Journal of Molecular Graphics and Modelling 28 (2009) 170–176

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

Molecular modeling of human cytochrome P450 2W1 and its interactions with substrates

Weihua Li [a], Yun Tang [b], Tyuji Hoshino [a,*], Saburo Neya [a]

[a] Graduate School of Pharmaceutical Sciences, Chiba University, Chiba 263-8522, Japan
[b] School of Pharmacy, East China University of Science and Technology, Shanghai 200237, China

ARTICLE INFO

Article history:
Received 20 March 2009
Received in revised form 10 June 2009
Accepted 13 June 2009
Available online 21 June 2009

Keywords:
CYP2W1
Homology modeling
Molecular docking
Molecular dynamics simulation
MM-GBSA

ABSTRACT

The human cytochrome P450 2W1 (CYP2W1) was categorized into the so-called "orphan" CYPs because of its unknown enzymatic function. However, recent studies showed that the recombinant CYP2W1 exhibited broad catalytic activity towards several chemicals. Furthermore, this enzyme was selectively expressed in some forms of cancers, whereas a very low expression was found in human normal issues. These render CYP2W1 as a potential drug target for cancer therapy. At present, however, little information is available on the active site topology and the substrate binding modes of CYP2W1. In this study, the three-dimensional model of CYP2W1 was constructed using the homology modeling method. Two known substrates, benzphetamine and indole, were then docked into the active site, and refined by molecular dynamics simulations. The interaction energy between the substrates and the enzyme was calculated and analyzed by using the MM-GBSA method. The results indicated that the constructed CYP2W1 model can account for the regioselectivity of this enzyme towards the known substrates and van der Waals interactions were the driving force for the substrate binding. Several key residues were identified to be responsible for the binding of indole and benzphetamine with CYP2W1. These findings provide useful information for the detailed characterization of the biological roles of CYP2W1 and structure-based drug design of this enzyme.

1. Introduction

Cytochromes P450 (CYP) are a superfamily of heme-containing enzymes, which play an important role not only in the biosynthesis and metabolism of many endogenous compounds like hormones and vitamins, but also in the metabolism and detoxification of a large number of xenobiotic chemicals such as drugs, carcinogens, and environmental pollutants [1]. There are 57 CYP isoenzymes found in humans. Most of human CYPs exhibit specific physiological functions, whereas some others presently have no assigned function and thus were ascribed to "orphan" CYPs [2].

CYP2W1 was considered as one of such "orphan" CYPs until recently. Guengerich and his coworkers recently reported that the recombinant human CYP2W1 exhibited broad catalytic activity towards several chemicals [3]. Their results demonstrated that CYP2W1 can catalyze benzphetamine *N*-demethylation (K_{cat}, 3.8 min^{-1}) and arachionic acid oxidation at a very low rate (~0.05 min^{-1}). It was also shown that CYP2W1 was able to catalyze activation of several procarcinogens, particularly polycyclic hydrocarbon diols [3]. Another group reported that CYP2W1 exhibited catalytic activity towards indole, 3-methylindole and chlozoxazone [4]. Further analysis suggested that the metabolite of indole by CYP2W1 was oxindole or isatin. Karlgren et al. [5] recently found that CYP2W1 was specifically expressed in human tumors and a very low expression of this enzyme was detected in normal adult tissues. For this reason, human CYP2W1 has been suggested to be a potential drug target for cancer therapy [6]. At present, however, how CYP2W1 binds and oxidizes substrates remains unclear. In addition, little information is available on the active site topology of CYP2W1. These have seriously hampered further characterization of biological roles and structure-based drug design of CYP2W1 for the gene-directed enzyme prodrug therapy (GDEPT) [6]. A detailed understanding of these questions requires accurate three-dimensional (3D) structures of CYP2W1. However, currently the 3D structure of CYP2W1 has not yet been determined by experimental approaches.

Homology modeling combined with docking and molecular dynamics simulations has been successfully applied in the 3D model construction of CYPs and in the prediction and rationale of interactions between ligands and CYPs [7,8]. To date, crystal structures of several mammalian CYPs have been determined through deleting their transmembrane N-terminus [9]. The known CYP structures have common secondary structural elements and considerable similarity among them. Therefore, these available

* Corresponding author. Tel.: +81 43 2902926; fax: +86 43 2902925.
E-mail address: hoshi@p.chiba-u.ac.jp (T. Hoshino).

1093-3263/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2009.06.002

structures provide good templates for computational modeling to construct 3D models of the mammalian CYPs whose structures are unknown. In this study, we firstly constructed the 3D model of CYP2W1 by using the homology modeling method. After that, two known substrates, indole and benzphetamine, were docked into the active site of the enzyme, followed by molecular dynamics (MD) refinement to explore the substrate specificity. Finally, the interaction energy of the substrates with CYP2W1 was calculated and analyzed by using the molecular mechanics-generalized born surface area (MM-GBSA) method. The results obtained in this study will be useful for the further characterization of the biological roles of CYP2W1 and the structure-based drug design targeting this enzyme.

2. Materials and methods

2.1. Homology modeling of human CYP2W1

The amino acid sequence of human CYP2W1 was obtained from the SwissProt database (access number Q8TAV3). The two crystal structures of CYP2C5 in complex with a substrate were taken from the protein data bank (PDB, entry code 1N6R and 1NR6) [10,11] and used as the template for building the initial 3D model of human CYP2W1. The sequence alignment of CYP2W1 with the templates was accomplished using ClustalW 1.83 [12] with a gap penalty of 10 and BLOSUM series weight matrix.

The Modeller 9v2 program [13] was employed to generate the initial 3D models of CYP2W1. Modeller generates the 3D models by optimization of molecular probability density functions. The optimization process consists of applying the variable target function as well as conjugated gradients and MD with simulated annealing. A set of 20 models of CYP2W1 were produced based on the resulting alignment obtained above. The outcomes were ranked based on the internal scoring function of Modeller. The top 5 models with high scores were validated by the Procheck [14] procedure and further evaluated by docking to test their capacity of accommodating benzphetamine, a relatively large substrate. After above steps of validation, a model was finally chosen for further refinement by energy minimization. The energy minimization was performed by using the Amber9 package [15]. The initial 3D model was solvated in water and optimized by gradually decreasing the restraint (see the MD simulation section for the detailed optimization protocol).

The optimized model was subjected to quality assessment with respect to its geometry and energy and used for the subsequent molecular docking. The Procheck and Verify3D [16] were utilized for geometric evaluation. Verify3D was used to assess the compatibility of an atomic model with its own amino acid sequence. A high Verify3D profile score indicates the high quality of a protein model. The Prosa program [17] was employed to evaluate the quality of consistency between the native fold and the sequence and examine the energy of residue–residue interactions using a distance-based pair potential. The energy was transformed to a score called *z*-score. Residues with negative *z*-score indicate reasonable side-chain interactions.

2.2. Molecular docking

The structures of benzphetamine ((*S*)-(+)-benzphetamine) and indole were built using InsightII (Accelrys Inc.) and geometrically optimized at the B3LYP/6-31G** level using Gaussian03 (Gaussian Inc.).

The substrates were docked into the active site of the refined CYP2W1 structure using AutoDock version 4.0 [18]. The AutoDock Tools kit was used to prepare the required structures for docking. The Gasteiger charges were added to both the protein and substrates. The center of the grid was placed on the Fe atom and a box with dimensions of $80 \times 60 \times 60$ points was generated. One hundred solutions were generated for each docking with the default setting and clustered with the RMSD cutoff set to 2.0 Å. The solution that has the lowest energy in the individual cluster was chosen for further analysis. The resulting CYP2W1–substrate complexes were further refined by MD simulations.

2.3. MD simulations

The MD simulations on the substrate-free CYP2W1 model, the CYP2W1–indole complex, and the CYP2W1–benzphetamine complex were performed using the Amber9 package [15]. The AMBER ff03 all atom force field [19] was used for the protein and the general AMBER force field [20] was used as the parameters for benzphetamine and indole. The force field parameters for heme were taken from previous work [21,22]. Atom charges of the substrates were obtained based on the optimized structures with restrained electrostatic potential (RESP) fitting procedure implemented in Amber9. The hydrogen atoms were added to the initial model using the xleap module in Amber9. To release the bad steric collision of atoms, the structures were optimized by steepest descent method for 100 steps, followed by the conjugated gradient method for 300 steps. The protein was then solvated with water in a truncated octahedron periodic box. The TIP3P model [23] was used for water, and counterions were added to neutralize the system.

Prior to the production stage, the following equilibration protocol was applied. First, the solvent was relaxed for 3000 steps by restraining the protein atomic positions with a harmonic potential. Second, the backbone atoms were kept restrained to allow side-chain atoms and solvent to move for 3000 steps with a force constant of 10 kcal mol^{-1} A^{-2}, followed by another 3000 steps with a force constant of 5 kcal mol^{-1} A^{-2}. Lastly, all atoms were allowed to freely move for 3000 steps without any restraint. After minimization, each system was heated up to 300 K over 40 ps using the NVT ensemble and equilibrated at 300 K for 50 ps. Finally, a 1.5-ns of the unrestrained MD run for each system was conducted at 1 atm and 300 K with the NPT ensemble. In all simulations, the SHAKE algorithm [24] was used to constrain the bonds involving hydrogen atoms. The cutoff value of 10 Å was applied to treat the nonbonding interactions and a time step of 2 fs was used.

2.4. Substrate–CYP2W1 interaction energy calculation and decomposition

To obtain a detailed view of the interaction between the substrates and the protein residues and identify certain key residues responsible for the substrate binding, the interaction energy was calculated and then decomposed into the contribution on each individual residue by the MM-GBSA method [25]. The interaction energy was calculated according to the following equation:

$$\Delta G_{binding} = \Delta G + \Delta G_{ele} + \Delta G_{gb} \quad (1)$$

where ΔG and ΔG_{ele} represent nonbonded van der Waals and electrostatic interaction between the substrate and each residue, respectively. The solvation energy (ΔG_{gb}) is composed of the polar contribution and the nonpolar part. The polar contribution was calculated with modified GB model described by Onufriev et al. [26] by using of the bondi radii set. The interior dielectric constant was set to 1.0, and the outer dielectric constant was set to 80.0. The nonpolar part was estimated based on the solvent-accessible surface area using the LCPO method [27]. 50 snapshots from the last 500 ps of MD simulations were extracted for the binding

energy calculations. The other parameters used for the MM-GBSA calculations were the same as those used in our previous work [28].

3. Results and discussion

3.1. Construction of CYP2W1 model

The first important step by using homology modeling is to select an appropriate template structure for constructing the target model. To date, several crystal structures of mammalian P450s have been determined (http://www.pdb.org). These structures generally share common secondary structure elements in spite of their different origins. A Blastp search (www.ncbi.nih.gov/blast) confirmed that several mammalian CYPs structures could serve as the potential template for building the CYP2W1 model. CYP2W1 has the sequence identity about 40% with several structure-known CYPs, such as rabbit 2C5 (40%), rabbit 2B4 (40%), human 2A6 (38%), human 2D6 (42%), and human 2C9 (37%). The appropriate template was chosen based on sequence similarity, residue completeness, and crystal resolution. Among the available potential templates, the crystal structures of CYP2C5 in complex with a substrate (PDB code: 1N6R and 1NR6) [10,11] were eventually selected as the template structure to construct the CYP2W1 model in terms of the earlier criteria. Although human P450 2D6 [29] has the highest sequence identity with 2W1, the relatively low crystal resolution (3.0 Å) and missing residue segments render it unsuitable as the template for constructing the 2W1 model.

In the crystal structures of CYP2C5 complexes, the N-terminal trans-membrane segment (residues 1–26) was deleted in order to facilitate its crystallization. Since the N-terminus does not affect the substrate binding, the corresponding first 28 residues were not modeled in the CYP2W1 model. The sequence alignment between 2W1 and 2C5 was displayed in Fig. 1. The majority of the structure is considered conserved except for the limited gap inserts. The resulting alignment was used to be the input file for Modeller to generate the initial 3D model of CYP2W1 using the fast simulated annealing procedure implemented in the Modeller program.

The initial 3D model of CYP2W1 was energy-minimized to release the bad atomic contact and tune unreasonable local structural conformations. After that, the refined CYP2W1 model was subjected to the quality assessment from both geometric and energetic aspects. The Procheck evaluation results showed that the *G*-factors score was −0.10 for the built 2W1 model. An overall *G*-factors score with greater than −0.5 is generally accepted as an indicator for a good quality model. There are 90.4% of residues located in the most favored regions, 8.0% in allowed regions, 1.1% in generously allowed regions, and 0.5% (2 residues) in disallowed regions, as depicted by Ramchandran plot. Fig. 2A shows the Verify3D profile for the refined CYP2W1 model. It illustrates that 91% of residues of the 2W1 model have a score over 0.2, a value used as judgment for a good model. Several sequences between residue 227 and residue 232 exhibit low scores, which are from the G′–G loop. The Prosa2003 analysis showed that *z*-score in the 2W1 model was negative in most residues except several residues in the G′–G loop (Fig. 2B). Since this loop is located at the external region of the model and has little effect on the substrate binding, no attempts were made to further refine the segment. For comparison, the Verify3D scores and *z*-scores of the templates were also presented. It can be seen from Fig. 2, the overall results of CYP2W1 were quite similar as those of the templates. Taken together, the

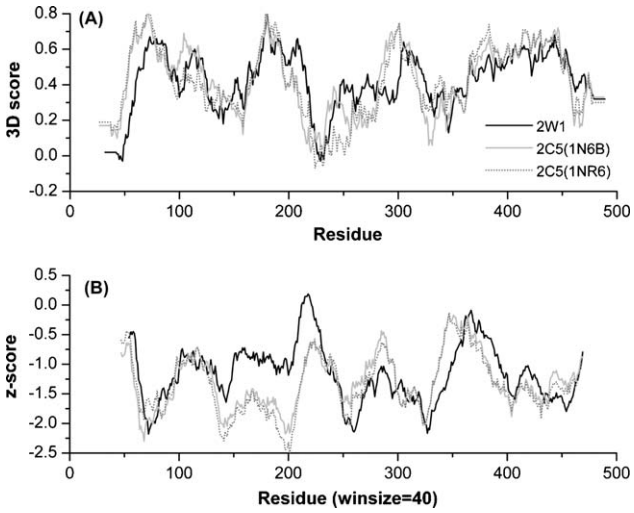

Fig. 2. (A) Verify3D score profile versus residue number; (B) Prosa *z*-score with respect to residue (winsize = 40). The scores of CYP2W1, 2C5 (PDB: 1N6R), and 2C5 (PDB: 1NR6) are shown in black solid lines, gray solid lines, and light gray dot lines, separately.

```
     .: **** *:*::**:  :  .: .:** ::** *******:** : **** *:*******.. *:*:*.* .:.*:: :
P2C5 GKLPPGPTPFPIIGNILQIDAKDISKSLTKFSECYGPVFTVYLGMKPTVVLHGYEAVKEALVDLGEEFAGRGSVPILEKV  106
P2W1 ARWPPGPRPLPLVGNLHLLRLSQQDRSLMELSERYGPVFTVHLGRQKTVVLTGFEAVKEALAGPGQELADRPPIAIFQLI  108

     .:* ** **..  *:  *:*:: :*:.:*:*:..: *:* :* :**  :*   .. *  *  :** **.*:  :::* .*
P2C5 SKGLGIAFSNAKTWKEMRRFSLMTLRNFGMGKRSIEDRIQEEARCLVEELRKTNASPCDPTFILGCAPCNVICSVIFHNR  186
P2W1 QRGGGIFFSSGARWRAARQFTVRALHSLGVGREPVADKILQELKCLSGQLDGYRGRP-FPLALLGWAPSNITFALLFGRR  187

     ***:*  *:.*:  ::* : :*.** **::* :* *   :  :*:.:*:: : :: ::   :: ::  :  .:*
P2C5 FDYKDEEFLKLMESLNENVRILSSPWLQVYNNFPALLDYFPGIHKTLLKNADYIKNFIMEKVKEHQKLLDVNNPRDFIDC  266
P2W1 FDYRDPVFVSLLGLIDEVMVLLGSPGLQLFNVYPWLGALLQ-LHRPVLRKIEEVRAILRTLLEARRPHVCPGDPVCSYVD  266

      **:  * ::    *:  . * .. *:. *******:**::: **: :**:* .*****::**:*  *:* ::*:. :**
P2C5 FLIKMEQENN--LEFTLESLVIAVSDLFGAGTETTSTTLRYSLLLLLKHPEVAARVQEEIERVIGRHRSPCMQDRSRMPY  344
P2W1 ALIQQGQGDDPEGLFAEANAVACTLDMVMAGTETTSATLQWAALLMGRHPDVQGRVQEELDRVLGPGRTPRLEDQQALPY  346

     *.**:**:**** *** ::*:..: *.:: .:::**** :*. ***** **. : .*  *:****** .*:* * : *:*
P2C5 TDAVIHEIQRFIDLLPTNLPHAVTRDVRFRNYFIPKGTDIITSLTSVLHDEKAFPNPKVFDPGHFLDESGNFKKSDYFMP  424
P2W1 TSAVLHEVQRFITLLP-HVPRCTAADTQLGGFLLPKGTPVIPLLTSVLLDETQWQTPGQFNPGHFLDANGHFVKREAFLP  425

     ****:*:**** *** ****:::.:**.::* .   *.* .** *..  .*.  * :  ** :*
P2C5 FSAGKRMCVGEGLARMELFLFLTSILQNFKLQS--LVEPKDLDITAVVNGFVSVPPSYQLCFIPI    487
P2W1 FSAGRRVCVGERLARTELFLLFAGLLQRYRLLPPPGVSPASLDTTPAR-AFTMRPRAQALCAVPR    489
```

Fig. 1. Sequence alignment of CYP2W1 with CYP2C5 from residues 29 to 489. The asterisk indicates an identical or conserved residue; a colon indicates a conserved substitution; a dot indicates a semi-conserved substitution.

built 3D model of CYP2W1 was of reasonable quality compared to the crystal structures of CYP2C5. This energy-minimized model was used for subsequent structural analysis and docking simulations.

3.2. Structural analysis of CYP2W1 model

As expected, the 3D structural model of CYP2W1 is overall similar to those of other CYPs reported previously, and exhibits a closed conformational form with the active site deeply buried into the center of structural fold, as shown in Fig. 3. The coordinates in PDB format of CYP2W1 model have been provided in the supporting material. The heme group is sandwiched between helices I and L (the secondary structure nomenclature is referred to the template 2C5). The active site above heme is mainly composed of hydrophobic residues. The phenyl rings of four hydrophobic phenylalanines (Phe105, Phe115, Phe116 and Phe475) point to the active site. Residues Phe105, Phe475, Met206 and Met327 constitute the roof of the active site. Many available mammalian CYP structures reveal that a highly conserved threonine is located in the middle of helix I. This conserved residue has been suggested to participate in the proton delivery and plays an important role in the dioxygen bond cleavage during the catalytic cycle [30]. A threonine, Thr300, is found to reside at the same position in CYP2W1. Thr300 forms a hydrogen bond with its adjacent Ala296. The hydrogen bond was stable during the 1.5-ns MD simulation, as shown in Fig. S1 of supporting material. This hydrogen bonding is also observed in other mammalian CYP structures, such as 2C5, 2C8, 2D6, 2E1 and 3A4.

An intriguing question about CYPs is how a ligand enters or exits the active site, which is deeply buried at the center of fold. Previous evidence from both crystal structures and theoretic simulations suggests that different CYPs have different ligand access channels [31–33]. For a given CYP, there might exist multiple channels for ligand passage. To identify possible substrate access channels in CYP2W1, we utilized the MOLE program [34] to search for the pathways. The detailed methodology behind this program was described in the literature [34]. Two possible pathways were found from the active site to the external surface in CYP2W1 (supporting material, Fig. S2). They correspond to pw2e and the solvent channel named by Wade and coworkers [31]. The pw2e is penetrating through the B′ helix/B–C loop. This channel was also suggested to be the possible substrate channel in CYP3A4 [32] and CYP51 [35]. The solvent channel is located between helix I, helix F and the loop preceding C-terminus. The question of which channel can be used for substrate passage warrants further investigation by other methods such as steered molecular dynamics simulation [32] and/or random accelerated molecular dynamics simulation [36].

Fig. 3. Ribbon schematic representations of the homology model of CYP2W1. Heme is shown as red stick. Major helices are labeled. Image was generated using PyMOL (http://www.pymol.org).

3.3. Binding modes of substrates with CYP2W1

The homology model of CYP2W1 was further evaluated by testing its capacity of accommodating the known substrates. To date, only two substrates, namely benzphetamine and indole, have their biochemical experimental data available for CYP2W1 [3,4]. Hence, these two known substrates were docked into the active site of the enzyme to explore the substrate binding modes. The initial binding modes were obtained by using AutoDock. The obtained complex models were then subjected to MD simulations, which has two purposes. One is to refine the substrate binding modes, since the docking does not take into account flexibility of the protein; the other is to examine the stability of the homology model, which is achieved by measuring the root-mean-square deviation (RMSD) of the protein backbone atoms with respect to the starting structure. For the purpose of comparison, the substrate-free CYP2W1 model was also performed by MD simulation. Fig. 4 shows the RMSD value variation with respect to the simulation time. The RMSD value of the indole–2W1 complex has a low fluctuation during the entire time process. For the two benzphetamine–2W1 complexes and the substrate-free system, the RMSD values display an increase in the first 300 ps. After 300 ps, the RSMD values slightly fluctuate at ~2.0 Å until 1.5 ns. Although the RMSD value in the benzphetamine–2W1 complex (binding mode 2) has a slight increase from 0.8 to 1.1 ns, it converges to about 2.0 Å after ~1.2 ns. The results suggested that the built 3D models were stable during the MD simulation. The average structures from the last 500 ps were used for the binding mode analysis.

Benzphetamine is an anti-obesity drug on the market and often used as the model substrate to probe the CYPs' function. It was reported that benzphetamine was metabolized by CYPs to several products via three different pathways: N-demethylation, aromatic hydroxylation, and N-debenzylation [37]. The major metabolite in human is 1-(*p*-hydroxyphenyl)-2-(*N*-benzylamino)propane, which is produced via aromatic hydroxylation (ring B hydroxylation) [38]. The recent study demonstrated that benzphetamine is also one of substrates of CYP2W1 [3]. The docking results showed that two distinct clusters were found for benzphetamine orientations in the active site of CYP2W1, which correspond to N-demethylation (binding mode 1) and aromatic hydroxylation (binding mode 2). The average binding free energies calculated by AutoDock for these two binding modes were −5.99 kcal/mol (binding mode 1) and

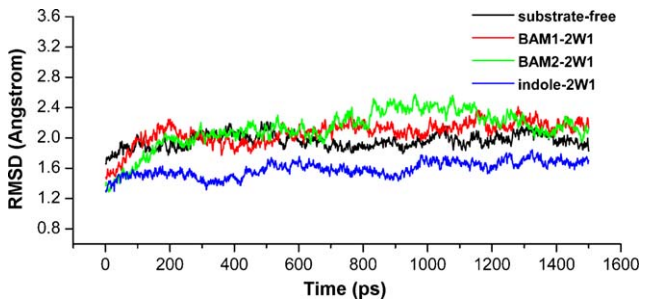

Fig. 4. RMSD variations of substrate-free CYP2W1 (black), BAM1–CYP2W1 (red), BAM2–CYP2W1 (green), and indole–CYP2W1 (blue) systems with respect to the homology model as a function of simulation time. BAM1 and BAM2 are the abbreviation for binding mode 1 and binding mode 2 of benzphetamine in the CYP2W1 active site.

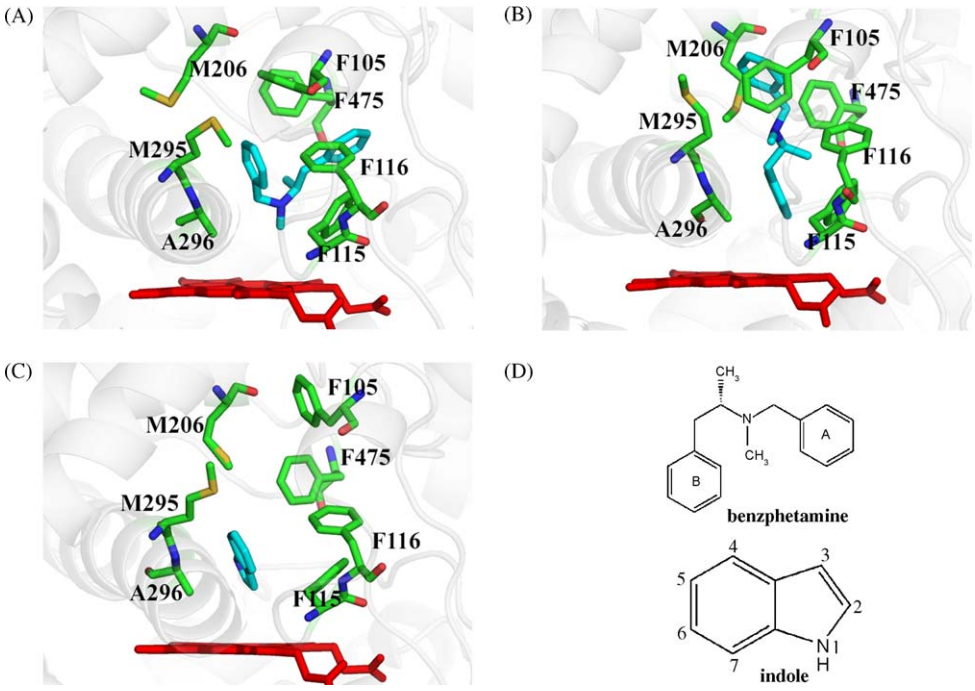

Fig. 5. A close view of the average structures of refined CYP2W1 with benzphetamine ((A): binding mode 1; (B): binding mode 2) and indole (C). The 2D structures of benzphetamine and indole are shown in (D). Heme is shown as red stick. Key residues are represented with green stick and substrates with cyan stick.

−6.65 kcal/mol (binding mode 2) respectively, which implies that the binding mode 2 is more stable than binding mode 1. Accordingly, both CYP2W1–benzphetamine complexes were subjected to MD refinement.

The average structures of CYP2W1–benzphetamine complexes after MD refinement are shown in Fig. 5A and B. Benzphetamine is stabilized by hydrophobic interactions in the active site. This is in line with the fact that the active site of CYP2W1 is mainly constituted by nonpolar residues. In the binding mode 1 of benzphetamine, the methyl group linked to the nitrogen atom points to the Fe atom, which facilitates *N*-demethylation of benzphetamine. The average distance between the methyl carbon atom and the Fe atom is about 4.6 Å after the substrate reaches a stable binding mode, as shown in Fig. 6. The phenyl ring A of benzphetamine in binding mode 1 is paralleling with helix I to form hydrophobic interactions with residues in the middle of helix I, whereas the phenyl ring B is sandwiched between Phe115 and Phe475, as depicted in Fig. 5A. In the binding mode 2, benzphetamine uses its phenyl ring B pointing to the Fe atom, which facilitates the aromatic hydroxylation. The phenyl ring A of benzphetamine extends to the roof of the active site and forms edge-to-face interactions with Phe105, as shown in Fig. 5B. We monitored the distance between the carbon atom at the para position of ring B and the Fe atom during the MD simulation, as shown in Fig. 6. The average distance maintained less than 6 Å, a value potentially for carbon oxidation by CYPs [39].

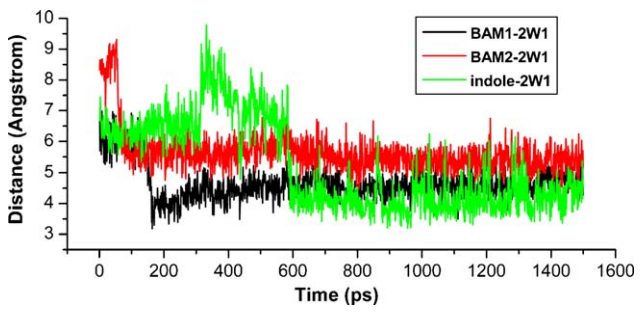

Fig. 6. Distances between Fe and the methyl C of benzphetamine (BAM1–2W1), between Fe and C4 on ring B of benzphetamine (BAM2–2W1), and between Fe and C2 of indole (indole–2W1) with respect to simulation time during MD simulations. BAM1 and BAM2 are the abbreviation for binding mode 1 and binding mode 2 of benzphetamine in the CYP2W1 active site.

Indole was known to be metabolized by several P450s, such as CYP2A6, 2C19, and 2E1 [40]. The C-2, C-3 and C-6 are the major oxidation sites by P450s. A recent study suggested that indole was also a substrate for CYP2W1 [4]. The major metabolite of indole by CYP2W1 was oxindole and/or isatin. The 3D conformation of indole bound to CYP2W1 is shown in Fig. 5C. The ring plane of indole is paralleling to helix I and forms hydrophobic interactions with the residues in the middle of helix I. The similar binding mode was also observed in the recently resolved crystal structure of indole–CYP2A13 [41]. The C-2 and C-3 atoms of indole point to the Fe atom, which facilitates to produce the major metabolites. The distance between C-2 and Fe was monitored as shown in Fig. 6. At ∼600 ps, the distance value has a jump, which indicates the substrate adjusted its orientation during the first 600 ps of the MD simulation. After this point, the distance fluctuates around ∼4 Å, a favorable value for carbon oxidation by CYPs. Compared with the distances of benzphetamine complexes, the distance value of indole with Fe exhibits a relatively large fluctuation. This may attribute to the small size of indole. On the other hand, it may account for the lower catalytic activity of CYP2W1 towards indole when compared to CYP2A6 [4].

3.4. Important residues contributing to substrate binding

To detect certain key residues responsible for substrate binding, the interaction energies between the substrates and CYP2W1 were calculated and decomposed on the individual residue using the MM-GBSA method. Table 1 shows the calculated binding energies including the total energies and the individual energy components. Seen from Table 1, van der Waals energy has the major contribution to the binding energy and is the major driving force

Table 1
Interaction energies of CYP2W1 complexed with substrates.

CYP2W1	ΔG_{ele}	ΔG	$\Delta G_{nonp,sol}$	$\Delta G_{ele,sol}$	$\Delta G^{cal}_{binding}$
BAM1[a]	−3.12	−40.53	−4.27	12.81	−35.12
BAM2	−0.64	−40.33	−4.80	10.47	−35.30
Indole	−3.29	−20.02	−2.36	12.88	−12.78

ΔG_{ele}, electrostatic interaction energy; ΔG, van der Waals interaction energy; $\Delta G_{nonp,sol}$, nonpolar contribution to solvation energy; $\Delta G_{ele,sol}$, polar contribution to solvation energy; $\Delta G^{cal}_{binding}$, calculated interaction energy.

[a] BAM is an abbreviation for benzphetamine. BAM1 and BAM2 indicate the binding modes 1 and 2 of benzphetamine in the CYP2W1 active site, respectively.

for the substrate binding. This is consistent with the fact that the active site of CYP2W1 is primarily composed of hydrophobic residues, especially the aromatic resides. The calculated binding energies of CYP2W1 with benzphetamine are −35.12 kcal/mol (binding mode 1) and −35.30 kcal/mol (binding mode 2), respectively. Both energies are lower than that of indole–2W1 complex, −12.78 kcal/mol. This is also in agreement with the experimental data. Previous experimental data [3,4] showed that the K_m value of indole is about 12-fold higher than that of benzphetamine. In general, the larger the apparent K_m is, the weaker the affinity of enzyme has for the substrate.

Table 2 lists the residues which have the major contribution to the binding energies (less than −0.1 kcal/mol). Most of the residues are located in the six substrate recognition sites (*SRS*s), defined by Gotoh [42]. These residues include Arg99, Ile102, Phe105, Ile114, Phe115, and Phe116 in the *SRS*-1; Met206, Leu209, and Le214 in the *SRS*-2; Val294, Met295 Ala296, Glu299, and Thr300 in the *SRS*-4; Leu361 and Val364 in the *SRS*-5; Phe475 and Thr476 in the *SRS*-6. Some of residues identified by current MM-GBSA calculations have been shown to affect the substrate specificities and/or the catalytic activity in the corresponding alignment positions of other CYPs. For examples, Phe120 in CYP2D6, corresponding to Phe115 of 2W1, was found to influence the selectivity and regiospecificity in substrate binding and catalysis [43]. Phe475 corresponds to Phe476 at the same position in 2C9. Phe476 of 2C9 was shown not only to guide substrate binding [44], but also to alter the kinetic parameters towards several substrates when mutated by other residues [45]. Recently, the MM-GBSA method has been applied to successfully predict the critical residues responsible for HIV drugs binding [46,47]. Based on these lines of evidence, the CYP2W1 residues identified in the current study are expected to affect the binding of indole and benzphetamine and/or catalytic activity of the enzyme towards these two substrates.

Table 2
Energy contributions of key residues to the binding energies.

Residue	ΔG_{bing} (kcal/mol)		
	BAM1[a]	BAM2	Indole
Arg99	−0.49	−0.43	−0.16
Ile102	−0.15	−0.26	−0.01
Phe105	−0.26	−1.0	0.00
Ile114	−0.19	−0.15	−0.10
Phe115	−2.36	−1.47	−0.67
Phe116	−0.56	−0.73	−0.06
Met206	−0.14	−1.54	−0.16
Leu209	−0.88	−1.03	−0.04
Leu214	−0.37	−1.04	−0.04
Val294	−0.23	−0.32	−0.13
Met295	−1.78	−2.23	−1.04
Ala296	−1.25	−1.20	−1.32
Glu299	−0.42	−0.08	−0.38
Thr300	−0.47	−0.70	−0.67
Leu361	−0.69	−0.39	−0.58
Val364	−1.54	−0.89	−0.15
Phe475	−1.95	−1.04	−0.52
Thr476	−0.56	−0.11	−0.31
Heme	−1.80	−1.46	−2.00

[a] BAM is an abbreviation for benzphetamine.

4. Conclusions

Previous studies suggested that CYP2W1 could be a potential drug target for cancer therapy. Lack of the structural information about this enzyme hinders the detailed characterization of its detailed biological functions and its application in the structure-based drug design. In this study, the 3D model of human cytochrome CYP2W1 was constructed by using homology modeling. To provide useful information to characterize the enzyme's function, two known substrates, benzphetamine and indole, were docked into the active site and then refined by MD simulations to explore the favorable binding modes. Two benzphetamine–CYP2W1 complexes were obtained, which facilitate *N*-demethylation and B-ring aromatic hydroxylation of benzphetamine, respectively. The binding modes of benzphetamine with CYP2W1 obtained in this study are in agreement with the regioselectivity of CYP2W1 and other P450s towards the substrate. In the indole–CYP2W1 complex, the C-2 and C-3 sites approach to the Fe atom, which corresponds to the major metabolites of indole by CYP2W1. The interaction energy between the substrates and CYP2W1 was calculated and analyzed using the MM-GBSA method. The energetic analysis revealed that van der Waals energy had the predominant contribution to the substrate binding with the enzyme. Several residues including Phe115, Met206, Met295, Ala296, Glu299, Leu361, Phe475 and Thr476, were identified to have great contributions to the binding of indole and benzphetamine. These key residues are expected to affect the catalytic activity and can be used as the candidates for further mutagenesis studies. The generated homology model is also useful for the structure-based drug design of CYP2W1.

Acknowledgments

We thank Dr. Zhongliu Wu for her invaluable discussion about CYP2W1. W. Li is supported by a Japanese Society for the Promotion of Science (JSPS) research fellowship for foreign researchers. This work was also supported by the Grant-in-Aid for Scientific Research from JSPS.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2009.06.002.

References

[1] S. Rendic, Summary of information on human CYP enzymes: human P450 metabolism data, Drug Metab. Rev. 34 (2002) 83–448.
[2] K. Stark, F.P. Guengerich, Characterization of orphan human cytochromes P450, Drug Metab. Rev. 39 (2007) 627–637.
[3] Z.L. Wu, C.D. Sohl, T. Shimada, F.P. Guengerich, Recombinant enzymes over-expressed in bacteria show broad catalytic specificity of human cytochrome P450 2W1 and limited activity of human cytochrome P450 2S1, Mol. Pharmacol. 69 (2006) 2007–2014.
[4] H. Yoshioka, N. Kasai, S. Ikushiro, R. Shinkyo, M. Kamakura, M. Ohta, K. Inouye, T. Sakaki, Enzymatic properties of human CYP2W1 expressed in *Escherichia coli*, Biochem. Biophys. Res. Commun. 345 (2006) 169–174.
[5] M. Karlgren, A. Gomez, K. Stark, J. Svard, C. Rodriguez-Antona, E. Oliw, M.L. Bernal, S. Ramon y Cajal, I. Johansson, M. Ingelman-Sundberg, Tumor-specific expression of the novel cytochrome P450 enzyme, CYP2W1, Biochem. Biophys. Res. Commun. 341 (2006) 451–458.
[6] M. Karlgren, M. Ingelman-Sundberg, Tumour-specific expression of CYP2W1: its potential as a drug target in cancer therapy, Exp. Opin. Ther. Targets 11 (2007) 61–67.
[7] C. de Graaf, N.P. Vermeulen, K.A. Feenstra, Cytochrome p450 in silico: an integrative modeling approach, J. Med. Chem. 48 (2005) 2725–2755.
[8] W. Li, Y. Tang, H. Liu, J. Cheng, W. Zhu, H. Jiang, Probing ligand binding modes of human cytochrome P450 2J2 by homology modeling, molecular dynamics simulation, and flexible molecular docking, Proteins 71 (2008) 938–949.

[9] E.F. Johnson, C.D. Stout, Structural diversity of human xenobiotic-metabolizing cytochrome P450 monooxygenases, Biochem. Biophys. Res. Commun. 338 (2005) 331–336.
[10] 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.
[11] M.R. Wester, E.F. Johnson, C. Marques-Soares, S. Dijols, P.M. Dansette, D. Mansuy, C.D. Stout, Structure of mammalian cytochrome P450 2C5 complexed with diclofenac at 2.1 Å resolution: evidence for an induced fit model of substrate binding, Biochemistry 42 (2003) 9335–9345.
[12] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucleic Acids Res. 25 (1997) 4876–4882.
[13] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815.
[14] 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.
[15] D.A. Case, T.E. Cheatham 3rd, T. Darden, H. Gohlke, R. Luo, K.M. Merz Jr., A. Onufriev, C. Simmerling, B. Wang, R.J. Woods, The Amber biomolecular simulation programs, J. Comput. Chem. 26 (2005) 1668–1688.
[16] D. Eisenberg, R. Luthy, J.U. Bowie, VERIFY3D: assessment of protein models with three-dimensional profiles, Methods Enzymol. 277 (1997) 396–404.
[17] M.J. Sippl, Recognition of errors in three-dimensional structures of proteins, Proteins 17 (1993) 355–362.
[18] R. Huey, G.M. Morris, A.J. Olson, D.S. Goodsell, A semiempirical free energy force field with charge-based desolvation, J. Comput. Chem. 28 (2007) 1145–1152.
[19] Y. Duan, C. Wu, S. Chowdhury, M.C. Lee, G.M. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J.M. Wang, P. Kollman, A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations, J. Comput. Chem. 24 (2003) 1999–2012.
[20] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, J. Comput. Chem. 25 (2004) 1157–1174.
[21] D.L. Harris, J.Y. Park, L. Gruenke, L. Waskell, Theoretical study of the ligand–CYP2B4 complexes: effect of structure on binding free energies and heme spin state, Proteins 55 (2004) 895–914.
[22] W. Li, H. Ode, T. Hoshino, H. Liu, Y. Tang, H. Jiang, Reduced catalytic activity of P450 2A6 mutants with coumarin: a computational investigation, J. Chem. Theory Comput. 5 (2009) 1411–1420.
[23] W.L. Jorgensen, J. Chandrasekhar, J. Madura, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.
[24] J.P. Rychaert, G. Ciccotti, H.J.C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, J. Comput. Phys. 23 (1977) 327–341.
[25] P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S.H. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D.A. Case, T.E. Cheatham, Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, Acc. Chem. Res. 33 (2000) 889–897.
[26] A. Onufriev, D. Bashford, D.A. Case, Exploring protein native states and large-scale conformational changes with a modified generalized born model, Proteins 55 (2004) 383–394.
[27] J. Weiser, P.S. Shenkin, W.C. Still, Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO), J. Comput. Chem. 20 (1999) 217–230.
[28] Y. Zhao, W. Li, J. Zeng, G. Liu, Y. Tang, Insights into the interactions between HIV-1 integrase and human LEDGF/p75 by molecular dynamics simulation and free energy calculation, Proteins 72 (2008) 635–645.
[29] P. Rowland, F.E. Blaney, M.G. Smyth, J.J. Jones, V.R. Leydon, A.K. Oxbrow, C.J. Lewis, M.G. Tennant, S. Modi, D.S. Eggleston, R.J. Chenery, A.M. Bridges, Crystal structure of human cytochrome P450 2D6, J. Biol. Chem. 281 (2006) 7614–7622.
[30] I.G. Denisov, T.M. Makris, S.G. Sligar, I. Schlichting, Structure and chemistry of cytochrome P450, Chem. Rev. 105 (2005) 2253–2277.
[31] V. Cojocaru, P.J. Winn, R.C. Wade, The ins and outs of cytochrome P450s, Biochim. Biophys. Acta 1770 (2007) 390–401.
[32] 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.
[33] W. Li, H. Liu, E.E. Scott, F. Grater, J.R. Halpert, X. Luo, J. Shen, H. Jiang, Possible pathway(s) of testosterone egress from the active site of cytochrome P450 2B1: a steered molecular dynamics simulation, Drug Metab. Dispos. 33 (2005) 910–919.
[34] M. Petrek, P. Kosinova, J. Koca, M. Otyepka, MOLE: a Voronoi diagram-based explorer of molecular channels, pores, and tunnels, Structure 15 (2007) 1357–1363.
[35] 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. USA 98 (2001) 3068–3073.
[36] K. Schleinkofer, Sudarko, P.J. Winn, S.K. Ludemann, R.C. Wade, Do mammalian cytochrome P450s show multiple ligand access pathways and ligand channelling? EMBO Rep. 6 (2005) 584–589.
[37] B.K. Muralidhara, L. Sun, S. Negi, J.R. Halpert, Thermodynamic fidelity of the mammalian cytochrome P450 2B4 active site in binding substrates and inhibitors, J. Mol. Biol. 377 (2008) 232–245.
[38] T. Inoue, S. Suzuki, The metabolism of 1-phenyl-2-(N-methyl-N-benzylamino)-propane (benzphetamine) and 1-phenyl-2-(N-methyl-N-furfurylamino)propane (furfenorex) in man, Xenobiotica 16 (1986) 691–698.
[39] C. de Graaf, C. Oostenbrink, P.H. Keizers, T. van der Wijst, A. Jongejan, N.P. Vermeulen, Catalytic site prediction and virtual screening of cytochrome P450 2D6 substrates by consideration of water and rescoring in automated docking, J. Med. Chem. 49 (2006) 2417–2430.
[40] E.M. Gillam, L.M. Notley, H. Cai, J.J. De Voss, F.P. Guengerich, Oxidation of indole by cytochrome P450 enzymes, Biochemistry 39 (2000) 13817–13824.
[41] B.D. Smith, J.L. Sanders, P.R. Porubsky, G.H. Lushington, C.D. Stout, E.E. Scott, Structure of the human lung cytochrome P450 2A13, J. Biol. Chem. 282 (2007) 17306–17313.
[42] 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.
[43] P.H. Keizers, B.M. Lussenburg, C. de Graaf, L.M. Mentink, N.P. Vermeulen, J.N. Commandeur, Influence of phenylalanine 120 on cytochrome P450 2D6 catalytic selectivity and regiospecificity: crucial role in 7-methoxy-4-(aminomethyl)-coumarin metabolism, Biochem. Pharmacol. 68 (2004) 2263–2271.
[44] 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.
[45] A. Melet, N. Assrir, P. Jean, M. Pilar Lopez-Garcia, C. Marques-Soares, M. Jaouen, P.M. Dansette, M.A. Sari, D. Mansuy, Substrate selectivity of human cytochrome P450 2C9: importance of residues 476,365, and 114 in recognition of diclofenac and sulfaphenazole and in mechanism-based inactivation by tienilic acid, Arch. Biochem. Biophys. 409 (2003) 80–91.
[46] T. Hou, W. Zhang, J. Wang, W. Wang, Predicting drug resistance of the HIV-1 protease using molecular interaction energy components, Proteins 74 (2009) 837–846.
[47] T. Hou, W.A. McLaughlin, W. Wang, Evaluating the potency of HIV-1 protease drugs to combat resistance, Proteins 71 (2008) 1163–1174.