

## **KEY RESIDUES CONTROLLING BINDING OF DIVERSE LIGANDS TO HUMAN CYTOCHROME P450 2A ENZYMES<sup>†</sup>**

**DeVore, N. M., Smith, B. D., Wang, J. L., Lushington, G. H., and Scott, E. E.**

From the Department of Medicinal Chemistry (NMD, BDS, and EES) and the Molecular  
Graphics and Modeling Laboratory (JLW and GHL), University of Kansas, 1251 Wescoe Hall  
Dr., Lawrence, KS 66045

**Running title:** Key ligand binding residues in CYP2A enzymes

**Corresponding author:** Emily E. Scott, Ph.D., Department of Medicinal Chemistry, University of Kansas, 1251 Wescoe Hall Dr., Lawrence, KS 66045, Tel. 785 864-5559; Fax. 785 864 5326; E-mail: [eescott@ku.edu](mailto:eescott@ku.edu)

**Counts:**

Number of text pages: 31

Number of tables: 4

Number of figures: 5

Number of references: 30

Number of words:

Abstract: 249

Introduction: 746

Discussion: 1375

**Non-standard abbreviations:** Cytochrome P450 (CYP)

## Abstract

Although the human lung cytochrome P450 2A13 (CYP2A13) and its liver counterpart cytochrome P450 2A6 (CYP2A6) are 94% identical in amino acid sequence, they metabolize a number of substrates with substantially different efficiencies. In order to determine differences in binding for a diverse set of cytochrome P450 2A ligands, we have determined the spectral binding affinities ( $K_D$ ) for nicotine, phenethyl isothiocyanate (PEITC), coumarin, 2'-methoxyacetophenone (MAP), and 8-methoxypsoralen (MOP). The differences in the  $K_D$  values for CYP2A6 vs. CYP2A13 ranged from 74-fold for 2'-methoxyacetophenone to 1.1-fold for coumarin, with CYP2A13 demonstrating the higher affinity. To identify active site amino acids responsible for the differences in binding of MAP, PEITC, and coumarin, ten CYP2A13 mutant proteins were generated in which individual amino acids from the CYP2A6 active site were substituted into CYP2A13 at the corresponding position. Titrations revealed that substitutions at positions 208, 300, and 301 individually had the largest effects on ligand binding. The collective relevance of these amino acids to differential ligand selectivity was verified by evaluating binding to CYP2A6 mutant enzymes that incorporate several of the CYP2A13 amino acids at these positions. Inclusion of four CYP2A13 amino acids resulted in a CYP2A6 mutant protein (I208S/I300F/G301A/S369G) with binding affinities for MAP and PEITC much more similar to those observed for CYP2A13 than for CYP2A6 without altering coumarin binding. Structure-based quantitative structure-activity relationship (QSAR) using COMBINE successfully modeled the observed mutant-ligand trends and emphasized steric roles for active site residues including four substituted amino acids and an adjacent conserved Leu370.

## Introduction

Xenobiotic-metabolizing P450 enzymes act on chemically diverse small molecules, often with overlapping specificities, but with the formation of distinct metabolites and/or metabolic rates. The only active human cytochromes P450 in the 2A subfamily are CYP2A13 and CYP2A6. CYP2A13 is primarily expressed throughout the respiratory system including nasal mucosa, trachea, and lung (Su et al., 2000b; Zhu et al., 2006), while CYP2A6 is primarily a hepatic enzyme (Yamano et al., 1990; Fernandez-Salguero et al., 1995). Coumarin 7-hydroxylation is a characteristic activity for the P450 2A subfamily. However, reports disagree on whether the catalytic efficiency for coumarin is 10-fold higher for CYP2A6 compared to CYP2A13 (He et al., 2004b) or similar (von Weymarn and Murphy, 2003). A number of other substrates are metabolized very differently by these two enzymes, including nicotine and the nicotine-derived compounds cotinine and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK). CYP2A13 metabolizes nicotine and cotinine with 23- and 15-fold higher catalytic efficiency ( $k_{cat}/K_m$ ) than CYP2A6, respectively (Bao et al., 2005). Additionally, the metabolism of NNK by CYP2A13 occurs at a catalytic efficiency 215-fold greater than CYP2A6 (He et al., 2004a). Both CYP2A13 and CYP2A6 have been implicated in tobacco-related lung cancers and have genetic polymorphisms reported to decrease cancer risk (Tan et al., 2001; Wang et al., 2003).

CYP2A13 also functions in the activation of several other procarcinogens or the metabolism of toxins. Aflatoxin B<sub>1</sub> is efficiently activated by CYP2A13 to a mutagenic epoxide, an activity similar to CYP1A2, but is not activated by CYP2A6 (He et al., 2006). Phenacetin, an antipyretic drug used for many years and later found to be a kidney toxin, is *O*-deethylated

efficiently by both CYP1A2 and CYP2A13 ( $k_{\text{cat}}/K_m$  0.36 for both enzymes), but not by CYP2A6 (Fukami et al., 2007; DeVore et al., 2008).

Diversity in the functions of CYP2A13 and CYP2A6 is not likely to originate with the global structure as the enzymes are 94% identical, differing by only 32 amino acids and the crystallographically-determined structures have an overall root mean square deviation for the C $\alpha$  carbons of only 0.5 Å (Smith et al., 2007). Ten of the 32 amino acid differences are in or near the active site and are likely to be directly related to the functional differences observed in substrate metabolism. We have previously established that differential binding and metabolism of the analgesic phenacetin in human CYP2A enzymes is the result of the disparate amino acids at positions 208, 300, 301, and 369 in concert with the effects of those four substitutions on adjacent conserved residues Phe209 and Leu370 (DeVore et al., 2008). The net steric effects of these differences alter active site morphology to determine CYP2A compatibility with phenacetin binding and *O*-deethylation. Other studies of human CYP2A enzymes have revealed that residues 117 and 372 are critical for both coumarin (He et al., 2004b) and NNK metabolism, and that the residue at position 208 is also critical for NNK metabolism (He et al., 2004a).

The goal of the present work is to determine if these same amino acids are responsible for differential binding of other ligands by human CYP2A enzymes. In several cases, other P450 proteins are known to adapt to different ligands in a more induced-fit mode of binding, frustrating attempts at prediction of drug binding and metabolism. If the relatively small CYP2A selectivity is based on a more steric-dominated, lock-and-key type model, then xenobiotic binding to these enzymes may be suitably evaluated using docking methodologies based on the known crystal structures. In this study, we have surveyed a series of five structurally diverse cytochrome P450 2A ligands, nicotine, phenethyl isothiocyanate (PEITC), coumarin, 2'-

methoxyacetophenone (MAP), and 8-methoxypsoralen (MOP). Coumarin, PEITC, and MAP were identified to have a range of differential affinities for the human CYP2A enzymes and were thus used to investigate the effects of mutations at ten active site positions differing between CYP2A6 and CYP2A13. These ligand binding studies identified a number of substitutions that changed the  $K_D$  significantly for at least one substrate. Substitutions combining three or four of these CYP2A13 amino acid residues into CYP2A6 indicate that the differences between CYP2A13 and CYP2A6 binding affinities observed for the ligands in this study are largely dependent on the same four residues identified in the phenacetin study: 208, 300, 301, and 369 (DeVore et al., 2008). In order to evaluate the roles that individual features of the active site have in modulating ligand binding affinity, we performed structure-based quantitative structure-activity relationship (QSAR) analysis via the COMBINE method, whose results were consistent with a steric binding rationale.

## Methods

*Materials:* Coumarin, phenethyl isothiocyanate (PEITC), 8-methoxypsoralen (MOP), (*S*)-nicotine, and 2'-methoxyacetophenone (MAP) were purchased from Sigma Aldrich (St. Louis, MO).

*Protein expression and purification:* The human CYP2A13 and CYP2A6 proteins used in these studies were all generated by deleting the N-terminal transmembrane sequence ( $\Delta 2-23$ ), altering several residues at the modified N-terminus (from 24-WRQRKSR-30 to 24-AKKTSSK-30) and adding four histidine residues at the C-terminus. All mutations were also made in this background. Many studies with several different membrane P450 enzymes have shown that truncated versions metabolize the same substrates as the full-length parent enzyme with the same regio- and stereoselectivity (von Wachenfeldt et al., 1997; Scott et al., 2001; Schoch et al., 2004; Wester et al., 2004). Given substantial differences in the availability and purity of full-length and truncated enzymes and differences in the requirement for lipid, direct comparisons of activity rates are difficult, but similar activities and binding have been reported (Wester et al., 2004; Yano et al., 2004). Use of the truncated versions allow a far greater yield of protein than could be attained with the full-length enzyme (Scott et al., 2001). These modifications allow generation of highly purified protein that retains its catalytic activity and in quantities that are sufficient for ligand binding and other biophysical studies. Such truncated CYP2A proteins were expressed with yields ranging from 320 – 1100 nmol/liter of culture at the microsome stage and further purified as described (Smith et al., 2007). The total yield of purified protein ranged from 17.4–155 nmol P450/liter of *E. coli* culture. Two mutant proteins, CYP2A6 I208S/I300F/G301A and CYP2A6 I208S/I300F/G301A/S369G had specific contents of  $\geq 3$ -3.5 nmol/mg, but all other

proteins had specific contents ranging from 10.4 – 19.3 nmol/mg. These CYP2A proteins were competent in the metabolism of phenacetin as previously reported (DeVore et al., 2008).

*Site-directed mutagenesis:* The mutations were generated using the modified CYP2A gene in an expression vector (pKK2A13dH) or (pKK2A6dH) as a template (13) and the Quikchange® Site-Directed Mutagenesis method (Stratagene, La Jolla, CA). Synthetic oligonucleotides were designed to yield the desired amino acid substitution, but also to contain silent restriction site mutations where possible to facilitate identification of mutated genes (Table 1). Oligonucleotides were synthesized by Genosys (Woodlands, TX).

*Ligand binding titrations:* Binding titrations with the ligands were conducted with purified P450 protein at 20°C using a UV-2101 UV-visible scanning spectrophotometer (Shimadzu Scientific Instruments, Inc., Columbia, MD). Protein was diluted to 1 µM in 100 mM potassium phosphate buffer, pH 7.4. Diluted protein was equally divided between two 1.0 ml quartz cuvettes (1-cm path length), and a baseline was recorded (300 to 500 nm). Freshly prepared aliquots of ligand dissolved in 100% ethanol were added to the sample cuvette. An equal volume of ethanol was added to the reference cuvette. Difference spectra were collected (300 to 500 nm) after an equilibration period. During the spectral titrations, the total amount of ethanol added did not exceed 2%. Binding to P450 was monitored as the absorbance difference ( $\Delta A$ ) between the minimum (~420 nm) and the maximum (~385 nm). The apparent binding constant ( $K_D$ ) and the maximum spectral change ( $\Delta A_{\max}$ ) were determined from

nonlinear least-squares regression fitting to the following equation:

$$\Delta A = \frac{\Delta A_{\max}}{2P} [P + S + K_D - \sqrt{(P + S + K_D)^2 - 4PS}]$$

where P is total P450 concentration and S is total

ligand concentration. Nonlinear regression was accomplished using Graphpad Prism 4 (Graphpad Software, San Diego, CA).



*Computational Methods:* In order to elucidate the role that individual residues play in determining the relative affinities of ligands as a function of active site structure, we performed structure-based quantitative structure-activity relationship (QSAR) analysis via the COMBINE method (Ortiz et al., 1995). Protein structures were generated from crystal structures of human CYP2A13 with indole bound (2P85; (Smith et al., 2007)) and human CYP2A6 complexed with methoxsalen (PDB code 1Z11; (Yano et al., 2005)). Mutations of CYP2A13 were performed *in silico* via the Biopolymer suite in Insight-II (Accelrys, San Diego, CA) and each unique mutant protein was allowed to relax via CHARMM simulations (Brooks et al., 1983) entailing a 100 step molecular mechanics preconditioning run to alleviate clashes induced by the point mutation, a 20 ps molecular dynamics (MD) warmup, a 20 ps thermal equilibration and a 100 ps analysis run. In all cases, the native ligand (i.e., indole and methoxsalen for CYP2A13 and CYP2A6 respectively) and crystallographic waters were retained throughout the simulation. All atomic charge and force field parameters corresponded to standard CHARMM terms (Brooks et al., 1983). For each unique protein, the conformation present at the end of the analysis run was used for ligand docking and subsequent COMBINE analysis.

In order to derive a bound conformer for coumarin bound to CYP2A6, the known crystal structure of this complex (PDB code 1Z10 (Yano et al., 2005)) was aligned to the relaxed CYP2A6 structure derived in the previous step. This alignment was effected via the Biopolymer module in SYBYL (version 8.0, Tripos Inc., St Louis, MO) with extra weight being assigned to alignment of the heme moiety (to focus the alignment within the active site region). For all other ligand-protein complexes, the relevant ligand (coumarin, MAP, or PEITC) was docked to each of the distinct proteins (CYP2A6, CYP2A13 and the L110V, A117V, S208I, A213S, F300I, A301G, M365V, L366I, G369S and H372R mutational variants of CYP2A13) via FlexX (Rarey

et al., 1996). In each case the protein structure was stripped of the native ligand and all waters except the one closest to the Asn297 side chain (due to observations in the CYP2A13 crystal structure that it served to bridge between indole and the Asn 297 side chain). Fifty poses were then requested per ligand, using default charge and conformational search setting. For mutant protein CYP2A13 A117V, docking simulations for the ligands MAP and coumarin did not achieve a reasonable bound structure, thus the initial guess positions of these two ligands were derived by aligning to the conformations attained for the CYP2A13 protein.

The set of COMBINE descriptors employed for QSAR analysis constituted of van der Waals (vdW) and electrostatics interaction terms between each ligand and each residue (i.e., including all amino acids, plus the heme and the one water) present in each of the 12 proteins. The QSAR model was then trained by partial least squares fitting as implemented in the Simca P program (Umetrics AB, Umea, Sweden) of a weighted linear combination of the COMBINE electrostatic and van der Waals parameters to experimental  $pK_D$  ( $= \log[K_D]$ ) data. By default, the top scoring ligand pose for each ligand/enzyme pair was selected for analysis, however in instances where substantial discrepancies were observed between the experimental  $pK_D$  value and that computed from leave-one-out cross-validated correlation analysis, lower-scoring poses were tested for improved fidelity. In the four cases where the top-scoring ligand conformer did not correlate well with experiment (coumarin and MAP binding to H372R and L366I), none of the lower scoring conformers achieved marked improvement, thus these cases were discarded as outliers. Overall, the use of van der Waals and electrostatic interaction parameters yielded a single COMBINE QSAR model that was reasonably able to predict the experimental binding affinities for 32 of the 36 possible combinations of the three ligands (MAP, PEITC, or coumarin)

with CYP2A6, CYP2A13, and the ten different CYP2A13 mutants ( $R^2 = 0.87$ ,  $Q^2_{\text{Loo}} = 0.60$ ; three components).

## Results

*Binding of structurally diverse ligands to CYP2A6 and CYP2A13 proteins:* Initially a series of five structurally diverse cytochrome P450 2A ligands were surveyed and binding affinities determined for both CYP2A6 and CYP2A13 enzymes. Coumarin was chosen because it is an often-used marker substrate for cytochromes P450 in the 2A subfamily (Fernandez-Salguero et al., 1995; He et al., 2004b; Kim et al., 2005). Nicotine (Bao et al., 2005; Murphy et al., 2005) and 2'-methoxyacetophenone (MAP) (Su et al., 2000a; von Weymarn et al., 2005) are also CYP2A substrates, while 8-methoxypsoralen (von Weymarn et al., 2005) and phenethyl isothiocyanate (PEITC) (von Weymarn et al., 2006) are inhibitors. Each of these ligands displayed type I binding spectra with a shift from ~420 nm to ~385 nm (Figure 1), indicating an increase in the high-spin, 5-coordinate state of the iron as water bound to the heme iron in the resting state is displaced upon ligand binding in the active site.

All five of the ligands surveyed bound more tightly to CYP2A13 than to CYP2A6 (Table 2), but with a substantial range in both overall affinities and in the differences in affinity between the two enzymes. The ligand that bound the tightest to both enzymes is 8-methoxypsoralen. In contrast, nicotine had the lowest affinity for both enzymes. MAP, PEITC, and coumarin had intermediate  $K_D$  values.

In terms of differential affinities, the compounds fall into three groups. First, coumarin binds equally well to both enzymes. Second, nicotine, 8-MOP, and PEITC have moderate selectivity for CYP2A13. The substrate nicotine binds to CYP2A13 4.7 times more tightly than to CYP2A6, while the inhibitors PEITC and MOP bind to CYP2A13 13- to 14-fold more tightly than to CYP2A6. The ligand with the largest difference in binding affinity is MAP, which binds 74-fold more tightly to CYP2A13 than to CYP2A6.

*Ligand binding to mutants of CYP2A13:* Of the five ligands surveyed above, MAP, PEITC, and coumarin were chosen to characterize mutant CYP2A proteins in which active site amino acids from the opposite CYP2A enzyme were substituted. These three ligands were chosen because they span the range of differences observed in  $K_D$  values, from MAP with its 74-fold differences in binding affinity, to PEITC with a moderate 14-fold difference in affinity, to coumarin which binds equally well to both proteins. The ten amino acids that were mutated were selected based on their location within the CYP2A6 and CYP2A13 active sites (Figure 2) and their effects on phenacetin binding and metabolism (DeVore et al., 2008).

MAP binding to the single site CYP2A13 mutant proteins revealed that four modifications resulted in substantially different  $K_D$  values compared to the parent CYP2A13 enzyme (Table 3). Substitution of Ile for Phe300 (F300I) resulted in a 15-fold increase in  $K_D$ , S208I had a 12-fold increase, and A301G 8-fold increase in  $K_D$ . A117V had a more moderate 4.1-fold decrease in affinity for MAP. The remaining six mutations had relatively little effect on MAP binding, although L366I improved binding affinity almost two-fold.

Binding of PEITC to CYP2A13 is characterized by a  $K_D$  of 0.43  $\mu$ M, approximately 14-fold higher affinity than CYP2A6 (Table 3). The single substitution of F300I in the CYP2A13 protein resulted in a 14-fold loss in PEITC affinity, such that its  $K_D$  is essentially that of CYP2A6. The mutation A301G, located directly adjacent to F300I, demonstrated a 7-fold increased  $K_D$ . The CYP2A13 mutants S208I, M365V, and G369S had  $K_D$  values that were increased 3 to 4-fold compared to the CYP2A13. The remaining single mutants yielded smaller changes in the PEITC binding affinity. Again, CYP2A13 L366I was the only protein to demonstrate an increase in binding affinity for PEITC, though the change was small.

In contrast to PEITC and MAP, which bind more tightly to CYP2A13 than CYP2A6, the  $K_D$  value for coumarin is  $\sim 2.9 \mu\text{M}$  for both CYP2A13 and CYP2A6. Again, the F300I substitution had the most dramatic effect on  $K_D$ , with a 5-fold increase. Two other substitutions, G369S and S208I, increased  $K_D$  values by 2- and 3-fold, respectively. The substitution H372R increased coumarin binding affinity by nearly 50% to a  $K_D$  of  $1.5 \mu\text{M}$ , while L366I increased coumarin binding affinity by 5-fold to  $0.57 \mu\text{M}$ .

*Ligand Binding to CYP2A6 Multiple Mutants:* The above studies revealed amino acids that decreased CYP2A13 affinity for MAP and PEITC (Figure 3), but it is much easier to disrupt ligand binding than to improve it. To verify that these same amino acids would increase CYP2A6 binding affinity for MAP and PEITC, several of these substitutions were incorporated into CYP2A6 and binding of the same three ligands was evaluated. Since mutations at 208, 300, and 301 had the largest negative effects on MAP and PEITC affinity for CYP2A13, the CYP2A6 triple mutant I208S/I300F/G301A was generated. This mutated protein had affinities for MAP and PEITC that were intermediate between CYP2A6 and CYP2A13 (Table 3). The affinity for MAP was increased by the most, but coumarin affinity was also decreased by two-fold. Since the effect of the substitution at position 208 was much greater for MAP and the effect of substitution at position 369 was greater for PEITC, a second CYP2A6 triple mutant was examined that consisted of the I300F/G301A/S369G substitutions. Compared to the previous CYP2A6 mutant, this protein had essentially the same affinity for MAP but greater affinity for PEITC. Finally, combination of these substitutions in the quadruple CYP2A6 mutant I208S/I300F/G301A/S369G yielded additional increases in affinity for MAP without further altering PEITC binding and maintaining the approximate  $K_D$  for coumarin.

*COMBINE Results:* The ability to engineer into CYP2A6 the capacity to bind phenacetin (DeVore et al., 2008) and the application of these same mutations to rationally increase CYP2A6 affinity for PEITC and MAP suggest that it may be possible to treat the CYP2A active site as a relatively static active site for docking ligands and predicting ligand binding affinities. In order to test this hypothesis, receptor-based QSAR studies were undertaken using the comparative binding energy (COMBINE) approach to identify and quantify individual protein/ligand interactions that contribute to the differing affinities for the three ligands with the different proteins under study (Figure 3). As described in Methods, the result was a model in which contributions from van der Waals and electrostatic interactions could be used to reasonably predict the observed binding affinities in most cases. The agreement between the binding affinities observed experimentally and those predicted by the COMBINE model had a strong correlation ( $R^2 = 0.87$ ; Figure 4). Thus, this model reproduces most of the observed affinity trends for the three different ligands across the 12 CYP2A parent and mutant enzymes under consideration (Table 4). It successfully predicts the two enzyme-ligand complexes with the lowest affinity (MAP interacting with CYP2A6 and with CYP2A13 F300I, respectively) and only narrowly interchanges the order of the next two (coumarin binding to CYP2A13 F300I and MAP binding to CYP2A13 S208I). It also correctly identifies the enzyme-ligand complex with the greatest affinity (PEITC interacting with CYP2A13 L366I), and yields predictions that differ by no more than half a logarithmic unit for 32 of the 36 ligand/enzyme complexes examined experimentally.

Analysis of the electrostatic features that contribute to the binding affinity trends reveal that by far the most significant are favorable coupling with the iron and porphyrin portions of the heme and the water molecule that putatively facilitates H-bond bridging to the Asn297 side chain

amide NH<sub>2</sub>. Each of the three ligands have a polar end that either binds to the water or to Asn297 directly, and a more hydrophobic end that couples with the heme via a loose cation- $\pi$  interaction. The strongly binding CYP2A13/PEITC conformer places its aromatic group in a position that is much more suitable for favorable interactions with the Fe than does CYP2A6/MAP (Figure 5A), which likely is a key factor behind the stronger activity of the former. Other electrostatic interactions are fairly modest in importance and are located proximal to the heme or fairly distant from the ligand ( $> 6.0$  Å) and thus have only small effects on the overall binding affinity predictions. In addition, almost all of the top-ten electrostatic contributors are conserved between CYP2A6 and CYP2A13 and thus are unlikely to discriminate in terms of ligand binding selectivity.

In terms of van der Waals contributions, the largest contribution to binding is also derived from the active site water molecule, but other important van der Waals contacts include amino acids that differ between the two enzymes or that are conserved but have notable interactions. In particular, the nonconserved Phe300, Ala301, and Ala117 are three of the top six most important protein/ligand interactions. In this case, all but one of these top six protein/ligand interactions confer a small penalty to ligand binding. From this, one would argue that substitutions of CYP2A6 residues that further occlude the volume of the active site (e.g., A117V, F300I) should generally lead to poorer ligand binding. This consistently appears to be the case, with the exception of coumarin binding to A117V. In this situation, docking reveals that coumarin is still small enough in proximity to position 117 as to engage in a favorable lipophilic interaction with the larger valine side chain without crossing over into the steric clash domain. One would conversely expect that mutations that substitute smaller amino acid side chains within the active site in place of larger ones would lead to improved ligand activity. In



the case of A301G, however, this is universally not the case. However, the modeling results suggest that while the alanine to glycine mutation itself opens up space in the active site closer to the polar end of the ligands, there is a slight shift of the I helix backbone and Phe300 toward the ligand that actually leads to slightly greater constriction in the area of the lipophilic group (Figure 5B). Comparison of the aligned CYP2A6 and CYP2A13 experimentally determined crystal structures, reveals a slight helical kinking that arises in the proximity of residue 301 that occludes the cavity by about  $\sim 0.2$  Å in the case of CYP2A6.

Finally, conserved amino acids Asn297, Leu370, and Phe480 also have significant steric roles in this model. Asn297, which is considered to be the lone available H-bonding site in the CYP2A active sites, has the second largest contribution for an amino acid. The one favorable protein steric contribution to ligand binding in the COMBINE model is the conserved amino acid Leu370, which is located near ligand hydrophobic groups. In this case, the best interpretation of its status as a favorable contact is that mild steric clashes with ligands are compensated for with favorable entropy, desolvation energy, or some combination of the two. The benefits of desolvation are obvious in that in all complexes the ligand places a hydrophobic group in the vicinity of Leu370's hydrophobic side chain. Entropy is also a plausible contributor in that Leu370 is easily the most mobile side chain observed in the COMBINE modeling studies. The extent of this variability in this side chain in our modeling studies is evident in Figure 5C, wherein the Leu370 side chain location and conformation is depicted for a series of relevant proteins. Leu370 is also located between Leu366 and His372. Explaining of the effects of mutations of these two residues (CYP2A13 L366I and H372R) with MAP and coumarin constitute the greatest shortcoming of the COMBINE model.

## Discussion

Examination of these five structurally diverse cytochrome P450 2A ligands revealed type I binding and a very large range in the affinities of CYP2A13 vs. CYP2A6. At the extremes, coumarin bound equally well to both proteins, while MAP bound more than 74 times more tightly to CYP2A13 than to CYP2A6.

Coumarin binds to CYP2A6 by hydrogen bonding with Asn297, adopting an orientation within a largely planar active site that positions C7 for hydroxylation (Yano et al., 2005), consistent with the only coumarin metabolite detected for CYP2A6. Asn297 and the planar active site are conserved in CYP2A13 (13), as is the generation of 7-hydroxycoumarin. Consistent with these similarities, the nearly identical  $K_D$  values for both proteins suggest that coumarin binds similarly in CYP2A13. However, CYP2A13 also forms a significant amount of 3,4 epoxide and smaller amounts of 6-hydroxy- and 8-hydroxycoumarin (von Weymarn and Murphy, 2003). This suggests that CYP2A13 binds coumarin via N297 hydrogen bonding, but is less constrained in its approach to the activated oxygen on the heme so that C6 or C8 can also be hydroxylated and also binds coumarin in a second, inverted orientation to expose C3/C4 for epoxidation.

MAP is *O*-demethylated by both enzymes to produce 2'-hydroxyacetophenone and formaldehyde, and its binding showed the largest difference in affinity among the ligands examined. Although kinetics are not available for MAP metabolism, Su *et al.* reported that at 1 mM substrate concentration, CYP2A13 metabolized MAP at a rate 4-fold faster than CYP2A6 (Su et al., 2000a). Thus, the scale of the differences in MAP affinity were unexpected based on the single productive orientation in the active site suggested by a single product and the preliminary metabolic rate information available.

The remaining ligands had smaller intermediate differences in affinity. 8-methoxypsoralen and PEITC are both mechanism-based inactivators (Koenigs et al., 1997; von Weymarn et al., 2005; von Weymarn et al., 2006) and had similar ~14- fold greater affinities for CYP2A13. Because the reactive metabolites are not known, it is difficult to predict the orientation of these ligands in the active site. (*S*)-Nicotine demonstrated ~5-fold preference for CYP2A13. Both CYP2A enzymes hydroxylate nicotine at the pyrrolidine 5' position and the methyl, though the latter is generated more readily by CYP2A13 (Murphy et al., 2005). Thus, nicotine likely binds similarly in both enzymes, orienting either the 5' or methyl carbons for oxidation, though the latter may be slightly favored in the case of CYP2A13.

Site-directed mutagenesis and binding titrations with MAP, PEITC, and coumarin revealed several amino acids responsible for disparate ligand  $K_D$  values. The three substitutions with the largest effects were F300I, A301G, and S208I. CYP2A13 F300I yielded the largest changes in ligand binding, 5- to 18-fold decreases for all three ligands (Figure 3), suggesting that this residue plays a central role in ligand binding. This is consistent with both the steric role suggested by the COMBINE model and a recent crystal structure of the CYP2A6 I208S/I300F/G301A/S369G quadruple mutant (DeVore et al., 2008). Comparison of the CYP2A structures reveal that I300 and F300 overlap at the alpha carbon atom, but that the phenyl ring is torsioned farther away from the bound ligand than the isoleucine side chain in both CYP2A13 and the CYP2A6 quadruple mutant (Yano et al., 2005; Smith et al., 2007; DeVore et al., 2008). The amino acid at this position is key for phenacetin (DeVore et al., 2008) and indole metabolism (Wu et al., 2005) in CYP2A enzymes. Overall, it is clear that this residue is critical for the higher affinity of CYP2A13 for PEITC and MAP, playing a key role in substrate selectivity.

Substitution of the adjacent CYP2A13 residue, A301, to glycine removes the side chain and causes 7- to 8-fold increases in the  $K_D$  for MAP and PEITC, respectively, but does not significantly alter the  $K_D$  for coumarin. This suggests that coumarin interacts differently with the residue at position 300 than does MAP or PEITC. Since MAP and PEITC bind more tightly to CYP2A13, interactions with this residue also contribute to ligand selectivity.

The S208I substitution has a 14-fold increase in the  $K_D$  for MAP, but much smaller ~3-fold effects on PEITC and coumarin binding. Since MAP is the ligand with the most contrast in affinity between the two CYP 2A enzymes, interactions with this residue may also be significant in ligand selectivity. The more moderate effects on coumarin binding are consistent with the results of He *et al.*, which showed that this mutation had little effect on the kinetics of coumarin 7-hydroxylation by CYP2A13 (He et al., 2004b). This mutation is also reported to decrease the catalytic efficiency of CYP2A13 NNK metabolism to both the keto aldehyde and keto alcohol products, neither of which are formed efficiently by CYP2A6 (He et al., 2004a).

COMBINE evaluation of the influence of electrostatics revealed that the extended heme system is found to exert the dominant influence, favoring ligands that orient negatively charged groups toward the positively charged iron. The key non-heme electrostatic interaction is predicted to be with a water molecule that was isolated as an apparent H-bond donor in the CYP2A13/indole crystal structure (Smith et al., 2007) and may help to bridge the ligand with the conserved Asn297 donor site, but whose position was not rigorously conserved during molecular dynamics relaxation of all mutants. Thus no significant electrostatic differences were identified between CYP2A6 and CYP2A13 that are likely to directly affect ligand binding. In contrast, the COMBINE model identified several differing amino acids at positions 300, 301, and 117 that do have important steric contributions. All three residues have substantial effects on not only the

ligand binding reported here, but also on phenacetin binding and metabolism. The CYP2A13 A117V increased phenacetin affinity 2-fold and activity by >5-fold, while the F300I and A301G mutations showed minimal phenacetin binding or metabolism. Structures of the CYP2A6 I208S/I300F/G301A/S369G quadruple mutant in complex with phenacetin indicated steric differences related to the positioning of the amino acids at 300 and 301 that are also likely relevant to the binding of the ligands examined here.

Analysis of the CYP2A6 and CYP2A13 crystal structures reveals that the difference in the position of Leu370 is a root mean squared deviation (RMSD) of 0.74 Å, as opposed to the global residue average of only 0.51 Å. Substantial further variation was also found when examining the different CYP2A13 mutants computationally in this study. This flexibility of Leu370 observed in these docking studies is further substantiated by the structure of the CYP2A6 I208S/I300F/G301A/S369G quadruple mutant (DeVore et al., 2008) which demonstrates that the position of Leu370 is altered in response to protein mutations, in this case probably primarily by the adjacent S369G mutation. Leu370 may well also be capable of differential adaptation to the presence of different ligands in ways that permit the alleviation of steric clashes. It is possible that a primary cause of the mutagenic influence of positions 366, 369, and 372 may be the effect of mutation on the conformation of Leu370, some of which were modeled poorly by the COMBINE model.

Combination of I300F, G301A, and I208S mutations in the CYP2A6 background was sufficient to substantially increase the affinity for MAP, but not for PEITC. Because 1) substitution at 369 had the next largest effects on PEITC binding to CYP2A13, 2) modeling suggested effects of mutation at 369 on the position of Leu370, and 3) this substitution made important contributions to the binding and metabolism of phenacetin, G369S was also

incorporated into CYP2A6, resulting in an enzyme whose overall binding affinities for all three ligands were much more similar to CYP2A13.

In conclusion, the goal of this study was to gain insight into the differences in structure-binding relationships between the human cytochrome CYP2A6 and CYP2A13 enzymes and their diverse ligands. Binding titrations to CYP2A13 and CYP2A6 mutant proteins in conjunction with QSAR identified key active site amino acids at positions 208, 300, 301, and 369. The development of a COMBINE QSAR model that substantially fit the trends observed in the experimental results suggests the differing CYP2A residues play primarily steric roles in modulating ligand affinity. Difficulties with predictions for a small subset of mutations adjacent to Leu370 likely occur from demonstrated variations in the positioning of Leu370. Thus, apart from the mobility of Leu370, these results suggest that the CYP2A active site can reasonably be treated as a steric docking problem for prediction of ligand affinities.

## Acknowledgments

Thanks are due to a number of students involved in constructing the mutants described in this work (Jenny Morrison, Chad Schroeder, Kyle Bailey, Matthew Axtman, and Melanie Blevins) and to Christopher Wood for assistance with purification.

## References

- Bao Z, He XY, Ding X, Prabhu S and Hong JY (2005) Metabolism of nicotine and cotinine by human cytochrome P450 2A13. *Drug Metabolism and Disposition* **33**:258-261.
- Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S and Karplus M (1983) CHARMM: A program for macromolecular energy, minimization and dynamics calculations. *J. Comp. Chem.* **4**:187-217.
- DeVore NM, Smith BD, Urban MJ and Scott EE (2008) Key Residues Controlling Phenacetin Metabolism by Human Cytochrome P450 2A Enzymes. *Drug Metabolism and Disposition: The Biological Fate of Chemicals* **36**:2582-2590.
- Fernandez-Salguero P, Hoffman SMG, Cholerton S, Mohrenweiser H, Raunio H, Rautio A, Pelkonen O, Huang JD, Evans WE, Idle JR and Gonzalez FJ (1995) A Genetic-Polymorphism in Coumarin 7-Hydroxylation - Sequence of the Human Cyp2a Genes and Identification of Variant Cyp2a6 Alleles. *American Journal of Human Genetics* **57**:651-660.
- Fukami T, Nakajima M, Sakai H, Katoh M and Yokoi T (2007) CYP2A13 metabolizes human CYP1A2 substrates, phenacetin and theophylline. *Drug Metab. Dispos.* **35**:335-339.
- He XY, Shen J, Ding X, Lu AY and Hong JY (2004a) Identification of critical amino acid residues of human CYP2A13 for the metabolic activation of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone, a tobacco-specific carcinogen. *Drug Metab. Dispos.* **32**:1516-1521.
- He XY, Shen J, Hu WY, Ding X, Lu AY and Hong JY (2004b) Identification of Val(117) and Arg(372) as critical amino acid residues for the activity difference between human CYP2A6 and CYP2A13 in coumarin 7-hydroxylation. *Arch. Biochem. Biophys.* **427**:143-153.
- He XY, Tang L, Wang SL, Cai QS, Wang JS and Hong JY (2006) Efficient activation of aflatoxin B1 by cytochrome P450 2A13, an enzyme predominantly expressed in human respiratory tract. *Int. J. Cancer.* **118**:2665-2671.
- Kim D, Wu ZL and Guengerich FP (2005) Analysis of coumarin 7-hydroxylation activity of cytochrome P450 2A6 using random mutagenesis. *Journal of Biological Chemistry* **280**:40319-40327.
- Koenigs LL, Peter RM, Thompson SJ, Rettie AE and Trager WF (1997) Mechanism-based inactivation of human liver cytochrome P450 2A6 by 8-methoxypsoralen. *Drug Metab Dispos* **25**:1407-1415.
- Murphy SE, Raulinaitis V and Brown KM (2005) Nicotine 5'-oxidation and methyl oxidation by P450 2A enzymes. *Drug Metab Dispos* **33**:1166-1173.
- Ortiz AR, Pisabarro MT, Gago F and Wade RC (1995) Prediction of drug binding affinities by comparative binding energy analysis. *Journal of Medicinal Chemistry* **38**:2681-2691.
- Rarey M, Kramer B, Lengauer T and Klebe G (1996) A fast flexible docking method using an incremental construction algorithm. *Journal of Molecular Biology* **261**:470-489.
- Schoch GA, Yano JK, Wester MR, Griffin KJ, Stout CD and Johnson EF (2004) Structure of human microsomal cytochrome P450 2C8. Evidence for a peripheral fatty acid binding site. *J Biol Chem* **279**:9497-9503.
- Scott EE, Spatzenegger M and Halpert JR (2001) A truncation of 2B subfamily cytochromes P450 yields increased expression levels, increased solubility, and decreased aggregation while retaining function. *Arch Biochem Biophys* **395**:57-68.



- Smith BD, Sanders JL, Porubsky PR, Lushington GH, Stout CD and Scott EE (2007) Structure of the human lung cytochrome P450 2A13. *J. Biol. Chem.* **282**:17306-17313.
- Su T, Bao Z, Zhang QY, Smith TJ, Hong JY and Ding X (2000a) Human cytochrome P450 CYP2A13: predominant expression in the respiratory tract and its high efficiency metabolic activation of a tobacco-specific carcinogen, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Cancer Res.* **60**:5074-5079.
- Su T, Bao ZP, Zhang QY, Smith TJ, Hong JY and Ding XX (2000b) Human cytochrome p450 CYP2A13: Predominant expression in the respiratory tract and its high efficiency metabolic activation of a tobacco-specific carcinogen, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Cancer Research* **60**:5074-5079.
- Tan W, Chen GF, Xing DY, Song CY, Kadlubar FF and Lin DX (2001) Frequency of CYP2A6 gene deletion and its relation to risk of lung and esophageal cancer in the Chinese population. *International Journal of Cancer* **95**:96-101.
- von Wachenfeldt C, Richardson TH, Cosme J and Johnson EF (1997) Microsomal P450 2C3 is expressed as a soluble dimer in Escherichia coli following modification of its N-terminus. *Arch Biochem Biophys* **339**:107-114.
- von Weyarn LB, Chun JA and Hollenberg PF (2006) Effects of benzyl and phenethyl isothiocyanate on P450s 2A6 and 2A13: potential for chemoprevention in smokers. *Carcinogenesis* **27**:782-790.
- von Weyarn LB and Murphy SE (2003) CYP2A13-catalysed coumarin metabolism: comparison with CYP2A5 and CYP2A6. *Xenobiotica* **33**:73-81.
- von Weyarn LB, Zhang QY, Ding X and Hollenberg PF (2005) Effects of 8-methoxypsoralen on cytochrome P450 2A13. *Carcinogenesis* **26**:621-629.
- Wang H, Tan W, Hao B, Miao X, Zhou G, He F and Lin D (2003) Substantial reduction in risk of lung adenocarcinoma associated with genetic polymorphism in CYP2A13, the most active cytochrome P450 for the metabolic activation of tobacco-specific carcinogen NNK. *Cancer Research* **63**:8057-8061.
- Wester MR, Yano JK, Schoch GA, Yang C, Griffin KJ, Stout CD and Johnson EF (2004) The structure of human cytochrome P450 2C9 complexed with flurbiprofen at 2.0-Å resolution. *J Biol Chem* **279**:35630-35637.
- Wu ZL, Podust LM and Guengerich FP (2005) Expansion of substrate specificity of cytochrome P450 2A6 by random and site-directed mutagenesis. *Journal of Biological Chemistry* **280**:41090-41100.
- Yamano S, Tatsuno J and Gonzalez FJ (1990) The CYP2A3 gene product catalyzes coumarin 7-hydroxylation in human liver microsomes. *Biochemistry* **29**:1322-1329.
- Yano JK, Hsu MH, Griffin KJ, Stout CD and Johnson EF (2005) Structures of human microsomal cytochrome P450 2A6 complexed with coumarin and methoxsalen. *Nat Struct Mol Biol* **12**:822-823.
- Yano JK, Wester MR, Schoch GA, Griffin KJ, Stout CD and Johnson EF (2004) The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-Å resolution. *J Biol Chem* **279**:38091-38094.
- Zhu LR, Thomas PE, Lu G, Reuhl KR, Yang GY, Wang LD, Wang SL, Yang CS, He XY and Hong JY (2006) CYP2A13 in human respiratory tissues and lung cancers: an immunohistochemical study with a new peptide-specific antibody. *Drug Metabolism and Disposition: The Biological Fate of Chemicals* **34**:1672-1676.

## Footnotes

<sup>†</sup>This work was supported by the National Institutes of Health [Grant GM076343 (EES)].

Portions of this work were presented in a meeting abstract at Experimental Biology 2007.

**Reprint request to:** Emily Scott, Ph.D., Department of Medicinal Chemistry, University of Kansas, 1251 Wescoe Hall Dr., Lawrence, KS 66045, eescott@ku.edu

## Figure Legends

Figure 1. UV-Vis difference spectra from titration of CYP2A13 with increasing concentrations of MAP. Sequential titration additions are ordered on spectral binding image from orange to indigo. Not all spectra are shown. Inset is the nonlinear regression analysis completed using Graphpad Prism 4 (Graphpad Software, San Diego, CA) from which  $K_D$  is obtained.

Figure 2. Crystal structure of CYP2A13 with indole in the active site (2P85; (Smith et al., 2007). The ten mutations highlighted in yellow. Figure generated using Pymol (DeLano, W. L. The PyMOL Molecular Graphics System (2003) DeLano Scientific, San Carlos, CA, USA).

Figure 3. Fold difference comparison of CYP2A13 mutant  $K_D$  values to CYP2A13  $K_D$ . Black bars represent the ligand MAP, light grey represents PEITC, and medium grey represents coumarin."CYP2A13  $K_D$  values are all arbitrarily set to 1 to facilitate comparison of mutation effects across ligands. All mutants whose  $K_D$  varies by 3-fold or greater from CYP2A13 are shown as the mean of two independent titration experiments.

Figure 4. Correlation plot for  $pK_D$  ( $= \log[K_D]$ ) values determined via the COMBINE-based QSAR model for CYP2A6, CYP2A13, and CYP2A13 mutants relative to the experimental measurements determined herein.

Figure 5. Results of COMBINE analysis. A. High and low affinity ligands depicted in the presence of electrostatic features as gauged from COMBINE analysis. The high-scoring complex corresponds to PEITC (CPK colored sticks with green carbon atoms) in its conformation as bound to CYP2A13, and the low-scoring complex entails MAP (CPK colored sticks, with orange carbon atoms) in the conformer as bound to 2A6. The protein surface entails a composite of the largely conserved CYP2A6 and CYP2A13 structures, with features colored as follows: blue, favorable electropositive; red, favorable electronegative; cyan, unfavorable electropositive; magenta, unfavorable electronegative. B. High and low affinity ligands PEITC

and MAP colored as in panel A, but in the presence of key protein steric features as gauged from COMBINE analysis. The ribbons and corresponding mesh represent the CYP2A13 parental protein (cyan ribbons and mesh) and the CYP2A13 A301G mutant protein (magenta ribbons and mesh). Voids calculated by VOIDOO. C. Conformational variation of Leu370 as a function of the CYP2A active site (CYP2A6, green; CYP2A13, cyan; CYP2A13/L366I, salmon; CYP2A13/G369S, blue; CYP2A13/H372R, magenta). The relative positions of bound ligands PEITC and MAP (colored as in panel A) are shown for reference.

## Tables

Table 1. Sequences are shown below for one of the two oligonucleotides used in the construction of each of the CYP2A13 mutations. For each mutation, the second oligonucleotide used was a perfect complement of the oligonucleotide sequence shown. Bold indicates changes from the CYP2A13 sequence. Underline indicates the location of the desired mutation. Italics indicate changes that alter a restriction site and were used to facilitate identification of plasmids containing the desired mutation.

Mutation	Sequence (5' to 3')	Restriction Site Altered
L110V	CC ACC TTC GAC TGG <u>GTC</u> TTC AAA GGC TAT GGC	delete SapI
A117V	GGC TAT GGC GTG <u>GTC</u> TTC AGC AAC GGG	add BbsI
S208I	CGC ATG ATG CTG GGA <u>ATC</u> TTC CAG TTC ACG GCA ACC	delete HindIII
A213S	GGA AGC TTC CAG TTC ACG <u>TCC</u> ACC TCC ACG GGG CAG C	new SalI
F300I	CC CTG AAC CTC TTC <u>ATT</u> GCG GGT ACC GAG ACC GTG AGC ACC	add KpnI
A301G	CC CTG AAC CTC TTC TTT <u>GCG</u> GGT ACC GAG ACC GTG AGC	add KpnI
M365V	C CAA AGA TTT GGA GAC <u>GTC</u> CTC CCC ATG GGT TTG G	add AatII
L366I	C CAA AGA TTT GGA GAC ATG <u>ATC</u> CCG ATG GGT TTG GCC C	delete NcoI
G369S	GGA GAC ATG CTC CCC ATG <u>AGC</u> TTA GCC CAC AGG GTC AAC AAG G	add Bpu1102I
H372R	CCC ATG GGT TTG GCC <u>CGC</u> AGG GTT AAC AAG GAC ACC AAG TTT CG	add HpaI

Table 2. Binding constants for various ligands to cytochrome 2A6 vs. cytochrome 2A13.

Values are the average of duplicate measurements.

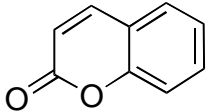
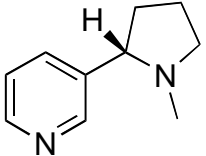
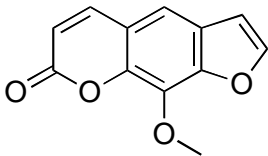
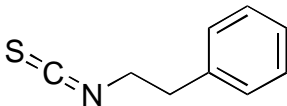
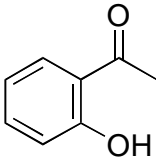
Ligand	Structure	$K_D$ ( $\mu$ M)		$\frac{K_D 2A6}{K_D 2A13}$
		2A6	2A13	
coumarin		3.1	2.7	1.1
nicotine		103.	22.	4.7
8-methoxypsoralen		1.3	<0.10	13
phenethyl isothiocyanate		6.2	0.43	14.
2'-methoxyacetophenone		74.	1.0	74.

Table 3.  $K_D$  values ( $\mu\text{M}$ ) for 2'-methoxyacetophenone (MAP), phenethyl isothiocyanate (PEITC) and coumarin binding to CYP2A6, CYP2A13, and mutants of CYP2A13 in which individual active site residues have been substituted with the residue found at the corresponding position in CYP2A6. Numbers in italics give the fold difference from the CYP2A13 value for the corresponding ligand. Values represent the mean of two titrations for CYP2A6, CYP2A13, and all mutants with  $\geq 3$ -fold difference from the parental enzyme.

Enzyme	MAP		PEITC		Coumarin	
	$K_D$ ( $\mu\text{M}$ )	$\frac{K_D}{K_D(2A13wt)}$	$K_D$ ( $\mu\text{M}$ )	$\frac{K_D}{K_D(2A13wt)}$	$K_D$ ( $\mu\text{M}$ )	$\frac{K_D}{K_D(2A13wt)}$
CYP2A6	74	74	6.1	14	3.1	1.1
CYP2A13	1.0	1.0	0.43	1.0	2.7	1.0
CYP2A13 mutants:						
L110V	1.6	1.6	0.46	1.1	1.7	0.63
A117V	4.1	4.1	1.2	2.8	1.8	0.67
S208I	12.0	12.	1.4	3.3	7.9	2.9
A213S	2.9	2.9	0.87	2.0	3.4	1.3
F300I	15.0	15	6.0	14.	14.0	5.2
A301G	8.2	8.2	3.1	7.2	4.2	1.6
M365V	1.7	1.7	1.7	4.0	5.0	1.9
L366I	0.56	0.56	0.34	0.79	0.57	0.21
G369S	1.6	1.9	1.8	4.2	6.0	2.2
H372R	0.77	0.77	1.0	2.3	1.5	0.56
CYP2A6 mutants:						
I208S/I300F/G301	12.0	12.	4.9	11	6.0	2.2
I300F/G301A/S369G	12.3	12	1.9	4.4	3.1	1.1
		3.5		4.8		0.85
I208S/I300F/G301A/S369 G	3.5		2.1		2.3	

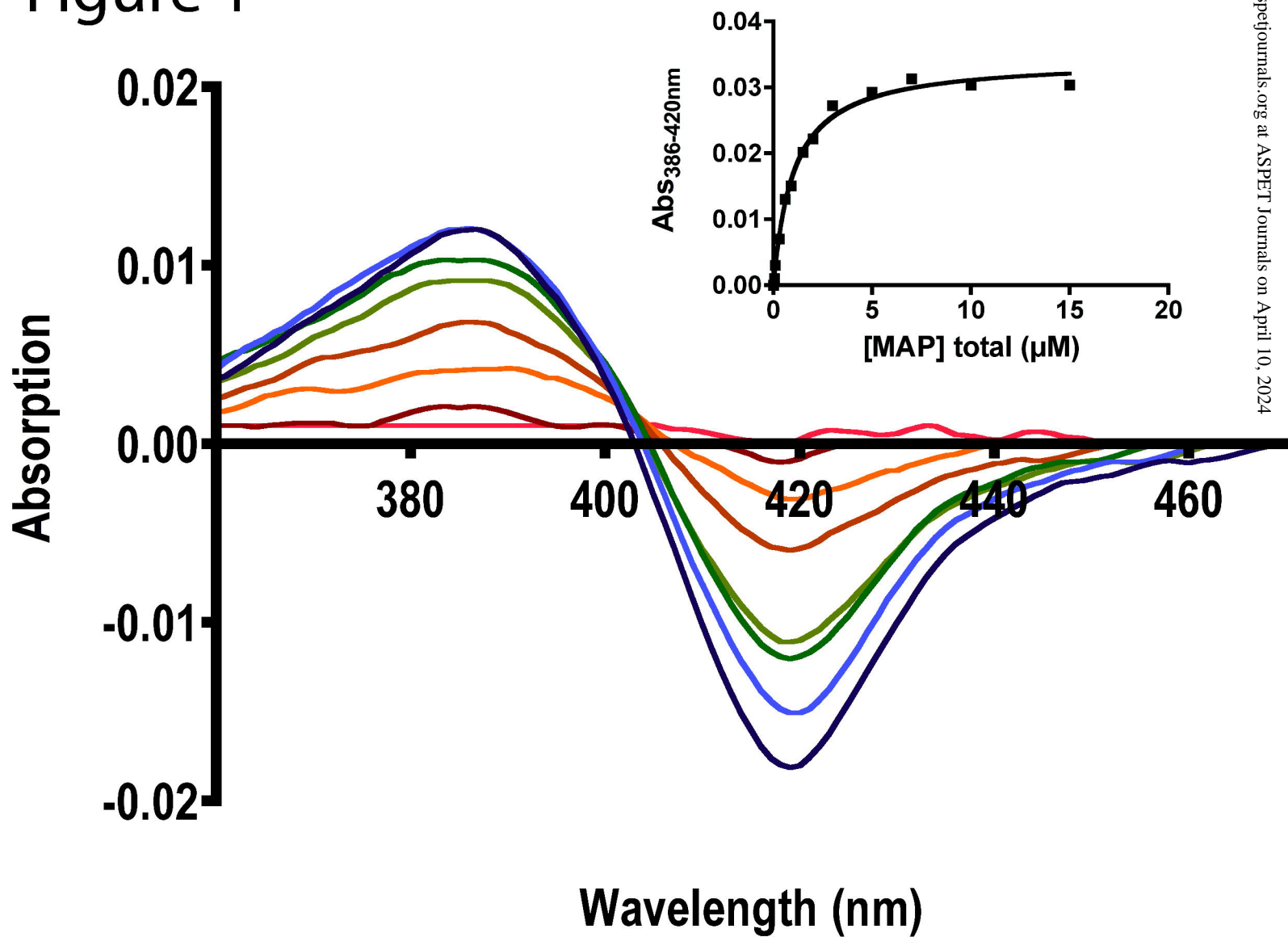
Table 4. Comparison of  $pK_D$  ( $=\log K_D$ ) values derived from the COMBINE model and from experimental titrations of MAP, PEITC and coumarin with CYP2A6, CYP2A13, and CYP2A13 mutants.

Enzyme	MAP		PEITC		Coumarin	
	Experimental	Predicted	Experimental	Predicted	Experimental	Predicted
CYP2A6	1.87	1.67	0.79	0.78	0.49	0.49
CYP2A13	0.02	0.49	-0.37	-0.22	0.43	0.55
CYP2A13						
mutants:						
L110V	0.20	0.36	-0.34	-0.32	0.23	0.09
A117V	0.61	0.45	0.08	0.24	0.26	0.10
S208I	1.1	1.19	0.15	0.05	0.90	0.79
A213S	0.46	0.57	-0.06	-0.24	0.53	0.71
F300I	1.2	1.26	0.78	0.97	1.1	0.94
A301G	0.91	0.66	0.49	0.62	0.62	0.59
M365V	0.23	0.34	0.23	0.12	0.70	0.60
L366I	-0.25	--	-0.47	-0.49	-0.24	--
G369S	0.20	0.50	0.26	0.16	0.78	0.56
H372R	-0.11	--	0.00	-0.19	0.18	--

--: omitted from model as described in methods.



Figure 1



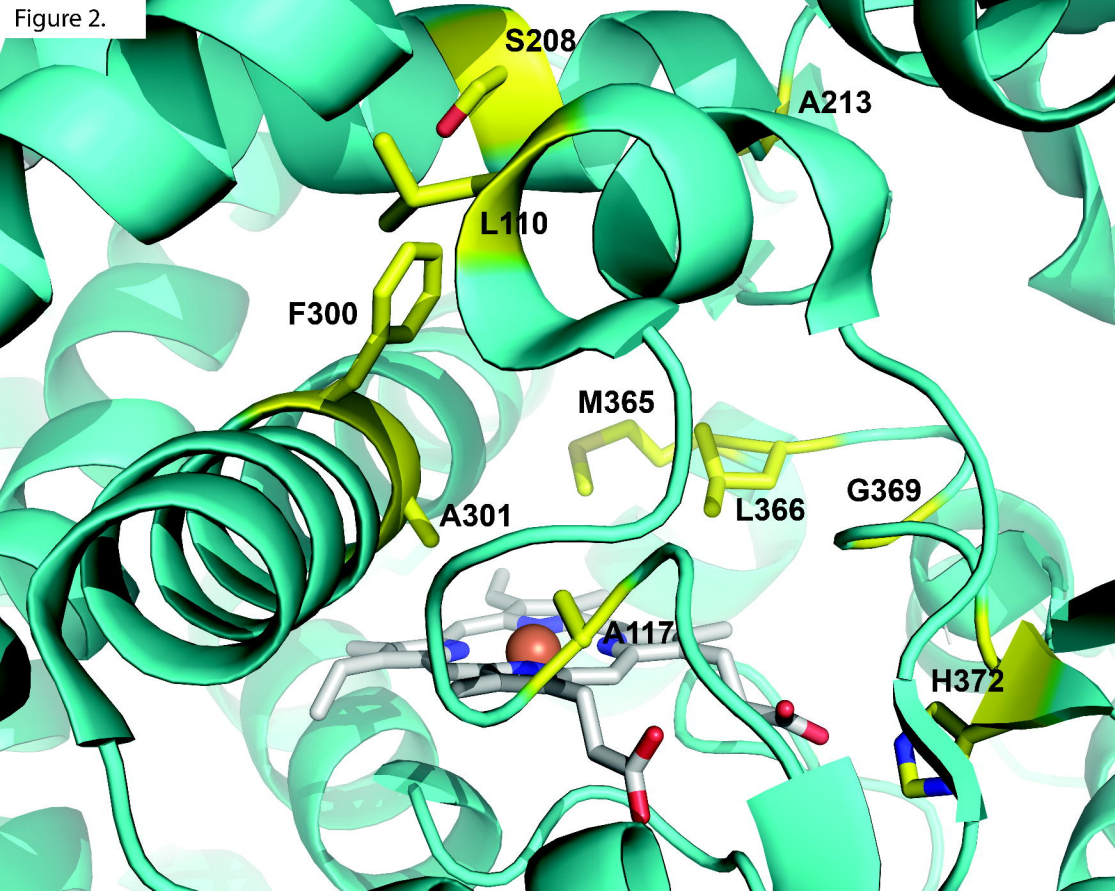


Figure 3.

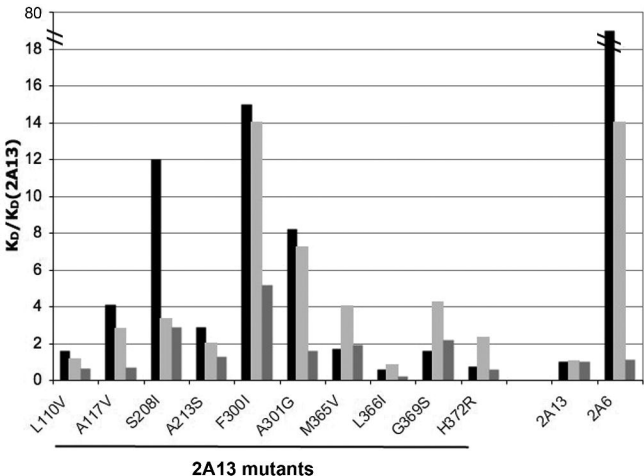


Figure 4.

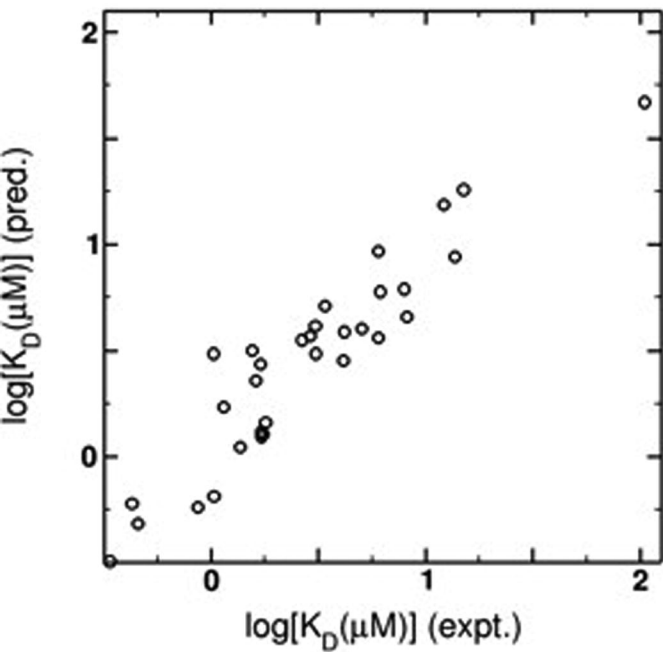


Figure 5, panels A,B, and C, top to bottom

