15}H_{12}Br_4O_5S$	623.83	622.7026	622.7019	1.12	8.414
TBBPA glucuronide	$C_{21}H_{20}Br_4O_8$	719.91	718.7780	718.7778	0.28	7.394
Tribromobisphenol A	$C_{15}H_{13}Br_3O_2$	463.97	462.8373	462.8368	1.08	10.738
Tribromobisphenol A glucuronide	$C_{21}H_{21}Br_3O_8$	640.00	640.8676	640.8670	0.94	7.410
TBBPA-glucuronide/sulfate	$C_{21}H_{20}Br_4O_{11}S$	719.96	798.7348	798.7334	1.75	6.599
TBBPA diglucuronide	$C_{27}H_{28}Br_4O_{14}$	895.94	894.8103	894.8095	0.89	6.218

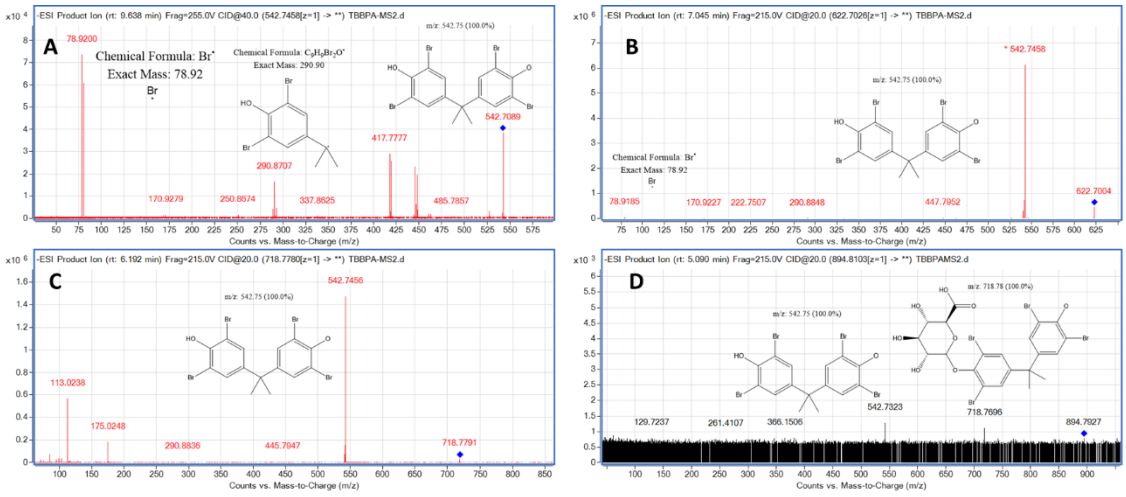


Figure B8. MS/MS spectrum examples of TBBPA and its metabolites (A) MS/MS spectrum of TBBPA; (B) MS/MS spectrum of TBBPA-S; (C) MS/MS spectrum of TBBPA-G; (D) MS/MS spectrum of TBBPA-SG.

Text B2. BPA Quantification in mice plasma samples

We first measured the authentic volume of mice plasma samples (V_{plasma}). If the V_{plasma} is less than 100 ul, saline (0.9%) were added to the sample until V_{plasma} = 100

ul; if the V_{plasma} is more than 100 ul, 100 ul sample were transferred to corresponding EP tube. 200 ul methanol (MeOH) were added in each sample for protein removal. The samples were then centrifuged at 13,000 rpm for 10 min in 15°C, and the supernatant was collected to 15 ml PP tube and dried with Nitrogen. Subsequent to an addition of 50 ul MeOH: H2O (1:1, v/v) solvent, the samples were further transferred to another new EP tube and centrifuged at 13,000 rpm for 10 min in 15°C. After spiking 500 ppb isotope internal standards (BPA-d16), the samples were collected for further analysis.

Analyses of mice plasma samples were performed using an ExionLC AD UHPLC system coupled with an AB Sciex Q TRAP 5500 MS (AB SCIEX, Concord, ON). Samples were analyzed with an Extend-C18 column (Narrow Bore RR 2.1 × 100 mm, 3.5 μm, 80 Å; Agilent Technologies) for MS analysis in MRM mode. The LC mobile phases were water containing 2 mM of ammonium acetate (A) and 100% methanol (B). The mobile phase flow rate was 0.5 mL/min and the gradient was as follows: 0-0.5 min: 90% of A; increased to 50% B at 1 min, 99% B at 7 min and maintained to 10 min. The mass spectrometer parameters were: CUR: 20 psi, IonSpray voltage: -4500 V, TEM: 550 °C, GS1: 55 psi, GS2: 55 psi. Table B5 listed the detailed MS/MS information of BPA and BPA-d16 while Figure B9 presented their representative chromatograms.

Table B5. Detailed MS/MS information of BPA and BPA-d16

	MRM transitions (m/z)	Declustering potential (DP, volts)	Collision energy (CE, volts)
BPA-d16	241.0/142.2	120	-40
BPA	226.8/132.9	-30	-34

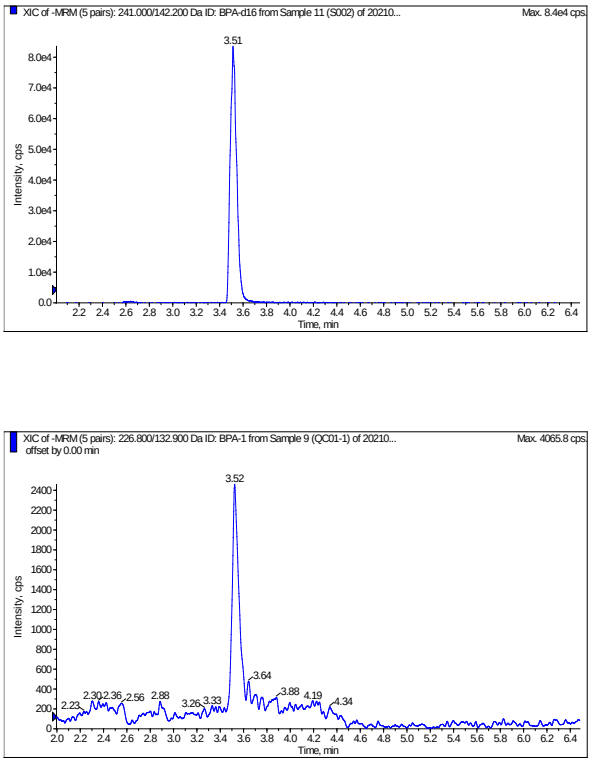


Figure B9. The representative chromatograms of BPA and BPA-d16.

Text B3. Reverse transcription polymerase chain reaction (RT-PCR) in chapter 5. Primers that were used in RT-PCR are listed in Table B6. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as the house-keeping gene for the expression normalization. Specific primer sets (listed in Table B6) used in this study for real-time PCR have been previously published and confirmed by NCBI PRIMER BLAST (Hanet et al. 2008).

Table B6. RNA primers

Gene Name	Forward primer 5'→3'	Reverse primer 5'→3'	Fragment size
SULT 1A1	GAGTTCAAAGCCCCAGGGATT	ACCTTGGCCATGTGGTAGAAG	197
SULT 1E1	AGAGGAGCTTGTGGACAGGA	GGCGACAATTTCTGGTTCAT	116
MDR 1	GTCCCAGGAGCCCATCCT	CCCGGCGTTGTCTCCATA	70
GADPH	GCCATCAATGACCCCTTCATT	TTGACGGTGCCATGGAATTT	91

Text B4. Reverse transcription polymerase chain reaction (RT-PCR) in chapter 6. The gene expression of UDP-Glucuronosyltransferases 1A1 (*Ugt1a1*), UDP-

Glucuronosyltransferases 2B1 (*Ugt2b1*) and Sulfotransferase (*Sult1a1*), which code the major enzymes of BPA biotransformation, were quantified using real-time PCR method. Total RNA from mouse livers was isolated using TRIzol reagent (Ambion, Thermofisher, MA, USA) and Homogenizer (Servicebio, China). Then 500 ng total RNA was reverse transcribed to cDNA using reverse transcription kit (Accurate Biology, China). RT-PCR assay was performed with SYBR Green Real-Time PCR Master Mix reagent (Accurate Biology, China) on LightCycler 96 Real-Time PCR System. The liver actin gene (Gapdh, LOT_H308KA9138) was used as the internal reference. The sequences of the primers used for RT-PCR are listed in Table B7.

Table B7. Primers used for RT-PCR analysis

	Name	Family	Forward	Reverse	Gene ID
1	*Ugt1a1*	UGT	TCTGCTTCTTCCGTACCTTCT	GCTTCAGGTGCTATGACCACAA	394436
2	*Ugt2b1*	UGT	GTGCTGGTGTGGCCTACAG	CAGAAGATATGAGAACGGTGACG	71773
3	*Sult1a1*	SULT	AACATGGAGCCCTTGCGTAAA	ATGAGCACATCATCAGGCCAG	20887
4	*Gapdh*		GGTTGTCTCCTGCGACTTCA	TGGTCCAGGGTTTCTTACTCC	

Text B5. Phenolic xenobiotic mixtures.

To better mimic the human exposure, we prepared a mixture consisting of 22 compounds at human-relevant concentrations. Table B8 listed the basic information of these compounds and Table B9 summarizes the maximum (MAX) and geometric mean (GM) concentration of those 22 compounds in human urinary or blood concentrations. These compounds are commonly found in our daily life and cover a relatively wide applications such as plasticizers, personal care products, phytoestrogens, sunscreen reagents, surfactants as well as flame retardants. Furthermore, most of these compounds have phenolic or benzoic acid function groups, which are very likely to compete with BPA during biotransformation. Most concentration in either urine or blood were determined by the Fourth National Report on Human Exposure to Environmental Chemicals in 2018 that consists of the cumulative biomonitoring data gathered from 1999-2000 through 2015-2016 (Prevention, 2018). The concentrations of other compounds that were not included in this report were referring to some earlier available studies (Zhao et al. 2017; Frederiksen et al. 2010; Högberg et al. 2007; S. I. Kim et al. 2015; Li et al. 2018;

Nagayama et al. 2000; Li et al. 2017; Li et al. 2015; Thomsen et al. 2002; Wang et al. 2015; Xiao et al., 2011). The "Maximum" concentrations of each compound were determined by averaging all detected maximum values in each study. The GM concentrations of compounds were determined by averaging the GM collected in each study.

Table B8. Summary of basic information of 22 compounds in the phenolic xenobiotic mixtures.

No.	Name	Cas-No.	Uses
1	2,4-dihydroxybenzophenone(DHB)	131-56-6	Ultraviolet (UV) light absorber
2	Bipshenol S (BPS)	80-09-1	Plasticizer
3	Triclocarban	101-20-2	Antiseptic in detergents and plastics.
4	Triclosan (TCS)	3380-34-5	Antiseptic
5	Butyl Paraben	94-26-8	Antiseptic and additives in cosmetics
6	Ethyl Paraben	120-47-8	Antifungal preservative and food addictive.
7	Methyl Paraben	99-76-3	Anti-fungal agent and food preservative.
8	Tris(1,3-dichloroisopropyl)phosphate (TDCPP)	13674-87-8	Flame retardants, pesticides, plasticizers, and nerve gases.
9	Triphenyl phosphate (TPhP)	115-86-6	Flame retardant and plasticizer.
10	Di-n-butyl phthalate (DBP)	84-74-2	Plasticizer.
11	Mono-n-butyl phthalate (MnBP)	131-70-4	Metabolite of di(n-butyl) phthalate
12	Benzylbutyl phthalate (BzBP)	85-68-7	Plasticizer
13	Mono-2-ethylhexyl phthalate (MEHP)	4376-20-9	Urinary phthalate metabolite.
14	Mono-benzyl phthalate (MBzP)	2528-16-7	Used as plasticizers
15	Di-isononyl phthalate (DiNP)	28553-12-0	Used to make plastics more flexible
16	Di-isodecyl phthalate (DiDP)	26761-40-0	Plasticizer
17	Genistein	446-72-0	An isoflavone
18	Perfluorooctanoic acid (PFOA)	335-67-1	Used in fire-fighting applications, cosmetics, greases and lubricants, paints, polishes and adhesives
19	Bisphenol A Diglycidyl Ether (BADGE)	1675-54-3	Used as constituent of epoxy resins
20	Bisphenol A bis(2,3-dihydroxypropyl) ether (BADGE-$2H_2O$)	5581-32-8	A component of commercial liquid epoxy resins commonly used in the food-packing industry and in dental sealants
21	Bisphenol-F-diglycidyl ether (BFDGE)	2095-03-6	Used as a material of interior coating for food cans
22	Tetrabromobisphenol A (TBBPA)	79-94-7	Flame retardant in electronics, paper, plastics, textiles, carpeting and furniture.

Table B9. Summary of the maximum (MAX), geometric mean (GM) concentration, sample type and size of 22 compounds from previous studies.

No.	Compound Name	Concentration in Urine (ng/mL) or Blood (ng/mL)		Sample Type	Sample Size (n)	Ref
		Maximum (MAX)	Geometric Mean (GM)			
1	2,4-dihydroxybenzophenone(DHB)	1911.7	21.9	Urine	15593	(Prevention, 2018)
2	Bipshenol S (BPS)	4.7	0.427	Urine	2682	(Prevention, 2018)
3	Triclocarban	24.5	0.200[c]	Urine	2686	(Prevention, 2018)
4	Triclosan (TCS)	599.7	13.8	Urine	15593	(Prevention, 2018)
5	Butyl Paraben	18.28	1.16[c]	Urine	13076	(Prevention, 2018)
6	Ethyl Paraben	99.62	5.44[c]	Urine	13076	(Prevention, 2018)
7	Methyl Paraben	1056.6	52.4	Urine	13076	(Prevention, 2018)
8	Tris(1,3-dichloroisopropyl)phosphate (TDCPP)[a]	8.60	0.856	Urine	2646	(Prevention, 2018)
9	Triphenyl phosphate (TPhP)	16.57	2.71	Blood	69	(Fanrong Zhao et al., 2017; Peng Li et al., 2017; Peng Li et al., 2015)
10	Di-n-butyl phthalate (DBP)	13.87	3.45	Blood	184	(Frederiksen et al., 2010; Högberg et al., 2007; X. Li, Sun, H., Yao, Y., Zhao, Z., Qin, X., Duan, Y., & Wang, L. , 2018)
11	Mono-n-butyl phthalate (MnBP)	110.2	16.8	Urine	21003	(Prevention, 2018)
12	Benzylbutyl phthalate (BzBP)	232.8	16.7	Urine	18462	(Prevention, 2018)
13	Mono-2-ethylhexyl phthalate (MEHP)	31.3	2.70	Urine	21003	(Prevention, 2018)
14	Mono-benzyl phthalate (MBzP)	70.0	8.30	Urine	21003	(Prevention, 2018)
15	Di-isononyl phthalate (DiNP)	13.8	2.00[c]	Urine	21003	(Prevention, 2018)
16	Di-isodecyl phthalate (DiDP)	20.9	2.67	Urine	13075	(Prevention, 2018)

17	Genistein	781.3	29.7	Urine	15689	(Prevention, 2018)
18	Perfluorooctanoic acid (PFOA)	10.7	3.50	Blood	14178	(Prevention, 2018)
19	Bisphenol A Diglycidyl Ether (BADGE)	2.50	0.500	Blood	223	(S. I. Kim et al., 2015)
20	Bisphenol A bis(2,3-dihydroxypropyl) ether (BADGE-2H_2O)	9.50	2.30	Blood	243	(S. I. Kim et al., 2015; L. Wang et al., 2015)
21	Bisphenol-F-diglycidyl ether (BFDGE)	180	26.2	Blood	20	(L. Wang et al., 2015)
22	Tetrabromobisphenol A (TBBPA)	7.58	0.68	Blood	265	(NAGAYAMA et al., 2000; Thomsen et al., 2002; Xiao et al., 2011)

Table B10. IC_{50} values and their 95% confidence intervals of *p,m,o*-cresol to BPA sulfate formation in incubation of HepG2 cell line and rat liver S9 fractions

	IC_{50} (μM) Intracellular	IC_{50} (μM) Extracellular	IC_{50} (μM) Average Cell	IC_{50} (μM) S9 fraction
p-Cresol	11.07 (7.94 – 15.15)	11.77 (9.62 – 14.37)	11.42	2.81 (1.41 –4.39)
m-Cresol	9.25 (7.78 – 11.03)	9.01 (8.19 – 9.92)	9.13	0.94 (0.24 – 2.07)
o-Cresol	3.12 (2.67 – 3.56)	2.52 (1.96 – 2.97)	2.82	0.23 (0.02 – 0.84)

Table B11. Summary of calculated binding energy of chemicals to two allosteric sites using molecular docking

No.	Compound	Binding energy to catechin-binding site (kcal/mol)	Binding energy to NSAID-binding site (kcal/mol)
1	BPA	-6.31	-2.54
2	*p*-Cresol	-4.69	-2.31
3	*m*-Cresol	-4.73	-2.40
4	*o*-Cresol	-4.50	-2.30
5	Acetaminophen	-4.06	-2.12
6	Catechin EGCG	-3.58	+2.76

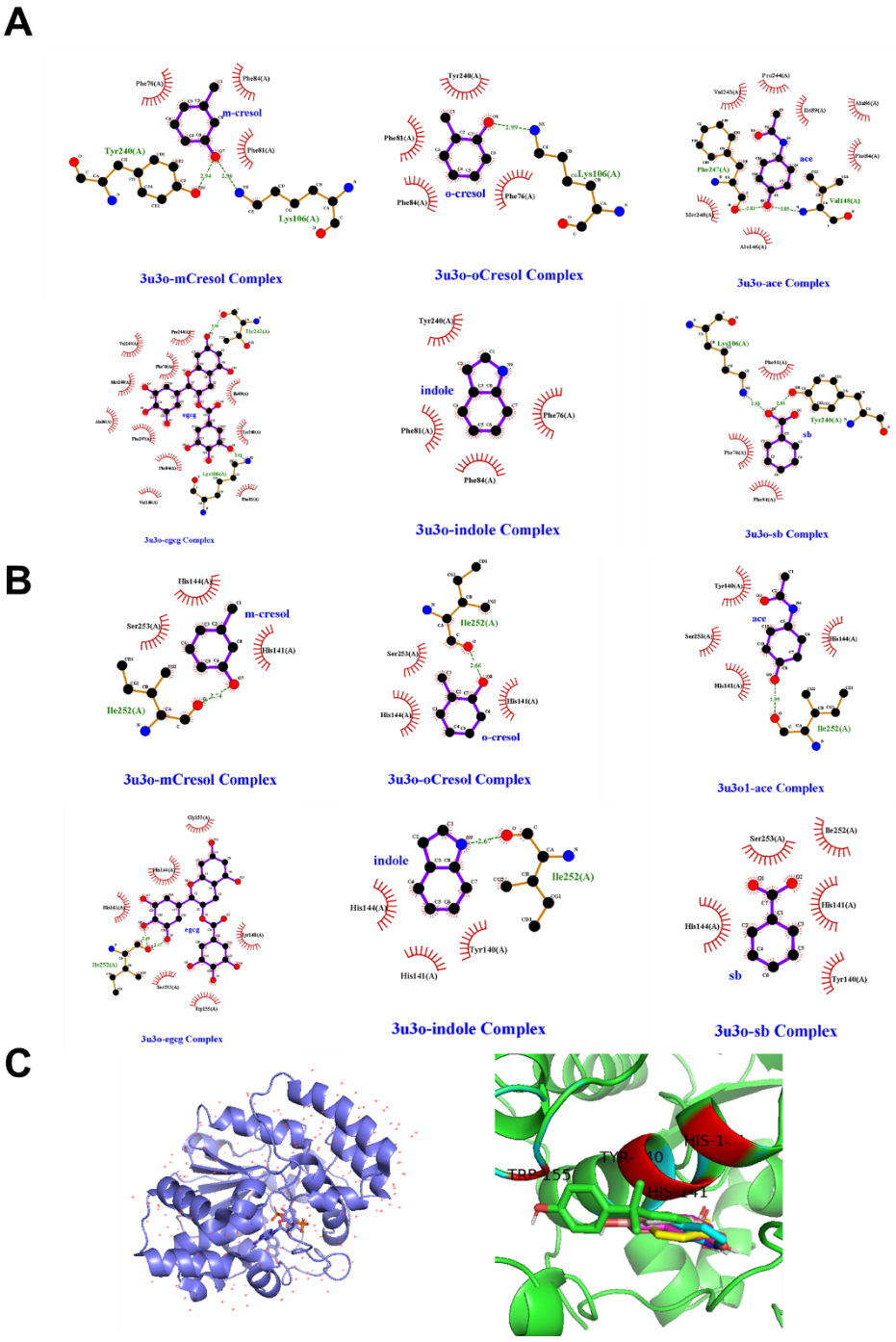


Figure B10. The docking analysis of cresol isomers, indole, benzoate acetaminophen and epigallocatechin gallate (EGCG) with two noninteracting allosteric sites of SULT 1A1. The docking sites showed as (A) The docking analysis of catechin-binding site with 6 compounds; (B) The NSAID-binding site with 6 compounds; (C) The protein structure (left) and the docking superimposed diagram (right) of SULT 1A1. The docking pockets of catechin-binding site and the NSAID-binding site are (0.375, 40, 40, 44, 22.191, 15.201, 16.633) and (0.375, 30, 20, 38, 5.894, 8.23, 8.532);

respectively.

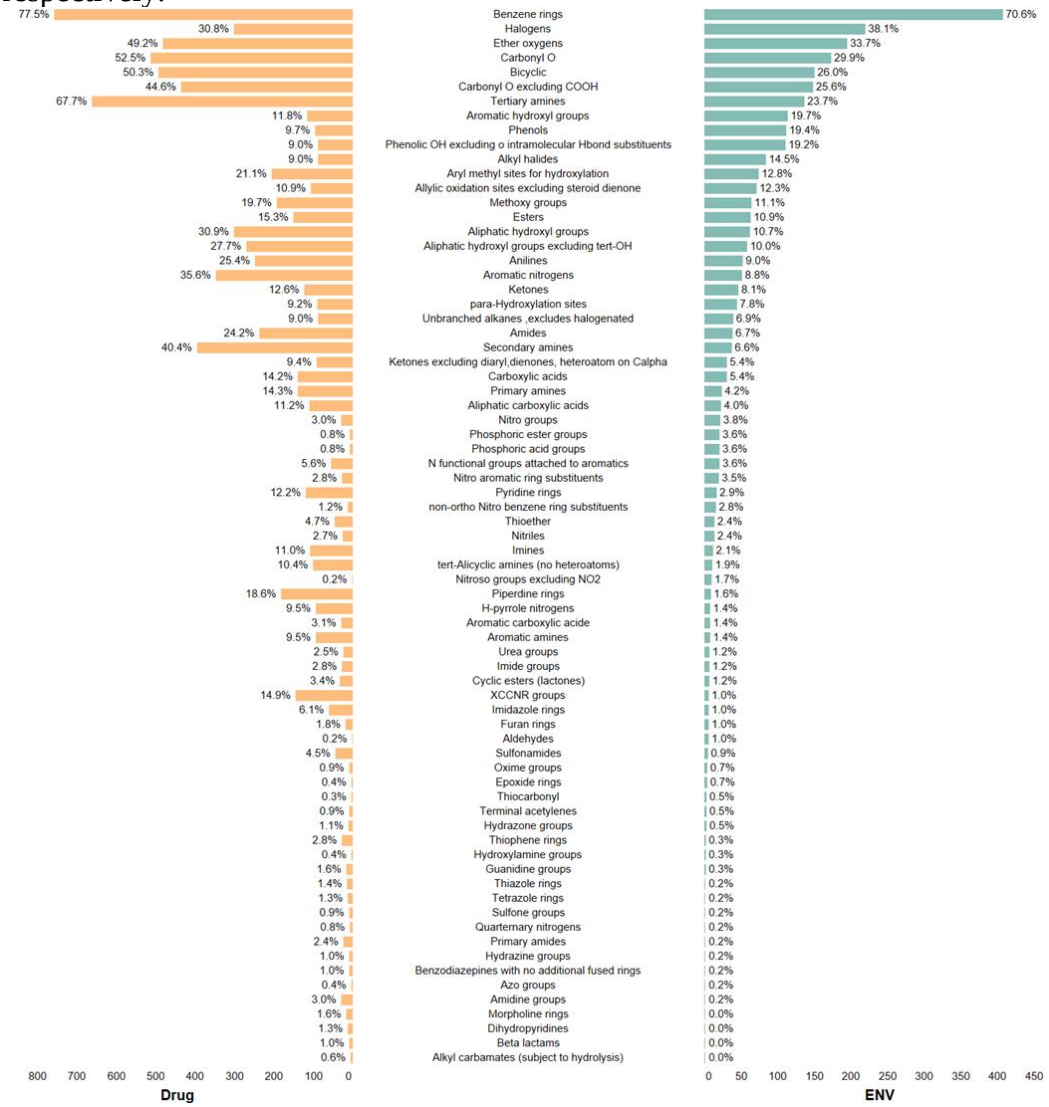


Figure B11. Fragment distribution of pharmaceuticals and environmental chemicals.