The Free Internet Journal
for Organic Chemistry

Review

Archive for
Organic Chemistry

Arkivoc **2019**, part iv, 0-0
to be inserted by editorial office

Computational methods for prediction of drug properties - application to Cytochrome P450 metabolism prediction

Mihai Burai Patrascu,[a] Jessica Plescia,[a] Amit Kalgutkar,[b] Vincent Mascitti,[b] and Nicolas Moitessier*[a]

[a] *Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec H3A 0B8, Canada*

[b] *Medicine Design, Pfizer Inc., 610 Main Street, Cambridge MA 02139, USA*

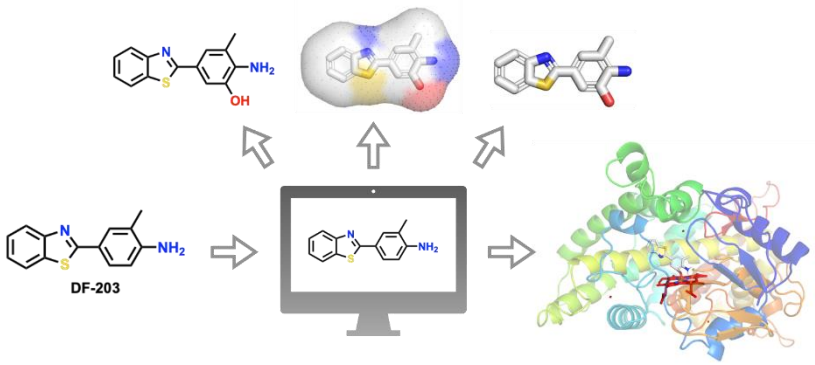
Email: nicolas.moitessier@mcgill.ca

Dedicated to Steve Hanessian, a great mentor and friend

Received 05-01-2019 **Accepted** 07-30-2019 **Published online** 10-07-2019

Abstract

Computational methods are becoming essential in the drug discovery world. Structure-based methods (i.e. docking), ligand-based methods, and machine learning are common practice. In this review, we present the major methods and their application to the prediction of cytochrome P450 (CYP)-mediated drug metabolism. More specifically, this mini-review is focused on the different methods used in predicting sites of metabolism (SoMs), and presents the advantages and disadvantages of various SoM prediction tools that are currently in use in both academia and industry.



Keywords: Sites of metabolism, machine learning, quantum mechanics, structure-based drug design, ligand-based drug design

Contents

1. Introduction

Adverse drug reactions and toxicity are among the major causes of attrition observed in the drug discovery and development processes. Despite major investments in toxicology and in clinical trials, severe and even fatal toxic effects have resulted in drug withdrawals. In fact, analysis revealed that 30% of the pharmaceutical attrition was driven by toxicity, with 90% of drug withdrawals and 33% of clinical phase terminations attributed to various forms of toxicity.[1-2] Although the primary causes of toxicity can be very different, the first-pass bioactivation by metabolic enzymes (phase 1 metabolism) such as cytochrome P450s (CYPs) is often an initiating step. Of drugs currently on the market, 75-90% are metabolized by one of the 57 human CYPs. Out of this set, six isoforms (CYP1A2, 2C9, 2C19, 2D6, 2E1 and 3A4), expressed mainly in the liver and in the gut, are responsible for over 90% of this oxidative metabolism and represent one of the main focus for medicinal chemists and pharmacologists when optimizing for drug-like properties.[3-5] In phase 2 metabolism, the drug (or the metabolite from phase 1) is conjugated to water-soluble moieties (e.g., glucuronic acid), which facilitates its excretion. The metabolites produced by these various processes have their own intrinsic pharmacological effect and toxicity that may differ from the parent drug. They can also exhibit high reactivity, leading to, for instance, hepatotoxicity and/or cancer, and are referred to as reactive metabolites (Figure 1).

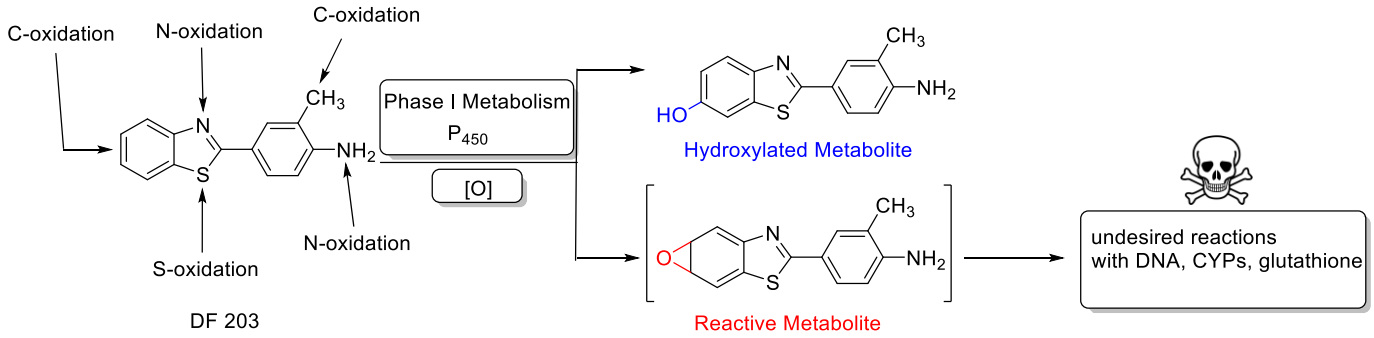


Figure 1. Potential metabolic pathways of anticancer drug prototype DF 203.[6-7]

Interestingly, about 75% of the drugs withdrawn due to adverse drug reactions were in fact activated into reactive metabolites.[2] Over the years, medicinal chemists have relied on structural alerts (functional groups known to possess high toxicity potentials), to flag drug candidates that could potentially form reactive metabolites. Thus, in their study of the top 200 drugs prescribed in the USA, Stepan et al. showed that approximately 80% of the drugs associated with toxicity contained structural alerts.[2] However, this empirical approach has limitations, as not all structural alerts lead to toxicity and additional parameters such as daily dose and body burden have to be taken into account. Moreover, a further significant limitation of using structural alerts is that the absence of these alerts is not indicative of compound benignity. As an alternative, the computational prediction of reactive metabolite might represent a more viable and efficient approach. Specifically, the ability to predict reactive metabolites relies on the ability to accurately predict the sites of metabolism (SoMs) on a molecule of interest. In this context, several methods that predict SoMs are discussed: ligand-based (quantum mechanics, machine learning), structure-based (docking, molecular dynamics), hybrid (both ligand- and structure-based) and rule-based methods. In this general mini-review, we will highlight advantages and disadvantages of each of these methods and we will provide representative examples of computational tools that fit in these categories. For further in-depth information about each method, the reader is encouraged to consult more specialized and focused reviews.[8-9]

2. CYP Metabolism Mechanism

CYP may contribute to the metabolism of xenobiotics in a number of ways. Among the common reactions are hydrogen abstraction – which may lead to *N*-dealkylation, alcohol and aldehyde oxidation among other reactions – sulfur oxidation, oxidative deboronation, and aromatic oxidation, to name a few.

3. Computational Methods

In the past two decades, a multitude of methods have been developed for predicting xenobiotic metabolism. Amongst these are ligand-based methods, which focus on the properties of ligands such as reactivity, topological and molecular descriptors, and activation energies.[10-17] These methods disregard the metabolizing enzymes, in terms of their polarization induced by proximity to amino acids, the ligand's binding mode within their active site, etc. These ligand-based methods range from computationally-intensive quantum mechanics (QM), to fast computations such as machine learning (ML) techniques. With the advent of a significant

increase in computational power for even a single user, these methods have become widely applicable. Opposite to ligand-based methods are structure-based methods, including docking and molecular dynamics (MD), which put an emphasis on the metabolizing enzymes themselves, including but not limited to binding poses, interactions with ligands, conformational changes, and reaction mechanisms.[18-22] Additionally, there exist expert-developed rule-based methods, based on years of *in vivo* and *in vitro* empirical data. Rule-based approaches consist of identifying a target fragment in a molecule of interest, followed by generation of a potential metabolic product and subsequent product search in a pre-existing catalogue. However, this approach gives rise to a plethora of potential products which might prove to be troublesome, as one needs to select an appropriate metabolic product from the generated ones.

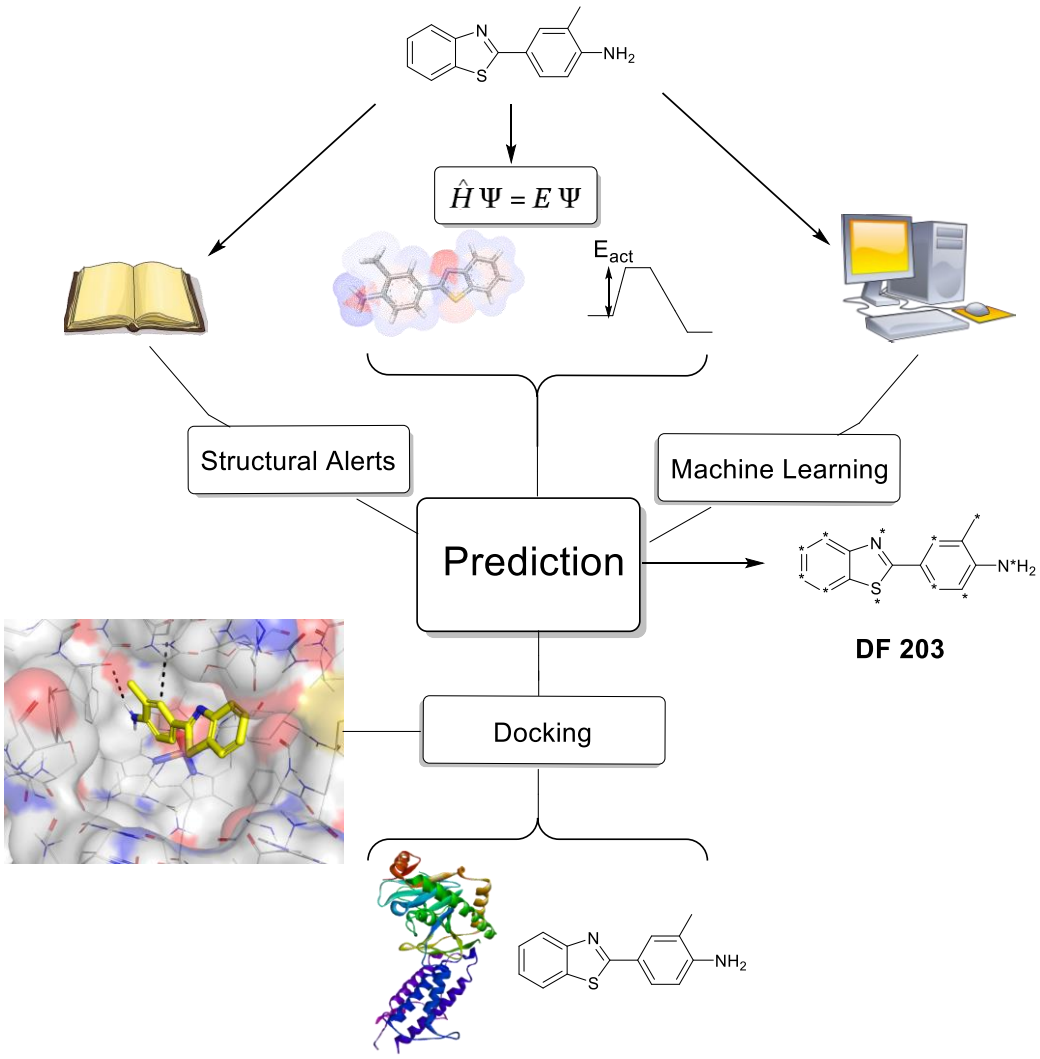


Figure 2. Summary of methods used to predict sites of metabolism.

3.1 Ligand-based methods

3.1.1 Quantum mechanics (QM). Since the 1950s the field of computational chemistry has been in a continuous expansion afforded by major breakthroughs, improved methodologies and increased computational power. Within the field of computational chemistry, QM approaches play a major role. The development of the Roothaan-Hall equations in 1951,[23-24] the Kohn-Sham equations in 1965[25] and the intermediate neglect of differential overlap (INDO) and neglect of diatomic differential overlap (NDDO)

methods of Pople in the 1970s[26-27] gave rise to some of the most popular computational methods in use today (semi-empirical methods, *ab initio* Hartree-Fock (HF) and Density Functional Theory (DFT)). These methods (discussed below) have proven to be indispensable in solving problems ranging from reaction mechanisms to molecular reactivity and predicting sites of metabolism. In the past decade or so, routine calculation (single point energies, geometry optimizations) using these methods have become more accessible to experimental chemists fuelled by the ongoing development of user-friendly QM packages such as ORCA[28] and GAMESS.[29] As such, computational chemistry is slowly becoming an integral part of an experimental chemist's toolbox.

3.1.2 Semi-empirical methods. Semi-empirical QM (SE-QM) methods are based on the Hartree-Fock formalism but fundamentally differ from HF by their use of empirical parameters and approximation (or omission) of electronic interactions, whose computation is the most expensive part of any HF or DFT calculation. To account for the lack of these interactions, the results of semi-empirical methods are empirically trained to predict experimental observations. By neglecting the electronic interactions, SE-QM methods are significantly less computationally expensive than HF or DFT; they have been used extensively in situations where the size of the molecule or molecular complex under scrutiny makes HF or DFT calculations intractable (e.g., greater than 200 atoms). Following the seminal work of Pople in the 1970s many SE-QM methods have been developed, with some of the most widely used models being AM1,[30] PM3,[31] PM6,[32] and RM1.[33] However, when using SE-QM methods one must be cautious - if the molecule of interest is not similar to those for which the method was parametrized the results can be both qualitatively and quantitatively wrong.

Limitations. In the field of drug discovery, SE-QM methods are particularly attractive for predicting sites of metabolism for both small and large drug molecules due to their speed and possibility of application on libraries of thousands of compounds. A major limitation of these QM-based properties is the lack of consideration of the enzyme as these methods are only considering the intrinsic reactivity of the small molecule. Thus, CYP selectivity cannot be predicted nor whether the ligands are actual substrates of any given CYP isoform.

Available methods. One method developed for small drug molecule metabolism around SE-QM is **CYPScore**,[10] currently in use at Bayer Schering Pharma among other places. **CYPScore** uses atomic reactivity descriptors obtained with the AM1 SE-QM model for seven CYP-catalyzed reactions, ranging from aliphatic hydroxylation to sulphur oxidation. These atomic descriptors include bond orders, solvent accessible surface area, atomic valence, atomic surface area, Coulson charge etc. and are used in the generation of an individual model for each type of catalyzed reaction. These models were validated on four differently designed datasets – in three out of four datasets over 60% of all major phase I metabolites were identified, furthermore in 70% of the compounds an active metabolite was found using the top 2 metabolites and 85% using the top 3 metabolites. One advantage of the **CYPScore** algorithm is that competing metabolic reactions are treated on the same reactivity scale, meaning that various metabolic positions and metabolites can be compared between molecules. With regards to availability, **CYPScore** is available as a free trial but ownership requires licensing. An improved version of **CYPScore**, called **MetScore**,[34] will be discussed in the machine learning section.

Another method based on the AM1 SE-QM model is **E_aMEAD**[13] (Activation energy of Metabolism reactions with Effective Atomic Descriptors), which predicts the activation energies E_a of four CYP-catalyzed reactions: aliphatic hydroxylation, N-dealkylation, O-dealkylation and aromatic hydroxylation.

The sites of metabolism with lowest E_a are the ones most likely to undergo CYP-mediated metabolism. To be able to predict the E_a of a given reaction empirical models were built using a set of compounds for which the E_a and atomic reactivity descriptors (effective atomic charge, effective atomic polarizability, and bond dipole moments) were computed using AM1. The choice of AM1 for predicting the E_a was based on the results obtained by Korzewka *et al*[35] which correlated well with experimental results while also being significantly

faster than DFT and other ab initio methods. **E_aMEAD** was shown to accurately predict E_a for all four reactions and was used to predict the sites of metabolism of 46 compounds oxidized by CYP3A4. The accuracy of predicting the correct metabolites was ~60% when considering the first two lowest E_as. However, while accuracy was shown for CYP3A4 (responsible for the metabolism of almost half of the drugs currently on the market), the rest of the major CYP isoforms were not tested. As such, if one desires to use **E_aMEAD** one should stay within the parameters of the published data. To the best of our knowledge, this program is not available for download.

If a versatile method is required for predicting the sites of metabolism on isoforms not considered by the methods presented above, then **StarDrop**[15] is a reliable option. The P450 metabolism module of **StarDrop** is based on pre-computed activation energies using SE-QM model AM1 calculations using high quality experimental and *ab initio* data. **StarDrop** can predict the reactivity of each site of metabolism while taking into account the molecular environment in approximately 1-2 minutes/compound on a single CPU. Based on these calculations, ligand-based models were built, with additional contributions from data correcting for steric and orientation effects. As such, along the seven CYP isoforms considered in the study, the accuracy of the method was between 82-91%. However, all seven isoforms are trained on relatively large sets (76-220 compounds) while the testing sets are fairly small (27-84 compounds), which might have an impact on the established accuracy. Moreover, while StarDrop is available as a free trial for anyone willing to try it out, ownership requires licensing.

3.1.3 DFT methods. While SE-QM methods are useful in obtaining fast and moderately accurate results, there are many cases in which higher accuracy is required due to the nature of the system under scrutiny. As a consequence, higher level methods have been developed, amongst which DFT is the most popular. Initially developed by Kohn and Hohenberg (and later by Kohn and Sham) in the 1960s, DFT fundamentally differs from SE-QM and HF methods in the fact that it only requires the electronic density for determining the ground state properties of many-electron systems. Although developed in the 1960s, DFT saw a rise in popularity in the field of computational chemistry only in the 1990s, with the ground-breaking work laid by Becke and others in the area of computing the contribution of electronic motions (exchange and correlation) to the total energy of the system, which includes the now famous B3LYP functional. Ever since, DFT has made major strides to overcome its limitations (e.g., describing intermolecular interactions, which is crucial in properly describing the behaviour of macromolecules).

Available tools. Considering the versatility, relative low-computational cost (although much higher than for SE-QM methods) and accuracy of DFT, the field of drug discovery slowly turned its attention to it, with several DFT-based methods being developed to predict sites of metabolism. One such method is **QMBO**,[12] which uses bond orders determined at the B3LYP/3-21G level of theory to compute C-H bond strengths. The relatively low level of theory used in the DFT computations makes **QMBO** calculations fairly fast. Based on these bond strengths, the weakest C-H bond is located and the hydrogen abstraction is predicted to take place at that bond. Moreover, a steric hindrance restriction on the reactivity is also considered through a penalty term (proportional to a function dependent on the solvent accessible surface area of each hydrogen atom), which is then applied to the bond strength. The method was tested on CYP3A4 and CYP2C9, for which it showed good prediction rates (~60% using the Top-1 metric and ~85% using the Top-3 metric) for both isoforms. Some disadvantages of this method include i) the small basis set used to compute the bond orders, which might give rise to inaccurate bond strengths, and ii) its applicability to only two of the six major CYP isoforms involved in xenobiotic metabolism.

Another method that showed excellent results on various CYP isoforms is **SMARTCyp**.[11, 36] Originally designed to be accurate on CYP3A4 and CYP2D6 , **SMARTCyp** is based on DFT-derived activation energies (B3LYP/6-

31G* for the ligand) of various ligand fragments which are then stored in a database, making **SMARTCyp** one of the fastest programs currently available for predicting CYP metabolism. **SMARTCyp** uses 2D structures that are then matched to existing fragments in the database of precomputed activation energies and, in conjunction with an accessibility parameter, makes predictions regarding the site of metabolism. A possible disadvantage of this method comes from the assumption that the fragment will be found within the database. If that is not the case, a prediction cannot be made. This concern was addressed in an enhanced version of **SMARTCyp** called **SMARTCyp 3.0**[37] which contains a similarity measure between the query molecule and the model fragment, along with a more comprehensive database of fragments. Moreover, the CYP isoform predictivity was expanded to include CYP2C9 as well. Across 394 compounds, using the Top-1 metric the accuracy of the original version of **SMARTCyp** was ~65%, while when using the Top-3 metric it increased to ~80%. Across 475 compounds, **SMARTCyp 3.0** showed an excellent area under the curve (AUC) for all three isoforms (3A4 – 0.73, 2D6 – 0.78, 2C9 – 0.76). If the relative low coverage of the CYP isoforms is not an issue, **SMARTCyp** is available for free at https://smartcyp.sund.ku.dk/mol_to_som.

3.2 Structure-based methods

Method: molecular docking. One branch within structure-based methods is molecular docking which entails using a macromolecular crystal structure (ideally co-crystallized with a ligand, either endogenous or exogenous), placing a ligand of interest, and observing the resultant ligand-target interactions within the binding site. Many software programs have been designed for the purpose of docking ligands of interest to proteins (e.g., AutoDock, FlexX, GLIDE, GOLD, FITTED, etc.).[18-22] Most programs consist of two steps/algorithms - conformational search and scoring. The former involves fitting the ligand into the active site while the later determines the most likely (e.g., lowest energy) ligand-protein conformation and/or predicts the binding free energy. Although scoring functions differ from program to program, they generally consist of sums of energy terms, such as the energies of hydrogen bonds or van der Waals interactions between the ligand and protein.
In terms of docking and scoring in programs predicting metabolism, there are some slight differences. The CYP in their reactive form feature an oxidized heme complex while the crystal structures are often in the resting state (coordinated to a water molecule). Ideally, scoring would therefore require incorporating interactions with this oxidized heme often not parameterized in docking programs.

Limitations. Of particular interest for molecular docking are the six CYP isoforms 1A2, 2C9, 2D6, 3A4, 2C19, and 2E1 (see Introduction). The most versatile of all these isoforms is CYP 3A4. The binding site of CYP 3A4 is very large and, virtually and biologically, can accommodate most ligands, including large macrolides such as Erythromycin (

Figure 3).[38] It is therefore difficult to use structure-based methods to assess whether or not a ligand is likely to be metabolized by CYP 3A4. Although it *is* known that a majority of xenobiotics are metabolized by this CYP,[38] it would be difficult to deduce from docking predictions. However, other isoforms such as 1A2, 2C9, 2C19, 2E1 and 2D6 allow for more effective use of structure-based methods, as their active sites are smaller and more selective in terms of which functional groups participate in stabilizing interactions with key residues.

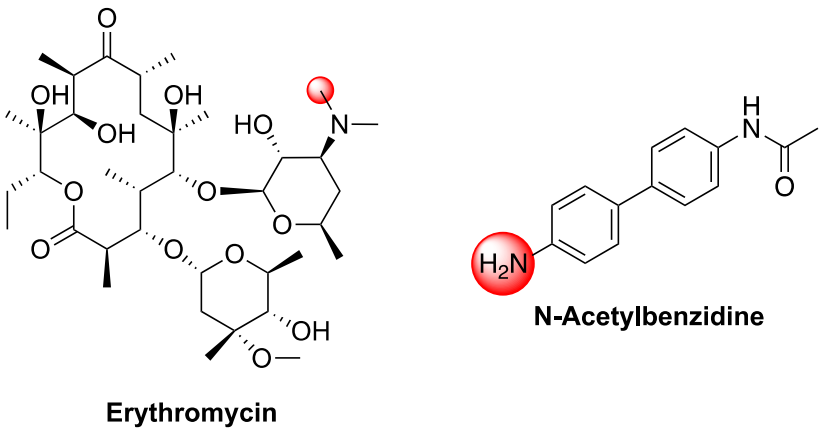


Figure 3. Known substrates of major CYP enzyme isoforms. Sites of metabolism highlighted in red.

For example, while CYP 1A2 is most reactive toward aromatic amines and nitrogen-containing heterocycles,[39] its active site is very narrow and contains mostly lipophilic and aromatic residues. Therefore, upon docking ligands of interest containing these functional groups it would be possible to see whether or not the enzyme active site would be likely to accommodate and complement the docked ligand. Such is the case with *N*-acetylbenzidine (

Figure 3), a flat aromatic lipophilic amine, which is a known substrate of CYP 1A2,[39] and fits well in the 1A2 active site upon docking. In the case of CYP 2D6, the binding site is narrow and therefore can accommodate flat lipophilic substrates. Furthermore, the substrate scope of CYP 2D6 tends towards lipophilic bases with aromatic rings.[40]

Figure 4 shows the docked pose of the CYP 2D6 substrate DF 203, a flat, basic, aromatic amine.

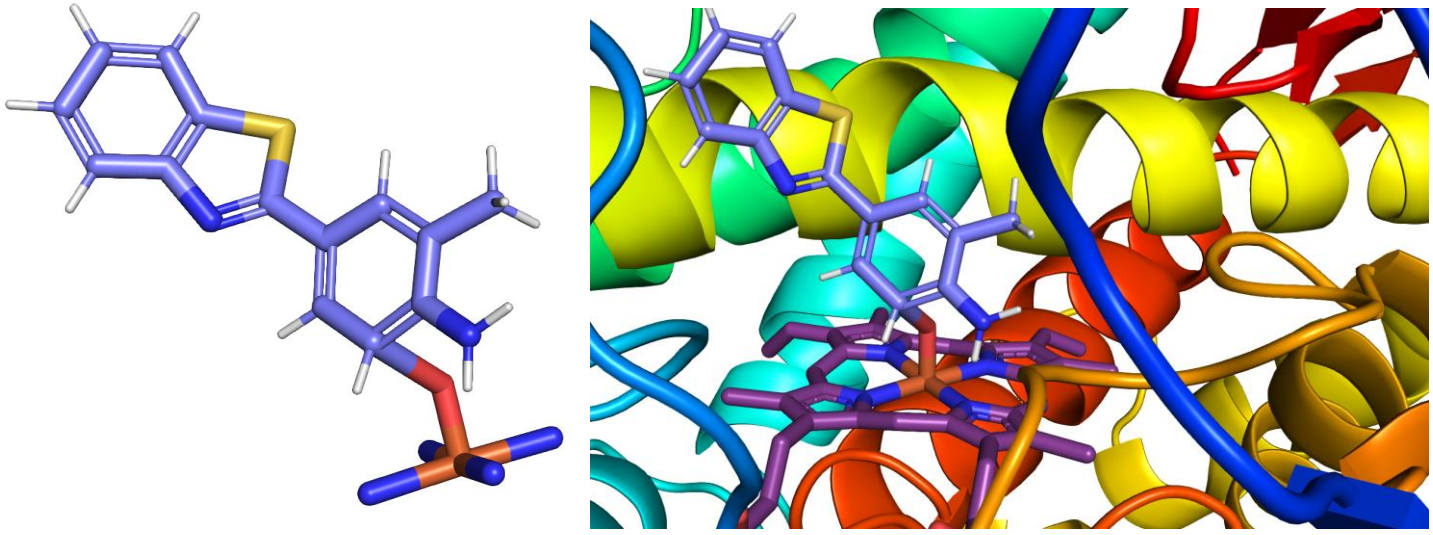


Figure 4. Left: DF 203's predicted site of metabolism, aromatic hydroxylation, shown bound to the heme iron; Right: DF 203 docked to the active site of its metabolizer CYP 2D6[41] using IMPACTS[42] on the FORECASTER platform.[43] Stabilizing interactions are shown with dashed lines (black).

One further limitation of docking-based methods is the failure of the software to account for reactivity of the cytochrome enzyme. Most reactivity predictions are ligand-based to determine the most likely SoM and are

incorporated into some docking programs.[42] However, the enzyme reactivity requires more time-consuming and computationally expensive studies, normally consisting of a combination of quantum mechanics and molecular mechanics.[44] In most medicinal chemistry endeavors, however, the potential accuracy of predictions is sacrificed for the convenience of a simple docking experiment.

3.3 Hybrid methods

In order to account for both the protein environment and enzyme reactivity, as well as for ligand reactivity, several tools have been developed that combine docking with ligand reactivity determination methods. Among these is the reactivity-binding model proposed by Jung *et al*.[45] who tried to predict the regioselectivity of biotransformations performed by CYP1A2. Their model proposed docking compounds to the crystal structure of CYP1A2, followed by determination of activation energies (based on AM1 calculations) of the metabolism reactions in which the predicted binding poses were involved. The energy terms obtained from docking and AM1 calculations were then used to compute the metabolic probability that a site would be preferred for a biotransformation. This approach provided excellent results, with the preferred site being predicted for 8 out of 12 compounds, although this set may be too small to draw conclusions. However, while this study provided an important insight into hybrid methods, it only involves one CYP isoform and a low number of tested compounds.

Another hybrid method, this time describing CYP3A4 metabolism, was developed by Oh *et al*.[45] The method, termed **MLite**, describes four CYP3A4-mediated reactions: aliphatic hydroxylation, *N*-dealkylation, *O*-dealkylation and aromatic hydroxylation. To describe the accessibility of a compound inside the active site of CYP3A4 Oh *et al.* implemented the ensemble catalyticphore-based docking method, along with quantum mechanical calculations for the ligand reactivity predictions. The method was trained on a small set of 47 molecules and tested on 25 – the success rate based on the top2 metric was 76%. **Metasite**, a hybrid method-based program developed by Cruciani *et. al.*,[46] allows the user to predict both the potential cytochrome enzyme (of CYP1A2, CYP2C9, CYP2C19, CYP2D6, or CYP3A4) and the potential SoM of a compound of interest. This program utilizes GRID flexible molecular interaction fields (GRID-MIFs) to characterize the CYP enzyme and GRID probe pharmacophore recognition to characterize the ligand of interest. **Metasite** assesses the compatibility of the enzyme-ligand pairs, in terms of accessibility in the enzyme active site and of reactivity of the ligand and provides the user with the predicted SoM. Upon testing this program against CYP substrates provided by several pharmaceutical companies, the predictions typically achieved high accuracy, ranging from 83-90%, with one 2D6-specific set achieving 62%.

Our research group has been involved in the development of our own hybrid method – **IMPACTS**.[42] This tool is based on our docking program FITTED[18], that takes care of providing an accurate binding pose of a compound of interest in the active site of a CYP isoform of interest. The docking is paired with pre-computed activation energies obtained at the B3LYP/6-31G* level of theory for a series of relevant fragments that are then assigned to the molecule of interest.[42] Across 4 CYP isoforms, **IMPACTS** has showed excellent accuracies, ranging from 75% (considering the top2 metric) for CYP3A4 to 82% for 2C9. Overall, on average, the accuracy of **IMPACTS** is 77%. Importantly, our tool was extensively tested on over 700 drugs and drug-like compounds, which makes it reliable for a broad range of compounds. Another important consideration of **IMPACTS** is that it is fully implemented within our drug-discovery platform **FORECASTER**, which is freely available for academia.[42, 47]

3.4 Rule-based / substrate-based methods

Methods. Over the years, experts in the fields of medicinal chemistry, biology, toxicology, and pharmacology have developed databases of structural alerts, or "warning" functional groups, via

comprehensive studies of hepatotoxic drugs and drug metabolites, among others. It is known that liver-related toxicity is associated with common toxicophores, including but not limited to alkynes and electron-rich aromatic heterocycles that either become oxidized into toxic metabolites or directly bind to the heme iron and inhibit the liver's cytochrome P450 enzymes.

Figure 5 gives three hepatoxic compounds: Gestodene, an oral contraceptive; OSI-930, a terminated oncology drug candidate; and L-739010, a terminated anti-asthma drug. All three were discovered to be CYP3A4 inhibitors.[48-49]

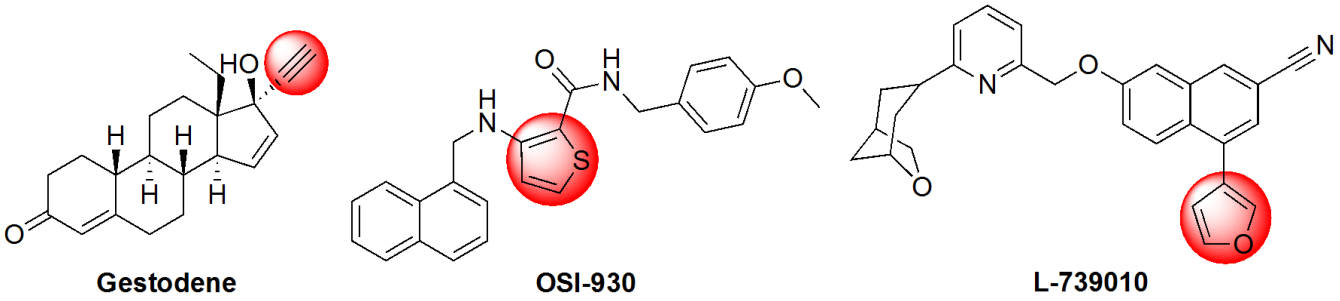


Figure 5. Selected hepatotoxic compounds, toxicophores highlighted in red.[48]

Although these functional groups, among many others, are present in many hepatotoxic drugs or pre-clinical drug candidates, a potential drug candidate containing a known potential toxicophore may not present risks, depending on the stereoelectronic environment of the motif, as well as the projected daily dose of the molecule. To aid medicinal chemists in designing drugs that are less likely to be toxic, several more advanced computational rule-based methods have been developed. This category of tools is normally referred to as "expert systems". [8] These programs are based on human knowledge of biotransformations instead of, for example, virtual ligand-CYP docking or computational predictions of relative atom energies. Predictive software based on experts' knowledge normally involves analyzing structural fragments of the drug candidate and searching a biotransformation dictionary of fragment metabolites, considering all possibilities. Examples of such software include **Meteor**,[50-53] which ranks metabolites by predicted probability, and **Metabolizer**,[54] which generate all possible metabolites of a drug candidate and includes predictions of non-human metabolism (e.g. bacteria). Despite the availability of these tools, using these expert-based predictive programs leads to a virtually endless library of possible drug metabolites. This overprediction of potential toxicity makes the lead optimization process much more difficult.

One additional method for SoM prediction is the alignment-based method. This alignment-based approach involves comparing the molecule of interest against a database of reference molecules for which the SoM is known. Sykes *et. al.* applied this method to study the reaction of cytochrome P450 2C9 with the common antibiotic Flurbiprofen. The authors overlaid Flurbiprofen with its known metabolic sites against a small data set of known CYP2C9 substrates and applied their alignment method to predict the known SoMs of the compounds in this data set. Their method utilized the alignment program **ROCS** and achieved high accuracy, ranging from 73-89%.[55] Another group, de Bruyn Kops *et. al.*,[56] expanded upon the Sykes study. They utilized the aforementioned alignment method and combined it with a reactivity-based prediction to account for intrinsic reactivity in compounds structurally different from the reference databases. This hybrid approach

improved predictive accuracy significantly, further demonstrating that these types of complex predictions require more than one source of calculation. [56]

3.5 Machine learning

3.5.1 Methods. Since the advent of Quantitative Structure-Activity Relationship (QSAR) methods in the 1960s, several groups have attempted to relate ADMET properties to simple molecular/atomic properties. LogP is an example of a widely used descriptor known to correlate with water solubility, and to some extent phase 1 metabolism and passive permeability. More advanced machine learning techniques (e.g., random forest, support-vector machine and artificial neural networks) have since then been considered.[57] **Random Forest (RF)** is a learning method first reported in the 1990s which produces sets of decision trees. In practice, a molecule can be given to these decision trees and be classified as either a CYP3A4 substrate or not. **Support-vector machine (SVM)** methods are other learning techniques which process data for classification (e.g, CYP substrate or not) and regression analysis. The most popular of the Artificial intelligence algorithms, **artificial neural networks (ANN)** is a set of algorithms mimicking the behaviour of biological neural networks. In practice, ANN are a set of matrices connecting neurons. These matrices are derived from experimental data so that giving an input (e.g., a 2D structure), the output (e.g. CYP3A4 substrate or not) can be accurately predicted. Over the past few years, ANN have been extensively used to predict ADMET properties.[58] Another approach that has been used in ADME prediction is Bayesian statistics. This approach enables a predictor to evolve or refine its models using additional data. As an example, if molecule A is oxidized by CYP3A4 and not molecule B, it is difficult from this data to predict whether C will be. If now, we just learn that A is aromatic and B is not, if C is aromatic the probability that it will be oxidized has raised.

3.5.2 Limitations. Although promising, these methods must be trained on high quality data in order to produce highly accurate predictions. However, the publicly available data may be too diverse to be useful. For example, CYP3A4 metabolism can be measured as the disappearance of the substrate in human liver microsomes (which may include several CYP isoforms) in the presence and absence of known CYP3A4 inhibitors and could also be measured in lysosomes containing a single CYP isoform. The condition of these assays may also differ from one laboratory to another and, as we discussed previously, would lead to non-uniform data. Unfortunately, a predictive method can rarely be better than the data used to train it. The size and diversity of the dataset are also critical for optimal applicability domain. For example, a model trained only of a set of chemical series (e.g., benzodiazepines, sulfamides,…) should not be expected to be predictive with very different chemical series (e.g., aminoglycosides and peptides). Too small of a dataset may lead to overtraining of the model and poor predictivity. To address these issues, training is carried out on atom environment rather than entire molecules, hence covering the space more efficiently and on carefully curated datasets such as the one reported by Zaretzki et al.[59]

3.5.3 Available tools. Among the tools developed using machine learning techniques is **MetScore**. Designed as an improved version of **CYPScore**, **MetScore** can accurately predict both phase 1 and 2 metabolism of xenobiotics. Relying on molecular representations built on quantum chemical partial charges that were used to build RF models, **MetScore** showed an excellent performance in predicting diverse phase 1 and 2 reactions (Matthews correlation coefficient of 0.61 for phase 1 and 0.76 for phase 2). Its versatility makes it an interesting SoM tool, however it is only available as an in-house tool for Bayer's research platform Plx. Another important tool is **Fast Metabolizer (FAME)** developed by Glen and co-workers with version 2 reported in 2017.[14, 60] **FAME** predicts the site of metabolism of drugs using random forest models trained on large dataset with an accuracy of 81% using the top-2 metrics and requiring only a 2D model of the substrate. **FAME** relies on a number of descriptors such as partial atomic charges, electronegativity, delocalizability **(FAME2.0)**, topological descriptors (distance between atoms) and hybridization. Interestingly, **FAME** is not limited to CYP-

mediated phase I metabolism but trained on various metabolism processes (e.g., oxidation, hydrolysis, reduction, acetylation, conjugation such as glucuronidation) and trained for species-dependent (human, dog or rat) metabolism. Applied to a CYP-mediated metabolism set, the performance of the first version on a CYP substrate set (73% with top-2 metrics) was below that of **SmartCYP** (79%). The second version showed enhanced accuracy (up to 94%). Other descriptors and classifiers were evaluated by Fu et al. and demonstrated the high level of accuracy that RF can provide.[61] Another example is **RS-Predictor** from Zaretzki *et al.* reported in 2011.[16] This program was trained using a combination of topological descriptors (motifs) and quantum descriptors (e.g., aromatic orbital electron density, electrophilicity index, interactions from Mulliken population analysis). RS-Predictor is an example of SVM-like program (it uses MIRank, a derivative of SVM). This program demonstrated an accuracy of 74.5% (top-2 metrics) outperforming **SMARTCyp** (72%) and **StarDrop** (59.2%) on an external set. A recently developed method by Finkelmann *et al*.[62] uses RF models built upon atomic descriptors that describe the electronic (through quantum mechanics) and steric environment of an atom and its precise location within a molecular environment. These models were tested on the same set used for **XenoSite**,[17] giving excellent results (top2 metric of 90.3% in leave-one-molecule-out cross-validation).
An example of the use of neural network for site of metabolism prediction is **XenoSite**. **XenoSite** neural network was trained using quantum descriptors, atomic and molecular descriptors (e.g., logP, water accessible surface area). This program was found to outperform all of the other 5 programs tested including **RS_Predictor, SMARTCyp** and **StarDrop**. Interestingly, although **XenoSite** was trained on CYP substrates, it was trained on substrates of different isoform separately implicitly considering the CYP isoform when making predictions.
Bayesian approaches have also been exploited for site of metabolism prediction as exemplified by **SOMP**[63-64] which was derived from PASS (prediction of activity spectra for substances). **SOMP** was trained to predict CYP-mediated SoM (1A2, 2C9, 2C19, 2D6 and 3A4) and phase II glucuronidation with accuracy ranging from 61% (top-2 metrics on 2D6) to 95% (UGT). This program is available for free on a server at http://www.way2drug.com/SOMP.
As most ligand-based methods, all of these machine learning models have a drawback. Although they can accurately predict the SoM, information provided to medicinal chemists is only about the SoM and not how the oxidation process occurs. While this information is useful to block given sites, knowing the binding in the P450s (e.g., through docking) may allow medicinal chemists to propose modifications away from the SoM which will disrupt the binding and modulate the metabolism.
As an advantage, these methods are very fast and allow interactive design.

4. Conclusions and Prospect

Several strategies have been envisioned to predict *in silico* drug properties, and cytochrome P450 enzyme metabolism prediction is no exception. As we discussed, docking-based methods, QM-based methods, machine learning and combination of these have been envisioned and led to several methods now available to the medicinal chemistry community. However, modeling this complex process remains of limited accuracy. On one side, the accuracy of ligand-based methods ignoring the binding process is limited. For example, the binding site specificities of some CYP isoforms may preclude reactions even at highly reactive substrate sites. On the other side, the docking-based methods are considering proximity to the heme to identify the site of metabolism. This approach assumes the binding is a static process. While a binding mode might be favored, it may present an unreactive site to the reactive heme. An alternative binding mode which may not be as

favored may present a highly reactive site and may be the most productive binding mode. This dynamics process led researchers to consider hybrid methods where enzyme binding and reactivity are considered. Although this approach is expected to improve accuracy, there are some remaining factors, such as enzyme flexibility, still poorly accounted for by the current methods and so is enzyme inhibition (programs assume small molecules are substrates). With this complexity in mind, machine learning was perceived as a promising alternative as training on isoform specific substrate sets should theoretically consider all the factors. However, the lack of sufficiently large and diverse substrate sets renders the development and testing difficult.
Despites these various limitations, dozens of methods are available. Overall, these tools demonstrate reasonable accuracies and can certainly be used to guide the design of drug candidates at the preclinical stage.

Acknowledgements

NM thanks NSERC for funding, JP thanks the faculty of science for a McGill University Molson and Hilton Hart Fellowship in Science.

References

1. Barton, P.; Riley, R. J. *Drug Discovery Today* **2016,** *21*, 72-81. https://doi.org/10.1016/j.drudis.2015.09.010
2. Stepan, A. F.; Walker, D. P.; Bauman, J.; Price, D. A.; Baillie, T. A.; Kalgutkar, A. S.; Aleo, M. D. *Chem. Res. Toxicol.* **2011,** *24*, 1345-1410. https://doi.org/10.1021/tx200168d
3. Dalvie, D.; Kalgutkar, A. S.; Chen, W. *Drug Metab. Rev.* **2015,** *47* , 56-70. https://doi.org/10.3109/03602532.2014.984813
4. Stjernschantz, E.; Vermeulen, N. P. E.; Oostenbrink, C. *Expert Opinion Drug Metab. Toxicol.* **2008,** *4*, 513-527. https://doi.org/10.1517/17425255.4.5.513
5. Guengerich, F. P. *Chem. Res. Toxicol.* **2008,** *21*, 70-83. https://doi.org/10.1021/tx700079z
6. Bradshaw, T. D.; Stevens, M. F.; Westwell, A. D. *Curr Med Chem* **2001,** *8*, 203-10. https://doi.org/10.2174/0929867013373714
7. Kashiyama, E.; Hutchinson, I.; Chua, M. S.; Stinson, S. F.; Phillips, L. R.; Kaur, G.; Sausville, E. A.; Bradshaw, T. D.; Westwell, A. D.; Stevens, M. F. G. *J. Med. Chem.* **1999,** *42*, 4172-84. https://doi.org/10.1021/jm990104o
8. Kirchmair, J.; Göller, A. H.; Lang, D.; Kunze, J.; Testa, B.; Wilson, I. D.; Glen, R. C.; Schneider, G. *Nature Reviews Drug Discovery* **2015,** *14*, 387. https://doi.org/10.1038/nrd4581
9. Kirchmair, J.; Williamson, M. J.; Tyzack, J. D.; Tan, L.; Bond, P. J.; Bender, A.; Glen, R. C. *J. Chem. Inf. Model.* **2012,** *52*, 617-648. https://doi.org/10.1021/ci200542m
10. Hennemann, M.; Friedl, A.; Lobell, M.; Keldenich, J.; Hillisch, A.; Clark, T.; Göller, A. H. *ChemMedChem* **2009,** *4*, 657-669

https://doi.org/10.1002/cmdc.200800384
11. Rydberg, P.; Gloriam, D. E.; Zaretzki, J.; Breneman, C.; Olsen, L. *ACS Med. Chem. Lett.* **2010,** *1*, 96-100. https://doi.org/10.1021/ml100016x
12. Afzelius, L.; Hasselgren Arnby, C.; Broo, A.; Carlsson, L.; Isaksson, C.; Jurva, U.; Kjellander, B.; Kolmodin, K.; Nilsson, K.; Raubacher, F.; Weidolf, L. *Drug Metab. Rev.* **2007,** *39*, 61-86. https://doi.org/10.1080/03602530600969374
13. Kim, D. N.; Cho, K.-H.; Oh, W. S.; Lee, C. J.; Lee, S. K.; Jung, J.; No, K. T. *J. Chem. Inf. Model.* **2009,** *49*, 1643-1654. https://doi.org/10.1021/ci900011g
14. Kirchmair, J.; Williamson, M. J.; Afzal, A. M.; Tyzack, J. D.; Choy, A. P. K.; Howlett, A.; Rydberg, P.; Glen, R. C. *J. Chem. Inf. Model.* **2013,** *53*, 2896-2907. https://doi.org/10.1021/ci400503s
15. Tyzack, J. D.; Hunt, P. A.; Segall, M. D. *J. Chem. Inf. Model.* **2016,** *56*, 2180-2193. https://doi.org/10.1021/acs.jcim.6b00233
16. Zaretzki, J.; Bergeron, C.; Rydberg, P.; Huang, T.-w.; Bennett, K. P.; Breneman, C. M. *J. Chem. Inf. Model.* **2011,** *51*, 1667-1689. https://doi.org/10.1021/ci2000488
17. Zaretzki, J.; Matlock, M.; Swamidass, S. J. *J. Chem. Inf. Model.* **2013,** *53*, 3373-3383. https://doi.org/10.1021/ci400518g
18. Corbeil, C. R.; Englebienne, P.; Moitessier, N. *J. Chem. Inf. Model.* **2007,** *47*, 435-449. https://doi.org/10.1021/ci6002637
19. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997,** *267*, 727-748. https://doi.org/10.1006/jmbi.1996.0897
20. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.* **2004,** *47*, 1739-1749. https://doi.org/10.1021/jm0306430
21. Forli, S.; Huey, R.; Pique, M. E.; Sanner, M. F.; Goodsell, D. S.; Olson, A. J. *Nature protocols* **2016,** *11*, 905-919. https://doi.org/10.1038/nprot.2016.051
22. Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.* **1996,** *261*, 470-489. https://doi.org/10.1006/jmbi.1996.0477
23. Roothaan, C. C. J. *Rev. Modern Phys.* **1951,** *23*, 69-89. https://doi.org/10.1103/RevModPhys.23.69
24. Hall, G. G.; Lennard-Jones J. E. *Proc. Roy. Soc. London. Ser. A. Math. Phys. Sci.* **1951,** *205*, 541-552. https://doi.org/10.1098/rspa.1951.0048
25. Kohn, W.; Sham, L. J. *Phys. Rev.* **1965,** *140*, A1133-A1138. https://doi.org/10.1103/PhysRev.140.A1133
26. Pople, J. A.; Beveridge, D. L.; Dobosh, P. A. *J. Chem. Phys.* **1967,** *47*, 2026-2033. https://doi.org/10.1063/1.1712233
27. Pople, J. A.; Beveridge, D. L., *Approximate Molecular Orbital Theory,* 1970.
28. Neese, F., Software update: the ORCA program system, version 4.0. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2018,** *8*, e1327. https://doi.org/10.1002/wcms.1327

29. Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery Jr, J. A. *J. Comp. Chem.* **1993,** *14*, 1347-1363. https://doi.org/10.1002/jcc.540141112
30. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985,** *107*, 3902-3909. https://doi.org/10.1021/ja00299a024
31. Stewart, J. J. P. *J. Comp. Chem.* **1989,** *10*, 221-264. https://doi.org/10.1002/jcc.540100209
32. Stewart, J. J. P. *J. Mol. Model.* **2007,** *13*, 1173-1213. https://doi.org/10.1007/s00894-007-0233-4
33. Rocha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J. P. *J. Comp. Chem.* **2006,** *27*, 1101-1111. https://doi.org/10.1002/jcc.20425
34. Finkelmann, A. R.; Goldmann, D.; Schneider, G.; Göller, A. H. *ChemMedChem* **2018,** *13*, 2281-2289. https://doi.org/10.1002/cmdc.201800309
35. Jones, J. P.; Mysinger, M.; Korzekwa, K. R. *Drug Metab. Disposition* **2002,** *30*, 7.
36. Rydberg, P.; Gloriam, D. E.; Olsen, L. *Bioinformatics* **2010,** *26*, 2988-2989. https://doi.org/10.1093/bioinformatics/btq584
37. Olsen, L.; Montefiori, M.; Tran, K. P.; Jørgensen, F. S. *Bioinformatics* **2019**, ASAP.
38. Kacevska, M.; Robertson, G. R.; Clarke, S. J.; Liddle, C. *Expert Opinion Drug Metab. Toxicol.* **2008,** *4*, 137-149. https://doi.org/10.1517/17425255.4.2.137
39. Lakshmi, V. M.; Zenser, T. V.; Davis, B. B. *Drug Metab. Disposition* **1997,** *25*.
40. Zhou, S.-F.; Liu, J.-P.; Lai, X.-S. *Cur. Med. Chem.* **2009,** *16*, 2661-2805. https://doi.org/10.2174/092986709788681985
41. Chua, M. S.; Kashiyama, E.; Bradshaw, T. D.; Stinson, S. F.; Brantley, E.; Sausville, E. A.; Stevens, M. F. G. *Cancer Res.* **2000,** *60*, 5196-203.
42. Campagna-Slater, V.; Pottel, J.; Therrien, E.; Cantin, L. D.; Moitessier, N. *J. Chem. Inf. Model.* **2012,** *52*, 2471-83. https://doi.org/10.1021/ci3003073
43. Therrien, E.; Englebienne, P.; Arrowsmith, A. G.; Mendoza-Sanchez, R.; Corbeil, C. R.; Weill, N.; Campagna-Slater, V.; Moitessier, N. *J Chem. Inf. Model.* **2012,** *52*, 210-24. https://doi.org/10.1021/ci2004779
44. Tyzack, J. D.; Kirchmair, J. *Chem. Biol. Drug Design* **2019,** *93*, 377-386. https://doi.org/10.1111/cbdd.13445
45. Jung, J.; Kim, N. D.; Kim, S. Y.; Choi, I.; Cho, K.-H.; Oh, W. S.; Kim, D. N.; No, K. T. *J. Chem. Inf. Model.* **2008,** *48*, 1074-1080. https://doi.org/10.1021/ci800001m
46. Cruciani, G.; Carosati, E.; De Boeck, B.; Ethirajulu, K.; Mackie, C.; Howe, T.; Vianello, R. *J. Med. Chem.* **2005,** *48*, 6970-6979. https://doi.org/10.1021/jm050529c
47. Moitessier, N.; Pottel, J. Molecular Forecaster, Inc. http://www.molecularforecaster.com (accessed 2019).
48. Smith, G. F. *Prog. Med. Chem.* **2011,** *50*, 1-47.
49. Palovaara, S.; Kivistiö, K. T.; Tapanainen, P.; Manninen, P.; Neuvonen, P. J.; Laine, K. *J. Clin. Pharmacol.* **2000**, 333-337.

https://doi.org/10.1046/j.1365-2125.2000.00271.x
50. Darvas, F.; Marakhazi, S.; Kormos, P.; Kulkarni, G.; Kalasz, H.; Papp, A., *Databases and High Throughput Testing During Drug Design and Development*. Blackwell Science, Ltd.: Gamburg, UK, 1999.
51. Ellis, L. B. M.; Hou, B. K.; Kang, W.; Wackett, L. P. *Nucleic Acids Res.* **2003,** *31*, 262-265. https://doi.org/10.1093/nar/gkg048
52. Greene, N.; Judson, P. N.; Langowski, J. J.; Marchant, C. A. *Environ. Res.* **1999,** *10*, 299-314. https://doi.org/10.1080/10629369908039182
53. Talafous, J.; Sayre, L. M.; Mieyal, J. J.; Klopman, G. *J. Chem. Inf. Model.* **1994,** *34*, 1326-1333. https://doi.org/10.1021/ci00022a015
54. ChemAxon. Metabolizer. https://docs.chemaxon.com/display/docs/Metabolizer.
55. Sykes, M. J.; McKinnon, R. A.; Miners, J. O. *J. Med. Chem.* **2008,** *51*, 780-791. https://doi.org/10.1021/jm7009793
56. de Bruyn Kops, C.; Friedrich, N. O.; Kirchmair, J., *J. Chem. Inf. Model.* **2017,** *57*, 1258-1264. https://doi.org/10.1021/acs.jcim.7b00165
57. Yang, X.; Wang, Y.; Byrne, R.; Schneider, G.; Yang, S. *Chem. Rev.* **2019**, ASAP.
58. Zhou, Y.; Cahya, S.; Combs, S. A.; Nicolaou, C. A.; Wang, J.; Desai, P. V.; Shen, J. *J. Chem. Inf. Model.* **2019**, *59*, 1005-1016. https://doi.org/10.1021/acs.jcim.8b00671
59. Zaretzki, J.; Rydberg, P.; Bergeron, C.; Bennett, K. P.; Olsen, L.; Breneman, C. M. *J. Chem. Inf. Model.* **2012,** *52*, 1637-1659. https://doi.org/10.1021/ci300009z
60. Šícho, M.; de Bruyn Kops, C.; Stork, C.; Svozil, D.; Kirchmair, J. *J. Chem. Inf. Model.* **2017,** *57*, 1832-1846. https://doi.org/10.1021/acs.jcim.7b00250
61. Fu, X.; He, S.; Du, L.; Lv, Z.; Zhang, Y.; Zhang, Q.; Wang, Y. *Biochem. Pharmacol.* **2018,** *152*, 302-314. https://doi.org/10.1021/acs.jcim.7b00250
62. Finkelmann, A. R.; Göller, A. H.; Schneider, G. *ChemMedChem* **2017,** *12*, 606-612. https://doi.org/10.1002/cmdc.201700097
63. Rudik, A. V.; Dmitriev, A. V.; Lagunin, A. A.; Filimonov, D. A.; Poroikov, V. V. *J. Chem. Inf. Model.* **2014,** *54*, 498-507. https://doi.org/10.1021/ci400472j
64. Rudik, A.; Dmitriev, A.; Lagunin, A.; Filimonov, D.; Poroikov, V. *Bioinformatics* **2015,** *31*, 2046-2048. https://doi.org/10.1093/bioinformatics/btv087

Authors' Biographies

Mihai Burai Patrascu received his B.Sc in Chemistry from Jacobs University Bremen (Bremen, Germany) in 2015. In the same year he moved to Montréal where he joined Prof. Moitessier's group at McGill University to pursue his PhD in Computational Chemistry. As a doctoral candidate his work focuses on both software development (drug discovery platform FORECASTER) and computational chemistry (nucleoside modeling, CYP-related metabolism, asymmetric catalysis), with a significant interest in interfacing organic and computational chemistry.

Jessica Plescia received her B.Sc. Hon. in Biochemistry from McGill University (Montréal, Canada) in 2014. She later joined the Moitessier lab in the chemistry department to pursue her doctorate in medicinal chemistry. As a doctoral candidate, her research focuses on computationally-aided design of anti-cancer therapeutics and the synthesis of chiral, bicyclic peptidomimetics as reversible covalent inhibitors of serine protease targets.

Amit Kalgutkar is a research fellow in the Medicine Design group at Pfizer. He received his Ph.D. degree from Virginia Tech, VA, and carried out post-doctoral research at Vanderbilt University, TN, prior to joining Pfizer. He has ~ 20 years' experience in drug discovery/development, spanning multiple therapeutic areas with over 15 investigational drug candidates nominated for clinical development. He has published over 160 peer-reviewed papers, reviews and book chapters and holds several patents. He is presently on the editorial boards of *Chemical Research in Toxicology* (American Chemical Society), *Drug Metabolism and Disposition* and *Xenobiotica*. Besides his current position in Pfizer, he is also an Adjunct Professor at the Department of Biomedical and Pharmaceutical Sciences, School of Pharmacy, University of Rhode Island.

Vincent Mascitti received his diploma in chemical engineering from the ECPM (Strasbourg, France). He completed his Ph.D. with Professor Hanessian (University of Montreal, Canada) on the total synthesis of natural products bearing deoxypropionate motifs (e.g., doliculide and borrelidin), and the synthesis of bioactive oligosaccharides. He did his postdoctoral studies in the laboratories of Professor E. J. Corey where he completed the first total synthesis of the ladderane-containing natural product pentacycloanammoxic acid. Vincent joined Pfizer in 2006, where as a medicinal chemist in the CVMED chemistry department, he contributed to various diabetes and obesity related projects. In particular, Vincent was the driving force behind the design and synthesis of SGLT2 inhibitor Ertugliflozin (PF-04971729). Ertugliflozin, along with its fixed dose combinations with metformin and dpp4 inhibitor sitagliptin, were developed in phase 3 in partnership between Pfizer and Merck and have been recently approved by the FDA and EMA as a family of medicines for type 2 diabetes treatment. These medicines are now marketed under the brand names Steglatro™, Segluromet™, and Steglujan™. Vincent also contributed to the discovery of novel ASGPr ligands that have been used in the deal between Pfizer and WaVe Life Sciences focused on the tissue selective delivery of stereo enriched oligonucleotides to the liver. Beyond small molecules drug discovery, Vincent also

has expertise in additional therapeutic modalities. In particular, he contributed to the design and evaluation of engineered CRISPR-Cas9 endonucleases for cell-type-specific gene editing. He is the (co)author of 60 publications and patent applications and is currently a Senior Director at Pfizer in the medicinal chemistry department.

Nicolas Moitessier received his PhD from Université Henri Poincaré (Nancy, France). In 1998, he moved to Montréal where he joined Prof. Hanessian's group for post-doctoral work. In 2001, he moved back to Nancy to start an academic career then back to Montréal in 2003 (McGill University). His current research interests integrate software development, computational chemistry and organic/medicinal chemistry. He has published nearly 100 papers, is a co-inventor on 2 patents (potential drugs) and has two registered copyrights (software). In 2010, he co-founded Molecular Forecaster Inc., a company distributing the computational platforms and providing service to the pharmaceutical and biotechnological industry. In 2016, he was appointed as an Editor of the European Journal of Medicinal Chemistry.

