REVIEW

WILEY

Editor's choice

Computational methods and tools to predict cytochrome P450 metabolism for drug discovery

Jonathan D. Tyzack[1] (ID) | **Johannes Kirchmair[2,3,4]** (ID)

[1]EMBL-EBI, Wellcome Genome Campus, Cambridge, UK

[2]Department of Chemistry, University of Bergen, Bergen, Norway

[3]Computational Biology Unit (CBU), University of Bergen, Bergen, Norway

[4]Center for Bioinformatics, Universität Hamburg, Hamburg, Germany

Correspondence
Jonathan D. Tyzack, EMBL-EBI, Wellcome Genome Campus, Cambridge, UK.
Email: tyzack@ebi.ac.uk

Abstract

In this review, we present important, recent developments in the computational prediction of cytochrome P450 (CYP) metabolism in the context of drug discovery. We discuss in silico models for the various aspects of CYP metabolism prediction, including CYP substrate and inhibitor predictors, site of metabolism predictors (i.e., metabolically labile sites within potential substrates) and metabolite structure predictors. We summarize the different approaches taken by these models, such as rule-based methods, machine learning, data mining, quantum chemical methods, molecular interaction fields, and docking. We highlight the scope and limitations of each method and discuss future implications for the field of metabolism prediction in drug discovery.

KEYWORDS

cytochrome P450, drug discovery, enzyme–ligand interaction, machine learning, metabolism, metabolite structures, prediction, reactivity, sites of metabolism

1 | INTRODUCTION

Understanding the metabolism of small molecules is of paramount importance to the drug discovery industry and mitigates the risk of costly late failure in drug development projects due to adverse ADMET properties. Modern experimental approaches enable the elucidation of ADMET properties at an unprecedented level of detail but remain costly and time-consuming, so it is desirable to have efficient and reliable in silico methods in place (Kirchmair et al., 2015; Wilson, 2014). The most effective computational approaches allow the profiling of large datasets and enable the interactive optimization of lead compounds but at vastly lower expense. In the context of metabolism prediction, in silico tools are most commonly used for predicting substrates and inhibitors of metabolic enzymes, sites of metabolism (SoMs that is, metabolically labile atom positions in the substrate at which

biotransformations are initiated) and structures of likely metabolites. These predictions can then be used as part of the multi-parameter optimization drug discovery process, helping to satisfy stability constraints, increase in vivo half-life and avoid toxic metabolites.

Metabolic enzymes and systems have evolved to provide defense against xenobiotics (foreign and potentially hazardous molecules in our environment such as toxins and poisons) by transforming them into more readily excretable metabolites (Testa, 2014; Tyzack, Furnham, Sillitoe, Orengo, & Thornton, 2017). Pharmaceuticals and other molecules encountered through the course of modern life fall under the remit of these metabolic processes which operate on them in two broad categories: phase I metabolism involves making the molecule more polar and hydrophilic; and phase II involves conjugation with endogenous hydrophilic compounds. The net result is that metabolism is responsible for

the clearance of about 75% of all drugs (Di, 2014) producing metabolites with different physicochemical, physiological, pharmacological, and toxicological properties (Kirchmair, et al., 2015; Kirchmair, Howlett, et al., 2013; Tyzack & Glen, 2014). Metabolism poses many challenges, but a more complete understanding also generates opportunities (Testa, Pedretti, & Vistoli, 2012), as summarized in Figure 1.

The most important enzymes in phase I belong to the cytochrome P450s (CYPs) since they produce the most first generation metabolites and have a high proportion of toxic/reactive metabolites (Testa et al., 2012). They are a family of heme-containing enzymes ubiquitously found in animals, plants, fungi, and bacteria where at least 57 CYP isoforms have been documented in humans. The different CYP isoforms exhibit varying pocket sizes, shapes, binding surfaces, and flexibility, giving them different substrate specificity profiles and directing metabolism toward different parts of small molecules (Leach & Kidley, 2014; Mustafa, Yu, & Wade, 2014; Testa, 2014). Some CYP isoforms have remarkable ligand promiscuity driven in part by the size and plasticity of their binding sites where significant flexibility and conformational change has been revealed with molecular dynamics simulations (Mustafa et al., 2014). CYPs can be classified into two major classes: those involved in xenobiotic detoxification found mainly in the liver (such as the CYP2 and CYP3 families); and those involved in the biosynthesis of endogenous compounds such as sterols, fatty acids, eicosanoids, and vitamins (Guengerich, Waterman, & Egli, 2016; Rendic & Guengerich, 2015).

There are many factors that make understanding the action of CYP enzymes in vivo challenging, including expression patterns, inhibition levels, and genetic polymorphisms. Expression patterns vary significantly across organs, the highest human concentrations being found in the liver and small intestine, but expression is also influenced by gender, age, disease, stress, lifestyle, diet, and medication (Testa, 2014). These factors will all influence in vivo CYP expression and the rate of drug clearance. Furthermore, CYP inhibition and induction can be hugely influential, such as the flavonoid CYP inhibitors found in grapefruit juice that can

result in higher drug concentrations than anticipated in the dosing regimen. Conversely, CYP induction can cause drug concentrations to fall below therapeutic levels, such as the dietary supplement St John's Wort, a potent inducer of CYP3A4 (Roby, Anderson, Kantor, Dryer, & Burstein, 2000). CYP genetic polymorphisms manifesting as loss or gain of function variants also need to be considered in the drug development process so that undue reliance is not placed on CYP isoforms that have known deficiencies in certain ethnic populations. These factors all contribute to an overall complex picture and make prediction of drug metabolism highly challenging but essential for drug discovery and development.

This review is intended to describe important contributions in the field of CYP metabolism prediction and cover more recent developments in this field but with a focus on methods that are freely available (Table 1). More comprehensive and complete lists of publications, software, and databases can be found elsewhere (Bezhentsev et al., 2016; Kar & Leszczynski, 2017; Kirchmair et al., 2015) with other reviews covering topics such as molecular dynamics, and QM modeling (Kirchmair et al., 2012; Shaik, Chen, Usharani, & Thiel, 2014; Williamson, 2014) and pharmacogenetics, pharmacoepigenetics, and clinical significance (Manikandan & Nagini, 2018) in more detail.

2 | METHODS FOR CYP SPECIFICITY PREDICTION

Understanding the specificity of individual CYP isoforms to bind small molecules can assist in the prediction of metabolic stability, enzyme inhibition, and drug–drug interactions. Docking methods explicitly model the substrate binding event, such as OpenVirtualToxLab (Vedani, Dobler, Hu, & Smieško, 2015) that uses flexible docking in combination with multi-dimensional QSAR to predict the inhibition of five of the main xenobiotic metabolizing CYPs. However, these methods are computationally expensive, presenting challenges when attempting to incorporate them into software tools. For this reason, machine learning methods such as neural networks (Dai, Xu, Xiong, Liu, & Wei, 2016), support vector machines (SVMs) (Daina, Michielin, & Zoete, 2017) and random forests (Hunt, Segall, & Tyzack, 2018) have evolved as the mainstay for predicting enzyme specificity as they allow the modeling of the complex nonlinear relationships observed in large collections of enzyme–substrate interaction data. They differ in terms of the data on which they are trained, the descriptors used to represent the data, the scope of predictions, and the type of machine learning methodology employed. Machine learning models produce accurate results quickly and so lend themselves for usage in web applications where results can be returned in real time. However, the quantity and quality of the available data is often a limiting factor
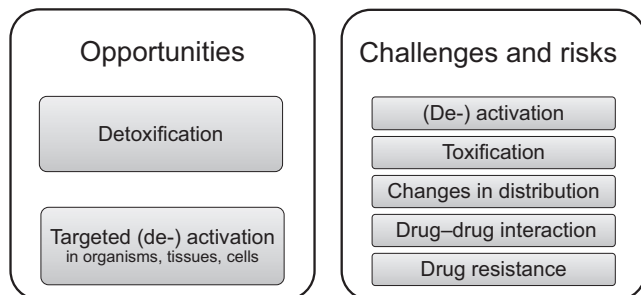

FIGURE 1 Opportunities, challenges and risks related to drug metabolism

TABLE 1 Free in silico models for the prediction drug metabolism

Tool name	Description	Available as a free		At URL
		Software package	Web service	
Prediction of CYP specificity				
OpenVirtualToxLab (Vedani et al., 2015)	• Combination of flexible docking with multi-dimensional QSAR • Predicts inhibitors and non-inhibitors of CYP 1A2, 2A13, 2C9, 2D6, and 3A4	x		http://www.biograf.ch/data/projects/OpenVirtualToxLab.php
SwissADME (Daina et al., 2017)	• Support vector machines • Predicts inhibitors and non-inhibitors of CYP 1A2, 2C9, 2C19, 2D6, and 3A4		x	http://www.swissadme.ch
CypRules (Shao et al., 2015)	• Learning base model • Predicts inhibitors and non-inhibitors of CYP 1A2, 2C9, 2C19, 2D6, and 3A4		x	https://cyprules.cmdm.tw
CypReact (Tian et al., 2018)	• Predicts substrates and non-substrates of CYP 1A2, 2A6, 2B6, 2C8, 2C9, 2C19, 2D6, 2E1, and 3A4	x		https://bitbucket.org/Leon_Ti/cypreact
Prediction of sites of metabolism				
SMARTCyp (Rydberg, Gloriam, & Olsen, 2010; Rydberg, Gloriam, Zaretzki, et al., 2010)	• Predicts SoMs for CYPs based on reaction energies derived from density functional theory	x	x	https://smartcyp.sund.ku.dk
SOMP (Rudik et al., 2015)	• Predicts SoMs for five major CYPs and for UDP-glucuronosyltransferases based on a Bayesian approach		x	http://www.way2drug.com/SOMP/
FAME 2 (Šícho et al., 2017)	• Predicts SoMs for CYPs with extremely randomized trees	x		https://www.zbh.uni-hamburg.de/forschung/acm/software_datasets.html
XenoSite (Matlock et al., 2015; Zaretzki et al., 2013)	• Predicts SoMs for CYPs with neural networks		x	http://swami.wustl.edu/xenosite
Prediction of metabolite structures				
SyGMa (Ridder & Wagener, 2008)	• Expert-curated rule set with empirical scoring of metabolites generated by phase I and phase II reactions	x[a]		https://github.com/3D-e-Chem/sygma
ToxTree (Patlewicz et al., 2008)	• Set of biotransformation rules for CYPs applied to SoMs predicted with SMARTCyp	x		http://toxtree.sourceforge.net
MetaTox (Rudik et al., 2017)	• Bayesian approach for five major CYPs and for UDP-glucuronosyltransferases		x	http://www.way2drug.com/mg/

[a] Available also as a KNIME node.

and determines the coverage and performance of these models (Kirchmair et al., 2015). So far they have not allowed the generation of sufficiently accurate regression models but only classification models (Gleeson et al., 2007).

When applying models for CYP specificity prediction, care must be taken to ensure that the predictive models are based on predicting enzyme–substrate interactions or enzyme–inhibitor interactions. SwissADME (Daina et al., 2017) is a web service that offers, among many other tools, SVM models for the prediction of inhibitors for the five major CYP isoforms (i.e., CYPs 1A2, 2C19, 2C9, 2D6, and 3A4). The classifiers were trained on data from the PubChem Bioassay 1851 dataset (Veith et al., 2009) using 50 molecular and physicochemical descriptors. CypRules (Shao et al., 2015) is another web service that predicts inhibitors and noninhibitors of the same major CYPs based on the same data. It utilizes decision trees in combination with the concept of information entropy.

In contrast, a downloadable java program called CypReact (Tian, Djoumbou-Feunang, Greiner, & Wishart, 2018) classifies molecules as substrate or non-substrate (but not inhibitor and non-inhibitor) for nine major xenobiotic metabolizing isoforms. It uses a machine learning method called LBM (learning base model) based on many physicochemical and structural properties but applies feature selection to avoid over-fitting. The training data of approximately one thousand compounds was compiled from several different sources.

There are also a number of commercial offerings for the prediction of enzyme inhibition and specificity. The StarDrop platform contains functionality called WhichP450 (Hunt et al., 2018) that identifies enzyme–substrate interactions for seven major xenobiotic metabolizing isoforms. It uses random forest models based on datasets manually extracted from the primary literature, where care was taken to annotate the training data with major and minor metabolizing isoforms for completeness. ADMET Predictor (Simulations Plus) has functionality to both predict inhibitors for five major drug metabolizing CYPs and substrates for nine CYP isoforms using data acquired from the BIOVIA Metabolite database, (BIOVIA Metabolite, n.d.) the Drugbank database (Wishart et al., 2018), and other public resources.

3 | METHODS FOR SITE OF METABOLISM PREDICTION

There are two components to model in order to predict SoMs: the reactivity and accessibility of atoms in a molecule. Modern methods for SoM prediction follow different concepts to model these two components falling into two categories: structure-based and ligand-based. Structure-based methods incorporate knowledge about the CYP enzyme and suffer from being computationally expensive. Consequently,

ligand-based methods trained on datasets of known (non-) binders and metabolic sites are far more commonly applied in prediction tools. Ligand-based methods can be further subdivided into the sub-categories data mining, expert knowledge, and machine learning. The models presented in the following sections address reactivity and accessibility using various combinations of structure- and ligand-based approaches.

3.1 | Reactivity

The reactivity of different molecular fragments toward CYP-mediated oxidation is the major determinant of SoMs. Reactivity can only really be classified as structural in terms of QM-MM simulations that model the reaction in silico but the extremely high computational cost limits this approach to very detailed studies of a particular pathway with a particular substrate (Shaik et al., 2014). Consequently, reactivity is nearly always ligand-based in the SoM prediction tools that will be described. Data mining, expert knowledge, and machine learning methods are almost exclusively ligand-based, involving searching annotated ligand datasets, applying knowledge of ligand metabolism, or learning from ligand datasets respectively.

3.1.1 | Reactivity: Simulation-based

QM methods allow the accurate investigation of reactivity but at substantial computational expense and it is common for investigators to carry out QM calculations on a predefined library of molecular fragments from which to match to query molecules. This technique was adopted as one component of SMARTCyp (Rydberg, Gloriam, & Olsen, 2010; Rydberg, Gloriam, Zaretzki, Breneman, & Olsen, 2010), which predicts SoMs based on activation energies calculated using density functional theory (DFT) calculations applied to a library of molecular fragments. The P450 Module within the commercial software STARDROP adopts a different approach (Tyzack, Hunt, & Segall, 2016), directly modeling the reaction pathway using parameterized semi-empirical AM1 QM methods. This enables activation energy estimates to be generated for each molecular fragment in the context of the molecular environment in which it resides, rather than treating each fragment as identical regardless of its neighboring atomic moieties. The semi-empirical QM methods use a methoxy radical to simplify the calculations since modeling the heme group and its protein environment is more akin to full QM-MM methods which are extremely computationally expensive and cannot be run routinely for SoM identification (Olsen, Rydberg, Rod, & Ryde, 2006; Shaik et al., 2014).

3.1.2 | Reactivity: Descriptor-based

Other methods include reactivity descriptors as a component to approximate the hydrogen and electron abstraction

processes fundamental to the CYP catalytic cycle but often require QM minimization to be carried out on a query molecule which carries significant computational expense. A number of studies have identified reactivity descriptors based on the energies of molecular orbitals (Mukherjee, Lal Gupta, & Jayaram, 2015; Tyzack, Williamson, Torella, & Glen, 2013) and hardness (Pragyan, Kesharwani, Nandekar, Rathod, & Sangamwar, 2014) as important metrics to help determine SoMs. The commercial software ADMET Predictor (Simulations Plus) uses reactivity descriptors based on Huckel charges and semi-empirical molecular orbital calculations among others, and MetScore (Göller, Finkelmann, Goldmann, & Schneider, 2018) is based on quantum chemical partial charges.

3.2 | Accessibility

Accessibility can be modeled either as structure-based, using (pseudo-) docking approaches to insert the molecule into the CYP binding pocket, or ligand-based, using molecular fingerprints to approximate steric hindrance and orientational concerns. These approaches will be discussed in more detail alongside the reactivity metric they supplement.

3.2.1 | Accessibility: Structure-based

Structure-based methods explicitly model the accessibility criteria by placing the molecule into the CYP binding pocket. The output from these methods is easy to interrogate as they generate binding poses that can be inspected by the user and are rooted in the physical reality of explicitly modeling the binding event. Today, crystal structures are available for almost all CYP isoforms relevant to xenobiotic metabolism (Guengerich, 2014; Oostenbrink, 2014). Malleability of these enzymes, their complex interplay with water, and the hydrophobic character of their—in part—very large binding sites pose significant challenges to the application of structure-based approaches, in particular docking. For this reason, structure-based methods are primarily applied for the detailed investigation of the interaction of CYPs with individual compounds rather than for the profiling of small-molecule libraries. In this setup, structure-based methods can be particularly valuable for rationalizing distinct biological properties of enantiomers.

A large number of studies employing docking algorithms to determine the SoMs of small molecules have been published. One example from 2015 used Autodock Vina (Trott & Olson, 2010) to insert substrates into an ensemble of CYP 2C9 structures (Kingsley, Wilson, Essex, & Lill, 2015) and gave overall predictions by coupling with reactivities from SMARTCyp. The authors highlighted the importance of considering flexibility and sampling conformational space by the improvement in going from docking into a single structure to

an ensemble of structures. Constrained docking has also been explored (Tyzack et al., 2013), reducing the size of the search space by fixing each ligand atom in turn in a position relevant to catalysis relative to the CYP heme and explicitly obtaining a score for each potential SoM. However, it is far more common for unconstrained docking to be applied, predicting SoMs as those that are in close proximity to the heme in the best pose(s), but placing reliance on the docking algorithm to fully explore conformational space rather than forcing it to consider each potential SoM as in the constrained workflow.

Another publication used the top three poses from in-house docking software combined with reactivity descriptors based on molecular orbitals to build a classifier with good performance (Mukherjee et al., 2015). The docking method carries considerable computational expense since its estimation of binding free energies requires the calculation of partial charges from QM software and energy minimization with MD software, but a good correlation to experimental binding free energies was observed. The docking software GLIDE has also been combined with reactivity descriptors based on hardness to identify metabolically labile sites (Pragyan et al., 2014).

There are also several commercial packages available that make SoM predictions using CYP structural information. MetaSite (Cruciani et al., 2005) was one of the pioneers of structure-based methods for SoM prediction and was developed into a commercial software package. It uses a pseudo-docking approach where steric and chemical properties of CYPs such as hydrogen bonding or hydrophobic regions are described by molecular interaction fields (MIFs), calculated by placing various chemical probes in a grid system embedded over the CYP structure. Potential SoMs can be identified in a target molecule by aligning to the MIFs, coupled with matching molecular fragments to precomputed reactivity scores from QM approaches. IDSite (Li, Schneebeli, Bylund, Farid, & Friesner, 2011) samples the conformational space of CYPs as part of a flexible docking procedure with Glide (Friesner et al., 2004).

3.2.2 | Accessibility: Ligand-based

Ligand-based approaches have the benefit of speed and have been shown to perform at least as well as structural methods, although sometimes with less interpretability. They provide indirect information about the active site of a protein on the basis of the location of documented SoMs in the molecular graph, a concept rather like describing a glove by the properties of hands that will fit into it. There are various ways to encode the accessibility of molecular fragments using a 2D molecular graph such as the simple SPAN predictor used in SMARTCyp (Rydberg, Gloriam, Zaretzki, et al., 2010) which measures the relative distance to the "edge" of a molecule and was used in conjunction with pre-computed

reactivity scores. A different concept is the use of 3D alignments to derive the SoMs for a compound of interest from reference compounds with annotated SoMs (de Bruyn Kops, Friedrich, & Kirchmair, 2017). Typically, methods evaluating the accessibility of atom positions are combined with reactivity metrics to produce predictive models such as the P450 Module within StarDrop which contains orientational and steric descriptors derived from manually curated datasets from the primary literature.

3.3 | Data mining

These methods encode the SoMs stored in metabolic databases in a manner that allows easy interrogation with molecular fragments in query molecules. One example is MetaPrint2D (Adams, 2010) that derives likelihoods of metabolic transformations for atoms in defined atomic environments by mining the BIOVIA Metabolite database, but is unfortunately no longer maintained. PROXIMAL (Yousofshahi, Manteiga, Wu, Lee, & Hassoun, 2015) uses a similar data mining of atomic sites approach but using the KEGG and DrugBank databases, where matches are scored by activity and abundance data obtained from the literature.

3.4 | Machine learning

Machine learning software tools are popular since they can be routinely applied to large datasets early in the drug discovery process (Xiong et al., 2018). However, the size and quality of available data and the features used to describe them are of paramount importance in the development of these approaches. Much data are held in-house within big pharmaceutical companies but there is a dearth of high-quality SoM data in the public domain where the Zaretzki dataset (Zaretzki, Matlock, & Swamidass, 2013) (recently revised by de Bruyn Kops et al., 2017) remains the most comprehensive. Also, many of the datasets only record the major sites and metabolites and often significant secondary or tertiary data points are overlooked, leading to pessimistic false positive rates being reported when a non-recorded SoM is predicted. Recent efforts at Optibrium to collate enlarged, high-quality datasets from the literature highlighted this issue where it was observed that papers from different groups studying the same metabolic process varied in the completeness of the metabolic sites recorded although this data remain in-house and forms the basis of the P450 module within StarDrop.

There can be a tendency with machine learning approaches to keep adding extra descriptors but this can lead to overfitting problems and hinder interpretability of the models. However, one study showed that 2D topological descriptors based on MOLPRINT2D fingerprints (Bender, Mussa, Glen, & Reiling, 2004) coupled with machine learning methods were sufficient to produce strong classification performance

(Tyzack, Mussa, Williamson, Kirchmair, & Glen, 2014). Descriptors based on 2D atomic neighborhoods also form the basis of SOMP (Rudik, Dmitriev, Lagunin, Filimonov, & Poroikov, 2014) (available as a web server; Rudik, Dmitriev, Lagunin, Filimonov, & Poroikov, 2015) which uses Bayesian classification models trained on data from the BIOVIA Metabolite database and confirms that successful prediction models can be built without having to resort to a plethora of QM based descriptors.

FAME 2 (Šícho, de Bruyn Kops, Stork, Svozil, & Kirchmair, 2017) is an accurate and robust approach to the prediction of SoMs related to CYP metabolism. It is built on the strong foundations laid by its predecessor, FAME (Kirchmair, Williamson, et al., 2013), but utilizes an extremely randomized tree algorithm (rather than random forest) and a new type of circular fingerprint. At the same time, FAME 2 maintains the overall design ethos of stripping away complexity where it is not justified in terms of performance. The more descriptive topological descriptors of FAME 2 produce models that generalize well, making them applicable to molecular structures distant from those represented by the training data, such as complex natural products. The model was shown to obtain competitive performance with computationally more expensive methods based on density functional theory (Finkelmann, Göller, & Schneider, 2017; Šícho et al., 2017).

Another example of a free SoM predictor is Xenosite (Zaretzki et al., 2013), which was built on RS-Predictor. Xenosite utilizes neural networks trained on topological, molecular, and quantum chemical descriptors coupled with SMARTCyp reactivities. In contrast to its predecessor, outputs are represented as probabilities rather than rank orderings. Xenosite is available as a web server (Matlock, Hughes, & Swamidass, 2015) and was extended to non-CYP applications such as UGT-mediated metabolism prediction (Dang, Hughes, Krishnamurthy, & Swamidass, 2016).

4 | METHODS FOR METABOLITE STRUCTURE PREDICTION

Methods for metabolite structure prediction are dominated by knowledge-based approaches. These approaches rely on cumulative knowledge acquired by drug discovery scientists over the years through observation and experimental work. Underlying these methods are a series of expert-curated rules that represent SoMs and metabolic transformations, but when applied iteratively they can lead to an unmanageable combinatorial explosion of predictions. SyGMa (Ridder & Wagener, 2008) combines an expert-curated rule set with empirical scoring to predict and rank the likely products of phase I and phase II metabolism. It was one of the earliest freely available offerings to generate metabolite

structures but remains a highly effective and useable tool. More recently, in addition to being made available as a software package, SyGMa has been released as a KNIME (Berthold et al., 2018) node as part of the 3D-e-Chem project (Kooistra et al., 2018). ToxTree (Patlewicz, Jeliazkova, Safford, Worth, & Aleksiev, 2008) includes a module for metabolite structure prediction which is based on SoMs predicted with SMARTCyp. A further example of a free model is the MetaTox web server (Rudik et al., 2017) which uses a Bayesian approach for the prediction of likely metabolites and additionally estimates the toxicity of predicted metabolites using QSAR models. MetaPrint2D-React is an extension to the SoM predictor MetaPrint2D (Adams, 2010) which predicts the structures of likely metabolites based on the occurrence ratio of biotransformations in the BIOVIA Metabolite database. Unfortunately, the MetaPrint2D-React web service is no longer maintained.

Commercial tools for specificity and SoM prediction usually have functionality to generate metabolite structures from their regioselectivity predictions, including StarDrop, ADMET Predictor, and Percepta. Also the previously discussed pseudo-docking method MetaSite includes an array of methods for the prediction and assessment of metabolite structures. A leading expert system for the prediction of likely metabolites and structures is Meteor Nexus (Marchant, Briggs, & Long, 2008), which mitigates the combinatorial explosion problem with the 2017 addition of a k-nearest neighbor approach to prioritize transformations toward those most commonly observed (Marchant, Rosser, & Vessey, 2017). Meteor also includes models for SoM prediction via a reimplementation of SMARTCyp and tools for linking mass spectrometry data with predicted metabolites of phase I and phase II biotransformations.

5 | CHALLENGES AND OUTLOOK

This review has demonstrated the number and diversity of different approaches to predict xenobiotic metabolism highlighting the fundamental importance of this topic to drug discovery. Machine learning methods are by far the most commonly adopted method to make predictions as they can be usefully applied to large, complex data at an early stage in a drug discovery project. It is hoped that this review will be useful in focusing readers' attention to those methods that can be readily applied in drug discovery programmes.

The users of metabolism prediction tools are primarily industry-based but due to confidentiality issues data can often not be submitted to web services across the internet. For the available tools to be useful to industry, they need to be distributed as packages that can be executed within the company's infrastructure. However half of all tools developed in academia are unfortunately only offered as web services.

This should be reconsidered to make software developed by academia more useable by industry.

The predictive performance of these methods is strong and is probably approaching the limit based on the quality of the available data. However, annotation (SoM assignment in particular) is time-consuming and requires expert knowledge hindering the availability of data in the public domain. The amount of data is growing slowly and would be greatly aided by the release of metabolic data by pharmaceutical companies but suffered a significant setback when BIOVIA decided to withdraw the Metabolite database from their product range which formed the basis of many applications.

Similarly, strong classification performance is often reported by new methods but it would be desirable if metabolism prediction tools gave some assessment of their domain of applicability or an indication of the confidence domain of predictions. This functionality is sometimes observed in commercial offerings such as the metabolic lability labels in StarDrop and the SoM propensity scores in ADMET Predictor but is often neglected in academic tools. Furthermore, the comparison of methods is often challenging because many studies do not disclose their training sets and the creation of benchmark datasets would make direct comparison much easier. When presenting performance, it would also be beneficial to routinely make comparisons to pure chance to fairly reflect the predictions being made.

We expect the momentum in metabolism prediction software to be maintained in the foreseeable future due to the importance of this topic to drug discovery but progress will inevitably be hindered by the lack of high-quality data. Currently, the focus has been on CYP metabolism due to its relevance to phase I drug metabolism but we expect that more attention will turn to enzymes involved in other clearance pathways such as glutathione transferases and sulfotransferases.

AUTHOR INFORMATION

Jonathan Tyzack was awarded a Ph.D. from The University of Cambridge for his studies into the computational prediction of xenobiotic metabolism by cytochrome P450 enzymes. He is currently a research fellow in structural bioinformatics at the European Bioinformatics Institute, part of the European Molecular Biology Laboratory, studying enzyme structure, function and evolution.

Johannes Kirchmair is an associate professor in bioinformatics at the Department of Chemistry and the Computational Biology Unit (CBU) of the University of Bergen. He is also a group leader at the Centre for Bioinformatics (ZBH) of the University of Hamburg. His research activities focus on the development and application of in silico methods for the prediction of biological activities, metabolic fate and toxicity of xenobiotics.

ORCID

Jonathan D. Tyzack [ID] https://orcid.org/0000-0003-4827-9023

Johannes Kirchmair [ID] https://orcid.org/0000-0003-2667-5877

REFERENCES

Adams, S. E. (2010). Molecular similarity and xenobiotic metabolism. Cambridge, UK: University of Cambridge. Retrieved from https://www.repository.cam.ac.uk/handle/1810/225225

Bender, A., Mussa, H. Y., Glen, R. C., & Reiling, S. (2004). Similarity searching of chemical databases using atom environment descriptors (MOLPRINT 2D): Evaluation of performance. *Journal of Chemical Information and Computer Sciences*, *44*(5), 1708–1718. https://doi.org/10.1021/ci0498719

Berthold, M. R., Cebron, N., Dill, F., Gabriel, T. R., Kötter, T., Meinl, T., … Wiswedel, B. (2018). *KNIME: The Konstanz information miner (version 3.5.3)*. Retrieved from https://www.knime.com

Bezhentsev, V. M., Tarasova, O. A., Dmitriev, A. V., Rudik, A. V., Lagunin, A. A., Filimonov, D. A., & Poroikov, V. V. (2016). Computer-aided prediction of xenobiotic metabolism in the human body. *Russian Chemical Reviews*, *85*(8), 854. https://doi.org/10.1070/RCR4614

BIOVIA Metabolite. (n.d**.**). San Diego, CA: BIOVIA.

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. *Journal of Medicinal Chemistry*, *48*(22), 6970–6979. https://doi.org/10.1021/jm050529c

Dai, H., Xu, Q., Xiong, Y., Liu, W.-L., & Wei, D.-Q. (2016). Improved prediction of michaelis constants in CYP450-mediated reactions by resilient back propagation algorithm. *Current Drug Metabolism*, *17*(7), 673–680. https://doi.org/10.2174/1389200217666160513144551

Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, *7*, 42717. https://doi.org/10.1038/srep42717

Dang, N. L., Hughes, T. B., Krishnamurthy, V., & Swamidass, S. J. (2016). A simple model predicts UGT-mediated metabolism. *Bioinformatics*, *32*(20), 3183–3189. https://doi.org/10.1093/bioinformatics/btw350

de Bruyn Kops, C., Friedrich, N.-O., & Kirchmair, J. (2017). Alignment-based prediction of sites of metabolism. *Journal of Chemical Information and Modeling*, *57*(6), 1258–1264. https://doi.org/10.1021/acs.jcim.7b00165

Di, L. (2014). The role of drug metabolizing enzymes in clearance. *Expert Opinion on Drug Metabolism & Toxicology*, *10*(3), 379–393. https://doi.org/10.1517/17425255.2014.876006

Finkelmann, A. R., Göller, A. H., & Schneider, G. (2017). Site of metabolism prediction based on ab initio derived atom representations. *ChemMedChem*, *12*(8), 606–612. https://doi.org/10.1002/cmdc.201700097

Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., … Shenkin, P. S. (2004). Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of Medicinal Chemistry*, *47*(7), 1739–1749. https://doi.org/10.1021/jm0306430

Gleeson, M. P., Davis, A. M., Chohan, K. K., Paine, S. W., Boyer, S., Gavaghan, C. L., … Albertson, N. (2007). Generation of in-silico cytochrome P450 1A2, 2C9, 2C19, 2D6, and 3A4 inhibition QSAR models. *Journal of Computer-Aided Molecular Design*, *21*(10–11), 559–573. https://doi.org/10.1007/s10822-007-9139-6

Göller, A. H., Finkelmann, A. R., Goldmann, D., & Schneider, G. (2018). MetScore: Site of metabolism prediction beyond CYP P450 enzymes. *ChemMedChem*, *13*(21), 2281–2289. https://doi.org/10.1002/cmdc.201800309

Guengerich, F. P. (2014). Experimental approaches to analysis of reactions of cytochrome P450 enzymes. In J. Kirchmair (Ed.), Drug metabolism prediction, Vol. 32 (pp. 199–220). Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA.

Guengerich, F. P., Waterman, M. R., & Egli, M. (2016). Recent structural insights into cytochrome P450 function. *Trends in Pharmacological Sciences*, *37*(8), 625–640. https://doi.org/10.1016/j.tips.2016.05.006

Hunt, P. A., Segall, M. D., & Tyzack, J. D. (2018). WhichP450: A multi-class categorical model to predict the major metabolising CYP450 isoform for a compound. *Journal of Computer-Aided Molecular Design*, *32*(4), 537–546.

Kar, S., & Leszczynski, J. (2017). Recent advances of computational modeling for predicting drug metabolism: A perspective. *Current Drug Metabolism*, *18*(12), 1106–1122.

Kingsley, L. J., Wilson, G. L., Essex, M. E., & Lill, M. A. (2015). Combining structure- and ligand-based approaches to improve site of metabolism prediction in CYP2C9 substrates. *Pharmaceutical Research*, *32*(3), 986–1001. https://doi.org/10.1007/s11095-014-1511-3

Kirchmair, J., Göller, A. H., Lang, D., Kunze, J., Testa, B., Wilson, I. D., … Schneider, G. (2015). Predicting drug metabolism: Experiment and/or computation? *Nature Reviews. Drug Discovery*, *14*(6), 387–404. https://doi.org/10.1038/nrd4581

Kirchmair, J., Howlett, A., Peironcely, J. E., Murrell, D. S., Williamson, M. J., Adams, S. E., … Glen, R. C. (2013). How do metabolites differ from their parent molecules and how are they excreted? *Journal of Chemical Information and Modeling*, *53*(2), 354–367. https://doi.org/10.1021/ci300487z

Kirchmair, J., Williamson, M. J., Afzal, A. M., Tyzack, J. D., Choy, A. P. K., Howlett, A., … Glen, R. C. (2013). FASt MEtabolizer (FAME): A rapid and accurate predictor of sites of metabolism in multiple species by endogenous enzymes. *Journal of Chemical Information and Modeling*, *53*(11), 2896–2907. https://doi.org/10.1021/ci400503s

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. https://doi.org/10.1021/ci200542m

Kooistra, A. J., Vass, M., McGuire, R., Leurs, R., de Esch, I. J. P., Vriend, G., … de Graaf, C. (2018). 3D-e-Chem: Structural cheminformatics workflows for computer-aided drug discovery. *ChemMedChem*, *13*(6), 614–626. https://doi.org/10.1002/cmdc.201700754

Leach, A. G., & Kidley, N. J. (2014). Cytochrome P450 substrate recognition and binding. In J. Kirchmair (Ed.), Drug metabolism prediction, Vol. 392 (pp. 103–132). Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA.

Li, J., Schneebeli, S. T., Bylund, J., Farid, R., & Friesner, R. A. (2011). IDSite: An accurate approach to predict P450-mediated drug

metabolism. *Journal of Chemical Theory and Computation*, 7(11), 3829–3845. https://doi.org/10.1021/ct200462q

Manikandan, P., & Nagini, S. (2018). Cytochrome P450 structure, function and clinical significance: A review. *Current Drug Targets*, 19(1), 38–54.

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. *Toxicology Mechanisms and Methods*, 18(2–3), 177–187. https://doi.org/10.1080/15376510701857320

Marchant, C. A., Rosser, E. M., & Vessey, J. D. (2017). A k-nearest neighbours approach using metabolism-related fingerprints to improve in silico metabolite ranking. *Molecular Informatics*, 36(3), 1600105. https://doi.org/10.1002/minf.201600105

Matlock, M. K., Hughes, T. B., & Swamidass, S. J. (2015). XenoSite server: A web-available site of metabolism prediction tool. *Bioinformatics*, 31(7), 1136–1137. https://doi.org/10.1093/bioinformatics/btu761

Mukherjee, G., Lal Gupta, P., & Jayaram, B. (2015). Predicting the binding modes and sites of metabolism of xenobiotics. *Molecular Biosystems*, 11(7), 1914–1924. https://doi.org/10.1039/C5MB00118H

Mustafa, G., Yu, X., & Wade, R. C. (2014). Structure and dynamics of human drug-metabolizing cytochrome P450 enzymes. In J. Kirchmair (Ed.), Drug metabolism prediction, Vol. 38 (pp. 75–102). Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA.

Olsen, L., Rydberg, P., Rod, T. H., & Ryde, U. (2006). Prediction of activation energies for hydrogen abstraction by cytochrome p450. *Journal of Medicinal Chemistry*, 49(22), 6489–6499. https://doi.org/10.1021/jm060551l

Oostenbrink, C. (2014). Structure-based methods for predicting the sites and products of metabolism. In J. Kirchmair (Ed.), Drug metabolism prediction, Vol. 4 (pp. 243–264). Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA.

Patlewicz, G., Jeliazkova, N., Safford, R. J., Worth, A. P., & Aleksiev, B. (2008). An evaluation of the implementation of the Cramer classification scheme in the Toxtree software. *SAR and QSAR in Environmental Research*, 19(5–6), 495–524. https://doi.org/10.1080/10629360802083871

Pragyan, P., Kesharwani, S. S., Nandekar, P. P., Rathod, V., & Sangamwar, A. T. (2014). Predicting drug metabolism by CYP1A1, CYP1A2, and CYP1B1: Insights from MetaSite, molecular docking and quantum chemical calculations. *Molecular Diversity*, 18(4), 865–878. https://doi.org/10.1007/s11030-014-9534-6

Rendic, S., & Guengerich, F. P. (2015). Survey of human oxidoreductases and cytochrome P450 enzymes involved in the metabolism of xenobiotic and natural chemicals. *Chemical Research in Toxicology*, 28(1), 38–42. https://doi.org/10.1021/tx500444e

Ridder, L., & Wagener, M. (2008). SyGMa: Combining expert knowledge and empirical scoring in the prediction of metabolites. *ChemMedChem*, 3(5), 821–832. https://doi.org/10.1002/(ISSN)1860-7187

Roby, C. A., Anderson, G. D., Kantor, E., Dryer, D. A., & Burstein, A. H. (2000). St John's Wort: Effect on CYP3A4 activity. *Clinical Pharmacology and Therapeutics*, 67(5), 451–457. https://doi.org/10.1067/mcp.2000.106793

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. https://doi.org/10.1021/acs.jcim.6b00662

Rudik, A. V., Dmitriev, A. V., Lagunin, A. A., Filimonov, D. A., & Poroikov, V. V. (2014). Metabolism site prediction based on xenobiotic structural formulas and PASS prediction algorithm. *Journal of Chemical Information and Modeling*, 54(2), 498–507. https://doi.org/10.1021/ci400472j

Rudik, A., Dmitriev, A., Lagunin, A., Filimonov, D., & Poroikov, V. (2015). SOMP: Web server for in silico prediction of sites of metabolism for drug-like compounds. *Bioinformatics*, 31(12), 2046–2048. https://doi.org/10.1093/bioinformatics/btv087

Rydberg, P., Gloriam, D. E., & Olsen, L. (2010). The SMARTCyp cytochrome P450 metabolism prediction server. *Bioinformatics*, 26(23), 2988–2989. https://doi.org/10.1093/bioinformatics/btq584

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 Medicinal Chemistry Letters*, 1(3), 96–100. https://doi.org/10.1021/ml100016x

Shaik, S., Chen, H., Usharani, D., & Thiel, W. (2014). QM/MM studies of structure and reactivity of cytochrome P450 enzymes: Methodology and selected applications. In J. Kirchmair (Ed.), Drug metabolism prediction, Vol. 311 (pp. 133–178). Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA.

Shao, C.-Y., Su, B.-H., Tu, Y.-S., Lin, C., Lin, O. A., & Tseng, Y. J. (2015). CypRules: A rule-based P450 inhibition prediction server. *Bioinformatics*, 31(11), 1869–1871. https://doi.org/10.1093/bioinformatics/btv043

Šícho, M., de Bruyn Kops, C., Stork, C., Svozil, D., & Kirchmair, J. (2017). FAME 2: Simple and effective machine learning model of cytochrome P450 regioselectivity. *Journal of Chemical Information and Modeling*, 57(8), 1832–1846.

Testa, B. (2014). Metabolism in drug development. In J. Kirchmair (Ed.), Drug metabolism prediction, Vol. 1 (pp. 1–26). Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA.

Testa, B., Pedretti, A., & Vistoli, G. (2012). Reactions and enzymes in the metabolism of drugs and other xenobiotics. *Drug Discovery Today*, 17(11–12), 549–560. https://doi.org/10.1016/j.drudis.2012.01.017

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. https://doi.org/10.1021/acs.jcim.8b00035

Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2), 455–461.

Tyzack, J. D., Furnham, N., Sillitoe, I., Orengo, C. M., & Thornton, J. M. (2017). Understanding enzyme function evolution from a computational perspective. *Current Opinion in Structural Biology*, 47, 131–139. https://doi.org/10.1016/j.sbi.2017.08.003

Tyzack, J. D., & Glen, R. C. (2014). Investigating and predicting how biology changes molecules and their properties. *Molecular Informatics*, 33(6–7), 443–445. https://doi.org/10.1002/minf.201400031

Tyzack, J. D., Hunt, P. A., & Segall, M. D. (2016). Predicting regioselectivity and lability of cytochrome P450 metabolism using quantum mechanical simulations. *Journal of Chemical Information and Modeling*, 56(11), 2180–2193. https://doi.org/10.1021/acs.jcim.6b00233

Tyzack, J. D., Mussa, H. Y., Williamson, M. J., Kirchmair, J., & Glen, R. C. (2014). Cytochrome P450 site of metabolism prediction from 2D topological fingerprints using GPU accelerated probabilistic classifiers. *Journal of Cheminformatics*, *6*, 29. https://doi.org/10.1186/1758-2946-6-29

Tyzack, J. D., Williamson, M. J., Torella, R., & Glen, R. C. (2013). Prediction of cytochrome P450 xenobiotic metabolism: Tethered docking and reactivity derived from ligand molecular orbital analysis. *Journal of Chemical Information and Modeling*, *53*(6), 1294–1305. https://doi.org/10.1021/ci400058s

Vedani, A., Dobler, M., Hu, Z., & Smieško, M. (2015). OpenVirtualToxLab–a platform for generating and exchanging in silico toxicity data. *Toxicology Letters*, *232*(2), 519–532. https://doi.org/10.1016/j.toxlet.2014.09.004

Veith, H., Southall, N., Huang, R., James, T., Fayne, D., Artemenko, N., … Auld, D. S. (2009). Comprehensive characterization of cytochrome P450 isozyme selectivity across chemical libraries. *Nature Biotechnology*, *27*(11), 1050–1055. https://doi.org/10.1038/nbt.1581

Williamson, M. J. (2014). Computational free energy methods for ascertaining ligand interaction with metabolizing enzymes. In J. Kirchmair (Ed.), Drug metabolism prediction, Vol. 45 (pp. 179–198). Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA.

Wilson, I. D. (2014). Metabolite detection and profiling. In J. Kirchmair (Ed.), Drug metabolism prediction, Vol. 76 (pp. 485–498). Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA.

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. https://doi.org/10.1093/nar/gkx1037

Xiong, Y., Qiao, Y., Kihara, D., Zhang, H.-Y., Zhu, X., & Wei, D.-Q. (2018). Survey of machine learning techniques for prediction of the isoform specificity of cytochrome P450 substrates. *Current Drug Metabolism*, *19*, 1–7. https://doi.org/10.2174/1389200219666181019094526

Yousofshahi, M., Manteiga, S., Wu, C., Lee, K., & Hassoun, S. (2015). PROXIMAL: A method for prediction of xenobiotic metabolism. *BMC Systems Biology*, *9*, 94. https://doi.org/10.1186/s12918-015-0241-4

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. https://doi.org/10.1021/ci400518g