Exploration of Enzyme−Ligand Interactions in CYP2D6 & 3A4 Homology Models and Crystal Structures Using a Novel Computational Approach

Britta Kjellander,*[,†,‡] Collen M. Masimirembwa,[†,§] and Ismael Zamora[||,⊥]

Discovery DMPK & Bioanalytical Chemisty Department, AstraZeneca R&D Mölndal, SE-431 81, Sweden, Department of Chemistry, Medicinal Chemistry, Göteborg University, SE-412 96 Gothenburg, Sweden, African Institute of Biomedical Science & Technology (AiBST), Harare, Zimbabwe, Lead Molecular Design, S.L., Vallés 96-102 (27), E-08190, San Cugat del Vallés, Spain, and Institution Municipal d'Investigació Medica (IMIM), Universitat Pompeu Fabra, Doctor Aiguader 80, 08003 Barcelona, Spain

Received December 20, 2006

New crystal structures of human CYP2D6 and CYP3A4 have recently been reported, and in this study, we wanted to compare them with previously used homology models with respect to predictions of site of metabolism and ligand−enzyme interactions. The data set consisted of a family of synthetic opioid analgesics with the aim to cover both CYP2D6 and CYP3A4, as most of these compounds are metabolized by both isoforms. The program MetaSite was used for the site of metabolism predictions, and the results were validated by experimental assessment of the major metabolites formed with recombinant CYP450s. This was made on a selection of 14 compounds in the data set. The prediction rates for MetaSite were 79−100% except for the CYP3A4 homology model, which picked the correct site in half of the cases. Despite differences in orientation of some important amino acids in the active sites, the MetaSite-predicted sites were the same for the different structures, with the exception of the CYP3A4 homology model. Further exploration of interactions with ligands was done by docking substrates/inhibitors in the different structures with the docking program GLUE. To address the challenge in interpreting patterns of enzyme−ligand interactions for the large number of different docking poses, a new computational tool to handle the results from the dockings was developed, in which the output highlights the relative importance of amino acids in CYP450-substrate/inhibitor interactions. The method is based on calculations of the interaction energies for each pose with the surrounding amino acids. For the CYP3A4 structures, this method was compared with consensus principal component analysis (CPCA), a commonly used method for structural comparison to evaluate the usefulness of the new method. The results from the two methods were comparable with each other, and the highlighted amino acids resemble those that were identified to have a different orientation in the compared structures. The new method has clear advantages over CPCA in that it is far simpler to interpret and there is no need for protein alignment. The methodology enables structural comparison but also gives insights on important amino acid substrate/inhibitor interactions and can therefore be very useful when suggesting modifications of new chemical entities to improve their metabolic profiles.

INTRODUCTION

Since the demonstration that cytochrome P450s play a major role in the metabolism of therapeutics, numerous in vitro and computational methods have been developed to predict metabolism issues associated with this family of enzymes. It is believed that these predictive tools will assist in the discovery of new chemical entities (NCEs) with favorable pharmacokinetic and toxicological profiles. Molecular understanding of the cytochrome P450s started with the successful cloning and heterologous expression of most human CYP isoforms important for drug metabolism in the 1990s. Mutagenesis studies started to yield important structure−function properties of the enzymes with respect to determinants of substrate/inhibitor selectivity and the site of metabolism. Successful crystallization of bacterial soluble CYP450s in the 1980s was a significant development in the quest to understand CYP450-inhibitor/substrate interactions. It spurred an enormous effort in homology modeling of human CYP450s using the bacterial crystal structures as templates. Most work was done on CYP450s in the 1, 2, and 3 families since they have been shown to be the main drug-metabolizing isoforms. Crystallization of mammalian membrane-bound CYP450s met with extreme difficulties until the 2000s.

Homology modeling, substrate/inhibitor docking into homology models, and pharmacophore and quantitative structure−activity modeling software were extensively applied toward predicting CYP450-mediated metabolism.[1,2] The computational modeling approaches have provided important insights to questions metabolism scientists want to answer: (a) predicting amino acids important for CYP-substrate/inhibitor interactions; (b) pharmacophore models of substrates and/or inhibitors of CYP450s; (c) prediction of the site of metabolism using a combination of homology

* Corresponding author tel.: +46 31 706 4262; fax: +46 31 776 3787; e-mail: Britta.kjellander@astrazeneca.com.
† AstraZeneca R&D Mölndal.
‡ Göteborg University.
§ AiBST.
|| Lead Molecular Design.
⊥ IMIM.

10.1021/ci600561v CCC: $37.00
Published on Web 03/24/2007

modeling, docking strategies, and pharmacophoric features of substrates of specific CYPs; (d) predicting the inhibitory potency of compounds on CYP450s; and (e) defining CYP450 selectivity for substrates and inhibitors.

The ability to make the above predictions is believed to enhance the capability of DMPK and medicinal chemistry scientists in designing NCEs and optimizing the structure of a potential NCE: (i) to improve metabolic stability, (ii) to avoid pathways leading to the production of potentially toxic metabolites, (iii) to abolish or reduce the inhibitory potential of compounds that could lead to drug−drug interactions, and (iv) to switch the metabolism of the compound from one isoform to another if the enzyme involved is polymorphic and the NCE has a narrow therapeutic index.

Despite the success scored with the current computational methods, many challenges and uncertainties have remained unresolved. Many scientists continued to believe that crystallization of the human CYP450s with and without ligands (substrates and inhibitors) would provide the ultimate answers. The crystallization of the first mammalian membrane-bound CYP450, a rabbit CYP2C5,[3] in 2000 was followed up by rapid success in crystallizing human CYP450s, 2C9,[4,5] 2C8,[6] 3A4,[7−9] and 2D6.[10] We are therefore at that crossroad situation where there is a need to compare the homology models of these human CYPs that have been used over the years with the currently available crystal structures.

In this study, we have compared homology models and crystal structures for CYP2D6 and CYP3A4 with respect to the prediction of the site of metabolism. The major metabolic routes of a set of opioid analgesics were experimentally determined, and the sites of metabolism were predicted using the program MetaSite.[11,12] We further wanted to explore the patterns of ligand−enzyme interactions in the different structures using the docking of enzyme substrates and inhibitors. In order to address the challenge in interpreting patterns of enzyme−ligand interactions for a large number of different docking poses, a computational tool to handle the results from the dockings was developed, which highlights the relative importance of amino acids in CYP450-substrate/inhibitor interactions. This analysis allows the comparison between the homology models and the crystal structures avoiding the complex, and usually subjective, step of structure alignment. The methodology also gives insights on important amino acid substrate/inhibitor interactions useful toward modifications of NCEs to improve their metabolic profiles.

MATERIALS

Computer Hardware and Software. The calculations were performed in a Linux environment on a HP workstation xw6200 with a Pentium IV processor. The software used was GRID version 22.2.2 and MetaSite version 2.7.5 (Molecular Discovery, http://moldiscovery.com), GOLPE version 4.5.12 (MIA, http://www.miasrl.com), and SYBYL version 7.1 (Tripos Associates Inc., St. Louis, http://www.tripos.com). 2D to 3D conversion was made with CORINA (Molecular Networks GmbH, http://www.mol-net.de). The energy calculations were performed using a script written in Perl language.

Bioanalytical Equipment and Reagents. The source of recombinant enzymes was bacterial membranes containing human cytochrome P450 (CYP2D6 and CYP3A4) coexpressed with NADPH-cytochrome P450 reductase supplied from Cypex Ltd., Scotland, U. K. The test compounds were purchased from Sigma Chemicals Co., St. Louis, Missouri (dextromethorphan, 3-methoxymorphinan, levallorphan, and codeine), ICN Biomedicals Inc., Ohio (dextrorphan), and Lipomed AG Arlesheim, Switzerland (norcodeine, dihydrocodeine, dihydronorcodeine, oxycodone, hydrocodone, morphine, etylmorphine, dihydromorphine, and nalorphine). NADPH was purchased from Sigma Chemicals Co, St. Louis, Missouri, acetonitrile from Rathburn Chemicals Ltd. Walkerburn, Scotland, and formic acid from Merck KGaA, Darmstedt, Germany.

Liquid chromatography−mass spectrometry (LC−MS) was used for the metabolite identification. The high-performance liquid chromatography (HPLC) system was an Agilent 1100 Series (Hewlett-Packard GmbH, Walbronn, Germany) coupled to a HTC PAL auto sampler (CTC Analytics AG, Zwingen, Germany). Analytical columns used for chromatographic separation of the metabolites were HyPurity C18 50X2.2, 5μ for dextromethorphan, dextrorphan, and 3-methoxymorphinan; HyPurity C18 100X2.5 5μ for levallorphan; and Zorbax XDB-C8 150X4.6 5μ for codeine, norcodeine, dihydrocodeine, dihydronorcodeine, oxycodone, hydrocodone, morphine, etylmorphine, dihydromorphine, and nalorphine. The mobile phases consisted of water and acetonitrile acidified with 0.1% formic acid. The HPLC system was connected to a Sciex API 4000 quadrupole mass spectrometer with an electrospray ionization interface (Applied Biosystems, Ontario, Canada). Positive-ion mass spectra were recorded over mass range m/z 80−400 followed by product ion scans on peaks of interest. The Analyst 1.4.1 software (Applied Biosystems) was used for analysis and storage of data.

Data Set. A set of opioid analgesics (Figure 1) was used for the analysis since most members of this family of compounds are known to be metabolized by CYP2D6 through O-demethylation and CYP3A4 through N-demethylation.[13−17] With this data set, both CYP3A4 and CYP2D6 are covered. The structures were drawn in ISIS and exported as SDF files. The 2D−3D conversion and the generation of stereoisomers were made with CORINA. Since the compounds are not very flexible, only one or two conformations are in most cases possible, which clearly simplified the computational work.

In addition to the opioid analgesics, a set of 15 known CYP3A4 inhibitors (Figure 2) was docked in CYP3A4 to get more structural diversity and a wider range of possible interactions than that obtained with the opioid family. This data set included more flexible molecules, and since the bioactive conformation was unknown, different conformations were generated with the random search module in Sybyl. The maximum number of conformations was set to 15, and the energy cutoff was 3 kcal/mol.

METHODS

Overview. The strategy and procedures used in this study are depicted diagrammatically in Scheme 1 and briefly summarized here. Major metabolic routes for a set of 14 opioid analgesics were determined experimentally in recombinant CYP2D6 and CYP3A4. The site of metabolism prediction software MetaSite was used on these compounds,

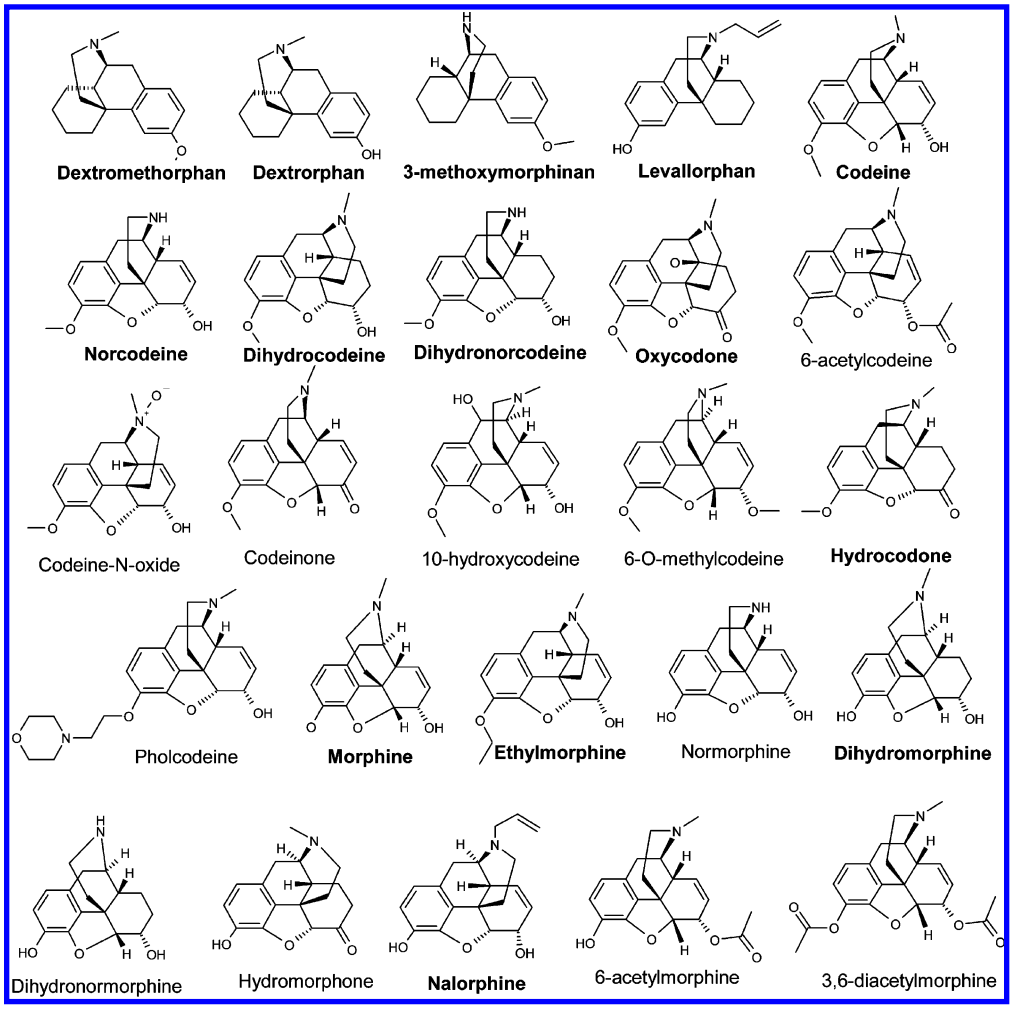


Figure 1. Structures of 25 opioid analgesics used in the analysis. Experimental determination of metabolite ID has been done on compounds with compound name in bold.

and the predictions were made on the different structures for the above-mentioned isoforms. For CYP2D6, a recently published crystal structure (PDB code 2f9q)[10] and a previously used homology model[18] were compared. For CYP3A4, three crystal structures, two without substrate or inhibitors present in the active site (PDB codes 1tqn and 1woe)[7,8] and one cocrystallized with erythromycin (PDB code 2j0d)[9] were compared with a homology model.[18] The different structures were superimposed using the backbone of the proteins and visually compared.

Further comparison was made from the different poses obtained from docking in the active site of the CYPs using the program GLUE. In order to enlarge the range of possible interactions, an extended data set with 25 opioid analgesics and known CYP3A4 inhibitors was used for the analysis. A method, which involves the calculation of interaction energies between ligands and amino acids in the active site from docking poses, was used to compare the different protein structures. For CYP3A4, four different structures were studied, and this method was compared with the results from using consensus principal component analysis (CPCA), a commonly used method for structural comparison. The analyses of docking poses by calculation of the accumulative interaction energy between the molecules and the amino acids in the binding site were further visualized with the multivariate data analysis method, principal component analysis (PCA).

Metabolite Identification. The metabolic pattern of the 14 opioid analgesics was investigated in recombinant enzymes CYP2D6 and CYP3A4 (Cypex). The compound (10 μM) was mixed with rCYP (100 pmol/mL) and a potassium phosphate buffer (0.1 M, pH 7.4). After a 10 min preincubation, the reaction was initiated by the addition of NADPH (1 mM) and incubated for 60 min. All incubations were performed at 37 °C. The reaction was quenched with two parts ice-cooled acetonitrile and 0.8% formic acid; the samples were centrifuged at 4000 rpm for 20 min at 4 °C and the supernatant diluted 1:2 in water prior to analysis.

The samples were analyzed on a Sciex API 4000 instrument. Initial full scans were obtained between *m*/*z* 80 and 400, and metabolites were identified by comparing $t = 0$ min samples with $t = 60$ min samples. Structural information was generated from the fragmentation pattern from product ion scans of the ions of interest with a collision energy of 45 eV. The interpretation of the fragment ions was based on previous studies on MS fragmentation of alkaloids.[19]

MetaSite. MetaSite is a tool for predicting the site of metabolism.[11,12,20] The methodology considers structural complementarity between the active site of the enzyme and the ligand as well as the reactivity of the ligand. To compare the enzyme and ligand, two sets of descriptors are calculated. For the enzyme, these are based on flexible GRID molecular interaction fields (GRID-MIFs). In order to generate the descriptor set of the ligand, each atom of the molecule is

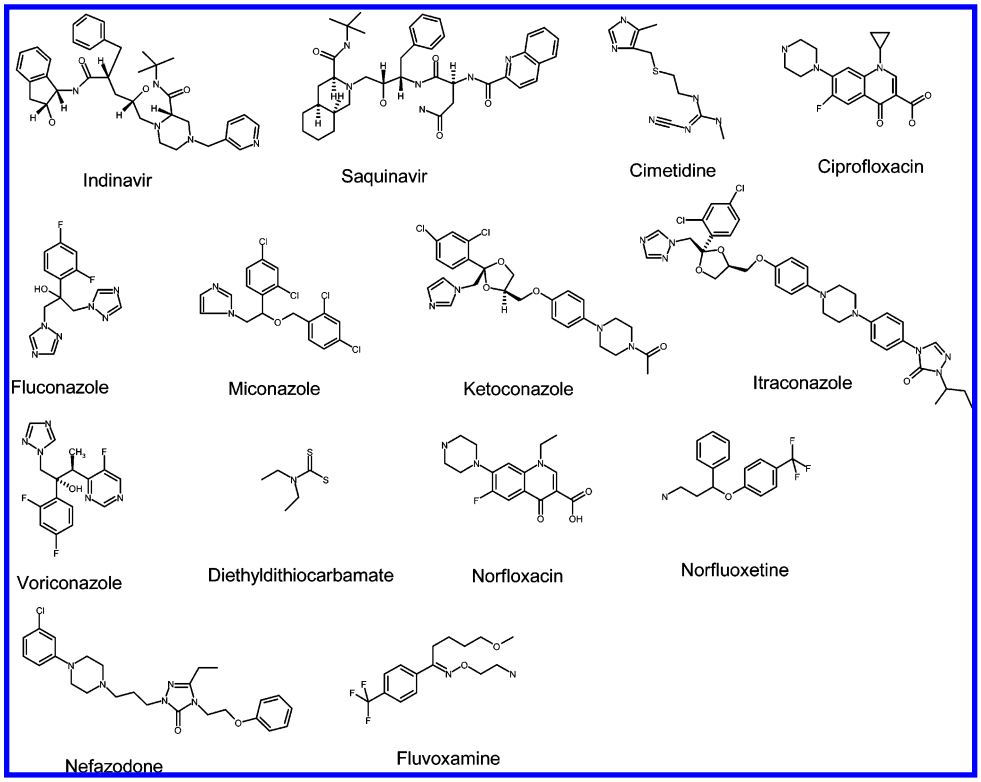
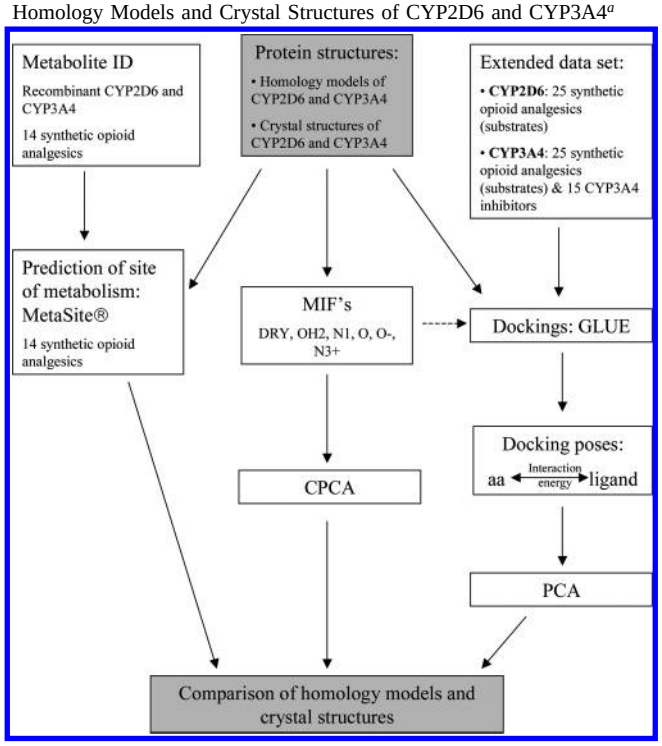


Figure 2. Structures of CYP3A4 inhibitors used in 3A4 docking analysis.

Scheme 1. Flow Chart of the Experimental Strategy Used to Compare Homology Models and Crystal Structures of CYP2D6 and CYP3A4[a]



[a] CPCA, consensus principal component analysis; aa, amino acids; PCA, principal component analysis; MIF, molecular interaction fields.

classified as a GRID probe and the distances between them are calculated. The resulting fingerprint of the ligand is compared with the description of the enzyme active site, and the most optimal orientations are obtained. On the basis of this, all atoms in the molecule are ranked according to their accessibility to the heme. In addition, a reactivity factor based on fragment recognition is added. In summary, the site of metabolism is described by a probability index which is the product of the similarity between the ligand and protein (i.e., accessibility of one atom to the heme) and the reactivity.

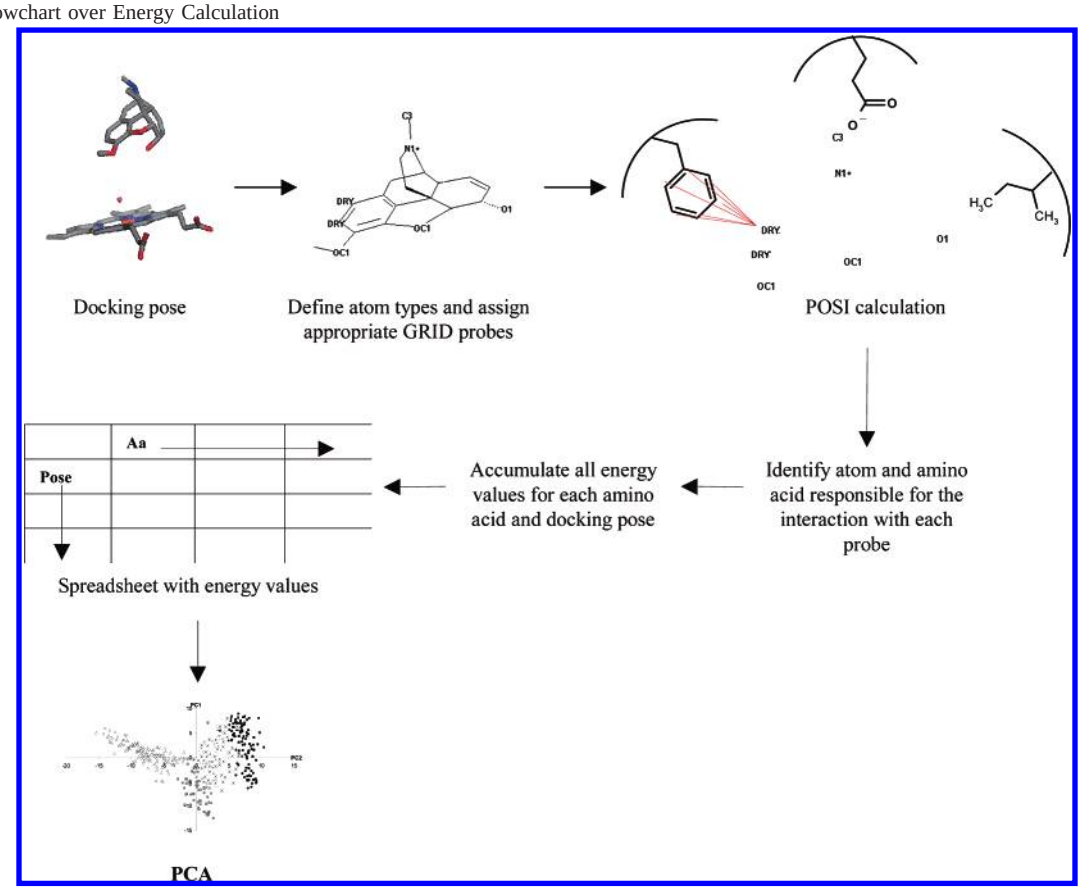
Table 1. Grid Probes Used for Calculation of Molecular Interaction Fields (MIFs) in CPCA and Dockings with GLUE

probe	chemical group	used in
OH2	water	CPCA/dockings
DRY	hydrophobic	CPCA/dockings
H	neutral hydrogen	dockings
N1	neutral flat NH (e.g., amide)	CPCA/dockings
N1/2/3+	sp^3 amine cation	CPCA/dockings
N:	sp^3 N with lone pair	dockings
O	sp^2 carbonyl oxygen	CPCA/dockings
O−	sp^2 phenolate oxygen	CPCA
O::	sp^2 carboxy oxygen	dockings
O1	alkyl hydroxy OH group	dockings
OC1	aromatic/aliphatic ether oxygen	dockings

For this analysis, the compounds were loaded into MetaSite as single conformers both as neutral species and protonated at the nitrogen atom according to the pK_a of the compounds. When analyzing the results, the top three averaged ranked solutions with reactivity correction enabled were taken into account.

Calculation of GRID Molecular Interaction Fields. The proteins were aligned and imported into Greater, which is an interface for the program GRID. The MIFs in the binding site were calculated by the GRID methodology.[21] In this program, a grid is built over the molecule, and in each grid point, probes representing different functional groups are placed and the interaction energy is calculated with the GRID force field. When analyzing the protein, a box size is defined including the heme and the active-site cavity. To get the same net charge of the different protein structures, they were neutralized by adding movable counterions (Cl^- or Na^+). The probes used to generate MIFs were DRY, N1, O, O−, N3+ and OH2 (see description in Table 1). The MIFs

Scheme 2. Flowchart over Energy Calculation



generated in GRID were imported into GOLPE and pretreated in order to focus the areas of interest. This pretreatment included (1) a maximum cutoff of 0 kcal/mol, considering only the favorable negative interaction energies, (2) a standard deviation cutoff, excluding variables with a variance less than 0.05 standard deviations, (3) the application of block unscaled weights, a scaling factor which gives all blocks the same variance and normalizes the interaction energies between the different probes, and (4) a cutout of the active site only including grid points within 3 Å from an ensemble of all docking poses achieved in GLUE (see below).

PCA and CPCA. PCA is a method used to handle multivariate data. It reduces the dimensionality of multidimensional data, while explaining as much of the variation as possible.[22] From the multidimensional space, principal components (PCs) are extracted, which are linear combinations of the original variables. The first PC is a line through the data space which best describes the variation of the data. The second PC is orthogonal to the first and explains as much of the variance that was not described by the first PC and so on. When two PCs are selected, two smaller matrixes, the score matrix and the loading matrix, are generated. The two-component score matrix is a plane generated by the first two components on which the objects can be projected (score plot). Objects that are close in the score plot have comparable variance and are therefore similar. The loading matrix relates each PC to the original variables (loading plot). The score plot and loading plot are related, and variables that have positive influence on an observation are positioned in the same place in the loading plot as the observation in the score plot. Here, we are dealing with negative energies (the more favorable the interaction, the more negative the value), and consequently, a variable which has strong influence on an observation is positioned with the same coordinates but with opposite signs in the loading plot.

CPCA is a method based on PCA and identifies selectivity between proteins based on differences in interaction patterns.[23] The CPCA uses MIFs calculated for different probes in the active site of the proteins. The data from each probe is organized as blocks in the matrix. That is, the MIFs are added side by side, giving an addition of new variables to the same object. The resulting matrix has one row for each protein and $n \times k$ columns where n is the number of blocks (probes) and k is the variables. The CPCA can be considered as a PCA in two levels. A superlevel, which is a consensus of all blocks, and the solution is identical to the regular PCA. At the second level, the block level, blocks are extracted from the superlevel, and it describes the variation each probe is responsible for. That is, the influence of each probe on the whole model. On this level, one can get scores and loadings for each probe. An additional solution of the CPCA is a super weight matrix, which gives the partition of each probe on the overall scores and relates the two levels of information.

Dockings with GLUE. The docking module GLUE is based on calculations of GRID-MIFs.[21] For the dockings in CYP2D6, the molecules were protonated since the electrostatic interaction between the basic nitrogen in the ligands and the positively charged Glu216 on top of the active site in CYP2D6 are believed to be important for the orientation of this set of compounds. When the results were compared with the metabolite identification data, the dockings with positively charged molecules gave better results compared

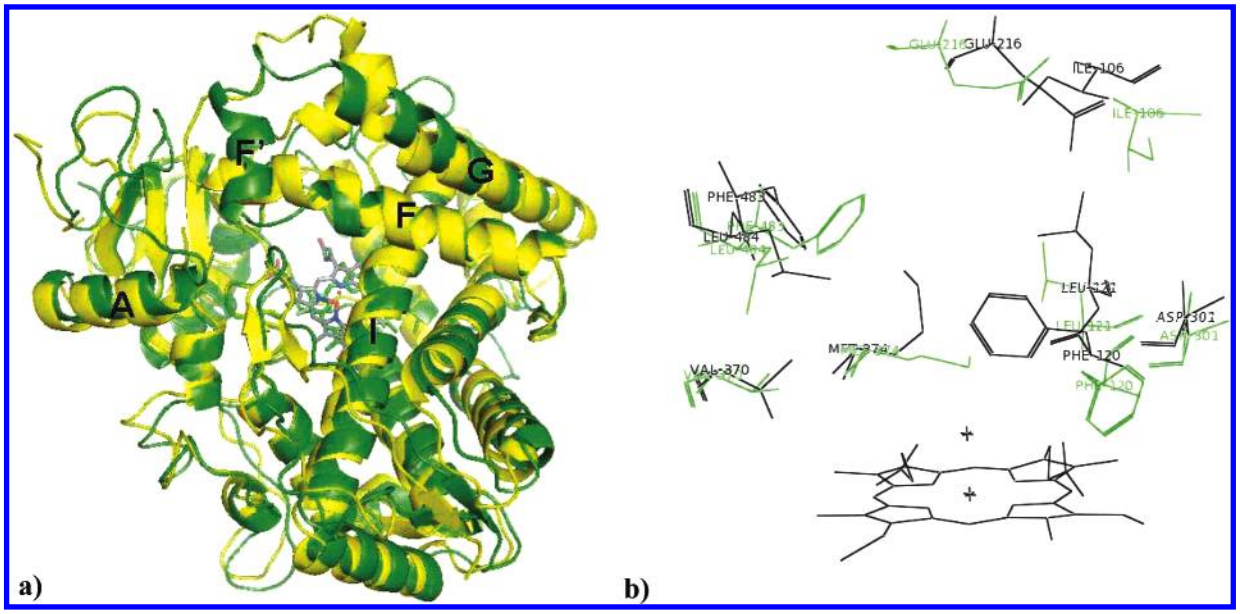


Figure 3. Comparison of CYP2D6 structures: (a) Secondary structure element in homology model (green) and 2f9q (yellow). (b) Selected amino acids in the active site of the homology model (green) and 2f9q (black).

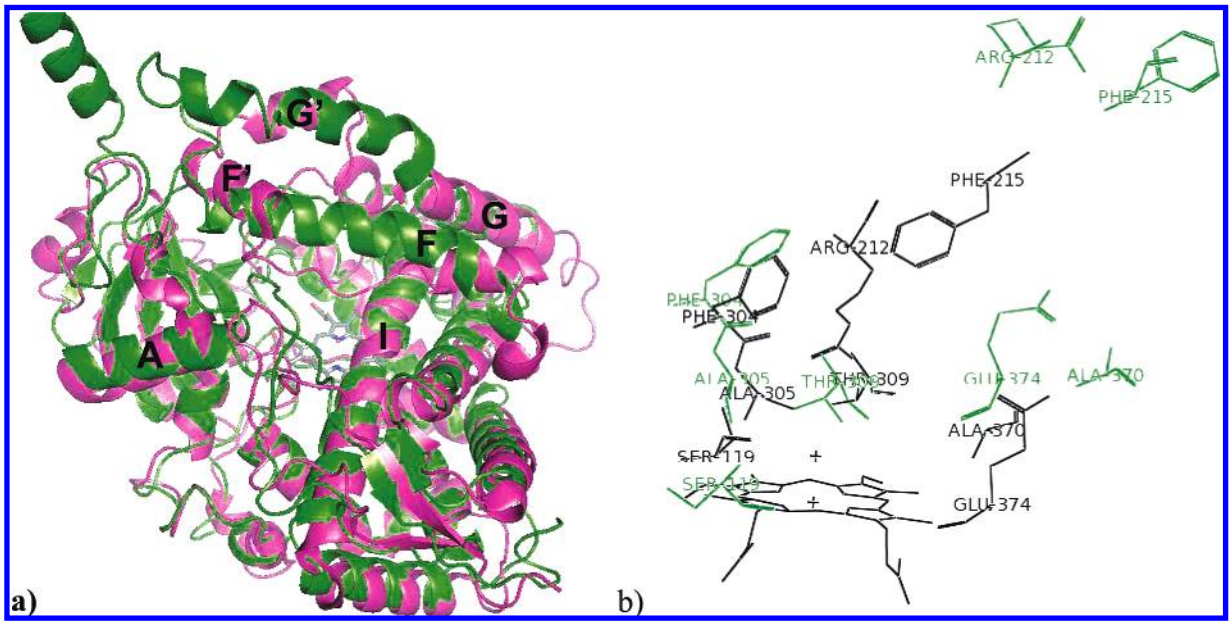


Figure 4. Comparison of CYP3A4 homology model and ligand free crystal structure (1tqn) a) Secondary structure element in homology model (green) and 1tqn (magenta). b) selected amino acids in the active site of the homology model (green) and 1tqn (black).

to dockings with neutral molecules. The importance of electrostatic interaction for substrate binding in this enzyme has also been previously reported.[24,25] Specific probes, which suited the data set of opioid analgesics, were chosen for analysis of the active site of the enzyme, in this case, H, OH2, DRY, N1+/N2+, and OC1 (see description in Table 1). For the binding affinity estimation, contributions from electrostatic interactions were included. Even though this set of compounds is not very flexible, the number of rotable bonds was set to three to provide for ligand flexibility. A maximum number of binding sites was set to 300, and an energy cutoff at −100 kcal/mol was included.

In contrast to CYP2D6, the dockings in CYP3A4 were performed on neutral molecules. This was based on the fact that CYP3A4 is mainly responsible for the N-demethylation of these compounds and positively charged nitrogen would be repelled by the charged iron-oxy species. That is, for N-demethylations, a basic nitrogen must somehow be deprotonated prior to oxidation to make this reaction possible. For the opioid analgesics, the same probes as for CYP2D6 were used, except for N1+/N2+, which were exchanged with N:/N1: (Table 1). Otherwise, the same settings were applied. For the set of CYP3A4 inhibitors, the default probes (H, OH2, DRY, N1, N+, O, O::, and O1) were used since these compounds are more diverse in structure. For the docking calculations, the number of rotable bonds was set to five to provide for flexibility of the ligand.

Since the aim was to identify sites of metabolism and orientation of a substrate in the active site, the resulting docking poses were filtered, and only those that were within 6 Å from any atom of the heme were selected for further analysis.

Energy Calculations. The active-site cavities were extracted and defined by the amino acid residues involved in

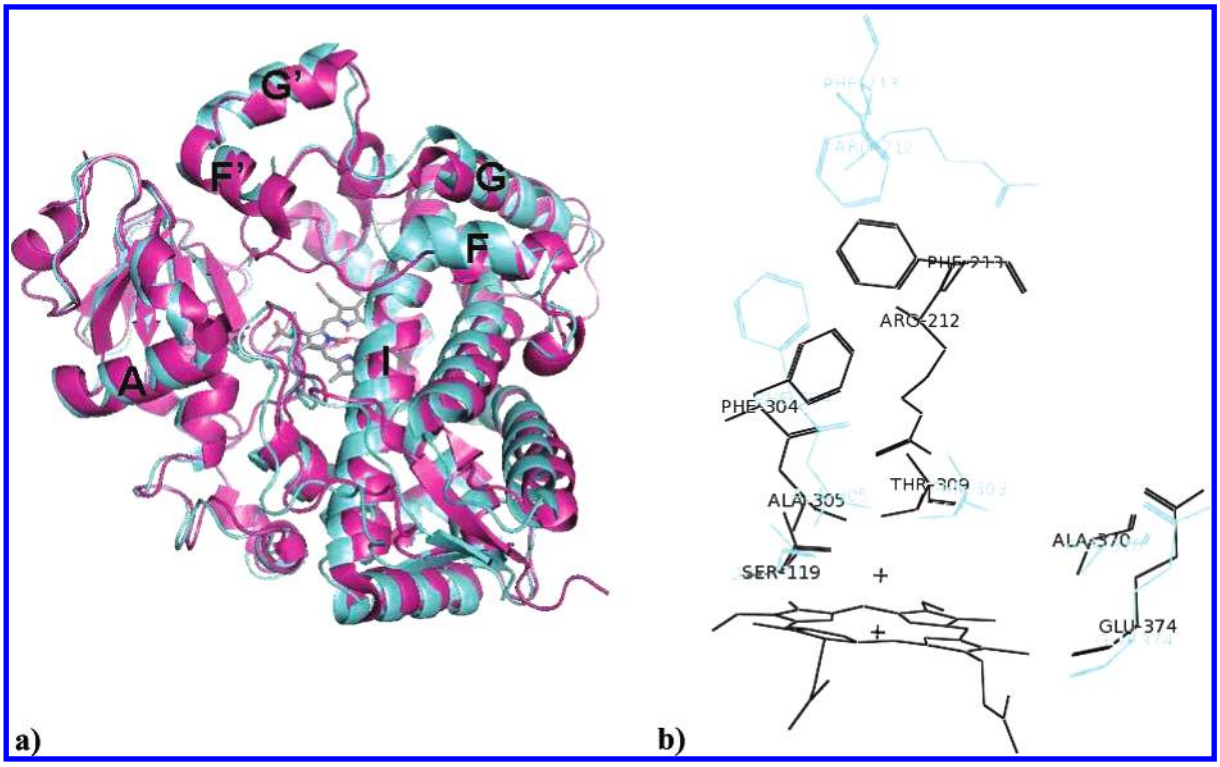


Figure 5. Comparison of CYP3A4 ligand-free (1tqn) and erythromycin complex (2j0d) crystal structure: (a) Secondary structure element in 1tqn (magenta) and 2j0d (cyan). (b) Selected amino acids in the active site of the 1tqn (black) and 2j0d (cyan).

any interaction with any of the selected docking poses. These amino acids were used for calculation of the interaction energies. The computations were made using the docking poses obtained from the program GLUE.

A flowchart of this procedure is depicted in Scheme 2. Each atom in the docked molecules was assigned to an appropriate GRID probe, that is, functional groups representing different physicochemical properties (see example for codeine in Scheme 2). This resulted in GRID probes that were located at the atom positions achieved from the docking poses. At these *x*, *y*, and *z* coordinates, the energy of interaction was calculated by GRID in the protein cavity. The calculations were performed with the POSI directive, which is described in more detail in the GRID user manual.[26] This directive optimizes the probe orientation with respect to the protein, and the obtained energy is the maximum possible interaction for the probe (atom), and it is not influenced by constraints such as bonds, angles, and neighboring atoms. After the POSI calculation, an analysis was done in order to identify which atoms in the protein are responsible for each interaction, and the energy values for all atoms of the same amino acid residue were summed up. This process was repeated for all atoms in the docking pose, accumulating the energy of interaction for each amino acid and resulting in one energy value for each docking pose and amino acid residue. The accumulated energy provided a good representation of interactions between two aromatic systems since this is in fact a hydrophobic interaction (as defined with the DRY probe in GRID) between six atoms from the phenyl group in the ligand and another six from a phenylalanine (Phe) in the protein. The output was a table with the amino acid residues as columns and the docking poses as rows, containing the accumulative energies between them (see example in Scheme 2).

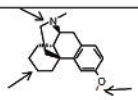
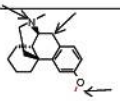
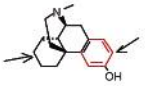
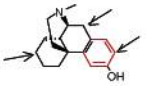
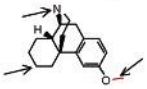
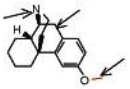
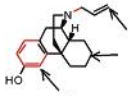
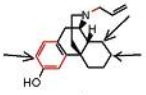
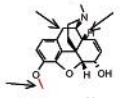
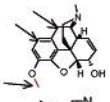
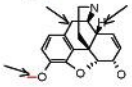
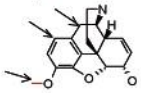
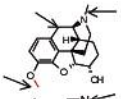
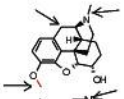
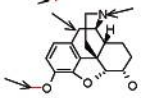
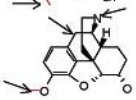
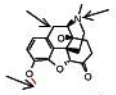
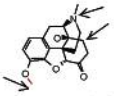
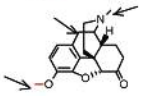
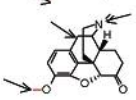
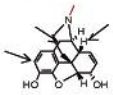
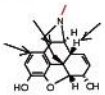
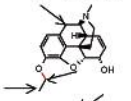
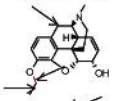
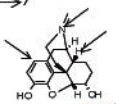
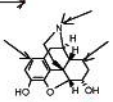
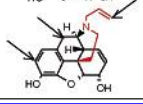
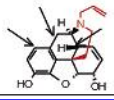
Data Analysis. The results from the energy calculations resulted in a large amount of data, which were further analyzed by PCA.

GOLPE was used for the PCA analysis of the interaction data obtained from the dockings. All docking poses from the different structures of each isoform were collected in a single table. The amino acids in the binding site represented the variables. Five principal components were extracted, but only the first two were used for analysis since the higher components did not give any additional interpretable information. The results were visualized in score and loading plots. Since the energy has a negative sign, the amino acids on one side of the PCA loading plot are important for the binding of substrate in the docking poses positioned on the opposite side of the score plot.

RESULTS

Visual Comparison of the CYP2D6 Protein Structures. By superimposing the structures using the backbones of the proteins and comparing the positions of the amino acids in the active site, some major differences could be identified. In the homology model, the hydrophobic pocket is spread out over the heme formed by, for example, the orientation of Ile106, Leu121, Leu213, Leu284, and Phe483. In the crystal structure, the different orientation of Ile106 and Phe483 makes the hydrophobic pocket more narrow and forms a channel from the heme to the surface. The orientation of the Phe120 into the active site is also contributing to the narrow shape of the pocket. Glu216 and Asp301 are believed to have a major impact on the orientation of the substrate due to charge−charge interactions. When the homology model and the crystal structure are compared, Glu216 is oriented toward the cavity in the homology model, while Asp301 does not differ much between the two structures.

Table 2. Results from Metabolite Identification and MetaSite Predictions in CYP2D6 Homology Model and Crystal Structure[a]

Compound name	Metabolic reaction	Homology model	2F9Q
Dextromethorphan	O-demethylation		
Dextrorphan	Aromatic hydroxylation		
3-Methoxymorphinan	O-demethylation		
Levallorphan	N-dealkylation, aromatic hydroxylation		
Codeine	O-demethylation		
Norcodeine	O-demethylation		
Dihydrocodeine	O-demethylation		
Dihydronorcodeine	O-demethylation		
Oxycodone	O-demethylation		
Hydrocodone	O-demethylation		
Morphine	N-demethylation (minor)		
Ethylmorphine	O-dealkylation		
Dihydromorphine	No metabolites		
Nalorphine	Hydroxylation		

[a] Experimentally determined major sites of metabolism are highlighted in red, and sites proposed by MetaSite are indicated by arrows.

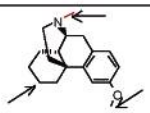
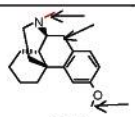
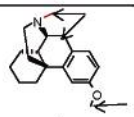
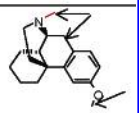
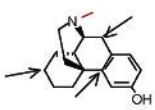
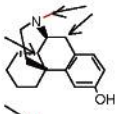
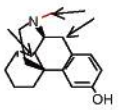
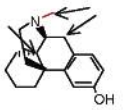
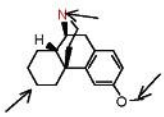
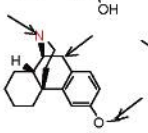
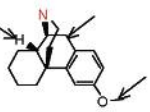
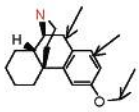
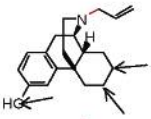
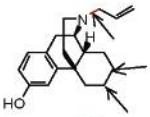
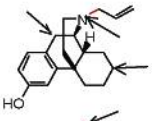
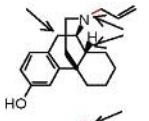
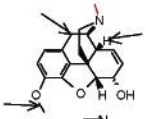
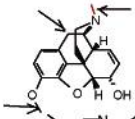
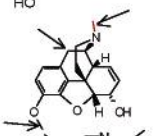
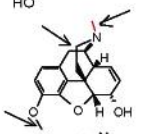
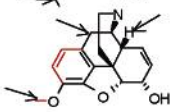
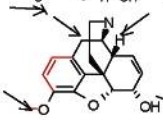
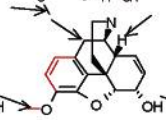
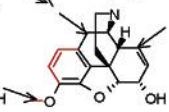
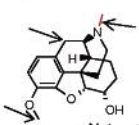
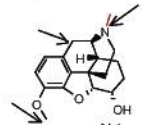
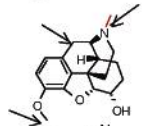
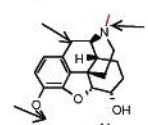
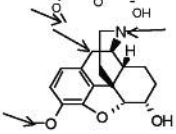
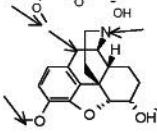
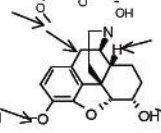
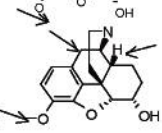
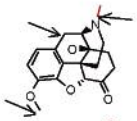
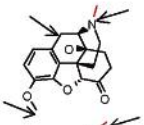
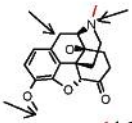
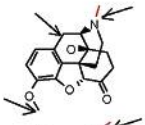
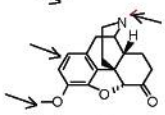
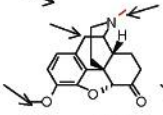
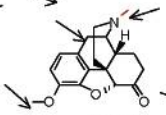
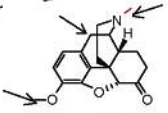
Calculations of MIFs with the N1+ probe show that the position of the Glu216 in the homology model contributes to a more widespread positively charged field within the cavity, while in the crystal structure, it is more concentrated to one side. The secondary elements of the two CYP2D6 structures show some overall shifts, but the major difference is located in the F−G region where the crystal structure has a long F helix connected to the G helix with a loop, while the homology model has a shorter F helix and an additional F′ helix in the loop between F and G helices. The differences in the CYP2D6 structures are visualized in Figure 3a and b.

Visual Comparison of the CYP3A4 Protein Structures. The two structures without cocrystallized ligand or inhibitor (1tqn and 1woe) are very similar, and the most pronounced difference is the orientation of Arg212, which points into the active site in the case of 1tqn. There are also some minor differences in the orientation of Lys173 and Arg255. The position of Arg212 in 1tqn results in a split hydrophobic field over the heme, while in the other structures, it is more compact. In 2j0d, the structure with erythromycin present in the binding site, the hydrophobic field indicates a larger pocket and a rather wide access channel from the surface. Crystal structure 2j0d and the homology model have an overall different conformation compared to the other two. In both of them, the Arg212 is oriented out of the active site. Concerning the secondary structure element, the two ligand-free crystal structures are very similar. Structure 2j0d is somewhat different from these two in which the entire F−G region is moved up resulting in a larger pocket.[9] An estimation of the solvent-accessible volume suggests an increase of the volume of the active site pocket to ∼2000 Å^3 compared to 950 Å^3 in the ligand-free structure.[9] The secondary structure of the homology model differs from the crystal structures mainly in the F−G region where it consists of two long helices (F and G) connected by a loop, while this region in the crystal structures has two additional shorter helices (F′ and G′). Differences in the secondary structure of the CYP3A4 structures are visualized in Figures 4 and 5.

Metabolite Identification and MetaSite Predictions. The results of the metabolite identification of the 14 opioid analgesics are in agreement with the literature data available.[13−17,27,28] In CYP2D6, the major metabolic pathway is O-demethylation for almost all compounds. The proposed metabolism of levallorphan and morphine, which lack the methoxy group, is through N-deallylation and N-demethylation, respectively. For morphine, this pathway produces a minor metabolite. Levallorphan is also metabolized by aromatic hydroxylation in addition to the N-deallylation. In CYP3A4, the compounds are metabolized by N-demethylation/N-deallylation. Some exceptions are in 3-methoxymorphinan, which undergoes N-oxidation, and norcodeine, which is metabolized on the aromatic ring by O-demethylation and hydroxylation. The proposed major sites of metabolism are highlighted in red in Tables 2 and 3. In addition to these metabolic reactions, minor metabolites in the form of hydroxylations and combinations of demethylation and hydroxylation are formed for several compounds. LC−MS/MS data for the compounds and major metabolites are available in the Supporting Information.

Except for the CYP3A4 homology model, the prediction rate in MetaSite is between 79 and 100% when considering the top three ranked solutions (Table 4). The predictions made with 2j0d result in the correct site of metabolism for 12 of 14 compounds, but the prediction rate for the homology model is only 50%. In most cases, the three crystal structures pick the same top three sites of metabolism, while the homology model differs in some predicted sites. For CYP2D6, the homology model and the crystal structure predictions are very similar and at a high rate (92% and 85%, respectively). The MetaSite-predicted sites for each structure are indicated with arrows in Tables 2 and 3. Table 4 summarizes the prediction rates.

Table 3. Results from Metabolite Identification and MetaSite Predictions in CYP3A4 Homology Model and Three Crystal Structures[a]

Compound name	Metabolic reaction	Homology model	2j0d	1tqn	1woe
Dextromethorphan	N-demethylation				
Dextrorphan	N-demethylation				
3-methoxymorphinan	N-hydroxylation				
Levallorphan	N-dealkylation				
Codeine	N-demethylation				
Norcodeine	O-demethylation +aromatic hydroxylation				
Dihydrocodeine	N-demethylation				
Dihydronorcodeine	No metabolites				
Oxycodone	N-demethylation				
Hydrocodone	N-demethylation				
Morphine	No metabolites				
Ethylmorphine	N-demethylation				
Dihydromorphine	Alifatic hydroxylation				
Nalorphine	N-dealkylation				

[a] Experimentally determined major sites of metabolism are highlighted in red, and sites proposed by MetaSite are indicated by arrows.

Dockings and Energy Calculations. The comparison based on interaction energies between ligands and amino acids in the active site reveals some differences between the CYP2D6 crystal structure and homology model. By looking at the spreadsheet with calculated energies, amino acids could be identified, which show a higher energy of

Table 4. Calculated Prediction Rates for MetaSite Using Crystal Structures and Homology Models of CYP2D6 and CYP3A4[a]

CYP3A4				CYP2D6	
homology model	2j0d	1tqn	1woe	homology model	2f9q
50% (58%)	86% (100%)	79% (92%)	79% (92%)	86% (92%)	79% (85%)

[a] The prediction rates were calculated by dividing the number of correct predicted compounds with the total number of compounds used in the study. The calculations were first made for all compounds and then done excluding compounds that do not undergo metabolism as indicated in parentheses.

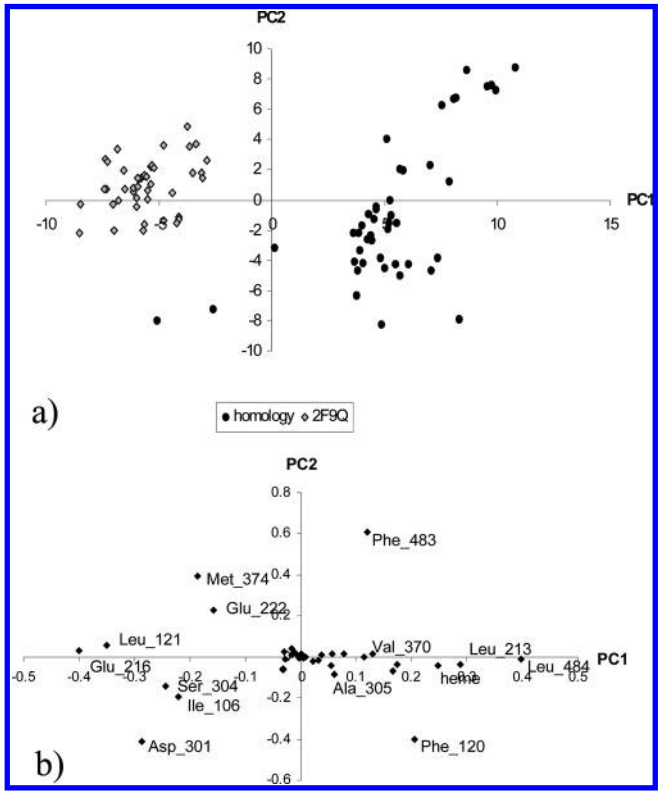


Figure 6. PCA results on energy calculations from dockings of opioid analgesics in CYP2D6 homology model (●) and crystal structure (gray ◇). (a) PCA score plot; (b) PCA loading plot.

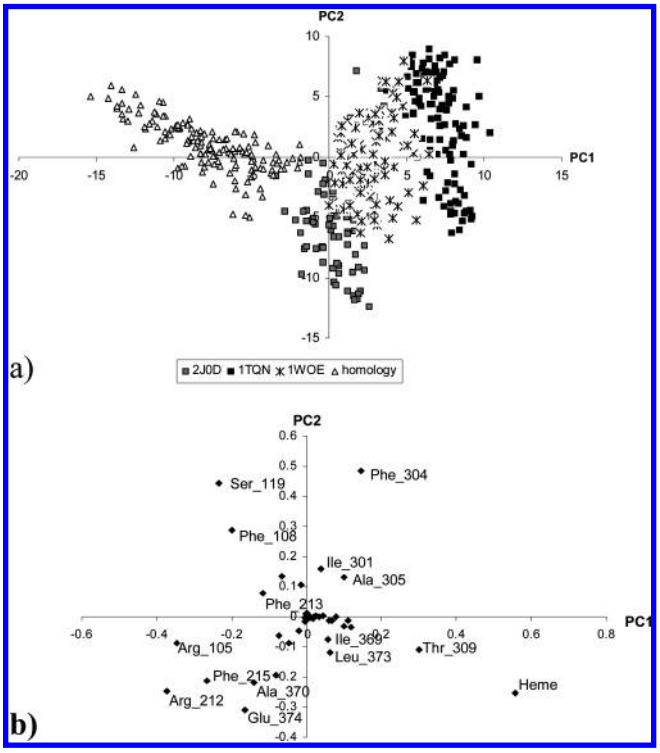


Figure 7. PCA results on energy calculations from dockings of opioid analgesics in CYP3A4 homology model (△) and crystal structures 2j0d (gray □), 1tqn (■), and 1woe (∗). (a) PCA score plot; (b) PCA loading plot.

interaction in one structure compared to the other. PCA analysis, however, made it easier to visualize and interpret these interactions (Figure 6). The homology model and the crystal structure are clearly separated over the first principal component in the PCA score plot. By correlating the score plot with the loading plot, these results suggest the following differences in the two structures: Leu121, Glu216, Ser304, Ile106, Asp301, and Met374 seem to have greater importance in interacting with the ligand in the homology model; that is, the energy of interaction is greater in the homology model. On the other hand, Phe483, Leu484, Leu213, Phe120, and Val370 have greater interaction energy with the ligands in the crystal structure. In general, the results indicate that the hydrophobic interactions are more important in the crystal structure. Some of the amino acids showing major differences are visualized in Figure 3b.

The results based on the dockings in CYP3A4 are shown in Figures 7 and 8. Similar results are achieved from dockings with the two different data sets, the opioid analgesics and the CYP3A4 inhibitors. In the PCA score plots, the four different structures are clearly separated in clusters and over the first principal component. Moreover, in the analysis of which amino acids differ between the different structures in regard to the interaction with the ligands, the results from both data sets are similar. The crystal structure 1tqn and the homology model are best separated over the first principal component, and the most discriminative interactions are due to Phe304, Thr309 amino acids, and the heme, which are more important for substrate binding in the homology model. On the contrary, for 1tqn, Arg212, Phe215, Ala370, and Glu374 are suggested to play an important role. The most pronounced difference is due to Arg212, and in this structure, this residue has a different orientation compared to the others. Ser119 and Phe304 seem to be selective for the substrate binding in 2j0d, and the other ligand-free crystal structure, 1woe, does not seem to have pronounced different interactions compared to the other structures. The amino acids highlighted in the text are visualized in Figures 5b and 6b.

CPCA. CPCA analysis aims at comparing the active sites of the CYP3A4 homology model with the crystal structures using molecular interaction fields in the postulated active sites of the enzymes. In CPCA, the superweights plot gives information on the partition of the different blocks (GRID probes) on the overall score (Figure 9a). In this plot, it is shown that all probes have a similar contribution to the first component but that the hydrophobic probe (DRY) differs

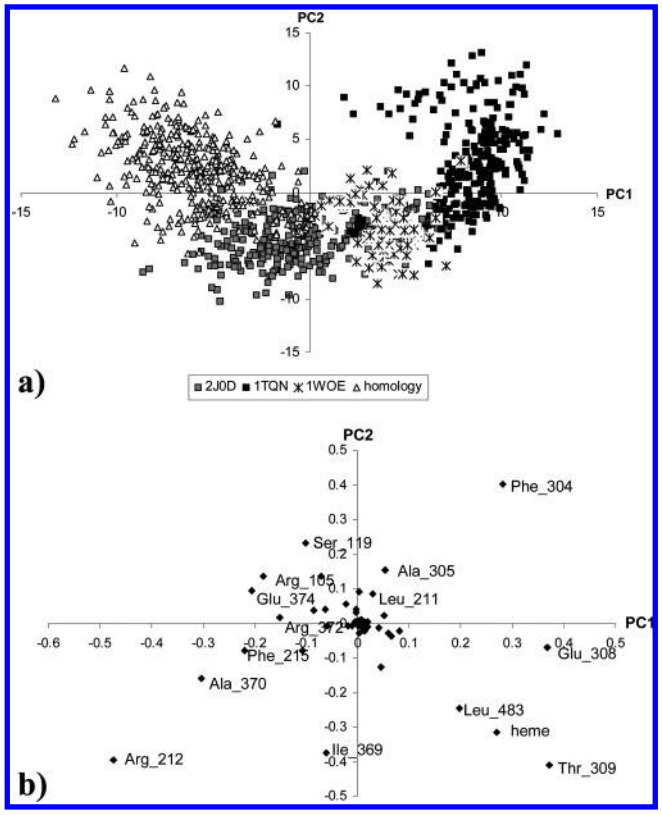


Figure 8. PCA results on energy calculations from dockings of 3A4 inhibitors in CYP3A4 homology model (△) and crystal structures 2j0d (gray □), 1tqn (■), and 1woe (∗). (a) PCA score plot; (b) PCA loading plot.

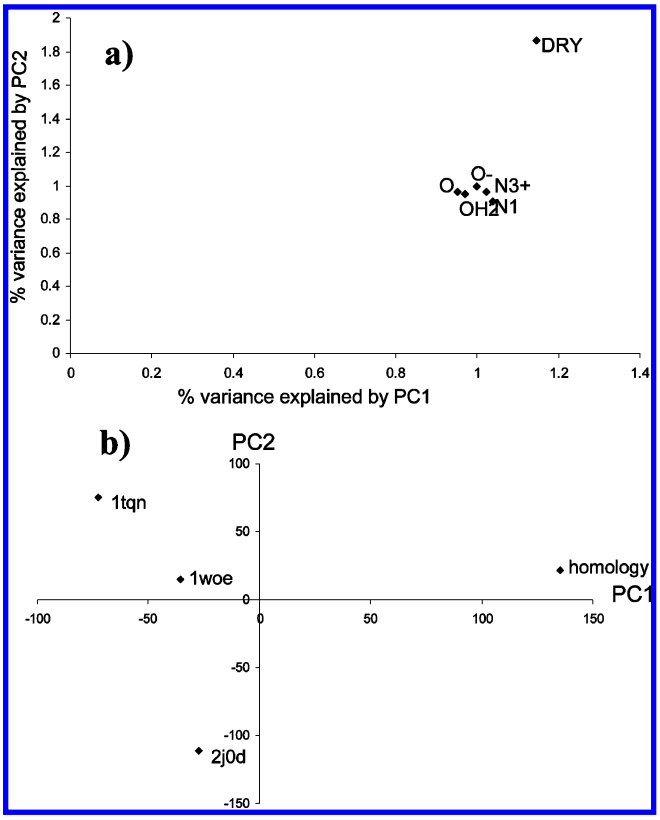


Figure 9. CPCA on all CYP3A4 structures based on molecular interaction fields. (a) Superweights plot describing the influence of the different probes. (b) PCA score plot showing the intercorrelation between the structures.

from the others over the second component. In the PCA score plot (Figure 9b), it is shown that the first component discriminates between the homology model and the crystal structures while the substrate-free crystal structure, 1tqn, and the erythromycin bound crystal structure, 2j0d, are separated over the second component. The relative positions of the different structures in the score plot resemble the pattern of the PCA score plot from the energy calculation methodology. On the block level, the most important discriminant loadings described by each probe are selected and projected back to the 3D structure of the protein. The molecular interaction fields corresponding to the loadings can be visualized, and the amino acids responsible for these discriminative fields can be identified. For each GRID probe the selective amino acid residues are identified for the four different protein structures. A summary of the amino acids in each structure is outlined in Table 5. Also present in the table are the corresponding secondary structure elements. The identification of selective amino acids in the CPCA is biased by the number points that are picked in the block level loading plot, and the results could therefore include some noise. CPCA however still picks the amino acids identified from the energy calculation method, for example, Phe108 and Phe304 in the 2j0d crystal structure; Ala370 and Arg212 in the ligand-free (1tqn) crystal structure; and Thr309, Glu308, and the heme in the homology model. The second ligand-free crystal structure (1woe) is positioned in the center of the PCA score plot, which means that the data are similar to the average, indicating that the structure is a mixture of the other three.

DISCUSSION

Efforts to predict P450-mediated metabolism are improving with the availability of crystal structures of P450 enzymes and different computational tools. The overall aim of this study was to compare previously used homology models with crystal structures for CYP2D6 and CYP3A4 using MetaSite and the docking program GLUE. A set of synthetic opioid analgesics was used for the analysis since these compounds are previously known to be metabolized mainly through N-demethylation by CYP3A4 and O-demethylation by CYP2D6. For the 14 compounds used, the choice of protein structure did not seem to influence the MetaSite predictions to any large extent. The CYP2D6 homology model and crystal structure gave similar results within the top three ranked positions for all compounds even though the active site environment differed in the orientation of some important amino acid residues, for example, Glu216. Nevertheless, the MetaSite procedure uses flexible GRID-MIFs where the amino acid side chains are allowed to move. In this case, it could be that different static views of the protein (homology model and crystal structure) could lead to a similar dynamic representation of the potential interaction sites. Since the reactivity correction was enabled, an explanation could also be that the reactivity is driving the results of this set of compounds. It is also worth mentioning that these compounds are rather small and compact and can probably adopt different orientations in the active site, which further emphasizes the importance of the reactivity. For CYP3A4, the three crystal structures gave more or less the same results, but the homology model differed and the prediction rate was lower compared to the other structures. The visual comparison of

Table 5. CPCA Results for the CYP3A4 Structures[a]

homology		2j0d		1tqn		1woe	
1°	2°	1°	2°	1°	2°	1°	2°
Asn104	B−C loop	Arg106	B−C loop	Asp76		Phe108	B−C loop
Arg105	B−C loop	Phe108	B−C loop	Phe108	B−C loop	Arg212	F−G loop
Val111	B−C loop	Phe213	F−G loop	Ile120	B−C loop	Phe213	F−G loop
Phe304	I helix	Phe241	G helix	Arg212	F−G loop	Phe215	F−G loop
Phe305	I helix	Phe304	I helix	Phe213	F−G loop	Phe304	I helix
Glu308	I helix	Ala305	I helix	Asp214	F−G loop	Ile369	K′-β1 loop
Thr309	I helix	Gly306	I helix	Phe304	I helix	Glu374	β1−2
Thr310	I helix	Tyr307	I helix	Ile369	K′-β1 loop	Gly481	C-term
Ser312	I helix	Glu308	I helix	Ala370	K′-β1 loop	Leu482	C-term
Val313	I helix	Thr309	I helix	Met371	K′-β1 loop		
Ile369	K′-β1 loop	Ala370	K′-β1 loop	Glu374	β1−2		
Arg372	K′-β1 loop	Met371	K′-β1 loop	Leu483	C-term		
Glu374	β1−2	Arg372	K′-β1 loop	Gln484	C-term		
Gly481	C-term	Leu373	β1−2				
Leu482	C-term	Glu374	β1−2				
Leu483	C-term	Gly481	C-term				
Heme	Heme	Leu482	C-term				

[a] A summary of the results for all six probes used in the analysis. Amino acid residues that are found to be selective for each structure and the corresponding secondary structure element.

the different structures showed that the homology model had an overall different conformation compared to the crystal structures with several amino acid residues pointing in other directions, resulting in pronounced differences in the secondary structure, above all in the F−G region. The homology model showed more distinct differences in these regions with extended helices and not the same helix−loop−helix motif.

The methodology that combines reactivity and protein−ligand complementarity found in MetaSite performed well and was not influenced by the choice of structure. Further studies to get a more detailed picture of how the different structures interacted with the ligand were done using the docking program GLUE. The achieved docking poses were analyzed with a new approach in which the interaction energy between the ligand and amino acid residues in the active site was calculated. This approach showed very promising results in the comparison of different protein structures without structural alignment. The structures were clearly clustered in the PCA score plot, suggesting that the different protein structures can be separated on the basis of which amino acid residues are relevant for the interaction with the ligand. Interestingly, the comparison of the two CYP2D6 structures identified differences in involvement of the two charged amino acids Glu216 and Asp301. These two residues have previously been extensively studied with site-directed mutagenesis to show their influence on the regioselectivity and substrate binding of CYP2D6 substrates.[24,25] Proteins where these acidic residues have been exchanged with neutral amino acids showed both increased binding constants and a decreased rate of oxidation for some common basic CYP2D6 substrates. In the crystal structure, these two amino acids appeared to be less important for interacting with the ligand, and instead the hydrophobic Phe483 residue had more impact. This residue has been proposed to be involved in substrate binding in homology models,[29] but this hypothesis could not be proved by site-directed mutagenesis and metabolism of the CYP2D6 probe substrate bufuralol.[30] On the other hand, removal of the aromatic character of Phe481 has been shown to impair the binding of CYP2D6 probe substrates.[31] The importance of Phe481 seems to depend on the ligand, and the binding is not influenced by this residue alone but a combination with the acidic residues. As observed in the results of the energy calculation, hydrophobic interactions dominate the binding in the crystal structure, and the identified residues Phe483, Leu213, Leu284, and Phe120 were in agreement with the amino acids that differed, in their position, between the structures. These residues are also responsible for the different shape of the hydrophobic pocket in the crystal structure. In the homology model, the acidic residues appeared to have more influence on binding. When functional studies made on CYP2D6 active-site residues are referred to, the effect of these differences in the two structures is difficult to interpret since both hydrophobic interactions and the electrostatic interactions are needed for binding.

In the case of CYP3A4, many functional studies have been done to explore the influence of the active-site amino acids on substrate binding and catalysis. Most of the amino acid residues investigated in site-directed mutagenesis experiments are important for the regioselectivity and metabolic profile of the substrates. For example, Ser119, Ile301, Phe304, Ala305, Ile369, and Leu373 have been shown to have an impact on the orientation of substrates like testosterone, 7-hexoxycoumarine, and midazolam, demonstrated by changes in the regioselective metabolism.[32−35] Some of these amino acids were identified from the energy calculation to differ between the four structures used in this study. As mentioned above, MetaSite is based on flexible GRID, and the results should be less influenced by these differences. From the energy calculation based on docking poses in the CYP3A4, the four structures were clearly clustered in the PCA score plot, suggesting that different amino acids are important for substrate binding in the different structures. The fact that the PCA results from docking the homogeneous data set of opioid analgesics are comparable with the results from the more diverse data set of CYP3A4 inhibitors is very encouraging, since this diminishes the risk that the results are dependent on the data set used. For the opioid analgesics, the homology model was separated from the crystal structures over the first principal component, which also was observed in the CPCA. This can also be linked to the results from MetaSite where the homology model predicted differently compared to the other structures. That is, the overall different

structure of the homology model is reflected in how the active-site residues are interacting with the ligand (energy calculations), the CPCA, is based on molecular interaction fields and consequently influences the site of metabolism predictions. CPCA is a method based on molecular interaction fields and multivariate data analysis, which was previously used for the structural comparison of protein structures.[23,36] Here, this method was used on the CYP3A4 structure to see if the results agreed with those achieved with the energy calculation method. The interpretation of the CPCA results is not totally objective, and the amount of selective amino acid identified could be biased by the number of points that are selected from the loadings. It is however still a valuable method in identifying amino acids, which could be interesting for further analysis. The residues identified for each protein structure from the energy calculation were also picked in the CPCA, Phe108 and Phe304 in 2j0d; Ala307, Glu374, and Arg212 in 1tqn; and Leu483, Thr309, Glu308, and the heme in the homology model. The score plot from the CPCA and the PCA score plot on the interaction energies also showed that the relative position of the structures was similar. This clearly indicates that the two methods used for structural comparison are comparable to each other. The energy calculation method has clear advantages since it is far simpler to interpret than CPCA and there is no need for protein alignment.

The basic idea of calculating the interaction energies from docking solutions could also be used for other purposes. Docking experiments result in a large amount of docking poses that have to be analyzed by eye. By the approach mentioned in this work, analysis could be simplified by just selecting one conformation/compound from each cluster to analyze, based on the fact that the poses that have similar positions in the score plot also have similar interaction patterns. This can also be used to identify compounds with the same interaction pattern as well as compounds that differ in how they interact with the protein. By inspecting the spreadsheet with the calculated interaction energies, amino acids that show overall high interaction energy can be identified, and this can guide the search for interactions that need to be diminished or increased. The method could then be used to find interactions that determine metabolic properties like enzyme inhibition or compound instability and help to design molecules, which lack these interactions, and in this way guide the design of NCEs away from these issues.

In conclusion, this study showed that the site of metabolism predictions with MetaSite was good for all structures used, and they did not differ from each other except for the CYP3A4 homology model, which had a lower prediction rate than the corresponding crystal structures. The energy calculation approach to analyze the docking results was shown to be a valuable method for structural comparison, and the results were comparable with structural differences in orientation of amino acids in a way similar to results from CPCA analysis. This method could also be very useful in the analysis of dockings and the identification of wanted/unwanted interactions and in this way guide the design of new drugs away from unfavorable interactions with CYP450.

ACKNOWLEDGMENT

We thank Professor Neal Castagnoli, Jr. for his excellent help in interpretation of mass spectra of dextromethorphan and levallorphan. We also thank AstraZeneca R&D, Mölndal for financing this Ph.D. project.

Supporting Information Available: LC−MS/MS data for compounds and major metabolites in the metabolite identification study. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Ekins, S.; de Groot, M. J.; Jones, J. P. Pharmacophore and Three-Dimensional Quantitative Structure Activity Relationship Methods for Modeling Cytochrome p450 Active Sites. *Drug Metab. Dispos.* **2001**, *29*, 936−944.

(2) Boyer, S.; Zamora, I. New Methods in Predictive Metabolism. *Mol. Diversity* **2002**, *5*, 277−287.

(3) Williams, P. A.; Cosme, J.; Sridhar, V.; Johnson, E. F.; McRee, D. E. Microsomal Cytochrome P450 2C5: Comparison to Microbial P450s and Unique Features. *J. Inorg. Biochem.* **2000**, *81*, 183−190.

(4) Wester, M. R.; Yano, J. K.; Schoch, G. A.; Yang, C.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. The Structure of Human Cytochrome P450 2C9 Complexed with Flurbiprofen at 2.0-Å Resolution. *J. Biol. Chem.* **2004**, *279*, 35630−35637.

(5) Williams, P. A.; Cosme, J.; Ward, A.; Angove, H. C.; Matak Vinkovic, D.; Jhoti, H. Crystal Structure of Human Cytochrome P450 2C9 with Bound Warfarin. *Nature* **2003**, *424*, 464−468.

(6) Schoch, G. A.; Yano, J. K.; Wester, M. R.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. Structure of Human Microsomal Cytochrome P450 2C8. Evidence for a Peripheral Fatty Acid Binding Site. *J. Biol. Chem.* **2004**, *279*, 9497−9503.

(7) Williams, P. A.; Cosme, J.; Vinkovic, D. M.; Ward, A.; Angove, H. C.; Day, P. J.; Vonrhein, C.; Tickle, I. J.; Jhoti, H. Crystal Structures of Human Cytochrome P450 3A4 Bound to Metyrapone and Progesterone. *Science* **2004**, *305*, 683−686.

(8) Yano, J. K.; Wester, M. R.; Schoch, G. A.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. 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.

(9) Ekroos, M.; Sjogren, T. Structural Basis for Ligand Promiscuity in Cytochrome P450 3A4. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13682−13687.

(10) Rowland, P.; Blaney, F. E.; Smyth, M. G.; Jones, J. J.; Leydon, V. R.; Oxbrow, A. K.; Lewis, C. J.; Tennant, M. G.; Modi, S.; Eggleston, D. S.; Chenery, R. J.; Bridges, A. M. Crystal Structure of Human Cytochrome P450 2D6. *J. Biol. Chem.* **2006**, *281*, 7614−7622.

(11) Zamora, I.; Afzelius, L.; Cruciani, G. Predicting Drug Metabolism: A Site of Metabolism Prediction Tool Applied to the Cytochrome P450 2C9. *J. Med. Chem.* **2003**, *46*, 2313−2324.

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

(13) Hutchinson, M. R.; Menelaou, A.; Foster, D. J. R.; Coller, J. K.; Somogyi, A. A. CYP2D6 and CYP3A4 Involvement in the Primary Oxidative Metabolism of Hydrocodone by Human Liver Microsomes. *Br. J. Clin. Pharmacol.* **2004**, *57*, 287−297.

(14) Benetton, S. A.; Borges, V. M.; Chang, T. K. H.; McErlane, K. M. Role of Individual Human Cytochrome P450 Enzymes in the in Vitro Metabolism of Hydromorphone. *Xenobiotica* **2004**, *34*, 335−344.

(15) Kirkwood, L. C.; Nation, R. L.; Somogyi, A. A. Characterization of the Human Cytochrome P450 Enzymes Involved in the Metabolism of Dihydrocodeine. *Br. J. Clin. Pharmacol.* **1997**, *44*, 549−555.

(16) Yue, Q. Y.; Sawe, J. Different Effects of Inhibitors on the O- and N-Demethylation of Codeine in Human Liver Microsomes. *Eur. J. Clin. Pharmacol.* **1997**, *52*, 41−47.

(17) Caraco, Y.; Tateishi, T.; Guengerich, F.; Wood, A. Microsomal Codeine N-Demethylation: Cosegregation with Cytochrome P4503A4 Activity. *Drug Metab. Dispos.* **1996**, *24*, 761−764.

(18) De Rienzo, F.; Fanelli, F.; Menziani, M. C.; De Benedetti, P. G. Theoretical Investigation of Substrate Specificity for Cytochromes P450 IA2, P450 IID6 and P450 IIIA4. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 93−116.

(19) Poeaknapo, C.; Fisinger, U.; Zenk, M. H.; Schmidt, J. Evaluation of the Mass Spectrometric Fragmentation of Codeine and Morphine after ^{13}C-Isotope Biosynthetic Labeling. *Phytochemistry* **2004**, *65*, 1413−1420.

(20) Zhou, D.; Afzelius, L.; Grimm, S. W.; Andersson, T. B.; Zauhar, R. J.; Zamora, I. Comparison of Methods for the Prediction of the Metabolic Sites for CYP3A4-Mediated Metabolic Reactions. *Drug Metab. Dispos.* **2006**, *34*, 976−983.

(21) Goodford, P. J. A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules *J. Med. Chem*. **1985**, *28*, 849−857.
(22) Wold, S.; Esbensen, K.; Geladi, P. Principal Component Analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37−52.
(23) Kastenholz, M. A.; Pastor, M.; Cruciani, G.; Haaksma, E. E.; Fox, T. GRID/CPCA: A New Computational Tool to Design Selective Ligands. *J. Med. Chem*. **2000**, *43*, 3033−3044.
(24) Paine, M. J.; McLaughlin, L. A.; Flanagan, J. U.; Kemp, C. A.; Sutcliffe, M. J.; Roberts, G. C.; Wolf, C. R. Residues Glutamate 216 and Aspartate 301 Are Key Determinants of Substrate Specificity and Product Regioselectivity in Cytochrome P450 2D6. *J. Biol. Chem.* **2003**, *278*, 4021−4027.
(25) Guengerich, F. P.; Hanna, I. H.; Martin, M. V.; Gillam, E. M. Role of Glutamic Acid 216 in Cytochrome P450 2D6 Substrate Binding and Catalysis. *Biochemistry* **2003**, *42*, 1245−1253.
(26) *GRID*, version 22; Molecular Discovery Ltd: Ponte San Giovanni, Italy. http://www.moldiscovery.com/docs/grid/ (accessed Feb 2007).
(27) von Moltke, L. L.; Greenblatt, D. J.; Grassi, J. M.; Granda, B. W.; Venkatakrishnan, K.; Schmider, J.; Harmatz, J. S.; Shader, R. I. Multiple Human Cytochromes Contribute to Biotransformation of Dextromethorphan in-Vitro: Role of CYP2C9, CYP2C19, CYP2D6, and CYP3A. *J. Pharm. Pharmacol*. **1998**, *50*, 997−1004.
(28) Lalovic, B.; Phillips, B.; Risler, L. L.; Howald, W.; Shen, D. D. Quantitative Contribution of CYP2D6 and CYP3A to Oxycodone Metabolism in Human Liver and Intestinal Microsomes. *Drug Metab. Dispos.* **2004**, *32*, 447−454.
(29) Modi, S.; Paine, M. J.; Sutcliffe, M. J.; Lian, L. Y.; Primrose, W. U.; Wolf, C. R.; Roberts, G. C. A Model for Human Cytochrome P450 2D6 Based on Homology Modeling and NMR Studies of Substrate Binding. *Biochemistry* **1996**, *35*, 4540−4550.
(30) Smith, G.; Modi, S.; Pillai, I.; Lian, L. Y.; Sutcliffe, M. J.; Pritchard, M. P.; Friedberg, T.; Roberts, G. C.; Wolf, C. R. Determinants of the Substrate Specificity of Human Cytochrome P-450 CYP2D6: Design and Construction of a Mutant with Testosterone Hydroxylase Activity. *Biochem. J*. **1998**, *331*, 783−792.
(31) Hayhurst, G. P.; Harlow, J.; Chowdry, J.; Gross, E.; Hilton, E.; Lennard, M. S.; Tucker, G. T.; Ellis, S. W. Influence of Phenylalanine-481 Substitutions on the Catalytic Activity of Cytochrome P450 2D6. *Biochem. J*. **2001**, *355*, 373−379.
(32) Khan, K. K.; Halpert, J. R. Structure−Function Analysis of Human Cytochrome P450 3A4 Using 7-Alkoxycoumarins as Active-Site Probes. *Arch. Biochem. Biophys*. **2000**, *373*, 335−345.
(33) Khan, K. K.; He, Y. Q.; Domanski, T. L.; Halpert, J. R. Midazolam Oxidation by Cytochrome P450 3A4 and Active-Site Mutants: An Evaluation of Multiple Binding Sites and of the Metabolic Pathway that Leads to Enzyme Inactivation. *Mol. Pharmacol*. **2002**, *61*, 495−506.
(34) Domanski, T. L.; Liu, J.; Harlow, G. R.; Halpert, J. R. Analysis of Four Residues within Substrate Recognition Site 4 of Human Cytochrome P450 3A4: Role in Steroid Hydroxylase Activity and alpha-Naphthoflavone Stimulation. *Arch. Biochem. Biophys.* **1998**, *350*, 223−232.
(35) He, Y. A.; He, Y. Q.; Szklarz, G. D.; Halpert, J. R. Identification of Three Key Residues in Substrate Recognition Site 5 of Human Cytochrome P450 3A4 by Cassette and Site-Directed Mutagenesis. *Biochemistry* **1997**, *36*, 8831−8839.
(36) Afzelius, L.; Raubacher, F.; Karlen, A.; Jorgensen, F. S.; Andersson, T. B.; Masimirembwa, C. M.; Zamora, I. Structural Analysis of CYP2C9 and CYP2C5 and an Evaluation of Commonly Used Molecular Modeling Techniques. *Drug Metab. Dispos*. **2004**, *32*, 1218−1229.

CI600561V