Published in final edited form as:

Eur J Med Chem. 2011 September ; 46(9): 3953–3963. doi:10.1016/j.ejmech.2011.05.067.

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

Matthew L. Danielson^a, Prashant V. Desai^b, Michael A. Mohutsky^b, Steven A. Wrighton^{b,c}, and Markus A. Lill^{a,*}

^a Department of Medicinal Chemistry and Molecular Pharmacology Purdue University 575 Stadium Mall Drive, West Lafayette, IN 47907, USA

^b Department of Drug Disposition Lilly Research Laboratories Eli Lilly and Company Lilly Corporate Center Indianapolis, IN 46285, USA

Abstract

Cytochrome P450 enzymes are responsible for metabolizing many endogenous and xenobiotic molecules encountered by the human body. It has been estimated that 75% of all drugs are metabolized by cytochrome P450 enzymes. Thus, predicting a compound's potential sites of metabolism (SOM) is highly advantageous early in the drug development process. We have combined molecular dynamics, AutoDock Vina docking, the neighboring atom type (NAT) reactivity model, and a solvent-accessible surface-area term to form a reactivity-accessibility model capable of predicting SOM for cytochrome P450 2C9 substrates. To investigate the importance of protein flexibility during the ligand binding process, the results of SOM prediction using a static protein structure for docking were compared to SOM prediction using multiple protein structures in ensemble docking. The results reported here indicate that ensemble docking increases the number of ligands that can be docked in a bioactive conformation (ensemble: 96%, static: 85%) but only leads to a slight improvement (49% vs. 44%) in predicting an experimentally known SOM in the top-1 position for a ligand library of 75 CYP2C9 substrates. Using ensemble docking, the reactivity-accessibility model accurately predicts SOM in the top-1 ranked position for 49% of the ligand library and considering the top-3 predicted sites increases the prediction success rate to approximately 70% of the ligand library. Further classifying the substrate library according to K_m values leads to an improvement in SOM prediction for substrates with low K_m values (57% at top-1). While the current predictive power of the reactivity-accessibility model still leaves significant room for improvement, the results illustrate the usefulness of this method to identify key protein-ligand interactions and guide structural modifications of the ligand to increase its metabolic stability.

© 2011 Elsevier Masson SAS. All rights reserved.

*Corresponding author: Phone: 1-765-496-9375, Fax: 1-765-494-1414, mlill@purdue.edu.

^cCurrent Address: W925 Miramar Drive East Troy, WI 53120, USA

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

CYP2C9; metabolism; docking; protein flexibility; computational chemistry

1. Introduction

Cytochrome P450 (P450s) are a superfamily of heme-thiolate proteins responsible for catalyzing the oxidation of many endogenous substrates and xenobiotics.[1] In humans it has been estimated that P450s metabolize approximately 75% of all marketed drugs, 95% of which are metabolized by CYP3A4, 2D6, 2C9, 2C19, and 1A2.[2] The P450 mediated metabolism of a potential drug molecule can often lead to poor exposure of the compound resulting in the inability to reach pharmacologic concentrations and thus limiting its use as a therapeutic agent. Furthermore, two substrates of the same P450 may compete for the binding site of the enzyme resulting in elevated levels of one of the compounds and a potentially dangerous drug-drug interaction (DDI).

Due to the significance of P450 enzymes in metabolizing drugs, predicting if a drug candidate will bind to P450s and where it will be metabolized are germane during the drug discovery and development process. Prior studies predicting site of metabolism (SOM) of compounds interacting with P450s have utilized a variety of computational methods. Techniques previously used include: quantum chemical calculations, pharmacophore models, quantitative structure-activity relationships (QSAR), docking, molecular dynamics (MD) simulations, and basic chemical rules.[3–18] Recent reviews[19–21] published on P450 SOM prediction provide a good summary of prior studies and techniques used.

Although previous studies have been performed to predict SOM, there is no consensus on which method performs “best”. Recent research suggests that estimating the intrinsic reactivity and accessibility of individual ligand atoms can be used to predict SOM.[7,11] In general, the intrinsic reactivity of ligand atoms is normally estimated using Hartree-Fock, AM1, or DFT quantum mechanical calculations[3,5,8–10,22–24] while solvent accessible surface area (SASA) calculations[11,12,19,25] or docking[4,6,7,11,14–16,19,25–27] are used to model the accessibility of each ligand atom relative to the reactive center of the enzyme.

Olsen and co-workers have recently published a rule based reactivity model known as the neighboring atom type (NAT) model.[11] The NAT model predicts the activation energy required to abstract a particular hydrogen atom from a ligand, a common P450 reaction mechanism, based on the chemical environment of atoms that surround the putative metabolized atom.[11] Each NAT rule is derived from high level DFT calculations, and the model uses SASA values to discriminate between sites predicted with the same reactivity value. In a study of 60 CYP1A2 substrates[11], the NAT model was combined with GOLD[28] docking to predict SOM. The model was capable of predicting experimentally known SOM within the top-3 predictions for 87% of the compounds.

When considering promiscuous enzymes such as CYP3A4, CYP2C9, and CYP2D6 several pieces of evidence suggest that protein flexibility plays an important role in their ability to bind a diverse set of ligands and potentially complicates SOM prediction in such systems. Comparing experimentally determined X-ray structures with different ligands bound to the same P450 enzyme demonstrates the need to account for protein flexibility during ligand binding. For example, changes in side chain rotamer states are widely observed and are evident when comparing the N-palmitoylmethionine[29] bound structure of the extensively studied bacterial enzyme P450 BM-3 with the same protein with palmitoleic acid bound[30].

Illustrated in Figure 1A, residue Arg47 is engaged in water mediated hydrogen bonding to the carboxyl group of palmitoleic acid, but would have a steric overlap with the additional methionine substitution in palmitoylmethionine, thus resulting in a different configuration of the side chain.

In addition to side chain rotamer changes discussed above, alternative loop conformations are observed with CYP119 when different ligands are bound (Figure 1B–C). Compared to the apo structure of CYP119 (F/G loop in orange)[31], the F/G loop adopts two distinct configurations when 4-phenylimidazole (ligand and loop in magenta)[32] or imidazole (blue)[32] is bound. With CYP3A4 significant protein flexibility (Figure 1C) occurs in the F/G portion of the protein (apo: orange[33], erythromycin bound: blue[34]) to accommodate erythromycin and part of the F-F' loop becomes disordered. This motion causes the solvent accessible volume of the binding site to approximately increase by a factor of two and can dramatically affect ligand binding. CYP3A4 also exhibits a protein breathing motion (Figure 1D) that increases the size of the binding pocket to accommodate two ketoconazole molecules[34] without significant conformational changes of the helices or loop regions composing the binding pocket.

Discussed in the publication of the only available CYP2D6 X-ray structure (an apo structure without a co-crystallized ligand) the crystallographers note that the binding site of the crystal structure is not large enough to accommodate known CYP2D6 ligands.[35] In a recent publication by Oostenbrink and coworkers[6], the significance of protein flexibility was explored when predicting SOM in CYP2D6. This group showed that ensemble docking dramatically increases the accuracy of SOM prediction in CYP2D6 substrates, thus demonstrating the importance of protein flexibility during ligand binding and SOM prediction.

In this article, a modified version of the NAT+docking model was applied to 75 CYP2C9 ligands to predict SOM. In an attempt to improve SOM prediction, ensemble docking was utilized to incorporate protein flexibility into the prediction. MD simulations were used to create an ensemble of diverse protein structures used for subsequent docking experiments. Comparison of the results of SOM prediction using static X-ray docking with ensemble docking indicates that protein flexibility may be important for a subset of the library in order to produce binding poses consistent with the experimental SOM data. However, this analysis also outlines the difficulty of predicting the relative free energies of binding for different binding poses, a prerequisite for accurate SOM prediction. Our approach differs from the model of Olsen et al.[11] by the addition of a solvent-accessibility term and the inclusion of backbone flexibility of the protein (see Methods section for details).

2. Results and Discussion

2.1 CYP2C9 Ligand Library

A diverse library of 75 substrates metabolized by CYP2C9, all with experimentally determined SOM, was prepared using Maestro (Schrödinger, LCC) and MacroModel.[36] Each substrate was classified into low (1–10 μ M), medium (10–60 μ M), medium-high (60–100 μ M), or high (>100 μ M) K_m class based on K_m values reported in the literature[16] or extracted from the University of Washington Metabolism and Transporter Drug Interaction database© (available at www.druginteractioninfo.org). The database includes information from peer-reviewed journal articles referenced in PubMed, drug interaction sections of new drug applications (NDAs), and product labels for recently approved drugs. Detailed in Table 1, the library is composed of 23 low K_m substrates, 20 medium K_m substrates, 2 medium-high K_m substrates, 11 high K_m substrates, and 19 substrates that lack kinetic data with

recombinant CYP2C9. Substrates with unknown K_m values are marked as not available (NA) in the “ K_m Class” column of Table 1.

2.2 NAT Model Alone

To test if the NAT model derived by Olsen and coworkers can predict CYP2C9 SOM alone, the model was applied to the ligand library. The model determined if an experimentally known SOM was within the top-1, top-2, and top-3 computationally predicted sites and these results defined the accuracy of the NAT model alone.[7,11] Consideration of the top-3 computationally predicted sites is sufficient to allow medicinal chemists to interrogate multiple sites that are predicted to be a SOM for a particular ligand. In addition, many ligands are known to possess multiple experimentally determined SOM that result in different metabolites.

Shown in the column “NAT alone” of Table 2, the model predicted an experimentally known SOM as the top-1 computational site for 33% of the ligand library. When considering the top-2 and top-3 predicted sites, the model correctly predicted a known SOM for 53% and 64% of the ligand library. To determine if the NAT model predicts SOM more accurately than random guessing, a hyper-geometric distribution was applied to compute the probability of randomly selecting a known SOM from all possible heavy atoms of the ligand. Comparing the columns “Random Pick” and “NAT alone” in Table 2 it is evident that the NAT model is significantly more accurate when predicting SOM compared to random selection.

2.3 Static X-ray Docking

To test if static X-ray docking with AutoDock Vina[37] alone was capable of predicting SOM, the ligand library was docked to the static CYP2C9 X-ray structure (PDB-code: 1R9O[38]). To be considered a feasible technique to predict SOM, the docking program must produce binding poses that orient the ligand with its experimentally known SOM within a reactive distance to the heme iron atom, the known cofactor responsible for catalysis in P450 enzymes. To test if AutoDock Vina alone could accurately predict SOM, similar criteria as described in section “NAT Model Alone” were applied. If the top-1, top-2, or top-3 AutoDock Vina scored binding poses orient an experimentally known SOM within 5 Å of the heme iron atom a positive prediction was recorded.

Static X-ray docking with AutoDock Vina was capable of producing at least one binding pose with an experimentally known SOM within reactive distance to the heme iron atom for 85% of the compounds in the ligand library. Using the above criteria to define successful SOM prediction, column “Static X-ray Docking::Vina” in Table 2 reveals that static X-ray docking with AutoDock Vina alone only predicted an experimentally known SOM within the top-1 ranked binding pose in 19% of the ligand library. Considering the top-2 and the top-3 AutoDock Vina ranked binding poses moderately increased the percentage of the ligand library with an experimentally known SOM correctly predicted to 24% and 29%.

While AutoDock Vina was capable of producing a bioactive conformation for 85% of the ligand library, the internal AutoDock Vina scoring function did not accurately predict SOM, evident from the 29% success rate. Furthermore, the 29% success rate within the top-3 predicted binding poses was less accurate than the 33% top-1 success rate of the NAT model alone and the 34% top-3 rate when randomly selecting the SOM. The problem of predicting SOM using only AutoDock Vina appears to be related to the difficulty of accurately scoring binding poses within the large and unspecific binding site of CYP2C9. For some ligands, water mediated interactions may contribute to the free energy of binding, and such interactions are not considered. In many instances, different binding poses of the same

ligand are predicted with similar binding energies making the selection of the top-ranked poses somewhat ambiguous. It is also important to note that different ligand configurations exist in the binding site of P450s naturally. The pose with the lowest predicted binding energy in the large unspecific binding site of P450s may not be consistent with the experimental SOM data. Thus, using only protein-ligand interaction energies as criteria to predict SOM appears to be insufficient.

To increase the accuracy of SOM prediction, a rescoring function was applied that combines a ligand reactivity model with docking similar to the one first proposed by Jung and coworkers and shown to be effective in CYP1A2.[7] Combining the AutoDock Vina binding pose score, the neighboring atom type (NAT) model, and a SASA term creates what is referred to in Table 2 as the Vina+NAT+SASA scoring function. To initially filter the large pool of binding poses produced for each ligand, only poses that place at least one ligand atom within reactive distance of the heme iron atom were subjected to rescoring by the Vina+NAT+SASA function.

Following rescoring, static X-ray docking accurately predicted an experimentally known SOM as the top-1 computationally predicted site for 44% of ligands (column “Static X-Ray Docking::Vina+NAT+SASA” in Table 2). Considering the top-2 and top-3 predicted sites increased the percentage of ligands with successfully predicted SOM to 59% and 73%. Comparing the columns “NAT alone” and “Static X-Ray Docking::Vina+NAT+SASA” in Table 2, AutoDock Vina static X-ray docking in combination with the Vina+NAT+SASA rescoring provides a 11% improvement in the top-1 predicted SOM compared to using the NAT model alone.

2.4 Ensemble Docking

Although static X-ray docking was capable of producing at least one bioactive conformation for 85% of the CYP2C9 ligand library, additional protein flexibility may be required to produce bioactive conformations for the remaining 15% of ligands. Shown to be effective for CYP2D6, ensemble docking has been used to incorporate protein flexibility and thereby improve SOM prediction.[6] To test if ensemble docking increases the number of ligands producing bioactive conformations in CYP2C9, an MD simulation was performed on the X-ray structure with the co-crystallized flurbiprofen removed from the binding site (creating a “pseudo-apo” simulation). The underlying principle of using the “pseudo-apo” MD simulation to generate alternative protein conformations is the population-shift mechanism[39–41] of protein flexibility coupled to ligand-binding. This theory assumes that all possible protein conformations are sampled by the apo protein and the ligand selects its associated holo protein conformation from the ensemble of protein structures during binding.

After the completion of the MD simulation, a diverse set of 10 “pseudo-apo” protein structures was extracted from the trajectory using QT clustering[42]. Ensemble docking with AutoDock Vina was then performed with identical docking parameters and settings as used in static X-ray docking. Operating under the assumption that introducing protein flexibility through ensemble docking would increase the number of ligands successfully docked in a bioactive conformation consistent with the experimental SOM data, the opposite trend was actually observed. Ensemble docking into the 10 MD protein structures did not increase the percentage of ligands with at least one conformation consistent with the SOM data compared to static X-ray docking (static 84% vs. ensemble 83%, data not shown). Examining the docking statistics in more detail, for several ligands, it was noted that docking into the static X-ray structure produced bioactive conformations while ensemble docking produced none. This information led to the hypothesis that such ligands may need additional protein flexibility upon binding (induced fit[43]), possibly a protein conformation

similar to the holo X-ray structure, to produce bioactive conformations. The 10 diverse structures collected from the “pseudo-apo” MD simulation may not represent such a structure. The average side chain heavy atom RMSD of all amino acids within the docking search space of each ensemble structure compared to the X-ray structure ranges from 0.9Å to 1.9Å. Considering each amino acid individually, the side chain heavy atom RMSD value of at least one residue in the binding site ranges from 2.0Å to 5.3Å compared to the X-ray structure. Thus, the ensemble used here lacks a structure that closely approximates the X-ray holo structure of CYP2C9 bound to flurbiprofen, a protein conformation likely relevant for other ligands in the library as well. The magnitude of the observed difference in side chain orientation compared to the holo X-ray structure is known to significantly affect docking results.[44–47]

To address the problem of including holo-like protein conformations, several options were explored: 1) Adding the holo X-ray protein conformation to the 10 apo-ensemble structures, 2) running a separate holo-structure MD simulation, QT clustering to produce 10 holo-ensemble structures, and subsequent docking, 3) adding the holo X-ray protein conformation to the 10 holo-structures, and 4) mixing the 10 apo-ensemble protein structures with the 10 holo-ensemble structures to create one large representative ensemble. Based on the results of each study, adding the X-ray protein conformation to the apo-ensemble was the most suitable option leading to the best overall results (data not shown).

Following the addition of the X-ray protein structure to the ensemble, the percentage of ligands docked with at least one bioactive conformation increased to 96% (column “Ensemble Docking” in Table 2). The additional ligands that were docked in a bioactive conformation spanned all K_m classes. Figure 2 displays two examples of ligands where static X-ray docking produced no bioactive conformation consistent with experimental SOM data while ensemble docking found a bioactive conformation. The included protein flexibility allows such ligands additional space and alternative protein contacts within the binding site and permits AutoDock Vina to produce bioactive binding poses.

Using the same criteria as before to determine if AutoDock ensemble docking alone can be used to predict SOM, the top-1 ranked binding pose determined with “Ensemble Docking::Vina” (Table 2) accurately predicted an experimentally known SOM in only 12% of the ligand library. The top-2 and top-3 AutoDock Vina ranked binding poses successfully predicted an experimentally known SOM for 13% and 20% of the ligands. Although ensemble docking increased the number of ligands docked in a bioactive conformation, column “Ensemble Docking::Vina” in Table 2 shows that the ability to accurately predict SOM with ensemble docking is decreased compared to static X-ray docking. The increased number of protein templates used in ensemble docking leads to a direct increase in the number of binding poses produced for each ligand in the library. As mentioned above, it is not feasible to use the AutoDock Vina scoring function to rank the binding poses in the promiscuous binding site of CYP2C9. As a result of the increased number of generated binding poses, more binding poses inconsistent with the experimental SOM data are ranked higher than bioactive poses, leading to the decrease in SOM prediction accuracy. Ensemble docking’s decreased ability to predict SOM illustrates a trend encountered throughout the study; used alone, AutoDock Vina is not a suitable tool to predict SOM due to the challenging task of scoring binding poses within the binding site and the fact that AutoDock Vina’s scoring function was not optimized for CYP2C9.

To address AutoDock Vina’s internal scoring issue, the Vina+NAT+SASA scoring function was applied to re-rank binding poses. Shown in column “Ensemble Docking::Vina+NAT+SASA” of Table 2, Vina+NAT+SASA rescoring significantly improved ensemble docking SOM prediction with an experimentally known SOM in the top-1 predicted position for 49%

of the ligand library. While this result is a significant improvement compared to using the internal AutoDock Vina's scoring function, it represents only a slight increase in accuracy compared to static X-ray docking results using the Vina+NAT+SASA scoring function. The slight increase in ensemble docking SOM prediction accuracy in the top-1 position (comparing the columns "Static X-Ray Docking::Vina+NAT+SASA" and "Ensemble Docking::Vina+NAT+SASA" in Table 2) can be attributed to the larger number of binding poses produced for each ligand in the library when additional protein templates are used during docking. While a detriment to SOM prediction with AutoDock Vina scoring alone (column "Ensemble Docking::Vina" in Table 2), the larger pool of binding poses produced by ensemble docking results in the identification of bioactive poses consistent with experimental SOM data for additional ligands that produced no bioactive poses in static X-ray docking.

With regard to the top-2 and top-3 computationally predicted sites, in 61% and 71% of the ligand library an experimentally known SOM was successfully predicted. Compared to the static docking results, the overall success of identifying an experimentally known SOM within the top-3 computationally predicted sites was not significantly improved using ensemble docking in combination with the Vina+NAT+SASA scoring function. However, ensemble docking provides additional important information as it generates bioactive poses for all except two ligands (96% of all substrates) compared to 85% of all substrates using static docking. This increase is scientifically meaningful and will improve SOM prediction if one can more accurately rank the additional bioactive poses produced by ensemble docking with an optimized scoring function. Future research aims to perform such an optimization using a scoring function shown to be successful in P450 enzymes.[48]

2.5 Classification of Substrates by K_m values

In addition to the cumulative CYP2C9 substrate library statistics presented in Table 2, the substrates were separated into four classes based on K_m values to determine if SOM prediction was more accurate for low K_m (good) substrates compared to high K_m (poor) substrates. Due to the limited number of substrates in the medium-high class, these compounds were added to the high K_m substrate class. In order to fairly compare the K_m class statistics with the cumulative ligand library, substrates lacking kinetic data were removed from the ligand library and the cumulative results were re-tabulated and are presented in Table 3. Comparing the percentages in column "NAT Alone" of Table 3 indicates that the NAT model yielded approximately the same level of accuracy in predicting SOM in the top-1 ranked position for low K_m substrates and is slightly more accurate for medium K_m substrates compared to the cumulative library results. For high K_m substrates, the NAT model was slightly less accurate in predicting SOM compared to cumulative library results. The top-1 computationally predicted position was an experimentally known SOM for 23% of the high K_m substrates compared to 32% for the cumulative library.

Considering the top-3 predicted sites, an experimentally known SOM was correctly identified in 54% of the high K_m substrates versus 66% for the cumulative ligand library. While the top-3 predicted sites were slightly less accurate for high K_m substrates, the opposite trend was observed for low and medium K_m substrates. A known SOM was correctly identified in the top-3 predicted sites for 70% of ligands in these classes compared to 66% for the cumulative ligand library. Based on this information, it was concluded that the NAT model is slightly more accurate in predicting SOM for low and medium K_m substrates, but there is no appreciable increase in the accuracy of predicting an experimentally known SOM in the top-1 predicted site. The model identified a known SOM for 30% of the low K_m substrates and 40% of the medium K_m substrates in the top-1 position compared to 32% of the cumulative library.

In contrast to the reactivity-only model (“NAT Alone”) results, a more significant difference between the K_m classes was identified for the top-1 predicted SOM using the reactivity-accessibility model in combination with ensemble docking (Column: “Ensemble X-Ray::Vina+NAT+SASA” in Table 3). Only 39% of compounds in the poor substrate class had the correct SOM predicted in the top-1 position while 57% of low K_m substrates and 60% of the medium K_m substrates were accurately predicted. This represents approximately a 10% increase in prediction accuracy of an experimentally known SOM in the top-1 position for low K_m and medium K_m compounds compared to applying our Vina+NAT+SASA scoring function on the cumulative ligand library. Comparing the columns “Ensemble Docking::Vina+NAT+SASA” and “NAT Alone” in Table 3 indicates that this method is more accurate in predicting an experimentally known SOM at the top-1 position for the class of substrates with low K_m values compared to the NAT model alone.

To summarize these results, both low and high K_m substrates were predicted with similar accuracy when using the reactivity model alone. Including protein-ligand interactions during the prediction process significantly improved SOM prediction. With respect to hydroxylation reactions, this suggests that the inherent reactivity of the substrate does not determine SOM alone and that inclusion of protein-ligand information in SOM prediction is particularly important for low K_m substrates. This conclusion is further supported by the fact that the predicted hydrogen abstraction energy as the rate limiting step of the chemical reaction is indistinguishable between the different K_m classes; the mean hydrogen abstraction energies for low K_m , medium K_m and high K_m substrates are 53.5 ± 1.5 kJ/mol, 49.7 ± 3.8 kJ/mol, and 53.0 ± 3.6 kJ/mol. Neglecting any atypical non-Michaelis-Menten contributions to the enzyme kinetics, it is assumed that the difference in K_m values between the classes is determined primarily by the difference in binding affinities of ligands.

To determine if high binding affinity originates from the hydrophobicity of a molecule rather than from specific interactions, the average logP values[49] were computed for the different K_m classes. Such an analysis revealed significant differences between the classes, the average logP value for low K_m , medium K_m , and high K_m substrates were 4.07 ± 0.28 , 3.50 ± 0.33 , and 3.10 ± 0.29 . Based on this analysis, it was concluded that the difference in K_m values partially originates from variations in the hydrophobicity of the molecules. Compounds with high logP value tend to have a higher binding affinity in the hydrophobic binding pocket of CYP2C9 and reside in the low K_m class more frequently than compounds with low logP values.

A clear separation was observed based on ligand size between the different K_m classes. The average molecular weight for low K_m , medium K_m , and high K_m substrates were 329.3 ± 14.7 , 339.5 ± 19.9 , and 238.3 ± 19.6 . In previous docking studies[50], a trend between docking success and the ratio between the ligand and binding pocket size was noted; if the ligand size was relatively small compared to the size of the binding pocket, many diverse binding poses were generated making the identification of the bioactive pose difficult. Thus, it is speculated that the bioactive binding mode of low K_m substrates (which on average are larger in size than high K_m substrates) is more readily identified during docking, thereby explaining the increase in correct SOM prediction using the Vina+NAT+SASA model for this class of substrates compared to high K_m substrates.

2.6 Structural Analysis

Based on the statistical analyses presented in Tables 2 and 3, SOM prediction that includes docking is more accurate than using the NAT reactivity model alone. However, due to the limited accuracy of the scoring function, the combination of docking and the reactivity model is still less than optimal for predicting SOM as only approximately 50% of the CYP2C9 library had an experimentally known SOM ranked in the top-1 predicted position.

Nevertheless, the inclusion of docking into the SOM prediction process has significant value as a descriptive tool that can be used as part of an *in silico/in vitro* assessment of metabolism. Including structural information during the prediction process provides an atomic picture of substrate-P450 interactions that can be used to identify critical contacts between the substrate and protein and guide structural modifications of the ligand to improve its metabolic stability. To illustrate this strategy, the protein-ligand interactions were analyzed for four substrates in their predicted bioactive conformation.

Shown in Figure 3A, the bioactive pose of flurbiprofen (cyan) is in agreement with the co-crystallized orientation (magenta) with an RMSD of 1.82Å. If one considers that the position of the fluorine substituent cannot be assigned with absolute certainty during the interpretation of the electron density due to its small size, the symmetrical rotation of the fluorinated ring leads to an RMSD of 1.28Å between the crystal structure and the bioactive pose. The flurbiprofen carboxylic group forms critical interactions with Arg108 and Asn204 while the aromatic rings form a pi-stacking interaction with Phe144. Tienilic acid, a compound of similar size as flurbiprofen and containing a carboxyl moiety, is oriented in the same region as the co-crystallized flurbiprofen and shares similar critical contacts with residues Arg108, Asn204, and Phe144 (Figure 3B). Figure 3C shows the predicted bioactive pose of fluvastatin (cyan), a compound larger than flurbiprofen but still containing a carboxylic moiety, is oriented differently in the binding site than the co-crystallized flurbiprofen. The increased size of fluvastatin leads to the carboxylic group being placed in a region of the binding site closer to the surface of the protein. This portion of the binding site could be the entrance channel for fluvastatin. Thus, the carboxylic-acid group of fluvastatin is likely to interact with water molecules but does not interact with the Arg108 or Asn204 residue. Illustrated in Figure 3D, losartan is oriented in the same region of the binding site as the co-crystallized flurbiprofen molecule, but forms different types of contacts with amino acids in the binding site. The tetrazole ring forms a cation-pi and strong electrostatic interaction with positively charged Arg108 residue but appears to have no direct interaction with Asn204. Losartan is further stabilized by pi-pi interactions with Phe100 and Phe114.

3. Conclusions

The goal of this study was to determine if SOM of a ligand library containing 75 diverse CYP2C9 substrates could be predicted using a combination of structure-based and reactivity-based computational techniques and whether the inclusion of protein flexibility improves the accuracy of SOM prediction. We demonstrated that a combined structure-/reactivity-based model was superior in predicting SOM than a reactivity- or structure-based model alone. Including protein flexibility in the combined model increased the percentage of the ligand library docked in a bioactive conformation from 85% to 96%. Although the number of substrates successfully docked increased, correctly ranking the additional bioactive poses proved problematic as the SOM prediction results of the top-1 position were similar in static and ensemble docking.

Further classifying the substrate library according to published K_m values demonstrated that the reactivity-accessibility model, in combination with docking, improved SOM prediction for substances with low K_m values. Using ensemble docking and the Vina+NAT+SASA scoring function, an experimentally known SOM was computationally predicted in the top-1 position for 57% of low K_m substrates, whereas only 39% of high K_m substrates had a known SOM in the same position.

While more accurate than the NAT reactivity-only model, it should be emphasized that docking in its current form is not ideal as a standalone SOM prediction tool. More than 80% of the CYP2C9 library did not have an experimentally known SOM in the top-1 predicted

position. However, docking is valuable as a descriptive tool that can be used as part of an *in silico/in vitro* assessment of metabolism. Including structural information pertaining to the protein-ligand complex not only improves SOM prediction, but also provides researchers in drug development a tool to explore hypotheses related to assessing key ligand-CYP2C9 interactions responsible for binding affinity. This information can then be utilized to rationally improve the metabolic stability of drug candidates not only by 'blocking' the SOM but also by rationally designing ligand modifications to alter its affinity for CYP2C9.

In the future, applying more advanced docking techniques in combination with a tailor-made scoring function for P450 enzymes and a more detailed reactivity model may further improve SOM prediction for CYP2C9 and other P450 isoforms. Equally important, additional experimental data such as X-ray co-crystal structures exploring CYP2C9-ligand interactions with a diverse set of compounds would provide a better picture of the induced-fit effect.

4. Methods

4.1 Preparation of the CYP2C9 Ligand Library

A diverse set of 75 substrates metabolized by CYP2C9 (with experimentally determined site of metabolism reported) was created by literature mining; 63 substrates were taken from a study done by Sykes and coworkers[16] and 12 compounds were extracted from the University of Washington Metabolism and Transporter Drug Interaction database© (available at www.druginteractioninfo.org). All possible stereoisomers and relevant protonation states were created for each ligand. A library containing a total of 147 compounds was created when different stereoisomers and protonation states were considered as unique compounds (e.g. a01 *r* stereoisomer, a01 *s* stereoisomer, a01 *r* stereoisomer deprotonated, and a01 *s* stereoisomer deprotonated represent four unique compounds in the library). A three dimensional representation of each ligand was created with Maestro (Schrödinger, LCC) and subjected to energy minimization using MacroModel[36].

4.2 Static X-ray Docking

The crystal structure of CYP2C9, PDB code 1R9O[38], was used in all studies performed for this publication. All resolved water molecules, cofactors other than the heme prosthetic group, and the co-crystallized flurbiprofen were removed from the structure prior to protein preparation and subsequent docking simulations. To optimize the protein hydrogen bonding network, the program Reduce[51] was used to assign histidine protonation states and to flip asparagine and glutamine side chains when needed.

An in-house PyMol plugin script[52] was used to automatically prepare the protein system for docking with the AutoDock Vina[37] program. First, the plugin utilizes the AutoDockTools scripts to create a .pdbqt file for every ligand in the library. The AutoDock Vina.pdbqt file is a modified PDB format that contains additional information such as atomic charges, atom type definitions, and topological information such as rotatable bonds and is required for AutoDock Vina docking.

In a subsequent step, the docking search space was visually defined using our in-house PyMol plugin. AutoDock Vina requires a rectangular box definition for search space and the plugin allows the user to visually determine the dimensions and placement of the box to define the search volume. The box was defined to contain the known binding site of the co-crystallized flurbiprofen and was large enough to include the complete F/G loop region in order to allow all ligands enough space to freely rotate during docking. The plugin then automatically uses AutoDockTools to create the receptor .pdbqt file needed for docking.

Default values were used for all docking parameters, the maximum 20 binding poses was outputted for each ligand, and a maximum energy difference between the best and worst binding pose was chosen to be 10 kcal/mol.

In all experiments three separate docking runs were performed as it was noted that automatic random seed assignment by AutoDock Vina can affect docking pose generation (data not shown). An in-house python program was developed to analyze the docking results, perform post-docking file manipulation, and provide summary statistics of the docking experiment. Following the completion of all three docking runs, the in-house python program was used to pool all binding poses produced for each ligand across all three docking runs. The pooled binding poses were subsequently clustered to form one single set of unique binding poses for each ligand. During the clustering procedure the binding pose with the most negative AutoDock Vina score was selected as the first unique binding pose. Subsequent binding poses were then added (based on increasing AutoDock Vina scores) to the unique set if they were at least 1.5Å root mean squared deviation (RMSD) different compared to all previously added unique binding poses.

For ligands with multiple protonation states or stereoisomers, the unique set of binding poses from all protonation states and stereoisomers were combined to form a single ensemble of unique top ranked binding poses (e.g. unique binding poses for a01 *r* stereo isomer, a01 *s* stereoisomer, a01 *r* stereoisomer deprotonated, and a01 *s* stereoisomer deprotonated were combined to form a single ensemble binding poses for ligand a01).

Our in-house python program was then used to determine the distance between the heme iron atom and the known reported SOM for each ligand. A successful prediction of SOM using docking occurs if the binding pose orients a known ligand SOM within 5.0Å of the heme iron atom. Following distance calculations, the python program then generates summary statistics for all ligands indicating if AutoDock Vina successfully predicts SOM within the top ranked binding poses. Filtering bioactive conformations, rescoring of bioactive poses with the Vina+NAT+SASA scoring function, and final Vina+NAT+SASA summary statistics are also automatically done inside the python program.

4.3 Ensemble Docking

Protein preparation was identical to the procedure described in the previous section “Static X-ray Docking”. To allow the static X-ray structure to relax and adopt alternative protein conformations for the subsequent docking studies, a MD simulation was performed using Amber9[53,54]. First, the prepared protein system was solvated in a 30Å explicit water cap centered on the position on the co-crystallized flurbiprofen molecule. In order to obtain a low energy conformation of the solvated system, 250 steps of steepest decent followed by 250 steps of conjugate gradient minimization was performed. To allow the water cap to further equilibrate and adopt a lower energy configuration, 250ps of position restrained MD was performed. All non water residues were restrained by a force constant of 5 kcal/mol per Å² while all water molecules were contained within the initial 30Å cap region by a 5 kcal/mol per Å² cap restraint. 2ns of subsequent production MD was performed with all protein residues beyond 25Å from the center of the water cap restrained by a force constant of 5 kcal/mol per Å². The simulation was run at a temperature of 300K and used a 2ps time step.

Following MD simulation of the X-ray structure, QT clustering[42] was applied to the production MD trajectory to obtain ten diverse protein structures sampled throughout the simulation. A maximum RMSD of 0.93Å between each cluster center and all other members of the cluster was automatically identified generating 10 clusters. The holo X-ray structure was added creating a pool of 11 protein structures to be used in ensemble docking.

Ensemble docking preparation was identical to the procedure described in “Static X-ray Docking” except that all 11 ensemble structures were automatically prepared for docking using our PyMOL plugin in combination with AutoDockTools. Identical docking search volume definitions, AutoDock Vina parameters, and three separate docking runs were applied to all ensemble structures.

Upon completion of the docking simulations, the in-house python program was again used to pool all binding poses produced for each ligand across all 11 ensemble structures (and all three docking runs). As described in “Static X-ray Docking”, the pooled binding poses were subsequently clustered to form one single set of unique binding poses for each ligand. In instances where the ligand has multiple protonation states or stereoisomers, the unique set of binding poses from each protonation state and stereoisomer was combined to form a single ensemble of top ranked binding poses. Ensemble docking summary statistics were automatically generated by our in-house python program.

4.4 Vina+NAT+SASA Scoring Function

Based on the reactivity-binding concept where SOM is predicted by a summation of the binding energy and activation energy, AutoDock Vina docking scores were used to estimate the binding energy and the NAT model was used to estimate the activation energy. This approach is similar to other previously published work[11], but was extended by adding a SASA term to create what is defined as the Vina+NAT+SASA scoring function. Based on the assumption that hydrogen atoms with $SASA < 8\text{\AA}^2$ lack sufficient exposure to be abstracted and therefore are not SOM[12], equation 1 was applied to derive the SASA term in the Vina+NAT+SASA scoring function.

$$SASA = k - \alpha \ln \frac{SASA}{SASA_0} \quad (1)$$

The constants $SASA_0 = 8\text{\AA}^2$, $k = 1 \text{ kcal/mol}$, and $\alpha = 1 \text{ kcal/mol}$ were used in all SASA term calculations. Utilizing the logarithmic function accessible hydrogen atoms were associated with a moderately favorable score while the overall score of hydrogen atoms lacking sufficient exposure becomes penalized. Hydrogen atom SASA calculations were performed with MSROLL[55] using a probe radius of $1.4\text{\AA}$ and default atomic radii. In cases where MSROLL predicts a hydrogen atom SASA value equal to zero, the SASA value was automatically reassigned to a value of $0.1\text{\AA}$ to allow the use of Equation (1) for all accessibility values.

Acknowledgments

The authors would like to thank Laura Kingsley for assistance in preparing figures for the manuscript. M.A.L. gratefully acknowledges funding from the National Institutes of Health (GM085604 and GM092855), the Purdue Research Foundation, and Eli Lilly and Company.

References

1. Wrighton SA, Van den Branden M, Ring BJ. The human drug metabolizing cytochromes P450. *J Pharmacokinet Biop.* 1996; 24:461–473.
2. Wienkers LC, Heath TG. Predicting in vivo drug interactions from in vitro drug discovery data. *Nature Rev Drug Discov.* 2005; 4:825–833. [PubMed: 16224454]
3. Cruciani G, Carosati E, De BB, 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. [PubMed: 16250655]

4. de GC, Vermeulen NP, Feenstra KA. Cytochrome p450 in silico: an integrative modeling approach. *J Med Chem.* 2005; 48:2725–2755. [PubMed: 15828810]
5. Hennemann M, Friedl A, Lobell M, Keldenich J, Hillisch A, Clark T, Goller AH. CypScore: Quantitative prediction of reactivity toward cytochromes P450 based on semiempirical molecular orbital theory. *ChemMedChem.* 2009; 4:657–669. [PubMed: 19243088]
6. Hritz J, de RA, 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. [PubMed: 18998665]
7. Jung J, Kim ND, Kim SY, Choi I, Cho KH, Oh WS, Kim DN, No KT. Regioselectivity prediction of CYP1A2-mediated phase I metabolism. *J Chem Inf Model.* 2008; 48:1074–1080. [PubMed: 18412330]
8. Kim DN, Cho KH, Oh WS, Lee CJ, Lee SK, Jung J, No KT. EaMEAD: Activation energy prediction of cytochrome P450 mediated metabolism with effective atomic descriptors. *J Chem Inf Model.* 2009; 49:1643–1654. [PubMed: 19545128]
9. Korzekwa KR, Jones JP, Gillette JR. Theoretical-Studies on Cytochrome-P-450 Mediated Hydroxylation - A Predictive Model for Hydrogen-Atom Abstractions. *J Am Chem Soc.* 1990; 112:7042–7046.
10. Olsen L, Rydberg P, Rod TH, Ryde U. Prediction of activation energies for hydrogen abstraction by cytochrome p450. *J Med Chem.* 2006; 49:6489–6499. [PubMed: 17064067]
11. Rydberg P, Vasanthanathan P, Oostenbrink C, Olsen L. Fast prediction of cytochrome P450 mediated drug metabolism. *ChemMedChem.* 2009; 4:2070–2079. [PubMed: 19852016]
12. Singh SB, Shen LQ, Walker MJ, Sheridan RP. A model for predicting likely sites of CYP3A4-mediated metabolism on drug-like molecules. *J Med Chem.* 2003; 46:1330–1336. [PubMed: 12672233]
13. Vermeulen NPE. Prediction of drug metabolism: The case of cytochrome P450 2D6. *Curr Top Med Chem.* 2003; 3:1227–1239. [PubMed: 12769702]
14. Zamora I, Afzelius L, Cruciani G. Predicting drug metabolism: A site of metabolism prediction tool applied to the cytochrome P4502C9. *J Med Chem.* 2003; 46:2313–2324. [PubMed: 12773036]
15. Tarcsay A, Kiss R, Keseru GM. Site of metabolism prediction on cytochrome P450 2C9: a knowledge-based docking approach. *J Comput Aid Mol Des.* 2010; 24:399–408.
16. Sykes MJ, McKinnon RA, Miners JO. Prediction of metabolism by cytochrome P450 2C9: alignment and docking studies of a validated database of substrates. *J Med Chem.* 2008; 51:780–791. [PubMed: 18237107]
17. Vasanthanathan P, Hritz J, Taboureau O, Olsen L, Jorgensen FS, Vermeulen NP, Oostenbrink C. Virtual screening and prediction of site of metabolism for cytochrome P450 1A2 ligands. *J Chem Inf Model.* 2009; 49:43–52. [PubMed: 19099399]
18. Li H, Sun J, Fan X, Sui X, Zhang L, Wang Y, He Z. Considerations and recent advances in QSAR models for cytochrome P450-mediated drug metabolism prediction. *J Comput Aid Mol Des.* 2008; 22:843–855.
19. Afzelius L, Arnby CH, 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, mechanistical insights, and future applications. *Drug Metab Rev.* 2007; 39:61–86. [PubMed: 17364881]
20. Crivori P, Poggesi I. Computational approaches for predicting CYP-related metabolism properties in the screening of new drugs. *Eur J Med Chem.* 2006; 41:795–808. [PubMed: 16644065]
21. Stjernschantz E, Vermeulen NP, Oostenbrink C. Computational prediction of drug binding and rationalisation of selectivity towards cytochromes P450. *Expert Opin Drug Metab Toxicol.* 2008; 4:513–527. [PubMed: 18484912]
22. Jones JP, Mysinger M, Korzekwa KR. Computational models for cytochrome P450: A predictive electronic model for aromatic oxidation and hydrogen atom abstraction. *FASEB J.* 2003; 17:A608.
23. Rydberg P, Ryde U, Olsen L. Prediction of activation energies for aromatic oxidation by cytochrome P450. *J Phys Chem A.* 2008; 112:13058–13065. [PubMed: 18986131]

24. Rydberg P, Ryde U, Olsen L. Sulfoxide, sulfur, and nitrogen oxidation and dealkylation by cytochrome P450. *J Chem Theory Comput.* 2008; 4:1369–1377.
25. Oh WS, Kim DN, Jung J, Cho KH, No KT. New combined model for the prediction of regioselectivity in cytochrome P450/3A4 mediated metabolism. *J Chem Inf Model.* 2008; 48:591–601. [PubMed: 18290640]
26. Prusis P, Afzelius L. Improvement of Site of Metabolism Predictions for CYP3A4 by Using Discriminant Analysis of Compound Preference of CYP3A4 X-Ray Structural Conformers and Subsequent Docking. *QSAR Comb Sci.* 2009; 28:865–868.
27. Zheng MY, Luo XM, Shen QC, Wang Y, Du Y, Zhu WL, Jiang HL. Site of metabolism prediction for six biotransformations mediated by cytochromes P450. *Bioinformatics.* 2009; 25:1251–1258. [PubMed: 19286831]
28. Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. *J Mol Biol.* 1997; 267:727–748. [PubMed: 9126849]
29. Hegde A, Haines DC, Bondlela M, Chen B, Schaffer N, Tomchick DR, Machius M, Nguyen H, Chowdhary PK, Stewart L, Lopez C, Peterson JA. Interactions of substrates at the surface of P450s can greatly enhance substrate potency. *Biochemistry.* 2007; 46:14010–14017. [PubMed: 18004886]
30. Li H, Poulos TL. The structure of the cytochrome p450BM-3 haem domain complexed with the fatty acid substrate, palmitoleic acid. *Nat Struct Biol.* 1997; 4:140–146. [PubMed: 9033595]
31. Park SY, Yamane K, Adachi S, Shiro Y, Weiss KE, Maves SA, Sligar SG. Thermophilic cytochrome P450 (CYP119) from *Sulfolobus solfataricus*: high resolution structure and functional properties. *J Inorg Biochem.* 2002; 91:491–501. [PubMed: 12237217]
32. Yano JK, Koo LS, Schuller DJ, Li H, Ortiz de Montellano PR, Poulos TL. Crystal structure of a thermophilic cytochrome P450 from the archaeon *Sulfolobus solfataricus*. *J Biol Chem.* 2000; 275:31086–31092. [PubMed: 10859321]
33. Yano JK, Wester MR, Schoch GA, Griffin KJ, Stout CD, Johnson EF. The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-Å resolution. *J Biol Chem.* 2004; 279:38091–38094. [PubMed: 15258162]
34. Ekroos M, Sjogren T. Structural basis for ligand promiscuity in cytochrome P450 3A4. *Proc Natl Acad Sci U S A.* 2006; 103:13682–13687. [PubMed: 16954191]
35. Rowland P, Blaney FE, Smyth MG, Jones JJ, Leydon VR, Oxbrow AK, Lewis CJ, Tennant MG, Modi S, Eggleston DS, Chenery RJ, Bridges AM. Crystal structure of human cytochrome P450 2D6. *J Biol Chem.* 2006; 281:7614–7622. [PubMed: 16352597]
36. Mohamadi F, Richards NGJ, Guida WC, Liskamp R, Lipton M, Caufield C, Chang G, Hendrickson T, Still WC. MacroModel - An Integrated Software System for Modeling Organic and Bioorganic Molecules Using Molecular Mechanics. *J Comput Chem.* 1990; 11:440–467.
37. Trott O, Olson AJ. Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J Comput Chem.* 2010; 31:455–461. [PubMed: 19499576]
38. Wester MR, Yano JK, Schoch GA, Yang C, Griffin KJ, Stout CD, Johnson EF. The structure of human cytochrome P450 2C9 complexed with flurbiprofen at 2.0-Å resolution. *J Biol Chem.* 2004; 279:35630–35637. [PubMed: 15181000]
39. Frauenfelder H, Sligar SG, Wolynes PG. The Energy Landscapes and Motions of Proteins. *Science.* 1991; 254:1598–1603. [PubMed: 1749933]
40. Tsai CJ, Kumar S, Ma B, Nussinov R. Folding funnels, binding funnels, and protein function. *Protein Sci.* 1999; 8:1181–1190. [PubMed: 10386868]
41. Ma B, Kumar S, Tsai CJ, Nussinov R. Folding funnels and binding mechanisms. *Protein Eng.* 1999; 12:713–720. [PubMed: 10506280]
42. Heyer LJ, Kruglyak S, Yooseph S. Exploring expression data: identification and analysis of coexpressed genes. *Genome Res.* 1999; 9:1106–1115. [PubMed: 10568750]
43. Koshland DE. Application of A Theory of Enzyme Specificity to Protein Synthesis. *Proc Natl Acad Sci U S A.* 1958; 44:98–104. [PubMed: 16590179]

44. Kua J, Zhang Y, McCammon JA. Studying enzyme binding specificity in acetylcholinesterase using a combined molecular dynamics and multiple docking approach. *J Am Chem Soc.* 2002; 124:8260–8267. [PubMed: 12105904]
45. McGovern SL, Shoichet BK. Information decay in molecular docking screens against holo, apo, and modeled conformations of enzymes. *J Med Chem.* 2003; 46:2895–2907. [PubMed: 12825931]
46. Murray CW, Baxter CA, Frenkel AD. The sensitivity of the results of molecular docking to induced fit effects: application to thrombin, thermolysin and neuraminidase. *J Comput Aided Mol Des.* 1999; 13:547–562. [PubMed: 10584214]
47. Hoffmann D, Kramer B, Washio T, Steinmetzer T, Rarey M, Lengauer T. Two-stage method for protein-ligand docking. *J Med Chem.* 1999; 42:4422–4433. [PubMed: 10543886]
48. Rydberg P, Gloriam DE, Zaretski 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.
49. Cheng T, Zhao Y, Li X, Lin F, Xu Y, Zhang X, Li Y, Wang R, Lai L. Computation of octanol-water partition coefficients by guiding an additive model with knowledge. *J Chem Inf Model.* 2007; 47:2140–2148. [PubMed: 17985865]
50. Warren GL, Andrews CW, Capelli AM, Clarke B, LaLonde J, Lambert MH, Lindvall M, Nevins N, Semus SF, Senger S, Tedesco G, Wall ID, Woolven JM, Peishoff CE, Head MS. A critical assessment of docking programs and scoring functions. *J Med Chem.* 2006; 49:5912–5931. [PubMed: 17004707]
51. Word JM, Lovell SC, Richardson JS, Richardson DC. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J Mol Biol.* 1999; 285:1735–1747. [PubMed: 9917408]
52. Lill MA, Danielson ML. Computer-aided drug design platform using PyMOL. *J Comput Aided Mol Des.* 2011; 25:13–19. [PubMed: 21053052]
53. Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ. The Amber biomolecular simulation programs. *J Comput Chem.* 2005; 26:1668–1688. [PubMed: 16200636]
54. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, Debolt S, Ferguson D, Seibel G, Kollman P. Amber, A Package of Computer-Programs for Applying Molecular Mechanics, Normal-Mode Analysis, Molecular-Dynamics and Free-Energy Calculations to Simulate the Structural and Energetic Properties of Molecules. *Comput Phys Commun.* 1995; 91:1–41.
55. Connolly ML. The molecular surface package. *J Mol Graph.* 1993; 11:139–141. [PubMed: 8347567]

Figure 1.

Examples of protein flexibility in cytochrome P450 enzymes: A) Changes in Arg47 side chain rotamer in P450 BM-3 depending on the bound ligand (palmitoleic acid and corresponding protein in blue: PDB-code: 1FAG[30]; N-palmitoylmethionine and corresponding protein in magenta: 1ZO9[29]). B) Alternative loop conformations are observed in CYP119 when different ligands are bound. Compared to the apo structure of CYP119 (F/G loop in orange: 1IO7[31]), the F/G loop adapts distinct configurations when 4-phenylimidazole (ligand and loop in magenta: 1F4T[32]) or imidazole (blue: 1F4U[32]) are bound. C) In CYP3A4 significant protein flexibility occurs in the F/G portion of the protein (apo: orange: 1TQN[33]; erythromycin bound: blue: 2J0D[34]) to accommodate erythromycin and part of the F-F' loop becomes disordered. This motion causes the solvent accessible volume of the binding site to significantly increase and can dramatically affect ligand binding. D) CYP3A4 exhibits a protein breathing motion increasing the size of the binding pocket to accommodate two ketoconazole (ligands in magenta, protein in blue: 2V0M[34]) molecules without significant conformational changes of the helices or loop regions composing the binding pocket (apo: orange: 1TQN[33]).

Figure 2.

Examples of ligand conformations consistent with SOM data (shown as green sticks) only found when using ensemble docking. Static X-ray docking (binding site amino acids shown as orange sticks) produces no bioactive poses. A) The ensemble docking protein template (shown in cyan lines) has a heavy atom side chain RMSD of 1.6Å RMSD compared to the X-ray structure. Labeled amino acids show significant RMSD differences compared to the X-ray structure leading to the docking of carvedilol. T301 side chain flip is necessary to form additional hydrogen bonds to the ligand stabilizing the bioactive conformation (hydrogen bonds shown in purple). B) The ensemble docking protein template (shown in cyan lines) has a heavy atom side chain RMSD of 1.8Å RMSD compared to the X-ray structure. Labeled amino acids show significant side chain flexibility critical to the docking of SCH-351125. A significant conformational change of F476 provides additional space for the ligand to fit into the active site and forms stabilizing hydrophobic and pi-pi interactions with the bromo-phenyl ring of SCH-351125. In both panels, the experimentally known SOM of the ligand and the heme iron atom are shown as spheres.

Figure 3.

Bioactive poses of four good CYP2C9 substrates and their critical interactions. The co-crystallized flurbiprofen is shown in magenta stick representation, critical protein residues and the heme are shown as green sticks, the bioactive pose of the ligand is shown as cyan sticks, and the ligand SOM is illustrated with a sphere. A) The flurbiprofen bioactive poses has an RMSD of 1.28Å (not considering the F atom) compared to the co-crystallized structure and has critical contacts with Arg108, Asn204, and Phe114. B) Tienilic acid is oriented in the same region as the co-crystallized flurbiprofen and shares similar critical contacts with residues Arg108, Asn204, and Phe114. C) Fluvastatin is oriented in a different region of the binding site than the co-crystallized flurbiprofen. The increased size of fluvastatin leads to the carboxylic group being placed in a region of the binding site closer to the surface of the protein where it does not interact with the charged residues Arg108 or Asn204. D) Losartan is oriented in the same region of the binding site as the co-crystallized flurbiprofen molecule but forms different types of interactions with the enzyme. The tetrazole ring forms a pi-cation and a strong electrostatic interaction with positively charged Arg108 and is further stabilized by pi-pi interactions with Phe100 and Phe114.

Table 1

List of CYP2C9 substrates used in this study. Each substrate was classified as low K_m (1), medium K_m (2), medium-high K_m (3), or high K_m (4) based on published K_m values. The rank where each method correctly identified a known SOM is shown for each substrate.

Substrate	Rank					K _m Class ^c
	NAT Alone	Static Docking		Ensemble Docking		
		Vina Score ^a	Vina+NAT+ SASA ^b	Vina Score ^d	Vina+NAT+SASA ^b	
Flurbiprofen	2	5	1	35	1	1
Tienilic acid	2	6	2	48	2	1
Fluvastatin	7	12	5	7	5	1
Indomethacin	1	7	1	5	1	1
Desogestrel	3	20	1	77	1	1
Losartan	1	NS	NS	112	1	1
Mefenamic acid	5	11	2	91	2	1
Celecoxib	3	6	1	11	1	1
17 alpha-ethinylestradiol	3	10	1	13	1	1
Galangin	1	1	1	3	1	1
Meloxicam	2	1	1	3	1	1
Diclofenac	1	10	2	81	2	1
Kaempferide	1	NS	NS	NS	NS	1
DMZ	5	10	2	20	3	1
Warfarin	4	14	6	34	4	1
Lornoxicam	2	13	1	24	1	1
Fluoxetine	1	3	1	1	1	1
Phenytoin	1	2	1	1	1	1
Phenprocoumon	5	23	3	42	4	1
Delta-9-tetrahydrocannabinol	2	1	1	1	1	1
alitretinoin (9-cis-retinoic acid)	11	3	4	3	7	1
delta-(9)-tetrahydrocannabinol	2	1	1	1	1	1
SCH 351125	5	NS	NS	NS	5	1

Substrate	Rank						K _m Class ^c
	NAT Alone	Static Docking		Ensemble Docking			
		Vina Score ^a	Vina+NAT+ SASA ^b	Vina Score ^a	Vina+NAT+SASA ^b		
Dibenzo[a,h]anthracene	1	3	1	9	1	2	
Sildenafil	2	32	1	33	1	2	
Acenocoumarol	6	9	3	13	4	2	
Clozapine	1	26	1	38	1	2	
Terbinafine	1	17	1	11	1	2	
Tenoxicam	2	15	2	51	2	2	
(s)-lansoprazole	4	1	3	1	2	2	
Seratroast	1	1	1	12	1	2	
Suprofen	2	1	1	23	1	2	
Torsemide	6	54	3	17	2	2	
TR-14035	3	NS	NS	3	1	2	
2-oxoquazepam	1	NS	NS	NS	NS	2	
rosiglitazone	1	44	1	90	1	2	
sulfinpyrazone sulfide	2	1	1	1	1	2	
trimethadione	1	1	1	12	1	2	
flunitrazepam	5	22	4	8	1	2	
muraglitazar	5	36	2	20	3	2	
perazine	1	81	1	133	1	2	
propofol	3	NS	NS	109	3	2	
progesterone	5	NS	NS	63	4	2	
Amitriptyline	1	7	1	9	1	3	
Tolterodine	15	4	9	3	10	3	
Naproxen	1	87	1	214	1	4	
Zolpidem	5	7	3	41	3	4	
Tolbutamide	2	17	2	17	2	4	
Piroxicam	2	15	2	32	1	4	
Chlorpropanide	4	NS	NS	NS	6	4	

Substrate	Rank						K _m Class ^c
	NAT Alone	Static Docking		Ensemble Docking			
		Vina Score ^a	Vina+NAT+ SASA ^b	Vina Score ^d	Vina+NAT+SASA ^b		
Methyl Eugenol	5	13	5	66	5	4	
N-dehydroxyzileuton	2	1	1	16	1	4	
Valproic acid	6	2	3	1	6	4	
Limonene	3	15	2	34	4	4	
Safrrole	4	9	4	48	4	4	
Ketamine	1	11	1	22	1	4	
GV150526	1	1	1	5	1	NA	
Mestranol	1	12	1	26	1	NA	
Etodolac	6	79	4	108	6	NA	
Acetclufenac	2	14	3	53	3	NA	
Methoxyychlor	1	6	1	8	1	NA	
Ibuprofen	8	1	3	21	5	NA	
Arachidonic acid	6	NS	NS	NS	NS	NA	
Zafirlukast	8	11	2	15	2	NA	
58C80	11	2	3	2	3	NA	
Linoleic acid	4	3	2	77	2	NA	
Hydromorphone	1	4	1	4	1	NA	
Carveditol	3	NS	NS	99	3	NA	
Montelukast	23	2	8	1	8	NA	
Hexobarbital	2	4	1	25	1	NA	
Rosuvastatin	1	NS	NS	NS	5	NA	
Tamoxifen	6	1	3	1	5	NA	
Zopiclone	1	26	1	29	1	NA	
Phenobarbital	1	1	1	9	1	NA	
s-MTPPA	3	16	3	43	4	NA	

^a Internal AutoDock Vina score alone used to predict SOM

^b Combination of AutoDock Vina score, NAT model, and SASA term used to predict SOM

^c K_m classification: 1-good (1–10 μM), 2-moderate (10–60 μM), 3-moderate-poor (60–100 μM), 4-poor (>100 μM)

NA: No K_m value available

NS: No bioactive docking pose generated

Table 2

Comparison of techniques used throughout this publication to predict SOM for 75 CYP2C9 substrates. Values shown represent the percentage of substrates with a known SOM correctly predicted among the top-1, top-2 and top-3 predicted SOM.

	Random Pick	NAT alone	Static X-Ray Docking		Ensemble Docking	
			Vina ^a	Vina+NAT+SASA ^b	Vina ^a	Vina+NAT+SASA ^b
Top-1	12%	33%	19%	44%	12%	49%
Top-2	23%	53%	24%	59%	13%	61%
Top-3	34%	64%	29%	73%	20%	71%
Successfully docked^c			85%		96%	

^a Internal AutoDock Vina score alone used to predict SOM

^b Combination of AutoDock Vina score, NAT model, and SASA term used to predict SOM

^c Percentage of ligands docked by AutoDock Vina in at least one bioactive conformation that is consistent with the experimentally observed SOM

Table 3

Classifying the ligand library according to K_m values, the ability of the NAT model alone to predict a known SOM within the top-1, top-2, and top-3 positions is compared to static and ensemble docking with the Vina +NAT+SASA scoring function.

K_m Class	Predicted Position	NAT alone	Static X-Ray Vina+NAT+SASA	Ensemble X-Ray Vina+NAT+SASA
Low	1	30%	52%	57%
	2	57%	70%	70%
	3	70%	74%	74%
Medium	1	40%	50%	60%
	2	60%	60%	75%
	3	70%	75%	85%
High	1	23%	31%	39%
	2	46%	54%	46%
	3	54%	69%	54%
Cumulative statistics^a	1	32%	44%	49%
	2	55%	59%	61%
	3	66%	73%	71%

^aThe 19 substrates lacking kinetic data in CYP2C9 were removed to produce modified cumulative library statistics.