Effect of the Arg456His mutation on the three-dimensional structure of cytochrome P450 1A2 predicted by molecular dynamics simulations

Y Watanabe¹², K Kato³, S Fukuyoshi², M Hiratsuka⁴, N Yamaotsu⁵, S Hirono⁵, H Gouda¹ and A Oda²,³,⁶*

¹ School of Pharmacy, Showa University, 1-5-8 Hatanodai, Shinagawa-ku, Tokyo 142-8555, Japan
² Faculty of Pharmacy, Institute of Medical, Pharmaceutical and Health Science, Kanazawa University, Kakuma, Kanazawa, Ishikawa 920-1192, Japan
³ Graduate School of Pharmacy, Meijo University, 150 Yagotoyama, Tempaku-ku, Nagoya, Aichi 468-8503, Japan
⁴ Graduate School of Pharmaceutical Sciences, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai, Miyagi 980-8577, Japan
⁵ School of Pharmacy, Kitasato University, 5-9-1 Sirokane, Minato-ku, Tokyo 108-8641, Japan
⁶ Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan

*E-mail: oda@meijo-u.ac.jp

Abstract. Cytochrome P450 1A2 (CYP1A2), one of the major drug-metabolizing enzymes among CYPs, has many variant alleles. The genetic polymorphism of CYP1A2 is thought to cause individual differences in the pharmacokinetics of medicines. CYP1A2.8 (the Arg456His mutant), a CYP1A2 variant, has decreased enzymatic activity. In our previous work, to understand why the Arg456His mutant lost its enzymatic activity, we constructed an Arg456His mutant with the hydrogen on the epsilon nitrogen of the histidine (HIE) and performed a 300-ns MD simulation. After the MD simulation, the Arg456His mutant with HIE showed large differences in static structure and flexibility compared with the wild type, which would cause the decreased activity. In the work described here, we constructed an additional Arg456His mutant with positive charged histidine and performed a 300-ns MD simulation to consider the effect of the protonation state of the histidine. Comparing these two Arg456His mutants revealed differences in their static structures, flexibilities, and interactions, suggesting that the protonation state of the His456 residue has a considerable influence on the physiological properties of CYPs. Our results indicate that the Arg456His mutation causes the CYP1A2 structure to unfold regardless of the protonation state of His456.

1. Introduction
Cytochrome P450 (CYP), a superfamily of heme-containing multifunctional enzymes, is prevalent in the natural world, existing in a wide range of organisms, such as bacteria, plants, and animals [1]. Although CYPs play many important roles in the synthesis of biologically active substances, metabolic activation of some mutagenic compounds, and detoxification of xenobiotics, CYPs have
attracted clinical attention mainly for their drug-metabolizing function [2-4]. Therefore, genetic polymorphism in CYPs is viewed as a pharmacotherapeutic challenge, as this variation can affect their efficacy and result in adverse side effects of medicine [5]. Major drug-metabolizing CYPs such as CYP2C9, CYP2D6, and CYP3A4/5 have many variant alleles, which are important from the viewpoint of pharmacokinetics [6]. In drug metabolism reactions, CYPs catalyze the monooxygenation of drugs using oxygen molecule. The proposed reaction mechanism by CYPs is shown in scheme 1 [7]. First, a water molecule, the sixth ligand of the heme iron is replaced by the substrate. The binding of the substrate changes the six-coordinate high-spin Fe$^{III}$ to the five-coordinate, low spin Fe$^{II}$. Next, first electron from an electron carrier, such as nicotinamide adenine dinucleotide phosphate reduces the five-coordinate Fe$^{III}$ to the Fe$^{II}$. Then, the binding of the oxygen molecule converts the five-coordinate Fe$^{II}$ to the six-coordinate Fe$^{III}$. After the reduction by second electron, two proton cleaves the O—O double bond and generate a water molecule and an active species so-called compound I. Although this oxidizing intermediate has not been elucidated experimentally, it is proposed that the oxygen atom is inserted in the substrate and generate the final oxidized product.

Thus, the catalytic reaction requires two electrons from its electron carrier and electron transfers are mediated by the membrane-bound flavoprotein cytochrome P450 redox partner (CPR) [8]. Thus, binding and interaction with CPR are necessary for the catalytic reaction of CYPs, and putative CPR-binding sites in CYPs have been studied [9-12]. Although the potential binding site is not yet verified, it is likely located proximal to the C, K, and L helices of the CYP protein. Moreover, in the control of the heme catalytic cycle, water molecules play important roles [8] as the sixth ligand for the heme and are required for heme inactivation. When the substrate binds to the active site and approaches the heme iron, CYPs undergo a conformational change to form a substrate-bound complex structure. Formation of this complex changes the spin state from a low spin (inactive) to a high spin (active). In the P450-BM3 X-ray crystal structure, water molecule is expected to participate in proton transfer for the hydrolysis reaction [13]. Schlichting et al. conducted cryo-crystallography of P450cam bound to the substrate camphor [7] and showed that water molecules formed hydrogen bond networks with some residues. Moreover, in the hydrolysis reaction of camphor by P450cam, Parasad et al. confirmed via NMR that the hydrogen atom added to camphor was derived from water molecules [14]. Crystal structure analyses revealed that some mammalian CYPs, including CYP2C19 (PDB ID: 4GQS) and CYP3A4 (PDB ID: 4I3Q), have water molecules in their active sites. However, due to the high mobility of water molecules, their functions are difficult to define using current experimental techniques.
CYP1A2, a major drug-metabolizing subtype among CYPs, has many variant alleles including some low-active mutants [15]. The genetic polymorphism of CYP1A2 may cause individual differences in the pharmacokinetics of clinically important medicines such as asthma remedies (theophylline), antipsychotics (amitriptyline, imipramine, and olanzapine), and antiarrhythmic drugs (mexiletine and propranolol). One of the CYP1A2 variants, CYP1A2.8 (the Arg456His mutant), has been reported to decrease enzymatic activity [15,16]. The Arg456His mutation is located at the KL loop proximal to Cys458, which is the fifth ligand for heme. According to previously reported sequence analyses of CYPs, the heme-binding domain is considered to be the most conserved region [17–19], and the amino acid sequence around the cysteine (FXXGXRXCG) is highly conserved in most CYPs [18]. In the crystal structure analyses of Hasemann et al., arginine was sometimes substituted by another basic residue, such as histidine in P450terp and P450cam, and lysine in P450BM3, suggesting that the basic residues (arginine/histidine/lysine) fold the Cys-pocket into position by interacting their positively charged side chains with the propionate group of heme [19]. In the natural human CYP1A2.8 mutant, the Arg456His substitution dramatically reduced the heme peak at 450 nm [16], indicating a disturbance of the normal function of heme. A quantum mechanics/molecular mechanics (QM/MM) study of the P450cam–camphor complex by Guallar et al. suggested that the His355 of P450cam, corresponding to the Arg456 of CYP1A2, forms a hydrogen bond with heme and contributes to the stabilization of the oxygenated form [20]. In our previous work, to understand why the Arg456His mutant lost its enzymatic activity, we constructed an Arg456His mutant with the hydrogen on the epsilon nitrogen of the histidine (HIE) and performed a 300-ns MD simulation [21]. HIE should be suitable for interaction with the surrounding residues, including the heme molecule. After the MD simulation, the Arg456His mutant with HIE showed large differences in static structure and flexibility compared to the wild type, which would explain its decreased activity. In the present study, in order to consider the effect of the protonation state of histidine, we constructed an additional Arg456His mutant structure with positively charged histidine (HIP), which would have more electrostatic interaction with the surrounding residues than HIE. We performed 300-ns MD simulations at 300 K for the Arg456His mutant with HIP and compared the results with those from the Arg456His mutant with HIE.

2. Methods
Initial 3D structures of the Arg456His mutant with HIE were constructed from the optimized structure of wild-type CYP1A2 obtained by 100-ns MD simulation [21], which was constructed from the crystal structure (PDB ID: 2HI4) [22]. Likewise, the initial structure of Arg456His with HIP was also constructed from the optimized wild-type structure. Hereafter we refer to these, the Arg456His mutant with HIE and the Arg456His mutant with HIP, as the Arg456His mutant and the Arg456Hip mutant, respectively. First, all crystal water molecules and the ligand were removed from the 2HI4 crystal structure of wild-type CYP1A2. To obtain the equilibrium state for the ligand-free form, a 100-ns MD simulation was performed at 300 K under constant pressure [21]. Using the energy-minimized structure of wild-type CYP1A2, initial structures of the Arg456His mutants were constructed by manually substituting Arg to His. Calculation settings were the same as used in our previous studies of the ligand-free CYP1A2 [21], CYP2B6 [23], CYP2C9, CYP2C19 [24], and CYP2D6 [25]. The system was solvated using the TIP3P model [26] and calculated under cyclic boundary conditions. The cyclic boundary box was spaced with at least 8Å margins from the protein surface, the cutoff distance for van der Waals interactions was set at 10 Å, and the particle mesh Ewald method was used for calculating electrostatic interactions [27]. Chloride ions were arranged to neutralize the system. For the force field parameter around the heme iron, our determined parameters were used, which represent the sextet state with five-coordinate iron(III) [28]. When the substrate is recognized by CYP, forming the sextet and five-coordinate iron is essential for the catalytic reaction. For the rest of the heme molecule, parameters determined by Giammona were used [29]. AMBER ff12SB force field equipped within AMBER12 package [30] and general AMBER force field (GAFF) [31] were used for amino acids and the ligand, respectively. The tleap module of AmberTools was used to supplement hydrogen atoms not
observed in the crystal structure. The time step of all MD simulations was 2 fs. The SHAKE algorithm was used to constrain bond stretching of hydrogen atoms [32].

First, the solvent was relaxed and the whole system was minimized, then temperature-increasing MD simulations were performed for 20 ps, with the temperature raised from 0 to 300 K. Equilibrating MD simulations were then performed for 300 ns under constant pressure at 300 K. For the NPT ensemble, the Langevin thermostat was used to maintain the temperature of the system. To validate the 3D structures of the Arg456Hie and Arg456Hip mutants after MD simulations, root mean square deviations (RMSDs) for the main-chain atoms were evaluated along the calculated MD trajectories. As reference structures for RMSDs, the 3D structures after temperature-increasing MD simulations were used. To investigate the structural flexibilities of residues, root mean square fluctuations (RMSFs) were calculated for all Cα atoms using the last 20-ns MD trajectories. As reference structures for RMSFs, the average structures of 280-300 ns trajectories were used. Calculations of RMSDs and RMSFs were performed using the cpptraj module of AmberTools12. The hydrogen bond networks were also analyzed using 280–300 ns trajectories. Regarding hydrogen bond criteria, the distance between two heavy atoms and the angle between the acceptor, hydrogen, and donor atoms were determined with cutoffs of 3 Å and 120°, respectively.

3. Results and discussion

We conducted 300-ns MD simulations for both Arg456His mutants and calculated the RMSDs to validate the 3D structures (figure 1). In our previous work [21], 100-ns MD simulation was conducted for the wild-type CYP1A2 structure and the optimized structure was used as the initial structures of two Arg456His mutants. As shown in fig. 2a, the RMSD of the wild-type CYP1A2 converged around 50-100 ns, suggesting the 3D structure of the wild-type CYP1A2 reached its equilibrium state by 100-ns MD simulation.

Figure 1. Root mean square deviations (RMSDs) of the main-chain atoms for (a) the wild-type CYP1A2, (b) the Arg456Hie mutant and (b) the Arg456Hip mutant.
Figure 2. Interaction around His456 in the initial and optimized structures. Initial structures of Arg456Hie (a) and Arg456Hip (b) mutants. Optimized structures of Arg456Hie (c) and Arg456Hip (d) mutants. Residues within 3 Å are labeled. Hydrogen bonds shown with black dashed line.

Table 1. Average distance between the propionate oxygen atom of the heme and the nearest nitrogen atom of the 456th residue, calculated using last 20 ns of trajectories.

| Protein name                               | Heavy atom               | Distance/ Å |
|--------------------------------------------|--------------------------|-------------|
| Wild type (the ligand-free form) [21]      | HEME@O1D–ARG456@NH1      | 5.07        |
| Wild type (the substrate-bound form) [33]  | HEME@O2D–ARG456@NH1      | 2.84        |
| Wild type (the inhibitor bound form) [33]  | HEME@O2D–ARG456@NH1      | 5.17        |
| Arg456Hie mutant                           | HEME@O1A–HIP@NE2         | 5.21        |
| Arg456Hip mutant                           | HEME@O1A–HIP@ND1         | 5.28        |

On the contrary, the RMSD of the Arg456Hie mutant repeated rise and fall around 1.1 to 2.6 Å. Although the RMSD of the Arg456Hip mutants was less than 2.3 Å and not largely fluctuated, it is also rising. From these results, it seems that 3D structures of the Arg456Hie and Arg456Hip mutants did not reach their equilibrium states by 300-ns MD simulations. Although the 3D structures of two Arg456His mutants were not yet relaxed by 300 ns MD simulations, these results suggest that the Arg456His mutation have large influence on the CYP1A2 structure independently of the His456 protonation state. Moreover, although the RMSD of the Arg456Hie mutant showed the existence of a metastable state during 100–145 and 200–250 ns [21], the RMSD of Arg456Hip did not show this
metastable feature. This suggests that the protonation state of His456 has an influence on both the 
stability and energy state of the CYP1A2 structure. Figure 2 shows the initial and optimized 3D 
structures of the Arg456Hie and Arg456Hip mutants. In the initial structures, the side chains of both 
Hie456 and Hip456 interacted with the propionate oxygen of heme (figure 2a and 2b). In the 
Arg456Hip mutant only, the imidazole side chain of Hip456 formed a hydrogen bond with the charged 
amino group of Lys106. On the other hand, in the optimized structures, neither Hie456 or Hip456 
interacted with the heme molecules. Moreover, although the imidazole side chain of Hie456 formed a 
hydrogen bond with Ser129 (figure 2c), that of Hip456 did not interact with any residues (figure 2d). 
Thus, Hie456 and Hip456 clearly had different interactions, which could spread across the 3D 
structures. In our previous work on wild-type CYP1A2, the interaction between Arg456 and heme was 
also investigated [21]. In the ligand-free form of CYP1A2, the distance between Arg456 and the heme 
was 5.07 Å, and the interaction was weak (table 1). However, in the substrate-bound CYP1A2, the 
distance between Arg456 and the heme was shortened to 2.84 Å, and the interaction became strong 
[33]. On the other hand, in the complex of CYP1A2 with the inhibitor, the distance between Arg456 
and the heme was over 5 Å [33]. These results suggest that the interaction between Arg456 and heme 
might be important to start the catalytic reaction in the substrate-bound form. At the least, the 

Figure 3. Superimposition of static structures for the Arg456Hie (blue) and Arg456Hip (pink) 
mutants. (a) Overall structures, (b) D and E helices, (c) KL loops and C helices, and (d) HI loops 
and I helices.
interaction between Arg456 and heme was one feature of the substrate-bound form. Therefore, it is not surprising that His456 did not interact with the heme in the ligand-free Arg456His mutants.

Figure 3 shows structural differences between the Arg456His and Arg456Hip mutants. After 300-ns MD simulations, RMSDs between the optimized structures for the Arg456His and Arg456Hip mutants were 2.39 Å, and the overall structures were not so different (figure 3a). However, large structural differences were observed in D helices (figure 3b), KL loops and C helices (figure 3c), and I helices (figure 3d). In the Arg456Hip mutants, the length of the D helix was shortened and a small β strand was formed in the DE loop (figure 3b). In the 3D structure of wild-type CYP1A2, the D helix is located at the protein surface and away from the Arg456His mutated position. This suggests that the protonation state of His456 affected the distant substructure. In addition, in the Arg456Hip mutant, the C helix was apart from the heme and partial β sheets were formed in the KL loop (figure 3c). The C helix reportedly constitutes a protein–protein interface with CPR [9–12]. Therefore, misalignment of the C helix would affect the electronic state of the heme and change its catalytic activity. Furthermore, structural differences were also observed in the I helices, considered to be among potential substrate recognition sites [17]. While the I helix of the Arg456Hip mutant was a straight α helix, that of the Arg456His mutant formed a loop and bent at the center. From this result, the protonation state of Arg456His seems to influence substrate recognition.

It is well known that CYP proteins have structural flexibility or conformational flexibility because CYP proteins undergo conformational change by ligand binding, so-called induced fit [34]. Our computational studies about CYP proteins suggested that the structural flexibility have large influence on the substrate recognition and the enzymatic activity [21, 23, 35]. Regarding the wild-type CYP1A2 obtained after 100 ns MD simulation [21], we calculated the RMSF using last 20 ns trajectory and observed highly flexible regions at the CD, FG, and HI loops (figure 4a and 5). To compare the structural flexibilities between the wild type and two Arg456His mutants, RMSFs were calculated using last 20 ns trajectories (figure 5 and table 2). Moreover, to examine whether the choice of the last 20 ns trajectory is adequate for RMSF calculation of two Arg456His mutants, RMSFs were also calculated using last 20, 40, and 60 ns (280–300, 260–300, and 240–300 ns) trajectories (figure 6). Values of high RMSF peaks at the BC, CD, FG, and HI loops were also tabulated in table 3. As for the Arg456Hip mutant (figure 6a), although the RMSF value of the BC loop was different by 0.73 Å when 280–300 and 240–300 ns trajectories were compared, overall RMSFs were well similar. On the other hand, regarding the Arg456His mutant (figure 6b), RMSF value of the CD loop was significantly higher by 1.29 Å using 240–300 ns trajectory than that of 280–300 ns trajectory. Furthermore, in the Arg456Hip mutant, the RMSF of 240–300 ns trajectory was relatively higher than that of 260–300 ns and 280–300 ns trajectories. Especially, the high RMSF peak at Pro183 was 2.97 Å in 240–300 ns trajectory, whereas the RMSF values were below 0.9 Å in 260–300 ns and 280–300 ns trajectories. Pro183 of the DE loop was exposed to the solvent in the wild type and the RMSF value was only 0.81 Å. This result suggests that 240–260 ns trajectory of Arg456Hip mutant includes some peculiar conformations. As shown in figure 6 and table 3, overall flexibilities of two Arg456His mutants did

### Table 2. Values of high RMSF peaks at BC, CD, FG, and HI loops in wild-type Arg456His mutants, calculated using last 20 ns of trajectories. Significantly high values shown in bold.

|         | Gly102 of BC loop | Ser156 of CD loop | Phe246 of FG loop | Gly299 of HI loop |
|---------|-------------------|-------------------|------------------|------------------|
| Wild type                        | 0.71              | 2.19              | 1.32             | 1.58             |
| Arg456His mutant                  | 0.75<sup>1</sup>  | 2.39              | 1.84             | 3.43<sup>2</sup> |
| Arg456Hip mutant                   | 1.80<sup>3</sup>  | 3.15<sup>3</sup>  | 1.26             | 1.51             |

<sup>1</sup> RMSF value at Asp103.
<sup>2</sup> RMSF value at Ser298.
<sup>3</sup> RMSF value at Ser155.
Figure 4. Overall structure of wild-type CYP1A2, from [21]. Highly flexible regions indicated with circles in (a); main substrate recognition sites shown with blue color in (b).

not so change during 240–300 ns trajectory though minor differences were observed at high RMSF peaks. These results would reflect that RMSDs of the Arg456Hie and Arg456Hip mutants have not yet converged by 300 ns simulation. Based on these results, in order to unify the calculation condition with the wild type, we selected the last 20 ns (280–300 ns) MD trajectory for the RMSF calculation and focused on them in following discussion.

As for the structural flexibility of wild-type CYP1A2, high RMSF peaks were observed at the CD, FG, and HI loops (figure 4a), with RMSF values of 2.19 Å at Ser156, 1.32 Å at Phe246, and 1.58 Å at Gly299 [21] (table 2 and figure 5). Therefore, it is not unexpected that these RMSFs were relatively high in the two Arg456His mutants (figure 5). Importantly, the CD loop leads to the C helix, which is thought to be a constituent of the interaction site with CPR. Moreover, the FG and HI loops link to the F and I helices, respectively, which are considered to be substrate recognition sites [17] (figure 4b). Thus, the flexibility of these loops is expected to be important for the catalytic function of CYP1A2. In the Arg456Hie mutants, the RMSF of the HI loop was extremely high (3.43 Å) (figure 5), while in the Arg456Hip mutant, the CD loops were so flexible that the RMSF value was 3.15 Å. Thus, neither the Arg456Hie nor the Arg456Hip mutants seem to have appropriate dynamics for the catalytic function. Moreover, an RMSF peak of 1.80 Å was observed at Gly102 of the BC loop in the Arg456Hip mutant, which was not observed in the wild type, with a value of 0.71 Å at Gly102. As the BC loop also constitutes part of the substrate recognition site, flexibility would impair the ability to recognize some substrates.

In order to identify the differences in interaction of Hie456 and Hip456, hydrogen bond frequency was calculated using the last 20 ns of trajectories (figure 7a and 7b). In the Arg456Hie mutants (figure 7a), the imidazole hydrogen of Hie456 formed hydrogen bonds with the hydroxyl group of Ser129 at a frequency of 19.3% (average distance: 2.89 Å). On the other hand, Hip456 did not interact with Ser129 in the Arg456Hip mutant (average distance: 6.2 Å). Although the Arg456Hip mutants could easily form hydrogen bonds by the charged imidazole, hydrogen bonds were not observed between Hip456 and any residue with a frequency of >1%. This result indicates that interactions of Hie456 and Hip456 were not the same as those of Arg456. In mammalian CYP structures, though arginine of the Cys-pocket is highly conserved [18], in some bacterial CYPs such as P450cam and P450terp, the corresponding arginine is substituted by histidine [19]. However, our results imply that His could not be used for the conserved arginine in the case of human CYP1A2.
Figure 5. Structural flexibilities for the wild type (black), Arg456Hie (blue) and Arg456Hip (magenta). Highly flexible loops are labeled. The mutated position was marked with a red dot.

Figure 6. Superimposition of structural flexibilities for Arg456Hie (top) and Arg456Hip (bottom) using last 20, 40, 60 ns MD trajectories. Highly flexible loops are labeled.

Table 3. Values of high RMSF peaks at BC, CD, FG, and HI loops in Arg456His mutants, calculated using last 20, 40, 60 ns of trajectories.

|                  | Gly102 of BC loop | Ser156 of CD loop | Phe246 of FG loop | Gly299 of HI loop |
|------------------|-------------------|-------------------|-------------------|-------------------|
| Arg456Hie mutant |                   |                   |                   |                   |
| 280-300 ns       | 0.75₁             | 2.39              | 1.84              | 3.43²             |
| 260-300 ns       | 0.83₁             | 2.44              | 1.72              | 3.06              |
| 240-300 ns       | 1.48              | 2.15              | 1.61              | 2.81              |
| Arg456Hip mutant |                   |                   |                   |                   |
| 280-300 ns       | 1.80              | 3.15³             | 1.26              | 1.51              |
| 260-300 ns       | 1.84              | 4.23³             | 1.73              | 1.67              |
| 240-300 ns       | 1.66              | 4.44⁴             | 1.61              | 1.69              |

₁ RMSF value at Asp103.
₂ RMSF value at Ser298.
₃ RMSF value at Ser155.
⁴ RMSF value at Ser154.
To study the cause of the differing interactions of Hie456 and Hip456, we compared the optimized structures including solvent water (figure 7c and 7d). In the Arg456Hie mutant, Hie456 interacted with one water molecule as well as Ser126 (figure 7c). On the other hand, Hip456 of the Arg456Hip mutant formed hydrogen bonds with two water molecules. This result implies that interaction with water molecules could change the hydrogen bond network of the Arg456Hie and Arg456Hip mutants.

Figure 7. Interacting residues and water molecules around Arg456 in the optimized structures. Distances between His456 and Ser126 were 2.89 Å in the Arg456Hie mutant (a) and 6.23 Å in the Arg456Hip mutant (b). Water molecules surrounding His456 in the Arg456Hie (c) and Arg456Hip (d) mutants. Residues and water displayed with ball-and-stick diagram, hydrogen bonds shown with black dashed line.

To study the cause of the differing interactions of Hie456 and Hip456, we compared the optimized structures including solvent water (figure 7c and 7d). In the Arg456Hie mutant, Hie456 interacted with one water molecule as well as Ser126 (figure 7c). On the other hand, Hip456 of the Arg456Hip mutant formed hydrogen bonds with two water molecules. This result implies that interaction with water molecules could change the hydrogen bond network of the Arg456Hie and Arg456Hip mutants.

4. Conclusion

Although we conducted 300-ns MD simulations for both Arg456Hie and Arg456Hip mutants, equilibrium structures could not be obtained, indicating that the Arg456Hie mutation has a large influence on the 3D structure of CYP1A2 independently of the protonation state. Moreover, the 3D structures of the Arg456His mutants were quite different from that of the wild-type CYP1A2, suggesting that the mutants could not maintain these structures and have the same enzymatic activities as the wild type. Finally, the Arg456Hie and Arg456Hip mutants were significantly different in their 3D structures, flexibilities, and interactions around His456. Therefore, the Arg456His mutation differently affect on the CYP1A2 structure because of the His456 protonation state. Although the RMSDs have not yet been converged by 300 ns MD simulations, the Arg456His mutation could largely affect the 3D structure of CYP1A2 regardless of the His456 protonation state.
Acknowledgments
This work was supported by a grant-in-aid for scientific research [17K08257] from the Japan Society for the Promotion of Science and Showa University Research Grant for Young Researchers.

References
[1] Nelson DR 1999 Cytochrome P450 and the individuality of species Arch. Biochem. Biophys. 369 1–10
[2] Nebert DW and Russell DW 2002 Clinical importance of the cytochromes P450 Lancet. 360 1155–62
[3] Anzenbacher P and Anzenbacherová E 2001 Cytochromes P450 and metabolism of xenobiotics Cell. Mol. Life. Sci. 58 737–47
[4] Guengerich FP 2008 Cytochrome P450 and chemical toxicology Chem. Res. Toxicol. 21 70–83
[5] Evans WE and Relling MV 1999 Pharmacogenomics: translating functional genomics into rational therapeutics Science. 286 487–91
[6] Bozina N, Bradamante V and Lovrić M 2009 Genetic polymorphism of metabolic enzymes P450 (CYP) as a susceptibility factor for drug response, toxicity, and cancer risk Arh. Hig. Rada. Toksikol. 60 217–42
[7] Groves JT 2005 Models and mechanisms of cytochrome P450 action. In: Cytochrome P450: Structure, Mechanism, and Biochemistry, ed P.R. Ortiz de Montellano (New York: Kluwer Academic/Plenum Publishers) pp.1–43
[8] Chang YT, Stiffelman OB, vakser IA, Loew GH, Bridges A and Waskell L 1997 Construction of a 3D model of cytochrome P450 2B4 Protein. Eng. 10 119–29
[9] Bridges A, Gruenke L, Chang YT, Vakser IA, Loew G and Waskell L 1998 Identification of the binding site on cytochrome P450 2B4 for cytochrome b5 and cytochrome P450 reductase J. Biol. Chem. 273 17036–49
[10] Hamdane D, Xia C, Im SC, Zhang H, Kim JJ and Waskell L 2009 Structure and function of a NADPH-Cytochrome P450 oxidoreductase in an open conformation capable of reducing cytochrome P450 J. Biol. Chem. 284 11374–84
[11] Shimizu T, Tateishi T, Hatano M and Fujii-Kuriyama Y 1991 Probing the role of lysines and arginines in the catalytic function of cytochrome P450d by site-directed mutagenesis. interaction with NADPH-cytochrome P450 reductase J. Biol. Chem. 266 3372–5
[12] Haines DC, Tomchick DR, Machius M and Peterson JA 2001 Pivotal role of water in the mechanism of P450BM-3 Biochemistry. 40 13456–65
[13] Schlichting I, Berendzen J, Chu K, Stock AM, Maves SA, Benson DE, Sweet RM, Ringe D, Petsko GA and Sligar SG 2000 The catalytic pathway of cytochrome p450cam at atomic resolution Science. 287 1615–22
[14] Prasad B, Mah DJ, Lewis AR and Plettner E 2013 Water oxidation by a cytochrome P450: mechanism and function of the reaction PLoS ONE. 25 e61897
[15] Ito M, Katono Y, Oda A, Hirasawa N and Hiratsuka M 2015 Functional characterization of 20 allelic variants of CYP1A2 Drug. Metab. Pharmacokinet. 30 247–52
[16] Saito Y, Hanioka N, Maekawa K, Isobe T, Tsuneto Y, Nakamura R, Soyama A, Ozawa S, Tanaka-Kagawa T, Jinno H, Narimatsu S and Sawada J 2005 Functional analysis of three CYP1A2 variants found in a Japanese population Drug. Metab. Dispos. 33 1905–10
[17] Gotoh O 1992 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 83–90
[18] Syed K and Mashele SS 2014 Comparative analysis of P450 signature motifs EXXR and CXG in the large and diverse kingdom of fungi: identification of evolutionarily conserved amino acid patterns characteristic of P450 family PLoS ONE. 9 e95616
[19] Hasemann CA, Kurumbail RG, Boddupalli SS, Peterson JA and Deisenhofer J 1995 Structure and function of cytochromes P450: a comparative analysis of three crystal structures
Structure. 3 41–62
[20] Guallar V, Baik MH, Lippard SJ and Friesner RA 2003 Peripheral heme substituents control the hydrogen-atom abstraction chemistry in cytochromes P450 Proc. Natl. Acad. Sci. U. S. A. 100 6998–7002
[21] Watanabe Y, Fukuyoshi S, Hiratsuka M, Yamaotsu N, Hirono S, Takahashi O and Oda A 2016 Prediction of three-dimensional structures and structural flexibilities of wild-type and mutant cytochrome P450 1A2 using molecular dynamics simulations J. Mol. Graph. Model. 68 48–56
[22] Sansen S, Yano JK, Reynald RL, Schoch GA, Griffin KJ, Stout CD and Johnson EF 2007 Adaptations for the oxidation of polycyclic aromatic hydrocarbons exhibited by the structure of human P450 1A2 J. Biol. Chem. 282 14348–55
[23] Kobayashi K, Takahashi O, Hiratsuka M, Yamaotsu N, Hirono S, Watanabe Y and Oda A 2014 Evaluation of influence of single nucleotide polymorphisms in cytochrome P450 2B6 on substrate recognition using computational docking and molecular dynamics simulation PLoS ONE 9 e96789
[24] Oda A, Yamaotsu N and Hirono S 2004 Studies of binding modes of (S)-mephénytoin to wild types and mutants of cytochrome P450 2C19 and 2C9 using homology modeling and computational docking Pharm. Res. 21 2270–8
[25] Fukuyoshi S, Kometani M, Watanabe Y, Hiratsuka M, Yamaotsu N, Hirono S, Manabe N, Takahashi O and Oda A 2016 Molecular dynamics simulations to investigate the influences of amino acid mutations on protein three-dimensional structures of cytochrome P450 2D6.1, 2, 10, 14A, 51, and 62, PLoS ONE 11 e0152946
[26] Joung IS and Cheatham TE III 2008 Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations J. Phys. Chem. B 112 9020–41
[27] Darden TA, York DM and Pedersen LG 1993 Particle mesh Ewald: An N ⋅log(N) method for Ewald sums in large systems J. Chem. Phys. 98 10089–92
[28] Oda A, Yamaotsu N and Hirono S 2005 New AMBER force field parameters of heme iron for cytochrome P450s determined by quantum chemical calculations of simplified models J. Comput. Chem. 26 818–26
[29] Giammona DA 1984 Ph.D. thesis, University of California, Davis.

[30] Case DA, Darden TA, CheathamTE III, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, Roberts B, Hayik S, Roitberg A, Seabra G, Swails J, Götz AW, Kolossváry I, Wong KF, Paesani F, Vanicek J, Wolf RM, Liu J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Cai Q, Ye X, Wang J, Hsieh M-J, Cui G, Roe DR, Mathews DH, Seetin MG, Salomon-Ferrer R, Saigi C, Babin V, Luchko T, Gurjarov S, Kovalenko A and Kollman PA AMBER 12 2012 University of California, San Francisco.
[31] Wang J, Wolf RM, Caldwell JW, Kollman PA and Case DA 2004 Development and testing of a general amber force field J. Comput. Chem. 25 1157–74
[32] Ryckaert JP, Ciccotti G and Berendsen HJC 1977 Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes J. Comput. Phys. 23 327–41
[33] Watanabe Y, Fukuyoshi S, Kato K, Hiratsuka M, Yamaotsu N, Hirono S, Gouda H and Oda A 2017 Investigation of substrate recognition for cytochrome P450 1A2 mediated by water molecules using docking and molecular dynamics simulations J. Mol. Graph. Model. 74 326–36
[34] Muralidhara BK, Negi S, Chin CC, Braun W, Halpert JR 2006 Conformational flexibility of mammalian cytochrome P450 2B4 in binding imidazole inhibitors with different ring chemistry and side chains. Solution thermodynamics and molecular modeling J. Biol. Chem. 281 8051–61

12