Development of a Computational Tool to Rival Experts in the Prediction of Sites of Metabolism of Xenobiotics by P450s

Valérie Campagna-Slater,[†,‡,||] Joshua Pottel,[†,||] Eric Therrien,[†] Louis-David Cantin,[§] and Nicolas Moitessier*,[†]

[†]Department of Chemistry, McGill University, 801 Sherbrooke St W, Montreal, QC, Canada H3A 0B8
[§]AstraZeneca R&D Montréal, 7171 Frédérick Banting, St. Laurent, Québec, H4S 1Z9, Canada

Ⓢ *Supporting Information*

ABSTRACT: The metabolism of xenobiotics—and more specifically drugs—in the liver is a critical process controlling their half-life. Although there exist experimental methods, which measure the metabolic stability of xenobiotics and identify their metabolites, developing higher throughput predictive methods is an avenue of research. It is expected that predicting the chemical nature of the metabolites would be an asset for designing safer drugs and/or drugs with modulated half-lives. We have developed IMPACTS (In-silico Metabolism Prediction by Activated Cytochromes and Transition States), a computational tool combining docking to metabolic enzymes, transition state modeling, and rule-based substrate reactivity prediction to predict the site of metabolism (SoM) of xenobiotics. Its application to sets of CYP1A2, CYP2C9, CYP2D6, and CYP3A4 substrates and comparison to experts' predictions demonstrates its accuracy and significance. IMPACTS identified an experimentally observed SoM in the top 2 predicted sites for 77% of the substrates, while the accuracy of biotransformation experts' prediction was 65%. Application of IMPACTS to external sets and comparison of its accuracy to those of eleven other methods further validated the method implemented in IMPACTS.
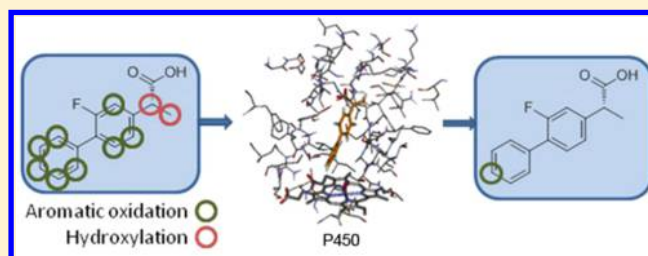
INTRODUCTION

The cytochrome P450s (CYPs) are a group of heme-containing enzymes involved in the metabolism of various xenobiotics and endogenous compounds. In particular, they are involved in the phase-I metabolism of most drugs currently on the market. A majority of these biotransformations are carried out by only 5 isoforms out of the 57 P450s in the human genome, namely CYP1A2, 2C9, 2C19, 2D6, and 3A4.[1−3] CYP-mediated chemical modifications of drugs affect their pharmacokinetic properties as microsomal stability often correlates with hepatic clearance, and hence with the half-life of drugs in the patient's body. In addition, the produced metabolites can themselves have a pharmacologic effect and intrinsic toxicity.[1] During the development stage, a number of chemical modifications of the lead compounds are often required to reach an acceptable pharmacokinetic profile and to produce a drug candidate. Thus, predicting the metabolic stability of drugs and the binding mode of small molecules in metabolic enzymes in a high throughput manner became a promising avenue of research. In fact, accurately predicting sites of metabolism (SoMs) and the binding mode of small molecules in metabolic enzymes could be useful to flag potential in vitro or in silico hits, help prioritize experiments, provide key insights enabling the design of more stable compounds with modulated half-life, predict metabolites that may induce toxicity (e.g., CYP1A2-mediated oxidation of aniline leads to carcinogenic metabolites[2]) or even investigate the polymorphism of CYP enzymes and their marked interindividual variability.[3] These multiple applications were the impetus for the development of computational approaches or protocols to predict P450 metabolism of small molecules,[4−6] with a great body of work from the Rydberg group.[7−9] These can be classified as ligand-based (e.g., quantitative structure–activity relationships,[10] pharmacophore, quantum mechanical-derived rules,[8,11] descriptors[12]), reactivity-based (e.g., calculation of activation energies of each potential reactive center by DFT or semiempirical calculations such as in CypScore or fragment recognition such as in SMARTCyp[8]), and structure-based (e.g., docking) methods.[5,13−18] A number of methods predicting SoMs have been devised but as stated in a recent perspective article,[6] *most consider a single aspect of the reaction*, as illustrated by ligand-based methods[11] which do not account for the recognition of the substrates by the CYP enzymes. In parallel, structure-based methods are often tested on a single CYP and their transferability to other CYP remains unknown.[19,20] Ultimately, it is expected that predictions would be more accurate if the method considered both CYP protein structures as well as ligand chemical reactivity. Approaches combining ligand reactivity and protein structures have been devised as illustrated by the pioneering work from Cruciani et al. (MetaSite[21]) and Oh and co-workers (MLite),[22] which use a nonatomistic representation of the enzymes.
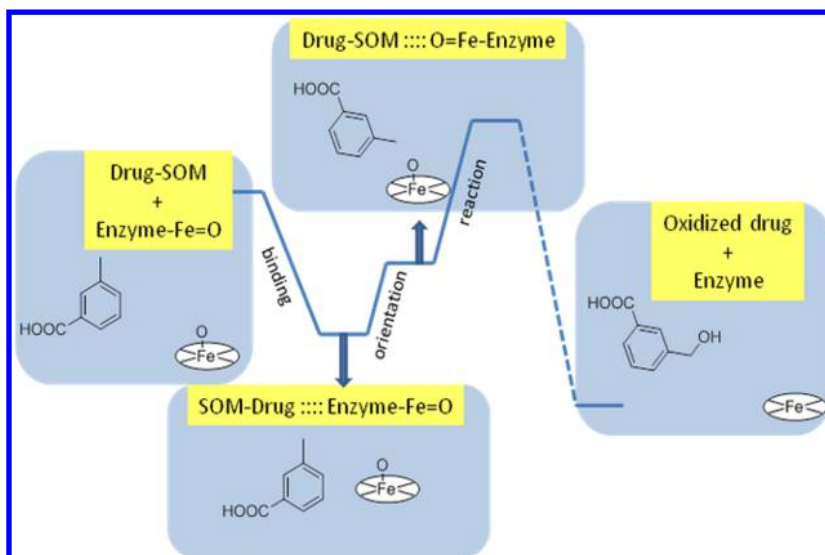
Figure 1. Investigated steps in the P450-mediated drug oxidation.

However, despite these efforts, little has been reported on the significance of the predictions. We report herein our efforts toward the development of a fully automated program that combines molecular docking, ligand reactivity estimation and transition state structure modeling to predict the SoM of drugs, with a focus not only on accuracy but also on significance of these predictions. Furthermore, our predictions were compared to those made by biotransformation experts, which revealed the usefulness of such a program.

■ THEORY AND IMPLEMENTATION

Docking and P450-Mediated Metabolism. Accurately docking small molecules to enzymes requires high resolution protein structures. As of today, crystal structures have been solved for about one-third of the human P450 isoforms (including four of the most important five listed above, with 2C19 not yet crystallized), making docking possible.[23−26] However, compared to traditional noncovalent drug docking, predicting P450 substrates and SoMs adds a level of difficulty. First, enzymatic catalysis does not only depend on protein–ligand noncovalent binding. After the substrate binding event, a reactive orientation of the substrate SoM with respect to the heme is required, allowing the oxidation reaction to take place. The efficiency of the biotransformation also depends on the intrinsic reactivity of the ligand reactive site. In several reported studies, putative substrates have been docked to a P450 structure of interest (either an X-ray structure or homology model[27]) using standard docking programs, and docking poses obtained have been used to predict the most likely metabolites based on distances to the heme group.[19] However, this approach does not take into account intrinsic reactivity of putative SoMs, does not model the chemical transformation, nor does it discriminate between inhibitors and substrates.

Docking and Drug Reactivity. In order to account for ligand reactivity while docking, rule-based approaches for predicting activation energies based on density functional theory (DFT) calculations, combined with docking to CYP1A2 were proposed to predict SoMs of substrates from both binding energy (based on docking scores) and intrinsic energy (obtained from a rule-based method).[28] Very recently, approaches considering docking to flexible P450[29] and ligand

reactivity to predict SoMs of drugs were disclosed.[7,30] Two other main limitations in accurately modeling CYPs, is their promiscuity due to receptor flexibility,[31,32] and the presence of water molecules, which may also be important in drug/CYP binding.[29] In fact, in a recent review, Tarcsay and Keserü listed three major issues to address for accurate docking-based approaches: poor scoring, protein flexibility, and the presence of water molecules.[5]

Docking, Drug Reactivity, and Transition State. To reproduce the process of CYP-mediated metabolism and consider both the thermodynamics of the ligand/enzyme binding and the thermodynamics of the chemical transformation, we thought to combine a docking program and a transition state (TS) modeling program. Docking of TS structures has been previously reported, although the method required development of a TS for each drug prior to the actual docking.[9] In contrast to other docking studies,[19,33] our method not only investigates the protein–ligand noncovalent binding ("binding" in Figure 1) but also imposes a proper orientation of the substrate with respect to the heme ("orientation" in Figure 1) and considers the intrinsic reactivity of the ligand reactive site ("reaction" in Figure 1).[34] This approach should provide a more accurate prediction of the activation energy including the distortion from optimal TS geometries. In addition, the absence of CYP-specific training throughout the development of this method and its validation on four very different CYPs assessed its transferability to other CYPs.

Although the complete CYP-mediated oxidation cycle includes several steps, the SoM is selected in the subprocess shown in Figure 1.

Development of IMPACTS. We have developed and implemented the framework into a SoM prediction program (IMPACTS, In-silico Metabolism Prediction by Activated Cytochromes and Transition States), which uses some modified routines of our FITTED docking program[35,36] to predict the CYP-mediated metabolism of small molecules (Figure 2). This fully automated program predicts the most likely site(s) of reaction and transition state (TS) structures of small molecules when reacting with the CYP heme when in the oxidized form..

Identifying SoMs and Their Reactivity. First, a database of fragments and their corresponding reactivity (in the form of
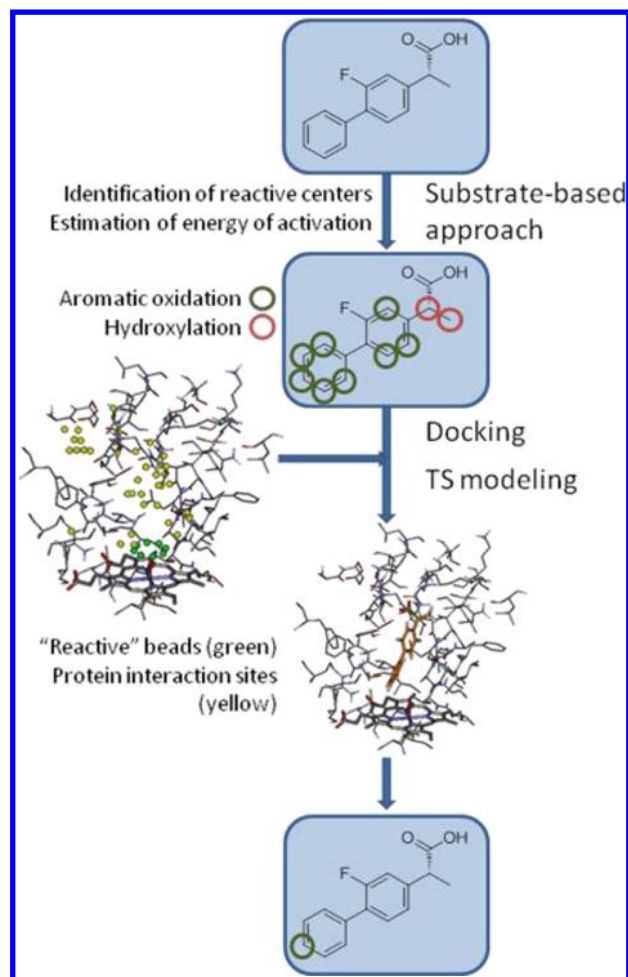
Figure 2. Fully automated protocol implemented in IMPACTS.

Figure 3. Correlation between activation energies (E_a) relative to benzene derived using the methoxy model and the full heme model.[38] $r^2 = 0.86$.

was implemented in IMPACTS to compute the effects of pairs. The computed relative activation energies are given in Table 1.

Figure 4. Correlation between activation energies (E_a) of bifunctionalized benzene derivatives (e.g., *m*-nitro, *p*-methoxy-benzene) relative to benzene and the sum of the relative energies of individual monofunctionalized benzene derivatives (e.g., sum of *m*-nitro-benzene and *p*-methoxy benzene). $r^2 = 0.96$.

activation energies, Supporting Information) was built. Previously computed data[11,37] was supplemented with additional density functional theory (DFT) calculations. For instance, exhaustive DFT calculations were carried out to include the impact of various electron-withdrawing and electron-donating groups on the reactivity of phenyl rings. The heme system was modeled using a simpler methoxy radical, and energies relative to benzene were computed. In order to test this model, correlation with activation energies obtained using the full heme model was computed (Figure 3). Substituents at all positions of the benzene (except that with methoxy) were considered.

We then considered pairs of substituents. All combinations of fifteen groups were carried out and only the minimum was considered for the same positions on mirror sides. The additive effects of these groups were clearly demonstrated. For instance, calculations indicated that the presence of both *p*-OMe and *o*-Me stabilized the radical transition state by 4.5 kcal/mol relative to unsubstituted benzene while *p*-OMe and *o*-Me alone induced a stabilization of 2.4 and 2.3 kcal/mol, respectively. A few exceptions arose with cases of π-stacking or hydrogen bonding. These were evaluated on an individual basis and a static correction factor could be applied to these cases in order to match the additive effect observed. Interestingly, we found that the stabilizing/destabilizing effect of pairs of groups (relative to unfunctionalized benzene) equals the sum of the effect of individual groups (Figure 4). As a result, a simple additive rule
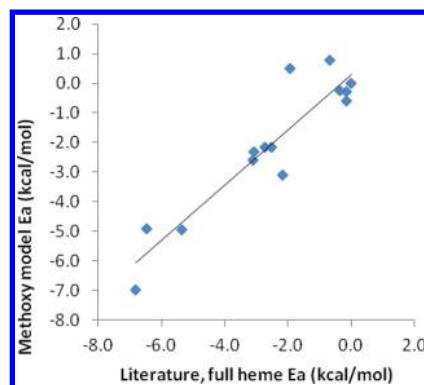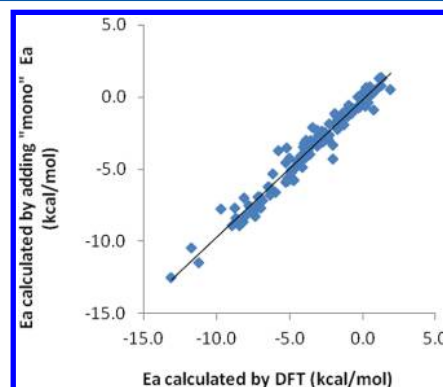
We then turned our attention to hydrogen abstraction. Similar to the aromatic oxidation, we have found that the methoxy model correlated well with the full heme model (Figure 5) and enabled the computation of a large number of derivatives. However, in contrast to aromatic oxidation, the saddle point was deemed a necessary calculation for hydrogen abstractions and was thus included. In this case, no additive effects were observed and each necessary pair had to be computed (Table 2). The activation energies were computed following eq 1.

$$E_a = E_{a(\text{derivative, methoxy model})} - E_{a(\text{ethyl, methoxy model})}$$
$$+ E_{a(\text{ethyl, full heme model})} \tag{1}$$

Next, our program initially developed to prepare ligands for docking,[35] SMART, was modified such that it can identify all putative reactive sites and stabilizing/destabilizing neighboring groups (e.g., *p*-OMe) in a small molecule. This is achieved by comparing the substrate to a database of fragments and groups including those in Tables 1 and 2, and assigning the corresponding activation energy. A routine identifying equivalent SoMs was also implemented into SMART (e.g., both

Table 1. Computed Activation Energies for Functionalized Benzene (Aromatic Oxidation)[b]

Fragment	Ortho	Meta	Para	Fragment	Ortho	Meta	Para
(benzene)	20.9[a]	20.9[a]	20.9[a]	(benzene)	20.9[a]	20.9[a]	20.9[a]
OH	-2.25	-0.35	-2.30	CF₃	0.04	0.62	0.65
OMe	-3.35	-0.44	-2.45	C(=O)NHMe	0.81	0.52	-0.21
O-C(=O)Me	-1.74	-0.45	-0.67	C(=O)O-Me	-1.27	0.40	-0.46
NH₂	-5.72	-1.72	-5.15	CN	-0.71	0.97	-0.67
NH-Me	-6.24	-0.36	-4.90	CH=CH-CH₃	-2.49	-0.91	-3.62
NMe₂	-7.33	-0.18	-5.28	Ph	-4.64	-2.02	-3.87
NH-Ph	-4.67	-0.87	-5.55	Cl	-0.65	0.65	-0.67
N(Ph)Me	-6.85	-1.04	-7.76	F	-0.96	0.34	-0.45
NH-C(=O)Me	-1.98	-0.13	-2.46	S-Me	-2.59	0.00	-3.08
NO₂	-0.76	0.96	0.29	pyridine	0.79	-0.43	1.27
Me	-2.29	-0.39	-0.91	indole	1: -4.84 3: -3.04	2: -1.69 4: -4.03	

[a]See ref 11. [b]Relative energies (kcal/mol).

ortho positions of a monosubstituted phenyl or atoms equivalent through planar symmetry).

IMPACTS. The activation energy values which are precomputed by SMART are read into IMPACTS (third term in eq 1). The drug/CYP complex is then built by docking the molecule into the CYP active site using a modified version of the hybrid matching algorithm (MA)/genetic algorithm (GA) from our docking program, FITTED.[35,40] When the MA

is applied within IMPACTS, a reactive group is positioned near the P450 heme while another two pharmacophoric groups are randomly selected and placed on complementary interaction site beads located in the CYP enzyme (Figure 2). The energy of the system is then computed (eq 2).

In addition to docking substrates, IMPACTS models the TS as a linear combination of the ground state noncovalently bound substrate and the covalently bound substrate (first two
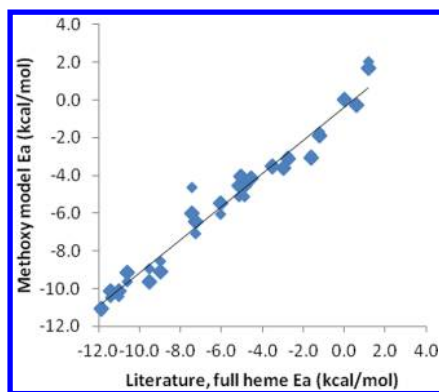
Figure 5. Correlation between activation energies relative to ethyl, isopropyl, and *tert*-butyl derived using the methoxy model and the full heme model.[11,39] $r^2 = 0.97$.

terms in eq 2). This approach originates from the hypothesis that the TS structure, according to the Hammond−Leffler principle, is located between the reactants and products and can indeed be described as a linear combination of the products and reactants. This approach has been successfully applied in the prediction of TSs with our ACE program, which predicts the stereochemical outcome of asymmetric reactions.[41,42] As in ACE, a weighing factor (λ), allows the user to shift the TS closer to the noncovalently or covalently bound structure following the chemical transformation (first two terms in eq 2). However, in contrast to ACE, IMPACTS can handle multiple potential reactive sites simultaneously, each leading to a specific TS with its corresponding list of interactions. Herein, an interaction list is defined as a series of bonds, angles, torsions, and out-of-plane terms used to describe the structure and energy of the ligand, together with the van der Waals and electrostatic interactions between the ligand and the protein. The mechanisms of commonly observed reactions[43] have been implemented in this first version of the program, including hydrogen abstraction[38] and aromatic oxidation.[38] The code has been written such that adding a new reaction is straightforward and can be done within a few minutes by the developers as long as the reaction mechanism is known.

$$E_{TS} = (1 - \lambda)E_{substrate} + \lambda E_{product} + E_a + E_{corr} + E_{dock}$$
$$(2)$$

$$Score_{TS} = E_a + S_{dock} \qquad (3)$$

With this method, the TS is located at the minimum of the linear combination of energy functions (Figure 6). However, as an artifact of the method, this minimum is associated with a nonzero energy (e.g., nearly 80 kcal/mol in Figure 6) which is very dependent on the reaction mechanism. As this value is much lower for hydrogen abstraction than for aromatic oxidation (Figure 7), the relative energies of these two reactions cannot be accurately compared unless correcting terms are included (fourth term in eq 2). To compute these correcting terms, a library of small fragment molecules was built and the energy terms associated with the TSs were computed separately using a modified version of IMPACTS and kept for reference.

Data Sets. In order to measure prediction accuracy, sets of substrates of CYP1A2, CYP2C9, CYP2D6, and CYP3A4 were assembled starting from previously reported sets.[7,19,30,44] In order to evaluate not only the accuracy but also the significance of the predictions, we went back to the primary literature which

revealed problems to be considered. First, data available for some compounds was conflicting. For example, the metabolism of Voriconazole by CYP2C9 and Selegiline by CYP2D6 has been observed by Hyland et al.[45] and Rittenbach et al.,[46] respectively, while recombinant CYP2C9 and CYP2D6 had no detectable Voriconazole oxidation activities and Selegiline oxidation activities, respectively, in studies from Murayama et al.[47] and Hidestrand et al.[48] In two separate reports, CYP2C9 was or not involved in the metabolism of Losartan.[49] Similarly, while CYP2C9 was found to be a major CYP for the metabolism of Sertraline in one report, it was also found to be reacting weakly in another.[50,51] Second, in several cases, drugs experimentally tested were racemic mixtures (e.g., Rosiglitazone[52]), and no data was given on the individual enantiomers. In addition, it is well-known that CYP-mediated metabolism may have some level of stereoselectivity (e.g., CYP2D6 preferentially hydroxylates (L)-Trimipramine and preferentially demethylates its enantiomer[53] and (R)-Bufuralol is preferentially hydroxylated by CYP2D6[54]). In some cases, the wrong enantiomer was given in the published set, and although this would not impact QSAR models, this would expectedly impact the apparent accuracy of any protein structure-based methods such as MetaSite or IMPACTS. Third, depending on the cells used to express the recombinant enzymes CYP1A2, CYP2D6, and CYP3A4 (human B-lymphoblast cells or baculovirus-infected insect cells), whether they were coexpressed with NADPH−CYP oxidoreductase or not, the azelastine N-demethylase activity of these three CYPs varied significantly (by more than 2 orders of magnitude). Thus, CYP2D6 has been found to be the most reactive CYP by Nakajima et al.[55] while CYP1A2 was identified as the most effective by Imai et al.[56] Fourth, the substrate concentration is also a factor as shown with the metabolism of FLU-1,[57] a metabolite of Fluvastatin. At a concentration of 1 μM, CYP3A4 was the major metabolizing enzyme and CYP2C9 did not react while CYP2C9 was the major metabolizer at a concentration of 100 μM. This same concentration dependence was observed for Sertraline.[51] In some reports, the reported data was unclear. For example, the controversy about the role of CYP3A4 in the metabolism of Ochratoxin A was mentioned[58] but not discussed and CMV423 was described as being oxidized by CYP1A2 > CYP3A4 > CYP2C9 > CYP2D6 although CYP1A2 was less reactive. In this report, the experimental data was presented as an odd result.[59] As an explanation of this variability of results, the role of DMSO as a CYP inhibitor was mentioned by Pearce and co-workers as the formation of some metabolites of Carbazepine were not observed by others who used DMSO-containing incubation mixtures.[60]

In some cases, several metabolites had been found but only one was investigated. For example, three major metabolites for DA-8159 have been found in rats.[61] However, a single one (N-dealkylation) has been investigated with human CYPs, identifying CYP3A4 as the major enzyme.[62] Although the 7-hydroxylation of Chlorpromazine was investigated in 2000,[63] it was not until 2010 that the other three major metabolites were investigated.[64] Similarly, Nitrendipine and Nifedipine which are dihydropyridines were aromatized by CYP3A4. However it was mentioned that "it must be concluded that that P4503A4 is able to oxidize other portion of some of these molecules".[65] Zolpidem was metabolized into 3 major metabolites (M3, ca. 65%, M4, ca. 25%, M11, ca. 10%), but the contributions of CYP1A2 (ca. 8%), 2C9 (ca. 31%), 2D6 (ca. 17%), 2C19 (ca. 2%), and 3A4 (ca. 40%) in the metabolism of Zolpidem has

Table 2. Computed Activation Energies for Hydrogen Abstraction (kcal/mol)

Fragment	$\Delta\Delta G$	Fragment	$\Delta\Delta G$	Fragment	$\Delta\Delta G$
Me–CH, H, H	21.12[a]	Me, Me–CH, H	18.39[a]	Me, Me–CH, Me	17.85[a]
Ph–CH, H, H	16.42	Ph, Me–CH, H	14.39	Ph, Me–CH, Me	16.18
vinyl–CH, H, H	16.00	vinyl, Me–CH, H	14.35	vinyl, Me–CH, Me	15.51
MeO–CH, H, H	15.04	MeO, Me–CH, H	13.91	MeO, Me–CH, Me	14.31
N–CH, H, H	10.32	N, Me–CH, H	9.63	N, Me–CH, Me	11.64
Ph–N–CH, H, H	11.50	Ph–N, Me–CH, H	12.02	Ph–N, Me–CH, Me	14.91
alkyne–CH, H, H	16.51	alkyne, Me–CH, H	14.75	alkyne, Me–CH, Me	14.82
O=S–CH, H, H	18.14	O=S, Me–CH, H	15.80	O=S, Me–CH, Me	16.07
O$_2$S–CH, H, H	25.53	O$_2$S, Me–CH, H	19.37	O$_2$S, Me–CH, Me	20.44
MeS–CH, H, H	14.05	MeS, Me–CH, H	13.21	MeS, Me–CH, Me	13.42
amide N–CH, H, H	15.23	amide N, Me–CH, H	16.57	amide N, Me–CH, Me	16.49
HO–CH, H, H	15.84	HO, Me–CH, H	14.43	HO, Me–CH, Me	13.77
Ph vinyl–CH, H, H	13.37	OMe vinyl–CH, H, H	11.81	diene–CH, H	12.56
N vinyl–CH, H, H	10.41	OMe, Ph–CH, H	11.43	Ph–N, CH, H	9.05
MeO, OMe–CH, H	11.77				

[a] See ref 11.

only been reported for M3. As a result, CYP1A2 may be producing more M4 than M3. Thus the role of CYP1A2 remains unclear. Similarly, CYP-mediated metabolism of

selective estrogen receptor modulators has been investigated revealing a large number of metabolites.[66] In the case where more than three metabolites are reported, we considered only
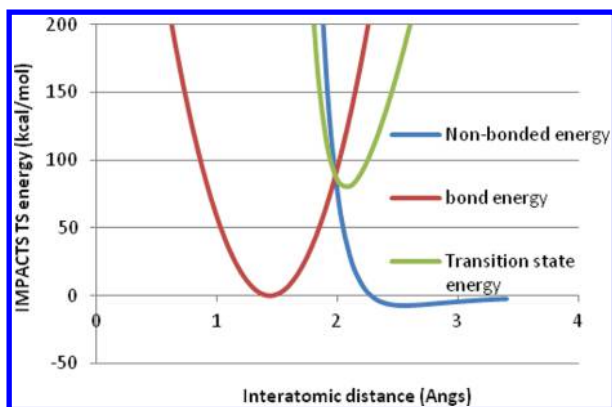
Figure 6. TS energy as a linear combination of bonded and non-bonded energies.
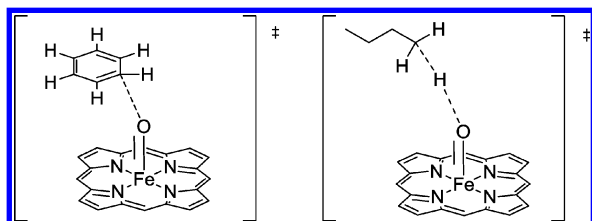
Figure 7. TS for aromatic oxidation and hydrogen abstraction. Side chains of the heme on the porphyrin ring are omitted for clarity.

the major three (e.g., metabolism of Ramelteon by CYP1A2[67]). Another issue is the use of animal models (rats and mice) as it has been found that these models were not always accurate to predict metabolism in humans (e.g., Midazolam in mice vs human[68]). Two additional points that readers should be aware of when developing such a set are given below. First, as the molecular weight of the different CYPs varies, the unit given for activity (whether pmol/min·**pmol** CYP or pmol/min·mg·**mg** CYP) is important as the perception of the role of each of the CYP may differ (see for example, Ketamine metabolism by CYP2B6, CYP3A4, and CYP2C9 using both units[69]). Second, the activity of the CYP is very sensitive to the isoform. *N,N*-Diethyl-*m*-toluamide is metabolized by CYP2D6*1 (Val374) while CYP2D6 (Met374) does not produce any detectable activity.[70] Data from in vivo studies should also be taken with care. For example, the metabolism in vitro of Gefitinib was found to produce a number of metabolites including the desmethyl derivative. The latter is a major metabolite found in human plasma but a minor metabolite in vitro.[71,72]

Considering all the collected information and the various factors described above, we curated the retrieved sets. As additional criteria, molecules that are too small and/or feature a single or only two potential reactive sites (e.g., butadiene monoxide) were excluded in order to avoid oversimplifying the testing set. Substrates such as Sulfinpyrazone sulphide, Capsaicin, Domperidone, and 2-*n*-propyl quinoline were also removed as many metabolites are not clearly identified (several possible regioisomers on aromatic rings). Finally, duplicates were identified and removed. For example, the metabolism of Selegiline,[48] Deprenyl,[73] and *N*-methyl,*N*-propargylphenyle-thylamine[46] was investigated independently although these three names refer to the same molecule. This led to sets of 137 CYP1A2 substrates, 128 CYP2C9 substrates, 157 CYP2D6 substrates, and 293 CYP3A4 substrates. These sets are supplied as Supporting Information.

RESULTS AND DISCUSSION

IMPACTS. In practice, IMPACTS identifies multiple potential reactive sites in a single ligand and creates interaction lists for each of these possible sites (Figure 8). The docking
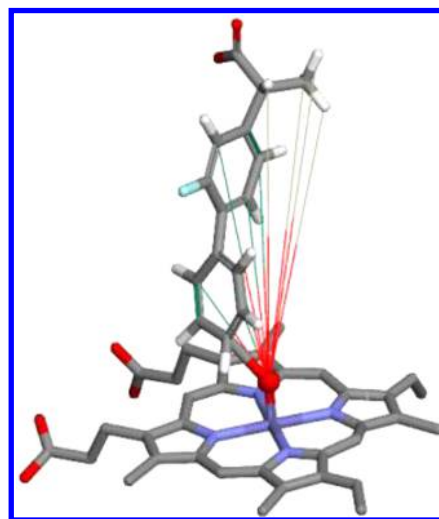
Figure 8. TS computed for oxidation of Flurbiprofen with CYP2C9. The various lines represent possible forming bonds (the protein is omitted for clarity).

procedure then docks the substrate and, for each pose, selects the reactive site closest to the ferryl oxygen as the reactive site. The interaction list corresponding to a bond formation at this particular reactive site is then used to compute forces and potential energy values of this particular TS. As with any docking program, a score is assigned to the proposed binding mode (referred to as a pose) and can be used to rank different poses. Within IMPACTS, the scoring function is composed of the noncovalent scoring function RankScore implemented in our docking program FITTED and the activation energy (eq 3).

Among the implemented transformations[43] are hydrogen abstraction[38] leading to either hydroxylation of alkyl chains (which also represents the first step in *N*-dealkylation and *O*-dealkylation)[74] or oxidation of aldehydes into carboxylic acids,[75] oxidative deboronation as observed with Bortezomib,[76] aromatic oxidation,[38] thioketone and thiophosphate oxidation,[77] double bond epoxidation, aromatic nitrogen (e.g., pyridine) oxidation, thioether oxidation into sulfoxide, sulfoxide oxidation into sulfone, and aniline oxidation. Other more unusual reactions such as P450-mediated conversion of nitriles to amides[78] and oxidative defluorination[72] have not yet been implemented.

Measuring Accuracy. One important factor to consider in these comparisons is that experimental data will report the most predominant metabolites from incubation samples. However, metabolites that can undergo sequential reactions or are too unstable to be isolated and may be missed. As we are looking at TSs critical for the regioselective oxidation, further rearrangements of an unstable reaction intermediate are not predicted by the program in its current version. For instance, the sulfur atom of the thiophene ring of drugs[79] can first be oxidized (Figure 9).[80] This is rapidly followed by a rearrangement and the formally observed metabolite is oxidized on the carbon adjacent to the sulfur. An alternative thiophene oxidation of Tienilic acid and Suprofen has been proposed
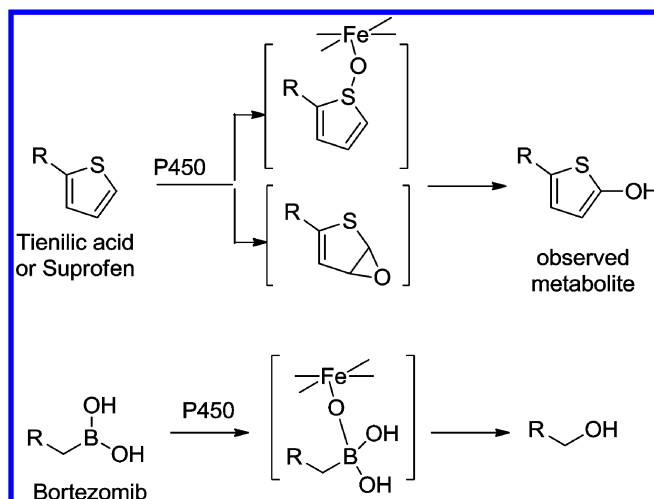
Figure 9. Multistep oxidation of drugs.

which goes through an epoxide formation.[81] Oxidative deboronation has also a well-defined mechanism (Figure 9). The empty orbital of the boronic acid (e.g., Bortezomib) plays a key role in the oxidation process. For our predictions to be deemed correct, oxidation of the sulfur atom or epoxidation of Tienilic acid and oxidation of the boron atom of Bortezomib should be proposed as the first transition state and not the oxidation of the adjacent carbons as observed.

Applications to CYP1A2, CYP2D6, CYP2C9, and CYP3A4 Substrates. In order to assess the accuracy of IMPACTS, testing sets of substrates of CYP1A2, 2C9, 2D6, and 3A4 were assembled. The sets of Sheridan et al., Danielson et al., Vasanthanathan et al.,[19] and Rydberg et al.[7] were either downloaded or rebuilt and further curated as described in the Theory and Implementation section. Although CYP2C19 is another major isoform, the lack of crystal structures precluded its use in this work. The success rate of IMPACTS is shown below using the metrics previously described[30] (Table 3).

Table 3. Accuracy[a] of IMPACTS in Predicting the Correct SoMs for Respective Data Sets

CYP	N^b	rand[c]	$E_a{}^d$	IMPACTS[e]	flexible[f]	experts[g]	best expert[h]
1A2	137	31	59	77		69 (5)	74
2C9	129	29	59	79–82	50	71 (7)	74
2D6	157	27	49	76		64 (4)	65
3A4	293	28	66	72–75	48	61 (6)	71
All 4	716	28	60	77	49	65 (5)	71

[a]Percent of molecules with an observed SoM in the predicted two SoMs. [b]Number of substrates in the set. [c]Random selection from the potential SoMs identified by IMPACTS. [d]Only the predicted reactivity of the potential SoMs is considered. [e]IMPACTS with a single crystal structure; a range is given if multiple structures were alternatively assessed. [f]IMPACTS when considering protein flexibility. [g]Average predictions by experts (standard deviation in brackets). [h]Best predictions from experts.

Predictions were made using IMPACTS on the sets of CYP1A2, CYP2C9, CYP2D6, and CYP3A4 substrates, and these predictions were defined as correct when one of the top two predicted SoMs had been experimentally observed (top-2 metrics, Table 3).[30] A single crystal structure of ligand-bound CYP1A2 (pdb code: 2hi4) and CYP2D6 (3qm4) are available

to date, while two and seven crystal structures of ligand-bound CYP2C9 (1r9o, 2og5) and CYP3A4 (1tqn, 1w0e, 1w0f, 1w0g, 2jog, 2vom, 3nxu), respectively, have been reported. In order to demonstrate the significance of the predictions, the potential SoMs were identified by IMPACTS and two were randomly selected (Table 3). We were pleased to see that the implementation of ligand reactivity significantly increased the overall accuracy from 28% (obtained by random selection) to 60%, and docking to rigid CYPs further increased this accuracy by another 17%. As expected, the accuracy with the most promiscuous CYP3A4 was the highest with the ligand potential SoM reactivity and increased only slightly when docking was considered. This contrasted with the data obtained with the other three "more specific" enzymes which revealed that overall using only the ligand-based rules implemented in IMPACTS was not sufficient. Pharmacophores used to identify potential SoMs for substrates of these three enzymes have been reported indicating that the substrate binding orientation has a significant role in the selection of the SoMs. With these four CYPs, the accuracy is well over 70%.

More unexpectedly, the accuracy of the predictions was lower when protein flexibility was considered. Other reports have shown that this factor improves the accuracy although only slightly.[20,30] A close look at the predictions reveals that the failures are more ligand-dependent (these ligands provide low accuracy regardless of the protein structure used) rather than protein structure-dependent (the failures are similar regardless of the CYP structure used). Thus, we believe that considering protein flexibility added noise to the calculations rather than improving the protein/substrate modeling. This data also demonstrated that either the promiscuity of the investigated CYPs is a great challenge or that the method as implemented has reached the limit of accuracy of the current potential energy function. We noted that most of the observed SoMs were found in the top four suggested (>90%). The other poorly predicted 10% included large and/or highly flexible molecules which are known to be problematic with docking programs or may require significant conformational changes in the protein.

Experts' Predictions. Despite the many reports on methods for SoM prediction,[6] none has specifically questioned the usefulness (accuracy and user-friendliness) of the current methods in the context of drug design, medicinal chemistry, and metabolism studies, with, to the best of our knowledge, a single study reporting the prediction of a single biotransformation expert on two medium-sized ($N = 39, 82$) sets.[82] Are these methods accurate enough to be useful? To address this critical issue, we challenged four medicinal chemists and two biotransformation experts each with over 10 years of experience. Although these six experimentalists may not be representative of the medicinal chemistry and biotransformation communities, the collected data (Table 3) is indicative of what could be considered an accurate and useful method. A web site has been set up to enable these experts to record their predictions on the four sets of substrates (in 2D), i.e., the same input given to IMPACTS. Their predictions were consistently lower than those by IMPACTS by at least 7%. The accuracy of the predictions of one of the experts were overall closer to that of IMPACTS, although still overall lower by more than 5%.

Interestingly, random selection provides an accurate prediction for as many as one-fourth of the substrates with the criteria used in Table 3. This accuracy rose to nearly 40% when the top-3 metrics was used.

Overall, this data revealed that our program will be useful for predicting SoMs of small molecules or even libraries of small molecules and for producing three-dimensional structures of the TSs of the small molecule/CYP complexes (Figure 10).
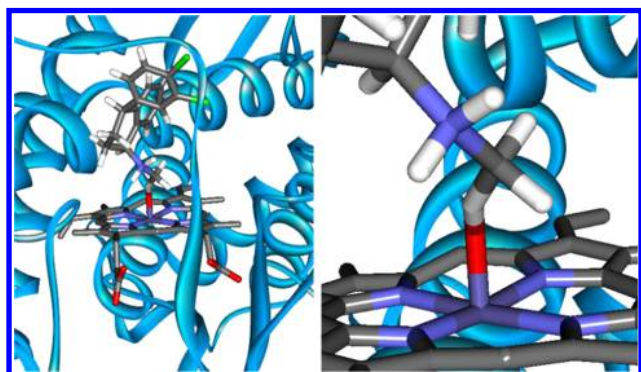

Figure 10. N-Demethylation of Sertraline by CYP2C9.

IMPACTS's Performance. In order to assess the accuracy of this first version of IMPACTS, we have retrieved sets of CYP substrates developed by Zaretzki et al.[83] and by Afzelius et al.[82] These sets were prepared and submitted to IMPACTS, and the predictions were compared to those reported using eleven other methods. These methods include academic (e.g., SmartCyp) and commercial programs (e.g., StarDrop), ligand-based, structure-based, and hybrid (i.e., both ligand and protein structure-based, MetaSite) methods. RS-Predictor makes use of ligand descriptors. A more thorough description of these methods and their use can be found in the original publications of Zaretzki et al. and Afzelius et al.[82,83] As was done by Afzelius et al., large and highly flexible substrates were not considered by our docking-based method IMPACTS. This preselection did not greatly affect the size of the CYP1A2, CYP2D6, and CYP2C9 substrate sets (one to four molecules per set, less than 2% of these substrates), but it did reduce the number of CYP3A4 substrates by 8% as shown in Table 4. As the sets are slightly different from one program to the next and as updated and/or improved versions of these methods may have been released since their use by Afzelius et al. and Zaretzki et al., this comparison should be considered with care and is

only used to illustrate the overall performance of IMPACTS and not to provide a ranking of methods. As can be seen in Table 4, this first version of IMPACTS stands well in this comparison, with accuracy equivalent to RS-Predictor and superior to all the other methods with CYP2C9 substrates. The accuracy with CYP1A2 substrates is also comparable to that obtained with RS-Predictor. However, IMPACTS was slightly less accurate than other methods for CYP2D6 and CYP3A4 substrates when looking at average accuracy.

In contrast to ligand-based methods, docking-based methods provide information on the binding mode of the substrates in the active site of CYPs. However, their major drawback[84] is the CPU time required to make predictions. In general, ligand-based predictions can be made within a second or less per compound. Analysis of the data generated to produce Table 4 revealed that 77% of the runs took 2 min or less per run and 97% took 10 min or less. 0.8% of the compounds, mostly 3A4 substrates, were not docked after an hour.

■ CONCLUSION

In conclusion, we have developed a preliminary version of a fully automated program, IMPACTS, for the prediction of the SoMs of drugs. Moreover, IMPACTS provides a 3D picture of the TS of the drug at the active site of cytochromes (Figure 10). In addition, this method does not require any training and should be applicable to other CYPs.[6] Knowledge about both the binding mode and the SoM may enable medicinal chemists to design modifications at locations other than the SoMs to disturb or enhance the substrate-CYP recognition. In contrast to other methods, ligand reactivity, binding affinity, and proper geometry of the TS are all considered in a single and fully automated run. Other limitations in modeling CYPs include the presence of water molecules, which may also be important in drug/CYP binding.[5,29] Strategies to overcome these challenges and improvements of the current activation energies are currently being investigated in the hope of further increasing the accuracy of IMPACTS. Finally, a comparison between the predictions generated by IMPACTS, those by random selection, those made by experts, and those made using eleven other methods demonstrated that IMPACTS can be extremely useful.

Table 4. Accuracy of IMPACTS and Eleven Other Methods in Predicting the Correct SoMs for External Data Sets

isozyme	1A2	2C9	2C9	2C9	2D6	2D6	3A4	3A4	3A4
N^a	271[c]	98[c]	128[c]	49[d]	134[c]	136[c]	321[c]	154[c]	65[d]
N^b	269[c]	96[c]	127[c]	45[d]	133[c]	132[c]	295[c]	144[c]	58[d]
IMPACTS	80.5	84.4	76.4	81.8	70.7	71.2	73.2	70.1	82.5
RS-Predictor (TOP QC SCR)[83]	83.0	81.6	79.7		86.6	78.7	85.7	72.7	
SMARTCyp[83]		67.7	66.9		48.5	68.1	73.1	77.2	
StarDrop[83]		77.4	78.4		81.5	69.2	77.5	66.9	
Schrödinger[83]		69.6	74.0		66.2	70.1	80.2	68.2	
Sheridan et al.[44,83]		72.4			71.9		77.4		
MetaSite[44,83]		68.8		91	65.4		61.8		87
MetaDock[82]				66					67
QMBO[82]				84					84
QMSpin[82]				78					78
MetaGlide[82]				67					65
SporCalc[82]				81					81

[a]Number of substrates used by Zaretzki et al. and Afzelius et al. [b]Number for substrates when large and flexible substrates are removed following the work of Afzelius et al. [c]Sets from Zaretzki et al., top-2 metrics was used.[83,82] [d]Sets from Afzelius et al., top-3 metrics was used.[82,83]

■ EXPERIMENTAL SECTION

Construction of the Testing Sets. The testing sets were built with care to reduce the noise in the prediction assessment. Here is a list of selection criteria that were used; a detailed discussion is provided in the Theory and Implementation section. (1) Small molecules such as butadiene monoxide with only one possible reactive site (i.e., a single double bond) have been removed from existing sets as they increase the apparent accuracy but reduce the significance of the predictions. (2) Eicosapentaenoic acid, Sulfinpyrazone sulphide, Capsaicin, Domperidone, and 2-*n*-propyl quinoline were also removed as many metabolites are not clearly identified (several possible regioisomers on aromatic rings). (3) Duplicates were identified and removed. Then, an asterisk was then added to the atoms being oxidized in order to eventually identify whether the predicted SoM is one of the observed SoMs. Sets are given in separate files in mol2 format.

Computation of the Activation Energies. All the quantum mechanical calculations were performed using DFT, more specifically the B3LYP functional (Unrestricted Hartree−Fock) and the 6-31G* basis-set. All calculations were performed in vacuum. The B3LYP calculations were performed using GAMESS-US v.Aug2011-64bit.

As the reacting methoxy group is not aligned with the benzene ring, the transition state is not perfectly symmetrical. Thus, for the monosubstituted benzene derivatives, the two meta and ortho positions are not equivalent. The lowest-in-energy of the two meta positions and highest-in-energy of the two optimized ortho positions were kept for each one. The highest-in-energy of the two minima was regarded as correct for ortho position as the difference in energy was often due to hydrogen-bonding effects which are artifacts of the methoxy models. In fact, hydrogen bonds would be limited within the heme system (steric block). In parallel, sterics are the driving force at the meta position and thus the lowest-in-energy was regarded as correct

Application of IMPACTS. Default parameters implemented in IMPACTS have been used. IMPACTS has been integrated into our platform FORECASTER[85] for user-friendliness. The user can draw the substrate into a 2D sketcher and select the CYP enzyme with which to predict the SoM. FORECASTER will take care of adding hydrogens, generating a 3D structure, and selecting the correct CYP files. IMPACTS and FORECASTER are accessible free of charge to academic users (www.fitted.ca).

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The sets of substrates. The program is available, upon request, from the authors (www.fitted.ca). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: nicolas.moitessier@mcgill.ca.

Present Address
‡National Research Council of Canada, 6100 Royalmount Avenue, Montreal, QC, H4P 2R2, Canada.

Author Contributions
‖The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. These authors contributed equally.

Notes
The authors declare no competing financial interest.

■ REFERENCES

(1) Wong, Y. C.; Qian, S.; Zuo, Z. Regioselective biotransformation of CNS drugs and its clinical impact on adverse drug reactions. *Exp. Opin. Drug Metab. Toxicol.* **2012**, *8*, 833−854.

(2) Shamovsky, I.; Ripa, L.; Börjesson, L.; Mee, C.; Nordén, B.; Hansen, P.; Hasselgren, C.; O'Donovan, M.; Sjö, P. Explanation for Main Features of Structure−Genotoxicity Relationships of Aromatic Amines by Theoretical Studies of Their Activation Pathways in CYP1A2. *J. Am. Chem. Soc.* **2011**, *133*, 16168−16185.

(3) He, S. M.; Zhou, Z. W.; Li, X. T.; Zhou, S. F. Clinical drugs undergoing polymorphic metabolism by human cytochrome P450 2C9 and the implication in drug development. *Curr. Med. Chem.* **2011**, *18*, 667−713.

(4) Zhang, T.; Chen, Q.; Li, L.; Liu, L. A.; Wei, D. Q. In silico prediction of cytochrome P450-mediated drug metabolism. *Comb. Chem. High Throughput Screening* **2011**, *14*, 388−395.

(5) Tarcsay, A.; Keserü, G. M. In silico site of metabolism prediction of cytochrome P450-mediated biotransformations. *Exp. Opin. Drug Metab. Toxicol.* **2011**, *7*, 299−312.

(6) Kirchmair, J.; Williamson, M. J.; Tyzack, J. D.; Tan, L.; Bond, P. J.; Bender, A.; Glen, R. C. Computational Prediction of Metabolism: Sites, Products, SAR, P450 Enzyme Dynamics, and Mechanisms. *J. Chem. Inf. Model.* **2012**, *52*, 617−648.

(7) Rydberg, P.; Vasanthanathan, P.; Oostenbrink, C.; Olsen, L. Fast prediction of cytochrome p450 mediated drug metabolism. *ChemMedChem* **2009**, *4*, 2070−2079.

(8) Rydberg, P.; Gloriam, D. E.; Olsen, L. The SMARTCyp cytochrome P450 metabolism prediction server. *Bioinformatics* **2010**, *26*, 2988−2989.

(9) Rydberg, P.; Hansen, S. M.; Kongsted, J.; Norrby, P.-O.; Olsen, L.; Ryde, U. Transition-State Docking of Flunitrazepam and Progesterone in Cytochrome P450. *J. Chem. Theory Comput.* **2008**, *4*, 673−681.

(10) Saraceno, M.; Massarelli, I.; Imbriani, M.; James, T. L.; Bianucci, A. M. Optimizing QSAR Models for Predicting Ligand Binding to the Drug-Metabolizing Cytochrome P450 Isoenzyme CYP2D6. *Chem. Biol. Drug Des.* **2011**, *78*, 236−251.

(11) Rydberg, P.; Gloriam, D. E.; Zaretzki, J.; Breneman, C.; Olsen, L. SMARTCyp: A 2D method for prediction of cytochrome P450-mediated drug metabolism. *ACS Med. Chem. Lett.* **2010**, *1*, 96−100.

(12) Zaretzki, J.; Bergeron, C.; Rydberg, P.; Huang, T.-w; Bennett, K. P.; Breneman, C. M. RS-Predictor: A New Tool for Predicting Sites of Cytochrome P450-Mediated Metabolism Applied to CYP 3A4. *J. Chem. Inf. Model.* **2011**, *51*, 1667−1689.

(13) Pelkonen, O.; Turpeinen, M.; Raunio, H. In vivo-in vitro-in silico pharmacokinetic modelling in drug development: Current status and future directions. *Clin. Pharmacokin.* **2011**, *50*, 483−491.

(14) Czodrowski, P.; Kriegl, J. M.; Scheuerer, S.; Fox, T. Computational approaches to predict drug metabolism. *Exp. Opin. Drug Metab. Toxicol.* **2009**, *5*, 15−27.

(15) De Graaf, C.; Pospisil, P.; Pos, W.; Folkers, G.; Vermeulen, N. P. E. Binding mode prediction of cytochrome P450 and thymidine kinase

protein-ligand complexes by consideration of water and rescoring in automated docking. *J. Med. Chem.* **2005**, *48*, 2308−2318.

(16) Stjernschantz, E.; Vermeulen, N. P. E.; Oostenbrink, C. Computational prediction of drug binding and rationalisation of selectivity towards cytochromes P450. *Exp. Opin. Drug Metab. Toxicol.* **2008**, *4*, 513−527.

(17) Vaz, R. J.; Zamora, I.; Li, Y.; Reiling, S.; Shen, J.; Cruciani, G. The challenges of in silico contributions to drug metabolism in lead optimization. *Exp. Opin. Drug Metab. Toxicol.* **2010**, *6*, 851−861.

(18) Sun, H.; Scott, D. O. Structure-based drug metabolism predictions for drug design. *Chem. Biol. Drug Des.* **2010**, *75*, 3−17.

(19) Vasanthanathan, P.; Hritz, J.; Taboureau, O.; Olsen, L.; Jorgensen, F. S.; Vermeulen, N. P. E.; Oostenbrink, C. Virtual screening and prediction of site of metabolism for cytochrome P450 1A2 ligands. *J. Chem. Inf. Model.* **2009**, *49*, 43−52.

(20) Moors, S. L. C.; Vos, A. M.; Cummings, M. D.; Van Vlijmen, H.; Ceulemans, A. Structure-Based Site of Metabolism Prediction for Cytochrome P450 2D6. *J. Med. Chem.* **2011**, *54*, 6098−6105.

(21) Cruciani, G.; Carosati, E.; De Boeck, B.; Ethirajulu, K.; Mackie, C.; Howe, T.; Vianello, R. MetaSite: Understanding Metabolism in Human Cytochromes from the Perspective of the Chemist. *J. Med. Chem.* **2005**, *48*, 6970−6979.

(22) Oh, W. S.; Kim, D. N.; Jung, J.; Cho, K. H.; No, K. T. New combined model for the prediction of regioselectivity in cytochrome P450/3A4 mediated metabolism. *J. Chem. Inf. Model.* **2008**, *48*, 591−601.

(23) Williams, P. A.; Cosme, J.; Ward, A.; Angove, H. C.; Vinkovi, D. M.; Jhoti, H. Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature* **2003**, *424*, 464−468.

(24) Williams, P. A.; Cosme, J.; Matak Vinkovi, D.; Ward, A.; Angove, H. C.; Day, P. J.; Vonrhein, C.; Tickle, I. J.; Jhoti, H. Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone. *Science* **2004**, *305*, 683−686.

(25) Wester, M. R.; Yano, J. K.; Schoch, G. A.; Yang, C.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. The structure of human cytochrome P450 2C9 complexed with flurbiprofen at 2.0-Å resolution. *J. Biol. Chem.* **2004**, *279*, 35630−35637.

(26) Coleman, S.; Linderman, R.; Hodgson, E.; Rose, R. L. Comparative metabolism of chloroacetamide herbicides and selected metabolites in human and rat liver microsomes. *Environ. Health Persp.* **2000**, *108*, 1151−1157.

(27) Kjellander, B.; Masimirembwa, C. M.; Zamora, I. Exploration of enzyme-ligand interactions in CYP2D6 & 3A4 homology models and crystal structures using a novel computational approach. *J. Chem. Inf. Model.* **2007**, *47*, 1234−1247.

(28) Rydberg, P.; Vasanthanathan, P.; Oostenbrink, C.; Olsen, L. Fast Prediction of Cytochrome P450 Mediated Drug Metabolism. *ChemMedChem* **2009**, *4*, 2070−2079.

(29) Hritz, J.; de Ruiter, A.; Oostenbrink, C. Impact of Plasticity and Flexibility on Docking Results for Cytochrome P450 2D6: A Combined Approach of Molecular Dynamics and Ligand Docking. *J. Med. Chem.* **2008**, *51*, 7469−7477.

(30) Danielson, M. L.; Desai, P. V.; Mohutsky, M. A.; Wrighton, S. A.; Lill, M. A. Potentially increasing the metabolic stability of drug candidates via computational site of metabolism prediction by CYP2C9: The utility of incorporating protein flexibility via an ensemble of structures. *Eur. J. Med. Chem.* **2011**, *46*, 3953−3963.

(31) Ekroos, M.; Sjögren, T. Structural basis for ligand promiscuity in cytochrome P450 3A4. *Proc. Natl. Acad. Sci. USA.* **2006**, *103*, 13682−13687.

(32) Guengerich, F. P. A malleable catalyst dominates the metabolism of drugs. *Proc. Natl Acad. Sci. USA* **2006**, *103*, 13565−13566.

(33) Ito, Y.; Kondo, H.; Goldfarb, P. S.; Lewis, D. F. V. Analysis of CYP2D6 substrate interactions by computational methods. *J. Mol. Graphics Modell.* **2008**, *26*, 947−956.

(34) Zhu, Y.; Silverman, R. B. Revisiting Heme Mechanisms. A Perspective on the Mechanisms of Nitric Oxide Synthase (NOS),

Heme Oxygenase (HO), and Cytochrome P450s (CYP450s). *Biochemistry* **2008**, *47*, 2231−2243.

(35) Corbeil, C. R.; Englebienne, P.; Moitessier, N. Docking ligands into flexible and solvated macromolecules. 1. Development and validation of FITTED 1.0. *J. Chem. Inf. Model.* **2007**, *47*, 435−449.

(36) Corbeil, C. R.; Moitessier, N. Docking Ligands into Flexible and Solvated Macromolecules. 3. Impact of Input Ligand Conformation, Protein Flexibility, and Water Molecules on the Accuracy of Docking Programs. *J. Chem. Inf. Model.* **2009**, *49*, 997−1009.

(37) Bathelt, C. M.; Ridder, L.; Mulholland, A. J.; Harvey, J. N. Mechanism and structure-reactivity relationships for aromatic hydroxylation by cytochrome P450. *Org. Biomol. Chem.* **2004**, *2*, 2998−3005.

(38) Rydberg, P.; Ryde, U.; Olsen, L. Prediction of Activation Energies for Aromatic Oxidation by Cytochrome P450. *J. Phys. Chem. A* **2008**, *112*, 13058−13065.

(39) Olsen, L.; Rydberg, P.; Rod, T. H.; Ryde, U. Prediction of Activation Energies for Hydrogen Abstraction by Cytochrome P450. *J. Med. Chem.* **2006**, *49*, 6489−6499.

(40) Corbeil, C. R.; Moitessier, N. Docking ligands into flexible and solvated macromolecules. 3. Impact of input ligand conformation, protein flexibility, and water molecules on the accuracy of docking programs. *J. Chem. Inf. Model.* **2009**, *49*, 997−1009.

(41) Corbeil, C. R.; Thielges, S.; Schwartzentruber, J. A.; Moitessier, N. Toward a computational tool predicting the stereochemical outcome of asymmetric reactions: Development and application of a rapid and accurate program based on organic principles. *Angew. Chem., Int. Ed.* **2008**, *47*, 2635−2638.

(42) Weill, N.; Corbeil, C. R.; De Schutter, J. W.; N., M. Toward a computational tool predicting the stereochemical outcome of asymmetric reactions: Development of the molecular mechanics-based program ACE and application to asymmetric epoxidation reactions. *J. Comput. Chem.* **2011**, *32*, 2878−2889.

(43) Guengerich, F. P.; Isin, E. M. Mechanisms of cytochrome P450 reactions. *Acta Chim. Slov.* **2008**, *55*, 7−19.

(44) Sheridan, R. P.; Korzekwa, K. R.; Torres, R. A.; Walker, M. J. Empirical Regioselectivity Models for Human Cytochromes P450 3A4, 2D6, and 2C9. *J. Med. Chem.* **2007**, *50*, 3173−3184.

(45) Hyland, R.; Jones, B. C.; Smith, D. A. Identification of the Cytochrome P450 Enzymes Involved in the N-Oxidation of Voriconazole. *Drug Metab. Dispos.* **2003**, *31*, 540−547.

(46) Rittenbach, K. A.; Holt, A.; Ling, L.; Shan, J.; Baker, G. B. Metabolism of N-methyl, N-propargylphenylethylamine: Studies with human liver microsomes and cDNA expressed cytochrome P450 (CYP) enzymes. *Cell. Mol. Neurobiol.* **2007**, *27*, 179−190.

(47) Murayama, N.; Imai, N.; Nakane, T.; Shimizu, M.; Yamazaki, H. Roles of CYP3A4 and CYP2C19 in methyl hydroxylated and N-oxidized metabolite formation from voriconazole, a new anti-fungal agent, in human liver microsomes. *Biochem. Pharmacol.* **2007**, *73*, 2020−2026.

(48) Hidestrand, M.; Oscarson, M.; Salonen, J. S.; Nyman, L.; Pelkonen, O.; Turpeinen, M.; Ingelman-Sundberg, M. CYP2B6 and CYP2C19 as the Major Enzymes Responsible for the Metabolism of Selegiline, a Drug Used in the Treatment of Parkinson's Disease, as Revealed from Experiments with Recombinant Enzymes. *Drug Metab. Dispos.* **2001**, *29*, 1480−1484.

(49) Yun, C. H.; Lee, H. S.; Lee, H.; Rho, J. K.; Jeong, H. G.; Guengerich, F. P. Oxidation of the angiotensin II receptor antagonist losartan (DuP 753) in human liver microsomes. Role of cytochrome P4503A(4) in formation of the active metabolite EXP3174. *Drug Metab. Dispos.* **1995**, *23*, 285−289.

(50) Kobayashi, K.; Ishizuka, T.; Shimada, N.; Yoshimura, Y.; Kamijima, K.; Chiba, K. Sertraline N-Demethylation Is Catalyzed by Multiple Isoforms of Human Cytochrome P-450 In Vitro. *Drug Metab. Dispos.* **1999**, *27*, 763−766.

(51) Obach, R. S.; Cox, L. M.; Tremaine, L. M. Sertraline is metabolized by Multiple Cytochrome P450 Enzymes, Monoamie oxidases, and Glucuronyl transferases in Human: an in Vitro Study. *Drug Metab. Dispos.* **2005**, *33*, 262−270.

(52) Baldwin, S. J.; Clarke, S. E.; Chenery, R. J. Characterization of the cytochrome P450 enzymes involved in the in vitro metabolism of rosiglitazone. *Br. J. Clin. Pharmacol.* **1999**, *48*, 424−432.

(53) Eap, C. B.; Bender, S.; Gastpar, M.; Fischer, W.; Haarmann, C.; Powell, K.; Jonzier-Perey, M.; Cochard, N.; Baumann, P. Steady state plasma levels of the enantiomers of trimipramine and of its metabolites in CYP2D6-, CYP2C19- and CYP3A4/5-phenotyped patients. *Ther. Drug Monit.* **2000**, *22*, 209−214.

(54) Narimatsu, S.; Takemi, C.; Tsuzuki, D.; Kataoka, H.; Yamamoto, S.; Shimada, N.; Suzuki, S.; Satoh, T.; Meyer, U. A.; Gonzalez, F. J. Stereoselective Metabolism of Bufuralol Racemate and Enantiomers in Human Liver Microsomes. *J. Pharmacol. Exp. Ther.* **2002**, *303*, 172−178.

(55) Nakajima, M.; Nakamura, S.; Tokudome, S.; Shimada, N.; Yamazaki, H.; Yokoi, T. Azelastine N-Demethylation by Cytochrome P-450 (CYP)3A4, CYP2D6, and CYP1A2 in Human Liver Microsomes: Evaluation of Approach to Predict the Contribution of Multiple CYPs. *Drug Metab. Dispos.* **1999**, *27*, 1381−1391.

(56) Imai, T.; Taketani, M.; Suzu, T.; Kusube, K.; Otagiri, M. In Vitro Identification of the Human Cytochrome P-450 Enzymes Involved in the N-Demethylation of Azelastine. *Drug Metab. Dispos.* **1999**, *27*, 942−946.

(57) Goda, R.; Nagai, D.; Akiyama, Y.; Nishikawa, K.; Ikemoto, I.; Aizawa, Y.; Nagata, K.; Yamazoe, Y. Detection of a New N-oxidized Metabolite of Flutamide N-[4-nitro-3-(trifluoromethyl)phenyl]-hydroxylamine, in Human Liver Microsomes and Urine of Prostate Cancer Patients. *Drug Metab. Dispos.* **2006**, *34*, 828−835.

(58) Simarro Doorten, A. Y.; Bull, S.; Van Der Doelen, M. A. M.; Fink-Gremmels, J. Metabolism-mediated cytotoxicity of ochratoxin A. *Toxicol. Vitro.* **2004**, *18*, 271−277.

(59) Bournique, B.; Lambert, N.; Boukaiba, R.; Martinet, M. In vitro metabolism and drug interaction potential of a new highly potent anti-cytomegalovirus molecule, CMV423 (2-chloro 3-pyridine 3-yl 5,6,7,8-tetrahydroindolizine 1-carboxamide). *Br. J. Clin. Pharmacol.* **2001**, *52*, 53−63.

(60) Pearce, R. E.; Vakkalagadda, G. R.; Steven Leeder, J. Pathways of carbamazepine bioactivation in vitro I. Characterization of human cytochromes P450 responsible for the formation of 2- and 3-hydroxylated metabolites. *Drug Metab. Dispos.* **2002**, *30*, 1170−1179.

(61) Choi, S. J.; Ji, H. Y.; Lee, H. Y.; Kim, D. S.; Kim, W. B.; Lee, H. S. In vitro metabolism of a novel phosphodiesterase-5 inhibitor DA-8159 in rat liver preparations using liquid chromatography/electrospray mass spectrometry. *Biomed. Chrom.* **2002**, *16*, 395−399.

(62) Ji, H. Y.; Lee, H. W.; Kim, H. H.; Kim, D. S.; Yoo, M.; Kim, W. B.; Lee, H. S. Role of human cytochrome P450 3A4 in the metabolism of DA−8159, a new erectogenic. *Xenobiotica* **2004**, *34*, 973−982.

(63) Yoshii, K.; Kobayashi, K.; Tsumuji, M.; Tani, M.; Shimada, N.; Chiba, K. Identification of human cytochrome P450 isoforms involved in the 7- hydroxylation of chlorpromazine by human liver microsomes. *Life Sci.* **2000**, *67*, 175−184.

(64) Wójcikowski, J.; Boksa, J.; Daniel, W. A. Main contribution of the cytochrome P450 isoenzyme 1A2 (CYP1A2) to N-demethylation and 5-sulfoxidation of the phenothiazine neuroleptic chlorpromazine in human liver-A comparison with other phenothiazines. *Biochem. Pharmacol.* **2010**, *80*, 1252−1259.

(65) Guengerich, F. P.; Brian, W. R.; Iwasaki, M.; Sari, M. A.; Baeaernhielm, C.; Berntsson, P. Oxidation of dihydropyridine calcium channel blockers and analogs by human liver cytochrome P-450 IIIA4. *J. Med. Chem.* **1991**, *34*, 1838−1844.

(66) Zhang, Z.; Chen, Q.; Li, Y.; Doss, G. A.; Dean, B. J.; Ngui, J. S.; Silva Elipe, M.; Kim, S.; Wu, J. Y.; DiNinno, F.; Hammond, M. L.; Stearns, R. A.; Evans, D. C.; Baillie, T. A.; Tang, W. In Vitro Bioactivation of Dihydrobenzoxathiin Selective Estrogen Receptor Modulators by Cytochrome P450 3A4 in Human Liver Microsomes: Formation of Reactive Iminium and Quinone Type Metabolites. *Chem. Res. Toxicol.* **2005**, *18*, 675−685.

(67) Obach, R. S.; Ryder, T. F. Metabolism of Ramelteon in human liver microsomes and correlation with the effect of fluvoxamine on ramelteon pharmacokinetics. *Drug Metab. Dispos.* **2010**, *38*, 1381−1391.

(68) Perloff, M. D.; von Moltke, L. L.; Court, M. H.; Kotegawa, T.; Shader, R. I.; Greenblatt, D. J. Midazolam and Triazolam Biotransformation in Mouse and Human Liver Microsomes: Relative Contribution of CYP3A and CYP2C Isoforms. *J. Pharmacol. Exp. Ther.* **2000**, *292*, 618−628.

(69) Hijazi, Y.; Boulieu, R. Contribution of CYP3A4, CYP2B6, and CYP2C9 Isoforms toN-Demethylation of Ketamine in Human Liver Microsomes. *Drug Metab. Dispos.* **2002**, *30*, 853−858.

(70) Usmani, K. A.; Rose, R. L.; Goldstein, J. A.; Taylor, W. G.; Brimfield, A. A.; Hodgson, E. In Vitro Human Metabolism and Interactions of Repellent N,N-Diethyl-m-Toluamide. *Drug Metab. Dispos.* **2002**, *30*, 289−294.

(71) McKillop, D.; McCormick, A. D.; Millar, A.; Miles, G. S.; Phillips, P. J.; Hutchison, M. Cytochrome P450-dependent metabolism of gefitinib. *Xenobiotica* **2005**, *35*, 39−50.

(72) Mckillop, D.; Mccormick, A. D.; Miles, G. S.; Phillips, P. J.; Pickup, K. J.; Bushby, N.; Hutchison, M. In vitro metabolism of gefitinib in human liver microsomes. *Xenobiotica* **2004**, *34*, 983−1000.

(73) Grace, J. M.; Kinter, M. T.; Macdonald, T. L. Atypical Metabolism of Deprenyl and Its Enantiomer, (S)-(+)-N,.alpha.-Dimethyl-N-Propynylphenethylamine, by Cytochrome P450 2D6. *Chem. Res. Toxicol.* **1994**, *7*, 286−290.

(74) Rydberg, P.; Ryde, U.; Olsen, L. Sulfoxide, Sulfur, and Nitrogen Oxidation and Dealkylation by Cytochrome P450. *J. Chem. Theory Comput.* **2008**, *4*, 1369−1377.

(75) Liu, X.; Wang, Y.; Han, K. Systematic study on the mechanism of aldehyde oxidation to carboxylic acid by cytochrome P450. *J. Biol. Inorg. Chem.* **2007**, *12*, 1073−1081.

(76) Larkin, J. D.; Markham, G. D.; Milkevitch, M.; Brooks, B. R.; Bock, C. W. Computational investigation of the oxidative deboronation of boroglycine, H2N-CH2-B(OH)2, using H2O and H 2O2. *J. Phys. Chem. A* **2009**, *113*, 11028−11034.

(77) Jayathirtha Rao, V.; Muthuramu, K.; Ramamurthy, V. Oxidations of thioketones by singlet and triplet oxygen. *J. Org. Chem.* **1982**, *47*, 127−131.

(78) Zhang, Z.; Li, Y.; Stearns, R. A.; Ortiz de Montellano, P. R.; Baillie, T. A.; Tang, W. Cytochrome P450 3A4-Mediated Oxidative Conversion of a Cyano to an Amide Group in the Metabolism of Pinacidil. *Biochemistry* **2002**, *41*, 2712−2718.

(79) Taguchi, K.; Konishi, T.; Nishikawa, H.; Kitamura, S. Identification of human cytochrome P450 isoforms involved in the metabolism of S-2-[4-(3-methyl-2-thienyl)phenyl]propionic acid. *Xenobiotica* **1999**, *29*, 899−907.

(80) Jean, P.; Lopez-Garcia, P.; Dansette, P.; Mansuy, D.; Goldstein, J. L. Oxidation of Tienilic Acid by Human Yeast-Expressed Cytochromes P-450 2C8, 2C9, 2C18 and 2C19. *Eur. J. Biochem.* **1996**, *241*, 797−804.

(81) O'Donnell, J. P.; Dalvie, D. K.; Kalgutkar, A. S.; Obach, R. S. Mechanism-based inactivation of human recombinant P450 2C9 by the nonsteroidal anti-inflammatory drug suprofen. *Drug Metab. Dispos.* **2003**, *31*, 1369−1377.

(82) Afzelius, L.; Arnby, C. H.; Broo, A.; Carlsson, L.; Isaksson, C.; Jurva, U.; Kjellander, B.; Kolmodin, K.; Nilsson, K.; Raubacher, F.; Weidolf, L. State-of-the-art tools for computational site of metabolism predictions: Comparative analysis, mechanistic insights, and future applications. *Drug Metabol. Rev.* **2007**, *39*, 61−86.

(83) Zaretzki, J.; Rydberg, P.; Bergeron, C.; Bennett, K. P.; Olsen, L.; Breneman, C. M. RS-predictor models augmented with SMARTCyp reactivities: Robust metabolic regioselectivity predictions for nine CYP isozymes. *J. Chem. Inf. Model.* **2012**, *52*, 1637−1659.

(84) Rydberg, P. Theoretical Study of the Cytochrome P450 Mediated Metabolism of Phosphorodithioate Pesticides. *J. Chem. Theory Comp.* **2012**.

(85) Therrien, E.; Englebienne, P.; Arrowsmith, A. G.; Mendoza-Sanchez, R.; Corbeil, C. R.; Weill, N.; Campagna-Slater, V.; Moitessier, N. Integrating Medicinal Chemistry, Organic/Combinatorial Chemistry, and Computational Chemistry for the Discovery of Selective

Estrogen Receptor Modulators with FORECASTER, a Novel Platform for Drug Discovery. *J. Chem. Inf. Model.* **2012**, *52*, 210−224.

2483

dx.doi.org/10.1021/ci3003073 | *J. Chem. Inf. Model.* 2012, 52, 2471−2483