

Review

Structural diversity of human xenobiotic-metabolizing cytochrome P450 monooxygenases[☆]

Eric F. Johnson^{a,*}, C. David Stout^b^a Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA 92037, USA^b Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

Received 25 August 2005

Available online 2 September 2005

Abstract

Cytochrome P450 monooxygenases provide important pathways for the metabolic clearance of drugs and toxins in humans. These enzymes are expressed from multiple genes and exhibit complex patterns of differential and overlapping substrate selectivity. Recent structures of microsomal P450s determined by X-ray crystallography have provided a structural basis for understanding differences in substrate recognition. This review will describe similarities and differences in the active site structures of four human microsomal cytochrome P450 monooxygenases, 2A6, 2C8, 2C9, and 3A4, that contribute extensively to drug and toxin metabolism.

© 2005 Elsevier Inc. All rights reserved.

Keywords: CYP2A6; CYP2C8; CYP2C9; CYP3A4; Structure; Cytochrome P450; Monooxygenase; Drug metabolism

The human cytochrome P450 (CYP) monooxygenases are a diverse group of enzymes encoded by 57 genes [1]. A number of these enzymes such as the sterol 14-demethylase, the cholesterol side-chain cleavage enzyme and the aromatase, catalyze relatively specific pathways in the biosynthesis of sterols and related molecules such as steroid hormones. The biosynthetic CYP monooxygenases are conserved among mammalian species and typically exhibit a high degree of substrate selectivity. However, the majority of CYP genes in mammalian genomes encode enzymes that oxidize structurally diverse substrates and provide an overlapping capacity to oxidize both foreign and endogenous compounds [2]. These substrates range from relatively small molecules such as ethanol to large antibiotics such as cyclosporine and erythromycin. Xenobiotic-metabolizing CYP enzymes play an important role in drug clearance and in the elimination of toxic compounds. However, they can also produce toxic metabolites that lead to adverse

consequences to chemical exposures. Additional toxicity issues can arise from elevated drug exposures due to genetic differences in CYP alleles that reduce drug clearance or resulting from CYP inhibition by other drugs when multiple drug therapies are used.

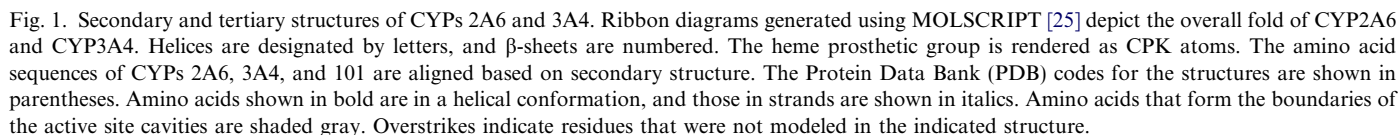
Recent advances in the modification, expression, and crystallization of microsomal CYPs have enabled the determination of structures for six mammalian cytochrome P450s by X-ray crystallography. These include structures for human CYPs 3A4 [3,4], 2C9 [5,6], 2C8 [7], and 2A6 [8] that together oxidize a major fraction of drugs used in clinical practice. On the occasion of the 50th anniversary of the discovery of monooxygenases, it is fitting to review the structural diversity of this important family of human monooxygenases. This brief overview will describe similarities and differences in the molecular structures of these enzymes.

Soluble, prokaryotic, and membranous, eukaryotic CYPs exhibit a conserved secondary structure and folding pattern [9] that is illustrated in Fig. 1 for two human CYPs, 3A4 and 2A6, which exhibit less than 40% amino acid identity. Helices are designated by letters beginning at the N-terminus of the catalytic domain as originally defined

[☆] Abbreviations: CYP, cytochrome P450; PDB, Protein Data Bank (<http://www.rcsb.org/pdb/>).

* Corresponding author. Fax: +1 858 784 7978.

E-mail address: johnson@scripps.edu (E.F. Johnson).



Eukaryotic P450s generally exhibit membrane targeting sequences that precede the catalytic domain shown in [Fig. 1](#). The majority of human P450s are targeted to the endoplasmic reticulum by a leader sequence that forms a transmembrane helix that anchors microsomal P450s to the cytoplasmic surface of the endoplasmic reticulum.

The polypeptide chains that form the catalytic domains of eukaryotic P450s are generally longer than those of prokaryotic P450s as illustrated by the sequence alignment of CYPs 3A4 and 2A6 with the prokaryotic camphor monooxygenase, CYP101, in Fig. 1. A large insertion is evident between helices F and G that typically exhibits one or two additional helices, F' and G'. These helices exhibit a

reverse amphipathicity with the hydrophobic surface of the helices exposed to the exterior of the protein. Experimental evidence suggests that these regions may provide a secondary membrane interaction site [11–14]. Additional sequence insertions are typically evident between helices J and K, between the end of strand 3 of β sheet 1 and helix L, and in the C-terminal region following helix L.

Although the overall structural organization of CYP2A6 and CYP3A4 is similar, significant differences are evident in the lengths of helices and loops as well as their placement, as would be expected from the low (<40%) amino acid sequence identity, Fig. 1. Generally, the most spatially conserved portions of P450 structures are helices E, I, J, K, and L as well as portions of β sheet 1 [9,15] that form the core of the protein. These regions maintain a conserved binding site for the heme prosthetic group that includes the conserved cysteine that forms the axial ligand to the heme iron. This structural conservation

is probably necessary to maintain the capacity of P450s to bind and reduce molecular oxygen to form the hypervalent heme iron oxo intermediate that oxidizes substrates.

Substrates bind in a cavity above the heme surface as depicted in Fig. 2. The heme forms the base of the active site cavity, and substrates must be positioned close to the reactive iron–oxo intermediate for oxidation. The regions that form the outer surfaces of the substrate binding cavity are generally more divergent between enzymes than other parts of the protein leading to differences in the sizes, shapes, and chemical features of the active sites that provide discrimination for different substrates. The outer surfaces of the active site cavity are formed by portions of β sheets 1 and 4, helices F–G, and the loop between helices B and C. The B–C loop may exhibit a helix, B', as seen in CYPs 2A6 and 2C8, or be less structured as seen for CYPs 3A4 and 2C9, Fig. 2. Differences in the lengths of helices and loops can be seen in the alignment of CYPs

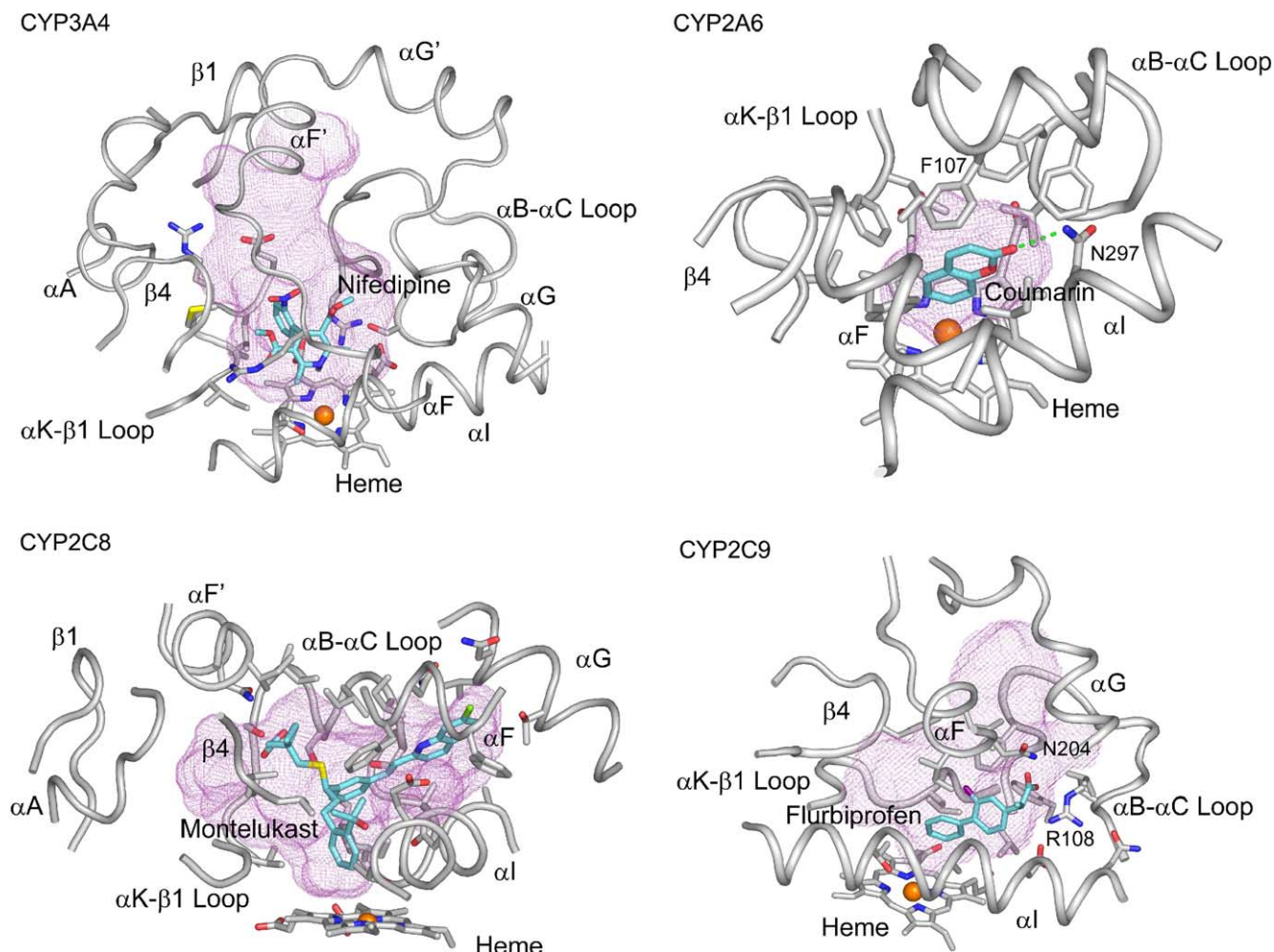


Fig. 2. Active site architectures of human CYPs 3A4, 2A6, 2C8, and 2C9. Cavity surfaces are rendered with a fine mesh calculated using VOIDOO [26] with a probe size of 1.4 Å. The surfaces were truncated at the ends of solvent channels in CYPs 2C8, 2C9, and 3A4 by placing dummy atoms at the exits to block the probe. Portions of each protein are rendered as gray ribbons with selected side-chains and substrates illustrated by stick figures. The heme prosthetic groups are depicted as stick figures with the iron shown as a sphere. Atoms are colored orange for iron, blue for nitrogen, red for oxygen, yellow for sulfur, green for chlorine, and purple for fluorine. Carbons are colored gray for the proteins and are colored cyan for nifedipine, coumarin, montelukast, and flurbiprofen. Molecular graphics were generated with PYMOL (<http://www.pymol.org>).

3A4 and 2A6 shown in Fig. 1. Notably, helices F and G are cantilevered over the top of the active site in the structures of the family 2 CYPs as well as in CYP 101, but the two helices do not extend across the top of the active site in the structure of CYP3A4, Fig. 1.

Amino acid side-chains that form the surface of the substrate binding cavities of CYPs 3A4, 2A6, and 101 are highlighted by gray shading in the structure based sequence alignment shown in Fig. 1. There is a good correspondence in the locations of these residues between CYP 101 and 2A6. These residues are localized to six regions termed Substrate Recognition Sites by Gotoh who correctly predicted this correspondence from a multiple sequence alignment of family 2 CYPs with CYP101, the first structurally characterized CYP [16]. The six regions correspond to the helix B–C loop, the C-terminal end of helix F, the N-terminal end of helix G, the middle portion of helix I, the loop between helix K and strand 3 of β sheet 1, and the turn in β sheet 4. However, this correspondence breaks down for CYP3A4 as well as other CYP enzymes with large active site cavities. The number of residues forming the active site surfaces of larger cavities is much greater, and they are distributed beyond the boundaries of the six regions defined based on the small active site architecture of CYP101 to include regions between helices F and G, the first turn in β sheet 1, and the region surrounding the N-terminal end of helix A.

CYP3A4 is the predominant enzyme in the clearance of about 50% of therapeutic drugs [2]. The active site cavity of CYP3A4 is relatively large contributing to its capacity to oxidize a variety of relatively large compounds, such as cyclosporine and erythromycin, which exceed the sizes of most drug-like molecules. However, CYP3A4 also oxidizes a number of smaller substrates, typified by nifedipine, midazolam, and testosterone that are often used as substrates in enzymic assays to characterize CYP3A4 expression. When the binding of nifedipine is modeled in the active site cavity of CYP3A4 in close proximity to the heme iron, a significant portion of the cavity remains unfilled by the substrate, and contacts between the substrate and enzyme are not as extensive as those seen for the other three enzymes in Fig. 2. Additionally, the unfilled portion of the cavity is sufficiently large to bind another substrate or effector molecule as suggested by kinetic [17,18] and fluorescence [19] studies. Multiple substrate/effector occupancy is thought to contribute to the homo- and hetero-tropic activation observed for the oxidation of testosterone, midazolam, nifedipine, and other substrates by CYP3A4.

The active site volume of CYP2A6 is roughly one-sixth that of CYP3A4. CYP2A6 is the principal nicotine oxidase in humans, and it also activates the tobacco derived carcinogens *N'*-nitrosonornicotine and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone to mutagenic products [2]. The relatively small size of typical substrates and the ability of these compounds to adopt almost planar conformations correspond well with the size and shape of the CYP2A6 active site cavity. This is illustrated in Fig. 2 for the substrate

coumarin, which was crystallized as a complex with CYP2A6 [8]. Coumarin binds to CYP2A6 with a relatively low apparent K_d , $<1 \mu\text{M}$. In contrast to nifedipine in the CYP3A4 active site, the coumarin molecule is highly constrained by the side chains of active site residues in CYP2A6, and water is completely displaced from the active site cavity when coumarin binds. With one exception, all of the residues contacting the coumarin molecule are hydrophobic. The upper surface of the cavity is defined by four large phenylalanine residues, and an edge-to-face interaction is apparent between coumarin and Phe107. Additionally, Asn297 hydrogen bonds to the carbonyl moiety of coumarin. These interactions favor an orientation of coumarin that positions carbon 7 close to the heme iron for efficient oxidation, and the enzyme displays a high degree of selectivity for the formation of 7-hydroxycoumarin, which is a marker reaction for CYP2A6 expression.

Family 2 CYPs account for roughly one-third of the CYP genes found in humans. These monooxygenases are divided into 13 subfamilies that exhibit roughly 70% or greater amino acid sequence identity across mammalian species. However, subfamilies exhibit only about 40–50% amino acid identity. Multiple enzymes are often found within CYP2 subfamilies in rats and mice that are encoded by gene clusters that reflect the independent duplication and divergence of genes within subfamilies during the radiation of mammalian species. In humans, many of the genes in these clusters are pseudogenes, and only the 2A and 2C clusters exhibit more than one functional gene [1].

The CYP2A6 gene is found in a gene cluster that contains two closely related genes for CYPs 2A7 and 2A13 [20]. These proteins exhibit greater than 90% amino acid sequence identity with each other. The catalytic activity of P450 2A7 remains uncharacterized, and expression of the enzyme as a functional protein has not been successful. CYPs 2A6 and 2A13 exhibit an overlapping substrate selectivity with distinct quantitative differences in kinetic parameters for shared substrates [2]. CYP2A13 exhibits 30 amino acid differences relative to P450 2A6 out of 494 residues. Several of these differences, V117A, I300F, G301A, and I366L, directly modify the active site cavity by altering 4 of the 12 amino acid side chains that directly contact the substrate in CYP2A6. Five additional substitutions, V110L, R208I, V365M, S369G, and H372R occur in close proximity to the active site and may alter the active site cavity indirectly.

The human CYP2C gene cluster encodes CYPs 2C8, 2C9, 2C18, and 2C19, and with the exception of 2C18, these enzymes contribute extensively to drug metabolism. The amino acid sequence identities of CYPs 2C9 and 2C19 are relatively high ($>90\%$), whereas CYP2C8 is more divergent and exhibits roughly 74% sequence identity with either 2C9 or 2C19. The three enzymes exhibit distinct substrate selectivity profiles [21,22].

The active site cavity of CYP2C8 is similar in size to that of CYP3A4, and the enzyme oxidizes a number of large substrates such as taxol and various statins that are also

oxidized by CYP3A4 [21]. However, the shape of the CYP2C8 substrate binding cavity is distinctly different from that of CYP3A4, Fig. 2. The cavity of CYP2C8 is less spacious close to the heme iron. As the cavity expands outward from the heme iron, it diverges in two distinct directions. The high affinity inhibitor, montelukast [23] complements this shape as is evident from a recently determined structure of CYP2C8 complexed with montelukast (unpublished), Fig. 2. The carboxylate moiety of montelukast interacts with the backbone of helix B–B' loop and occupies one branch of the Y-shaped cavity. This region contains a number of polar residues, is well solvated, and is connected by a solvent channel to the exterior of the protein making it well suited for polar moieties. The other, longer branch of the “Y” is more hydrophobic as is the region near the heme iron. Montelukast makes extensive hydrophobic contacts with these regions of the protein. In contrast, the active site cavity of CYP3A4 is “bowl-like” near the heme iron with a wide channel to the exterior of the protein.

The active site cavity of 2C9 is larger than that of CYP2A6 but smaller than that of CYP2C8. The structure of CYP2C9 complexed with flurbiprofen (PDB: 1R9O) has been determined [6], and it exhibits several differences from the CYP2C8 structure. These include a more extended conformation of the regions between helices F and G, and between helices B and C. The tubular, substrate binding cavity of CYP2C9 exhibits open solvent channels to the exterior of the protein at each end. One channel exits over helix I between helix F and β -sheet 4. The other channel runs between the turn in β -sheet 4 and the helix B–C loop, expands, and finally exits between helix G and the helix B–C loop, Fig. 2. CYP2C9 is the predominant enzyme in the oxidation of *S*-warfarin and phenytoin, two drugs with low margins of safety. CYP2C9 allelic variation in humans underlies the slow clearance of these drugs in some individuals and is associated with a higher incidence of adverse reactions. Additionally, CYP2C9 efficiently oxidizes several lipophilic anions such as the non-steroidal anti-inflammatory drugs, naproxen, ibuprofen, flurbiprofen, indomethacin, and similar compounds [22]. The active site cavity of CYP2C9 is relatively large compared to the size of flurbiprofen and other non-steroidal anti-inflammatory drugs. Flurbiprofen occupies a corner of the cavity where the carboxylate moiety of the substrate exhibits a favorable charge–charge interaction with Arg108 [6]. Occupation of this site by the substrate is also favored by hydrogen bonding interactions with Asn204 as well as interactions between hydrophobic side-chains and the hydrophobic portion of the substrate. Charge stabilization by R108 probably allows the enzyme to sequester other small lipophilic anions in the larger active site cavity and position them sufficiently close to the heme iron for efficient oxidation. The portion of the active site cavity extending over the surface of the heme and under helix F also allows the enzyme to accommodate larger substrates and inhibitor molecules such as benzbromarone

and troglitazone. The additional space may also accommodate another substrate or effector molecule as suggested by kinetic models for the activation of flurbiprofen oxidation by dapsone [24].

Structures of a mutated form of CYP2C9 with and without warfarin bound in the active site (PDB 1OG2 and 1OG5) have also been determined [5]. The warfarin is bound in a distal portion of active site cavity from the heme iron (not shown), and the active site cavity of the mutant CYP2C9 catalytic domain is much larger than that seen for the structure of the flurbiprofen complex with the native catalytic domain (PDB 1R9O). The structures of the mutant CYP2C9 more closely resemble those of CYP2C8, and the mutations alter the sequence of the divergent helix F to helix G region of CYP2C9 to more closely resemble that of CYP2C8 and other CYP2C enzymes.

In summary, each enzyme exhibits an active site architecture that differs in shape, size, and the chemical characteristics of the active site residues. These properties of the active sites fit well with the profiles of substrates that are oxidized by each enzyme, and the structures should facilitate the use of *in silico* approaches to predict substrate binding. Structures of additional complexes can be expected to better map the nature of adaptive changes that occur when substrates bind and further facilitate structure based predications of substrate binding. However, the four structures described here as well as those determined for rabbit CYPs 2B4 and 2C5 represent only a small fraction of the total mammalian complement of >50 CYPs. Although the available structures provide a basis for generating homology models for some of these enzymes, the differences exhibited by closely related enzymes such as CYP2C8 and CYP2C9 suggest that it may be difficult to model the substrate binding cavities of other enzymes with a high degree of certainty. Additional, experimentally determined structures should better delineate differences between CYP families and subfamilies.

Acknowledgments

The authors' studies of human drug-metabolizing P450s are supported by NIH Grant GM031001 and by Pfizer Global Research and Development.

References

- [1] D.R. Nelson, D.C. Zeldin, S.M. Hoffman, L.J. Maltais, H.M. Wain, D.W. Nebert, Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice variants, *Pharmacogenetics* 14 (2004) 1–18.
- [2] F.P. Guengerich, Human cytochrome P450 enzymes, in: P.R. Ortiz de Montellano (Ed.), *Cytochrome P450: Structure mechanism and biochemistry*, Kluwer Academic/Plenum Publishers, New York, 2005, pp. 377–530. This is a recent, comprehensive review of the human CYPs with an extensive bibliography.
- [3] J.K. Yano, M.R. Wester, G.A. Schoch, K.J. Griffin, C.D. Stout, E.F. Johnson, The structure of human microsomal cytochrome P450 3A4

- determined by X-ray crystallography to 2.05-Å resolution, *J. Biol. Chem.* 279 (2004) 38091–38094.
- [4] P.A. Williams, J. Cosme, D.M. Vinkovic, A. Ward, H.C. Angove, P.J. Day, C. Vonnrhein, I.J. Tickle, H. Jhoti, Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone, *Science* 305 (2004) 683–686.
 - [5] P.A. Williams, J. Cosme, A. Ward, H.C. Angove, V.D. Matak, H. Jhoti, Crystal structure of human cytochrome P450 2C9 with bound warfarin, *Nature* 424 (2003) 464–468.
 - [6] M.R. Wester, J.K. Yano, G.A. Schoch, C. Yang, K.J. Griffin, C.D. Stout, E.F. Johnson, The structure of human microsomal cytochrome P450 2C9 complexed with flurbiprofen at 2.0 Å resolution, *J. Biol. Chem.* 279 (2004) 35630–35637.
 - [7] G.A. Schoch, J.K. Yano, M.R. Wester, K.J. Griffin, C.D. Stout, E.F. Johnson, Structure of human microsomal cytochrome P450 2C8. Evidence for a peripheral fatty acid binding site, *J. Biol. Chem.* 279 (2004) 9497–9503.
 - [8] J.K. Yano, M.H. Hsu, K.J. Griffin, C.D. Stout, E.F. Johnson, Structures of human microsomal cytochrome P450 2A6 complexed with coumarin and methoxsalen, *Nat. Struct. Mol. Biol.* (2005). Available online ahead of the printed version.
 - [9] T.L. Poulos, E.F. Johnson, Structures of cytochrome P450 enzymes, in: P.R. Ortiz de Montellano (Ed.), *Cytochrome P450: Structure, Mechanism, and Biochemistry*, third ed., Kluwer Academic/Plenum Publishers, New York, 2005, pp. 87–114.
 - [10] T.L. Poulos, B.C. Finzel, I.C. Gunsalus, G.C. Wagner, J. Kraut, 2.6 Å crystal structure of the *Pseudomonas putida* cytochrome P-450, *J. Biol. Chem.* 260 (1985) 16122–16130.
 - [11] P.A. Williams, J. Cosme, V. Sridhar, E.F. Johnson, D.E. McRee, The crystallographic structure of a mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity, *Mol. Cell* 5 (2000) 121–132.
 - [12] S. Izumi, H. Kaneko, T. Yamazaki, T. Hirata, S. Kominami, Membrane topology of guinea pig cytochrome P450 17 α revealed by a combination of chemical modifications and mass spectrometry, *Biochemistry* 42 (2003) 14663–14669.
 - [13] H. Deng, J. Wu, S.P. So, K.H. Ruan, Identification of the residues in the helix F/G loop important to catalytic function of membrane-bound prostacyclin synthase, *Biochemistry* 42 (2003) 5609–5617.
 - [14] M.J. Headlam, M.C. Wilce, R.C. Tuckey, The F-G loop region of cytochrome P450_{scc} (CYP11A1) interacts with the phospholipid membrane, *Biochim. Biophys. Acta* 1617 (2003) 96–108.
 - [15] J. Mestres, Structure conservation in cytochromes P450, *Proteins* 58 (2005) 596–609.
 - [16] O. Gotoh, Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences, *J. Biol. Chem.* 267 (1992) 83–90.
 - [17] K.R. Korzekwa, N. Krishnamachary, M. Shou, A. Ogai, R.A. Parise, A.E. Rettie, F.J. Gonzalez, T.S. Tracy, Evaluation of atypical cytochrome P450 kinetics with two-substrate models: evidence that multiple substrates can simultaneously bind to cytochrome P450 active sites, *Biochemistry* 37 (1998) 4137–4147.
 - [18] W.M. Atkins, Non-Michaelis-Menten kinetics in cytochrome P450-catalyzed reactions, *Annu. Rev. Pharmacol. Toxicol.* 45 (2005) 291–310.
 - [19] M.J. Dabrowski, M.L. Schrag, L.C. Wienkers, W.M. Atkins, Pyrene-pyrene complexes at the active site of cytochrome P450 3A4: evidence for a multiple substrate binding site, *J. Am. Chem. Soc.* 124 (2002) 11866–11867.
 - [20] S.M. Hoffman, D.R. Nelson, D.S. Keeney, Organization, structure and evolution of the CYP2 gene cluster on human chromosome 19, *Pharmacogenetics* 11 (2001) 687–698.
 - [21] R.A. Totah, A.E. Rettie, Cytochrome P450 2C8: substrates, inhibitors, pharmacogenetics, and clinical relevance, *Clin. Pharmacol. Ther.* 77 (2005) 341–352.
 - [22] A.E. Rettie, J.P. Jones, Clinical and toxicological relevance of CYP2C9: drug-drug interactions and pharmacogenetics, *Annu. Rev. Pharmacol. Toxicol.* 45 (2005) 477–494.
 - [23] R.L. Walsky, R.S. Obach, E.A. Gaman, J.P. Gleesen, W.R. Proctor, Selective inhibition of human cytochrome P450 2C8 by montelukast, *Drug Metab. Dispos.* 33 (2005) 413–418.
 - [24] J.M. Hutzler, M.J. Hauer, T.S. Tracy, Dapsone activation of CYP2C9-mediated metabolism: evidence for activation of multiple substrates and a two-site model, *Drug Metab. Dispos.* 29 (2001) 1029–1034.
 - [25] P.J. Kraulis, MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures, *J. Appl. Cryst.* 24 (1991) 946–950.
 - [26] G.J. Kleywert, T.A. Jones, Detection, delineation, measurement and display of cavities in macromolecular structures, *Acta Cryst. D* 50 (1994) 178–185.