COMPARISON OF METHODS FOR THE PREDICTION OF THE METABOLIC SITES FOR CYP3A4-MEDIATED METABOLIC REACTIONS

Diansong Zhou, Lovisa Afzelius, Scott W. Grimm, Tommy B. Andersson, Randy J. Zauhar, and Ismail Zamora

Department of Drug Metabolism and Pharmacokinetics, AstraZeneca Pharmaceuticals, Wilmington, Delaware (D.Z., S.W.G.); Department of Chemistry & Biochemistry, the University of Sciences in Philadelphia, Philadelphia, Pennsylvania (D.Z., R.J.Z.); Department of Drug Metabolism and Pharmacokinetics & Bioanalytical Chemistry, AstraZeneca R&D, Molndal, Sweden (L.A., T.B.A.); Division of Molecular Toxicology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden (T.B.A.); Lead Molecular Design, Barcelona, Spain (I.Z.); and GRIB-IMIM, Barcelona, Spain (I.Z.)

ABSTRACT:
Predictions of the metabolic sites for new chemical entities, synthesized or only virtual, are important in the early phase of drug discovery to guide chemistry efforts in the synthesis of new compounds with reduced metabolic liability. This information can now be obtained from in silico predictions, and therefore, a thorough and unbiased evaluation of the computational techniques available is needed. Several computational methods to predict the metabolic hot spots are emerging. In this study, metabolite identification using MetaSite and a docking methodology, GLUE, were compared. Moreover, the published CYP3A4 crystal structure and computed CYP3A4 homology models were compared for their usefulness in predicting metabolic sites. A total of 227 known CYP3A4 substrates reported to have one or more metabolites adding up to 325 metabolic pathways were analyzed. Distance-based fingerprints and four-point pharmacophore derived from GRID molecular interaction fields were used to characterize the substrate and protein in MetaSite and the docking methodology, respectively. The CYP3A4 crystal structure and homology model with the reactivity factor enabled achieved a similar prediction success (78%) using the MetaSite method. The docking method had a relatively lower prediction success (~57% for the homology model), although it still may provide useful insights for interactions between ligand and protein, especially for uncommon reactions. The MetaSite methodology is automated, rapid, and has relatively accurate predictions compared with the docking methodology used in this study.

The cytochromes P450 (P450s) belong to a superfamily of heme-containing enzymes that metabolize a wide range of therapeutic agents. CYP3A4 is the most abundant human hepatic P450 isofrom and is responsible for the metabolism of about 50% of known drugs with wide structural diversity (Guengerich, 1999). From a drug development perspective, a new drug candidate that is susceptible to CYP3A4 metabolism may not be able to reach the target at an effective concentration or may be involved in drug-drug interactions when coadministered with other CYP3A4 substrates or inhibitors. In silico prediction of the primary metabolic site from the molecular structure may assist in the identification of the metabolite formed through the metabolic reaction, or could be used to guide the synthesis of metabolically stable compounds or direct the metabolism to certain P450 enzymes. Experimental studies at early drug discovery stage on metabolic reactions of new chemical entities can usually allocate the metabolic site(s) to a certain fragment of the molecule. To guide chemists to synthesize new compounds with improved metabolic properties, information on the exact metabolic site is preferred. This information may be obtained from in silico predictions, and therefore, we need a thorough and unbiased evaluation of the computational techniques available.

Mainly four approaches are considered in the literature to address the prediction of the metabolic sites.

1. Molecular orbital calculations (De Groot et al., 1999; Singh et al., 2003) can be used to estimate the energy necessary to abstract a hydrogen atom from the substrate. The energy for each possible metabolic site in combination with surface area exposure of the hydrogen atom was used to rank the different sites for the CYP3A4 metabolic liability (Singh et al., 2003). However, this approach does not consider the specific ligand-enzyme interactions and therefore overlooks the specificity of P450 enzymes. Hydrogen abstraction is the only metabolic mechanism considered in this method.

2. Molecular docking in combination with scoring functions has been applied to predict the metabolic sites for CYP2D6 substrates like codeine and 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (Kirton et al., 2002). Although this method theoretically optimizes favorable interactions between ligand and enzyme to estimate the binding mode, predictions heavily rely on a well defined pharmacophore of the active site. Therefore, the availability of a validated crystal structure or homology model of the P450 enzyme is a prerequisite for reliable predictions. Moreover, there are other factors such as protein flexi-
bility and water molecule interactions, which could make these methods difficult to use.

3. Data mining, a probabilistic scoring method based on fragment analysis of the substrate structures corresponding to the fragments observed in databases, recently has also yielded useful results in prediction of metabolic site (Hasselgren-Armbjörnsson et al., 2005). In this case, no guidance on major metabolic pathways involved is needed, but this also limits interpretation.

4. GRID molecular interaction fields (MIFs) (Goodford, 1985) have been used to derive parameters that describe the ligand-protein interactions, and this methodology has been implemented in MetaSite. In MetaSite, the active site of the protein is characterized using MIF interactions with different probes, and a set of distance-based descriptors is calculated by taking the oxygen atom in the heme as reference. The different atoms of the substrate are ranked in agreement with the similarity to the protein, and the top ranked positions are selected as the possible metabolic sites (Zamora et al., 2003; Cruciani et al., 2005). A well-defined active site is also essential for this methodology. However, when this methodology was first developed, the human P450 crystal structures were not publicly available and predictions were derived using homology models based on alignments of bacterial P450 structural models. The recently available wild-type CYP3A4 crystal structures (Williams et al., 2004; Yano et al., 2004) provide the opportunity to improve this methodology.

The aim of the present work was to analyze the use of the available CYP3A4 crystal structures in the prediction of metabolic sites of structurally diverse drugs with known metabolic profiles using both the docking/scoring and the MetaSite methodologies, and comparing the results with the ones obtained from P450 homology models. It is not within the scope of this study to predict the metabolic rate, or to differentiate the major or minor metabolic pathways. Other information, such as empirically determined data ($K_a$ and/or $V_{max}$) would probably be needed to develop such a model.

**Materials and Methods**

**Equipment, Software, and Database.** All calculations were done on the Linux (Red Hat 8.0) operating system on a 1.8 GHz Pentium IV. The programs used in the molecular modeling were GRID, GLUE, PENGUINS, MetaSite (Molecular Discovery Ltd., Middlesex, UK), GOLPE (Multivariate Infometric Analysis S.r.l., Perugia, Italy), and CONFORT (Tripos Inc., St. Louis, MO). Chemical structures were imported from the ISIS-BASE database or drawn using ISIS-Draw (MDL Information Systems Inc., San Leandro, CA).

**Proteins and Substrates.** A CYP3A4 crystal structure with a resolution of 2.8 Å was used in this study (PDB 1W0E). This is a wild-type enzyme, except that the N-terminal membrane insertion peptide has been removed to increase solubility for crystallization. There was no substrate or inhibitor bound in the active site of this crystal structure. A CYP3A4 homology model was generated by comparative modeling based on multiple bacterial P450s (PDB codes are 2bmh, 3cpp, 1cpt, and 1oxa) using Modeller (De Rienzo et al., 2000).

A total of 227 CYP3A4 substrates were collected from the literature (Rendic and Di Carlo, 1997) and the MDL metabolite database. Most of the well known CYP3A4 substrates, such as midazolam, nefilipine, and testosterone, were included in this data set. The substrates are reported to have one or more CYP3A4-mediated metabolites, adding up to 325 metabolic pathways. Most of the compounds (165) have only one metabolite catalyzed by CYP3A4, whereas 62 compounds have two or more metabolites. The two-dimensional chemical structures of the 227 compounds were drawn or imported from several databases and converted to three-dimensional structures using PENGUINS. A maximum of 50 diverse conformers within an energy window of 10 kcal/mol were generated for each substrate using CONFORT.

**Principal Component Analysis (PCA).** PCA (Pastor and Cruciani, 1995) was used to compare the active sites of the CYP3A4 crystal structure, the homology model based on four cytochrome bacterial structures, and the four proteins used to build the homology model. The PDB codes for these bacterial crystal structures are 2bnh, 3cpp, 1cpt, and 1oxa for CYP2E1, CYP2C9, and CYP2C19, respectively. The molecular interaction fields generated in GRID force field were used to characterize the active sites of all proteins. A GRID box (25 × 25 × 25 Å) was used to include the superimposed active sites of all three-dimensional structures. In our model, the heme was considered as part of the protein and was not treated differently except where the ferryl oxygen was defined as a dummy atom that is unable to form hydrogen bonds with any probe. Similar treatment was applied in the later calculations. The active site of each structure was characterized using 10 GRID probes (DRY, C3, N1, NH, N=, O, NMb, O", O", and OH) in flexible mode (MOVE = 1) with a grid step size of 1 Å. Hydrophobic interactions were calculated with the DRY probe, whereas the steric interactions were calculated with C3 and NMb probes. N1", O" and O" probes are charged, and N1, NH=, O, and OH probes are polar.

GRID force field calculations were then imported into GOLPE, and the following pretreatment was done before PCA: 1) the maximum cutoff was set to zero to consider only the favorable interactions (negative energy values); 2) block unscaled weights were used to normalize the interaction energies between the different probes; and 3) variables with values smaller than 0.01 kcal/mol and those with a standard deviation below 0.02 kcal/mol were removed to increase signal to noise ratio. PCA provided loading and score plots with insight into the topology and chemical identity difference between the structures.

**MetaSite Methodology Using GRID Descriptors.** MetaSite was first developed and applied to CYP2C9 and its substrates (Zamora et al., 2003). This methodology is improved by the inclusion of a reactivity factor in the current study and is used to characterize both the CYP3A4 crystal structure and the homology model. Both the protein active site and the ligand are represented by selected distance-based descriptors using the molecular interaction fields computed by GRID. The best match between active site and metabolic site is chosen based on similarity and, optionally, with atom reactivity (Fig. 1).

**Protein Treatment.** The MIFs in the active site of CYP3A4 were generated using four probe types: hydrophobicity (DRY probe), hydrogen-bond donor (amide nitrogen N1 probe), hydrogen-bond acceptor (carbonyl oxygen O probe), and electrostatic property (positively charged N1", N2", N1" probes and negatively charged O", COO", and N" probes) with a grid step size of
 descriptors were compared using the Carbo´ similarity index (Amat and Carbo´-a distance-based fingerprint for each possible metabolic site, both sets of interaction points in the protein and the structure of substrate was described as metabolic site in a substrate are generated.

the different preclassified atoms were computed and transformed into grouped between the possible metabolic sites, such as hydrogen or nitrogen atoms, and capabilities using the Tripos force-field atom type definitions. The distances hydrogen-bond donor, hydrogen-bond acceptor, and electrostatic interaction were then classified into four categories according to their hydrophobic, hydrogen-bond donor/acceptor and electrostatic interaction capabilities and geometry. When a pharmacophore was recognized, the ligand was aligned in the enzyme cavity and an energy computation followed. If there were any conflict contacts between the ligand and the protein, an induced fit process was started to accommodate the substrate in the protein cavity. The same process was repeated for all possible four-point pharmacophore templates.

The metabolic potential of each atom in the substrate was estimated using (1) a probability function based on the distance between each atom and the reactive center of the protein, and (2) a probability function based on the energy of interaction as computed by GLUE. Distance-based probability was calculated considering a Gaussian distribution for the difference in the distance between each substrate atom and the fixed ferryl oxygen (2 Å above heme) at the active center of the enzyme, with the optimal distance found in some crystallographic structures, 2.6 Å (Schlichting et al., 2000).

The energy probability was calculated based on the Boltzmann distribution for each docking solution. The product of distance probability for each atom in each docking solution and energy probability for each docking solution yield a value for each atom that was used to rank the probable metabolic site. The predicted metabolic site was identified using the ranking position for each atom evaluated from the best docking conformer (lowest energy) and from an ensemble of all docking solutions.

Results

PCA. PCA was used to evaluate the topological difference of the active sites of the CYP3A4 crystal structure, homology model, and active site of CYP3A4 using the homology model and the available crystal structure (PDB 1W0E), with and without crystallographic water molecules. GLUE, a GRID-based docking program, was used to analyze the ligand-receptor interaction and to perform the docking experiments.

In GLUE, the active sites were mapped using hydrophobic, hydrogen-bond donor/acceptor and electrostatic probes. All possible tetrahedra obtained from four minimal energy points from GRID are computed. These four-point pharmacophores derived from the interaction for the active sites were then used as templates to compare with the ligand.

Similarly, GLUE identified the polar and hydrophobic heavy atoms of the ligand and calculated all possible tetrahedra between these atoms. The atomic positions of the different conformers for each potential substrate were compared with the pharmacophores based on the hydrophobic, hydrogen-bond donor/acceptor and electrostatic interaction capabilities and geometry. When a pharmacophore was recognized, the ligand was aligned in the enzyme cavity and an energy computation followed. If there were any conflict contacts between the ligand and the protein, an induced fit process was started to accommodate the substrate in the protein cavity. The same process was repeated for all possible four-point pharmacophore templates.

29 crystallographic water molecules have been determined in the CYP3A4 crystal structure used in this study. In the process of P450-mediated metabolism, one water molecule is usually generated (Guenigerich, 1999). Water molecules may also play an important role by forming certain hydrogen bonds to hold the substrate with proper orientation toward the heme (Wester et al., 2003). Therefore, CYP3A4 crystal structures with or without these crystallographic water molecules were considered separately. Overall, six CYP3A4 protein structures or models were explored in this study: 1) homology model in GRID rigid mode, 2) homology model in GRID flexible mode, 3) crystal structure with water in GRID rigid mode, 4) crystal structure with water in GRID flexible mode, 5) crystal structure without water in GRID rigid mode, and 6) crystal structure without water in GRID flexible mode.

The following MIF treatment was similar to the one previously published (Zamora et al., 2003). In brief, the regions close to the binding site, but not accessible to the substrates, were removed from the analysis in an automatic cut-out process. Finally the distances were calculated between the selected GRID points and the fixed ferryl oxygen at the reactive center of the enzyme. The distances were grouped at a resolution of 0.8 Å and plotted as correlograms, which were compared with the distance-based descriptor of the substrate.

Substrate Treatment. The compounds were built or imported as two-dimensional structures. Substrate conformation sampling is critical to simulate flexible interaction between the substrate and CYP3A4 enzyme; therefore, conformation search followed by energy minimization was performed in CONFORT for each CYP3A4 substrate. The atoms of a CYP3A4 substrate were then classified into four categories according to their hydrophobic, hydrogen-bond donor, hydrogen-bond acceptor, and electrostatic interaction capabilities using the Tripod force-field atom type definitions. The distances between the possible metabolic sites, such as hydrogen or nitrogen atoms, and the different preclassified atoms were computed and transformed into grouped variables with a resolution of 0.8 Å. Four sets of fingerprints for each possible metabolic site in a substrate are generated.

Substrate-Protein Comparison. Once the protein interaction profiles were transformed into distances from the active site of the protein and the structure of substrate was described as a distance-based fingerprint for each possible metabolic site, both sets of descriptors were compared using the Carbó similarity index (Amat and Carbó-Dorch, 1999). Four similarity indexes were obtained for each possible metabolic site in a substrate according to hydrophobic interactions, hydrogen-bond donor/acceptor interactions, and charge-charge interactions. The atomic position with the highest similarity score will be the one that has the best complementarity with the protein and, theoretically, the enzyme will orient the compound with this atom toward the heme of CYP3A4.

Reactivity. In addition to the similarity comparison, a substrate fragment recognition factor called “reactivity” has been implemented in this methodology. A database of different small fragments with precalculated reactivity values has been generated and applied in MetaSite. Each fragment was considered as a participant in oxidative reactions, and a reactivity value was assigned to each atom regarding the liability toward the oxidative reaction. When a fragment in the molecule under study is recognized as one in this database, all atoms in the fragment are assigned to that reactivity value.

The final ranking for potential metabolic site is the product of protein effect (computed on the basis of the similarity analysis) and atomic reactivity effect (computed using the fragment-based approach).

Docking/Scoring Methodology. All the substrates were docked into the active site of CYP3A4 using the homology model and the available crystal structure (PDB 1W0E), with and without crystallographic water molecules. GLUE, a GRID-based docking program, was used to analyze the ligand-receptor interaction and to perform the docking experiments.

In GLUE, the active sites were mapped using hydrophobic, hydrogen-bond donor/acceptor and electrostatic probes. All possible tetrahedra obtained from four minimal energy points from GRID are computed. These four-point pharmacophores derived from the interaction for the active sites were then used as templates to compare with the ligand.

Similarly, GLUE identified the polar and hydrophobic heavy atoms of the ligand and calculated all possible tetrahedra between these atoms. The atomic positions of the different conformers for each potential substrate were compared with the pharmacophores based on the hydrophobic, hydrogen-bond donor/acceptor and electrostatic interaction capabilities and geometry. When a pharmacophore was recognized, the ligand was aligned in the enzyme cavity and an energy computation followed. If there were any conflict contacts between the ligand and the protein, an induced fit process was started to accommodate the substrate in the protein cavity. The same process was repeated for all possible four-point pharmacophore templates.

The metabolic potential of each atom in the substrate was estimated using (1) a probability function based on the distance between each atom and the reactive center of the protein, and (2) a probability function based on the energy of interaction as computed by GLUE. Distance-based probability was calculated considering a Gaussian distribution for the difference in the distance between each substrate atom and the fixed ferryl oxygen (2 Å above heme) at the active center of the enzyme, with the optimal distance found in some crystallographic structures, 2.6 Å (Schlichting et al., 2000).

The energy probability was calculated based on the Boltzmann distribution for each docking solution. The product of distance probability for each atom in each docking solution and energy probability for each docking solution yield a value for each atom that was used to rank the probable metabolic site. The predicted metabolic site was identified using the ranking position for each atom evaluated from the best docking conformer (lowest energy) and from an ensemble of all docking solutions.

PCA. PCA was used to evaluate the topological difference of the active sites of the CYP3A4 crystal structure, homology model, and
four bacterial crystal structures. Crystal structures with water molecules were excluded for this analysis because the corresponding active sites were smaller compared with the other models and would skew the results. The PCA scores plot is shown in Fig. 2, and the relative distance between individual data points along each principal component is an indication of the extent of the difference between the active sites of the CYP3A4 crystal structure, the homology model, and the four bacterial structures. There are four groups of these models: CYP3A4 crystal structure (group 1); CYP3A4 homology model and CYP<sub>terp</sub> (group 2); CYP<sub>cam</sub> and CYP<sub>eryF</sub> (group 3); and CYP<sub>Bm3</sub> (group 4). The CYP3A4 crystal structure differs from the other structures analyzed as indicated by its principal component scores in the scores plot. The calculated root-mean-square deviation between the CYP3A4 crystal structure and the homology model is 5.52 Å, which also reflects the dissimilarity between these two structures. CYP<sub>Bm3</sub> was separated from CYP<sub>cam</sub> and CYP<sub>eryF</sub> by the second principal component. The active site of the CYP3A4 homology model is more similar to that of the CYP<sub>terp</sub> structure, and also more similar.
to the active site of the CYP3A4 crystal structure than to the other three bacterial forms (CYPBM3, CYPcam, and CYPeryF), as indicated by their proximity along the principal component 1 score in the scores plot.

MetaSite Methodology Using GRID Descriptors. The validation of the prediction of the metabolic site was based on counting the number of experimentally reported metabolic pathways that were found among the first, second, and third sites as ranked by the methodology used. A total of 227 reported CYP3A4 substrates with 325 metabolic pathways were studied; 165 compounds had only one reported metabolite, whereas 62 compounds had multiple metabolic sites. When considering the multiple-metabolite substrates, two schemes were used to evaluate the prediction methods: 1) each reaction was considered independently, so that 325 pathways were considered, and 2) substrates with multiple metabolic sites were analyzed separately, and the substrate was classified as well predicted if at least one pathway was among the first-, second-, or third-ranked site proposed by this methodology. Because there are no data available about the rate of formation for each of the metabolites reported, the prediction for one of them was enough to classify the compound.

The predictions for all 325 metabolic pathways using MetaSite methodology and different protein models are shown in Fig. 3, and the prediction success using either the top-ranked site or the top three sites for the different models is presented in Table 1. In the case of considering both protein similarity and atomic reactivity, the crystal structure with the crystallographic water molecules model yielded the lowest prediction success of 58% when using the top three ranking positions and when applied to predict 325 metabolic pathways, whereas all the other models had similar predictive success, an average of 70%. Of these 29 crystallographic water molecules, 15 waters were inside the defined grid box, but only 5 water molecules can be considered as being inside the binding site. Using the crystal structure with these five water molecules included always led to the least successful performance and will not be considered when computing average prediction success. When the reactivity factor was not considered, the average prediction success decreased from 70% to 39%.

The MetaSite methodology yielded an average prediction success of 75% when using the first three ranked positions and when applied to the 165 substrates that had only one metabolite reported (and where both protein complementarity and reactivity were considered). This methodology also achieved an average prediction success of 86% when at least one metabolic site was well predicted by one of the first three ranked positions, and applied to the 62 substrates with multiple sites. The overall prediction success for all compounds when at least one metabolic site was predicted among the top three ranked sites by this methodology was 78% when reactivity was enabled.

For the 325 reactions under study, 43% (139 reactions) were aliphatic or aromatic hydroxylation, 27% (89 reactions) were N-dealkylation, and 7% (22 reactions) were O-dealkylation. These are three major metabolic pathways for CYP3A4 substrates, and they were well predicted using the flexible mode for the homology model or crystal structure with reactivity option enabled (Fig. 4A). Other reaction types such as reductions (6 reactions), epoxidations (5 reactions), and N-hydroxylations (2 reactions) were not well predicted. This could be due to the fact that these reactions are uncommon; therefore, the reactivity factor disfavors these kinds of reactions. In the model with crystal structure, flexible GRID mode and atomic reactivity option enabled, the method predicted correctly (among the

---

**Table 1**

Prediction success (percentage) for all CYP3A4 metabolic pathways found using either the top-ranked site or the top three sites for the different models using MetaSite methodology, the homology model in rigid mode, the homology model in flexible mode, the crystal structure in rigid mode, and the crystal structure in flexible mode.

<table>
<thead>
<tr>
<th>Models</th>
<th>Prediction Success for Top-Ranked Sites</th>
<th>Prediction Success for Top Three Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reactivity On</td>
<td>Reactivity Off</td>
</tr>
<tr>
<td>Crystal flexible</td>
<td>41</td>
<td>20</td>
</tr>
<tr>
<td>Crystal rigid</td>
<td>41</td>
<td>21</td>
</tr>
<tr>
<td>Homology flexible</td>
<td>42</td>
<td>22</td>
</tr>
<tr>
<td>Homology rigid</td>
<td>42</td>
<td>20</td>
</tr>
</tbody>
</table>

---

**Fig. 4.** A, the prediction success corresponding to various types of reactions for all 325 CYP3A4-mediated metabolic pathways by the homology model (gray bar) and the crystal structure (white bar) in flexible mode and reactivity factor enabled using MetaSite methodology. The number in the parentheses is the number of each reaction evaluated in this study. B, the prediction success corresponding to various types of reactions for all CYP3A4-mediated metabolic pathways by the homology model (gray bar) and the crystal structure (white bar) using docking methodology.
top three ranked sites): 70 of 98 aliphatic hydroxylations (71%), 23 of 41 aromatic hydroxylations (56%), 76% of N-dealkylation, and 100% of O-dealkylation.

The prediction success corresponding to various types of reactions for the 165 substrates with only one metabolic site by the homology and crystal models in flexible mode and with reactivity option enabled are shown in Fig. 5. The prediction success for 57 hydroxylated substrates, 53 N-dealkylated substrates, and 12 O-dealkylated substrates was 75%, 85%, and 100%, respectively by using the homology model.

Sixty-two multimetabolite substrates have 160 metabolic pathways, including 82 hydroxylations, 36 N-dealkylations, and 10 O-dealkylations. The prediction success using the crystal structure model was 60%, 67%, and 100% for hydroxylation, N-dealkylation, and O-dealkylation, respectively.

The rigid and flexible GRID modes were evaluated for all substrates. The reactions that were accurately predicted by the CYP3A4 crystal structure without added water molecules in rigid or flexible modes and with reactivity enabled were compared, and no significant difference was observed. The calculated active sites of CYP3A4 crystal structure in its rigid mode and with reactivity enabled could correctly predict the metabolic sites of 183 substrates, whereas in its flexible mode, 173 of a total of 227 substrates were well predicted. The calculated active sites of CYP3A4 crystal structure and the homology model are presented in Fig. 6.

A number of metabolic pathways can be well predicted using this docking approach. For example, 4-hydroxylation is the major pathway reason why the use of the crystal structure in either the flexible or the rigid mode had no clear impact on the accuracy of predictions.

Generally, when the atomic reactivity ranking was not considered, the prediction success decreased for all models. To evaluate the importance of reactivity in the methodology for the metabolic site prediction, the prediction success of the model constructed using the crystal CYP3A4 structure without water in GRID flexible mode was compared either with reactivity enabled or not. When reactivity was enabled, 173 of 227 substrates were well predicted compared with 103 of 227 when reactivity was disabled. Seventeen substrates were only well predicted in the model without the reactivity factor used, and 6 of these 17 substrates are metabolized through uncommon reactions, such as epoxidation, reduction, or dehydrogenation. The model without reactivity tends to have better prediction for uncommon reactions, and reactivity factor may disfavor the rank of uncommon reactions in this method. However, the uncommon reactions are only a small portion of all reactions mediated by CYP3A4, and the models with reactivity factor always yield better overall prediction success than the models without reactivity.

Docking/Scoring Approach. The first observation when analyzing the docking solutions is that not all substrates could be docked into the CYP3A4 active site. Docking ability also depended on the protein model used. Two of 227 substrates could not be docked into the CYP3A4 homology model, whereas 18 substrates could not be docked into the CYP3A4 crystal structure without the crystallographic water molecules, and 78 could not be docked into the CYP3A4 crystal structure when water molecules were included. Therefore, only 323, 298, and 205 metabolic pathways were evaluated for homology, crystal structure without water, and crystal structure with water, respectively.

A number of metabolic pathways can be well predicted using this docking approach. For example, 4-hydroxylation is the major pathway
for CYP3A4-mediated metabolism of alprazolam, whereas 1'-hydroxylation is relatively minor (Williams et al., 2002). The lowest energy docking result showed that the distance between the C4 position of alprazolam and the ferryl oxygen was 3.35 Å, whereas the distance between the C1' position and oxygen was 4.47 Å, and both metabolites are possible.

Multiple docking solutions were available for most of the CYP3A4 substrates. Similar to the MetaSite procedure, the evaluation of the docking results was based on the number of solutions that predicted correctly the metabolic site in the first-, second-, and third-ranked positions.

Two kinds of analysis were performed depending upon whether the best docking solution or all docking solutions were included. In the case of considering the best-docking (lowest-energy) solution for each substrate, only 17% to 27% of the metabolic reactions were correctly predicted as possible metabolic sites among the first three selections (Table 2). When all docking solutions are analyzed, it is considered a successful prediction if at least one of the docking solutions exhibits correct orientation. In this latter case, the homology model achieved the best prediction (57%), whereas the docking based on the crystal structure without water molecules yielded 47% prediction success, and the structure with water molecules gave the lowest success (27%). Again, the crystal structure with the water molecule model will not be considered in the following analysis.

When the substrates that only had one metabolite were analyzed, the prediction success was 63% and 53% for the CYP3A4 homology model and crystal structure when all docking solutions were considered. In the case of multiple-metabolite substrates, the prediction success increased to 82% and 74%, respectively, when at least one metabolite was correctly predicted.

The prediction success corresponding to the different reaction types found by docking into the CYP3A4 homology model and the crystal structure are also presented in Table 2. Three major metabolic pathways were analyzed: hydroxylation yielded a prediction success of 52% for the crystal structure-based model and 58% for the homology based one. O-dealkylation reactions gave a prediction success of 38% and 54% for the crystal- and homology-based docking, respectively, and finally, the O-dealkylation reactions yield 62% for the crystal structure docking and 73% for the homology docking. Some reactions such as reductions (two of three reactions), epoxidations (three of three reactions), and N-hydroxylations (one of two reactions), which were not well predicted with MetaSite methodology, could be reasonably predicted with the docking approach using the CYP3A4 homology model.

### Discussion

CYP3A4 plays a major role in drug metabolism in humans, and understanding the CYP3A4-mediated metabolism is especially relevant in drug discovery. Two computational techniques, the MetaSite and docking methodologies, for the prediction of the metabolic sites have been successfully applied and evaluated for a set of CYP3A4-mediated substrates using both the CYP3A4 crystal structure and a homology model. The predicted metabolic sites using the docking methodology in conjunction with the CYP3A4 homology model matched, in 69% of the cases, with the experimentally reported metabolic sites. Both the CYP3A4 crystal structures and the homology models achieved similar prediction success by using the MetaSite method; i.e., the predicted metabolic sites agreed with experimental results in 78% of the cases. The MetaSite methodology correctly predicted 1'- and 4-hydroxymidazolam, 6β-hydroxytestosterone, and dehydronifedipine in its top three rankings, whereas the docking methodology also identified metabolic sites of midazolam and testosterone but failed to predict dehydronifedipine. In addition, the MetaSite methodology can be considered as an automated method after the active site properties have been mapped. Another disadvantage of the docking method is that some compounds could not be docked into the enzyme active site. This might be because of deficiencies of the docking algorithm or the flexibility of the ligands. Overall, the MetaSite methodology can be used as a metabolic site prediction tool at the early drug discovery stage in terms of both speed and accuracy.

One aspect to be considered in the evaluation of any method for predicting the metabolic site is the approach used to establish an objective criterion to measure its predictive power. In this study, the number of metabolic reactions that are well predicted considering the first-, second-, and third-ranked atomic positions was used as a measurement of the predictive capabilities of the method. Nevertheless, this prediction success will depend on the number of metabolic pathways reported for one substrate. When each reaction is considered independently of the substrate, the ranking position could underestimate the prediction power. For example, a substrate with four metabolic sites would always have one site that would be misclassified, since we considered only the first three ranked positions. Therefore, the first three ranked positions for each reaction independent of substrate and for at least one reaction reported for a substrate were both used as a predictive measurement in this study. The first measurement represents the lower predictive power limit and the second one gives the higher predictive power limit of each methodology.

Both methodologies in this study evaluate the interactions between protein and ligands; therefore, the protein model is critical in deter-
mining the quality of predictions. The wild-type CYP3A4 crystal structure (PDB code: 1W0E) without bound ligand, used in this study, is very similar to a CYP3A4 crystal structure published by another group (PDB code: IQTN). Both structures have a cluster of phenylalanine residues that lies above the active site, which makes a relatively small active site for CYP3A4. CYP3A4 sometimes displays cooperative behavior with the binding of substrates (Domanski et al., 2001; Tang and Stearns, 2001; Galetin et al., 2003; He et al., 2003), and this is generally rationalized by a flexible and large CYP3A4 active site, which can accommodate multiple substrates. However, published CYP3A4 crystal structures, in contrast to CYP2B6 (Scott et al., 2004), have little conformational change in the ligand-free and ligand-bound forms. One possible reason might be that the ligand in the crystal is relatively small and does not cause dramatic conformational change of CYP3A4. Large ligands, when used in crystallization, either stay outside of the active site pocket (progestrone) or have yet to be co-crystallized with the protein (erythromycin) (Williams et al., 2004; Yano et al., 2004). All information suggests that the current CYP3A4 crystal structure might be only one of many available conformations for the enzyme. This might be the reason that the CYP3A4 crystal structure did not exhibit any advantage over the homology model in predicting the metabolic site. With the MetaSite methodology, this could also be related to the protein treatment, in which the interaction profile was compressed to a distance-based descriptor and the impact of the protein/substrate complementarity decreased. Because the docking technique depends much more on the protein structure than does MetaSite methodology, the difference of the prediction success from docking between the homology model and the crystal structure was larger (~10%) in the prediction using the top three ranking positions.

In an effort to model the flexibility of CYP3A4, we tried to take advantage of a special mode of GRID field calculation (MOVE = 1), also termed the flexible mode, in this study. Theoretically, the protein is more flexible and can accommodate different substrates; therefore, the flexible mode model could provide more accurate predictions. However, in contrast to the big impact on the prediction success for CYP2C9 substrates (Zamora et al., 2003), there is no clear effect of this mode for the CYP3A4 substrates. One possible reason might be that the GRID flexible mode can only capture the side chain movement of certain amino acids, but not the major conformational changes, such as helix movement. More extensive calculations, perhaps involving backbone motion, might be required to simulate the flexibility of the protein and the interaction between CYP3A4 and the ligands.

The water molecules can play a significant role in ligand-protein binding (Wester et al., 2003), but it is still challenging to include water during the automated docking process. The docking studies have shown that including the fixed water molecules can either increase or decrease the docking accuracy (Osterberg et al., 2002; De Graaf et al., 2005). We have shown that, even if only five water molecules are inside the defined binding site, the models including these fixed waters generated the lowest prediction success in both methodologies. Water molecules may mediate the substrate-enzyme interaction differently compared with the ones obtained from the crystal structures. The positions of water molecules might depend on the specific substrate in the active site. The use of fixed positions for the water molecules obtained from a crystal structure would therefore not represent all the possibilities. As in this study, it is not practical to specify the water molecule position(s) in the active site for each individual substrate when attempting to examine a large number of compounds. The structures excluding these waters would thus be better models in the metabolic sites prediction.

In summary, this study has presented the docking and the MetaSite methodologies to successfully predict the metabolic site for CYP3A4 substrates under various conditions. Each method has its own advantages, and with proper application, alone or in combination with each other, these methods should be of great help in identifying the metabolic sites, elucidating metabolite structures, and guiding chemical programs to synthesize compounds with improved metabolic properties.

References