A 3D Model of CYP1B1 Explains the Dominant 4-Hydroxylation of Estradiol

Toshimasa Itoh,‡ Hitomi Takemura,§ Kayoko Shimoi, and Keiko Yamamoto

Laboratory of Drug Design and Medicinal Chemistry, Showa Pharmaceutical University, 3-3165 Higashi-Tamagawagakuen, Machida, Tokyo 194-8543, Japan, Faculty of Human Health Sciences, Matsumoto University, 2095-1 Niimura, Matsumoto 390-1295, Japan, and Institute for Environmental Sciences, University of Shizuoka, 52-1 Yada, Suruga, Shizuoka 422-8526, Japan

Received February 6, 2010

CYP1A1 and CYP1A2 exhibit catalytic activity predominantly for the 2-hydroxylation of estradiol, whereas CYP1B1 exhibits catalytic activity predominantly for 4-hydroxylation of estradiol. To understand why CYP1B1 predominantly hydroxylates the 4-position of estradiol, we constructed three-dimensional structures of CYP1A1 and CYP1B1 by homology modeling, using the crystal structure of CYP1A2, and studied the docking mode of estradiol with CYP1A1, CYP1A2, and CYP1B1. The results demonstrated that two particular amino acid residues for each CYP, namely Thr124 and Phe260 of CYP1A2, Ser122 and Phe258 of CYP1A1, and Ala133 and Asn265 of CYP1B1, play an important role in estradiol recognition.

INTRODUCTION

The endogenous estrogens, 17β-estradiol (E2) and estrone (E1), have long been recognized as prime risk factors for the development of breast cancer.1 It is now known that oxidative estrogen metabolism is a critical factor in the development of human breast cancer.2,3 Figure 1 shows the metabolic pathway of the main estrogen, E2, in the mammary gland. Metabolism is regulated by the oxidizing enzymes CYP1A1 and CYP1B1 and by catechol-O-methyltransferase (COMT). CYP1A1 exhibits catalytic activity predominantly for the 2-hydroxylation of E2 to yield 2-hydroxyestradiol (2-OHE2).4,5 2-OHE2 is rapidly methylated by COMT to produce 2-methoxyestradiol (2-MeOE2), which is known to be noncarcinogenic and to inhibit the proliferation of cancer cells. CYP1B1 exhibits catalytic activity predominantly for the 4-hydroxylation of E2 to yield 4-hydroxyestradiol (4-OHE2).4,5 4-OHE2 is mainly oxidized by peroxidases to produce estradiol-3,4-semiquinone (E2-3,4-SQ) and estradiol-3,4-quinone (E2-3,4-Q). E2-3,4-Q forms a quinone-DNA adduct, causing mutations and initiating carcinogenesis. Recently, a quinone metabolite, E2-3,4-Q, was identified as a potential risk factor in breast cancer.6

Liehr and Ricci reported that 4-hydroxylation of E2 is a marker for human mammary tumors.7 In the breast, 4-hydroxylation is the dominant metabolic pathway for E2 because the expression level of CYP1B1 is higher than that of CYP1A1. In addition, CYP1B1 is expressed more highly in breast cancer tissue than in normal breast tissue.8,9 Consequently, the concentration of 4-OHE2 in human breast cancer is higher than that in normal tissue.3 Therefore, a CYP1B1 inhibitor might be promising as a drug for the prevention and treatment of breast cancer. To understand the molecular basis of carcinogenesis induced by E2, it is important to know why CYP1A1 and CYP1B1 predominantly hydroxylate the C2- and C4-position, respectively, of E2.

CYP1A2 is the principal family 1 enzyme expressed in the human liver and contributes to the hepatic metabolism of E2 to produce 2-OHE2. Recently Sansen et al. reported the X-ray crystal structure of human CYP1A2.10 In this article, we report the three-dimensional (3D) structures of CYP1A1 and CYP1B1 constructed by homology modeling using the crystal structure of CYP1A2 as a template. We also report results of the docking study of E2 into CYP1A2, CYP1A1, and CYP1B1 that indicates structural reasons why CYP1A1 and CYP1A2 predominantly hydroxylate the C2-position, whereas CYP1B1 hydroxylates the C4-position of E2.

MATERIALS AND METHODS

Sequence Alignment. We aligned the three human CYPs responsible for estrogen metabolism, CYP1A1, CYP1B1, and CYP1A2, using the ClustalW2 program (http://www.ebi.ac.uk/Tools/clustalw2/).11 A part of the computationally derived alignment was modified manually (see the Results Section).

Modeling. Based on the alignment, we constructed a 3D model of CYP1B1 and CYP1A1 by sequence replacement using the SYBYL Biopolymer program12 (Tripos Inc.) and the atomic coordinates (Arg34–Ser513) of the crystal structure of CYP1A2 (PDB; 2HI4) as a template.10 The procedure was as follows: First of all, the structure of residues in the structurally conserved regions, consisting of helices, β-strands, and loops with the same number of amino acid residues as the template, was constructed. Side chain conformational angles in the residues are retained as far as possible. The positions of the hydrogens added are determined by data in SYBYL’s internal tables for bond lengths and angles. Second, other parts of the CYP structure were constructed manually using the Coot 0.5.2 program13 by applying torsion and Ramachandran restraints. Then, heme
was merged into the protein to occupy the same position as the heme of the template protein, CYP1A2. Finally, energy minimization of the constructed structure was performed until the energy gradient was lower than 0.1 kcal/(mol Å) using the Tripos force field. The validity of the CYP1A1 and CYP1B1 models was evaluated using PROCHECK in the CCP4 program.

Docking. Docking was performed using Surflex Dock in Sybyl 8.0 (Tripos). The crystal structure of CYP1A2 (PDB: 2HI4) and the CYP1A1 and CYP1B1 models constructed above were used as the CYP protein structure. Surflex Dock program uses an empirical scoring function (based on the Hammerhead docking system) that has been updated and reparameterized with additional negative training data, along with a search engine that relies on a surface-based molecular similarity method. We applied the “hard grid treatment” method in which side chains of the active site residues are fixed, because we considered that flexibility of those side chains is minimum judging from the clear electron density map and the low B-factors of the CYP1A2 crystal structure. More than 40 substrate structures were obtained for each CYP. Substrate structures with the top 20 scores were selected, and then structures whose A-ring was positioned within 5.5 Å from the heme iron atom in CYP were extracted. Energy minimization (Tripos force field) of each substrate-CYP complex was performed to give the final complex model.

RESULTS AND DISCUSSION

Sequence Alignment. Figure 2 shows the final alignment of the CYP1A1, CYP1A2, and CYP1B1 sequences. A part of the computationally derived alignment was modified manually. Campos-Mollo reported that the F261L mutation in CYP1B1 causes primary congenital glaucoma (PCG) without affecting the stability of the CYP, suggesting that F261 is a critical residue for expression of the function. We thought that F261 faces the surface of the active site of CYP1B1. Therefore, we manually modified the computationally derived alignment to align F261 of CYP1B1 with F256 of CYP1A2.

Compared with template CYP1A2, CYP1B1 has four deletion and two insertion segments, and CYP1A1 has one deletion and one insertion segment. Sequence identity of CYP1A1 with CYP1A2 is 72%, whereas that of CYP1B1 with CYP1A2 is 37%. Highly conserved and functionally important residues in the CYP superfamily are: (1) three absolutely conserved residues, ExxR in the K helix and Cys just before the L helix (for example E387, R390, and C470 in CYP1B1); (2) the consensus sequence (A/G)(G/S)x(E/D)T in the center of the I helix; and (3) the consensus sequence F(G/S)xGx(H/R)xCxGxx(I/L/F)(A/S) containing the Cys (C470 in CYP1B1) responsible for heme binding. As shown in Figure 2, these important segments of the three CYP sequences correctly align with each other. In addition, the segments of the CYP1B1 and CYP1A1 sequences corresponding to the secondary structure elements of the CYP1A2 crystal structure do not have any gaps or insertions (Figure 2), indicating that most of the secondary structure is conserved.

Modeling. The constructed 3D models of CYP1A1 and CYP1B1 are shown in Figure 3. The N-terminal regions of CYP1A1 (Met1-Pro31) and CYP1B1 (Met1-Gln42) were not constructed because the coordinates of the corresponding region of CYP1A2 (Met1-Pro33) were not determined. We evaluated the model structure by using the PROCHECK in the CCP4 program. A Ramachandran plot showed that 99.1% of the residues in CYP1A1 and 98.1% of the residues in CYP1B1 are in either the most favored or allowed regions. The 3D structures of CYP1A1 and CYP1B1 were similar to those of other CYPs reported to date. The structures contain 12 α-helices and 4 β-sheets and have several additional helices, as does the template CYP1A2. In our models, Glu(E)374, Arg(R)377, and Arg(R)431 of CYP1A1 (Figure 3b) and Glu(E)387 and Arg(R)390 and Arg(R)444 of CYP1B1 (Figure 3c) form the ERR triad that is the conserved folding core in the CYP1 family. In the CYP1B1 model, Ala330 and Thr334 in helix I are highly conserved; the hydroxyl group of the Thr334 side chain forms a hydrogen bond with the backbone carbonyl oxygen of Ala330. Thr334, which is located near the heme, is thought
to be involved in the oxygen activation mechanism. The same structure is observed in the CYP1A1 model.
modes, we termed the complexes shown in Figure 5d (Figure 5d and e). To distinguish the two types of binding one binding mode was found in CYP1A2 and CYP1A1 modes were found in CYP1B1 (Figure 5f and i), while only the structure prior to oxidation at the A-ring. Two binding conformers were thought to be adequate representations of the substrate-recognition sites (SRS) proposed by Gotoh. It should be noted that all residues whose side chain faces the substrate-binding pocket are located at the six substrate-recognition sites of CYP1A1 and CYP1B1 active sites are 600 and 550 Å³, respectively, which are larger than that of CYP1A2 (470 Å³). It should be noted that all residues whose side chain faces the substrate-binding pocket are located at the six substrate-recognition sites (SRS) proposed by Gotoh. The identities of those residues between CYP1B1 and CYP1A1, CYP1B1 and CYP1A2, and CYP1A1 and CYP1A2 are 74, 63, and 68%, respectively. As shown in Figure 4, residues whose side chain faces the top or bottom of the substrate-binding pocket are highly conserved. In contrast, residues whose side chain faces a lateral part of the substrate-binding pocket are variable. These results imply that nonconserved residues at lateral positions, such as Thr124, Phe260, and Asn312 of CYP1A2, Ser122, Phe258, and Leu312 of CYP1A1, and Ala133, Asn265, and Thr325 of CYP1B1, play an important role in substrate recognition.

Docking. We obtained more than 40 E₂–CYP complex structures for each CYP, from which we selected the E₂ conformers with the top 20 scores because they covered all the representative conformations of E₂. Figures 5a–c show the 20 conformations of E₂ for each CYP. We found that there are five binding modes toward CYP1A2 and CYP1A1 (Figure 5a and b) and nine modes toward CYP1B1 (Figure 5c). These representative five binding modes toward CYP1A2 and CYP1A1 and the nine binding modes toward CYP1B1 are designated in Figure 5a–c as five and nine blue arrows, respectively, which point to the oxygen of the 3-hydroxyl group of E₂. To investigate the positional selectivity of hydroxylation at the A-ring, we chose conformers having an A-ring carbon within 5.5 Å of the heme Fe atom. These conformers were thought to be adequate representations of the structure prior to oxidation at the A-ring. Two binding modes were found in CYP1B1 (Figure 5f and i), while only one binding mode was found in CYP1A2 and CYP1A1 (Figure 5d and e). To distinguish the two types of binding modes, we termed the complexes shown in Figure 5d–f and i “methyl up” and “methyl down” complexes, respectively, based on the direction of the 18-methyl group of E₂.

In each “methyl up” complex, appropriate hydrophobic interactions between E₂ and the CYP were observed, but no hydrogen bonds were apparent (Figure 5d–f). Interestingly, in the “methyl down” complex (Figure 5i), a hydrogen bond between the 17-hydroxyl group of E₂ and Asn265 of CYP1B1 was observed in addition to hydrophobic interactions. Docking analysis was performed to better understand why E₂ cannot adopt a “methyl down” conformer in CYP1A2 and CYP1A1. E₂ was extracted from the “methyl down” complex of CYP1B1 and E₂ (Figure 5i) and merged into CYP1A2 and CYP1A1 (Figure 5g and h). These superimpositions indicated that Thr124 and Phe260 of CYP1A2 and Ser122 and Phe258 of CYP1A1 are sufficiently close (less than 3.2 Å) to cause steric clash with the B-ring and the 17-hydroxyl group of the substrate (Figure 5g and h). These results indicate that Thr124 and Phe260 of CYP1A2, Ser122 and Phe258 of CYP1A1, and Ala133 and Asn265 of CYP1B1 are critical residues for E₂ recognition.

Several papers concerning the importance of various residues for CYP1A1 activity were reported. Guengerich’s group generated random libraries of human CYP1A2 with substrate-recognition sequences mutated and obtained mutants that alter the intervening steps leading to the formation of the activated Michaelis complex. Szklarz’s group performed modeling and mutation studies of CYP1A1 and indicated the importance of Val382 for recognition of the side chain of resorufin derivatives. In addition, Thr124 of CYP1A2 and Ser122 of CYP1A1 were demonstrated to play an important role for alkoxyresorufin metabolism. This demonstration is in agreement with our suggestion that Thr124 of CYP1A2, Ser122 of CYP1A1, and Ala133 of CYP1B1 are important residues for substrate recognition.

It is clear that the “methyl up” conformations are oxidized at the C2-position (Figure 5d–f), whereas the “methyl down” conformations are oxidized at the C4-position (Figure 5i). The docking study suggests that CYP1A2 and CYP1A1 preferentially generate 2-OHE₂ rather than 4-OHE₂ and that CYP1B1 generates both 2-OHE₂ and 4-OHE₂. This suggestion is supported by a report from Lee et al. that CYP1A1 and CYP1A2 produce 2-OHE₂ and 4-OHE₂ in a ratio of >10:1, whereas CYP1B1 produces 2-OHE₂ and 4-OHE₂ in a ratio of 1:3. Preference for 4-hydroxylation by CYP1B1 can be explained by the hydrogen bond between the 17-hydroxyl...
group and Asn265; only the “methyl down” conformer forms this favorable hydrogen bond in CYP1B1. Thus, our docking model is consistent with the specific 4-hydroxylation of E2.

Lee et al. reported that the rank order of the reaction rate for E2 is 1A2 > 1A1 > 1B1. We suggest that one of the reasons for this order is the entropic advantages of CYP1A2 and CYP1A1. Since the shape of the CYP1B1 substrate-binding pocket is rounder than those of CYP1A2 and CYP1A1, as shown in Figure 5a–c, CYP1B1 allows multiple conformations of E2 to be accommodated. This is consistent with the docking simulation results described above, in which nine binding modes appeared in CYP1B1 but only five binding modes appeared in CYP1A2 and CYP1A1.

CONCLUSIONS

The preference for 2-hydroxylation of E2 by CYP1A2 and CYP1A1 was investigated by docking analysis of E2, and the analysis suggested that only the “methyl up” complex is allowed. The analysis also suggested that the preference for 4-hydroxylation of E2 by CYP1B1 is due to both “methyl up” and “down” complexes being acceptable but with “methyl down” being preferred due to the presence of a hydrogen bond between E2 and CYP. Furthermore, Thr124 and Phe260 of CYP1A2, Ser122 and Phe258 of CYP1A1, and Ala133 and Asn265 of CYP1B1 were indicated to be critical residues for E2 recognition.

REFERENCES AND NOTES


(12) SYBYL Molecular Modeling Software, version 8.0; Tripos Inc.: 1699 South Hanley Road, St. Louis, MO 63144–2913.


(19) Zhou, S. F.; Yang, L. P.; Wei, M. Q.; Duan, W.; Chan, E. Insights into the binding of inhibitors and both small molecular weight and fatty acid substrates. *J. Biol. Chem.* 2008, 283, 33698–33707.


(30) Yun, C. H.; Miller, G. P.; Guengerich, F. P. Rate-determining steps in phenacetin oxidations by human cytochrome P450 1A2 and selected P450s. *Biochemistry* 2000, 39, 11319–11329.


