Journal of Molecular Graphics and Modelling 28 (2010) 524–532

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Molecular modeling study on orphan human protein CYP4A22 for identification of potential ligand binding site

Poornima Gajendrarao [1], Navaneethakrishnan Krishnamoorthy [1], Sugunadevi Sakkiah, Prettina Lazar, Keun Woo Lee *

Department of Biochemistry, Division of Applied Life Science (BK21 Program), Environmental Biotechnology National Core Research Center (EBNCRC), Plant Molecular Biology and Biotechnology Research Center (PMBBRC), Gyeongsang National University, Jinju 660-701, Republic of Korea

ARTICLE INFO

Article history:
Received 6 July 2009
Received in revised form 26 November 2009
Accepted 30 November 2009
Available online 4 December 2009

Keywords:
Cytochrome P450 4A22
Homology modeling
Molecular dynamics simulation
Molecular docking
Arachidonic acid
Erythromycin
Ligand binding site

ABSTRACT

A molecular structure is an essential source to identify ligand binding sites in orphan human cytochrome P450 4A22 (CYP4A22) that belongs to family 4, which is known to be involved in the regulation of blood pressure. Thus, a homology model has been constructed for CYP4A22 and refined by molecular dynamics simulation (MDS). Subsequently, molecular docking was performed with possible substrates, arachidonic acid (essential fatty acid, AA) and erythromycin (therapeutic drug, ERY). These complexes were also subjected to MDS, which helped in predicting the energetically favorable binding sites for these ligands. Putative substrate recognition sites (SRSs) of this protein provide highly hydrophobic binding pockets for the target ligands. A few key ligand binding residues identified in this study indicates that they could also play a major role in ligand-channeling (F122, L132 and C230). Furthermore, it appears that they might serve critical support for the catalytic reaction center (E321, F450, P449 and R455). Structural analysis of channels proposed that the conformational changes might have originated from the active site upon ligand binding and transferred to the rest of the protein via SRSs, which could thereby regulate the channels in CYP4A22. Most of our prediction results are supported by other research groups. In summary, the first molecular modeling study of CYP4A22 yields structural knowledge, which would be helpful to design structure-based-drugs and functional experiments for the target protein.

1. Introduction

Cytochrome P450s (CYPs) are well-known for their monooxygenase reaction. These are heme containing enzymes and are involved in phase I metabolism. They also participate in a wide variety of reactions like oxidative, peroxidative and reductive metabolic biotransformation and metabolize endogenous compounds such as fatty acids, and prostaglandins, and exogenous compounds such as therapeutic drugs and xenobiotics [1,2]. This superfamily has received wide attention as they metabolize the majority of the marketed drugs. For instance, seven of these superfamily members metabolize more than 90% of the marketed drugs [3]. However, one quarter of the CYP family remains orphan, which means that their function, expression sites and substrate information are not yet clearly understood. Thirteen members of this superfamily are classified as orphans by Guengerich [4] including a family 4 CYP member A22 (CYP4A22).

The amino acid sequence of the orphan human protein CYP4A22 has been identified recently [5]. The three-dimensional structure of this protein is not yet known. Experimental studies of CYP4 family members previously proposed that they can metabolize arachidonic acid (AA) like endogenous compounds to its ω-hydroxy product, which are connected with hypertension [6]. In general, CYP family members share their substrates with one another. A protein from the same family 4, CYP4F11, metabolizes the therapeutic drug erythromycin (ERY) more efficiently and provides a large open access channel for ERY [7]. However, to date, information regarding the structure, ligand binding site and channels are not available for CYP4A22.

Structural information might help us to understand the ligand interaction and channeling of CYP4A22. Thus, we have built a homology model based on the known crystal structures and employed molecular dynamics simulation (MDS) for model refinement, molecular docking, and MDSs coupled with energy calculations for CYP4A22–AA and CYP4A22–ERY complexes. This helped in finding an energetically favorable binding site for AA and

* Corresponding author. Tel.: +82 55 751 6276; fax: +82 55 752 7062.
E-mail address: kwlee@gnu.ac.kr (K.W. Lee).
[1] These authors contributed equally to this work.

1093-3263/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2009.11.010

ERY, key ligand-interacting residues and other significant roles from the predicted homology model of CYP4A22. In addition, we propose possible channels of CYP4A22, key residues at both the ends of each channel, and structural and conformational changes in the channel-forming regions (B/C loop and helices F′, G′ and I).

2. Materials and methods

2.1. Construction of the homology model

Homology modeling was used to build the model for orphan human protein CYP4A22. The protein sequence (519 amino acids) was retrieved from UniProt (http://www.uniprot.org) (UniProtKB accession number: Q5TCH4). The Discovery Studio V2.0 (DS) [8] was used for homology model construction (Accelrys, San Diego, USA). "Blast search" within DS was utilized to select the templates. The top-ranked three templates were selected for the model building. Protein data bank (PDB, http://www.rcsb.org/pdb/) identifications of the templates are 1TQN, 2IJ2, and 3C6G, which are the high resolution crystal structures of CYP 3A4 [9], BM3 [10], and 2R1 [11], respectively. The templates were aligned with the target and examined for conserved sequence. The aligned sequences were taken for the model construction and built using "build homology model" within DS. The coordinates for heme was obtained from 1TQN and positioned as in the template. The resulting model was subjected to MDS using GROMACS V3.3 [12] package for the structural refinement.

2.2. MDS of CYP4A22 model

Energy minimization was carried out for the constructed CYP4A22 model by the steepest descent algorithm with a tolerance of 2000 kJ/mol/nm using the GROMACS package. MDS for the energy minimized structure was done using GROMOS96 [13] force field. The model was solvated in a cubic box with the size of 1.2 nm. The SPC3 [14] water model was used to create aqueous environment. Periodic boundary conditions were applied in all directions and the system was neutralized by adding ~6 Cl ions, replacing the water molecules. The resulting system contains ~100, 110 atoms. A twin range cutoff was used for long-range interactions: 0.9 nm for van der Waals interactions and 1.4 nm for electrostatic interactions. All bond lengths were constrained with the LINCS [15] algorithm. The SETTLE [16] algorithm was applied to constrain the geometry of water molecules. The energy minimized system was subjected to 50 ps equilibration. This pre-equilibrated system was subsequently used in the 3 ns production MDS with a time-step of 2 fs at constant temperature (300 K), pressure (1 atm) and number of particles, without any position restraints [17]. The snapshots were collected at every 5 ps and analyzed using GROMACS analysis tools. The lowest potential energy conformation was selected from the 3 ns MDS trajectory and further refined by energy minimization for molecular docking. The refined model was validated with ProSA [18] *z*-score and PROCHECK [19] Ramachandran plot.

2.3. Molecular docking and MDSs of CYP4A22–ligand complexes

The DS package was used to dock AA and ERY to our refined model. One of the preferred ligands is a long-chain ω-6 fatty acid with a hydrophobic tail, AA (Fig. 1A) and the other is a large hydrophilic drug molecule, ERY (Fig. 1B). These ligand molecules were built using ISIS draw, and optimized using "Prepare Ligands" of the DS for docking. The optimized ligand molecules were docked into the refined protein model using "LigandFit" of the DS. The binding sites for these molecules were selected based on the ligand binding pocket of the templates [9–11]. For each ligand, 50 poses

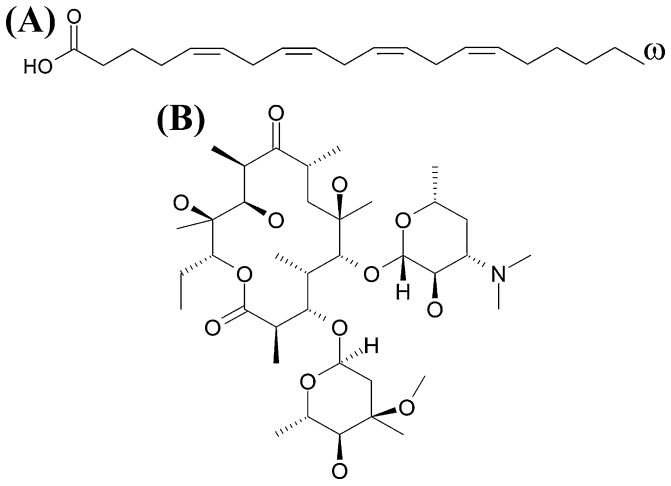

Fig. 1. Two ligand structures, which are used for MDSs. (A) Arachidonic acid and (B) Erythromycin.

were generated and scored using DS scoring functions that include Ligscore1, Ligscore2, PLP1, PLP2 and PMF. Among these poses, the most suitable docking mode for AA and ERY along with a high score from consensus scoring functions was finally selected. Since the hydroxylation site of AA is known [6], the distance between the heme and the ω-hydroxylation site of AA was also considered while selecting the lowest energy conformation with a high consensus score. Whereas ERY is a deduced substrate for CYP4A22, the lowest energy conformation along with high consensus score was preferred. Energy minimization and MDSs for both protein–ligand complexes were also carried out to permit flexible interactions and to observe the ligand's sustainability in the binding pocket using GROMACS with the same MDS parameters described previously. The PRODRG [20] web server was used to construct the ligand's topologies for MDS. In order to identify the key ligand-interacting residues, the lowest energy structures from the 3 ns MDS were used in the DS for calculating the interaction energy. The molecular graphical representations were prepared using PyMOL program (www.pymol.org).

3. Results

3.1. Sequence analysis of CYP4A22

The sequence conservation and the signature motifs of CYP4A22 were examined using multiple sequence alignment with templates (Fig. 2). The sequence identity between the target and the template structures 1TQN, 2IJ2 and 3C6G were 26%, 27% and 25%, respectively and the sequence similarity was >45%, which is reasonable for model building. As stated by Lewis et al. [21], several signature motifs were found to be conserved in CYP4A22. The structural motifs of CYP4A22 are as follows: from the N-terminal region of the protein, pentapeptides in the C helix (WxxxR) and in the I helix (xGxxT), a tetrapeptide in the K helix (ExxR), a dodecapeptide prior to the L helix (ZxxPxxZxPxxZ) and a decapeptide between L helix and the dodecapeptide (FxxGxxxCxG), where, Z could be any aromatic amino acid and x could be any residue. Among these signature motifs, FxxGxxxCxG is a very essential one, since it is the characteristic motif for the CYP superfamily, which includes a conserved cystine residue that ligates to the Fe of the heme. The basic residue, R455, from the same signature motif is found to be conserved well in all three templates, and it interacts with the propionates of heme in several CYPs. The other motifs were also substantially conserved

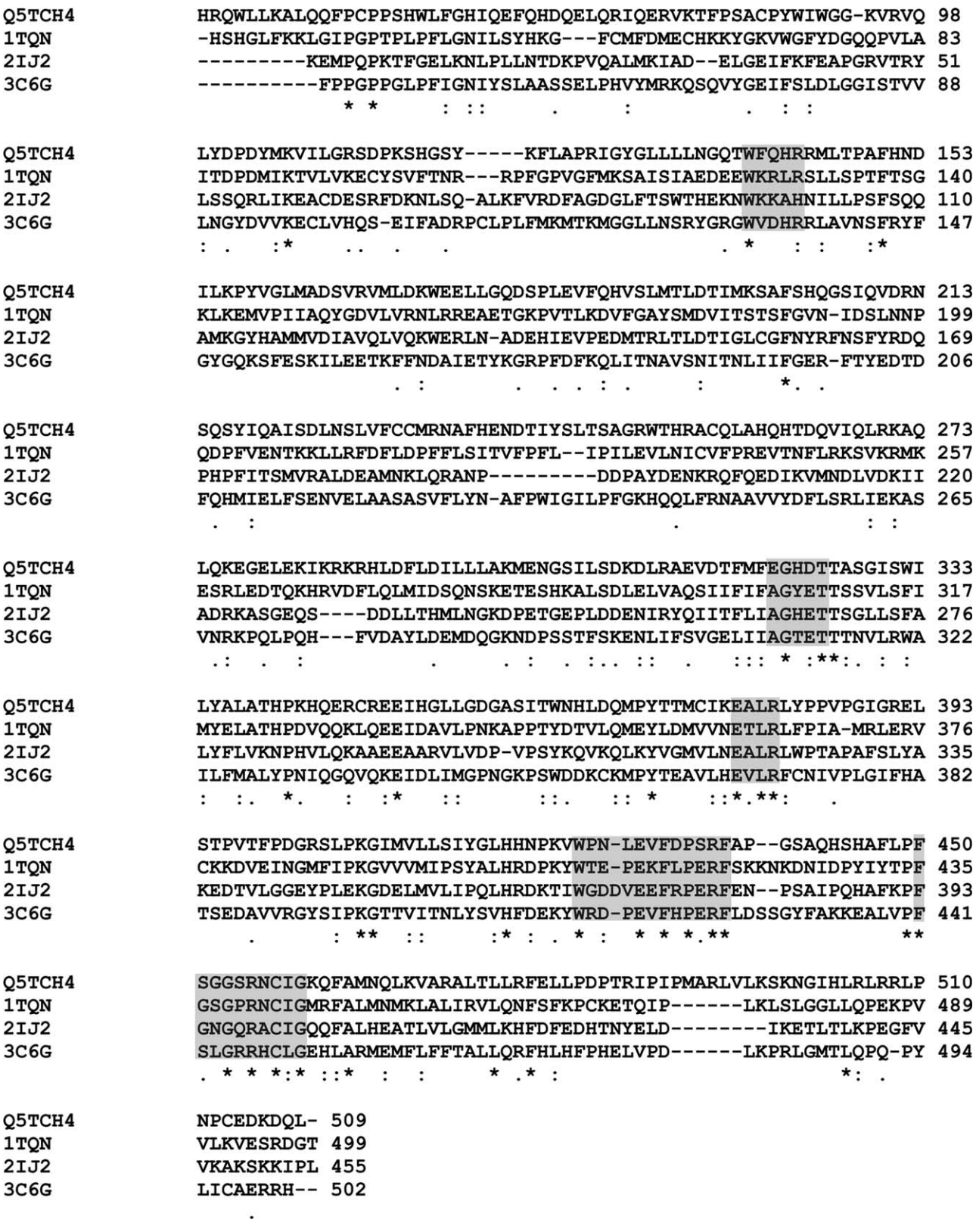

Fig. 2. Multiple sequence alignment. Target (CYP4A22: Q5TCH4) and templates sequences alignment is represented using ClustalW program for the intention of clarity. Here, asterisks indicate the residues conserved in all four structures. Signature motifs of CYP superfamily are highlighted by gray shade.

with the required features except I helix motif, in which an acidic residue, E is present in the first position instead of a basic residue A or G. The conservation of sequence elucidates that CYP4A22 model construction based on this alignment is reliable.

3.2. CYP4A22 model refinement by MDS

The constructed model was subjected to MDS, in order to assess the stability of the model and to find the energetically favorable

structure for further docking study. Our MDS trajectory-based analysis shows that the potential energy of the model gradually decreased from −1372500 kJ/mol to −1376900 kJ/mol, which indicates that the model is energetically stable during MDS (Fig. 3A). Structural stability of the constructed model during the MDS was examined using the root mean square deviation (RMSD). Fig. 3B shows the RMSD plot for the protein Cα-atoms with reference to the initial structure, as a function of time. The first 1000 ps were considered as an equilibration period. There was a gradual rise in the RMSD till ∼0.35 nm followed by a plateau where the model becomes stable. Distinguishing the flexible regions of the protein could help in understanding the stability of the protein. To examine the flexible regions of the model, we have generated the average root mean square fluctuation (RMSF) plot for the Cα-atoms with respect to the residues. Here, those residues, which deviated more than 0.25 nm were considered as highly flexible elements of this protein (Fig. 3C). As anticipated, the loop regions of the model, B/C, D/E, G/H, H/I, J/K and β4/β5 contributes to increasing fluctuations. Loops B/C and β4/β5 belong to SRS-1 and 6, while G/H and H/I loops are near to SRS-3 and 4. Substrate recognition sites (SRSs) of CYP4A22 are described in the structural features section. The flexibility of other regions was lower than 0.2 nm, which clearly shows that the protein structural core is well-constructed and the model is stable in the course of MDS. Results of the trajectory-based analysis also indicate that the quality of the generated model is reliable.

The energetically most favorable conformation was selected from the 3 ns MDS and subjected to further energy minimization for the fine refinement. This optimized model was validated using the ProSA *z*-score. The required negative scores for all residues were obtained, which further confirmed that the model is reliable (Fig. 4). The Ramachandran plot Φ/Ψ distribution of backbone conformational angles for each residue of the refined structure revealed that 80.0%, 15.2%, and 1.9% of the residues are located in the most favorable, additionally allowed and generously allowed regions, respectively. Moreover, overall the PROCHECK *G*-factor value −0.3 for the refined model also indicates the higher quality of the constructed model. The ProSA and PROCHECK results reveal that the optimized model is satisfactory, and is thus considered as a reliable source for the rest of the study.

3.3. Structural features of CYP4A22 model

Our model consists of helices A–L, five antiparallel β sheets and their connecting loops (Fig. 5A). The structures of the CYP superfamily deposited in the PDB have missing N-terminus residues, as a result, the constructed model also lack the first 39 residues at the N-terminal region. Comparison of the model with the templates clearly shows that the overall CYP fold is preserved in CYP4A22. In addition, the model contains several structurally conserved regions: helices D, E, I, J and L. It also consists of structurally variable regions that are known as plastic regions of other CYPs: N-terminal helices A, B, F, F′, G and G′ and their connecting loops. The heme is positioned in between two structurally conserved helices, I and L and its N-terminal loop. The propionate side chains of the heme interact with the protein by making hydrogen bonds with the residues L132 from the B/C loop, W139 and R143 from the C helix, E321 from the I helix, highly conserved G452 and R455 from the N-terminal loop of L helix, and G490 from β4–β5 loop (Fig. 5B).

SRSs of the CYP4A22 are predicted based on Gotoh's (1992) proposal [22] (Fig. 5C). SRS-1, SRS-2 and SRS-3 are observed in the B/C loop, F helix and G helix, respectively. As expected, structurally conserved I helix contains SRS-4. SRS-5 lies in the loop that connects helix K and β3 and a region between β4 and β5 forms

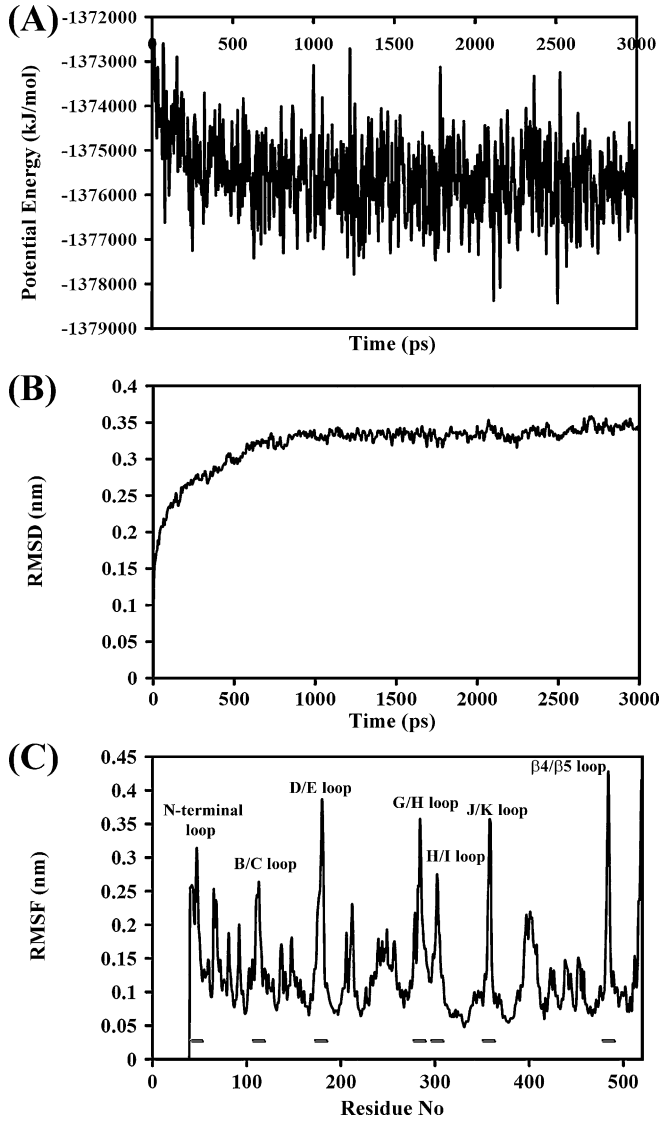

Fig. 3. MDS trajectory-based analyses for the model refinement. (A) Potential energy of CYP4A22 in the course of MDS, (B) RMSD of the Cα-atoms with respect to their initial structure shows stable nature of the model after the initial equilibration time, and (C) Average RMSF plot shows flexible loops and conserved core region of the protein.

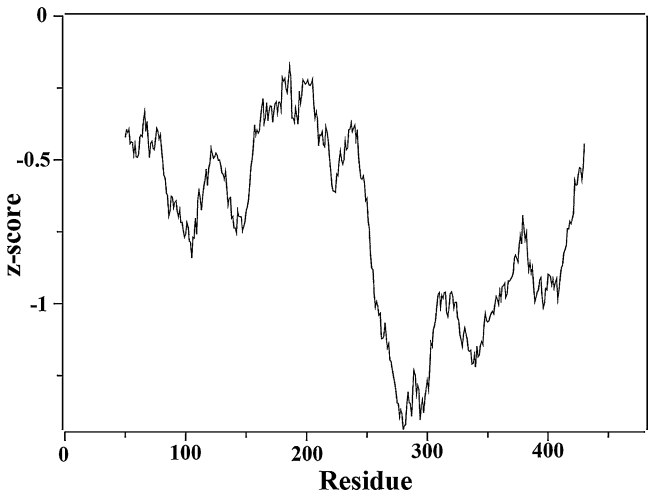

Fig. 4. Quality assessment of the model. ProSA negative *z*-score plot with respect to residues depicts the quality of the optimized model is satisfactory.

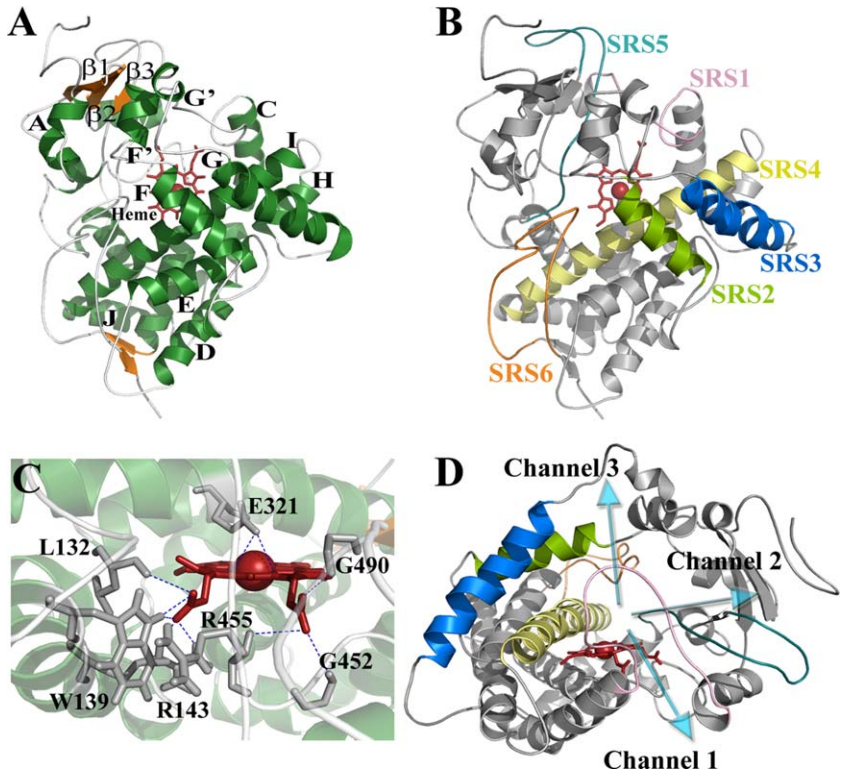

Fig. 5. Structural features, SRSs, and possible channels of CYP4A22. (A) Structural elements of the constructed CYP4A22 model, where heme is shown in sticks (firebrick), α-helices and ß-strands are indicated by green and orange, respectively and labeled. (B) CYP4A22 residues that interact with heme by making hydrogen bonds (blue dashes). (C) Substrate recognition sites of CYP4A22: SRS-1; light pink, SRS-2; light green, SRS-3 blue; SRS-4; pale yellow, SRS-5; teal and SRS-6; orange. SRS regions are depicted based on Gotoh et al. (1991). (D) Proposed channels of CYP4A22.

SRS-6. Possible channels of CYP4A22 have also been predicted and represented in Fig. 5D according to CYP3A4. Channel-1 of CYP4A22 travels through the B/C loop. Channel-2 forms between β2, the B/C loop and the end of F′ helix. Channel-3 egresses between the F′, G′ and I helices. In the optimized model, several structural features are well conserved, which suggests that the model is appropriate for subsequent docking studies.

3.4. Key residues responsible for ligand interaction

To find the potential ligand binding site, we employed MDS for the protein–ligand complexes. The distance between the center of mass for heme and ligands (heme ↔ AA ω-tail, since it is a linear molecule) were calculated in order to demonstrate the AA and ERY accommodation near the reaction center (Fig. 6). The ω-tail of AA was slightly away from the heme in the CYP4A22–AA conformation that obtained from DS with high consensus score. It also stayed away in the initial few picoseconds of MDS (Fig. 6 and S1). However, it shortly found a place near the catalytic reaction center, which could favor hydroxylation of AA. The distance plateau elucidates that it resides parallel to heme for the rest of the MDS. In contrast, ERY stably maintained its place in the binding site whose sugars were away from the Fe atom of the heme. The distance plateau between the heme and ligands suggest that they could adopt well in the binding pocket of CYP4A22.

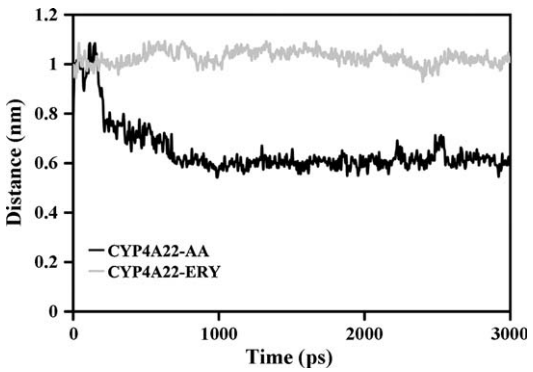

Fig. 6. Accommodation of ligands in the binding pocket. Distance between the ligands and the heme with respect to their initial structure.

Key residues of CYP4A22, which are responsible for ligand binding, were identified by calculating the residue interaction energy for the lowest energy conformations of CYP4A22–ligand complexes. Residues with interaction energy lower than −2 kcal/mol were considered to be crucial for ligand binding. The residue interaction energies for the key residues are listed in Tables 1 and 2 including van der Waals, electrostatic and total energies. Table 1 shows that the contribution of van der Waals energy is important for AA binding, which represents the hydrophobic nature of

Table 1
Residue interaction energies for CYP4A22–AA complex.

Residue	E_{vdW}[a] (kcal/mol)	E_{ele}[b] (kcal/mol)	E_{total}[c] (kcal/mol)
R455	−6.173	−2.238	−8.411
L133	−3.301	−3.252	−6.553
E321	−2.887	−2.094	−4.981
G388	−5.156	−0.25	−5.406
G118	−2.552	−2.096	−4.648
F122	−1.790	−1.403	−3.193
D324	−2.392	−0.023	−2.415
Y120	0.069	−2.206	−2.137
S454	−1.993	−0.035	−2.028

[a] vdW: van der Waals interaction energy.
[b] ele: electrostatic interaction energy.
[c] total: total interaction energy.

Table 2
Residue interaction energies for CYP4A22–ERY complex.

Residue	E_{vdW}[a] (kcal/mol)	E_{ele}[b] (kcal/mol)	E_{total}[c] (kcal/mol)
S451	−7.904	−0.046	−7.95
F229	−4.944	−2.673	−7.617
D324	−6.389	−0.066	−6.455
F450	−4.996	−0.086	−5.082
V228	−0.996	−3.686	−4.682
A313	−3.881	−0.273	−4.153
I389	−2.888	−1.093	−3.981
A124	−3.516	−0.032	−3.548
L132	−2.069	−0.206	−2.275
N500	−2.139	−0.135	−2.273
C230	−1.487	−0.663	−2.151
P449	−2.044	−0.021	−2.066
A235	−1.954	−0.08	−2.033

[a] vdW: van der Waals interaction energy.
[b] ele: electrostatic interaction energy.
[c] total: total interaction energy.

binding with CYP4A22. It is interesting that the identified key residues G118, Y120, F122, L133, E321, D324 and G388 are mostly from SRS-1, SRS-4 and SRS-5. Two other ligand-interacting residues, S454 and R455, are located in the cysteine containing loop between β3 and the L helix. The channel-1 forming B/C loop (SRS-1) is interacting with AA by four residues G118, Y120, F122 and L133. The residues E321 and D324 in the highly conserved I helix (SRS-4; channel-3) are also interacting with AA. In addition to hydrophobic interactions, hydrogen bonds between protein and ligand were also identified (Fig. 7A). CYP4A22 anchors AA using two polar residues S119 and N135. The backbone oxygen atom and nitrogen side chains of the residues created hydrogen bonds with the hydroxyl oxygen atom of the AA that is located opposite to the ω-end. It is remarkable that two hydrogen bond forming residues are located in the channel-forming B/C loop. The results of interaction energy calculations suggest the possible binding site for AA in CYP4A22. They also provide the architecture of the AA binding pocket, which appears to be constructed by arranging the hydrophobic interactions at one side of the ligand binding pocket with the polar contacts on the other side where the interactions are mostly from SRSs.

There were two reasons for selecting ERY as one of the ligands for our molecular docking and MDS studies. Firstly, recent studies indicate that the CYP4 family members can also metabolize therapeutic drugs like ERY [7]. Secondly, ERY was identified as one of the efficient substrates for CYP4F11. Moreover, CYP4A22 and CYP4F11 share 40% sequence identity. Since CYP family members share their substrates with one another, ERY could also be a ligand for CYP4A22. Therefore, we performed MDS with interaction energy calculation for the CYP4A22–ERY complex to identify the binding site and the key residues that interact with ERY.

The binding pocket for ERY in CYP4A22 is also mainly constructed by van der Waals interactions. The key residues (Table 2) are mostly found to be located in SRS-1–2, 4–6 (A124, L132, V228, F229, C230, A235, A313, D324, I389 and N500) and a few from the conserved loop that contains cysteine (P449, F450 and S451). Since ERY is relatively larger than AA, it interacts with several residues of CYP4A22. It is comprehensible that these residues are highly hydrophobic in nature where the interaction energy is largely obtained from van der Waals interactions. Significantly, half of the ERY interacting residues, A124, L132, V228, F229, C230, A235, A313 and D324 are from the B/C and F/F′ loops and the F′ and I helices. In general, these are the key structural elements for the formation of channels 1–3 and belong to SRS-1–4 in CYP4A22. The rest of the ERY interacting residues are from SRS-5 and 6 (I389: SRS-5 and N500: SRS-6) and from the cysteine containing loop P449, F450 and S451. Five hydrogen bonds were observed between CYP4A22 and ERY, in which three residues created hydrogen bonds with ERY from three different directions (Fig. 7B). A charged residue R126 formed three hydrogen bonds with the oxygen of ERY D-desoamine group. On the other side, G501 formed a hydrogen bond with the oxygen of the L-cladinose group. Surprisingly, a charged residue E321, which formed a hydrogen bond with heme in the ligand-free state (Fig. 5C), also formed hydrogen bonds with both heme and ERY in the CYP4A22–ERY complex (Fig. 7B). These residues, R126 and E321, are also from SRS-1 and 4, respectively. Results of residue interaction energy unveil the key residues, which are important for interaction with this therapeutic drug. Furthermore, the architecture of SRSs provides a possible binding site for ERY, which is stabilized by more hydrophobic and few polar interactions.

3.5. Key residues in putative channels of CYP4A22

The available crystal structures of various CYPs indicate the importance of the channels [23,24] that are required for substrate access and egress, whose active site is buried. Nevertheless, several MDS studies of CYPs revealed the dynamical changes of these channels that are not seen in the crystal structure. They also described how dynamic motions can cause conformational changes that are essential to understand ligand-channeling in proteins [25–27]. A recent report suggested six pairs of residues, situated at both the ends of each channel, which are important for channeling in the metyrapone-bound CYP3A4 structure [28]. The corresponding six pairs in CYP4A22 are: F122 and L133, L123 and L132 in channel-1, K94 and F122, K94 and T241 in channel-2, and C230 and L258, and C230 and F320 in channel-3. We found that in

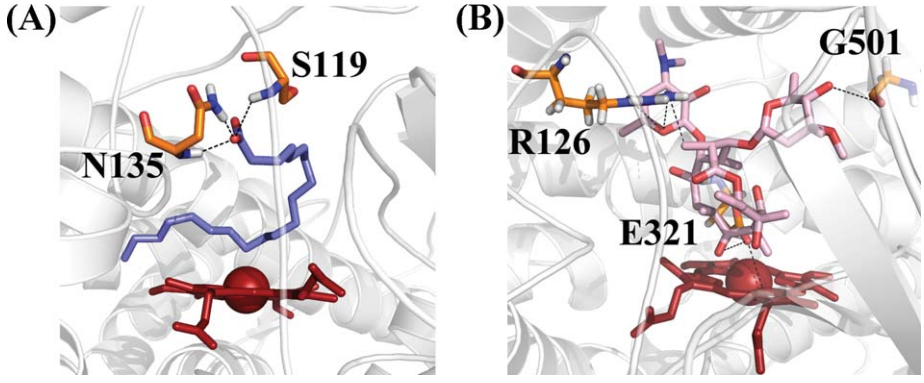

Fig. 7. Hydrogen bonds in protein–ligand complexes. (A) Possible binding mode of AA and its hydrogen bond interactions (black dashes) with CYP4A22 residues side chains. (B) Interaction between ERY and the residues side chains and its mode of binding. Here, heme (firebrick), AA (violet) and ERY (light pink) are shown in sticks.

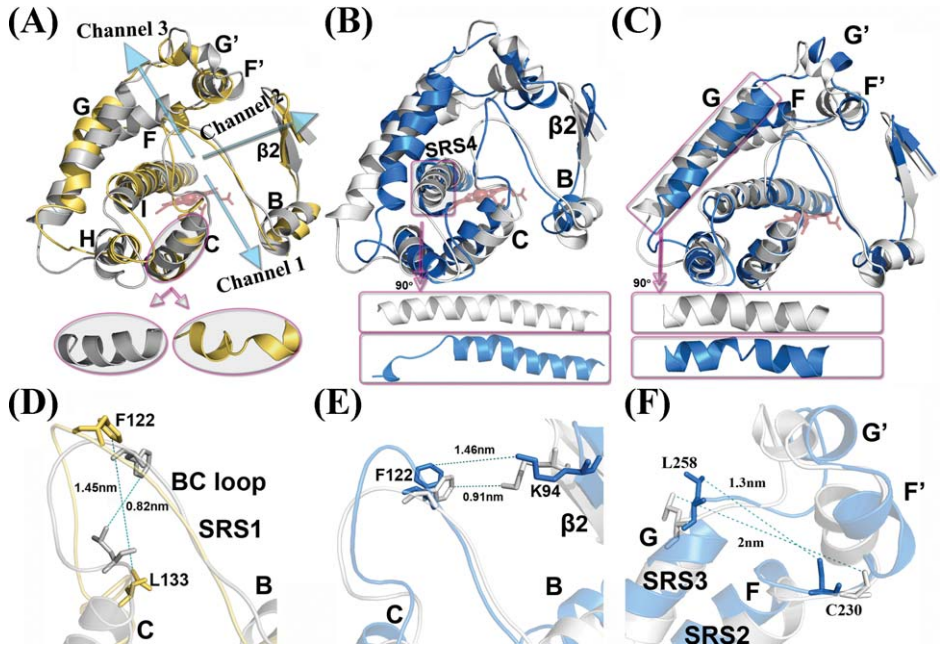

Fig. 8. Structural and conformational changes in the channel-forming regions. (A) Changes in structure of channel-forming regions in CYP4A22–AA complex at 1445 ps (yellow) and the C helix structural change magnified in ellipses. (B) and (C) Structural and conformational changes in CYP4A22–ERY (blue) complex at 2555ps (B) and at 2025 ps (C). Structural changes on I and G helices clearly showed in rectangles. (D) Distances between entrance key residues of channel-1 in CYPA22-AA complex. (E) and (F) Distances between key residues at channel-2 and channel-3 in CYPA22-ERY, respectively. Here, snapshots from MDS are superimposed with their initial structures (gray) and entrance key residues are shown in sticks.

CYP4A22, two of the corresponding residues are conserved (T241 and F320) and a few other residues are directly in contact with ligands (F122, L132, L133 and C230) (Tables 1 and 2). In addition, the rest of the ligand-interacting residues are mostly located near the key residues that form channels. Interaction energy analysis also proved that the ligand-interacting residues are mostly from SRSs, which are building blocks for channels. These results indicate that the six pairs of residues in CYP4A22 have a major role to play in ligand-channeling.

3.6. Structural and conformational changes of channel-forming regions

To understand the changes in the structures of the channels, we examined the representative snapshots from MDSs that showed higher and lower distances between key residues in the channels. Fluctuation of the residues in CYP4A22–ligand complexes was analyzed using RMSF calculation (Figs. 8 and 9). The distance between the key residues (F122 ↔ L133) at the entrance of channel-1 in CYP4A22–AA increased from 0.82 nm to 1.45 nm at 1445 ps (Fig. 8D), in which, the B/C loop deviated ~0.6 nm. Helix C was broken into two short helices and stretched out (Fig. 8A). Also, a large coordinate shift was seen on the other channel-forming helices (F′, G′ and I (SRS-4)) and their neighboring helices (B, F (SRS-2), G (SRS-3) and H). The RMSF plot of CYP4A22–AA (Fig. 9) also agrees with the structural observation data. These results indicate the extensive conformational and structural changes on the channel-forming elements of the CYP4A22–AA complex, especially at channel-1.

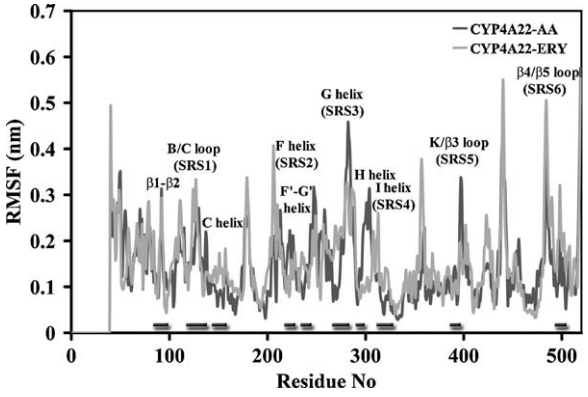

Fig. 9. Residues fluctuations of ligand-bound CYP4A22. RMSF of the protein with ligand in the course of MDSs shows flexible loops and SRSs, which are building blocks of channels in CYP4A22.

At 2555 ps, in channel-2 of the CYP4A22–ERY complex, the distance between key residues (K94 ↔ F122, at B/C loop (SRS-1) and ß2 sheet) increased from 0.91 nm to 1.46 nm (Fig. 8E). Excluding B/C loop motion, we found that the N-terminal of the I helix (SRS-4) unfolded into a loop (Fig. 8B). Interestingly, helices, which displayed large deviation in the CYP4A22–AA complex, also deviated in the CYP4A22–ERY complex. The average RMSF plot also suggests the same and illustrated clearly the high deviation on B/C loop and N-terminal region of the I helix (Fig. 9). In contrast to other channels, channel-3 in the CYP4A22–ERY complex shows a decrease in the distance between their key residues (C230 ↔ L258) at the channel gateway. The lowest distance observed was at 2025 ps (Fig. 8F). Surprisingly, the distance between the key residues of channel-3 reduced from 2.0 nm to 1.3 nm. The G helix was broken into two short helices and the N-terminal of the I helix was transformed into a loop (Fig. 8C). Thus, the results of the CYP4A22–ERY complex elucidate the conformational and structural changes on the channel-forming helices and their neighboring regions. These changes in the channel-forming regions of CYP4A22–ligand complexes suggest that the ligand binding may have influence on the regulation of channels with the help of SRSs.

4. Discussion

The present study has focused on constructing a stable CYP4A22 model, illustrating the binding site for AA and ERY, and predicting possible structural and conformational changes in

the channel-forming regions by utilizing various molecular modeling techniques. In the first part of this study, a reliable CYP4A22 model with a well-conserved core region was obtained by homology modeling and refined by MDS. In the next part, using MDS with energy calculation, we found the possibility of essential fatty acid and therapeutic drug binding in the highly hydrophobic pocket of CYP4A22 and the key residues interacting with the ligands. Finally, we have shown ligand-induced changes in the channel-forming regions of CYP4A22.

Interaction energy analysis of both the complexes implies that the van der Waals energy is a principal one to contribute in the CYP4A22–ligand interaction (Tables 1 and 2). This finding is in line with many CYP studies proposing a highly hydrophobic nature of the active site [29–32]. As the ligands are entirely different in structure, their mode of binding and the number of protein–ligand interactions varies greatly. These findings also coincide with the proposed results from a molecular modeling study of CYP3A4, which suggests that because of the larger size of ERY, it interacts with more active site residues than progesterone [33].

In order to anchor the anti-cancer drug (epipodophyllotoxin) in the active site, CYP3A4 used many residues including R105, S119, I120, F213 and A305 [34], which correspond to Y120, L132, L133, C230 and E321, respectively in CYP4A22. All the corresponding residues are in contact with docked ligands either by hydrophobic interactions or hydrogen bonding (Tables 1 and 2 and Fig. 7: Y120, L133 for AA and L132, C230, E321 for ERY). The origin of the ligand-interacting residues of CYP4A22 showed that they were mostly from SRSs (CYP4A22–AA: SRS-1, 4–5 and CYP4A22–ERY: SRS-1–2, 4–6). It has been found to be in agreement with the results of the recent MDS studies of CYP2D6, which indicates that the active site for propranolol (beta blocker used for the treatment of hypertension) is formed by SRS-1–2, 4–6 [35,36]. In CYP4A22, AA and ERY are surrounded mostly by the SRS residues that are highly hydrophobic in nature. The involvement of corresponding residues (Y120, L133, L132, C230, and E321) in the formation of active site supports that the identified binding sites could be the most potential interacting sites for these ligands.

Several CYPs studies proposed that SRSs are the major building blocks for channels including one of the most recent MDS studies of CYP3A4 [28]. In CYP4A22–AA complex, F122 plays a vital role in both channel-1 and -2, and forms a hydrophobic interaction with AA (Fig. 8D and E). On the other hand, four residues (G118, Y120, F122 and L133) from the B/C loop and two residues (E321 and D324) from the I helix interact with AA (Table 1). In the CYP4A22–ERY complex, two residues (A124 and L132) from the B/C loop and two residues (A313 and D324) from the I helix are in contact with ERY (Table 2). The B/C loop forms a channel in several CYPs including CYP3A4. Moreover, it was reported that channel-3 of CYP3A4 forms between helices F′, G′ and I [28]. Thus, the identified ligand-interacting residues here signify that they may have a possible dual role to play in ligand anchoring and passage.

Equivalents of a few ligand-interacting residues of CYP4A22 have been involved in various critical roles in other CYPs. A phylogenetically conserved F393 in CYPBM3 plays a vital role in controlling the reaction of the Fe atom of heme [37]. Interestingly, the corresponding residue in CYP4A22 (F450) from the signature motif (FxxGxxxCxG) forms a hydrophobic interaction with ERY (Table 2) and is located near Fe. To preserve the F450 in place, it is sandwiched between P449 and Q461. Surprisingly, this arrangement in CYP4A22 is identical with CYPBM3 (P392–F393–Q403). In addition, the P392 of CYPBM3 participates in the electron-transfer pathway [38]. The corresponding residue in CYP4A22, P449 is involved in a hydrophobic interaction with ERY. In various CYPs, the propionates of the heme (3A4 (R440) [29], eryF (R293), BM3 (R398) and cam (R299) [39]) are stabilized by interacting with the arginine residue. R455 has been conserved in CYP4A22 and found to form hydrogen bonds with heme propionates (Fig. 5B), and interacts with AA (Table 1). In the CYP3A4 carbamazepine epoxidation mechanism, F304 and A305 were located at the nearest position to the Fe atom and it is noted that they hold the substrate near the heme Fe atom for the reaction [40]. E321 of CYP4A22 is corresponding to A305 in CYP3A4. It is also conserved in several CYP4 family members and recognized to ligate with heme macrocycle [41]. This acidic residue in CYP4A22 is found to be located in close proximity to the heme Fe atom with its associated F320 and forms hydrogen bonds with ERY and the heme macrocycle (Fig. 7B). Altogether, these findings imply that the ligand-interacting residues of CYP4A22 may also play a significant role in supporting the enzyme reaction center for hydroxylation.

Structural analyses of protein–ligand complexes clearly illustrated the ligand-induced conformational changes in the channel-forming regions. Moreover, the distance between channel key residues in both the ligand-bound CYP4A22 complexes varies from 0.8 nm to 2 nm during MDS (Fig. 8). These results indicate that the long aliphatic chain molecules like AA and large molecules like ERY might bring more structural changes in the channel regions of CYP4A22. Recent crystal data for CYP3A4 supports our result, in which it displayed dramatic conformational changes due to the binding of ERY [42]. Likewise, induction of conformational changes upon ligand binding has been reported in many CYP family members such as CYP 2B4 [23,43], 2C5 [44], 2C9 [45], and 3A4. These findings also support our speculation.

The study of CYP family 4 established that the CYP4F11 model could provide a more open access channel for ERY than CYP4F3A [7]. It appears that to attain the larger passage in CYP4A22, these ligands might have been inducing large conformational changes on channel regions via SRSs. This speculation is clearly supported by studies of several CYPs, where conformational changes, unfolding of helices and large movements of channel-forming elements regulate the opening or closing of the ligand routes [46–48]. Thus, the observed conformational and/or structural changes (helix to coil) of this protein upon ligand binding suggest that they could play a major role in ligand-channeling.

5. Conclusions

To the best of our knowledge, a first molecular modeling report of orphan human protein CYP4A22 has been presented here based on the reliable built homology model and MDS study in complexes with AA and ERY. This study demonstrates the intermolecular interactions and spatial arrangement of binding pockets for AA and ERY. The illustrated ligand-binding mode and architecture of the active site could provide invaluable information to design an activator or inhibitor for CYP4A22. Interacting key residues infer the substrate specificity of the CYP4A22 active site when the ligands have different physical properties like molecular geometry, hydrophobicity (AA) and hydrophilicity (ERY). Furthermore, these residues elucidate that they might also play other critical roles, mainly in ligand-channeling and could help the reaction center of the protein. Thus, the key residues that are unveiled from the present study could be potential candidates for site directed mutagenesis studies. Structural and conformational changes of possible channels in CYP4A22, appear that, they originated from the active site upon ligand binding and transferred to the rest of the regions through neighboring SRSs and thereby could regulate the opening and closing of channels.

Acknowledgements

The authors are extremely gratifying to receive critical suggestions from Mr. Sundarapandian Thangapandian and Miss. Shalini John (GNU, Korea) during this manuscript preparation. All

students were supported by a scholarship from the BK21 Program, the Ministry of Education, Science and Technology, Korea and this work was supported by grants from the MEST/NRF for the Environmental Biotechnology National Core Research Center (grant #: 20090091489).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2009.11.010.

References

[1] P. Ortiz de Montellano, Cytochrome P450 Structure, Mechanism, and Biochemistry, Plenum Press, New York, 1995.
[2] D.R. Nelson, L. Koymans, T. Kamataki, J.J. Stegeman, R. Feyereisen, D.J. Waxman, M.R. Waterman, O. Gotoh, M.J. Coon, R.W. Estabrook, I.C. Gunsalus, D.W. Nebert, P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature, Pharmacogenetics 6 (1996) 1–42.
[3] D.A. Smith, B.C. Jones, Speculations on the substrate structure-activity relationship (SSAR) of cytochrome P450 enzymes, Biochem. Pharmacol. 44 (1992) 2089–2098.
[4] F.P. Guengerich, Z.L. Wu, C.J. Bartleson, Function of human cytochrome P450s: characterization of the remaining orphans, Biochem. Biophys. Res. Commun. 338 (2005) 465–469.
[5] A. Bellamine, Y. Wang, M.R. Waterman, J.V. Gainer 3rd, E.P. Dawson, N.J. Brown, J.H. Capdevila, Characterization of the CYP4A11 gene, a second CYP4A gene in humans, Arch. Biochem. Biophys. 409 (2003) 221–227.
[6] M.H. Hsu, U. Savas, K.J. Griffin, E.F. Johnson, Human cytochrome P450 family 4 enzymes: function, genetic variation and regulation, Drug Metab. Rev. 39 (2007) 515–538.
[7] A. Kalsotra, C.M. Turman, Y. Kikuta, H.W. Strobela, Expression and characterization of human cytochrome P450 4F11: putative role in the metabolism of therapeutic drugs and eicosanoids, Toxicol. Appl. Pharmacol. 199 (2004) 295–304.
[8] Discovery Studio, 2.0 User Guide, 2005, Accelrys Inc., San Diego, CA, USA.
[9] 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.
[10] H.M. Girvan, H.E. Seward, H.S. Toogood, M.R. Cheesman, D. Leys, A.W. Munro, Structural and spectroscopic characterization of P450 BM3 mutants with unprecedented P450 heme iron ligand sets. New heme ligation states influence conformational equilibria in P450 BM3, J. Biol. Chem. 282 (2007) 564–572.
[11] N. Strushkevich, S.A. Usanov, A.N. Plotnikov, G. Jones, H.W. Park, Structural analysis of CYP2R1 in complex with vitamin D3, J. Mol. Biol. 380 (2008) 95–106.
[12] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: fast, flexible and free, J. Comput. Chem. 26 (2005) 1701–1718.
[13] W.F. van Gunsteren, S.R. Billeter, A.A. Eising, P.H. Hunenberger, P. Kuger, A.E. Mark, W.R.P. Scott, I.G. Tironi, The GROMOS96 Manual and User Guide, Biomos, Zurich, Switzerland, 1996.
[14] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans, Interaction models for water in relation to protein hydration, in: B. Pullman, D. Reidel (Eds.), Intermolecular Forces, Reidel Publishing Company, Dordrecht, 1981, pp. 331–342.
[15] B. Hess, H. Bekker, J. Fraaije, H.J.C. Berendsen, A linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[16] S. Miyamoto, P.A. Kollman, SETTLE—an analytical version of the shake and rattle algorithm for rigid water models, J. Comput. Chem. 13 (1992) 952–962.
[17] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[18] M.J. Sippl, Recognition of errors in three-dimensional structures of proteins, Proteins 17 (1993) 355–362.
[19] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Cryst. 26 (1993) 283–291.
[20] A.W. Schuettelkopf, D.M.F. van Aalten, PRODRG—a tool for high-throughput crystallography of protein–ligand complexes, Acta Crystallogr. D 60 (2004) 1355–1363.
[21] D.F. Lewis, E. Watson, B.G. Lake, Evolution of the cytochrome P450 superfamily: sequence alignments and pharmacogenetics, Mutat. Res. 410 (1998) 245–270.
[22] 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.
[23] E.E. Scott, Y.A. He, M.R. Wester, M.A. White, C.C. Chin, J.R. Halpert, E.F. Johnson, C.D. Stout, An open conformation of mammalian cytochrome P450 2B4 at 1.6-Å resolution, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 13196–13201.
[24] L.M. Podust, T.L. Poulos, M.R. Waterman, Crystal structure of cytochrome P450 14alpha-sterol demethylase (CYP51) from *Mycobacterium tuberculosis* in complex with azole inhibitors, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 3068–3073.
[25] M.D. Paulsen, R.L. Ornstein, Dramatic differences in the motions of the mouth of open and closed cytochrome P450BM3 by molecular dynamics simulations, Proteins 21 (1995) 237–243.
[26] Y.T. Chang, G.H. Loew, Molecular dynamics simulations of P450 BM3-examination of substrate-induced conformational change, J. Biomol. Struct. Dyn. 16 (1999) 1189–1203.
[27] S.K. Ludemann, V. Lounnas, R.C. Wade, How do substrates enter and products exit the buried active site of cytochrome P450cam? 2. Steered molecular dynamics and adiabatic mapping of substrate pathways, J. Mol. Biol. 303 (2000) 813–830.
[28] W. Li, H. Liu, X. Luo, W. Zhu, Y. Tang, J.R. Halpert, H. Jiang, Possible pathway(s) of metyrapone egress from the active site of cytochrome P450 3A4: a molecular dynamics simulation, Drug Metab. Dispos. 35 (2007) 689–696.
[29] P.A. Williams, J. Cosme, D.M. Vinkovic, A. Ward, H.C. Angove, P.J. Day, C. Vonrhein, I.J. Tickle, H. Jhoti, Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone, Science 305 (2004) 683–686.
[30] S. Masuda, D.E. Prosser, Y.D. Guo, M. Kaufmann, G. Jones, Generation of a homology model for the human cytochrome P450, CYP24A1, and the testing of putative substrate binding residues by site-directed mutagenesis and enzyme activity studies, Arch. Biochem. Biophys. 460 (2007) 177–191.
[31] N.E. Hadjokas, R. Dai, F.K. Friedman, M.J. Spence, B.J. Cusack, R.E. Vestal, Y. Ma, Arginine to lysine 108 substitution in recombinant CYP1A2 abolishes methoxyresorun metabolism in lymphoblastoid cells, Br. J. Pharmacol. 136 (2002) 347–352.
[32] J.K. Yano, T.T. Denton, M.A. Cerny, X. Zhang, E.F. Johnson, J.R. Cashman, Synthetic inhibitors of cytochrome P450 2A6: inhibitory activity, difference spectra, mechanism of inhibition, and protein cocrystallization, J. Med. Chem. 49 (2006) 6987–7001.
[33] S. Wanchana, F. Yamashita, M. Hashida, QSAR analysis of the inhibition of recombinant CYP3A4 activity by structurally diverse compounds using a genetic algorithm-combined partial least squares method, Pharm. Res. 20 (2003) 1401–1408.
[34] M.K. Julsing, N.P. Vasilev, D. Schneidman-Duhovny, R. Muntendam, H.J. Woerdenbag, W.J. Quax, H.J. Wolfson, I. Ionkova, O. Kayser, Metabolic stereoselectivity of cytochrome P450 3A4 towards deoxypodophyllotoxin: in silico predictions and experimental validation, Eur. J. Med. Chem. 43 (2008) 1171–1179.
[35] C. de Graaf, C. Oostenbrink, P.H.J. Keizers, B.M.A. van Vugt Lussenburg, R.A.B. van Waterschoot, R.A. Tschirret-Guth, J.N.M. Commandeur, N.P.E. Vermeulen, Molecular modeling-guided site-directed mutagenesis of cytochrome P450 2D6, Curr. Drug. Metab. 8 (2007) 59–77.
[36] Y. Ito, H. Kondo, P.S. Goldfarb, D.F.V. Lewis, Analysis of CYP2D6 substrate interactions by computational methods, J. Mol. Graph. Model. 26 (2008) 947–956.
[37] T.W. Ost, A.W. Munro, C.G. Mowat, P.R. Taylor, A. Pesseguiero, A.J. Fulco, A.K. Cho, M.A. Cheesman, M.D. Walkinshaw, S.K. Chapman, Structural and spectroscopic analysis of the F393H mutant of flavocytochrome P450 BM3, Biochemistry 40 (2001) 13430–13438.
[38] F. Sevrioukova, H.Y. Li, H. Zhang, J.A. Peterson, T.L. Poulos, Structure of a cytochrome P450-redox partner electron-transfer complex, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 1863–1868.
[39] G.M. Keseru, B. Volk, G.T. Balogh, Cytochrome P450 catalyzed nitric oxide synthesis: a theoretical study, J. Biomol. Struct. Dyn. 17 (2000) 759–767.
[40] M. Hata, Y. Tanaka, N. Kyoda, T. Osakabe, H. Yuki, I. Ishii, M. Kitada, S. Neya, T. Hoshino, An epoxidation mechanism of carbamazepine by CYP3A4, Bioorg. Med. Chem. 16 (2008) 5134–5148.
[41] C. Colas, P.R. Ortiz de Montellano, Autocatalytic radical reactions in physiological prosthetic heme modification, Chem. Rev. 103 (2003) 2305–2332.
[42] M. Ekroos, T. Sjogren, Structural basis for ligand promiscuity in cytochrome P450 3A4, Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 13682–13687.
[43] E.E. Scott, M.A. White, Y.A. He, E.F. Johnson, C.D. Stout, J.R. Halpert, Structure of mammalian cytochrome P450 2B4 complexed with 4-(4-chlorophenyl)imidazole at 1.9-Å resolution: insight into the range of P450 conformations and the coordination of redox partner binding, J. Biol. Chem. 279 (2004) 27294–27301.
[44] M.R. Wester, E.F. Johnson, C. Marques-Soares, P.M. Dansette, D. Mansuy, C.D. Stout, Structure of a substrate complex of mammalian cytochrome P450 2C5 at 2.3-Å resolution: evidence for multiple substrate binding modes, Biochemistry 42 (2003) 6370–6379.
[45] P.A. Williams, J. Cosme, A. Ward, H.C. Angove, D. Matak Vinkovic, H. Jhoti, Crystal structure of human cytochrome P450 2C9 with bound warfarin, Nature 424 (2003) 464–468.
[46] H. Li, T.L. Poulos, Conformational dynamics in cytochrome P450-substrate interactions, Biochimie 78 (1996) 695–699.
[47] H. Li, T.L. Poulos, The structure of the cytochrome P450BM-3 haem domain complexed with the fatty acid substrate, palmitoleic acid, Nat. Struct. Biol. 4 (1997) 140–146.
[48] J.K. Yano, L.S. Koo, D.J. Schuller, H. Li, P.R. Ortiz de Montellano, T.L. Poulos, Crystal structure of a thermophilic cytochrome P450 from the archaeon *Sulfolobus solfataricus*, J. Biol. Chem. 275 (2000) 31086–31092.