Journal of Computer-Aided Molecular Design, 11 (1997) 265–272.

J-CAMD 390

Molecular modeling of cytochrome P450 3A4

Grazyna D. Szklarz* and James R. Halpert

Department of Pharmacology and Toxicology, College of Pharmacy, University of Arizona, Tucson, AZ 85721, U.S.A.

Received 10 September 1996
Accepted 13 January 1997

Keywords: Cytochrome P450 3A4; Molecular modeling; Substrate docking; Structure–function relationships; Human drug metabolism

Summary

The three-dimensional structure of human cytochrome P450 3A4 was modeled based on crystallographic coordinates of four bacterial P450s: P450 BM-3, P450cam, P450terp, and P450eryF. The P450 3A4 sequence was aligned to those of the known proteins using a structure-based alignment of P450 BM-3, P450cam, P450terp, and P450eryF. The coordinates of the model were then calculated using a consensus strategy, and the final structure was optimized in the presence of water. The P450 3A4 model resembles P450 BM-3 the most, but the B' helix is similar to that of P450eryF, which leads to an enlarged active site when compared with P450 BM-3, P450cam, and P450terp. The 3A4 residues equivalent to known substrate contact residues of the bacterial proteins and key residues of rat P450 2B1 are located in the active site or the substrate access channel. Docking of progesterone into the P450 3A4 model demonstrated that the substrate bound in a 6β-orientation can interact with a number of active site residues, such as 114, 119, 301, 304, 305, 309, 370, 373, and 479, through hydrophobic interactions. The active site of the enzyme can also accommodate erythromycin, which, in addition to the residues listed for progesterone, also contacts residues 101, 104, 105, 214, 215, 217, 218, 374, and 478. The majority of 3A4 residues which interact with progesterone and/or erythromycin possess their equivalents in key residues of P450 2B enzymes, except for residues 297, 480 and 482, which do not contact either substrate in P450 3A4. The results from docking of progesterone and erythromycin into the enzyme model make it possible to pinpoint residues which may be important for 3A4 function and to target them for site-directed mutagenesis.

Introduction

Cytochromes P450 3A are important mammalian enzymes that biotransform a vast array of clinically, physiologically, and toxicologically important compounds. The catalytic versatility and high capacity of these enzymes suggests that they are an ideal clearance pathway for drugs. Human P450 3A4 is a major hepatic and intestinal enzyme responsible for monooxygenation of drugs as well as carcinogens [1]. The clinical significance of this enzyme stems from both its ability to metabolize a large number of therapeutic agents and its high expression level in the liver of most humans. Moreover, intestinal 3A4 expression accounts for significant 'first pass' metabolism of orally administered medications [2].

Like the majority of P450s, 3A4 exhibits broad substrate specificity and is able to metabolize a variety of structurally unrelated compounds. Examples include macrolide antibiotics such as erythromycin [3] and troleandomycin, cyclosporin A [4], nifedipine [5], and the steroids testosterone, androstenedione and progesterone [6]. Interestingly, the activity of P450 3A4 is affected by flavonoids, e.g. α-naphthoflavone (α-NF), which is the most effective activator of this enzyme [7]. The susceptibility to modulation by flavonoids is a distinctive property of P450 3A enzymes, which can be activated or inhibited by various flavonoid compounds.

Despite the wealth of information on the regulation and substrate specificity of cytochrome P450 3A4, structure–function analysis has not been rigorously approached. There are, for example, no site-directed mutagenesis data that would pinpoint key substrate contact residues. Due to the importance of P450 3A4 for human drug metabolism, it is of great interest to understand how this enzyme functions, in particular the nature and structural requirements of its substrate binding site. These studies

*To whom correspondence should be addressed.

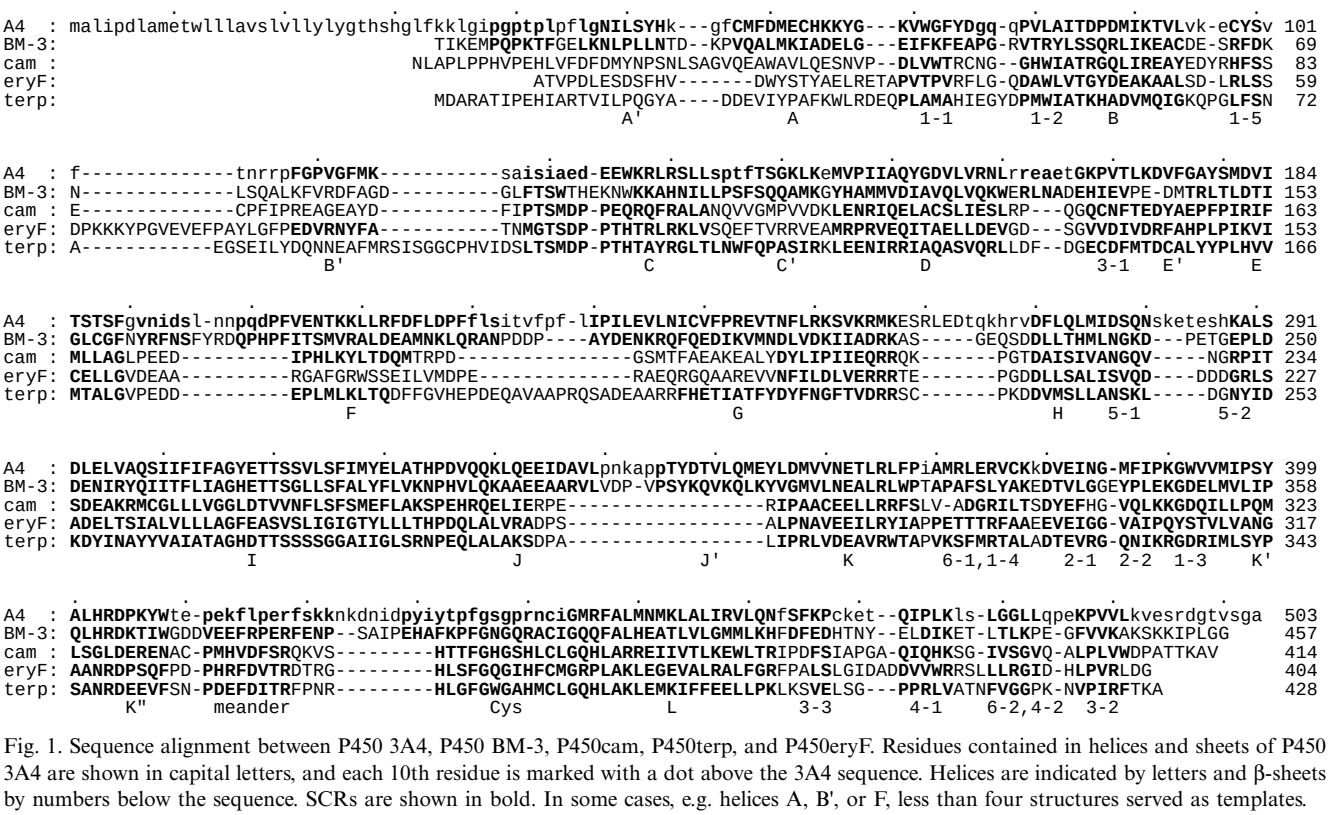

Fig. 1. Sequence alignment between P450 3A4, P450 BM-3, P450cam, P450terp, and P450eryF. Residues contained in helices and sheets of P450 3A4 are shown in capital letters, and each 10th residue is marked with a dot above the 3A4 sequence. Helices are indicated by letters and β-sheets by numbers below the sequence. SCRs are shown in bold. In some cases, e.g. helices A, B', or F, less than four structures served as templates.

will be facilitated by information about the three-dimensional (3D) structure, which should allow a detailed examination of enzyme–substrate interactions and help to predict the possible metabolic fate of drugs or carcinogens.

Although none of the mammalian P450s have been crystallized so far, four bacterial P450 crystal structures have been solved. These include P450cam (P450 101), a camphor hydroxylase from *Pseudomonas putida* [8,9], P450 BM-3 (P450 102), a fatty acid monooxygenase from *Bacillus megaterium* [10], P450terp (P450 108), an α-terpineol monooxygenase from *Pseudomonas sp*. [11], and P450eryF, a 6-deoxyerythronolide B hydroxylase from *Saccaropolyspora erythrea* [12]. Although the bacterial enzymes display low (15–20%) amino acid sequence identity with eukaryotic P450s, the regions of the sequence containing the heme binding site, the oxygen binding site, and the site of interactions with redox partners are highly conserved in all cytochromes P450 [13,14]. Moreover, a comparison of known structures indicates that the topology of all the enzymes is similar, especially in the heme-binding core region [12,15].

Based on the premise of structural homology among cytochromes P450, molecular models of various mammalian enzymes have been constructed. The initial models, based on the structure of P450cam, include P450 1A1 [16], 2B1 [17], 2D6 [18], 19A1 (aromatase) [19], and P450scc [20]. More recent models were based on the structure of P450 BM-3 [21], or those of the three enzymes, P450cam, P450 BM-3, and P450terp. The latter approach was utilized to model aromatase [22] and P450 2B1 [23]. In the case of P450 3A4, only its active site geometry was modeled based on the structure of P450cam [24,25]. In view of availability of additional P450 crystal structures, a new model for this enzyme seems to be in order. Such a model should include structural information from all known enzymes and lead to a better accuracy of the model structure.

Therefore, in the present investigation, we constructed a model of P450 3A4, which was built based on four known crystal structures, P450cam, P450 BM-3, P450terp, and P450eryF, using consensus modeling methods. The model can be used to explain substrate specificity and to relate enzyme function to its structure. For example, the active site of the P450 3A4 model can accommodate two known substrates, progesterone and the much larger erythromycin.

Methods

Molecular modeling – General

The crystallographic coordinates of P450cam were obtained from the Brookhaven Protein Databank [9], those of P450 BM-3 [10] and P450terp [11] from Dr. Julian A. Peterson (University of Texas Southwestern Medical Center at Dallas, TX, U.S.A.) and those of P450eryF [12] from Dr. Thomas L. Poulos (Department

of Molecular Biology and Biochemistry, University of California, Irvine, CA, U.S.A.). The sequence of P450 3A4 was from SwissProt database (accession number P08684) [26]. The P450 3A4 structure was modeled using INSIGHTII/CONSENSUS software (Biosym/MSI, San Diego, CA, U.S.A.) on a Silicon Graphics workstation. Energy minimization and molecular dynamics calculations were performed using the DISCOVER program (Biosym/MSI) with the consistent valence force field. The parameters for heme and ferryl oxygen were as described by Paulsen and Ornstein [27,28].

Sequence alignment

To align the P450 3A4 sequence to those of P450 BM-3, P450cam, P450terp, and P450eryF, we utilized the structure-based alignment proposed by Hasemann et al. [15], which aligns the sequences of the first three enzymes. The sequence of P450eryF was aligned to the others based on the alignment between P450cam and P450eryF [12]. This multiple alignment was then refined manually with HOMOLOGY (Biosym/MSI), and the structurally conserved regions (SCRs) of the proteins were determined based on rms deviation of the backbone for each pair of structures. The sequence of P450 3A4 was aligned to those of the reference proteins using the results of secondary structure prediction with the PHD method [29,30] from PredictProtein server [31]. Structurally conserved regions of 3A4 were then determined based on scores of Dayhoff's mutation matrix and the Engleman and Steitz hydrophobicity index for each 3A4–reference protein pair (i.e. 3A4–P450 BM-3, 3A4–P450cam, 3A4–P450eryF, and 3A4–P450terp). For some regions, less than four proteins served as templates.

Modeling P450 3A4 structure

After the SCRs were determined, coordinates of the model were calculated using the CONSENSUS program, which employs distance geometry calculations, as described previously for P450 2B1 [23]. The obtained structure was then verified with the PROFILES-3D program (Biosym/MSI), which measures the compatibility between the protein sequence profile and its 3D profile [32,33], and with the ProStat utility of the HOMOLOGY program, which checks bond lengths and angles.

Fig. 2. Ribbon representation of the P450 3A4 model. Heme is shown in red, and the residues corresponding to those of substrate contact residues in bacterial proteins and key residues of P450 2B1 are shown in purple. All these residues are listed in Table 1.

Based on the verification results, the model was further refined using the BIOPOLYMER program (Biosym/MSI). Modifications were required mainly in the regions of junctions between SCRs and loops. These regions were modified manually using torsions on the protein backbone. To avoid changes to the rest of the model structure, the C-terminal peptide bond of such a region was broken before torsions were initiated, and, after torsions were cleared, the atoms were bonded again. Subsequently, the conformation of the side chains within the modified segment was corrected using rotamers. Also included were side chains of the residues outside that segment which overlapped with the residues of the segment modified. In addition to the refinement of the misfolded regions, the N-terminus was modeled as a helix. Modified regions were then minimized, side chains first, then all atoms, using the steepest descent method and harmonic potential to a maximum gradient of 5 kcal mol^{-1} $Å^{-1}$. This was followed by conjugate gradient minimization of all atoms to a maximum of 1 kcal mol^{-1} $Å^{-1}$. The non-bond cutoff was 8 Å and the dielectric constant was 1.0. During the refinement procedure, the structure was checked periodically using PROFILES-3D and ProStat, as described above.

The most difficult to model was helix B' and the preceding loop (β1-5–B' loop), since this region is the most variable among the four crystal structures. In the initial model, helix B' was modeled after that of P450 BM-3. However, erythromycin, a substrate of 3A4, could not be docked into the active site, indicating that BM-3 was not a good template for this region. In fact, this substrate did not fit into an active site of other crystallized P450s, except for P450eryF. Therefore, the refinement of the model included modification of helix B' to resemble that of P450eryF. The β1-5–B' loop was then constructed to link the helix with β-sheet 1-5. This loop was optimized with one cycle of combined molecular dynamics and minimization, as described previously [17]. In the last step, heme was added to the P450 3A4 model, and amino acids neighboring heme residue were minimized as described above.

The final refinement of the model involved minimization of the whole structure in the presence of water. Water molecules were from soaking the protein using a sphere of 25 Å and a layer of 3 Å. The minimization was thus performed on a protein–water association, with heme and Cys^{442} fixed, using at first the steepest descent method until the gradient was less than 5 kcal mol^{-1} $Å^{-1}$, and then conjugate gradients to a maximum of 1 kcal mol^{-1} $Å^{-1}$ [17].

Docking of the substrate into the active site

Two substrates of P450 3A4 were docked into the enzyme model, progesterone and erythromycin. The progesterone structure was obtained previously [17,23]. Erythromycin was obtained by modifying the crystal structure of 6-*O*-methylerythromycin A [34] available from the Cambridge Structural Database. In the case of progesterone, the substrate was placed in a 6β-reactive binding orientation with the oxidation site fixed at 5.6–6 Å from the heme iron, and the C_6-$H_{6\beta}$ bond aligned with ferryl oxygen, heme iron and sulfur of Cys^{442}, as described earlier [17,23]. This results in a hydrogen-bonding distance between ferryl oxygen and the hydrogen atom to be abstracted from the substrate. For erythromycin, the N-atom was placed at a distance of 4.7 Å from the heme iron and, as for progesterone, aligned with Fe and S of Cys^{442}. This distance allows for van der Waals contacts between ferryl oxygen and the nitrogen of the substrate, and thus enables erythromycin N-demethylation. Docking of this substrate was also used to evaluate the initial model of P450 3A4, as described in the preceding section.

The substrate was docked into the active site of the 3A4 model using molecular dynamics. For these simulations, certain atoms of the substrate were fixed, while the rest of the molecule, along with the side chains of protein residues within 5 Å from the substrate, was allowed to move. In the case of progesterone the atoms fixed were C_6 and $H_{6\beta}$, while for erythromycin these were the nitrogen atom and the carbon atoms directly bonded to it. In each case, the system was first minimized using the steepest descent method and harmonic potential, with a non-bond cutoff of 10 Å, to a maximum gradient of 5 kcal mol^{-1} $Å^{-1}$. For the subsequent molecular dynamics simulations, the leap-frog algorithm was used, the system was equilibrated for 0.1 ps, and the simulations were continued for 1 ps at 300 K using 1 fs time steps. The system was then minimized again using conjugate gradients to a maximum gradient of 1 kcal mol^{-1} $Å^{-1}$, first the substrate molecule only, and then the side chains of protein residues neighboring the substrate. The non-bond interaction energy between the substrate, progesterone or erythromycin, and the protein, both electrostatic and van der Waals forces, was evaluated with the DOCKING module of the INSIGHTII package.

Results and Discussion

Modeling mammalian cytochromes P450 has often been applied to study structure–function relationships of these enzymes, especially in conjunction with site-directed mutagenesis experiments. The validity of this approach was demonstrated for aromatase [35] and P450 2A4 [36, 37], as well as for P450 2B1 [17,23]. At present, the crystal structures of four bacterial P450s are known, and, although they share only about 20% sequence identity, the enzymes display a similar overall topography [10–12, 15]. This is in agreement with a classical 'rule' which predicts that the tertiary structure of the protein is conserved much better than its sequence. Accordingly, the comparison of known P450 structures indicates that all

P450s share a highly conserved 'core structure' [12,15,38]. Moreover, based on sequence alignments, the overall P450 fold can be extended to other cytochromes P450, including mammalian enzymes. Thus, in accordance with the above premises, we have built a model of P450 3A4, an important human drug-metabolizing enzyme*.

Modeling of P450 3A4 and general features of the model

The sequence alignment used for modeling cytochrome P450 3A4 is shown in Fig. 1. This alignment is similar to that proposed by Hasemann et al. [15], but with minor differences. These include the alignment of helix B' and the B'-C loop, helix F and β-sheets 6-2, 4-2 and 3-2 at the C-terminal end of the protein. These changes were introduced based on the alignment scores and the structural requirements of model building. The P450 3A4 model is shown in Fig. 2.

The consensus strategy applied to model P450 3A4 makes it possible to use one or more reference proteins for modeling SCRs, and the coordinates of the model are weighted averages of the coordinates of the reference proteins. Thus, in the P450 'core region', usually all reference proteins were used as templates. The core structure includes very highly conserved elements in the neighborhood of heme, such as the central and the C-terminal end of helix I, β-sheet 6-1, the Cys-pocket, helices L and J, and somewhat less conserved β-sheets 1-1, 1-2, 1-3 and 1-4, the C, E and K helices, β-sheets 2-2, 6-2, and the meander region [15]. Other regions of P450s are less structurally conserved, and the most flexible regions are represented by the amino termini and helices A, B', F, and G, which are involved in substrate recruitment and binding [15,35]. To model these regions of P450 3A4, often less than four reference P450s were used as templates, based on mutation matrix and hydrophobicity scores for a 3A4–bacterial protein alignment. For example, helix A of P450 3A4 was modeled based solely on the coordinates of P450 BM-3, while helix B' was based on the coordinates of P450eryF. The latter region is characterized by high structural variability among known P450s, and the 3A4 modeling was aided by docking of its substrate, erythromycin, into the active site of the initial model. P450eryF is the only enzyme of known structure able to accommodate large substrates, such as 6-deoxyerythronolide B, due to the rotation of helix B' leading to the enlargement of the active site compared with that of P450cam, P450 BM-3 or P450terp [12]. In the case of helix F, its N-terminus in P450 3A4 was modeled based on the coordinates of P450 BM-3, P450cam, and P450terp, while the C-terminus was based solely on the coordinates of P450 BM-3. Conversely, the N-terminal end of helix G was modeled based on the coordinates of P450 BM-3, while the C-terminus was based on the coordinates of all four P450s. However, at the C-terminus, helix G in P450 3A4 was extended one turn as compared to other P450s, in agreement with the secondary structure prediction for this region.

TABLE 1
COMPARISON OF ACTIVE SITE RESIDUES OF P450 3A4 WITH THOSE OF P450 2B1 AND CRYSTALLIZED P450 ENZYMES ACCORDING TO THE ALIGNMENT IN FIG. 1

3A4	2B1[a]	BM-3	cam	eryF	terp
Arg[105]	(A102)	Q73	F87	G77	I78
Met[114]	Y111	A82	Y96	F86	F87
Ser[119]	I114	T88	T101	G91	T103
Asp[214]	F206	A184	M184	I174	F187
Phe[215]	(S207)	M185	T185	L175	F188
Asp[217]	L209	K187	P187	M177	V190
Ala[297]	I290	Y256	R240	S233	A259
Ile[301]	S294	T260	L244	V237	A263
Phe[304]	(F297)	I263	V247	L240	T266
Ala[305]	(A298)	A264	G248	A241	A267
Thr[309]	T302	T268	T252	A245	T271
Ala[370 b]	V363	A328	V295	P288	V314
Leu[373 b]	V367	F331	D297	T291	F317
Leu[479]	I477	L437	I395	L391	F414
Gly[480]	G478	T438	V396	L392	V415
Leu[482]	I480	K440	G398	G394	G417

[a] All residues except those in parentheses were studied experimentally in 2B enzymes, such as 2B1, 2B4, 2B5, or 2B11.
[b] Verified by site-directed mutagenesis (Y.-A. He, unpublished observations).

P450 BM-3 belongs to class II P450s which require an FAD/FMN-containing NADPH-P450 reductase for catalysis, as does P450 3A4, in contrast to class I enzymes that require both an iron–sulfur protein (ferredoxin) and an FAD-containing NAD(P)H-ferredoxin reductase. The latter group includes soluble bacterial enzymes P450cam, P450terp, and P450eryF. The proximal face of P450 BM-3 contains a proposed redox partner binding region, and its bowl-like shape results from the presence of helix J' and the C-terminally extended meander, in contrast to other P450 structures [10,15,38]. In the P450 3A4 model, this region resembles that of BM-3 with respect to both the 3D structure and the presence of charged residues. In structurally known P450s, helices B, C, J, J', K, L, the C-terminal part of the meander and the heme-binding region are likely to be involved in redox partner binding [15,38]. Accordingly, in P450 3A4 these regions may likewise bind the reductase.

The N-terminus of P450 3A4 (residues 8–29) was modeled as a helix, as indicated by the secondary structure prediction. There are no equivalents for that segment in the known bacterial proteins, since all of them are soluble enzymes. In contrast, P450 3A4 is a microsomal protein which contains a membrane-spanning region typical for eukaryotic P450s. In accordance with the currently accepted 'partial immersion' model, microsomal P450s

*The coordinates of the model are available from the corresponding author. E-mail: Szklarz@tonic.pharm.arizona.edu.

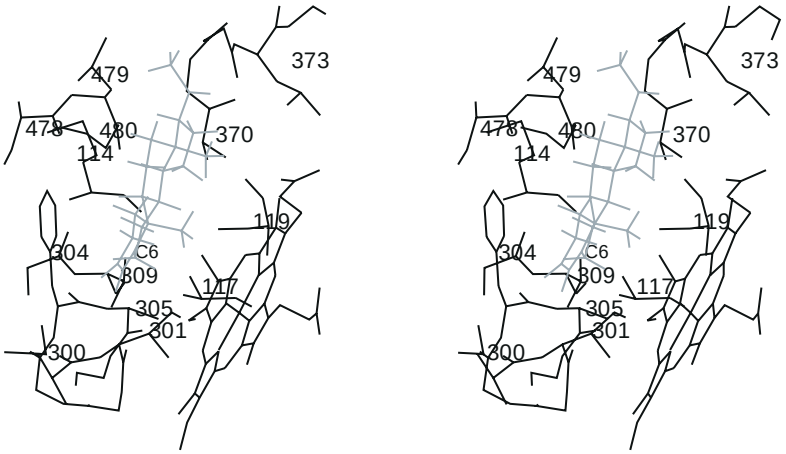

Fig. 3. Progesterone docked into the active site of the P450 3A4 model in a 6β-binding orientation. The substrate is shown in a dotted line, with all hydrogens displayed. The 3A4 residues shown are within 5 Å from the substrate.

contain one transmembrane helix which serves as a membrane anchor, and the bulk of the protein lay outside the lipid bilayer of the ER [39]. However, some of the recent findings indicate that, at least in the case of P450 2B4, the protein is deeply immersed in the membrane [40]. In the case of P450 3A4, its surface is largely hydrophilic or neutral, with the exception of the substrate docking region/mouth of the access channel, which includes a highly hydrophobic F-G loop and portions of β-sheets 1 and 2, and the hydrophobic redox binding pocket. A similar surface charge distribution was also found in the model of P450 aromatase [22]. Evidently, the location of P450 3A4 within the endoplasmic reticulum membrane will require further studies.

Active site of P450 3A4 and substrate binding

As in other P450s, the active site of 3A4 is bounded at the bottom by the heme, and on the sides by helix I, the B'-C loop, and regions of β-sheets 6, 1-4 and 4. Further outward from heme are helices B', F, and G [12,15,38]. The active site of P450 3A4 is quite large due to the location of helix B' away from heme, similar to that of P450eryF. However, the β1-5–B' loop is much shorter than its equivalent in P450eryF.

Table 1 lists active site residues of P450 3A4 equivalent to known substrate contact residues of the bacterial proteins based on the alignment in Fig. 1, and those corresponding to key residues of P450 2B1, a rat enzyme extensively studied in our laboratory, based on the previously published alignment [23]. All of these residues are located in the active site or substrate access channel of P450 3A4, as shown in Fig. 2. They represent potential substrate contact residues, and thus only some of them may interact with a given substrate, depending upon its size and orientation in the active site.

Two different substrates, the steroid progesterone and

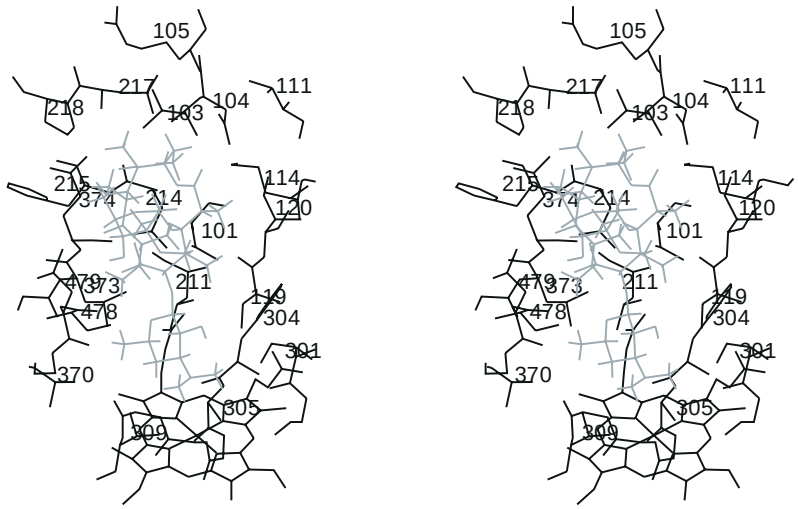

Fig. 4. Erythromycin docked into the active site of the P450 3A4 model. The substrate is shown in a dotted line, with all hydrogens displayed. The 3A4 residues shown are within 5 Å from the substrate.

the macrolide antibiotic erythromycin, were docked into the active site of P450 3A4. Progesterone is hydroxylated by 3A4 at several positions, but the major product is 6β-OH progesterone [41]. The substrate bound in this orientation is within 4 Å from the side chains of residues 114, 119, 301, 304, 305, 309, 370, 373, and 479 (Fig. 3), and is stabilized in the active site of P450 3A4 mainly by hydrophobic interactions. These interactions were also found to be predominant in the binding of steroids in P450 2B1 [17,23]. Due to its larger size, erythromycin contacts more active site residues than progesterone (Fig. 4). The residues within 4 Å from erythromycin are 101, 104, 105, 114, 119, 214, 215, 217, 218, 301, 304, 305, 309, 370, 373, 374, 478, and 479. As with progesterone, erythromycin is stabilized mainly through hydrophobic interactions.

A comparison of residues listed in Table 1 with those which contact progesterone bound in a 6β-orientation (Fig. 3) shows that all residues which interact with this substrate are included in the table. However, only some residues from Table 1 may interact with progesterone. In the case of erythromycin, most putative substrate contact residues, with the exception of Val^{101} and Ser^{478}, are listed in Table 1. It is also of interest to compare substrate contact residues of P450 3A4 with key residues of P450 2B enzymes, since their structure–function relationships have been extensively studied by site-directed mutagenesis and molecular modeling (Ref. 23 and references cited therein; Refs. 42 and 43]. Most 3A4 residues listed in Table 1 may interact with one and/or the other substrate (Figs. 3 and 4), with the exception of residues Ala^{297}, Gly^{480} and Leu^{482}, equivalent to Ile^{290}, Gly^{478} and Ile^{480} of 2B1, respectively. Thus, the importance of these three residues may be specific to 2B enzymes.

The 3A4 residues that contact progesterone and/or erythromycin will be further studied by site-directed mutagenesis, as has been done for P450 2B1 [17,23]. The importance of some of these residues, such as Ala^{370} and Leu^{373}, has recently been confirmed by site-directed mutagenesis (You Ai He, unpublished observations). The model can thus guide experimental studies, and these can in turn help to further refine the 3A4 structure.

An interesting question concerning P450 3A4 is the activation of the enzyme by α-NF, the mechanistic basis of which is poorly understood. Hypotheses include: (i) α-NF increases the affinity of P450 for the substrate [44]; (ii) α-NF increases the affinity of P450 for reductase [45]; (iii) α-NF binds in the same pocket as the substrate and increases coupling efficiency [7]; or (iv) α-NF is an allosteric effector binding at a distinct site and causing a conformational change in the substrate binding pocket [7,44,46]. With the 3A4 model, it will be possible to explore some of these hypotheses. In particular, hypothesis (iii), which assumes simultaneous binding of substrate and effector in the active site, can now be tested. Since the active site of 3A4 is quite large and can bind a substrate such as erythromycin, it may also be able to accommodate a steroid and α-NF simultaneously.

Conclusions

We have constructed a molecular model of cytochrome P450 3A4, an important human drug-metabolizing enzyme. The structure was built using a consensus strategy, based on the crystal coordinates of four known bacterial enzymes: P450 BM-3, P450cam, P450terp, and P450eryF. The 3A4 model retains the core structure characteristic for cytochromes P450, but there are variations in less conserved regions. The model is most similar to the structure of P450 BM-3, another class II P450, but the B' helix resembles that of P450eryF. The active site is thus enlarged, analogous to P450eryF, which enables the enzyme to metabolize large substrates, such as erythromycin. The accuracy of homology models increases with the number of structures solved. In particular, the P450 3A4 model would have been difficult to build without the coordinates of the recently solved P450eryF, which provided a key to the location of helix B' in human 3A4.

The P450 3A4 model, into which a substrate can be docked and enzyme–substrate interactions examined, makes it possible to pinpoint specific substrate contact residues, which are likely to be important for enzymatic activity. As shown in earlier studies (e.g. Refs. 17 and 23), the substrate can be contacted by various amino acids, depending upon its size and orientation in the active site. Many of the proposed key residues have counterparts in 2B1, a well-studied rat enzyme, but there are a number of residues that are unique to 3A4. These residues would not have been identified without molecular modeling of the 3A4 structure and docking of enzyme substrates.

Useful information can thus be obtained from homology models concerning amino acids important for substrate and/or inhibitor binding. The models can subsequently guide site-directed mutagenesis experiments as a means of experimental verification of model-derived predictions. Thus, homology modeling of eukaryotic enzymes remains a viable approach for studying structure–function relationships of cytochromes P450. In particular, the 3A4 model should prove instrumental for increasing our understanding of the function of this important human enzyme.

Acknowledgements

This work was supported by Grant ES03619 and Core Center Grant ES06694 from the National Institutes of Health. Molecular modeling studies were performed at the Molecular Modeling Facility of the Southwest Environmental Health Sciences Center (SWEHSC) at the University of Arizona, Tucson, AZ, U.S.A. The authors thank Dr. William Remers for reviewing the manuscript.

References

1 Guengerich, F.P., In Ortiz de Montellano, P.R. (Ed.) Cytochrome P450: Structure, Mechanism and Biochemistry, Plenum, New York, NY, U.S.A., 1995, pp. 473–535.
2 Kolars, J.C., Lown, K.S., Schmiedlin-Ren, P., Ghosh, M., Fang, C., Wrighton, S.A., Merion, R.M. and Watkins, P.B., Pharmacogenetics, 4 (1994) 247.
3 Brian, W.R., Sari, M.-A., Iwasaki, M., Shimada, T., Kaminsky, L.S. and Guengerich, F.P., Biochemistry, 29 (1990) 11280.
4 Aoyama, T., Yamano, S., Waxman, D.J., Lapenson, D.P., Meyer, U.A., Fischer, V., Tyndale, R., Inaba, T., Kalow, W., Gelboin, H.V. and Gonzalez, F.J., J. Biol. Chem., 264 (1989) 10388.
5 Guengerich, F.P., Martin, M.V., Beaune, P.H., Kremers, P., Wolff, T. and Waxman, D.J., J. Biol. Chem., 261 (1986) 5051.
6 Waxman, D.J., Attisano, C., Guengerich, F.P. and Lapenson, D.P., Arch. Biochem. Biophys., 263 (1988) 424.
7 Shou, M., Grogan, J., Mancewicz, J.A., Krausz, K.W., Gonzalez, F.J., Gelboin, H.V. and Korzekwa, K.R., Biochemistry, 33 (1994) 6450.
8 Poulos, T.L., Finzel, B.C., Gunsalus, I.C., Wagner, G.C. and Kraut, J., J. Biol. Chem., 260 (1985) 16122.
9 Poulos, T.L., Finzel, B.C. and Howard, A., J. Mol. Biol., 195 (1987) 687.
10 Ravichandran, K.G., Boddupalli, S.S., Hasemann, C.A., Peterson, J.A. and Deisenhofer, J., Science, 261 (1993) 731.
11 Hasemann, C.A., Ravichandran, K.G., Peterson, J.A. and Deisenhofer, J., J. Mol. Biol., 236 (1994) 1169.
12 Cupp-Vickery, J.R. and Poulos, T.L., Struct. Biol., 2 (1995) 144.
13 Nelson, D.R. and Strobel, H.W., J. Biol. Chem., 263 (1988) 6038.
14 Nelson, D.R. and Strobel, H.W., Biochemistry, 28 (1989) 656.
15 Hasemann, C.A., Kurumbail, R.G., Boddupalli, S.S., Peterson, J.A. and Deisenhofer, J., Structure, 3 (1995) 41.
16 Zvelebil, M.J.J.M., Wolf, C.R. and Sternberg, M.J.E., Protein Eng., 4 (1991) 271.
17 Szklarz, G.D., Ornstein, R.L. and Halpert, J.R., J. Biomol. Struct. Dyn., 12 (1994) 61.
18 Koymans, L.M.H., Vermeulen, N.P.E., Baarslag, A. and Donné-Op den Kelder, G.M., J. Comput.-Aided Mol. Design, 7 (1993) 281.
19 Laughton, C.A., Zvelebil, M.J.J.M. and Neidle, S., J. Steroid Biochem. Mol. Biol., 44 (1993) 399.
20 Vijayakumar, S. and Salerno, J.C., Biochim. Biophys. Acta, 1160 (1992) 281.
21 Lewis, D.F.V., Xenobiotica, 25 (1995) 333.
22 Graham-Lorence, S., Amarneh, B., White, R.E., Peterson, J.A. and Simpson, E.R., Protein Sci., 4 (1995) 1065.
23 Szklarz, G.D., He, Y.A. and Halpert, J.R., Biochemistry, 34 (1995) 14312.
24 Lewis, D.F.V. and Moereels, H., J. Comput.-Aided Mol. Design, 6 (1992) 235.
25 Ferenczy, G.G. and Morris, G.M., J. Mol. Graph., 7 (1989) 206.
26 Gonzalez, F.J., Schmidt, B.J., Umeno, M., McBride, O.W., Hardwick, J.P., Meyer, U.A., Gelboin, H.V. and Idle, J.R., DNA, 7 (1988) 79.
27 Paulsen, M.D. and Ornstein, R.L., Proteins, 11 (1991) 184.
28 Paulsen, M.D. and Ornstein, R.L., J. Comput.-Aided Mol. Design, 6 (1992) 449.
29 Rost, B. and Sander, C., J. Mol. Biol., 232 (1993) 584.
30 Rost, B. and Sander, C., Proteins, 19 (1994) 55.
31 Rost, B., Sander, C. and Schneider, R., Comput. Appl. Biosci., 10 (1994) 53.
32 Bowie, J.U., Luthy, R. and Eisenberg, D., Science, 253 (1991) 164.
33 Luthy, R., Bowie, J.U. and Eisenberg, D., Nature, 365 (1992) 83.
34 Iwasaki, H., Sugawara, Y., Adachi, T., Morimoto, S. and Watanabe, Y., Acta Crystallogr., C49 (1993) 1227.
35 Graham-Lorence, S., Khalil, M.W., Lorence, M.C., Mendelson, C.R. and Simpson, E.R., J. Biol. Chem., 266 (1991) 11939.
36 Iwasaki, M., Darden, T.A., Pedersen, L.G., Davis, D.G., Juvonen, R.O., Sueyoshi, T. and Negishi, M., J. Biol. Chem., 268 (1993) 759.
37 Iwasaki, M., Darden, T.A., Perker, C.E., Tomer, K.B., Pedersen, L.G. and Negishi, M., J. Biol. Chem., 269 (1994) 9079.
38 Graham-Lorence, S. and Peterson, J.A., FASEB J., 10 (1996) 206.
39 Black, S.D., Martin, S.T. and Smith, C.A., Biochemistry, 33 (1994) 6945.
40 Miller, J.P., Herbette, L.G. and White, R.E., Biochemistry, 35 (1996) 1466.
41 Waxman, D.J., Lapenson, D.P., Aoyama, T., Gelboin, H.V., Gonzalez, F.J. and Korzekwa, K., Arch. Biochem. Biophys., 290 (1991) 160.
42 Szklarz, G.D., He, Y.Q., Kedzie, K.M., Halpert, J.R. and Burnett, V.L., Arch. Biochem. Biophys., 327 (1996) 308.
43 He, Y.Q., Szklarz, G.D. and Halpert, J.R., Arch. Biochem. Biophys., 335 (1996) 152.
44 Schwab, G.E., Raucy, J.L. and Johnson, E.F., Mol. Pharmacol., 33 (1988) 493.
45 Huang, M.-T., Johnson, E.F., Muller-Eberhard, U., Koop, D.R., Coon, M.J. and Conney, A.H., J. Biol. Chem., 256 (1981) 10897.
46 Ueng, Y.-F., Shimada, T., Yamazaki, H. and Guengerich, F.P., Chem. Res. Toxicol., 8 (1995) 218.