THE HUMAN CYTOCHROME P450 2A FAMILY: FUNCTIONAL COMPARISONS AND IDENTIFICATION OF AMINO ACIDS ESSENTIAL FOR SUBSTRATE RECOGNITION

BY

©2008
Natasha M. DeVore

B.A., Evangel University, 2006

Submitted to the Department of Medicinal Chemistry and the Faculty of the Graduate School of the University of Kansas in partial fulfillment of the requirements for the degree of Master's of Science.

_____
Chairperson – Emily Scott

_____
Committee members: Thomas Prisinzano

_____
Audrey Lamb

Date Defended: 9/16/08

The Thesis Committee for Natasha M. DeVore certifies that
this is the approved version of the following thesis:

THE HUMAN CYTOCHROME P450 2A FAMILY:
COMPARISONS AND IDENTIFICATION OF AMINO ACIDS ESSENTIAL FOR SUBSTRATE RECOGNITION

Committee:

_____
Chairperson – Emily Scott

_____
Thomas Prisinzano

_____
Audrey Lamb

Date Approved:

Abstract

The goal of this research was to identify the differential structure-activity relationships between cytochromes P450 (CYP) 2A13 and 2A6 and their substrates. Cytochromes P450 2A13 and 2A6 are very closely related, having 94% amino acid sequence identity. Both proteins metabolize drugs, toxins, and procarcinogens, including the nicotine derivative 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) and aflatoxin B_1. Yet CYP2A13 has a much higher rate of metabolism than CYP2A6 for both compounds. Combined with its increased affinity and expression in the respiratory tract, CYP2A13 is a candidate for breakdown of NNK into the reactive metabolites that can form DNA adducts in the lung, leading to the development of lung cancer (1). Thus, understanding differential metabolism is important. Due to the high identity of CYP2A13 and CYP2A6, we hypothesized that a few amino acids found within the active site are responsible for CYP2A substrate specificity. Of the 32 amino acids which differ between CYP2A13 and CYP2A6, ten are located in regions of the protein that often determine substrate recognition in cytochromes P450 (2). Crystallization of CYP2A13 by the Scott lab verified that these amino acids were in or near the active site. The goal of this project was to determine which of these ten amino acids were vital for substrate selectivity in the CYP2A enzymes.

Several different methods were employed throughout this study including site-directed mutagenesis, ligand binding assays, metabolism assays, molecular modeling, and crystallography. Site-directed mutagenesis was used to construct ten

single residue CYP2A13 mutants with each mutation substituting the residue found in CYP2A6 in that position. The consequences of these ten mutations to the binding of coumarin, 2′-methoxyacetophenone, phenyl isothiocyanate, and phenacetin were determined using spectral ligand binding assays. The amino acids in positions 117, 208, 300, 301, 365, and 369 caused significant changes in ligand affinity. To determine if these amino acids were also essential for substrate metabolism, phenacetin was used as a structural probe since CYP2A6 converts phenacetin to acetaminophen with decreased catalytic efficiency compared to CYP2A13 (3). Employing this assay, four mutations (S208I, F300I, A301G, and G369S) were identified that diminished CYP2A13 phenacetin O-deethylation to near CYP2A6 activity. Two mutants, CYP2A13 A117V and L366I, both increased enzyme activity more than four-fold. Construction of the reverse mutant series, a CYP2A6 protein incorporating the residues found at the corresponding active site positions in CYP2A13, confirmed that positions 208, 300, 301, and 369 were jointly responsible for phenacetin metabolism and binding. While all single 2A6 mutants had very low-level activity, the double, triple, and a quadruple mutant with changes at these four positions increasingly conferred phenacetin metabolism to CYP2A6.

A structural basis for the effects of these four mutations on phenacetin binding and metabolism was explored with molecular docking studies using Surflex-Dock (4). The results suggested that the ability of CYP2A13 to bind and metabolize phenacetin was due to steric variations among key residues in the CYP2A13 and CYP2A6 active

sites. This was confirmed with the crystal structure of the CYP2A6 I208S/I300F/G301A/G369S complexed with phenacetin.

Finally, this study characterized the effects of naturally occurring CYP2A13 polymorphisms on the binding of the well-known CYP2A ligand coumarin (5). All of these polymorphisms were located on the exterior portions of the protein. None of the polymorphisms examined resulted in a significant change in ligand affinity, indicating that having one of these variants would not substantially alter an individual's ability to metabolize CYP2A13 substrates.

Thus, the research presented in this thesis provides structural and functional insight into CYP2A13 and CYP2A6 substrate selectivity, which is largely modulated by the steric effects mediated by the differential amino acids at positions 208, 300, 301, and 369.

Please Note:

The CYP2A13 polymorphism data presented in this paper was previously published as a joint publication from the Scott and Murphy labs in 2007 by Kari E. Schlicht, Natasha M. Michno, Brian D. Smith, Emily E. Scott, and Sharon E. Murphy in *Xenobiotica* as a "Functional characterization of CYP2A13 polymorphisms." The phenacetin binding, metabolism, docking, and the crystal structure of CYP2A13 I208S/I300F/G301A/S369G are included in the article "Key Residues Controlling Phenacetin Metabolism by Human Cytochrome P450 2A Enzymes" currently in press by *Drug Metabolism and Distribution*. The article is by Natasha M. DeVore, Brian D. Smith, Michael J. Urban, and Emily E. Scott.

References

1. Su, T., Bao, Z. P., Zhang, Q. Y., Smith, T. J., Hong, J. Y., and Ding, X. X. (2000) *Cancer Res.* **60**(18), 5074-5079
2. Zhang, J. Y., Wang, Y., and Prakash, C. (2006) *Curr Drug Metab* **7**(8), 939-948
3. Bieche, I., Narjoz, C., Asselah, T., Vacher, S., Marcellin, P., Lidereau, R., Beaune, P., and de Waziers, I. (2007) *Pharmacogenetics and Genomics* **17**(9), 731-742
4. Nishimura M Fau - Yaguti, H., Yaguti H Fau - Yoshitsugu, H., Yoshitsugu H Fau - Naito, S., Naito S Fau - Satoh, T., and Satoh, T. (0031-6903 (Print))
5. Schlicht, K. E., Michno, N., Smith, B. D., Scott, E. E., and Murphy, S. E. (2007) *Xenobiotica*, 1-11

Acknowledgements

Thanks are due to Jenny Morrison, Chad Schroeder, and Melanie Blevins for construction of the CYP2A13 mutants and to Linda Blake, Naseem Nikaeen, Matthew Axtman, Agnes Walsh, and Kyle Bailey for construction of the CYP2A6 mutants during their respective rotations in the Scott lab. Thanks are also due to Christopher Wood for his assistance in expressing and purifying some of the CYP2A13 mutants and Michael Urban for his assistance in expressing and purifying some of the CYP2A6 mutants. Full length human CYP1A2 was a generous gift from Dr. Fred Guengerich (Vanderbilt University School of Medicine, Nashville, TN). I would also like to thank all of the past and present Scott Lab members, Naseem Nikaeen, Eric Carrillo, Anu Metha, Agnes Walsh, Patrick Porubsky, Kathleen Meneely, Linda Blake, and Brian Smith, for their support and encouragement over the past two years.

Thank you to Dr. Lamb and Dr. Prisinzano for agreeing to serve on my committee. I am especially grateful to Dr. Scott for all of her advice and guidance throughout these past two years of graduate school. Thank you to the NIH for funding provided through grant GM 076343.

In addition, I would like to thank my husband Matt and dog Romulus for helping to keep me sane and for their constant encouragement.

Table of Contents

List of Figures

List of Tables

List of Schemes

Chapter 1.

Introduction

Introduction: Cytochromes P450

Cytochrome P450 enzymes are the primary enzymes responsible for Phase I metabolism of drugs and other xenobiotics. In this process, a foreign small molecule is modified to decrease hydrophobicity. There are a number of reactions catalyzed by P450s that increase the hydrophilicity of a substrate molecule, including both oxidative and reductive pathways. In addition to xenobiotics, cytochrome P450 substrates include sterols, fatty acids, eicosanoids, and vitamins (1). However, the cytochromes P450 that metabolize xenobiotics are of particular interest due to their importance in drug metabolism and their possible role activating protoxins and procarcinogens.

Cytochromes P450 were first identified by Garfinkel and Klingenberg in 1958 by noting a characteristic absorption maximum at 450 nm on a UV spectra after the addition of carbon monoxide into a solution of buffer containing the heme protein and a strong reducing agent such as dithionite (Figure 1.1) (2). Since that time, much research has been conducted into the expression, structure, and function of cytochrome P450 enzymes.

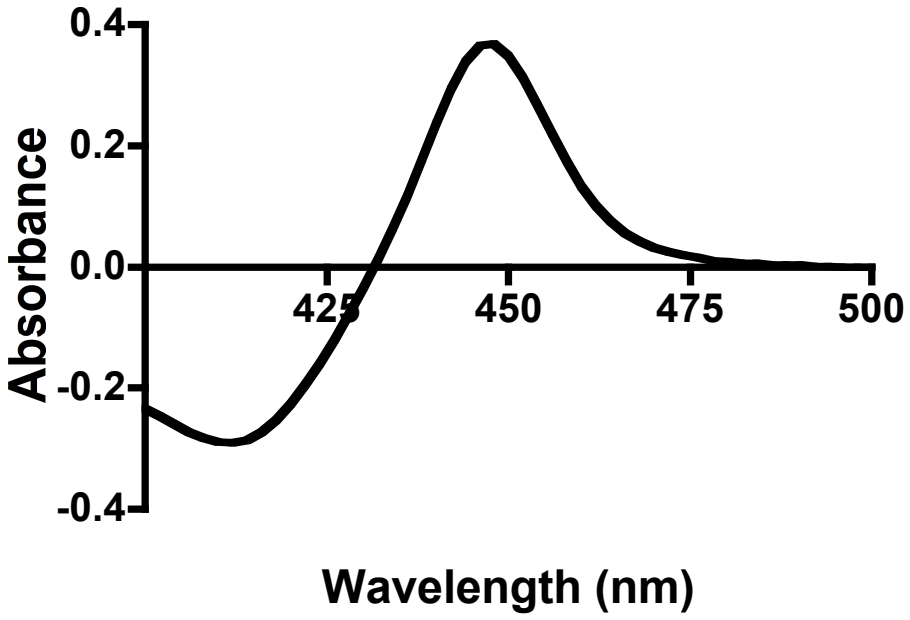


Figure 1.1. P450 carbon monoxide difference assay spectra. When CO is bubbled into a solution of P450 containing dithionite, CO binds to the P450 heme resulting in a spectral shift from 420 nm to 450 nm.

Human cytochromes P450 are primarily found in the liver, although they are also found in the kidney, respiratory system, gonads, pancreas, spleen, gastrointestinal tract, and skin (3). The expression location and amount of cytochrome P450 protein produced is highly dependent on the enzyme in question and often differs in response to xenobiotic exposure (4,5). Therefore, levels are often highly variable between individuals.

Cytochrome P450 Nomenclature

Cytochromes P450 are divided into 135 families. Each family is defined by having at least 40% amino acid identity among members (6,7). Separate families are indicated by Roman numerals. Families 1, 2, and 3 are involved in the metabolism of xenobiotics and, as such, are particularly important in drug metabolism studies (Figure 1.2) (1,7). Each family is sub-divided into subfamilies that share at least 55% identity and are represented by capital letters (7). Individual P450 isozymes are then assigned a second Roman numeral, indicative of the particular enzyme. For example, CYP2A13 belongs to family 2 and subfamily A. If a variant allele exists for an established cytochrome, it is named by a star followed by a number, for example CYP3A5*2 (8).

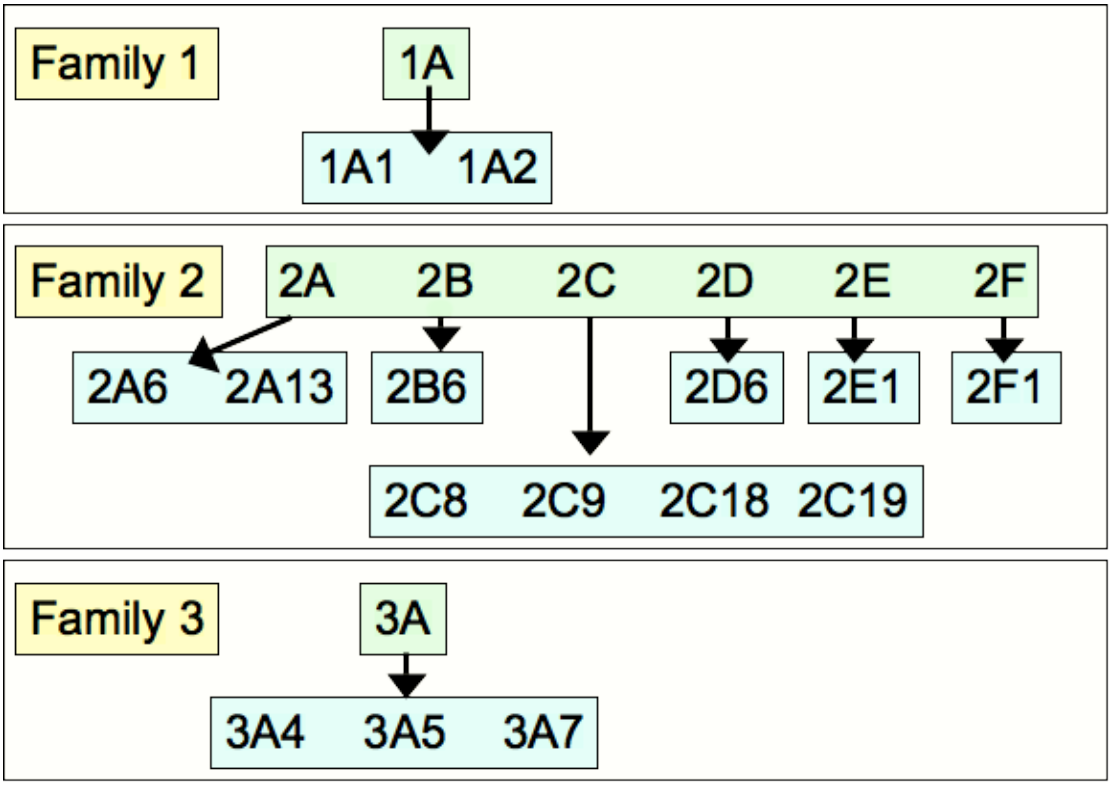


Figure 1.2. The three human cytochrome P450 families involved in xenobiotic metabolism.

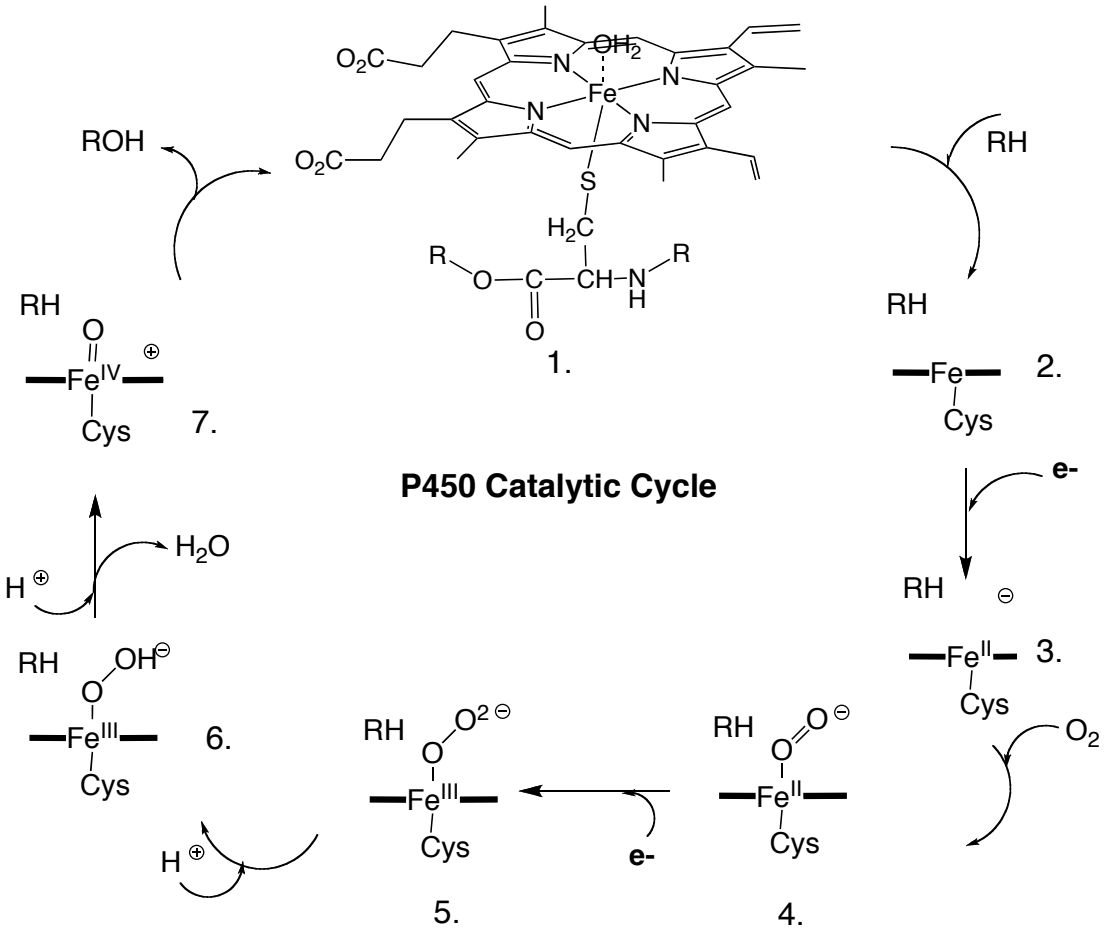
The Cytochrome P450 Catalytic Cycle

The xenobiotic cytochrome P450 enzymes are able to metabolize a substrate via a metal ion catalysis cycle. The primary features of this cycle are conserved between cytochromes P450. The main features of the catalytic cycle include: substrate binding, oxygen binding, oxygen scission, oxygen insertion into the substrate, and product release (Scheme 1.1).

The Human Cytochrome P450 2A subfamily

There are three cytochrome P450 enzymes belonging to the human CYP2A subfamily: CYP2A6, CYP2A13, and CYP2A7. All three members of the 2A family are located on human chromosome 19 in a cluster that also includes the genes for CYP2F1, CYP2B7, and CYP2B6 (Figure 1.3) (9,10). The CYP2A family is known for its ability to metabolize xenobiotics, and its substrates include several toxins and procarcinogens: aflatoxin B1, 4-aminobiphenyl, and the nicotine derivative 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK).

CYP2A6 is the most studied of the three enzymes because it constitutes 5-10% of the total microsomal cytochrome P450 in the human liver (11,12). The CYP2A6 enzyme was identified along with two variants and CYP2A7 in 1989 by *Yamano et al* (12). CYP2A6 was identified as the main enzyme responsible for coumarin 7-hydroxylation in humans while the two variants identified at the same time had severely reduced coumarin 7-hydroxylation levels and CYP2A7 had no activity (12). Coumarin 7-hydroxylation was at higher levels in human liver than in several animal species and coumarin was identified as a potential cytochrome P450



Scheme 1.1. The catalytic cycle of cytochromes P450. 1) A water molecule is bound to the Fe atom. 2) A ligand binds in the active site of the P450, displacing the water molecule. 3) An electron from NADPH is delivered to the iron (by NADPH cytochrome P450-oxidioreductase) reducing the iron to the ferrous form. 4) Molecular oxygen binds to the reduced iron atom. 5) A second electron is transferred to the system by either P450-oxidoreductase or cytochrome b_5 which reduces the complex to a peroxoanion intermediate. 6) The terminal oxygen atom binds a proton. 7) A second proton is bound to the terminal oxygen atom, which then leaves as water. The other atom of oxygen is left bound to the iron, probably as compound I. At this point the remaining oxygen is able to abstract a hydrogen from the substrate, and then rebound to the substrate as a hydroxyl group. The substrate may then undergo further rearrangement to form the final product.

substrate years before CYP2A6 was even discovered (13). The level of 2A6 expression in human liver samples is highly variable, consistent with the observed variation in coumarin 7-hydroxylation between human liver microsomes (12-14). In addition to the metabolism of coumarin, CYP2A6 is also able to metabolize a diverse range of substrates including indoles, nicotine, the gasoline additive methyl tert-butyl ether (MTBE), and N-nitrosodiethylamine (NDEA) (11,15). Known inhibitors of CYP2A6 include several isothiocyanates, the monoamine oxidase inhibitor tranylcypromine, 8-methoxysporalen, and the antifugal ketoconazole (16-18).

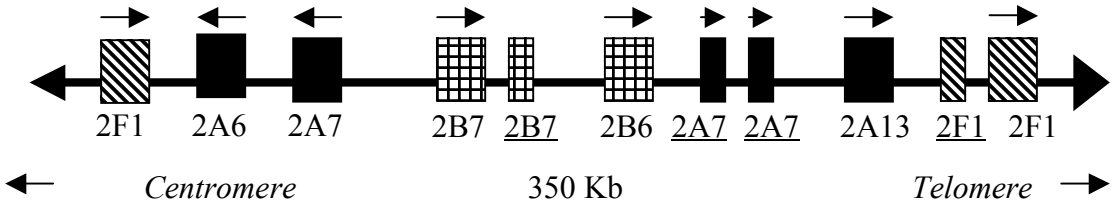


Figure 1.3. Organization of the human cytochrome P450 family 2 genes on chromosome 19. The diagram above is a 350 Kb cluster with the centromere and telomere labeled, and the direction of transcription indicated by arrows above the respective genes. Underlines indicate pseudogenes. This figure was redrawn from Fernadex-Salguero *et al.* (1995).

CYP2A7 was identified at the same time as CYP2A6 (14). CYP2A7 has a 94% sequence identity with CYP2A6, differing by 31 amino acids out of 494, and is located next to CYP2A6 on chromosome 19 (Figure 1.3) (10,12,14). Two psuedogenes have been identified that are alternate spliced forms of CYP2A7 (9). Several variants of CYP2A6 have also been identified which seem to have resulted from gene crossover events between the CYP2A6 and CYP2A7 genes (9). Early efforts at the expression of CYP2A7 were inconclusive as to whether or not CYP2A7 mRNA could produce active protein. Yamano and coworkers expressed CYP2A7 in the vaccinia virus (12). However, only small amount of immunostained protein band was generated as identified by SDS page gel using a rat CYP2A2 antibody (12). In addition, cells infected with this recombinant virus after 24 hours failed to show the characteristic CYP450 Soret shift from 420 nm to a maxima of ~450 nm upon the addition of dithionite and carbon monoxide (Figure 1.1). Ding and coworkers cloned CYP2A7 cDNA into a $pCMV_4$ vector and transfected it into COS cells (14). They were able to obtain a CYP2A7 protein with a molecular mass of 49kDa as indicated by Western-blot analysis, but addition of coumarin to the cells failed to produce a measurable amount of 7-hydroxycoumarin, the distinctive coumarin metabolite usually produced by CYP2A enzymes (14). Ding and coworkers did not report CO difference assay results, so it is impossible to know if folded, active CYP2A7 was present. Since these initial studies, very little research has been conducted on CYP2A7 and so the activity of the protein remains unknown. CYP2A7 is currently labeled as an orphan human cytochrome P450 (1).

The CYP2A13 gene was identified when the 2A genes were fully sequenced in 1995 (9,10). CYP2A13 has a 94% amino acid identity with CYP2A6, differing by 32 amino acids out of 494. The enzyme was not characterized until nearly five years later when Su *et al.* determined that CYP2A13 was primarily expressed within the respiratory tract, with the highest level of mRNA present in nasal mucosa, and had activity toward NNK, coumarin, and several other CYP2A6 substrates (19). Since that time, several more substrates have been identified. Isothiocyanates and 8-methoxysporalen are potent inhibitors for both CYP2A13 and CYP2A6 (17,18,20). Expression levels of CYP2A13 within the respiratory system vary between individuals (21). Primary interest in the enzyme stems from its location in the respiratory system and ability to metabolize NNK and aflatoxin B_1 to carcinogens with a higher efficiency than CYP2A6.

Polymorphisms

Each of the CYP2A family enzymes have several polymorphisms. Many of the CYP2A6 polymorphisms vary with ethnicity and may contribute to both smoking behavior and cancer incidence (22-25). Several of the CYP2A6 polymorphisms are alterations of a single amino acid, yet these polymorphisms can result in significantly reduced nicotine metabolism (Table 1.1). Although less is known about CYP2A7, several splice variants of both CYP2A6 and CYP2A7 have been identified that have either no or severely reduced catalytic activity (14,23).

The polymorphisms of CYP2A13 have been studied less than those of CYP2A6. Fujieda and coworkers identified 18 novel non-synonymous

polymorphisms in CYP2A13, however, the functional significance of the many of the single amino acid polymorphisms identified was unknown (26,27). One aspect of my research has been a collaboration to identify the functional significance of seven CYP2A13 polymorphisms (R257C, F453Y, R494C, D158E, V323L, R101Q, 133_134Thr insertion) using NNK and coumarin metabolism assays and coumarin spectral binding assays (28). Prior to our investigation, only the CYP2A13 variants R257C and R101E had been characterized as purified proteins (27,29).

Table 1.1. CYP2A6 non-synonymous polymorphisms and known functional effects.

Allele	Protein Change	Functional Change
CYP2A6*2	L160H	Alters coumarin metabolism to coumarin 3-hydroxylation instead of coumarin 7-hydroxylation (30)
CYP2A6*3	CYP2A6/CYP2A7 hybrid	Decreased coumarin metabolism *in vitro* (31)
CYP2A6*4	CYP2A6 deleted (32)	Unknown
CYP2A6*5	G479V	Decreased coumarin metabolism *in vitro* (31)
CYP2A6*6	R128Q	Decreased coumarin metabolism (33)
CYP2A6*7	I471T	Decreased activity for nicotine and coumarin (34)
CYP2A6*8	R485L	No effect on coumarin or nicotine metabolism (34)
CYP2A6*9	A-1013G and T-48G	Reduced coumarin metabolism (35)
CYP2A6*10	I471T and R485L	Absent or reduced catalytic activity for nicotine and coumarin (34)
CYP2A6*11	S224P	Reduced tegafur metabolism (36)
CYP2A6*12	CYP2A6/CYP2A7 hybrid	Reduced coumarin metabolism (37)
CYP2A6*13	T48G and G5R	Unknown
CYP2A6*14	S29N	Unknown
CYP2A6*15	T48G and K194E	Unknown
CYP2A6*16	R203S	No effect on coumarin or nicotine metabolism (38)
CYP2A6*17	V365M	Decreased nicotine and coumarin metabolism (38,39)
CYP2A6*18	Y392F	Decreased coumarin and tegafur clearance (40)
CYP2A6*19	Y392F and I471T	Decreased metabolic activity for coumarin and tegafur (40)
CYP2A6*20	Frame Shift, truncation	No activity for coumarin or nicotine (41)
CYP2A6*21	K476R	No affect on nicotine metabolism (42)
CYP2A6*22	D158E and L160I	Unknown
CYP2A6*23	R203C	Reduced nicotine and coumarin metabolism (38)

Substrate Selectivity

Consistent with the close amino acid identity, CYP2A13 and CYP2A6 have a substantial overlap in substrate selectivity. Both proteins are able to metabolize a wide range of substrates including coumarin, nicotine, aflatoxin B_1, and the nicotine procarcinogen NNK (11,43,44). Many CYP2A6 and CYP2A13 substrates are planar cyclic molecules, but there are a number of exceptions including methyl t-butyl ether, N-nitrosodiethylamine (NDEA), and halothane (2,11). Although there is considerable substrate overlap, the difference in the rate of metabolism can be considerable for individual substrates between CYP2A6 and CYP2A13. Depending on the report, CYP2A13 has a catalytic efficiency for NNK that is 75 - 215 times greater than that of CYP2A6 (19,45). Furthermore, CYP2A13 has 22- and 15-fold greater catalytic efficiency for nicotine C-oxidation and cotinine 3′-hydroxylation than CYP2A6, respectively (46). There is also considerable variation in the metabolism of coumarin, hexamethylphosphoramide, N,N-dimethylaniline, 2′-methoxyacetophenone, and N-nitrosomethylphenylamine (19). Furthermore, although both CYP2A6 and CYP2A13 metabolize aflatoxin B_1, only CYP2A13 metabolizes aflatoxin B_1 into a mutagenic epoxide (43).

Two substrates, the antipyretic drug phenacetin and procarcinogen 4-aminobiphenyl, are metabolized by CYP2A13 while CYP2A6 barely has detectable activity with these substrates (47,48). CYP2A13 is not the only cytochrome P450 that metabolizes phenacetin and 4-aminobiphenyl, as the CYP family 1 also metabolizes these substrates. CYP2A13 metabolizes 4-aminobiphenyl with a 8-fold

decrease in catalytic efficiency and phenacetin with a 2.5-fold higher efficiency compared to CYP1A2 (47,48). The overlap in substrate metabolism between CYP2A6, CYP2A13, and CYP1A2 is of particular interest because CYP2A13 and CYP1A1, closely related to CYP1A2, are both found primarily in the lung (4). In contrast, CYP2A6 and CYP1A2 are expressed primarily in the liver.

CYP2A6 and CYPA13 Structure and Active Site

The overall structure of CYP2A13 closely resembles that of CYP2A6, with an average root mean square deviation of 0.5 Å for the Cα atoms (49). Both enzymes contain twelve major α helices, which are labeled A-L (Figure 1.4). The I helix spans the length of the entire cytochrome P450 structure, with a small kink in the helix near the iron protoporphyrin IX prosthetic group. The iron atom of the protoporphyrin IX coordinates to C439, located in the K′-L loop, just below the heme. Several additional hydrogen bonds and hydrophobic interactions with the surrounding loop and helix residues occur to stabilize the location of the heme. The active site is found above the heme. Although direction in an enzyme is arbitrary, throughout the rest of this thesis the floor of the active site will refer to the heme, while the ceiling of the active site is formed I helix and F helix residues as well as a few residues from the K-K′ and B′-C loop. The walls of the active site are formed by the I helix, and the K- K′ and B′-C loops.

Some differences are seen in the in the active site architecture of the two CYP2A enzymes (49). These differences account for an active site in CYP2A13 which is 15-20% larger than the active site of CYP2A6, which is one of the smallest

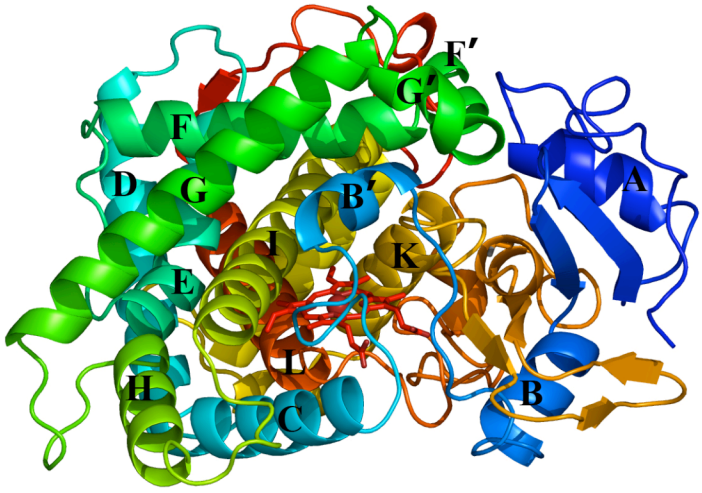


Figure 1.4. The crystal structure of human cytochrome P450 2A13 (PDB 2P85). Helices are labeled A-L. The heme is colored red.

identified P450 active sites (49). The active site volume for CYP2A13 is reported as 307 $Å^3$, a volume intermediate to that of CYP2A6 (260 $Å^3$) and CYP1A2 (375 $Å^3$) (49,50). CYP2A13 and CYP2A6 both have relatively small active sites which are planar, highly hydrophobic, and compact, therefore complimenting aromatic substrates (49,50). There are a number of amino acids located in or near the active site that differ between CYP2A6 and CYP2A13 (Figure 1.4). Crystal structures suggest that these differences could alter ligand binding and orientation within the two active sites (49). Although the crystal structures for CYP2A13 and CYP2A6 do not have the same ligand present, the coumarin bound in CYP2A6 and the indole bound in the CYP2A13 structure are bound in a similar location with respect to the heme and hydrogen bonding to N297 but adapt different planes (49). Several of the amino acids within the active site could be responsible for the ligand reorientation (Figure 1.4). These include the nearby amino acids at positions 208, 300, 301, 117, 365, and 366. Of particular interest is residue 300. In CYP2A13, phenylalanine 300 is positioned away from the ligand, while the isoleucine in CYP2A6 is directed into the active site (49). The amino acid at position 301 also differs, an alanine in CYP2A13 and a glycine in CYP2A6, potentially resulting in increased steric hindrance for ligand binding and orientation within the CYP2A13 active site.

Mutagenesis studies have shown that substitution of the alanine found at position 117 in CYP2A13 to the valine found in CYP2A6 creates a CYP2A13 mutant with increased activity for formation of 7-hydroxycoumarin, decreased activity for the formation of the NNK keto aldehyde metabolite, and increased metabolism of

aflatoxin B_1 to the harmful epoxides (43-45). Located on one wall of the CYP2A13 and CYP2A6 active site, it appears that the difference in space occupied by the alanine in CYP2A13 and the valine in CYP2A6 substantially affects the rate of metabolism for these two substrates (43-45,49). Site directed mutagenesis studies at the 208 position indicate that this residue may contribute to substrate specificity as studies have indicated that the 208 position is important for the metabolism of NNK but not coumarin (44,45).

Previous site directed mutagenesis studies also indicated that the histidine to arginine change at residue 372 in CYP2A13 decreases the substrate k_{cat} for couarmin 7-hydroxylation, decreases the NNK metabolism to both the aldehyde and keto metabolites, and abolishes the production of aflatoxin B_1 epoxides (43-45). The 372 position is further away from the substrate in the active site (Figure 1.4).

Active site differences between the amino acids at positions 365 and 366 could also affect substrate selectivity. These two positions are located across from the I helix (Figure 1.5). The methonine at position 365 in CYP2A13 is much longer than the corresponding valine in CYP2A6, resulting in the greater projection of this residue into the active site cavity in CYP2A13 (49). The leucine found in CYP213 also projects more into the active site than the isoleucine side chain at this position in CYP2A6. Thus, reorientation of the ligand in the active site as a result of several amino acid differences, including those at positions 208, 300, 301, 365, and 366 may mediate the observed differences in substrate binding and metabolism.

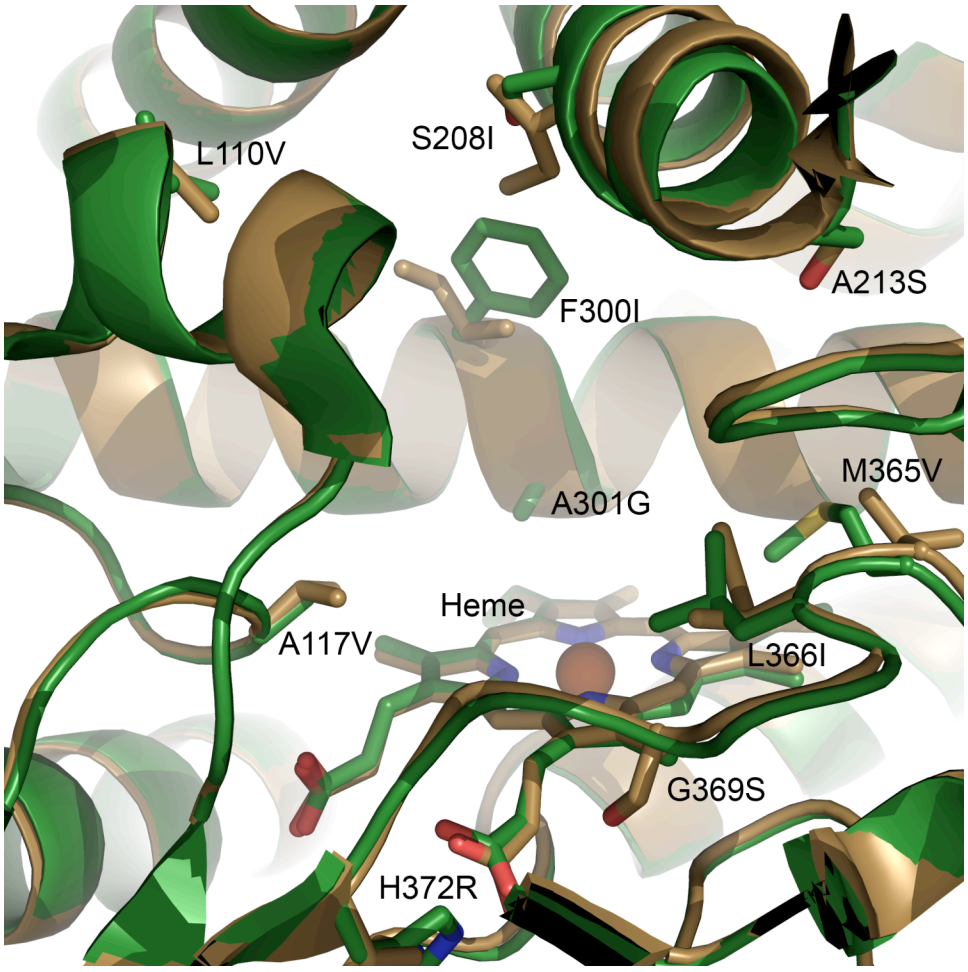


Figure 1.5. Structure of CYP2A13 (green, PDB 2P85) and CYP2A6 (brown, PDB 1Z10) overlaid with ligands removed. Residues that differ between CYP2A13 and CYP2A6 are represented by sticks and labeled (e.g. A117V – Ala found in CYP2A13, Val found in CYP2A6).

Project Goals and Hypothesis

The main goal of this study is to determine the structural divergence in the two very closely related human cytochromes P450 CYP2A13 and CYP2A6 responsible for enzyme function differences. We hypothesize that amino acids located near or in the active site that differ between CYP2A13 and CYP2A6 were most likely to control functional differences. To test this we have employed site-directed mutagenesis to exchange the amino acids that differ between the enzyme active sites and characterized the mutants for their ligand binding and metabolism abilities. We first characterized binding with a set of four CYP2A ligands, and then focused on the metabolism of phenacetin, a substrate metabolized by CYP2A13 and not by CYP2A6. Structural insight into the significance of various mutations was acquired by docking and X-ray crystallography. Collectively the crystal structures, phenacetin metabolism data, and ligand binding information provide valuable structural and function insight into the CYP2A proteins.

In addition, several polymorphisms that were previously identified for CYP2A13 were characterized for their coumarin binding affinity as a part of this project (26-29,51). Identifying and studying the effects of CYP2A polymorphisms on enzyme function is important to accessing nicotine and NNK metabolism and cancer risk.

Experimental Design

Rational Site-directed Mutagenesis. In order to determine the amino acids responsible for substrate selectivity, a series of ten mutants were constructed that

were identical to CYP2A13 except for one amino acid at a position where CYP2A13 and CYP2A6 did not have the same residue. These CYP2A13 mutants were: L110V, A117V, S208I, A213S, F300I, A301G, M365V, L366I, G369S, and H372R. This initial series of ten mutants was later expanded to include ten reverse single point mutations in the CYP2A6 background (V110L, V117A, I208S, S213A, I300F, G301A, V365M, I366L, S369G, and R372H) and several CYP2A6 multiple mutants combining these individual alterations.

At the initiation of this project, no crystal structures of the CYP2A enzymes were available, so the ten residues to mutate were selected based on sequence alignments between CYP2A13 and CYP2A6 and location within the previously postulated substrate recognition sites (SRS) in the cytochrome family 2 enzymes (Figure 1.6) (53). Residues found in these recognition sites are affiliated with differences in substrate selectivity in other cytochrome P450 enzymes. We hypothesized that residues within substrate recognition sites that differed between the two CYP2A enzymes may mediate the variation in CYP2A function.

During the course of this study, crystal structures for both CYP2A13 and CYP2A6 were determined by the Scott lab and the Johnson lab, respectively (49,54). The proximity of the residues selected for this study to the active site is verified by these structures (Figure 1.7). Two residues that were mutated (L110 and A117) are located in SRS-1. L110 is located above the active site, while A117 forms part of the active site wall. The residues S208 and A213 are both located in SRS-2. Residue A213 is located at the terminus of the F helix while S208 is located in the F helix

above F300. Residues F300 and A301 were also mutated and both are located in SRS-4 on the I helix. All of the other mutations made are located in SRS-5 and form part of the active site cavity across from the I helix. These residues are M365, L366, G369, and H372. H372 is located further from the active site, and is involved in the hydrogen bonding to the heme proprionate.

The wild type and mutant CYP2A13 and CYP2A6 proteins were expressed in *E. coli* and purified by nickel affinity and carboxymethyl sephrose ion exchange chromatography. Differences in ligand binding (phenacetin, 2`-methoxyacetophenone, phenethyl isothiocyanate, and coumarin) and phenacetin metabolism studies indicated residues that are involved in CYP2A affinity and metabolism. The effects of these mutations were interpreted in conjunction with molecular docking experiments. Finally, the structure of the quadruple mutant CYP2A6 I208S/I300F/G301A/S369G complexed with phenacetin was determined.

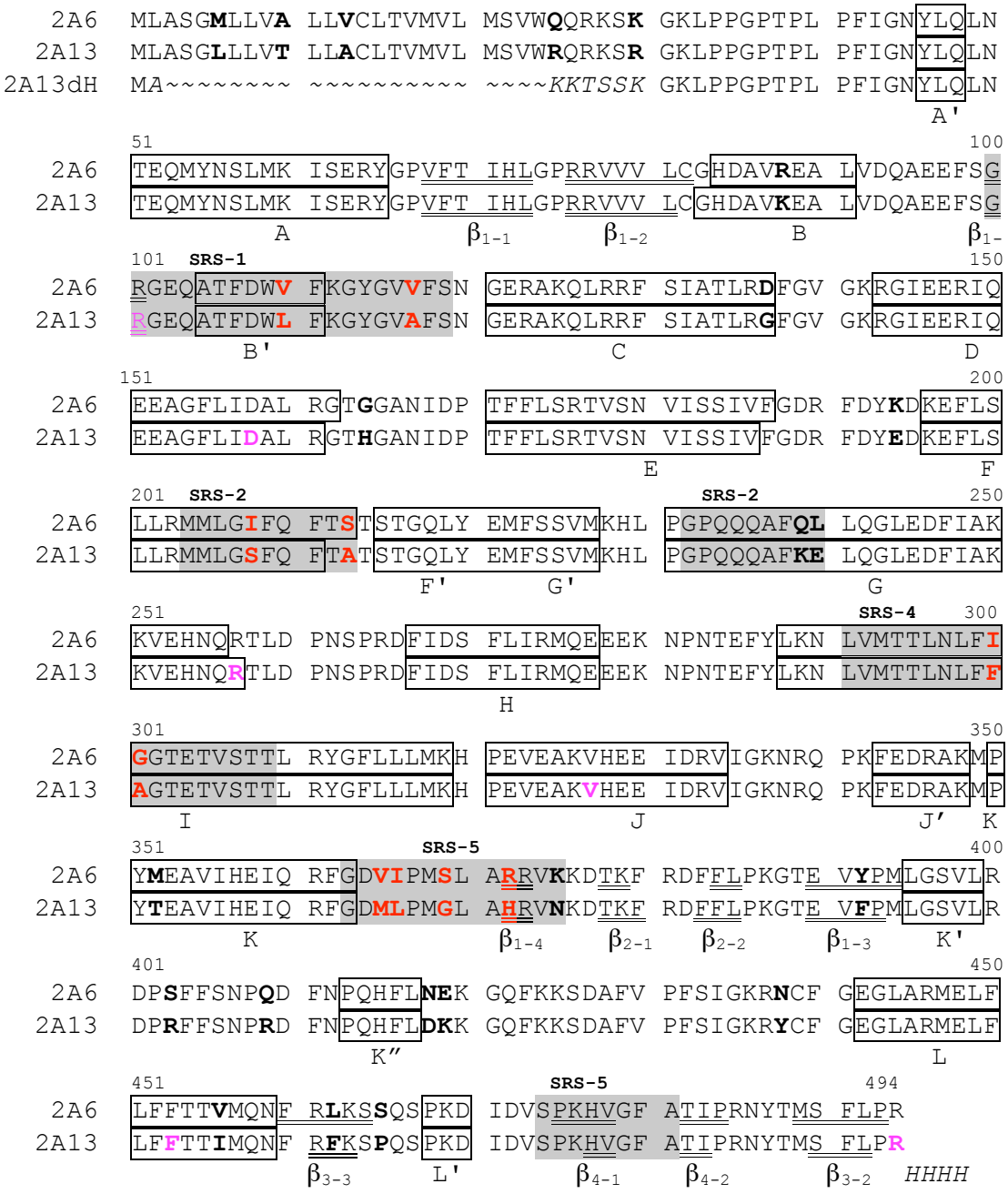


Figure 1.6. P450 2A6 and P450 2A13 amino acid sequence alignment completed with CL Workbench. P450 2A13dH (third line on first row of sequence alignment) is the sequence for the truncated CYP2A13 enzyme, ~ indicates a deletion. The 32 amino acid differences are highlighted in bold. The ten residues mutated in this study are colored red. The site of polymorphisms studied in CYP2A13 are colored pink. The secondary structures of CYP2A6 (PDB 1Z10) and CYP2A13 (PDB 2P85) were determined by DSSP (55). Boxes and double underline indicate α helices and β sheets, respectively. Italics indicate N- and C-terminal modifications of CYP2A13 to create the truncated, His-tagged form used for expression, purification, and crystallization. The substrate recognition sites are highlighted in grey.

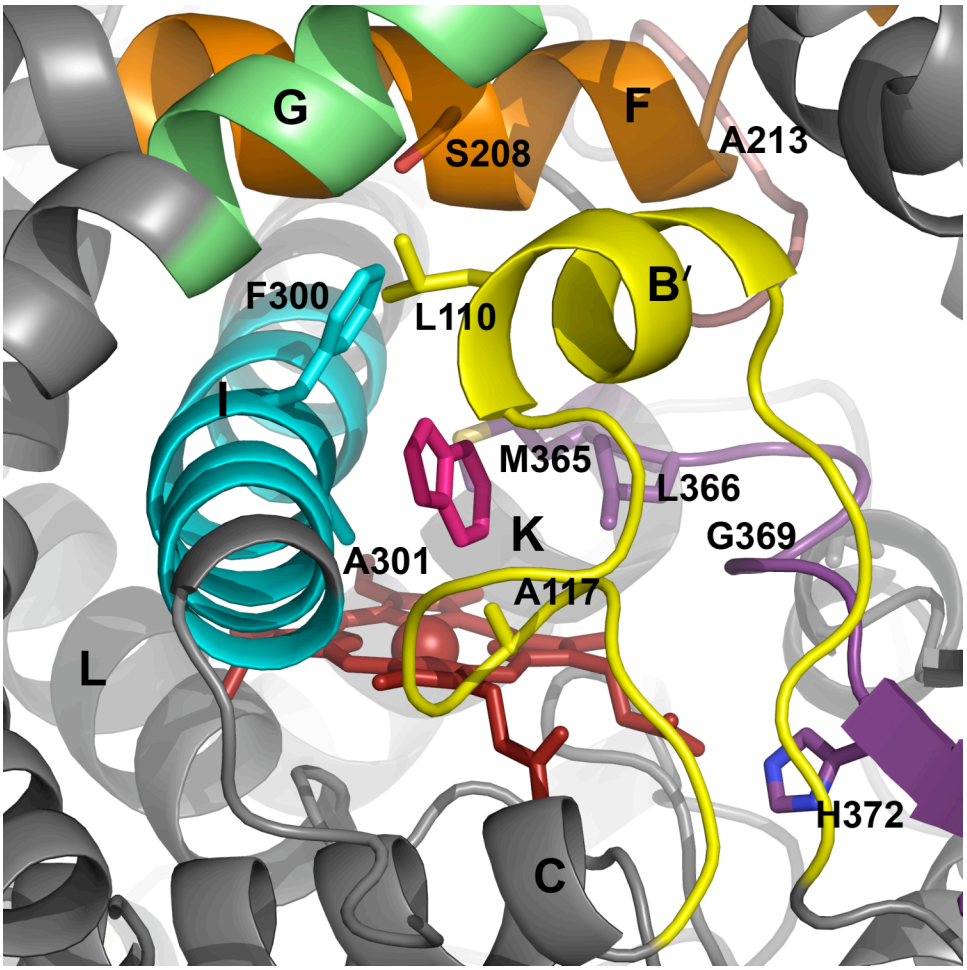


Figure 1.7. The active site of CYP2A13 (PDB 2P85) with indole in the active site (dark pink). The ten residues mutated in this study are labeled and shown as sticks. Substrate recognition sites are SRS-1 (yellow), SRS-2 (orange), SRS-3 (green), SRS-4 (cyan), SRS-5 (purple), and SRS-6 (pink). The heme group is colored red. Major helices are labeled.

References

1. Stark, K., and Guengerich, F. P. (2007) *Drug Metab. Rev.* **39**(2-3), 627-637
2. (2005) *Cytochrome P450: Structure, Mechanism, and Biochemistry*, Third Edition Ed., Kluwer Academic/ Plenum Publishers, New York
3. Lewis, D. F. V. (1996) *Cytochromes P450: Structure, Function and Mechanisms*, T. J. Press (Padstow) Ltd., Bristol
4. Bieche, I., Narjoz, C., Asselah, T., Vacher, S., Marcellin, P., Lidereau, R., Beaune, P., and de Waziers, I. (2007) *Pharmacogenetics and Genomics* **17**(9), 731-742
5. Nishimura M Fau - Yaguti, H., Yaguti H Fau - Yoshitsugu, H., Yoshitsugu H Fau - Naito, S., Naito S Fau - Satoh, T., and Satoh, T. (0031-6903 (Print))
6. Lewis, D. F. V. (2001) *Guide to cytochromes P450 structure and function*, Taylor & Francis, New York
7. Nelson, D. R. *D - 9214969* (- 1064-3745 (Print)), T - ppublish
8. Oscarson, M., and Ingelman-Sundberg, M. (2002) *Drug Metab Pharmacokinet* **17**(6), 491-495
9. Fernandezsalguero, P., Hoffman, S. M. G., Cholerton, S., Mohrenweiser, H., Raunio, H., Rautio, A., Pelkonen, O., Huang, J. D., Evans, W. E., Idle, J. R., and Gonzalez, F. J. (1995) *Am. J. Hum. Genet.* **57**(3), 651-660
10. Hoffman, S. M., Fernandez-Salguero, P., Gonzalez, F. J., and Mohrenweiser, H. W. (1995) *J. Mol. Evol.* **41**(6), 894-900
11. Le Gal, A., Dreano, Y., Lucas, D., and Berthou, F. (2003) *Toxicol. Lett.* **144**(1), 77-91
12. Yamano, S., Tatsuno, J., and Gonzalez, F. J. (1990) *Biochemistry (Mosc).* **29**(5), 1322-1329
13. Pelkonen, O., Sotaniemi, E. A., and Ahokas, J. T. (1985) *Br. J. Clin. Pharmacol.* **19**(1), 59-66
14. Ding, S., Lake, B. G., Friedberg, T., and Wolf, C. R. (1995) *Biochem. J.* **306 (Pt 1)**, 161-166
15. Wu, Z. L., Podust, L. M., and Guengerich, F. P. (2005) *J. Biol. Chem.* **280**(49), 41090-41100
16. Draper, A. J., Madan, A., and Parkinson, A. (1997) *Arch. Biochem. Biophys.* **341**(1), 47-61
17. von Weymarn, L. B., Chun, J. A., and Hollenberg, P. F. (2006) *Carcinogenesis* **27**(4), 782-790
18. von Weymarn, L. B., Chun, J. A., Knudsen, G. A., and Hollenberg, P. F. (2007) *Chem. Res. Toxicol.*
19. Su, T., Bao, Z. P., Zhang, Q. Y., Smith, T. J., Hong, J. Y., and Ding, X. X. (2000) *Cancer Res.* **60**(18), 5074-5079
20. von Weymarn, L. B., Zhang, Q. Y., Ding, X., and Hollenberg, P. F. (2005) *Carcinogenesis* **26**(3), 621-629
21. Zhang, X., D'Agostino, J., Wu, H., Zhang, Q. Y., von Weymarn, L., Murphy, S. E., and Ding, X. (2007) *J. Pharmacol. Exp. Ther.* **323**(2), 570-578

22. Nakajima, M. (2007) *Curr Opin Mol Ther* **9**(6), 538-544
23. Malaiyandi, V., Sellers, E. M., and Tyndale, R. F. (2005) *Clin. Pharmacol. Ther.* **77**(3), 145-158
24. Kubota, T., Nakajima-Taniguchi, C., Fukuda, T., Funamoto, M., Maeda, M., Tange, E., Ueki, R., Kawashima, K., Hara, H., Fujio, Y., and Azuma, J. (2006) *Pharmacogenomics J* **6**(2), 115-119
25. Nakajima, M., Fukami, T., Yamanaka, H., Higashi, E., Sakai, H., Yoshida, R., Kwon, J. T., McLeod, H. L., and Yokoi, T. (2006) *Clin. Pharmacol. Ther.* **80**(3), 282-297
26. Fujieda, M., Yamazaki, H., Kiyotani, K., Muroi, A., Kunitoh, H., Dosaka-Akita, H., Sawamura, Y., and Kamataki, T. (2003) *Drug Metab Pharmacokinet* **18**(1), 86-90
27. Wang, S. L., He, X. Y., Shen, J., Wang, J. S., and Hong, J. Y. (2006) *Toxicol. Sci.* **94**(1), 38-45
28. Schlicht, K. E., Michno, N., Smith, B. D., Scott, E. E., and Murphy, S. E. (2007) *Xenobiotica*, 1-11
29. Zhang, X., Su, T., Zhang, Q. Y., Gu, J., Caggana, M., Li, H., and Ding, X. (2002) *J. Pharmacol. Exp. Ther.* **302**(2), 416-423
30. Hadidi, H., Zahlsen, K., Idle, J. R., and Cholerton, S. (1997) *Food Chem. Toxicol.* **35**(9), 903-907
31. Oscarson, M., McLellan, R. A., Gullsten, H., Agundez, J. A., Benitez, J., Rautio, A., Raunio, H., Pelkonen, O., and Ingelman-Sundberg, M. (1999) *FEBS Lett.* **460**(2), 321-327
32. Ariyoshi, N., Sekine, H., Saito, K., and Kamataki, T. (2002) *Pharmacogenetics* **12**(6), 501-504
33. Kitagawa, K., Kunugita, N., Kitagawa, M., and Kawamoto, T. (2001) *J. Biol. Chem.* **276**(21), 17830-17835
34. Xu, C., Rao, Y. S., Xu, B., Hoffmann, E., Jones, J., Sellers, E. M., and Tyndale, R. F. (2002) *Biochem. Biophys. Res. Commun.* **290**(1), 318-324
35. Kiyotani, K., Yamazaki, H., Fujieda, M., Iwano, S., Matsumura, K., Satarug, S., Ujjin, P., Shimada, T., Guengerich, F. P., Parkinson, A., Honda, G., Nakagawa, K., Ishizaki, T., and Kamataki, T. (2003) *Pharmacogenetics* **13**(11), 689-695
36. Daigo, S., Takahashi, Y., Fujieda, M., Ariyoshi, N., Yamazaki, H., Koizumi, W., Tanabe, S., Saigenji, K., Nagayama, S., Ikeda, K., Nishioka, Y., and Kamataki, T. (2002) *Pharmacogenetics* **12**(4), 299-306
37. Oscarson, M., McLellan, R. A., Asp, V., Ledesma, M., Bernal Ruiz, M. L., Sinues, B., Rautio, A., and Ingelman-Sundberg, M. (2002) *Hum. Mutat.* **20**(4), 275-283
38. Ho, M. K., Mwenifumbo, J. C., Zhao, B., Gillam, E. M., and Tyndale, R. F. (2008) *Pharmacogenet Genomics* **18**(1), 67-75
39. Fukami, T., Nakajima, M., Yoshida, R., Tsuchiya, Y., Fujiki, Y., Katoh, M., McLeod, H. L., and Yokoi, T. (2004) *Clin. Pharmacol. Ther.* **76**(6), 519-527

40. Fukami, T., Nakajima, M., Higashi, E., Yamanaka, H., Sakai, H., McLeod, H. L., and Yokoi, T. (2005) *Drug Metab. Dispos.* **33**(8), 1202-1210
41. Fukami, T., Nakajima, M., Higashi, E., Yamanaka, H., McLeod, H. L., and Yokoi, T. (2005) *Biochem. Pharmacol.* **70**(5), 801-808
42. Al Koudsi, N., Mwenifumbo, J. C., Sellers, E. M., Benowitz, N. L., Swan, G. E., and Tyndale, R. F. (2006) *Eur. J. Clin. Pharmacol.* **62**(6), 481-484
43. He, X. Y., Tang, L., Wang, S. L., Cai, Q. S., Wang, J. S., and Hong, J. Y. (2006) *Int. J. Cancer* **118**(11), 2665-2671
44. He, X. Y., Shen, J., Hu, W. Y., Ding, X., Lu, A. Y., and Hong, J. Y. (2004) *Arch. Biochem. Biophys.* **427**(2), 143-153
45. He, X. Y., Shen, J., Ding, X., Lu, A. Y., and Hong, J. Y. (2004) *Drug Metab. Dispos.* **32**(12), 1516-1521
46. Bao, Z., He, X. Y., Ding, X., Prabhu, S., and Hong, J. Y. (2005) *Drug Metab. Dispos.* **33**(2), 258-261
47. Fukami, T., Nakajima, M., Sakai, H., Katoh, M., and Yokoi, T. (2006) *Drug Metab. Dispos.*
48. Nakajima, M., Itoh, M., Sakai, H., Fukami, T., Katoh, M., Yamazaki, H., Kadlubarz, F. F., Imaoka, S., Funae, Y., and Yokoi, T. (2006) *Int. J. Cancer* **119**(11), 2520-2526
49. Smith, B. D., Sanders, J. L., Porubsky, P. R., Lushington, G. H., Stout, C. D., and Scott, E. E. (2007) *J. Biol. Chem.*
50. Sansen, S., Yano, J. K., Reynald, R. L., Schoch, G. A., Griffin, K. J., Stout, C. D., and Johnson, E. F. (2007) *J. Biol. Chem.* **282**(19), 14348-14355
51. Herr, D., Bettendorf, H., Denschlag, D., Keck, C., and Pietrowski, D. (2006) *Arch. Gynecol. Obstet.* **274**(6), 367-371
52. Kimura, M., Yamazaki, H., Fujieda, M., Kiyotani, K., Honda, G., Saruwatari, J., Nakagawa, K., Ishizaki, T., and Kamataki, T. (2005) *Drug Metab. Dispos.* **33**(9), 1361-1366
53. Gotoh, O. (1992) *J. Biol. Chem.* **267**(1), 83-90
54. Yano, J. K., Hsu, M. H., Griffin, K. J., Stout, C. D., and Johnson, E. F. (2005) *Nat Struct Mol Biol* **12**(9), 822-823
55. Kabsch, W., and Sander, C. (1983) *Biopolymers* **22**(12), 2577-2637

Chapter 2.

Methods and Procedures

Introduction

Several of the methods used to compare CYP2A13 and CYP2A6 are common to multiple parts of this study. To avoid redundancy, these methods and procedures are presented in this chapter and will be referenced throughout the rest of this thesis. The majority of these methods and procedures are well established in the cytochrome P450 field, although details in the purification and assay protocols have been optimized.

Methods and procedures necessary to the overall study, for example, protein engineering and site directed mutagenesis, which were not actually performed by me are described with a short description of the general method. Purification of cytochrome b_5, an accessory protein necessary for the phenacetin metabolism assay, is not described because a prior lab member purified this protein (1).

Cytochrome P450 Modifications and Mutations

Modifications: The CYP2A13 and CYP2A6 proteins were altered to form CYP2A13dH and CYP2A6dH, referred to CYP2A13 and CYP2A6 throughout. The N-terminal transmembrane sequence (Δ2-23) was truncated and four charged residues were added to the N-terminus in order to increase the expression yield and decrease the hydrophobic nature of the protein. In addition, four histidine residues were added to the C-terminus to facilitate protein purification with a nickel affinity column (2). All of the CYP2A enzymes used in this study were truncated and histidine tagged.

Mutations: The mutations used in this study were made by swapping non-identical amino acids found in one CYP2A enzyme into the other. For example, the CYP2A13 mutations were produced using the modified CYP2A13 protein as a template and substituting the non-corresponding CYP2A6 residues. Multiple mutations were made additively, using a single or double mutant as the template. All of the mutants were constructed using the QuikChange® kit (Stratagene, La Jolla, CA) (3). The desired mutation is generated by annealing two synthetic oligonucleotide primers containing the desired mutation to opposite strands of the DNA vector (pKK2A13dH or pKK2A6dH) (4). If possible, primers were designed with incorporation or deletion of a distinctive restriction site so the desired mutation could be easily detected by restriction digest analysis (Table 2.1). Thermocycling was used to denature the DNA double stranded vector, anneal the primers to the DNA vector, and then to extend the primers (4). This resulted in a double stranded mutant plasmid with staggered nicks. Methylated template and hemi methylated hybrid wild type/mutant DNA were destroyed with *Dpn* I endonuclease (4). The mutant plasmid DNA was transformed into XL1-Blue supercompetent cells for nick repair and replication (4). The presence of the mutation was identified by restriction digest followed by gel electrophoresis. Inclusion of the desired mutation and the absence of additional alterations in the mutated DNA were verified by DNA sequencing of the complete gene.

Table 2.1. Sequences 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 wild type sequence. Underline indicates the location of the desired amino acid substitution. Italics indicate synonymous nucleotide changes that alter a restriction site and were used to facilitate identification of plasmids containing the desired mutation.

CYP2A13 Mutation	**Sequence (5' to 3')**	**Restriction Site Altered**
L110V	CC ACC TTC GAC TGG ***G***TC TTC AAA GGC TAT GGC	delete SapI
A117V	GGC TAT GGC GTG G***TC*** TTC AGC AAC GGG	add Bbs I
S208I	CGC ATG ATG CTG GGA A***T***C TTC CAG TTC ACG GCA AC C	delete HindIII
A213S	GGA AGC TTC CAG TTC ACG ***TCG*** ACC TCC ACG GGG CAG C	new SalI
F300I	CC CTG AAC CTC TTC **A**TT GCG GG***T*** AC***C*** GAG ACC GTG AGC AC C	add KpnI
A301G	CC CTG AAC CTC TTC TTT G**G**G GG***T*** AC***C*** GAG ACC GTG AGC	add KpnI
M365V	C CAA AGA TTT GGA GAC **G**T***C*** CTC CCC ATG GGT TTG G	add AatI I
L366I	C CAA AGA TTT GGA GAC ATG **A**TC CC***G*** ATG GGT TTG GCC C	delete NcoI
G369S	GGA GAC ATG CTC CCC ATG **A**G**C** TT***A*** GCC CAC AGG GTC AAC AAG G	add Bpu1102I
H372R	CCC ATG GGT TTG GCC C**G**C AGG GT***T*** AAC AAG GAC ACC AAG TTT C G	add HpaI

Table 2.2. Sequences for one of the oligonucleotides used in the construction of each CYP2A6 mutation(s). Bold indicates changes from the wild type 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. Multiple mutations were made by sequential addition of changes using these same oligonucleotides.

CYP2A6 Mutation	**Sequence (5' to 3')**	**Restriction Site Altered**
V110L	GCC ACC TTC GAC TGG **C**T**G** TT***T*** AAA GGC TAT GGC	add DraI
V117A	GGC TAT GGC GTG G**C**A TTC AGC AAC GGG	add BbsI
I208S	GC ATG ATG CTA GGA A***G***C TTC CAG TTC ACG TCA ACC TCC	delete HindIII
S213A	GGA ATC TTC CAG TTC ACG **G**C**G** AC***G*** TCC ACG GGG CAG C	add AatII
I300F	CG TTG AAC CTC TTC **A**T**C** GGG GG***T*** AC***T*** GAG ACC GTG AGC	add KpnI
G301A	CG TTG AAC CTC TTC ATT G**C**G GG***T*** AC***C*** GAG ACC GTC AGC	add KpnI
V365M	CAA AGA TTT GGA GAC **A**TG ATC CCC ATG AGT TTG	none
I366L	C CAA AGA TTT GGA GAC GT***C*** **C**T**G** CCC ATG AGT TTG GCC C	add AatII
S369G	C GTG ATC CCC ATG **G**GT TTG G*C*C CGC AGA GTC AAA AAG G	delete NcoI
R372H	CCC ATG AGT TTG GCC C**A**C AGA GT***G*** AAA AAG GAC ACC	delete DraIII
I300F/G301A	CC ACG TTG AAC CTC TTC **T**T**C** G**C**G GG***T*** AC***C*** GAG ACC GTC AGC ACC	add KpnI
V365M/I366L	CAA AGA TTT GGA GAC **A**TG **C**T**G** CCC ATG AGT TTG GCC CGC	none

Table 2.3. The sequence for the forward primers used to generate polymorphic variants of CYP2A13. The reverse primers were 100% complementary. Alterations from the wild type nucleotide sequence are shown in bold. Underlines indicate the location of the desired mutation. Italics indicate additional silent changes that either restriction sites introduced or deleted to facilitate identification of plasmids containing the desired polymorphism.

Allele	Protein Changes	Sequence (5' to 3')	Restriction Site Altere d
CYP2A13*2	R257C	G GTG GAG CAC AAC CAG **T**GC AC***T*** CT***A*** GAT CCC AAT TCC CC	add XbaI
CYP2A13*4	R101Q	CAG GCT GAG GA***A*** TTC AGC GGG C**A**A GGC GAG CAG G	add EcoRI
CYP2A13*5	F453Y	GCC AGA ATG GAG CT***G*** TTT CTC TTC T**A**C ACC ACC ATC	delete SacI
CYP2A13*6	R494C	GC TTC CTG CCC **T**GC CAT CAC CAC CAT TGA ***CG***T CCG GCT GAG G	add EcoRV
CYP2A13*8	D158E	C TTC CTC ATC GA**A** GCC CTC CG***C*** GGC_ACG CAC GG	add SacII
CYP2A13*9	V323L	G AAG CAC CCA GAG **C**T**C** GAG GCC AAG GTC C	add XhoI

Cytochrome P450 Expression and Purification

Transformation: Ampicillin resistant plasmids encoding the truncated, His-tagged variants of the cytochrome P450 enzymes were transformed into *E. coli* tetracycline resistant TOPP-3 cells (Stratagene, La Jolla, CA) (2). The transformation was completed by incubating 100 µL of *E. coli* TOPP3 cells with 0.5 µL of the desired P450 plasmid (pKKP450dH) (~ 50 ng/mL) on ice for 30 minutes. The cells were then heat shocked in a 42°C water bath for 30-45 seconds before the addition of 1.7 µL of a 1:10 dilution of 100% β-mercaptoethanol (BME) and 1 mL of super optimal catabolite repressive (SOC) media. Samples were incubated at 37°C for one hour while shaking at 250 rpm. The samples were centrifuged at briefly to pellet the cells at 10,000 g. All but 100 µL of the supernatant was removed. The pellet was resuspended and plated on to Luria-Bertani media (LB) plates containing 1 µL of 50 mg/mL ampicillin per 1 mL of agar and incubated overnight at 37°C.

Expression: One colony of transformed TOPP3/pKKP450dH was selected and used to grow a small-scale culture by inoculating a test tube containing 5 mL of LB media, 5 µL of 50 mg/mL ampicillin, and 10 µL of 12.5 mg/mL tetracycline. The culture was incubated at 37°C for 7-8 hours with shaking. After incubation, a 200 mL medium-scale starter culture was inoculated with 50 µL from the 5 mL culture or a few cells from a glycerol stock (culture from the last expression stored at -80°C in a 1:5 ratio with 100% glycerol) along with 200 µL of 50 mg/mL ampicillin and 400 µL of 12.5 mg/mL tetracycline. The medium-scale starter cultures were grown overnight at 37°C with shaking at 250 rpm.

Eighteen one-liter flasks containing 250 mL of Terrific Broth (TB media with 25 mL of TB salts) were each inoculated with 15 mL of the medium-scale overnight culture and 250 μL of 50 mg/mL ampicillin. The cells were grown to an OD_{600} of between 1.0-1.5 at 37°C with shaking at 250 rpm. Once the cells reached this absorbance, 20 mg aminolevulinic acid (ALA) and 60 mg isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to each 250 mL culture. ALA is an essential component required to synthesize the heme porphyrin ring found in CYP450s. IPTG was used to induce gene expression because it prevents the *lac* repressor from interfering with the transcription of the gene. After the addition of ALA and IPTG, the incubation temperature was lowered to 30° C and the shaking to 190 rpm. Cells were grown for three days. Once the incubation period was completed, cells were harvested by centrifugation at 6371 g for ten minutes. Cell pellets were collected and stored at -80° C for later purification.

Purification: The purification of cytochrome P450 enzymes began with the extraction of the protein from the whole bacteria cells. The frozen cell pellet was thawed, and re-suspended with 200 mL of resuspension buffer (20 mM potassium phosphate buffer (KPi), pH 7.4, 20% glycerol). Lysozyme (Sigma) was added to a final concentration of 0.3 mg/mL to initiate hydrolysis of peptidoglycan in the bacterial cell wall, and the solution was stirred at 4°C. After 30 minutes, 200 mL of 4° C distilled water was added to the mixture in order to increase the hypotonic nature of the solution. The mixture stirred for ten more minutes and was centrifuged at 8671 g for 15 minutes. After centrifugation, the cell pellet was flash frozen in a dry ice-

ethanol slurry. The pellet was resuspended in 100 mL of resuspension buffer (500 mM KPi pH 7.4, 300 mM NaCl, 20% glycerol) and homogenized with a homogenizer on ice. The suspension was sonicated on ice for three periods of 30 seconds. The combined methods of flash freezing, homogenization, and sonication completed cell lysis. A second centrifugation at 8671 g removed cellular debris and the detergent Cymal-5 was added (at a final concentration of 4.8 mM) to solubilize the membrane-bound CYP450s. The solution was stirred at 4°C for approximately one hour to ensure maximal CYP450 solubility. To separate the solubilized proteins from the other cell components, the solution was centrifuged at 70,000 g for one hour.

The crude protein solution was purified using nickel nitrilotriacetic acid (NiNTA) metal affinity and carboxymethyl (CM) sepharose chromatography on a Fast Protein Liquid Chromatography system (FPLC). The NiNTA column (QIAGEN) was able to significantly increase protein purity because of the four His residues engineered onto the carboxyl terminus (His-tag) of our P450 proteins. The His-tag of the protein complexes with the Ni^{2+} bound to the NTA column matrix (5). The sample was loaded onto the NiNTA column at a 1 mL/min flow rate. Non-binding impurities were removed from the column with a two-column volume wash of loading buffer (100 mM KPi pH 7.4, 20% glycerol, 0.2 M NaCl, and 4.8 mM Cymal-5). Proteins with weak affinity were washed off with a second wash additionally containing 8 mM histidine. The His-tagged P450 proteins were eluted from the column by displacement with 80 mM histidine. The sample was diluted five-fold with CM wash buffer (5 mM KPi pH 7.4, 1 mM EDTA, 20% glycerol, 4.8

mM Cymal-5) to reduce the salt concentration before loading on the ion exchange column.

The diluted protein was purified using ion exchange chromatography. Protein was loaded onto a CM column, a cation exchanger (6). Cytochrome P450 2A6 and 2A13 both have a predicted isoelectric point of ~9.2 (ExPASy tools), which means that at the pH of 7.4, the protein will have a net positive charge. After protein was bound to the negatively charged CM column, ten column volumes of CM wash buffer was used to remove impurities. The protein was eluted from the column with a buffer containing salt (50 mM KPi pH 7.4, 1 mM EDTA, 20% glycerol, 0.5 M NaCl).

The purification was monitored throughout by comparing ~417/280 nm absorbance. The final concentration of the protein was determined by both an absolute and CO difference spectra (7). The protein was divided into aliquots and stored at -80°C.

Rat NADPH Reductase Expression and Purification

Expression: Cells were grown from a HMS174(DE3)/pETOR262 glycerol stock. Twenty five µL of the glycerol stock was inoculated into 100 µL of LB media with 0.1 mL of kanamycin (50 mg/mL) in a 250 mL flask and grown overnight at 37° C with 250 rpm shaking. The next morning, 0.5 mL of the freshly grown culture was inoculated into 250 mL of TB media, 25 mL of TB salts, and 0.25 mL of kanamycin (50 mg/mL). A total of 12 flasks (3 Liters) were grown at 37° C at 250 rpm for four hours or until the OD_{550} absorption was greater than 0.8. Then, IPTG (0.5 mM final concentration) was added to each flask. After the addition of IPTG, cells were grown

at 30° C at 190 rpm for 48 hours. Cells were harvested by centrifugation at 6371 g for 15 minutes at 4° C.

Purification: Thawed cells were resuspended in 150 mL of resuspension buffer (75 mM Tris pH 8.0, 0.25 M sucrose, and 0.25 mM EDTA pH 8.0). Lysozyme (3 mg) was added to the resuspended cells and the solution was stirred for 30 minutes at 4° C to initiate cell lysis. The spheroplasts were pelleted by centrifugation at 4424 g at 4° C for 30 minutes. The pellets were resuspended in 100 mL of affinity buffer (50 mM Tris pH 7.7, 0.1 mM EDTA pH 8.0, 0.05 mM DTT, 10% glycerol, and 0.1% triton X-100). The samples were hemolyzed with a tissue grinder and then sonicated for ten seconds ten times. The mixture was stirred at 4° C for three hours. The solution was centrifuged at 80,000 g for 45 minutes.

The crude supernatant was purified by DE-52 anion exchange chromatography. The DE-52 column was washed with 500 mL of affinity buffer at 80-100 mL/min following sample loading. Gradient elution was performed with 1200 mL of affinity buffer with a concentration ranging from 0 – 350 mM KCl at a flow rate of 1 mL/min. All of the yellow fractions were pooled and FMN added to a final concentration of 1 μM.

The pooled fractions were loaded at a flow rate of 0.16 mL/min onto a 2′5′-ADP-Sepharose column equilibrated with 120 mL of affinity buffer. The column was washed with 40 mL of affinity buffer at 0.2 mL/hr. The sample was eluted with 10 mL of 2 mM 2′AMP in affinity buffer at 0.16 mL/min. All of the yellow fractions were pooled.

The yellow fractions containing reductase were dialyzed (10 kDa molecular weight cutoff) against 2 liters of dialysis buffer (30 mM KPi pH 7.7, 20% glycerol, 0.1 mM EDTA, and 0.002 mM FMN) for at least 8 hours.

The concentration of reductase present after dialysis was determined by a time course reaction measuring the amount of reductase required to catalyze the reduction of one micromole cytochrome c per minute. In a cuvette, 10 µL of cytochrome c (5 mM, Sigma) and 2 µL of a 1:50 dilution of reductase from the dialysis were combined to a total volume of 990 µL. The reaction was initiated by the addition of 10 µL of NADPH (10 mg/mL). The absorption at 550 nm was measured for five minutes. The reductase concentration was determined in units/mL (Absorption change/minutes*Dillution factor) where 1 nmol of reductase is ~4 units.

Carbon Monoxide Difference Assay (7)

Instrument: UV-2101 UV/visible scanning spectrophotometer (Shimadzu Scientific Instruments, Inc., Columbia, MD)

Procedure: An unknown concentration of protein sample was diluted 1:9 into buffer to a total volume of 900 µL in a quartz cuvette. A few crystals of dithionite were added to the cuvette and the cuvette was inverted before being placed into the spectrophotometer. The sample was zeroed and then a baseline was taken from 500 - 400 nm. The sample was removed from the spectrophotometer and 50 bubbles of carbon monoxide added. After the addition of carbon monoxide, the cuvette was inverted to mix and placed back in the spectrophotometer. The sample was scanned from 500-400 nm. Scans were taken every few minutes until the increase in

absorbance at 450 nm was maximal, as indicated by more than two scans at the same absorbance or a noticeable drop in absorbance.

Analysis: Concentration of P450 protein present was determined using Beer's law: $c = \frac{A}{\varepsilon l}$ where c is concentration in μM, l is the pathlength (1 cm), A is absorption, and ε is the absorptivity coefficient which is equal to 0.091 $\mu M^{-1}cm^{-1}$.

Ligand Binding Assay(7,8)

Instrument: UV-2101 UV/visible scanning spectrophotometer (Shimadzu Scientific Instruments, Inc., Columbia, MD)

Procedure: Protein samples were diluted to a concentration of 1 μM in 100 mM potassium phosphate buffer (pH 7.4) and were divided equally between reference and sample 1 mL quartz cuvettes. Concentrated ligand stocks were prepared in 100% ethanol. The sample baseline was taken from 500 – 300 nm before beginning the titration. Ligand was added to the sample cuvette, mixed, and allowed to equilibrate for 8 minutes. An equal amount of solvent was added to the reference cuvette for every addition of ligand to the sample cuvette. The total amount of solvent was less than 2% for each titration. After each ligand addition, a difference spectrum was recorded from 500 – 300 nm. Additions were continued until there was no change in absorption band observed upon addition of ligand.

Analysis: Graphpad Prism 4 (Graphpad Software, San Diego, CA) is used to determine the K_d and ΔA_{max} with nonlinear least-squares regression fitting to the equation: $\Delta A = \frac{\Delta A_{max}}{2P}[P + S + K_s - \sqrt{(P + S + K_s)^2 - 4PS}]$.

Phenacetin Metabolism Assay

Instrument: Prominence HPLC System (Shimadzu Scientific Instruments, Inc., Columbia, MD)

Reconstitution of Enzymes: CYP2A13, CYP2A6, and mutant enzymes were reconstituted with rat NADPH reductase (purified as described previously) and cytochrome b_5 (purified by Brian Smith, a previous lab member) for 20 minutes at room temperature before use. A 1:6:2 ratio of CYP450 enzymes to reductase to cytochrome b_5 was used for all metabolism experiments since this ratio gave optimal metabolism (*data not shown*).

Phenacetin O-Deethylation: Reconstituted protein system (30 pmol CYP450) was added to 100 mM potassium phosphate buffer (pH 7.4) with 0 – 250 μM phenacetin to total volume of 300 μL. Phenacetin *O*-deethylation activity was determined using a phenacetin concentration of 250 μM and a final NADPH (Sigma) concentration of 500 μM. The reactions were initiated with the addition of 30 μL of NADPH followed by incubation in a water bath at 37ºC for 30 minutes. Reactions were terminated with 15 μL 60% perchloric acid, denaturing all the enzymes present. Samples were centrifuged for five minutes at 4300 g to remove precipitated protein and the supernatant was retained for analysis by HPLC. HPLC analysis was completed using a 4.6 x 250 mm Phenomenex Luna C18 column with a flow rate of 1 mL/min and an isocratic mobile phase consisting of 8% acetonitrile with 50 mM potassium phosphate solution (pH 4.9). Acetaminophen standards were prepared in the same manner as the samples except for the absence of the reconstituted protein system. The lack of

reconstituted protein system did not affect the standard curve as indicated by preliminary experiments (data not shown).

Analysis: Graphpad Prism 4 (Graphpad Software, San Diego, CA) was used to determine the K_m and k_{cat} with nonlinear least-squares regression fitting to the equation: $Activity = \frac{k_{cat}}{2P}[P + S + K_m - \sqrt{(P + S + K_m)^2 - 4PS}]$.

References

1. Holmans, P. L., Shet, M. S., Martinwixtrom, C. A., Fisher, C. W., and Estabrook, R. W. (1994) **- 312**(- 2), - 565
2. Smith, B. D., Sanders, J. L., Porubsky, P. R., Lushington, G. H., Stout, C. D., and Scott, E. E. (2007) *J. Biol. Chem.*
3. Schlicht, K. E., Michno, N., Smith, B. D., Scott, E. E., and Murphy, S. E. (2007) *Xenobiotica*, 1-11
4. *QuikChange II Site-Directed Mutagenesis Kit Instruction Manual*, Stratagene, La Jolla, CA
5. (2003) *The QIAexpressionist: A handbook for high-level expression and purification of 6xHis-tagged proteins*, 5 Ed., QIAGEN Distributors, Valencia, Ca
6. (2000) *Ion Exchange Chromatography: Principles and Methods*, AA Ed., BioProcess
7. Schenkman, J. B. a. I. J. (2006) Spectral Analyses of Cytochromes P450. In: Phillips, I. R. a. E. A. S. (ed). *Methods Mol. Biol.*, Second Edition Ed., Humana Press Inc., Totowa, NJ
8. Yano, J. K., Hsu, M. H., Griffin, K. J., Stout, C. D., and Johnson, E. F. (2005) *Nat Struct Mol Biol* **12**(9), 822-823

Chapter 3.

Ligand Binding Studies

Introduction

Ligand binding studies have been utilized to study P450s since the 1960s (1-3). Schenkman *et al*. noted that the addition of various ligands into a cuvette containing liver microsomes resulted in spectral changes (*2*). Three distinct types of spectral changes are recognized: type I, type II, and reverse type I (Figure 3.1) (1,2).

Type I spectra are identified by the formation of a peak at ~385 nm when a substrate is bound and a trough at ~420 nm (Figure 3.1, Panel A). The majority of type I spectra shifts are the result of substrate binding. This type of spectral change reflects a shift from low-spin, 6-coordinate iron state when a water molecule is bound to the iron to high-spin, 5-coordinate iron when a ligand binds to the active site of the protein, displacing the water molecule (1,4).

Reverse type I spectra have a peak formation at ~420 nm and a trough at ~385 nm (Figure 3.1, Panel B) (*1*). Ligands causing reverse I spectra may bind in such a way as to increase the water coordination to the heme iron or they may bind at a different location in the active site than type I ligands (5). Typical reverse type I ligands are hard to characterize because a reverse type I ligand may shift spectra at higher ligand concentrations or may not be a reverse type I when bound to a different CYP450. Phenacetin is frequently mentioned as a reverse type I binder in early literature, although it is also reported as causing a type I spectral shift (5).

The third category of spectral binding shift is the type II, characterized by a peak at ~425-435 nm and a broad trough at 390-410 nm (Figure 3.1, Panel C). Type II ligands are generally nitrogen-containing inhibitors that can replace the water molecule bound to the heme iron, thus stabilizing a low-spin, 6-coordinate iron.

In this study, we utilized spectral binding shifts to survey a series of four structurally diverse cytochrome P450 2A ligands and determine the binding affinities for both CYP2A6 and CYP2A13 enzymes. These four ligands were coumarin, 2'-methoxyacetophenone (MAP), phenethyl isothiocyanate (PEITC), and phenacetin. Additionally, ten active site residues that differ between CYP2A13 and CYP2A6 were systematically examined for their contributions to ligand binding. These ligand binding studies identified a number of substitutions that changed the K_D significantly for one or more substrates. This allowed us to determine if all four ligands interact with the same amino acid residues.

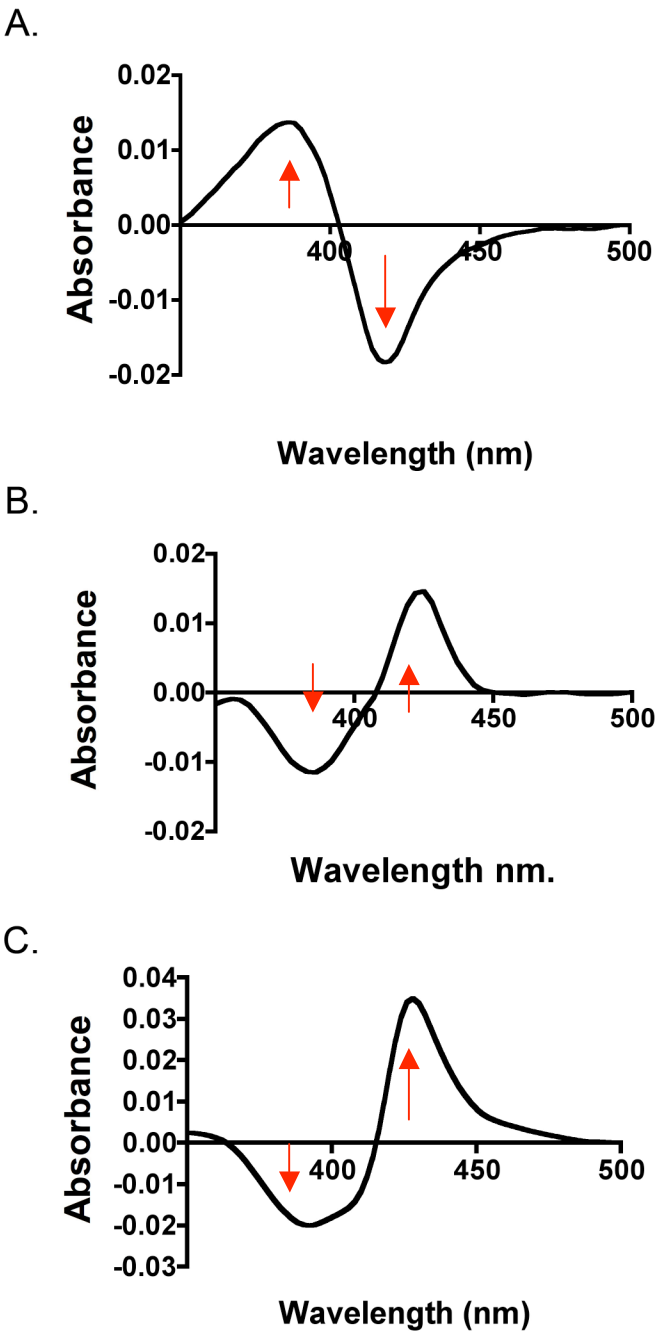


Figure 3.1. The three types of spectral binding spectra are illustrated above. A. Type I spectra of phenacetin binding to CYP2A13. B. Reverse Type I spectra of agroclavine binding to rat liver microsomes, based on figure 1 from Jefcoate (1). C. Type II spectra of pilocarpine binding to CYP2A6 I208S/I300F/G301A/S369G.

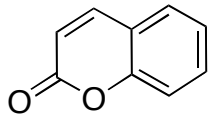
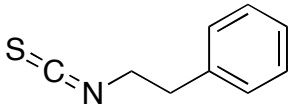
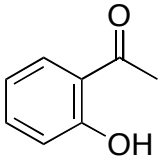
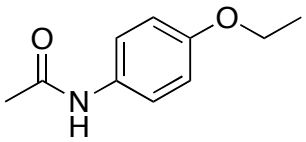
Methods

All proteins used in this chapter were expressed and purified according to the protocols found in chapter 2. The procedure for the spectral ligand binding assay is also found in chapter 2.

Results

Binding of structurally diverse ligands to CYP2A6 and CYP2A13 wild type proteins: The 2A ligands coumarin, 2'-methoxyacetophenone (MAP), phenethyl isothiocyanate (PEITC), and phenacetin all displayed type I binding spectra with a shift from ~420 nm to ~385 nm (Figure 3.1, Panal A). All of the ligands surveyed bound more tightly to CYP2A13 than to CYP2A6 (Table 3.1). The ligand with the largest difference in binding affinity is phenacetin, which binds with a K_D of ~34 μM for CYP2A13 but does not cause a spectral shift in CYP2A6 at 250 μM phenacetin. The ligand MAP had the second largest difference in binding affinity, it binds 87-fold more tightly to CYP2A13 than to CYP2A6. CYP2A13 has an affinity of 0.85 μM for this ligand, while the K_D for CYP2A6 is 74 μM. The inhibitor PEITC bound to CYP2A13 14-fold more tightly than to CYP2A6 with K_D values in the sub-micromolar range for 2A13 and low micromolar range for CYP2A6. In contrast to these three ligands, the binding affinity of both CYP2A enzymes appeared to be very similar for coumarin with K_D values ~3 μM.

Table 3.1. Binding constants for various ligands to cytochrome 2A6 vs. cytochrome 2A13. Values are in parenthesis indicate individual trials, value below is the average of the duplicate measurements. A third trial was preformed in some cases. All trials had R^2 goodness of fit values >0.99.

Ligand	**Structure**	**K_D (μM)** **2A6**	**2A13**	$\frac{K_D 2A6}{K_D 2A13}$
coumarin		(2.7, 3.4) 3.1	(2.0, 3.3) 2.7	1.1
phenethyl isothiocyanate		(5.4, 6.9) 6.2	(0.31, 0.54) 0.43	14.
2'-methoxyacetophenone		(65, 69, 104) 74.	(1.0, 0.66) 0.85	87.
phenacetin		N.D.	(30, 35, 37) 34.	N.D.

N.D. - K_D not determined due to little or no spectra shift.

Ligand binding to mutants of CYP2A13: After determining dissociation constants for both CYP2A13 and CYP2A6 wild type enzymes, we proceeded to characterize the ten CYP2A13 proteins into which single active site residues from CYP2A6 were substituted. As with the wild type enzymes, the most pronounced differences in binding affinity were with phenacetin as the ligand (Table 3.2). The wild type CYP2A13 had a K_D value of 34 μM (Figure 3.2, panel A), while a binding constant could not be determined for CYP2A6 (Figure 3.2, panel B). Two CYP2A13 mutants, F300I and A301G, reduced phenacetin binding to the point where a K_D value could not be determined (Figure 3.2, Panel C). More than a three-fold decrease in phenacetin affinity resulted with the mutation of residues 365, 369, and 208 (Figure 3.3). The substitution of Ile for Leu366 led to the only mutant with a substantial increase in binding affinity with a K_D of ~2 μM (Table 3.2). All other mutations had little affect on phenacetin ligand binding affinity.

MAP binding to the CYP2A13 mutants revealed that four modifications resulted in substantially different K_D values compared to the wild type CYP2A13 enzyme. Substitution of isoleucine for phenylalanine 300 (F300I) resulted in a >17-fold increase in K_D to 15 μM, from the wild type affinity of 0.85 μM (Figure 3.3). S208I had a 14-fold increase to 12 μM. A301G was characterized as binding MAP with a K_D of 8.2 μM, a decrease in affinity of more than 9-fold (Figure 3.3). A117V had a more moderate 4.8-fold decrease in affinity for MAP. The remaining six mutations had relatively little effect on MAP binding, although L366I slightly improved binding affinity (0.56 μM).

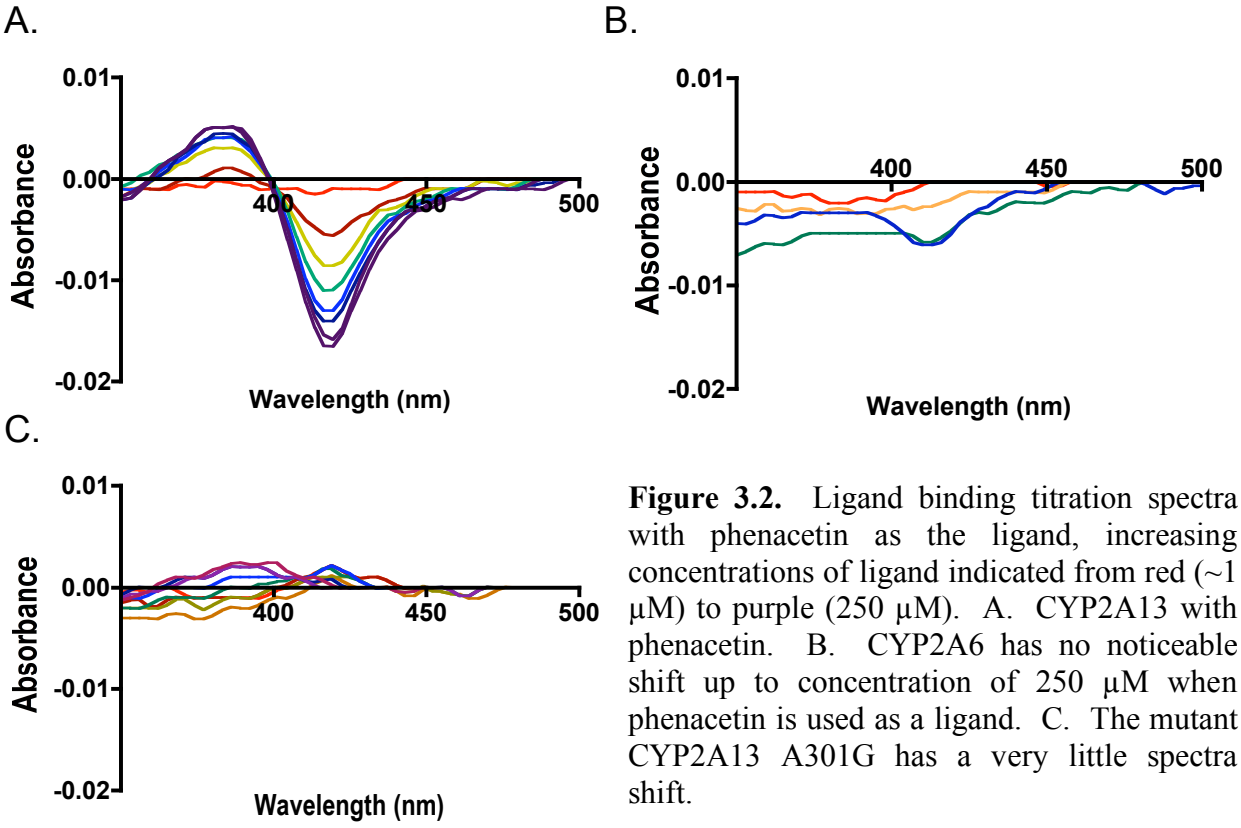


Figure 3.2. Ligand binding titration spectra with phenacetin as the ligand, increasing concentrations of ligand indicated from red (~1 µM) to purple (250 µM). A. CYP2A13 with phenacetin. B. CYP2A6 has no noticeable shift up to concentration of 250 µM when phenacetin is used as a ligand. C. The mutant CYP2A13 A301G has a very little spectra shift.

Table 3.2. K_D values for wild type CYP2A enzymes and the ten CYP2A13 mutants. All trials represent the mean of duplicate trials (values from individual trials in parenthesis) for wild type enzymes and for all mutants with a greater than 3-fold difference in K_D from wild type CYP2A13. The KD of the mutant CYP2A13 enzymes with phenacetin were only done once, these results are preliminary.

Enzymes	MAP K_D (μM)	PEITC K_D (μM)	Coumarin K_D (μM)	Phenacetin K_D (μM)
2A6	(65, 69, 104)	(5.4, 6.9)	(2.7, 3.4)	
	74.0	6.2	3.1	N.D.
2A13	(1.0, 0.66)	(0.31, 0.54)	(2.0, 3.3)	(30, 35, 37)
	0.85	0.43	2.7	34.
2A13 Mutants:				
L110V		(0.43, 0.48)		
	1.6	0.46	1.7	18.
A117V	(3.7, 4.6)			
	4.1	1.2	1.8	20.
S208I	(13, 11)	(0.99, 1.7)	(8.0, 7.9)	
	12.	1.4	7.9	110.
A213S				
	2.9	0.87	3.4	65.
F300I	(16, 14)	(6.2, 5.8)	(14.1, 13.1)	
	15.	6.0	14.	N.D.
A301G	(8.3, 8.0)	(2.0, 4.0)		
	8.2	3.1	4.2	N.D.
M365V		(1.6, 1.8)	(4.4, 5.6)	
	1.7	1.7	5.0	130.
L366I			(0.35, 0.80)	
	0.56	0.34	0.57	2.3
G369S		(1.8, 1.8)	(5.5, 6.6)	
	1.6	1.8	6.0	97.
H372R				
	0.77	1.0	1.5	20.

N.D. - K_D not determined due to little or no spectra shift.

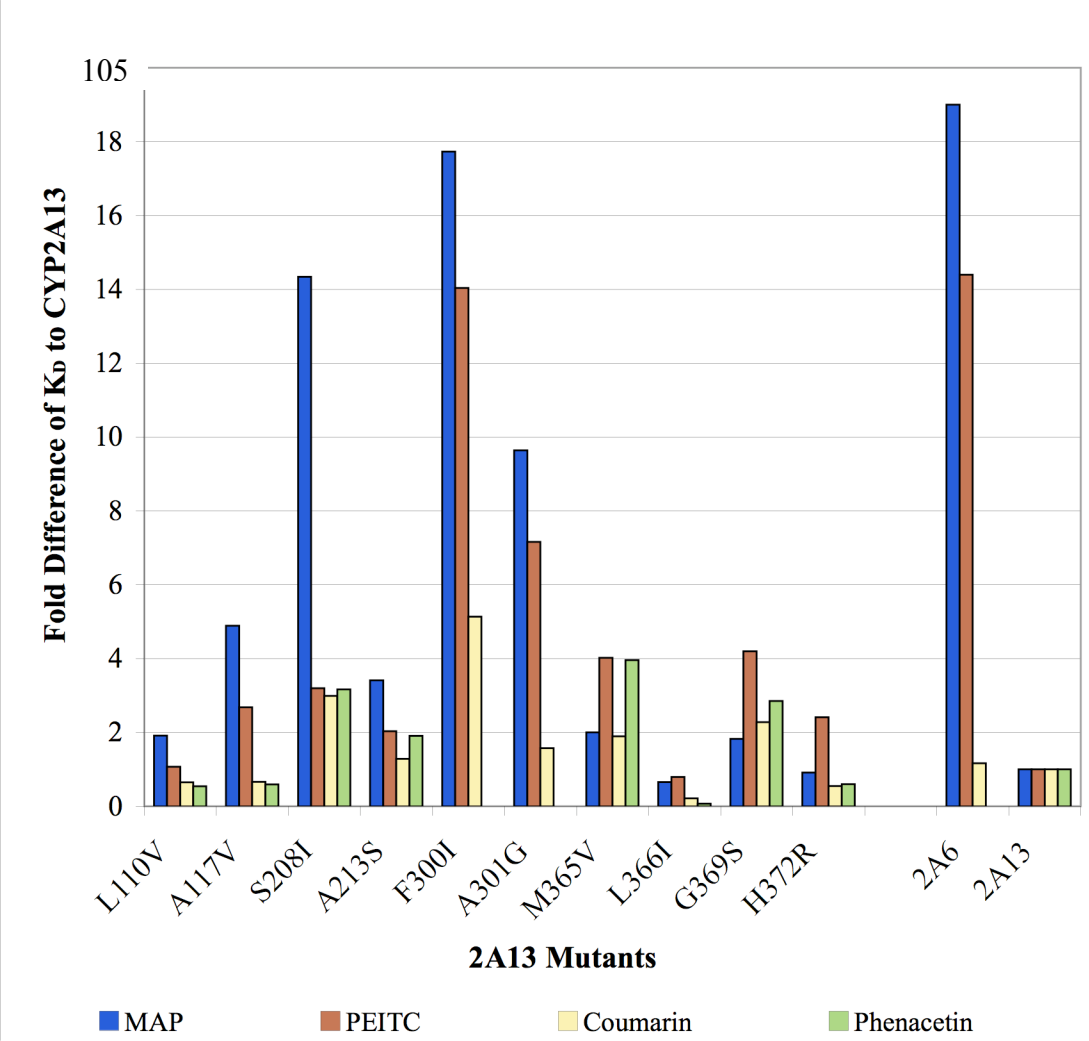


Figure 3.3. Fold-difference (K_D of Mutant/K_D of CYP2A13) of the ten CYP2A13 mutants compared to wild type CYP2A13 and CYP2A6.

Binding of PEITC to wild type CYP2A13 was characterized by a K_D of 0.43 μM, approximately 14-fold higher affinity than CYP2A6 (K_D 6.2 μM) (Figure 3.3). The single substitution of F300I in the CYP2A13 protein resulted in a 14-fold loss in PEITC affinity, such that its K_D (6.0 μM) is essentially that of CYP2A6 wild type. The mutation A301G, located directly adjacent to F300I, demonstrated a 7-fold increase in K_D to 3.1 μM. The mutants A117V, S208I, M365V, and G369S had K_D values that were increased 3 to 4-fold compared to the wild type CYP2A13. The remaining single mutants yielded smaller changes in the PEITC binding affinity. The substitution 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 ~2.9 μM for both CYP2A13 and CYP2A6. Again, the F300I substitution had the most dramatic effect on K_D, with a 5-fold increase to 14 μM. Two other substitutions, G369S and S208I, increased K_D values by 2- and 3-fold to 6.0 μM and 7.9 μM, respectively. The substitution H372R increased coumarin binding affinity by nearly 50% to a K_D of 1.5 μM, while L366I increased coumarin binding affinity by 5-fold to 0.57 μM.

Discussion

The goal of our present study was to understand the differences in structure-function relationships between the human cytochrome CYP2A6 and CYP2A13 enzymes and their ligands and to identify individual amino acid residues that mediate those binding affinities. Initially a series of four 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 (6-8). 2'-methoxyacetophenone (MAP) (9,10) is also a CYP2A substrate, while phenethyl isothiocyanate (PEITC) is an inhibitor (11). Phenacetin was selected as a ligand because it is a differential substrate of CYP2A13 but not CYP2A6 (12). Examination of these structurally diverse cytochrome P450 2A ligands revealed a very large range in the affinities of CYP2A13 vs. CYP2A6. At the extremes, coumarin bound equally well to both proteins, while phenacetin bound to CYP2A13 but did not cause a spectral shift in CYP2A6 (Table 3.1).

A structure of CYP2A6 has been determined with the strictly planar coumarin compound in the active site (13). This structure shows that the coumarin ketone oxygen hydrogen bonds with N297 within the largely planar active site and positions C7 for hydroxylation. Indeed, 7-hydroxycoumarin in the only metabolite detected for CYP2A6. Both the hydrogen bonding residue and the planarity of the active site are conserved in the 2A13 structure (13). These structural similarities, the nearly identical K_D values, and the formation of the 7-hydroxycoumarin metabolite by both

proteins suggest that coumarin would bind similarly in CYP2A13. However, Weymarn and Murphy observed that while 7-hydroxycoumarin is the major coumarin metabolite (43%) for CYP2A13, a significant amount of the 3,4 epoxide is also observed (30%), as well as smaller amounts of 6-hydroxy- and 8-hydroxycoumarin (14). In side-by-side studies, none of these products were observed for CYP2A6. This suggests that, despite the nearly identical K_D values, that in CYP2A13, coumarin 1) binds by hydrogen-bonding with N297, but is less constrained in its approach to the activated oxygen on the heme so that the C6 and C8 on either side of C7 can also be hydroxylated and 2) also binds in a second, inverted orientation to expose the C3/C4 carbons for formation of the epoxide.

MAP binding showed a large difference between CYP2A13 (K_D 0.85 μM) and CYP2A6 (K_D 74 μM). Enzymatically MAP is *O*-demethylated by both enzymes to produce 2'-hydroxyacetophenone and formaldehyde. 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 (15). Thus, the substantial differences in MAP affinity were not expected based on the single productive orientation in the active site indicated by a single product and the preliminary metabolic rate information available.

The inhibitor PEITC had a 14-fold greater affinity for CYP2A13 compared to CYP2A6. PEITC can inhibit without enzymatic activation, but appears to also inactivate via covalent modification of the CYP2A6 and CYP2A13 proteins (11).

The K_I values for inhibition of CYP2A13 with PEITC is 10-fold lower than for CYP2A6.

Phenacetin was selected as a ligand because Fukami *et al.* determined CYP2A13 metabolizes phenacetin to acetaminophen while CYP2A6 had very low activity (12). CYP2A13 metabolized phenacetin O-deethylation with higher activity than CYP1A2, the cytochrome P450 commonly attributed to phenacetin metabolism. CYP2A13 had a binding affinity for phenacetin of 34 μM, while the binding affinity of CYP2A6 could not be determined as no discernable spectra shift was evident at the highest concentration tested (250 μM phenacetin).

To determine the amino acids responsible for such disparate ligand K_D values, active site residues that differ between CYP2A13 and CYP2A6 were systematically examined for their contributions to ligand binding. Ten positions were selected for examination in this study because they either line the CYP2A13 active site or are immediately adjacent to residues that do (Chapter 2) (16). The resulting engineered CYP2A13 proteins were characterized using ligand binding titrations with phenacetin, MAP, PEITC, and coumarin.

A single mutation, phenylalanine to isoleucine, at the first shell active site position (300), yielded 5- to >17-fold decreases in the binding affinity for the four cytochrome P450 2A13 ligands examined. The loss in binding affinity for phenacetin was so substantial that binding affinity could not be determined (Table 3.2). Despite the ligand structural diversity, phenacetin, MAP, PEITC, and coumarin binding were all significantly perturbed by this single substitution. This indicates that this

phenylalanine plays a central role in the active site. Comparison of the CYP2A6 and CYP2A13 structures reveals that I300 and F300 overlap at the alpha carbon atom, but that the phenyl ring is rotated farther away from the bound ligand than the isoleucine side chain (Figure 1.4) (13,16). However, the two proteins were crystallized with different ligands, which could affect the side chain positioning. No data is presently available regarding the effects of this mutation on metabolism of the substrates examined here, although Guengerich and coworkers (17) found that both the residues at positions 300 and 297 were critical for CYP2A6 metabolism of indole. Overall, it is clear that this residue is critical for the higher affinity of CYP2A13 with all four ligands in this study.

Substitution of the adjacent CYP2A13 residue, A301, to glycine removes the side chain and caused spectral sifts that were so diminished with phenacetin that a K_D could not be determined. The CYP2A13 A301G mutation also had a 7- to 9-fold increase in the K_D for MAP and PEITC, respectively, but does not significantly alter the K_D for coumarin (Figure 3.3). This suggests that coumarin interacts differently with the residue at position 301 than does phenacetin, MAP, or PEITC. Since MAP and PEITC bind more tightly to CYP2A13, interactions with this residue may contribute to ligand selectivity. Again, there are no data available to determine the effects of this substitution on substrate metabolism.

The S208I substitution has a 14-fold increase in the K_D for MAP, but much smaller ~3-fold effects on phenacetin, PEITC, and coumarin binding (Figure 3.3). The more moderate effects on coumarin binding are consistent with the results of He

et al., who showed that this mutation had little effect on the kinetics of coumarin 7-hydroxylation by CYP2A13 (8). Of note, this mutation is 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 (18). The CYP2A13 wild type structure shows that S208 forms part of the C-terminus of the F helix (Figure 1.4). Several bacterial and mammalian cytochrome P450 structures have indicated that the N-terminus of the F helix can be fairly mobile and may form part of the small molecule entrance/exit channel (19-21). Thus the position of this residue could change depending on both the position and the identity of the ligand.

Substitutions of A117V, M365V, and G369S had more moderate effects on binding affinity for one or two substrates (Figure 3.3). A117V yielded a 4.8-fold increase in MAP K_D, but much smaller changes for phenacetin, PEITC, or coumarin. Although this mutation was not reported to significantly alter the K_m for coumarin metabolism, mutation did substantially alter the k_{cat}, such that this single mutation interconverted the CYP2A13 k_{cat} for coumarin (0.31 $\mu M^{-1}min^{-1}$) and CYP2A6 (2.6 $\mu M^{-1}min^{-1}$) (8). Similarly, CYP2A13 A117V altered the catalytic efficiency of NNK metabolism by alterations in the V_{max}, but not the K_m (18). These results suggest that for some ligands like MAP, the residue at position 117 has a moderate influence on binding affinity. While, for other ligands like coumarin residue 117 is more critical in the difference in metabolic rates than the difference in binding affinities for the two enzymes. In a similar manner, substitution of M365V and G369S altered the affinities for PEITC and phenacetin, but had much smaller effects on MAP and

coumarin binding. This is in agreement with the CYP2A13 structure that shows A117, M365 and G369 located more peripherally to the indole ligand in the CYP2A13 binding cavity while F300 and A301 are both centrally located and likely to cause a change in binding regardless of the ligand (Figure 1.4) (16).

The L366I mutation was the only change to increase binding affinity for ligands (Figure 3.3). The K_D for phenacetin was reduced 93% from 34 μM to 2.3 μM and coumarin was reduced 80% from 2.7 μM to 0.57 μM. The reductions for PEITC and MAP were smaller. Thus, the substitution of the leucine found in CYP2A13 for the valine in CYP2A6 must result in an active site that is able to better accommodate a variety of ligands.

Finally, the H372R substitution is notable because, although this mutation is reported to have significant effects on the k_{cat} for coumarin and thus the catalytic efficiency (22), no significant changes were observed in K_D for coumarin, MAP, PEITC, and phenacetin. A separate report indicated that this mutation caused very large alterations in the K_m for NNK metabolism to the keto aldehyde product and eliminated the keto alcohol metabolite (18). Therefore, this residue may be relevant to NNK binding, but not the ligands examined in this study. It is also possible that the H372R mutation may not have a direct effect on ligand binding, but is capable of altering metabolism through a different means.

The goal of this study was to gain insight into the differences in structure-function relationships between the human cytochrome CYP2A6 and CYP2A13 enzymes and their ligands. Binding titrations of chemically diverse ligands to

CYP2A13 and CYP2A6 were used in conjunction with site-directed mutagenesis to identify key active site amino acids. Overall, the experimental spectral titrations identify F300I as the most important amino acid difference between CYP2A13 and CYP2A6. F300I altered binding of coumarin, MAP, PEITC, and phenacetin. A301G altered MAP, PEITC, and phenacetin binding, while S208I altered MAP and phenacetin binding. Mutations at positions 117, 365, 366, and 369 had smaller effects on the binding affinities of one or two substrates.

Conclusion

In conclusion, residues at positions 300, 301, and 208 appear to be the most critical in the determining the selectivity of ligand binding, while 117, 365, and 369 have smaller effects. Mutation of the residue 366 increased the binding affinity for all ligands. The other positions examined, 110, 213, and 372 yielded K_D values that were relatively little changed by comparison with wild type CYP2A13. These ligand binding assays provide a greater understanding of CYP2A active site structure activity relationships.

References

1. Jefcoate, C. R. (1978) *Methods Enzymol.* **52**, 258-279
2. Schenkma.Jb, Remmer, H., and Estabroo.Rw. (1967) *Mol. Pharmacol.* **3**(2), 113-&
3. (2005) *Cytochrome P450: Structure, Mechanism, and Biochemistry*, Third Edition Ed., Kluwer Academic/ Plenum Publishers, New York
4. Locuson, C. W., Hutzler, J. M., and Tracy, T. S. (2007) *Drug Metab. Dispos.* **35**(4), 614-622
5. Schenkman, J. B., Cinti, D. L., Orrenius, S., Moldeus, P., and Kraschnitz, R. (1972) *Biochemistry (Mosc).* **11**(23), 4243-4251
6. Fernandezsalguero, P., Hoffman, S. M. G., Cholerton, S., Mohrenweiser, H., Raunio, H., Rautio, A., Pelkonen, O., Huang, J. D., Evans, W. E., Idle, J. R., and Gonzalez, F. J. (1995) *Am. J. Hum. Genet.* **57**(3), 651-660
7. Kim, D., Wu, Z. L., and Guengerich, F. P. (2005) *J. Biol. Chem.* **280**(48), 40319-40327
8. He, X. Y., Shen, J., Hu, W. Y., Ding, X., Lu, A. Y., and Hong, J. Y. (2004) *Arch. Biochem. Biophys.* **427**(2), 143-153
9. Su, T., Bao, Z., Zhang, Q. Y., Smith, T. J., Hong, J. Y., and Ding, X. (2000) *Cancer Res.* **60**(18), 5074-5079
10. von Weymarn, L. B., Zhang, Q. Y., Ding, X., and Hollenberg, P. F. (2005) *Carcinogenesis* **26**(3), 621-629
11. von Weymarn, L. B., Chun, J. A., and Hollenberg, P. F. (2006) *Carcinogenesis* **27**(4), 782-790
12. Fukami, T., Nakajima, M., Sakai, H., Katoh, M., and Yokoi, T. (2006) *Drug Metab. Dispos.*
13. Yano, J. K., Hsu, M. H., Griffin, K. J., Stout, C. D., and Johnson, E. F. (2005) *Nat Struct Mol Biol* **12**(9), 822-823
14. von Weymarn, L. B., and Murphy, S. E. (2003) *Xenobiotica* **33**(1), 73-81
15. Su, T., Bao, Z. P., Zhang, Q. Y., Smith, T. J., Hong, J. Y., and Ding, X. X. (2000) *Cancer Res.* **60**(18), 5074-5079
16. Smith, B. D., Sanders, J. L., Porubsky, P. R., Lushington, G. H., Stout, C. D., and Scott, E. E. (2007) *J. Biol. Chem.*
17. Wu, Z. L., Podust, L. M., and Guengerich, F. P. (2005) *J. Biol. Chem.* **280**(49), 41090-41100
18. He, X. Y., Shen, J., Ding, X., Lu, A. Y., and Hong, J. Y. (2004) *Drug Metab. Dispos.* **32**(12), 1516-1521
19. Scott, E. E., He, Y. A., Wester, M. R., White, M. A., Chin, C. C., Halpert, J. R., Johnson, E. F., and Stout, C. D. (2003) *Proc. Natl. Acad. Sci. USA* **100**(23), 13196-13201
20. Scott, E. E., White, M. A., He, Y. A., Johnson, E. F., Stout, C. D., and Halpert, J. R. (2004) *J. Biol. Chem.* **279**(26), 27294-27301
21. Li, H., and Poulos, T. L. (1997) *Nat. Struct. Biol.* **4**(2), 140-146

22. He, X. Y., Shen, J., Hu, W. Y., Ding, X., Lu, A. Y., and Hong, J. Y. (2004) *Arch. Biochem. Biophys.* **427**(2), 143-153

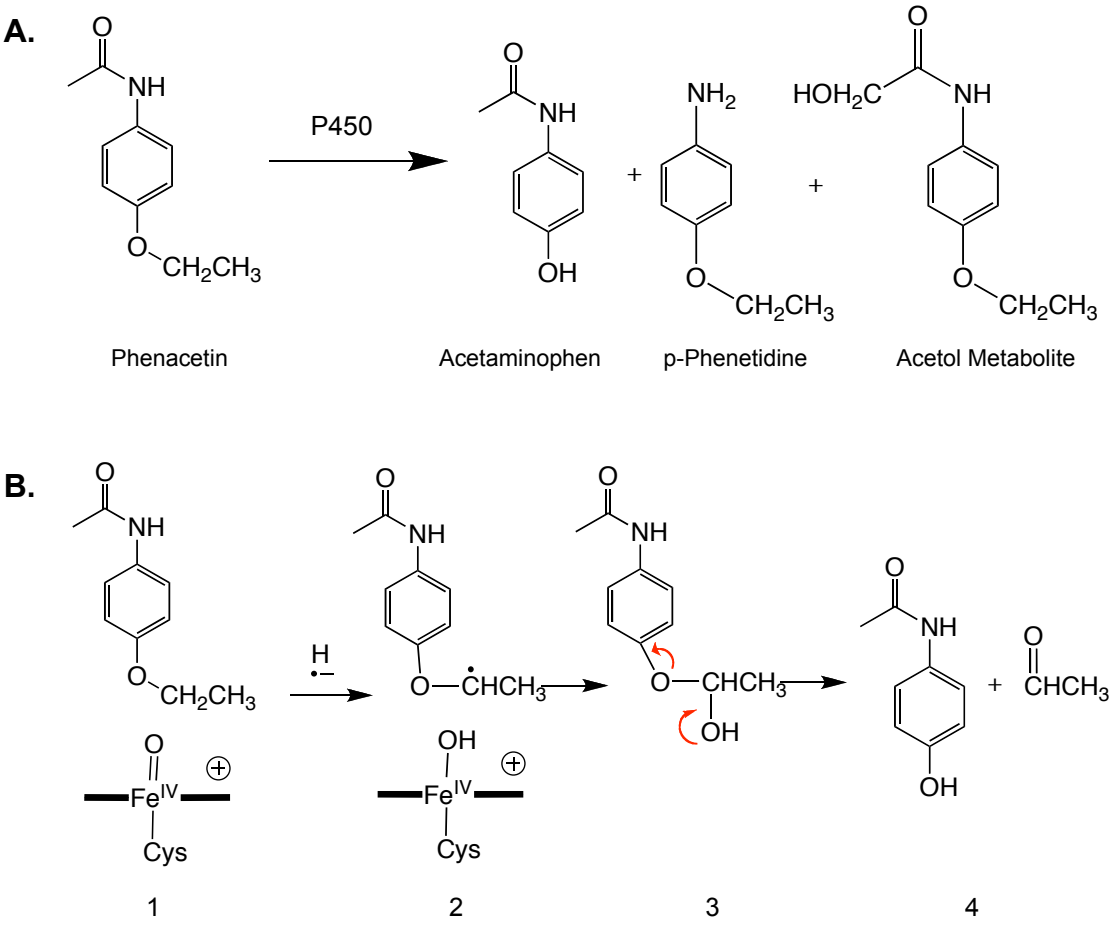
Chapter 4

Metabolism Assay with Phenacetin as a Substrate

Introduction

In the previous chapter, ligand binding studies identified several CYP2A13 residues the mutation of which resulted in significant changes in ligand binding affinity for MAP, PEITC, coumarin, and phenacetin. Since binding affinity does not always correlate directly to altered substrate metabolism, the next phase in the project was to determine whether the same amino acids also alter substrate metabolism. Phenacetin was selected as the substrate used to probe for residues contributing to the metabolic differences between CYP2A13 and CYP2A6 because we observed the largest difference in CYP2A ligand binding with phenacetin.

The primary metabolite of phenacetin is acetaminophen via *O*-deethylation, the commonly used analgesic and antipyretic (Scheme 4.1, panel A and B). Phenacetin was used as an analgesic and antipyretic, both alone and in combination with aspirin and caffeine, between 1887 and 1983 (1). After 1983, phenacetin was removed from market because of implications in human nephropathy (2). Phenacetin has also been implicated in the development of tumors in the urinary tract and nasal cavity of rats (3,4). Due to its connection with cancer and nephropathy, the metabolism of phenacetin has been studied extensively. Phenacetin is also *N*-deacetylated to the toxic metabolite *p*-phenetidine and hydroxylated to form an acetol metabolite (Scheme 4.1, panel A) (5-7). Enzymes from the cytochrome P450 1A family have previously



Scheme 4.1. A. Phenacetin is metabolized by several cytochrome P450s. The major product is acetaminophen via *O*-deethylation. *p*-Phenetidine and an acetol secondary metabolite have also been identified as P450 metabolites in previous. B. The phenacetin *O*-deethylation reaction as catalyzed by cytochrome P450 enzymes. 1) A hydrogen atom is abstracted from the subterminal alkyl carbon by activated oxygen of the cytochrome P450. 2) The hydroxyl group on the heme iron is then transferred to the sub-terminal carbon that had the hydrogen abstracted. 3) Rearrangement occurs leading to the formation of acetaminophen and acetyl aldehyde.

been identified as the major enzymes responsible for phenacetin *O*-deethylation to acetaminophen (5,8,9). Other cytochrome P450s such as CYP2C9, CYP2A19, and CYP2E1 have also been reported to metabolize phenacetin, although with much higher K_m values than CYP1A2 (9-11).

Fukami and coworkers recently established that CYP2A13 catalyzes phenacetin *O*-deethylation with an even higher k_{cat} (2.9 ± 0.3 pmol/min/pmol) and a lower K_m (21.8 μM) than CYP1A2 (1.9 ± 0.1 pmol/min/pmol and 29.2 μM, respectively) (12). In contrast, CYP2A6 has been reported to have a very high K_m (4000 μM) for phenacetin *O*-deethylation (9-11). Our own spectral binding analysis determined that CYP2A13 binds phenacetin with a K_D of ~34 μM, while CYP2A6 does not cause a spectral shift. The efficient metabolism of phenacetin by CYP2A13 and poor metabolism by CYP2A6 made phenacetin an ideal metabolite to identify amino acids responsible for differential substrate recognition in the CYP2A enzymes.

We hypothesized that the amino acids identified with roles in phenacetin binding would also be involved in differential phenacetin metabolism. The same ten mutants used in the spectral ligand binding study were tested for their ability to catalyze phenacetin *O*-deethylation compared to both wild type CYP2A13 and CYP2A6 enzymes. Since it is often easier to destroy enzyme activity than to create enzyme functionality where none existed previously, we verified the importance of these amino acids to phenacetin metabolism by introducing the CYP2A13 amino acids at these positions into CYP2A6. Because CYP2A13's ability to metabolize phenacetin is likely the cumulative result of a number of substitutions, we also

constructed CYP2A6 mutants using combinations of CYP2A13 mutants that reduced the phenacetin *O*-deethylation metabolism the most.

Methods

All proteins used in this chapter were expressed and purified according to the protocol found in chapter 2. The procedure for the phenacetin metabolism assay is also found in chapter 2. When full length CYP450 was assayed, the method was slightly modified. In order to stabilize the full length CYP450 enzyme the lipid dilauroylphosphatidylcholine (DLPC) was added to the reconstituted protein system in addition to NADPH cytochrome P450 reductase, cytochrome b_5, and P450 protein.

We determined the lower limit of detection for the phenacetin metabolism assay to be 0.026 pmol/min/pmol. The lower limit of detection was determined by the mean of 12 sample blanks (no acetaminophen present in sample) multipled by two times the standard deviation of those samples. Kinetic parameters were determined with a range of (0.5 – 250 μM phenacetin) for all proteins with a phenacetin *O*-deethylation activity greater than 0.7 pmol/min/pmol in the 250 μM activity studies.

Results

Comparison of CYP2A13 single mutations at 250 μM Phenacetin: The phenacetin *O*-deethylation activity was determined at 250 μM for all ten CYP2A13 mutants and for CYP2A13 and CYP2A6. The activity of CYP2A13 was ~3.4 pmol/min/pmol (Table 4.1). In contrast, the activity of CYP2A6 was barely above our lower limit of detection at 0.08 pmol/min/pmol.

Two mutants, A117V and L366I, resulted in five-fold and three-fold increases in activity, respectively. However, the other eight mutations resulted in decreases in phenacetin *O*-deethylation (Figure 4.1). The single mutations L110V, S208I, F300I, A301G, M365V, and G369S all resulted in phenacetin metabolism that is less than 15% of the wild type CYP2A13 activity. The residues at position 300, 301, 208, and 369 had the largest reductions in activity. The A213S and H372R mutations had a less pronounced decrease in activity with rates around ~1 pmol/min/pmol.

Table 4.1. Phenacetin *O*-deethylation rates and comparisons for CYP2A13, and CYP2A13 mutants. All assays used 250 μM phenacetin.

Protein	Rate of product formation (pmol/min/pmol)[a]	$\frac{Activity_{Mutant}}{Activity_{CYP2A13}}$
CYP2A13	3.4 ± 0.6	1.0
L110V	0.40 ± 0.07[b]	0.1
A117V	18 ± 2 [b]	5
S208I	0.35 ± 0.06 [b]	0.1
A213S	1.5 ± 0.5 [b]	0.4
F300I	0.07 ±0.01 [b]	0.02
A301G	0.18 ± 0.03 [b]	0.05
M365V	0.5 ± 0.1 [b]	0.1
L366I	11 ± 1 [b]	3
G369S	0.4 ± 0.1 [b]	0.1
H372R	1.06 ± 0.04 [b]	0.3

[a]Data represent the average of two independent experiments each conducted in triplicate.
[b]Statistically significant differences were calculated using a t test for the rate of product formation with wild type CYP2A13 versus mutant rate of product formation and are indicated with asterisks (*,$p < 0.0001$).

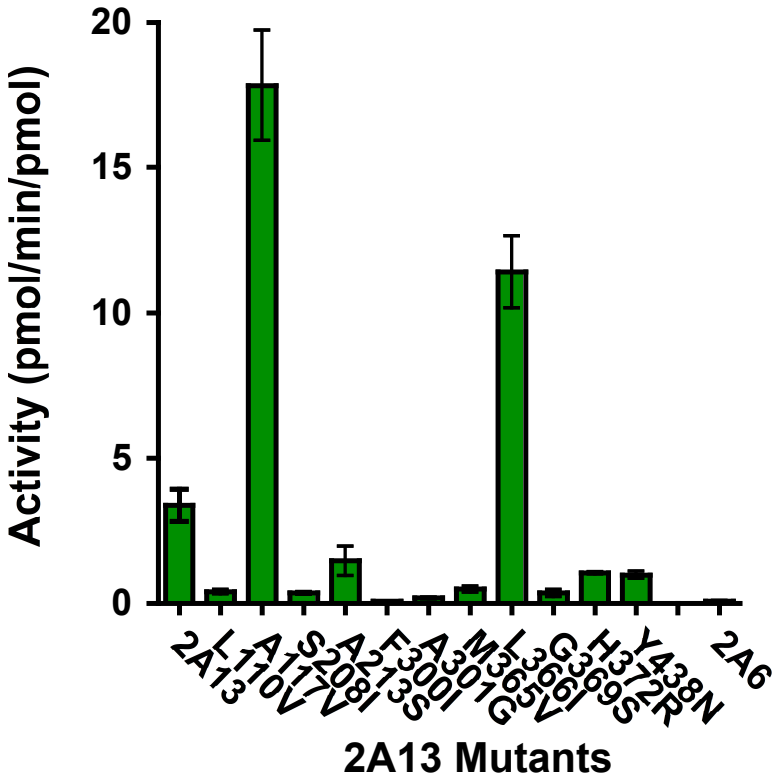


Figure 4.1. Effects of single mutations on CYP2A13 phenacetin *O*-deethylation rates.

Comparison of CYP2A6 mutations at 250 μM Phenacetin: The majority of CYP2A6 single mutants had very little phenacetin *O*-deethylation activity with the exception of V110L, I300F, and S369G, which had a slightly increased phenacetin *O*-deethylation activity of (0.2-0.3 pmol/min/pmol) (Table 4.2). This suggested that phenacetin *O*-deethylation activity in the CYP2A family was the aggregate result of several amino acid substitutions between CYP2A13 and CYP2A6.

Several CYP2A6 multiple mutants were constructed to explore this hypothesis. The double mutant CYP2A6 V365M/I366L was generated because of the unique placement of these two substitutions and the opposite effect observed at these positions with the CYP2A13 single mutants. However, this double mutant had activity similar to wild type CYP2A6. The double mutant CYP2A6 I300F/G301A showed a moderate increase in activity to 0.63 pmol/min/pmol (Table 4.2). The addition of I208S or S369G substitutions to the I300F/G301A mutant led to further increases in activity: 0.8 pmol/min/pmol and 1.0 pmol/min/pmol, respectively (Table 4.2). The inclusion of all four mutations (2A6dH I208S/I300F/G301A/S369G) led to a mutant with activity of ~3 pmol/min/pmol, similar to the activity of CYP2A13. The additive effects of these mutations can be easily visualized in Figure 4.2.

Table 4.2. Phenacetin O-deethylation rates by CYP2A6 and CYP2A6 mutants. All assays used 250 μM phenacetin.

Protein	Rate of product formation (pmol/min/pmol)[a]	$\frac{Activity}{Activity_{CYP2A6}}$
2A6	0.08 ± 0.04	1.00
V110L	0.3 ± 0.1[d]	3.9
V117A	< det. lim.[b]	N.A.[c]
I208S	< det. lim.[b]	N.A.[c]
S213A	0.08 ± 0.00	1.0
I300F	0.21 ± 0.04[d]	2.7
G301A	0.07 ± 0.01	0.8
V365M	0.13 ± 0.03[d]	1.7
I366L	< det. lim.[b]	N.A.[c]
S369G	0.21 ± 0.05[d]	2.6
R372H	0.14 ± 0.09	1.7
V365M/I366L	0.12 ± 0.04	1.5
I300F/G301A	0.63 ± 0.09[e]	7.9
I208S/I300F/G301A	0.8 ± 0.1[e]	10
I300F/G301A/S369G	1.02 ± 0.09[e]	13
I208S/I300F/G301A/S369G	3.0 ± 0.3[e]	38

[a]Data represent the average of two independent experiments each conducted in triplicate.
[b]Value below the detection limit of our system (0.026 pmol/min/pmol).
[c]Not applicable.
[d]Statistically significant differences were calculated using a t test for the rate of product formation with wild type CYP2A6 versus mutant rate of product formation ($p < 0.05$).
[e]Statistically significant differences were calculated by t test for the rate of product formation with wild type CYP2A6 versus mutant rate of product formation ($p < 0.0001$).

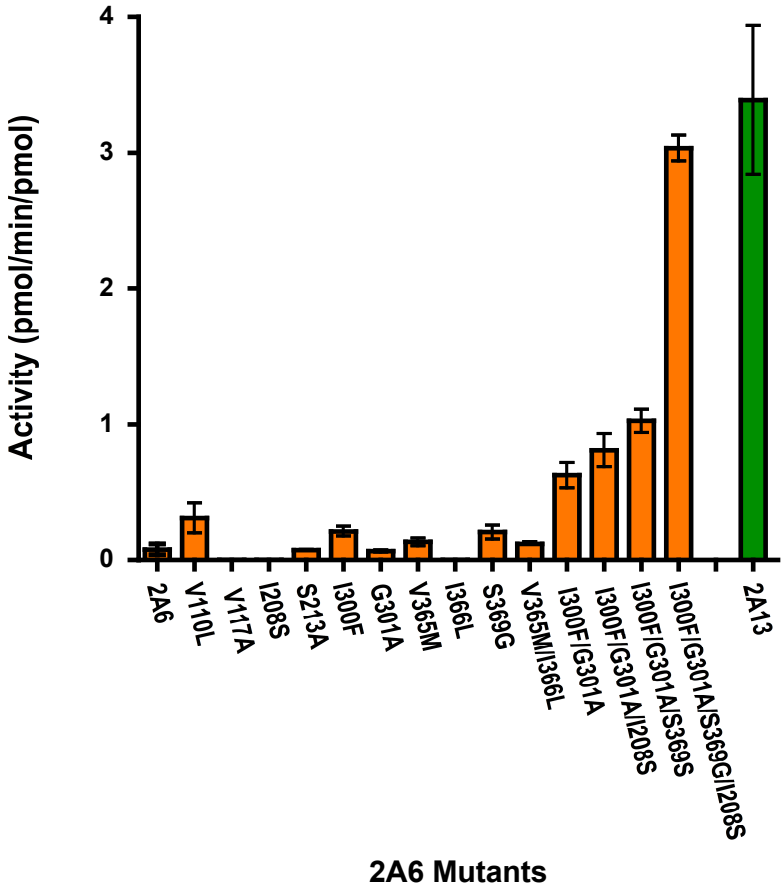


Figure 4.2. Effects of CYP2A6 mutations on phenacetin *O*-deethylation rates (250 µM phenacetin).

Phenacetin O-deethylation Kinetic Parameters: The kinetic parameters between CYP2A13 mutant enzymes varied considerably (Table 4.3). As expected, CYP2A13 A117V and L366I mutations that caused increased activity at 250 μM phenacetin also had increased k_{cat} values (Figure 4.3, panel A). However, the K_m for A117V and L366I were also both increased compared to wild type CYP2A13 (Table 4.3). This resulted in the mutant A117V having a catalytic efficiency 1.5-fold greater than CYP2A13 while the L366I mutation had a catalytic efficiency identical to CYP2A13 (Table 4.3). Three other CYP2A13 mutants, A213S, M365V, and H372R had k_{cat} values below that of wild type CYP2A13 (Figure 4.3, panel B). One of these mutations, H372R, had a K_m value significantly decreased to ~5 μM; however, a corresponding decrease in k_{cat} value resulted in catalytic efficiency identical to CYP2A13 (Table 4.3).

Each of the CYP2A6 single mutants had activities at 250 μM phenacetin that were too low to determine reliable kinetic parameters. However, several of the CYP2A6 multiple mutants had sufficient activity to determine kinetic parameters (Table 4.3). The CYP2A6 triple mutants I208S/I300F/G301 and I300F/G301A/S369G both had k_{cat} values of ~0.9 pmol/min/pmol, significantly increased from CYP2A6, and equivalent to several of the CYP2A13 single mutants (Figure 4.3, panel C). In contrast, the K_m value of the triple mutant CYP2A6 I300F/G301A/S369G was 4.9 μM, or half that of CYP2A13, while the triple mutant CYP2A6 I208S/I300F/G301A had a K_m of 39 μM, a four-fold increase from CYP2A13.

Table 4.3. Kinetic Parameters for CYP1A2, CYP2A13, and mutants of CYP2A13 and CYP2A6[a].

Enzyme	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m ($\mu M^{-1}min^{-1}$)	$\frac{(k_{cat}/K_m)_{Mutant}}{(k_{cat}/K_m)_{2A13}}$
1A2[b]	27 ± 2[d]	9.7 ± 0.3 [d]	0.4	1
2A13	11 ± 2	3.8 ± 0.1[d]	0.4	1
A117V	48 ± 6[d]	29 ± 1 [d]	0.6	1.5
A213S	21 ± 3[d]	2.14 ± 0.06 [d]	0.1	0.3
M365V	20 ± 5[d]	0.93 ± 0.05 [d]	0.05	0.1
L366I	36 ± 5[d]	13.7 ± 0.6 [d]	0.4	1
H372R	5 ± 1[d]	1.91 ± 0.06 [d]	0.4	1
2A6	N.D.[c]	N.D.[c]	N.D.[c]	N.D.[c]
I208S/I300F/G301A	39 ± 7[d]	0.98 ± 0.06 [d]	0.03	0.08
I300F/G301A/S369G	4.9 ± 0.5[d]	0.89 ± 0.01 [d]	0.2	0.5
I208S/I300F/ G301A/G369S	10.3 ± 0.8	2.88 ± 0.04 [d]	0.3	0.8

[a]Data represent the average of two independent experiments each with all concentrations conducted in triplicate.
[b]This is the full length enzyme.
[c]N.D., Reliable kinetic values could not be determined for enzymes with activities <0.7 pmol/min/mol at 250 μM phenacetin (see Tables 2 and 3)
[d]Statistically significant differences were calculated using a t test for the K_m and k_{cat} of wild type CYP2A13 versus the mutant kinetic parameters ($p < 0.001$).

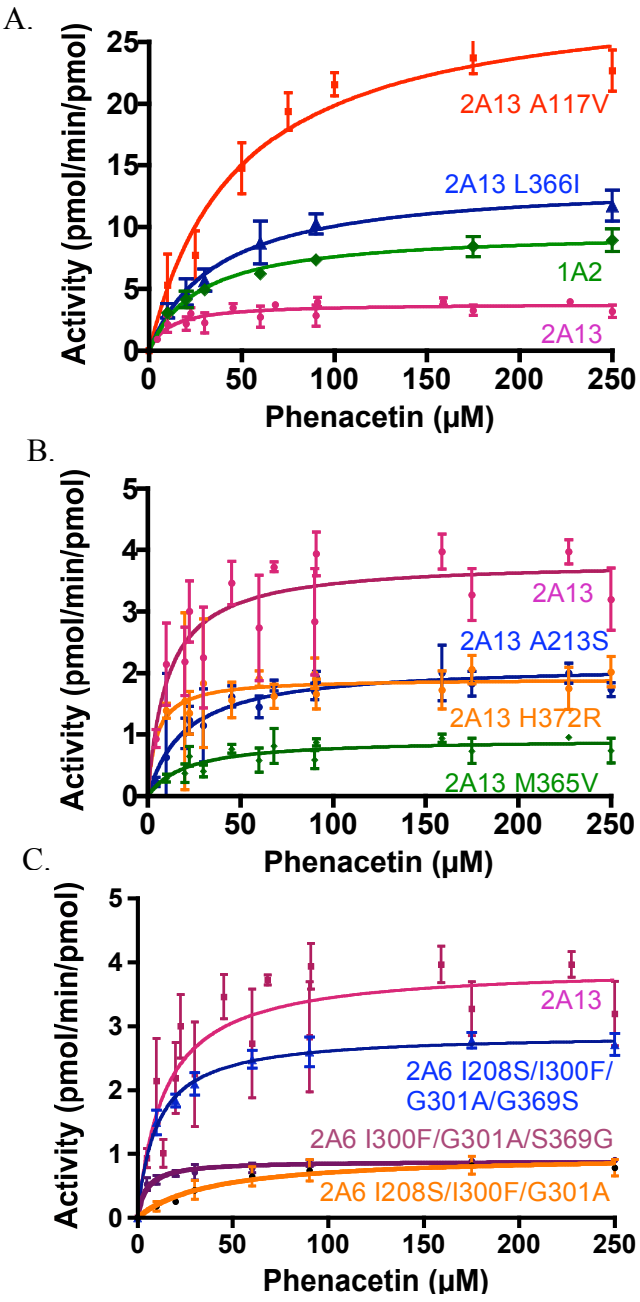


Figure 4.3. Phenacetin metabolism by CYP1A2, CYP2A13, mutant proteins. Curves represent a global fit of two independent experiments with each point performed in triplicate.

As a result, the catalytic efficiency of the triple mutant containing I208S was much lower than that of the triple mutant containing S369G. The quadruple mutant CYP2A6 I208S/I300F/G301A/S369G had a catalytic efficiency very similar to CYP2A13, with an almost identical K_m and a slightly reduced k_{cat} (Figure 4.3, panel C).

Since phenacetin has been used as a marker substrate for CYP1A2, the full-length version of CYP1A2 (generous gift from Dr. Guengerich) was characterized in addition to the CYP2A13 and CYP2A6 mutant proteins. Both CYP1A2 and CYP2A13 had the same catalytic efficiency of ~0.4 $\mu M^{-1}min^{-1}$ although their catalytic parameters varied considerably (Table 4.3). CYP1A2 had a k_{cat} of 9.7, approximately three times greater than the k_{cat} of CYP2A13. However, CYP1A2 also had a K_m value that was three times greater than CYP2A13 (Figure 4.3, panel A).

Discussion

The metabolism assay for phenacetin *O*-deethylation verified that amino acids at positions 208, 300, 301, and 369 are important for phenacetin metabolism. Mutation of the 300 or 301 amino acids alone nearly eliminated phenacetin metabolism in CYP2A13, while mutations at positions 110, 208, and 369 all significantly reduced phenacetin metabolism activity such that kinetic parameters could not be determined for any of these mutants. Other mutant enzymes either reduced phenacetin metabolism to a lesser degree, as is the case for mutations at 213, 365, and 374, or increased the rate of phenacetin metabolism (117 and 366).

Substitution of single CYP2A13 residues into CYP2A6 had very little effect on phenacetin metabolism activity. Therefore, we concluded that the ability of CYP2A13 to metabolize phenacetin was likely the result of several combined differences from the CYP2A6 active site. In order to explore this possibility, a series of multiple mutations were generated that involved mutating residues 208, 300, 301, and 369 (those with the lowest activity in CYP2A13 at 250 μM phenacetin). Incorporation of these residues resulted in additive increases in phenacetin activity. The two triple mutant proteins (CYP2A6 I208S/I300F/G301A and CYP2A6 I300F/G301A/S369G) both had significant phenacetin activity of ~1 pmol/min/pmol at 250 μM phenacetin. Although both mutant enzymes had reduced catalytic efficiency compared to CYP2A13, the triple mutant 2A6 I300F/G301A/S369G had a K_m value half that of the wild type CYP2A13. The quadruple mutant containing mutations at residues 208, 300, 301, and 369 resulted in an enzyme with a K_m nearly identical to that of CYP2A13 and a k_{cat} ~80% of CYP2A13. The fact that substitution of four CYP2A13 residues into CYP2A6 creates an enzyme with phenacetin *O*-deethylation activity and kinetic parameters very similar to CYP2A13 indicates that these four residues are largely responsible for the ability to metabolize phenacetin in human CYP2A enzymes.

Conclusion

We hypothesized that a few amino acids may be responsible for CYP2A substrate selectivity. Consistent with this hypothesis, the phenacetin metabolism data indicates that just four amino acids, 208, 300, 301, and 369, seem to mediate phenacetin metabolism in the CYP2A enzyme family. Single amino acid substitutions at these positions in CYP2A13 resulted in enzymes with little to no phenacetin metabolism activity. More than one of these substitutions are necessary for phenacetin metabolism as illustrated by the observations that none of the CYP2A6 single mutants had significant increases in phenacetin metabolism compared to the wild type CYP2A6 enzyme. Increases in phenacetin metabolism were observed when multiple mutations are made at these positions. Mutations at amino acid 300 and 301 alone resulted in a CYP2A6 enzyme with phenacetin metabolism activity ~8-fold greater than the wild type CYP2A6 (Table 4.2). The quadruple mutant CYP2A6 I208S/I300F/G301A/S369G had activity ~38-fold greater than the CYP2A6 enzyme. Further more, the quadruple mutant had a K_m nearly identical to that of CYP2A13 and a catalytic efficiency ~80% of CYP2A13. These results indicate that the substitution of just four amino acids is sufficient to change the active site environment of CYP2A6 enough to accommodate phenacetin as a ligand. A logical next step in this analysis was molecular modeling studies with phenacetin and the known CYP2A13 crystal structure.

References

1. (2005) *NTP Report on Carcinogens*, Eleventh Edition Ed., U.S. Department of Health and Human Services
2. Flower, R. J., S. Moncada, and J.R. Vane. (1985) Analgesic-antipyretics and anti-inflammatory agents. In: A.G. Gilman, L. S. G., T.W. Rall, and G. Murad (ed). *Goodman and Gilman's The Pharmacological Basis of Therapeutics*, 7 Ed., Macmillan Publishing Co., New York
3. Fischbach, T., and Lenk, W. (1985) *Xenobiotica* **15**(11), 915-927
4. Fischbach, T., and Lenk, W. (1985) *Xenobiotica* **15**(2), 149-164
5. Yun, C. H., Miller, G. P., and Guengerich, F. P. (2000) *Biochemistry (Mosc).* **39**(37), 11319-11329
6. Hinson, J. A. (1983) *Environ. Health Perspect.* **49**(0091-6765 (Print)), 71-79
7. Hinson Ja Fau - Mitchell, J. R., and Mitchell, J. R. (1976) *Drug Metab. Disposition* **4**(0090-9556 (Print)), 430 - 435
8. Tassaneeyakul, W., Birkett, D. J., Veronese, M. E., McManus, M. E., Tukey, R. H., Quattrochi, L. C., Gelboin, H. V., and Miners, J. O. (1993) *J. Pharmacol. Exp. Ther.* **265**(1), 401-407
9. Venkatakrishnan, K., von Moltke, L. L., and Greenblatt, D. J. (1998) *J. Pharm. Sci.* **87**(12), 1502-1507
10. Kobayashi, K., Nakajima, M., Oshima, K., Shimada, N., Yokoi, T., and Chiba, K. (1999) *Drug Metab. Dispos.* **27**(8), 860-865
11. Minn, A. L., Pelczar, H., Denizot, C., Martinet, M., Heydel, J. M., Walther, B., Minn, A., Goudonnet, H., and Artur, Y. (2005) *Drug Metab. Dispos.* **33**(8), 1229-1237
12. Fukami, T., Nakajima, M., Sakai, H., Katoh, M., and Yokoi, T. (2006) *Drug Metab. Dispos.*

Chapter 5.

Molecular Modeling of Phenacetin in CYP2A13

Introduction: Molecular Modeling in Cytochromes P450

Molecular modeling has been a powerful tool in understanding protein function. The most common types of molecular modeling applied to cytochrome P450s include homology modeling for cytochromes P450 for which a crystal structure is not available, molecular docking to approximate ligand orientation, and 3D quantitative structure-activity relationship (QSAR) models (1-7). Since the xenobiotic-metabolizing cytochrome P450 enzymes each bind a number of different substrates with relatively low affinity (μM), molecular modeling of cytochromes P450 has some unique challenges. Many cytochrome P450 substrates are metabolized at more than one site, suggesting multiple binding orientations, this further emphasizes the difficulties of molecular docking with cytochromes P450. Despite the low substrate affinity and multiple binding orientations, molecular modeling still offers a way to visualize possible ligand orientations within an active site. For this reason, we applied molecular docking simulations to CYP2A13 using phenacetin as the ligand.

We selected Surflex Dock to implement docking simulations. Surflex Dock utilizes Hammerhead's empirical scoring function (8) and incremental construction of the ligand around flexible bonds leading to high affinity docked conformers of phenacetin. The first step is the construction of a protomol or prototypical molecule that represents the active site of the enzyme as mapped by probes complementary to

the active site surfaces. For example, molecule moieties that can act as a hydrogen bond donor (NH), acceptor (OH), or hydrophobic regions (CH_4) are matched to amino acids found in the active site. The active site cavity can be defined either as the largest cavity detected in the enzyme or by using the coordinates of a ligand in the active site defined by a crystal structure. Docked conformers are then incrementally constructed into this protomol and scored on the basis of hydrophobic, polar, repulsive, and entropic interactions as well as solvation between pairs of exposed protein-ligand atoms with a van der Waals surface distance of less than 2 Å (www.tripos.com).

Methods: Surflex Docking

The CYP2A13 enzyme used in the docking study was prepared from the crystallographic coordinates of PDB 2P85 (9) by protonating and adding Gasteiger-Marsili charges in SYBYL (10). Only molecule A was used. Both orientations of the indole active site ligand and waters (except for the water hydrogen bonded to N297) were deleted. The indole ligands (indole 501 and 507 as numbered in PDB 2P85) were prepared by protonating and adding Gateiger-Marsili charges and were saved as separate files for use in identifying the active site cavity. The ligand phenacetin was generated starting from the coordinates of indole ligand, protonated in SYBYL (10), and structurally optimized using the Tripos Force Field (11) (Gasteiger-Marsili electrostatics (12) with an 8.0 Å nonbonding interaction cutoff; default convergence thresholds).

Three different docking simulations were completed using Surflex Dock. The only difference between the three simulations was the method used to define the active site before generating the protomol file. The first simulation utilized indole 501, the second used indole 507, and the third used Surflex Dock's automatic cavity detection (Figure 5.1). The protomol files using indole 507 and automatic cavity detection were the most similar, and differed from the protomol file generated with indole 501 by only one additional hydrogen bond acceptor probe. This small difference appeared to be significant in the docking of phenacetin. The results from the protomol file generated by indole 507 and automatically gave very similar conformers of phenacetin. Several of these conformers were hydrogen bonded to N297, and approximately half the conformers were consistent with the generation of the metabolite acetaminophen. In contrast, the conformers generated using indole 501 were not inconsistent with acetaminophen generation. Additionally, only one of the conformers in the indole 501 trial had an orientation allowing a hydrogen bond between phenacetin and N297, a residue that has been shown in crystal structures of both CYP2A13 and CYP2A6 to form hydrogen bonds with several different ligands (9,13). Since the results from the automatically generated protomol file and the protomol file generated with indole 507 were consistent with the known metabolite and were in strong agreement, the results for indole 507 were analyzed further for structural insights into phenacetin binding.

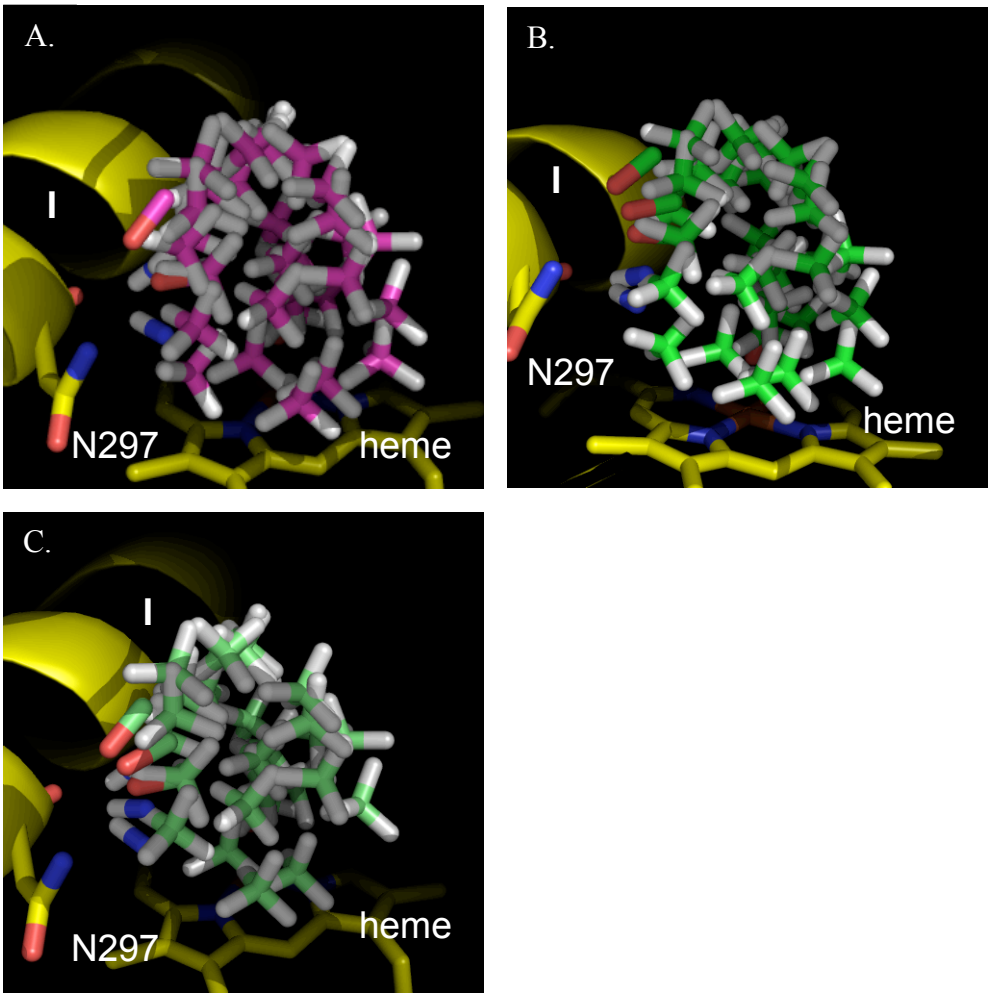


Figure 5. 1. Protomol files generated by Surflex Dock. A. The protomol generated by indole 501. B. The protomol generated by indole 507. C. The protomol generated on automatic cavity detection. Protomols in panels B and C have three hydrogen bond acceptor probes, while the protomol in panel A only has two. The acceptor probes (OH) have a red oxygen, the donor probes (NH) have a blue nitrogen, and the CH4 probes are pink in panel A and green in panels B and C.

Results

Docking of phenacetin into CYP2A13 (PDB 2P85) and examination of the top ten scoring conformations from the protomol generated from indole 507 revealed two general types of orientations in the active site (Table 5.1). Five conformers were in positions oriented with the ether toward the heme iron, consistent with the observed acetaminophen metabolite (Figure 5.2, panel A). In contrast, the second type of orientation had the amide and ethoxy ends swapped, orientating the amide toward the heme (Figure 5.2, panel B). All of the conformations were within 1 log unit from one another in their total Surflex score, which is a compilation of protein-ligand atomic surface distance interactions including steric, polar, entropic, and solubility (Table 5.1).

The five conformers oriented with the ether toward the heme iron primarily orient the phenacetin phenyl ring alongside the I helix between A301 and F300 and perpendicular to the heme, with the amide oxygen within hydrogen bonding distance to N297. Two of these conformers adopt an identical orientation with the carbon adjacent to the ether oxygen oriented for hydrogen abstraction and rebound during *O*-deethylation (H-Fe, ~2.9 Å, pink sticks in Figure 5.2, panel A). The third conformer (blue sticks in Figure 5.2, panel A) overlaps in terms of the phenyl ring and amide components, but differs in the conformation of the flexible alkyl chain. Simple rotation about the phenyl-oxygen bond in this conformer would also orient the hydrogens of the subterminal carbon for *O*-dealkylation. The fourth conformer places the phenyl group alongside the I helix and differs from the other three in that the

terminal carbon is closest to the iron. The fifth conformation places the phenyl ring nearly parallel to the heme with the ether oxygen directly over the iron and maintains a hydrogen bond to the N297 residue (Figure 5.2, panel B).

The five phenacetin conformers with the amide end oriented towards the heme iron belong to two subtypes. In two conformers the ether oxygen can hydrogen bond with N297 (hydrogen bonding distance 2.75-3.22 Å), whereas in the other three orientations distance from N297 was greater than 3.5 Å (Figure 5.3, panel C). These conformers would be more consistent with N- or C-hydroxylation on the amide end of phenacetin. This would lead to metabolites that have been previously identified for CYP1A2 (14) or in rats, rabbits, and hamsters (15-17), but have not been attributed to human 2A enzymes.

Table 5.1. Orientations and log K_D values for the top ten scoring conformers.

Conformation	**Scoring Funtion** (k_D)**	**Orientation for metabolism**	**H-bond to N297 (Å)**	**Metabolite Proposal from Orientation**
1	3.53	ether group toward heme	3.22	Acetaminophen
2	3.26	ether group toward heme (H 4.43 Å from heme)	2.68	Acetaminophen
3	3.18	amide group toward heme	2.75	Acetol
4	3.00	amide group toward heme	*	Acetol
5	2.97	amide group toward heme	*	Acetol
6	2.87	amide group toward heme	2.75	Acetol
7	2.74	ether group toward heme	2.75	Acetaminophen
8	2.74	ether group toward heme	2.75	Acetaminophen
9	2.73	amide group toward heme	2.75	Acetol
10	2.63	ether group toward heme	*	Acetaminophen

*A distance of greater than 3.5 Å to N297 is used to define the upper limit in hydrogen bonding.
** Scoring function includes terms to account for hydrophobic, polar, repulsive, entropic, and solvation of exposed protein and ligand atoms. The term is given in -log(K_D) and is a measure of binding affinity.

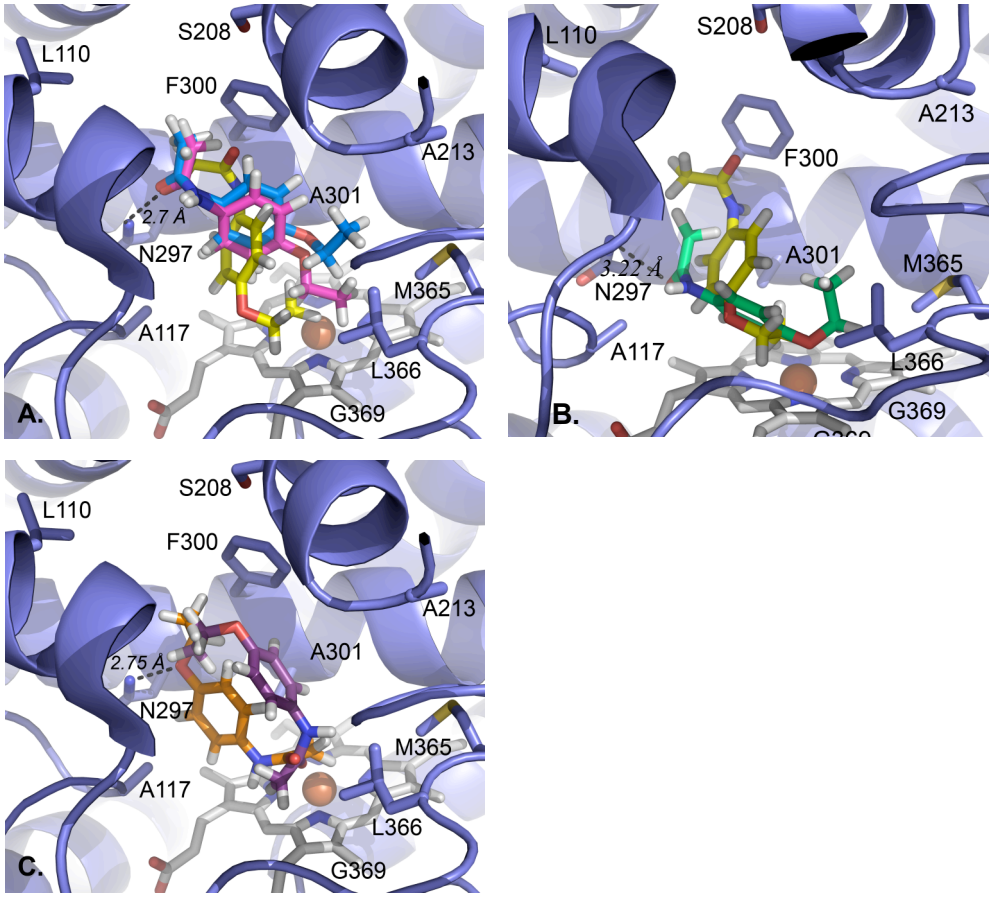


Figure 5.2. Molecular docking of phenacetin into the active site of CYP2A13 (blue, PDB 2P85) leads to two types of orientations. In all pictures, the heme is white with red iron and the view is in the active site with the I helix at the back. Black dashed lines indicate hydrogen bonding to N297. A. Orientations of phenacetin that place the ether oxygen toward the iron and would result in acetaminophen as the metabolite (pink, blue, and yellow). Phenacetin with pink carbons (two conformers) represents the orientation most optimally oriented for hydrogen abstraction from the alkyl subterminal carbon and *O*-deethylation to form acetaminophen. B. Two additional orientations that place the ether group by the heme (green and yellow). C. Orientations of phenacetin with the amide near the heme (orange and purple). The orange conformer represents two conformers and the purple conformer represents the other three conformations.

Discussion

Three of the top ten conformers in our molecular modeling simulation were very consistent with the phenacetin metabolite observed for CYP2A13 (Figure 5.2, panel A). We cannot exclude the possibility that the other conformers observed in our molecular modeling simulation are not viable orientations for metabolism because many cytochromes P450 are able to metabolize a given substrate into more than one metabolite. This has been observed for CYP1A2 with phenacetin (14). Using our current method of metabolite detection, absorbance, only acetaminophen was observed. However, it is possible that by altering our detection method or using more sensitive techniques we could observe other minor metabolites. This is one possible area for further research. For the purposes of identifying possible reasons for phenacetin metabolism and binding by CYP2A13 but not CYP2A6, we analyzed the conformer most consistent with generation of the acetaminophen metabolite.

We aligned the structure of CYP2A6 (molecule A) to the CYP2A13 (molecule A) in using secondary structure matching (SSM) alignment in Coot (18), permitting a graphical overlay in Pymol (Figure 5.3). This way potential structural reasons for phenacetin binding differences in CYP2A13 and CYP2A6 could be observed. Experimentally, we have determined the four residues 208, 300, 301, and 369 mediate phenacetin metabolism in CYP2A13. This was further established in that the CYP2A6 quadruple mutant containing the CYP2A13 amino acids at these positions had phenacetin *O*-deethylation activity and kinetic parameters similar to those of CYP2A13.

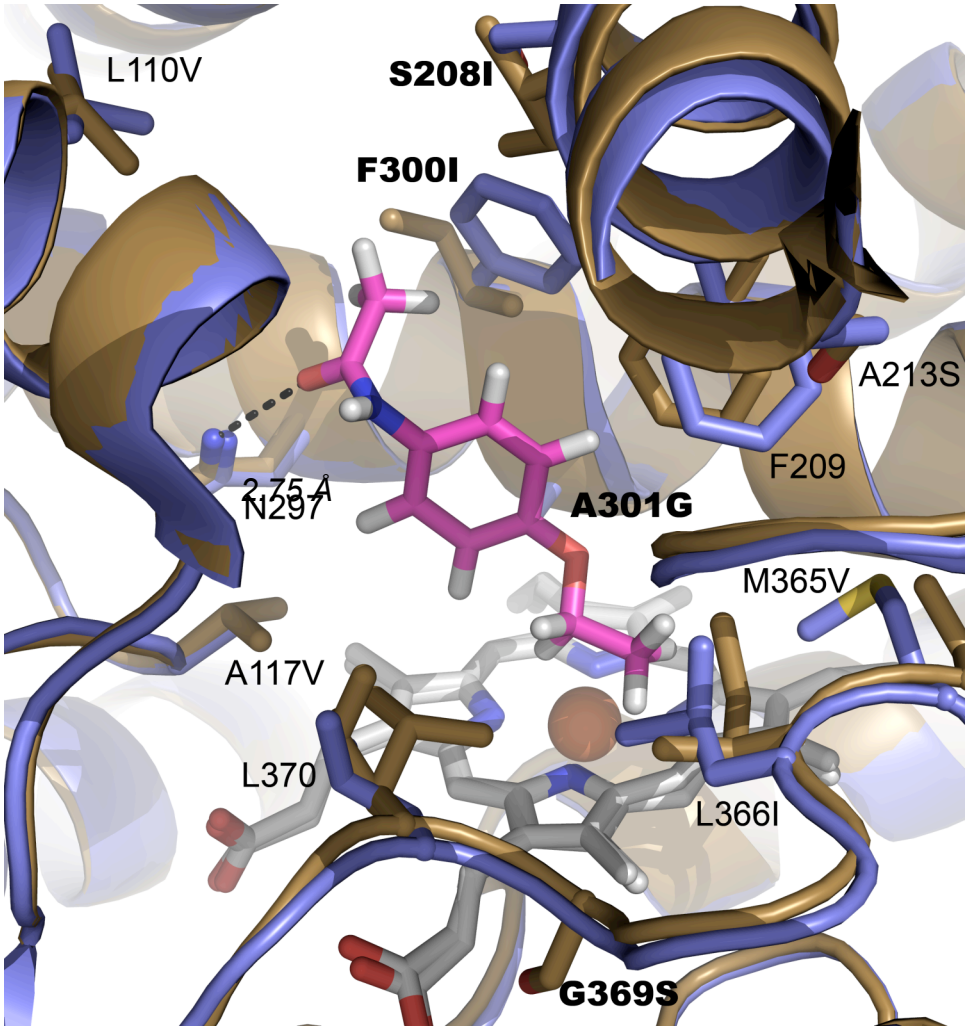


Figure 5.3. Comparison of the active site of CYP2A6 (brown, PDB 1Z10) and the active site of CYP2A13 (blue, PDB 2P85) containing the docked phenacetin conformer most ideally located for *O*-dealkylation (pink). The four residues shown experimentally to be most important for phenacetin metabolism and binding are labeled in bold. The hydrogen bond to Asn297 is illustrated by a dashed black line.

Two of these residues, 300 and 301, are located on the I helix near the ligand. In CYP2A13, residue 300 is a phenylalanine whereas it is an isoleucine in CYP2A6. Although the phenylalanine is larger is size, its planarity places the side chain further away from the active site, leading to more space within the CYP2A13 active site cavity compared to CYP2A6 (Figure 5.4).

CYP2A13 contains an alanine at residue 301, which is also larger than the glycine found in CYP2A6. The slightly larger alanine at 301 may improve hydrophobic interactions between the enzyme and ligand (Figure 5.3).

The third residue that affects substrate enzyme interactions is residue 208. In CYP2A6 this residue is a bulky isoleucine and it is directed towards the 300 residue within the active site cavity (Figure 5.3). However, in CYP2A13 a much smaller serine is present. The substitution to the smaller serine in CYP2A13 may permit the planer phenylalanine found at position 300 to locate with the phenyl group in the volume occupied by the isoleucine in CYP2A6. Orientation of the phenylalanine side chain away from the active site may not be possible without the substitution at residue 208 to a smaller amino acid. This residue may also be involved in the reorientation of the adjacent conserved phenylalanine 209 (9). The conserved F209 forms part of the active site cavity ceiling, and from an overlay of the CYP2A active sites, it is apparent that this residue's position can substantially increase or decrease active site volume (Figure 5.4).

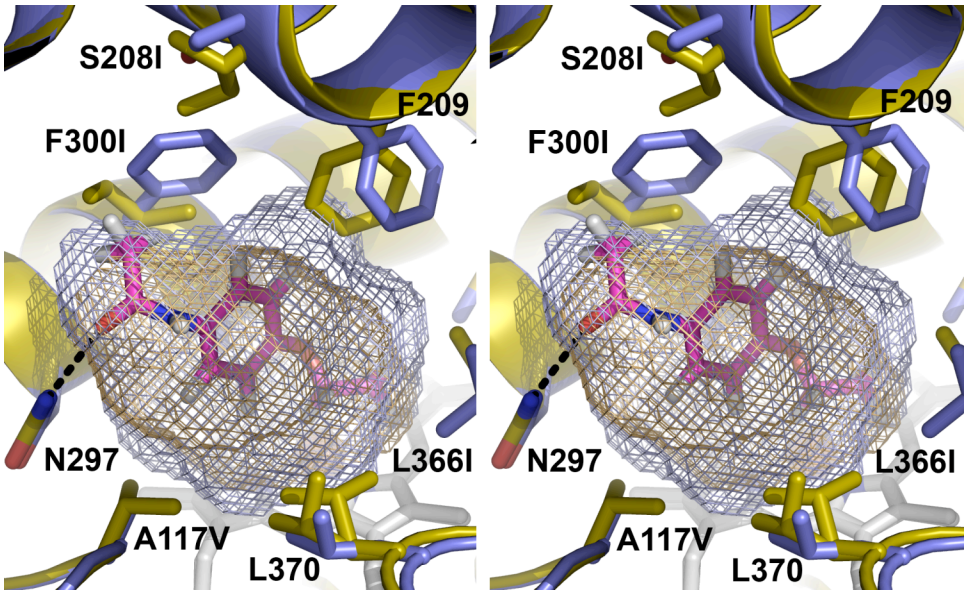


Figure 5.4. Stereo view comparison of the active site cavities for CYP2A6 (PDB 1Z10, gold protein and gold mesh) and CYP2A13 (2P85, blue protein and blue mesh), as determined using Voidoo with a 1.4 $Å^3$ probe (http://xray.bmc.uu.se/). In the CYP2A13 active site the phenacetin conformation most ideally oriented for *O*-dealkylation is shown (sticks with pink carbons). Steric clashes between CYP2A6 and the phenacetin are evident, particularly with CYP2A6 residues 300, 209, and 370. These clashes are relieved by substitution or repositioning of the CYP2A13 amino acids at those positions, which alter the active site volume in ways that would facilitate phenacetin binding in this orientation.

Reasons for the importance of G369S to phenacetin binding and metabolism are likely indirect. Neither residue appears to change the size of the CYP2A active site cavity as the serine side chain found in CYP2A6 is located below the plane of the heme and is not part of the active site wall. In CYP2A6, the serine side chain does participate in a hydrogen bond to one of the propionate oxygens that the glycine in CYP2A13 is unable to fulfill. However, in CYP2A13 this same interaction is conserved by a water molecule hydrogen bonding to the propionate oxygen (9,13). The water molecule is stabilized by additional hydrogen bonding to the L370 backbone amide and carbonyl. It is possible that instead the substitution to the smaller, more flexible glycine found in CYP2A13 may allow the adjacent conserved residue L370 more flexibility. In all of the CYP2A13 molecules, L370 is found oriented toward the protein backbone, away from the active site cavity. However, in CYP2A6 the predominant position of the L370 side chain appears to be toward the active site cavity (13,19). Only molecule A of the CYP2A6 coumarin structure (PDB 1Z10) has an alternate conformation placing the leucine side chain toward the protein backbone (13,19).

Other residues that have notable structural differences between the active sites of CYP2A6 and CYP2A13 include 117, 366, and 365. Mutation of these residues did alter the activity of CYP2A13 mutants. In particular, the mutations A117V and L366I actually caused increased phenacetin binding affinity and metabolism. The increased size of the valine at position 117 may encourage hydrophobic interaction and/or optimal substrate positioning. In contrast, the L366I mutation may actually

result in more space available to the ligand because of a change in the side chain orientation. The amino acid found at position 366 forms part of the active site wall for both CYP2A enzymes; however, the placement of the methyl group off of the beta carbon in the isoleucine side chain found in CYP2A6 allows for an increase in active site volume available to ligand (Figure 5.3). Mutation of residue 365 to the smaller valine found in CYP2A6 decreased both metabolism and binding affinity. V365 in CYP2A6 is not a part of the active site cavity while the larger methionine found at this position in CYP2A13 forms part of the active site wall.

The hypothesis that changes in the active site size may be largely responsible for the difference in phenacetin recognition between CYP2A13 and CYP2A6 is further supported by comparing the overall active site volume (Figure 5.4). The active site of CYP2A13 (370 Å) is larger than that of CYP2A6 (260 Å) (9,13). In particular, an increase in space available to the ligand by the mutations can be seen by residue 300, 209, and also to a lesser degree near residues 370, 366, and 365 (Figure 5.4). Of particular note is that the model phenacetin molecule does not completely fit into the active site volume of CYP2A6 near residue 300. The difference in volume and steric clash around residue 300 indicates that the difference in CYP2A phenacetin metabolism may be largely due to steric constraints in the CYP2A6 active site (Figure 5.4).

Conclusion

The molecular docking of phenacetin into CYP2A13 gives some structural insight as to why residues 208, 300, and 301 are important for phenacetin metabolism. Examination of the active site void volumes of both CYP2A enzymes with one of the phenacetin conformers clearly shows that the phenacetin molecule cannot fit into the CYP2A6 active site in the conformation that was favored in the CYP2A13 modeling study (Figure 5.4). Closer examination of the active site void overlay in Figure 5.4 indicates that the active sites seem to have the greatest variation by the amino acids at position 117, 209, 300, 366, and 370. It is likely that substitution at residue 208 or 369 result in indirect effects on the position of other amino acid side chains, possibly the orientation of residues 209 and 370. It may be possible to determine the affect of the mutations at residues 208, 300, 301, and 369 with a crystal structure of CYP2A6 containing these mutations.

References

1. de Graaf, C., Oostenbrink, C., Keizers, P. H., van Vugt-Lussenburg, B. M., van Waterschoot, R. A., Tschirret-Guth, R. A., Commandeur, J. N., and Vermeulen, N. P. (2007) *Curr Drug Metab* **8**(1), 59-77
2. de Groot, M. J., Ackland, M. J., Horne, V. A., Alex, A. A., and Jones, B. C. (1999) *J. Med. Chem.* **42**(9), 1515-1524
3. Li, W., Tang, Y., Liu, H., Cheng, J., Zhu, W., and Jiang, H. (2008) *Proteins* **71**(2), 938-949
4. Hudelson, M. G., Ketkar, N. S., Holder, L. B., Carlson, T. J., Peng, C. C., Waldher, B. J., and Jones, J. P. (2008) *J. Med. Chem.* **51**(3), 648-654
5. Prasad, J. C., Goldstone, J. V., Camacho, C. J., Vajda, S., and Stegeman, J. J. (2007) *Biochemistry (Mosc).* **46**(10), 2640-2654
6. Lewis, D. F., and Lake, B. G. (2002) *Xenobiotica* **32**(7), 547-561
7. Rahnasto, M., Wittekindt, C., Juvonen, R. O., Turpeinen, M., Petsalo, A., Pelkonen, O., Poso, A., Stahl, G., Holtje, H. D., and Raunio, H. (2007) *Pharmacogenomics J*
8. Jain, A. N. (1996) *J. Comput. Aided Mol. Des.* **10**(5), 427-440
9. Smith, B. D., Sanders, J. L., Porubsky, P. R., Lushington, G. H., Stout, C. D., and Scott, E. E. (2007) *J. Biol. Chem.*
10. Associates', T. T. (2006) *In.* **St. Louis, MO**
11. Clark, M., III, R. D. C., and Opdenbosch, N. V. (1989) *J. Comput. Chem.* **10**, 982-1012
12. Gasteiger, J., and Marsili, M. . (1980) *Tetrahedron* **26**, 3219-3228
13. Yano, J. K., Hsu, M. H., Griffin, K. J., Stout, C. D., and Johnson, E. F. (2005) *Nat Struct Mol Biol* **12**(9), 822-823
14. Yun, C. H., Miller, G. P., and Guengerich, F. P. (2000) *Biochemistry (Mosc).* **39**(37), 11319-11329
15. Hinson, J. A. (1983) *Environ. Health Perspect.* **49**(0091-6765 (Print)), 71-79
16. Fischbach, T., and Lenk, W. (1985) *Xenobiotica* **15**(11), 915-927
17. Fischbach, T., and Lenk, W. (1985) *Xenobiotica* **15**(2), 149-164
18. Emsley, P., and Cowtan, K. (2004) *Acta crystallographica* **60**(Pt 12 Pt 1), 2126-2132
19. Yano, J. K., Denton, T. T., Cerny, M. A., Zhang, X., Johnson, E. F., and Cashman, J. R. (2006) *J. Med. Chem.* **49**(24), 6987-7001

Chapter 6.

Structure of CYP2A6 I208S/I300F/G301A/S369G complexed with Phenacetin

Introduction

The molecular modeling study discussed in chapter 5 offered potential insight into why the residues at 208, 300, 301, and 369 are important for CYP2A phenacetin metabolism. However, due to the nature of molecular modeling with cytochromes P450 and because molecular modeling may not accurately predict all the potential structural changes mediated by multiple mutations, we decided to pursue a crystal structure of the quadruple mutant. We obtained a 2.15 Å crystal structure of the mutant CYP2A6 I208S/I300F/G301A/S369G protein with phenacetin in the active site. This structure, combined with the known structures of the CYP2A6 and CYP2A13 wild type enzymes, verifies the structural basis for the roles of amino acids at positions 208, 300, 301, and 369 in phenacetin metabolism. This structure also validates the molecular modeling study.

Crystallography Methods

Purification Alterations: The CYP2A6 I208S/I300F/G301A/G369S protein used for crystallography was purified as described in chapter 2, except that 1% ANAPOE-35 (Anatrace, Maumee, OH) was used instead of 4.8 mM Cymal-5 as the purification detergent (Anatrace, Maumee, OH). The change in purification detergent was necessary to generate diffraction quality crystals, and lengthened the growth time from five hours for crystals purified in Cymal-5 to three days in ANAPOE-35. This

protein was additionally purified on a Superdex 200 16/60 gel filtration column (GE Healthcare, Uppasala, Sweden) in buffer A (50 mM KPi, pH 7.4, 20% glycerol, 0.5 M NaCl, and 1 mM EDTA) following the CM column to ensure protein was in a monodisperse, monomeric state. Fractions containing CYP2A6 I208S/I300F/G301A/G369S were pooled and concentrated by centrifugal ultrafiltration with 5 mM phenacetin in buffer A. At this concentration, a partial spectra shift from the soret of ~416 nm to ~392 nm was observed. Due to solubility problems, we were unable to add more than 5 mM of phenacetin to the protein. Prior to crystallization, a final concentration of 0.2% ANAPOE-X-405 (Anatrace, Maumee, OH) was added to the concentrated protein.

Protein Crystallization, Data Collection, and Structure Determination: The hanging drop diffusion method was used to grow crystals of 100 μM CYP2A6 I208S/I300F/G301A/G369S with 5 mM phenacetin. A 2μL of the protein solution (1 μL 100 μM CYP2A6 I208S/I300F/G301A/G369S and 1 μL of the equilibration solution) was equilibrated against 750 μL of equilibration solution (30% PEG 3350, 0.10 M Tris, pH 8.5, and 0.2 M ammonium sulfate). The protein crystals grew in a mixture of plates and rods that appeared to be mature after four days at 20ºC. Data was collected on one medium sized rod. Prior to data collection, crystals were flash cooled in liquid nitrogen after being immersed in a cryoprotectant consisting of 350 μL of synthetic mother liquor, 350 μL of buffer A, and 300 μL of 100% ethylene glycol. A data set was collected from a single crystal at beamline 9-2 at the Stanford Synchrotron Radiation Laboratory (Stanford, CA) using a 0.98 Å wavelength and

temperature of 100 K (Figure 6.1). The best diffraction data was collected by centering the X-ray beam on the side of the rod protruding from the loop (Figure 6.1). One hundred and four degrees with 1º oscillations and 5 second exposure/oscillation were collected (Figure 6.1). Data processing was completed using Mosflm (1) and Scala (2). The CYP2A6 I208S/I300F/G301A/G369S mutant structure was solved by molecular replacement with the program PHASER (3) using CYP2A6 complexed with coumarin (PDB 1Z10, molecule B) with the ligand and water molecules removed as the search model. Model building and refinement were done iteratively using COOT (4) and REFMAC5 (5) in the CCP4 suite (6). The final model had an R value of 0.21 and R_{free} value of 0.27 with four protein monomers and four heme prosthetic groups in the asymmetric unit. Molecules A, B, and D of the asymmetric unit contained well-defined electron density above the heme prosthetic group, but molecule C did not, so only three molecules of phenacetin were modeling in. Data collection and processing statistics are provided in Table 6.1.

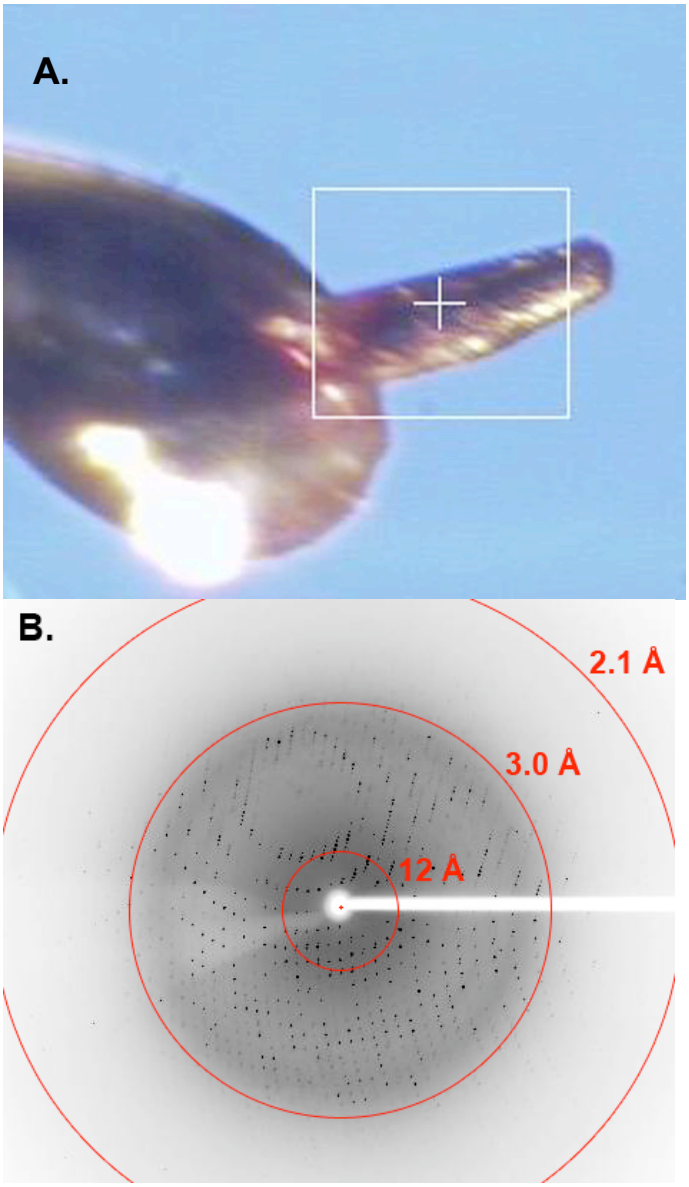


Figure 6.1. Data collection images. A. Snapshot of the rod shaped CYP2A6 I208S/I300F/G301A/S369G crystal that the data set was collected on in the loop during data collection. The white box around the protruding end of the crystal indicates the area of the crystal within the X-ray beam. Data was collected for a total of 104 degrees. B. An x-ray diffraction images at zero degree oscillation.

Table 6.1. Data collection and refinement statistics.

	CYP2A6 I208S/I300F/G301A/S369G with Phenacetin
Data collection	
Space group	$P2_1$
Cell dimensions	
a, b, c (Å)	70.8, 159.1, 104.0
α, β, γ (°)	90.0, 92.1, 90.0
Resolution[a] (Å)	87.04 - 2.15 (2.21 - 2.15)
Total observations[a]	270986 (19484)
Unique reflections[a]	121786 (8964)
Completeness[a] (%)	97.8 (97.5)
R_{sym}[a,b]	0.059 (0.357)
I/σI[a]	10.5 (1.6)
Refinement	
Resolution (Å)	87.04 - 2.15
No. reflections	115652
R/R_{free} (%)	21.1/26.9
No. atoms	
Protein	15029
Ligand	39
Heme	172
Water	297
B-factors (average)	
Protein	39.2
Ligand	66.6
Heme	27.2
Water	35.4
Rms. deviations	
Bond lengths (Å)	0.014
Bond angles (°)	1.5

[a] Values in parentheses are for highest-resolution shell.

[b] $R_{sym} = \Sigma_{hkl}\Sigma_i \left| I_i(hkl) - \langle I(hkl) \rangle \right| / \Sigma_{hkl}\Sigma_i I_i(hkl)$

Structure Validation and Analysis: The CYP2A6 I208S/I300F/G301A/G369S structure was validated using the programs WHAT-IF (7) and PROCHECK (8). The Ramachandran plot showed 90.7% of residues in the most favored regions, 8.6% in the additional allowed regions, 0.4% in the generously allowed regions, and 0.2% in the disallowed regions. F42 was in the disallowed region for all four molecules, but was clearly defined in the $2|F_o|-|F_c|$ electron density map. This was consistent with previous structures of CYP2A6, all of which contained F42 in the disallowed region (*9, 10*). All voids used in structural analysis were calculated using VOIDOO (11) with a probe radius of 1.4 Å and a grid mesh of 0.3 Å.

Results: Structure Analysis

The overall tertiary structure of the CYP2A6 quadruple mutant was very similar to both wild type CYP2A6 and CYP2A13 with root mean squared deviations (r.m.s.d.) of the Cα atoms of 0.39 and 0.47 Å, respectively. These r.m.s.d. variations were comparable to the r.m.s.d. between the four molecules of the CYP2A6 quadruple mutant asymmetric unit (0.29 to 0.48 Å). The greatest variation in tertiary structure was observed in the solvent exposed loops.

All four mutations were clearly evident in the electron density. The mutated side chains F300 and A301 were in the same orientation in all four monomers. Both F300 and A301 are found on the I helix. F300 forms the top portion of the active site, whereas A301 forms part of the side (Figure 6.2, panel A). The S208 mutation is directly above F300. The S208 side chain is rotated toward the active site in molecule A, while in molecules B, C, and D it is angled away from the active site

(Figure 6.2, panel B). The rotation of this side chain does not appear to alter the active site volume as molecules A and C had nearly identical active site volumes.

In addition, electron density corresponding to phenacetin was present in three of the four protein active sites. This density indicated phenacetin oriented with the acetamide directed toward residue 300 and the ethoxy group oriented toward the heme iron, consistent with the experimentally observed *O*-deethylation (Figure 6.2, panel A). Phenacetin appears to have some flexibility in orientation as the molecule varied somewhat between all three molecules, in terms of the plane of the phenyl ring and the flexible ethoxy tail (Figure 6.1, panel B). There is, however, a conserved hydrogen bond (2.9-3.14 Å) between N297 for each molecule and the amide oxygen of phenacetin and the distance from the subterminal carbon and the iron was maintained at 4-4.5 Å in each monomer.

The orientation of the L370 side chain varied between molecules. Molecule A and C had the leucine side chain oriented so that it is rotated away from the phenacetin (Figure 6.2, panel B). In monomer A, the terminal carbon of phenacetin ethoxy chain projects toward this side chain and the repositioning of L370 makes additional space available to the ligand. This side chain had an alternate conformation only in monomer A of four monomers in the CYP2A6 structure complexed with coumarin (PDB 1Z10) (10). Due to the rotation of L370 away from the active site, the active site volumes for molecules A and C are 300 $Å^3$ and 297 $Å^3$, respectively. In contrast, the active site volume is reduced in molecules B (270 $Å^3$) and D (272 $Å^3$) (Figure 6.2, panel C). Corresponding to the reduced volume

available, in molecule B and D the terminal carbon of the phenacetin ethoxy side chain projects toward the I helix. Replacement of the adjacent S369 by the glycine present in CYP2A13 may result in additional flexibility for L370 in the CYP2A6 mutant and the CYP2A13 wild type enzyme.

The active site of the quadruple mutant is primarily hydrophobic, with the hydrophilic residue N297 as the most orientating feature. Despite the small active site volume, phenacetin has some degree of flexibility within the active site, as indicated by its different orientation in the three protein molecules.

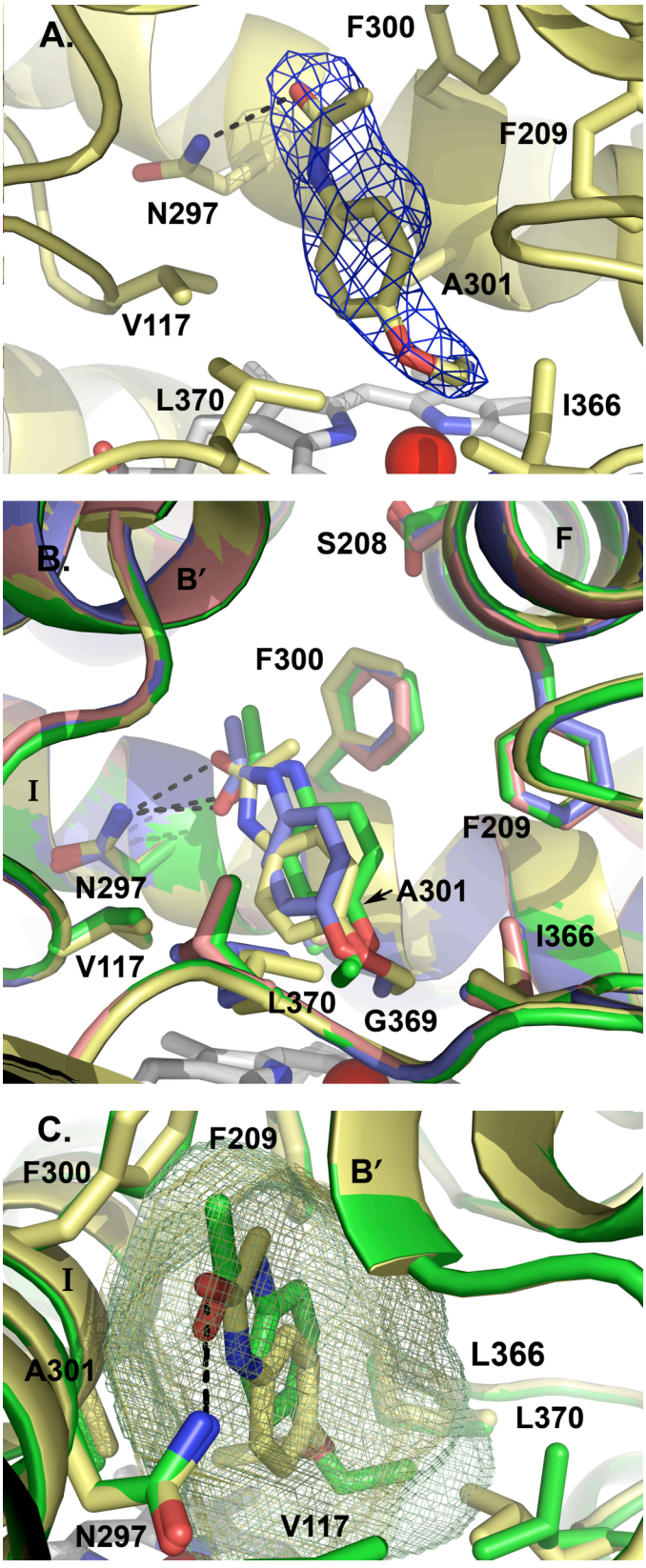


Figure 6.2. The CYP2A6 I208S/I300F/G301A/G369S structure. A. Electron density surrounding the phenacetin ligand found in molecule B of the CYP2A6 I208S/I300F/G301A/G369S structure ($2|F_o|-|F_c|$). B. Structural overlay of the non-identical protein molecules found in the CYP2A6 I208S/I300F/G301A/G369S asymmetric unit, highlighting the differences between the three phenacetin molecules and L370 positions. Molecules A (green), B (yellow), and D (blue) contain one molecule of phenacetin, while molecule C (pink) did not have electron density for phenacetin. C. Active site volume comparison between molecule A (green mesh, 300 $Å^3$) and molecule B (yellow mesh, 270 $Å^3$). The difference in active site volume among the four CYP2A6 quadruply-mutated protein molecules appears to depend largely on the position of L370. Black dashed lines indicate hydrogen bonding to N297.

Comparison of the CYP2A6 I208S/I300F/G301A/G369S Structure to the Structures of Wild Type CYP2A6 and CYP2A13

Insight into how mutations at residues 208, 300, 301, and 369 can significantly alter CYP2A functionality is gained by comparing the CYP2A13, CYP2A6, and the CYP2A6 I208S/I300F/G301A/G369S active sites (Figure 6.3). CYP2A6 and CYP2A13 have some of the smallest human P450 active sites, with a volume of 214-230 $Å^3$ for the four CYP2A6 molecules (PDB 1Z10) and a volume of ~304 $Å^3$ (PDB 2P85) for CYP2A13 (*12*). The four quadruple mutant molecules had intermediate volumes from 270 – 300 $Å^3$, depending primarily on the L370 conformation.

The residues 300 and 301 are critical in phenacetin binding and metabolism, consistent with their close interaction with phenacetin in the crystal structure. The planer phenylalanine at position 300 in CYP2A13 and the quadruple mutant results in additional active site space in this region to accommodate the phenacetin acetamide. In contrast, the bulkier isoleucine in CYP2A6 actually results in a decrease in active site volume available to the ligand because the isoleucine side chain decreases the active site area available to ligand near residue 300. This is evident when viewing the active site cavities as a steric clash occurs between CYP2A6 I300 and the terminal carbon of the amide end of phenacetin (Figure 6.4).

Both the CYP2A13 and CYP2A6 quadruple mutant structures place the methyl side chain of alanine at position 301 toward the face of phenacetin just proximal to the ring.

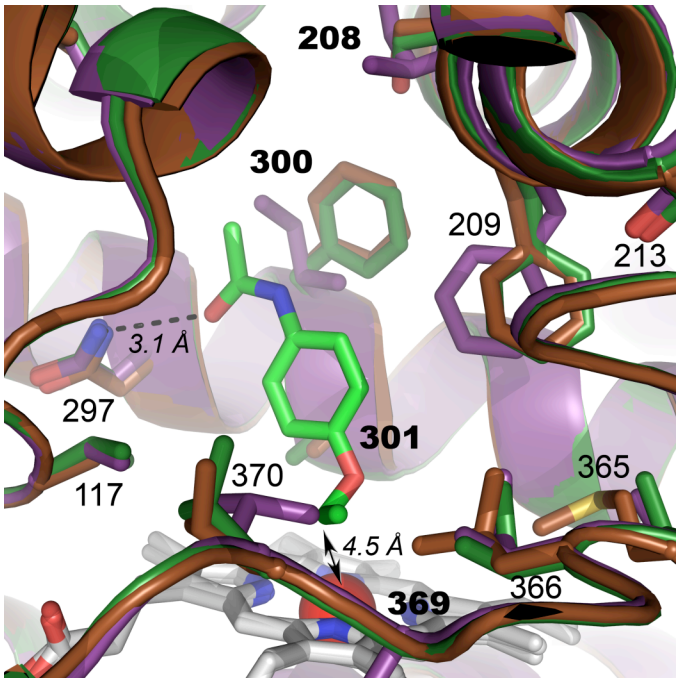


Figure 6.3. Active site comparison of CYP2A6 (purple, PDB 1Z10, molecule B), CYP2A6 I208S/I300F/G301A/G369S with phenacetin (green, molecule A), and CYP2A13 (brown, PDB 2P85, molecule A). The four residues mutated in the CYP2A6 quadruple mutant, identical to the residues found in CYP 2A13, are labeled in bold. The hydrogen bond from phenacetin to N297 is shown as a dashed black line.

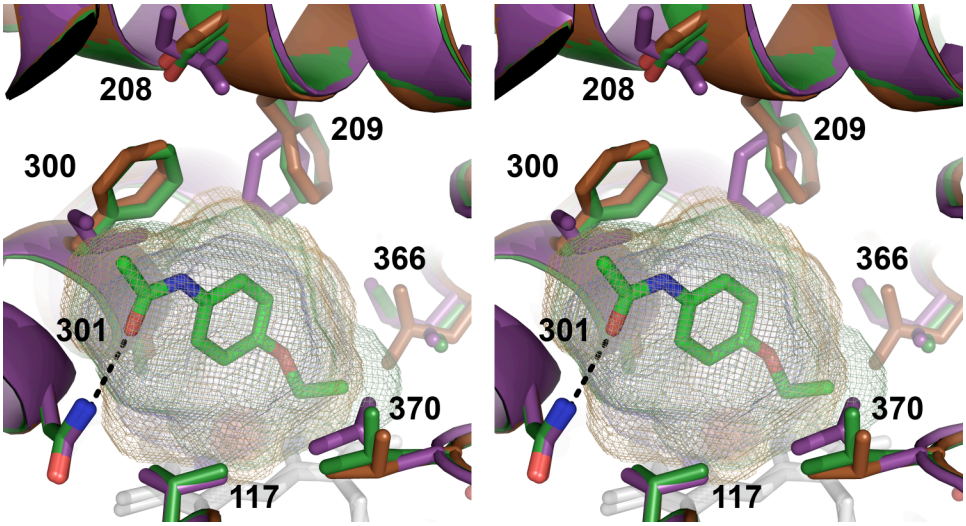


Figure 6.4. Comparison of the active site cavities of CYP2A6 (purple protein and mesh, PDB 1Z10), CYP2A13 (brown protein and mesh, PDB 2P85), and CYP2A6 I208S/I300F/G301A/G369S (green protein and mesh) in stereo view. Steric clashes between the phenacetin and purple CYP2A6 residues I300, F209, and L370 are evident.

It might seem counterintuitive that the smaller glycine does not improve substrate metabolism and binding. However, it may be that the glycine found in CYP2A6 removes necessary packing interactions to stabilize phenacetin in the active site and adds additional flexibility to the I helix that is not favorable.

The third mutated residue essential for phenacetin metabolism is residue 208. In CYP2A13 and the quadruple mutant, this residue is a serine and located in helix F directly above residue 300. There is a substantial reduction in size of the side chain compared to the isoleucine found in CYP2A6, and furthermore, S208 is torsioned away from the active site in the majority of the quadruple mutant molecules. This also creates space that may facilitate phenyl ring of F300. This would indirectly increase the space for available for phenacetin binding. In addition, the residues at position 208 and 300 may alter the orientation of the adjacent conserved F209 (Figure 6.3). The orientation of F209 in the CYP2A6 quadruple mutant structure is identical to the position of F209 in CYP2A13, placing the phenylalaine away from phenacetin and thus creating an active site volume that better accommodates phenacetin (Figure 6.4).

The fourth residue, 369, is located on the opposite side of the active site cavity. The serine found at this position in CYP2A6 interacts with the heme propionate (Figure 1.4). In CYP2A13 and the CYP2A6 quadruple mutant, a glycine is found at this position. A water molecule instead mediates the contact maintained with the heme propionate by the serine in CYP2A6. Therefore, the residue at position 369 appears to have an indirect role in phenacetin binding by modulating the

positioning of the adjacent L370 residue. In CYP2A6 the adjacent residue L370 usually projects into the active site where phenacetin would otherwise bind while this side chain is oriented along the loop forming the wall of the active site in CYP2A13. In two of the four quadruple mutant molecules, the L370 side chain is in the CYP2A13 position. In the 1Z10 structure of CYP2A6, there is one molecule with this second CYP2A13-like orientation of L370, but the lack of steric conflicts when glycine is at 369 may facilitate the repositioning of L370 as seen in CYP2A13. Ultimately, this may promote phenacetin binding and metabolism because the volume is more compatible with phenacetin binding and metabolism (Figure 6.4).

The CYP2A6 I208S/I300F/G301A/G369S crystal structure also gives structural insight into the two CYP2A13 mutations, A117V and L366I, that resulted in substantial increases in phenacetin k_{cat} above that of the CYP2A13 wild type enzyme. The first, CYP2A13 A117V, resulted in a 40% decrease in K_D, a 4.5-fold increase in K_m, and an overall two-fold increase in catalytic efficiency, suggesting multiple effects in binding and catalysis. Although the quadruple mutant's valine side chain projects further into the active site toward the phenacetin than the CYP2A13 alanine, the reduction in volume in this portion of the active site may be beneficial in orientating the phenacetin within the active site (Figure 6.3). The L366I mutation had a dramatically increased binding affinity, but an overall catalytic efficiency similar to wild type CYP2A13, with ~3.5-fold increases in both K_m and k_{cat}. The residue 366 is located close to residues 365 and 369 near the alkyl chain of

phenacetin. Mutation from the leucine found in CYP2A13 to the isoleucine in CYP2A6 would also likely increase the size of the active site cavity. This is evident in the substantial increase in active site volume near this residue in the quadruple mutant compared to CYP2A13 (Figure 6.4).

Validity of Molecular Modeling Experiments

The large degree of flexibility for phenacetin in the active site is reminiscent of the various orientations obtained in the molecular modeling of phenacetin docking into CYP2A13. The molecular modeling placed the phenacetin molecule with either the ethoxy or the acetamide end orientated toward the heme with equal probability. When refining the CYP2A6 I208S/I300F/G301A/S369G crystal structure, we modeled the phenacetin molecule in both directions. After multiple rounds of refinement, the phenacetin density clearly favored orientation of the ligand with the ethoxy end of phenacetin toward the heme. However, we cannot exclude the possibility that some phenacetin molecules can bind with the acetamide end toward the heme as suggested by the molecular modeling study. The molecular modeling program did capture several of the most defining features of phenacetin binding in the orientation observed in the crystal structure.

A majority of the top ten orientations in the molecular modeling study maintained a hydrogen bond with N297, which was observed in all three phenacetin molecules in the crystal structure. In addition, the relative location of phenacetin is consistent with that of the phenacetin molecules in the crystal structure (Figure 6.5). The plane of the phenyl ring also varied between the three molecules of the

asymmetric unit, indicating that more than one plane is allowed for ligand orientation within the active site.

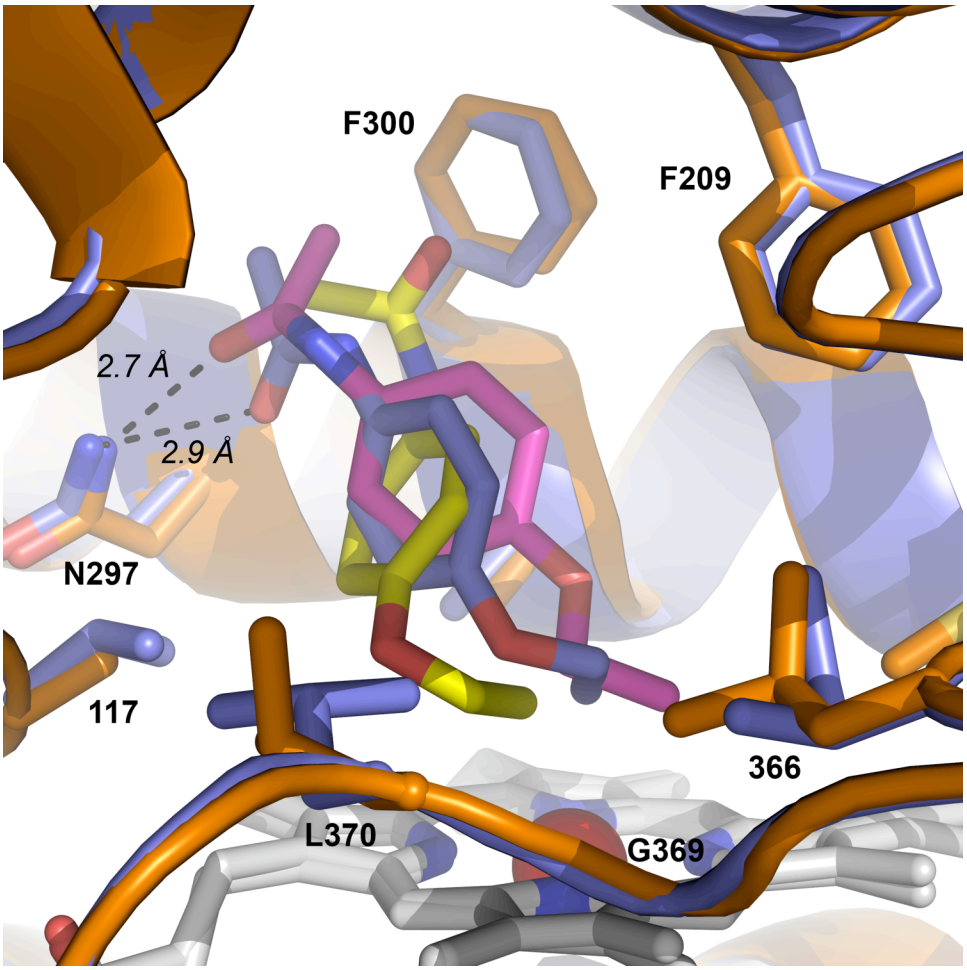


Figure 6.5. The overlay of the CYP2A6 I208S/I300F/G301A/S369G mutant molecule D (blue) and the CYP2A13 molecule A (orange). The phenacetin molecule found in the CYP2A6 quadruple mutant is also colored blue. Two of the phenacetin conformers from the CYP2A13 molecular modeling simulation were also included for comparison (yellow and pink). Hydrogen bonds are indicated by black dashed lines.

Conclusions

The crystal structure of CYP2A6 I208S/I300F/G301A/S369G complexed with phenacetin provided experimental validation into the importance of residues 208, 300, 301, and 369 on phenacetin binding and metabolism by human CYP2A enzymes. Comparison to the active sites of CYP2A13 and CYP2A6 crystal structures suggest that the ability of CYP2A enzymes to metabolize phenacetin is largely the culmination of steric effects of the substituted side chains themselves and their effects on the positions of adjacent side chains F209 and L370. Although many of the same structural conclusions could be drawn by examination of the molecular docking results, obtaining the crystal structure was important confirmation and allowed for direct analysis of the four key amino acid substitutions (13). As hypothesized, the substitution of the CYP2A13 amino acids found at positions 208, 300, 301, and 369 into CYP2A6 resulted in a mutant enzyme with an active site structurally similar to CYP2A13.

References

1. Leslie, A. G. (2006) The integration of macromolecular diffraction data, *Acta crystallographica 62*, 48-57.
2. Evans, P. (2006) Scaling and assessment of data quality, *Acta crystallographica 62*, 72-82.
3. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *Journal of Applied Crystallography 40*, 658-674.
4. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta crystallographica 60*, 2126-2132.
5. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta crystallographica 53*, 240-255.
6. (1994) The CCP4 suite: programs for protein crystallography, *Acta crystallographica 50*, 760-763.
7. Vriend, G. (1990) WHAT IF: a molecular modeling and drug design program, *J Mol Graph 8*, 52-56, 29.
8. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *Journal of Applied Crystallography 26*, 283-291.
9. Sansen, S., Hsu, M. H., Stout, C. D., and Johnson, E. F. (2007) Structural insight into the altered substrate specificity of human cytochrome P450 2A6 mutants, *Arch. Biochem. Biophys. 464*, 197-206.
10. Yano, J. K., Hsu, M. H., Griffin, K. J., Stout, C. D., and Johnson, E. F. (2005) Structures of human microsomal cytochrome P450 2A6 complexed with coumarin and methoxsalen, *Nat Struct Mol Biol 12*, 822-823.
11. Kleywegt, G. J., and Jones, T. A. (1994) Detection, delineation, measurement and display of cavities in macromolecular structures, *Acta crystallographica 50*, 178-185.
12. Porubsky, P. R., Meneely, K.M., Scott, E.E. . Structures of Human Cytochrome P450 2E1: Insights into the Binding of Inhibitors and Both Small Molecular Weight and Fatty Acid Substrates, *J. Biochem. To be Published*.
13. Jain, A. N. (2003) Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine, *J. Med. Chem. 46*, 499-511.

Chapter 7.

Spectral Binding Assays of CYP2A6 Multiple Mutants

Introduction

To determine if the four amino acids important for mediating phenacetin binding and metabolism in CYP2A enzymes (chapters 3-6) also collectively mediate the selectivity of other CYP2A substrates, spectral ligand binding assays were conducted using the CYP2A6 triple and quadruple multiple mutants with coumarin, 2′-methoxyacetophenone (MAP), phenethyl isothiocyanate (PEITC) and phenacetin as ligands. These ligands were selected so the K_D values could be compared to those determined for the single CYP2A13 mutants in chapter 3.

Methods

All proteins used in this chapter were expressed and purified according to the protocol found in chapter 2. The procedure for the spectral ligand binding assay is also found in chapter 2.

Results and Discussion

Spectral Binding Assays of CYP2A6 Multiple Mutants: The largest variation in CYP2A binding affinity was with phenacetin as a ligand. Phenacetin does not cause a significant spectral shift in CYP2A6 at concentrations up to 250 μM, while CYP2A13 has a dissociation constant of 34 μM (Figure 3.2). However, both CYP 2A6 I300F/G301A/S369G and quadruple CYP2A6 I208S/I300F/G301A/S369G mutants had phenacetin dissociation constants below that of the wild type CYP2A13. The CYP2A6 I208S/I300F/G301A mutant had a K_D of 63 μM, approximately two-

fold higher than wild type CYP2A13 (Table 7.1). This indicates that substitution at 369 has a greater effect on phenacetin binding than substitution at position 208. Unlike phenacetin metabolism, which required all four mutations to reach phenacetin *O*-deethylation metabolism levels similar to those of CYP2A13, the mutation of position 208 to the CYP2A13 residue is not required to attain phenacetin binding affinity similar to CYP2A13 and appears to be somewhat detrimental.

The ligand 2′-methoxyacetophenone also had a substantially different binding affinity for CYP2A13 (0.85 μM) compared to CYP2A6 (74 μM). The CYP2A6 triple and quadruple mutants all had binding affinities that were reduced from the wild type CYP2A6, with a dissociation constant of ~ 10-12 μM for the triple mutants and a K_D value of 3.5 μM for the quadruple mutant (Table 7.1).

CYP2A13 and CYP2A6 have a 14-fold difference in binding affinities for phenethyl isothiocyanate with dissociation constants of 0.43 μM and 6.1 μM, respectively. The CYP2A6 I208S/I300F/G301A mutant had slightly increased binding affinity compared to wild type CYP2A6 with a K_D of 4.9 μM (Figure 7.1). The CYP2A6 I300F/G301A/S369G and CYP2A6 I208S/I300F/G301A/S369G mutants both had binding affinities ~ 2 μM, a value intermediate to the wild type enzymes.

Table 7.1. K_D values for wild type CYP2A enzymes and CYP2A6 multiple mutants with four ligands. K_D values for independent trials are in parenthesis, average of the two trials is below. Phenacetin K_D values represent preliminary data from one trial.

Cytochrome P450	MAP K_D (μM)	PEITC K_D (μM)	Coumarin K_D (μM)	Phenacetin K_D (μM)
2A6	(65, 69, 104)	(5.4, 6.9)	(2.7, 3.4)	N.D.[a]
	74	6.1	3.1	
2A13	(1.0, 0.66)	(0.31, 0.54)	(2.0, 3.3)	(30, 35, 37)
	0.85	0.43	2.7	34
2A6 mutants:				
I208S/I300F/G301	(8.2, 10.7)	(5.8, 4.0)	(6.0, 4.6)	
	9.5	4.9	5.3	63
I300F/G301A/S369G	(12.6, 12.0)	(2.3, 1.5)	(2.6, 3.6)	
	12.3	1.9	3.1	13
I208S/I300F/G301A/S369G	(2.4, 4.5)	(2.6, 1.5)	(2.5, 2.1)	21
	3.5	2.1	2.3	

[a] N.D. Not determined as only a reduced spectral shift was detected, which did not allow determination of K_D see Figure 3.2.

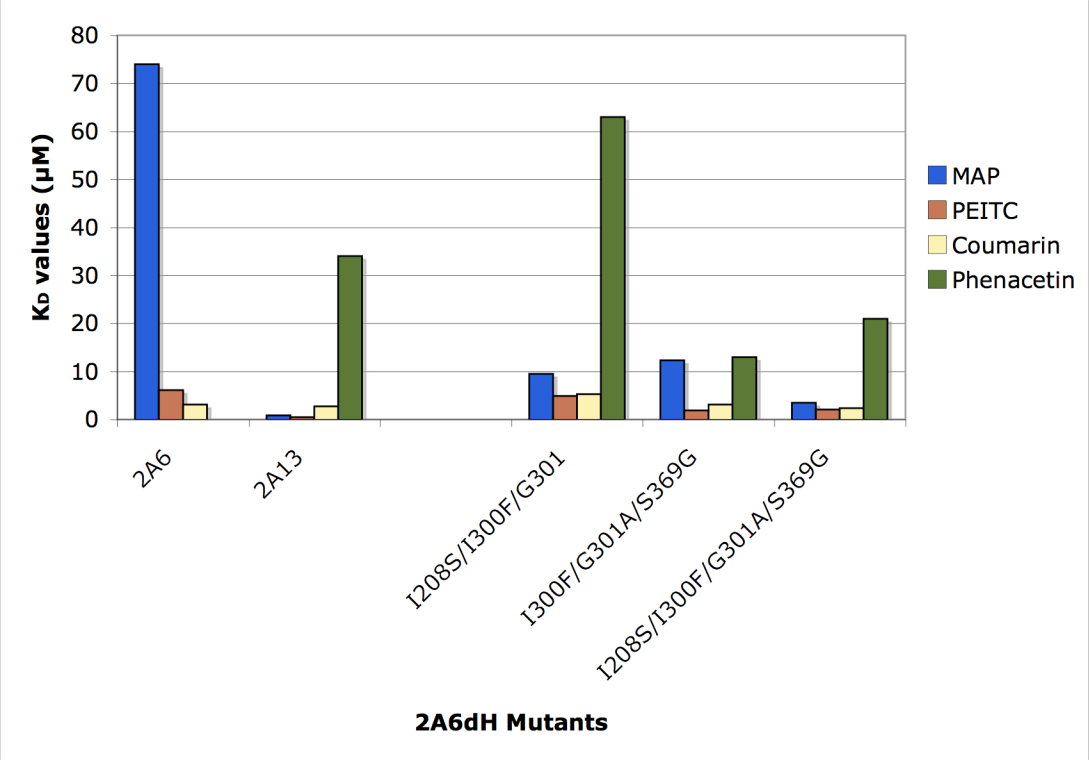


Figure 7.1. Binding affinity comparison of the two wild type CYP2A enzymes and three of the CYP2A6 multiple mutants.

The binding affinity of both CYP2A13 and CYP2A6 is ~3 μM for coumarin. Not surprisingly, both CYP2A6 I300F/G301A/S369G and CYP2A6 I208S/I300F/G301A/S369G mutants had a binding affinity for coumarin similar to the wild type enzymes (Table 7.1). The CYP2A6 I208S/I300F/G301A mutant had a slightly decreased binding affinity, but this was less than two-fold change, with a K_D of 5.3 μM. It is notable that the single S208I mutation in CYP2A13 also caused a decrease in coumarin binding affinity.

Conclusion

These spectral ligand binding titrations indicate that the four amino acids that are essential for CYP2A phenacetin binding and metabolism are also important for binding affinity for PEITC and MAP. The quadruple mutant CYP2A6 I208S/I300F/G301A/S369G had ligand binding affinities for MAP, PEITC, and phenacetin more similar to CYP2A13 than to the wild type CYP2A6 enzyme (Figure 7.1). This is consistent with the observations made during analysis of the CYP2A6 I208S/I300F/G301A/S369G crystal structure that revealed the quadruple mutant active site was similar to CYP2A13 (Figure 6.4). However, there were some differences between the mutant CYP2A6 I208S/I300F/G301A/S369G active site void and the active site void for CYP2A13. In particular, the region around residue 117 has more active site volume in CYP2A13 than in CYP2A6 I208S/I300F/G301A/S369G. Although this decrease in active site volume may have been beneficial for phenacetin metabolism, as indicated by the increased k_{cat} in the

mutant CYP2A13 A117V (Figure 4.3), more space may be required in this portion of the active site for other ligands.

Based on this spectral binding data, it may be that to completely convert the binding affinity of CYP2A6 to that of CYP2A13, additional residue substitution may be required based on the specific ligand. In chapter 3, it is apparent that the extent of the effects of mutating some residues in CYP2A13 was ligand dependent. Although for phenacetin, the substitution of just three amino acids, those at positions 300, 301, and 369, were required to convert the CYP2A6 enzyme to a binding affinity that was actually greater than that of CYP2A13, for other ligands the inclusion of other important substitutions may be necessary to gain increased binding affinity conversion. For example, with the CYP2A13 mutant spectral binding studies, mutation of residue 117 in CYP2A13 to the valine found in CYP2A6 significantly decreased MAP binding affinity (Table 3.2). The inclusion of this mutation in the CYP2A6 quadruple mutant could potentially result in an enzyme capable of binding MAP tighter. Furthermore, residue 208, which was important for phenacetin metabolism, did not seem to have as much of an effect on phenacetin binding. This indicates that some substitutions that affect substrate metabolism may do so in ways other than directly mediating substrate binding. In order to have a complete understanding of the substrate specificity for these other ligands in light of the CYP2A active site structure, metabolism assays could be pursued in addition to the substrate binding studies.

The study described in the chapter substantiates the hypothesis that only a few amino acids are largely responsible for substrate selectivity differences in CYP2A enzymes. For the differentially binding ligands assayed, the four mutations that were identified as essential for pheneacetin binding and metabolism also led to significant increases in binding affinity, approaching the K_D values observed for CYP2A13. Thus, it is likely that the four residues identified (208, 300, 301, and 369) are also key residues that mediate CYP2A substrate specificity for other ligands. Further study would be required to determine whether these residues are also essential for selective metabolism of these substrates and whether other residues besides these four are important to substrate specificity in a ligand dependent manner.

Chapter 8.

A Functional Study of CYP2A13 Polymorphisms

Introduction

We have established in prior chapters that the difference in CYP2A substrate selectivity is primarily mediated by a few amino acids in or near the active site that differ between the two CYP2A enzymes. However, single amino acid variants in regions of the protein not near the active site can also alter protein stability and the overall catalytic rates. Several naturally occurring polymorphisms have been shown to alter substrate metabolism in CYP2A6 (1-21). These CYP2A6 polymorphisms can result in reduced or eliminated rates of substrate metabolism, likely due to indirect effects on the protein activity or its stability. Despite the amount of research available on CYP2A6 polymorphisms, very few studies were available on CYP2A13 polymorphisms. A study conducted by Wang and coworkers utilizing Flp-In in the Chinese hamster ovary cell line expressing CYP2A13 was one of the few studies exploring the effects of CYP2A13 variants (Table 8.1) (22). That study evaluated nine single amino acid variants and determined that there was not significant variation from wild type CYP2A13 for aflatoxin B_1 or NNK-induced cell toxicity except for the variant R101Q, which was identified as a null variant of CYP2A13 (22). The Scott and Murphy labs collaborated to study the functional consequences of six CYP2A13 polymorphisms (R101Q, D158E, R257C, V323L, F453Y, and R434C) using purified CYP2A13 enzyme and analysis of coumarin binding and the α-hydroxylation of the procarcinogens

Table 8.1. The known functional changes of CYP2A13 polymorphisms.

Allele	Protein Change	Functional Change
CYP2A13*2	A25N	Unknown
	R257C	Decreased coumarin 7-hydroxylation and NNK metabolism (24) No change in AFB1 and NNK toxicity (22)
CYP2A13*3	133_134 T insert.	Unknown
CYP2A13*4	R101Q	Inactive enzyme (22)
CYP2A13*5	F453Y	No change in AFB1 and NNK toxicity (22)
CYP2A13*6	R494C	No change in AFB1 and NNK toxicity (22)
CYP2A13*7	R101 X	Inactive enzyme (25)
CYP2A13*8	D158E	No change in AFB1 and NNK toxicity (22)
CYP2A13*9	V323L	No change in AFB1 and NNK toxicity (22)

4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) and N′-nitrosonornicotine (NNN) and 7-hydroxylation of coumarin (23).

Methods

Each of the CYP2A13 polymorphisms (R101Q, D158E, R257C, V323L, F453Y, R434C, and the T133_134 insertion) were designed, expressed, and purified in the truncated, His tagged form according to the procedures found in chapter 2. Active protein could not be obtained due to protein degradation during purification for the CYP2A13 R101Q and T133_134 insertion variants. The inclusion of protease inhibitors in the purification led to the isolation of very little CYP2A13 R101Q enzyme, and this protein was inactive according to the CO difference assay. The other five polymorphic CYP2A13 proteins were expressed at similar levels as CYP2A13 wild type enzyme and yielded active protein, as indicated by the CO difference assay after the two-column purification. Prior lab members Brian Smith and Jason Sanders were responsible for the purification of these variant proteins.

Results and Discussion

CYP2A13 Polymorphism Spectral Binding Affinity and Metabolism of Coumarin: CYP2A13 and the five variants all had very similar binding affinities for coumarin (Table 8.2). The K_D value for CYP2A13 wild type was determined to be 2.4 μM. None of the mutants had statistically different K_D values. The Murphy lab determined the kinetic parameters for coumarin 7-hydroxylation. None of the variants had a significant change in catalytic efficiency (Table 8.2).

Table 8.2. Coumarin binding and coumarin 7-hydroxylation by CYP2A13 and five variants.

	Binding [a]		**Coumarin 7-hydroxylation** [b]		
Enzyme	**K_D (µM)**	**A_{max}**	**K_m (µM)**	***K_{cat}* (pmol/min/pmol)**	***K_{cat}*/K_m**
2A13	2.4 ± 0.7	0.037 ± 0.009	8 ± 2	0.57 ± 0.03	0.08
R257C	2.4 ± 0.9	0.059 ± 0.005	16 ± 2[c]	0.52 ± 0.02	0.03
F453Y	2.5	0.071	9 ± 1	0.51 ± 0.02	0.06
R494C	2.1	0.054	13 ± 2[c]	0.66 ± 0.03[c]	0.05
D158E	2.4 ± 0.6	0.069 ± 0.003	6 ± 1	0.15 ± 0.006[c]	0.03
V323L	3 ± 2	0.059 ± 0.004	12 ± 1[c]	0.52 ± 0.02	0.04

[a] Mean K_D and A_{max} values are based on three independent determinations; single determinations are given for R453Y and R494C

[b] Kinetic parameters presented as weighted means ± SE determined from three independent experiments conducted in duplicate, (four independent experiments for D148E)

[c] Significance compared to CYP2A13 as determined by , $p < 0.05$

CYP2A13 Polymorphism Metabolism of NNN and NNK: The Murphy lab also determined kinetic parameters for NNN and NNK metabolism (Table 8.3). CYP2A13 catalyzes the 5′-hydroxylation of (S)-NNN to a lactol with a K_m of 110 ± 10 μM and a k_{cat} of 3.3 ± 0.1 min^{-1}. All of the mutant variants had catalytic efficiencies for NNN at approximately the CYP2A13 wild type level although there was some variation in the individual kinetic parameters. For NNK metabolism, each of the polymorphic variants had slightly increased K_m values compared to the wild type CYP2A13 enzyme. In addition, the D158E and V323L polymorphisms had decreased k_{cat} values. This led to three polymorphic variants with NNK catalytic efficiency decreased by more than 2-fold: D158E, R257C, and V323L.

Table 8.3. Hydroxylation of (S)-NNN and NNK by CYP2A13 and polymorphic variants.

Enzyme	(*S*)-NNN 5′-hydroxylation			NNK α-hydroxylation		
	K_m[a] (μM)	V_{max}[a] (pmol/min/pmol)	V_{max}/K_m	K_m[b] (μM)	V_{max}[b] (pmol/min/pmol)	V_{max}/K_m
2A13dH	110 ± 10	3.3 ± 0.1	0.03	110 ± 10	7.0 ± 0.6	0.07
R257C	67 ± 6[c]	1.88 ± 0.06[c]	0.03	250 ± 50[c]	6.5 ± 0.9	0.03
F453Y	69 ± 8[c]	1.57 ± 0.06[c]	0.02	160 ± 30	5.9 ± 0.6	0.04
R494C	100 ± 10	2.5 ± 0.1[c]	0.03	140 ± 30	5.6 ± 0.6	0.04
D158E	70 ± 10[c]	1.7 ± 0.1[c]	0.02	150 ± 30	2.2 ± 0.3[c]	0.02
V323L	85 ± 9	3.2 ± 0.1	0.04	170 ± 50	4.2 ± 0.6[c]	0.03

Conclusion

None of the polymorphic variants of CYP2A13 in this study had substantial effects on coumarin binding and metabolism, or on NNK or NNN metabolism that would likely lead to significant *in vitro* effects. There was a slight decrease in coumarin metabolism catalytic efficiency for CYP2A13 R257C and D158E. There was also a 2-fold decrease in NNK metabolism catalytic efficiency for CYP2A13 D158E, R257C, and V323L.

The small change in catalytic efficiency for coumarin, NNK, and NNN with the polymorphisms is not surprising considering their locations in the CYP2A13 enzyme. The only mutation near the active site was R101Q, which in the crystal structure interacts with a heme proprionate oxygen (Figure 8.1, panel A). Therefore, it is logical that the mutation of this residue could result in an unstable enzyme, as observed in our attempts to purify the enzyme. Each of the other mutations were located either below the heme plane or away from the active site in solvent exposed areas of the protein (Figure 8.1, panel B). The insertion of a threonine between residues 133 and 134 is into a helix located beneath the heme.

While it is possible that changes in amino acids away from the active site can still affect enzyme function, it appears that the majority of CYP2A13 polymorphisms have little affect on CYP2A13 metabolism. Of the mutations we studied, R101Q and the threonine 133_134 insertion are likely to have the greatest effect on NNK metabolism, and therefore possible cancer risk, because both resulted in inactive

protein. The reduction in inactive CYP2A13 protein would theoretically result in a decrease in the toxic metabolites produced from NNK.

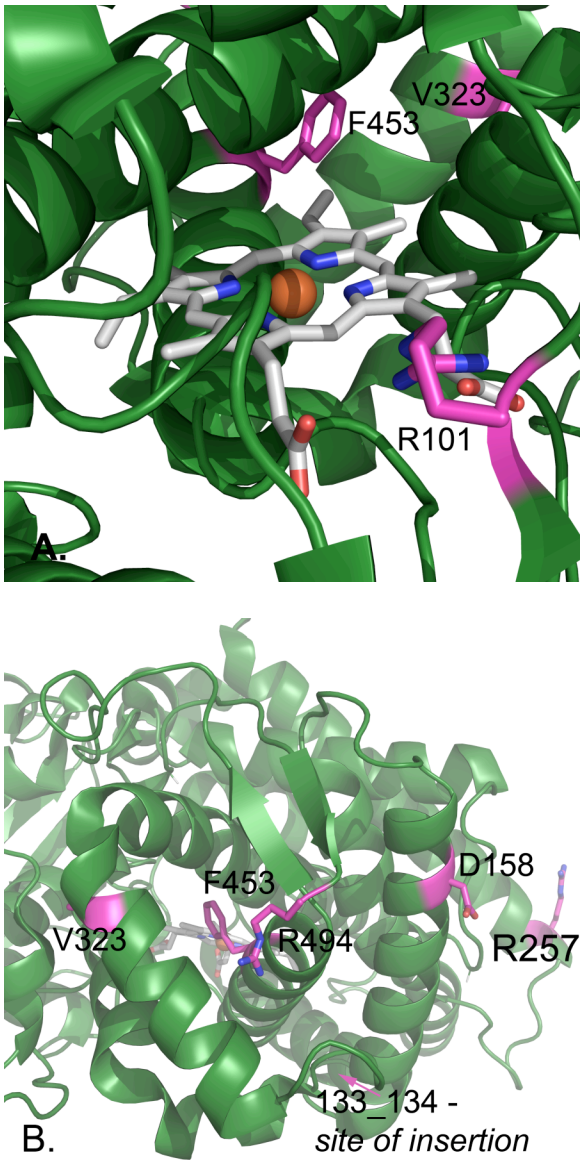


Figure 8.1. Cytochrome P450 2A13 (green, PDB 2P85) with the location of polymorphic variants with side chains shown and colored pink. A. Only one of the six polymorphic variants in this study was located within the active site cavity, this was the mutation of an arginine to a glutamine at position 101. B. The location of the other six polymorphisms are on the exterior part of CYP2A13.

References

1. Al Koudsi, N., Mwenifumbo, J. C., Sellers, E. M., Benowitz, N. L., Swan, G. E., and Tyndale, R. F. (2006) *Eur. J. Clin. Pharmacol.* **62**(6), 481-484
2. Ariyoshi, N., Sekine, H., Saito, K., and Kamataki, T. (2002) *Pharmacogenetics* **12**(6), 501-504
3. Brown, P. J., Bedard, L. L., Reid, K. R., Petsikas, D., and Massey, T. E. (2007) *Drug Metab. Dispos.* **35**(11), 2086-2094
4. Daigo, S., Takahashi, Y., Fujieda, M., Ariyoshi, N., Yamazaki, H., Koizumi, W., Tanabe, S., Saigenji, K., Nagayama, S., Ikeda, K., Nishioka, Y., and Kamataki, T. (2002) *Pharmacogenetics* **12**(4), 299-306
5. Ding, S., Lake, B. G., Friedberg, T., and Wolf, C. R. (1995) *Biochem. J.* **306 (Pt 1)**, 161-166
6. Fernandezsalguero, P., Hoffman, S. M. G., Cholerton, S., Mohrenweiser, H., Raunio, H., Rautio, A., Pelkonen, O., Huang, J. D., Evans, W. E., Idle, J. R., and Gonzalez, F. J. (1995) *Am. J. Hum. Genet.* **57**(3), 651-660
7. Fukami, T., Nakajima, M., Higashi, E., Yamanaka, H., McLeod, H. L., and Yokoi, T. (2005) *Biochem. Pharmacol.* **70**(5), 801-808
8. Fukami, T., Nakajima, M., Higashi, E., Yamanaka, H., Sakai, H., McLeod, H. L., and Yokoi, T. (2005) *Drug Metab. Dispos.* **33**(8), 1202-1210
9. Fukami, T., Nakajima, M., Yoshida, R., Tsuchiya, Y., Fujiki, Y., Katoh, M., McLeod, H. L., and Yokoi, T. (2004) *Clin. Pharmacol. Ther.* **76**(6), 519-527
10. Hadidi, H., Zahlsen, K., Idle, J. R., and Cholerton, S. (1997) *Food Chem. Toxicol.* **35**(9), 903-907
11. Kimura, M., Yamazaki, H., Fujieda, M., Kiyotani, K., Honda, G., Saruwatari, J., Nakagawa, K., Ishizaki, T., and Kamataki, T. (2005) *Drug Metab. Dispos.* **33**(9), 1361-1366
12. Kitagawa, K., Kunugita, N., Kitagawa, M., and Kawamoto, T. (2001) *J. Biol. Chem.* **276**(21), 17830-17835
13. Kiyotani, K., Yamazaki, H., Fujieda, M., Iwano, S., Matsumura, K., Satarug, S., Ujjin, P., Shimada, T., Guengerich, F. P., Parkinson, A., Honda, G., Nakagawa, K., Ishizaki, T., and Kamataki, T. (2003) *Pharmacogenetics* **13**(11), 689-695
14. Kubota, T., Nakajima-Taniguchi, C., Fukuda, T., Funamoto, M., Maeda, M., Tange, E., Ueki, R., Kawashima, K., Hara, H., Fujio, Y., and Azuma, J. (2006) *Pharmacogenomics J* **6**(2), 115-119
15. Malaiyandi, V., Sellers, E. M., and Tyndale, R. F. (2005) *Clin. Pharmacol. Ther.* **77**(3), 145-158
16. Nakajima, M. (2007) *Curr Opin Mol Ther* **9**(6), 538-544
17. Nakajima, M., Fukami, T., Yamanaka, H., Higashi, E., Sakai, H., Yoshida, R., Kwon, J. T., McLeod, H. L., and Yokoi, T. (2006) *Clin. Pharmacol. Ther.* **80**(3), 282-297

18. Oscarson, M., McLellan, R. A., Asp, V., Ledesma, M., Bernal Ruiz, M. L., Sinues, B., Rautio, A., and Ingelman-Sundberg, M. (2002) *Hum. Mutat.* **20**(4), 275-283
19. Oscarson, M., McLellan, R. A., Gullsten, H., Agundez, J. A., Benitez, J., Rautio, A., Raunio, H., Pelkonen, O., and Ingelman-Sundberg, M. (1999) *FEBS Lett.* **460**(2), 321-327
20. Wang, H., Tan, W., Hao, B., Miao, X., Zhou, G., He, F., and Lin, D. (2003) *Cancer Res.* **63**(22), 8057-8061
21. Xu, C., Rao, Y. S., Xu, B., Hoffmann, E., Jones, J., Sellers, E. M., and Tyndale, R. F. (2002) *Biochem. Biophys. Res. Commun.* **290**(1), 318-324
22. Wang, S. L., He, X. Y., Shen, J., Wang, J. S., and Hong, J. Y. (2006) *Toxicol. Sci.* **94**(1), 38-45
23. Schlicht, K. E., Michno, N., Smith, B. D., Scott, E. E., and Murphy, S. E. (2007) *Xenobiotica*, 1-11
24. Zhang, X., Su, T., Zhang, Q. Y., Gu, J., Caggana, M., Li, H., and Ding, X. (2002) *J. Pharmacol. Exp. Ther.* **302**(2), 416-423
25. Cauffiez, C., Pottier, N., Tournel, G., Lo-Guidice, J. M., Allorge, D., Chevalier, D., Migot-Nabias, F., Kenani, A., and Broly, F. (2005) *Xenobiotica* **35**(7), 661-669

Chapter 9.

Conclusion

The goal of this project was to determine the structural rationale for functional differences between the CYP2A13 and CYP2A6 enzymes that share 94%, sequence identity. Our hypothesis was that non-identical amino acids located in or near the active site of CYP2A13 and CYP2A6 most likely controlled functional differences. A number of techniques including site-directed mutagenesis, spectral binding assays, and metabolism assays were used to determine that four amino acids (208, 300, 301, and 369) largely mediate the substrate selectivity of phenacetin in the CYP2A enzymes. The determination of a crystal structure of the quadruple mutant 2A6 I208S/I300F/G301A/S369G verified that mutation of these four residues led to a mutant enzyme with an active site similar to CYP2A13 in terms of the orientations of both the mutated and conserved side chain and overall volume. Spectral binding assays of single and multiple CYP2A13 and CYP2A6 mutants with MAP, PEITC, and coumarin indicate that the residues are also likely to be important for selective binding of other ligands.

This project provided new information about the active sites of CYP2A6 and CYP2A13. We now know several residues that are key for differential ligand interaction and substrate metabolism. One of the long-term goals of the Scott lab is the design of an inhibitor specific for CYP2A13. The understanding of the cytochrome 2A family gained through this structure-function study will prove

valuable in our efforts to design an inhibitor for CYP2A13 that will have very low affinity for CYP2A6 and other cytochrome P450 enzymes.

We also examined some of the naturally occurring human CYP2A13 polymorphic variants in collaboration with the Murphy lab. Our results indicated that for coumarin binding and metabolism, and for NNN and NNK metabolism, CYP2A13 polymorphisms appear to have very little affect on substrate affinity or metabolism. This suggests that the single polymorphic variants studied to date likely contribute little to NNK-related cancer susceptibility. The two polymorphisms that could not be purified, the R101Q mutation and the insertion of a threonine between residues 133 and 134, may have an effect on an individual's susceptibility to cancer as these polymorphisms appear to lead to inactive CYP2A13.

Although this collection of research has provided valuable insight into the CYP2A family, it also generates new questions. We have shown that four residues primarily mediate selectivity for phenacetin; however, further work remains to determine if residues 208, 300, 301, and 369 also mediate catalytic selectivity of other substrates, particularly substrates of physiological importance like NNK. In addition, although we have studied CYP2A6 and CYP2A13, the closely related CYP2A7 is currently classified as an orphan cytochrome P450. With the knowledge of CYP2A enzymes gained through this study, it may be possible to determine how the CYP2A7 structure and function (or lack thereof) relates to CYP2A13 and CYP2A6.