ENGINEERING BACTERIAL CYTOCHROME P450 BM3 INTO A
PROTOTYPE WITH HUMAN P450 ENZYME ACTIVITY USING INDIGO FORMATION

Sun-Ha Park, Dong-Hyun Kim, Dooil Kim, Dae-Hwan Kim, Heung-Chae Jung, Jae-Gu Pan, Taeho Ahn, Donghak Kim, and Chul-Ho Yun*

School of Biological Sciences and Technology, Chonnam National University, Gwangju 500-757, Republic of Korea (S.-H.P., Do.-H.K., Da.-H.K., C.-H.Y.); Systems Microbiology Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 305-806, Republic of Korea (Doo.K., H.-C.J., J.-G.P.); Department of Biochemistry, College of Veterinary Medicine, Chonnam National University, Gwangju 500-757, Republic of Korea (T.A); Department of Biological Sciences, Konkuk University, Seoul, 413-701, Republic of Korea (Don.K.)
Running title: Bacterial P450 BM3 mutants with human P450 enzyme activity

Correspondence: Prof. Chul-Ho Yun, School of Biological Sciences and Technology, Chonnam National University, Gwangju 500-757, Republic of Korea; Tel: 82-62-530-2194; Fax: 82-62-530-2199; E-mail: chyun@jnu.ac.kr

Number of text pages: 19
Number of tables: 1
Number of figures: 3
Number of references: 37
Number of words in Abstract: 239
Number of words in Introduction: 500
Number of words in Results and Discussion: 3228
Supplemental Data: 2 supplemental tables and 7 supplemental figures

Abbreviations used: P450 or CYP, cytochrome P450; CPR, NADPH-P450 reductase; CYP102A1, cytochrome P450 BM3; PCR, polymerase chain reaction; IPTG, isopropyl-β-D-thiogalactopyranoside; δ-ALA, δ-aminolevulinic acid; HPLC, high performance liquid chromatography; MD, molecular dynamics; TTNs, total turnover numbers; 7-OH coumarin, 7-hydroxycoumarin; 3-OH coumarin, 3-hydroxycoumarin; HPLC, high performance liquid chromatography; LC/MS/MS, liquid chromatography-tandem mass spectrometry; PhOD, phenacetin O-deethylation; EROD, 7-ethoxyresorufin O-deethylation
Abstract

Human P450 enzymes metabolize a variety of endogenous and xenobiotic compounds including steroids, drugs, and environmental chemicals. Here we examine the possibility that bacterial P450 BM3 (CYP102A1) mutants with indole oxidation activity have the catalytic activities of human P450 enzymes. Error-prone polymerase chain reaction was carried out on the heme domain coding region of the wild-type gene to generate a CYP102A1 DNA library. The library was transformed into Escherichia coli for expression of the P450 mutants. A colorimetric colony-based method was adopted for primary screening of the mutants. When the P450 activities were measured at the whole-cell level, some of the blue colonies, but not the white colonies, possessed apparent oxidation activity toward coumarin and 7-ethoxycoumarin, which are typical human P450 substrates that produce fluorescent products. Coumarin is oxidized by the CYP102A1 mutants to produce two metabolites, 7-hydroxycoumarin and 3-hydroxycoumarin. In addition, 7-ethoxycoumarin is oxidized to 7-hydroxycoumarin and 3-hydroxy, 7-ethoxycoumarin by O-deethylation and 3-hydroxylation reactions, respectively. Highly active mutants are also able to metabolize several other human P450 substrates, including phenacetin, ethoxyresorufin, and chlorzoxazone. These results indicate that indigo formation provides a simple assay for identifying CYP102A1 mutants with a greater potential for human P450 activity. Furthermore, our computational findings suggest a correlation between the stabilization of the binding site and the catalytic efficiency of CYP102A1 mutants towards coumarin: the more stable the structure in the binding site, the lower the energy barrier and the higher the catalytic efficiency.
Introduction

Cytochrome P450 (P450, also known as CYP for specific isoforms) enzymes constitute a large family of enzymes that are remarkably diverse oxygenation catalysts found in archaea, bacteria, fungi, plants, and animals. P450s function primarily in the oxidation of various xenobiotics and endogenous compounds, especially in mammals. Due to the catalytic diversity and broad substrate range of P450s, they are attractive biocatalyst candidates for the production of fine chemicals, including pharmaceuticals. Despite the potential use of mammalian P450s in various biotechnology fields, they are not suitable as biocatalysts because of their low stability and low catalytic activity (Yun et al., 2007). P450 BM3 (CYP102A1) from *Bacillus megaterium* was the first P450 discovered to be fused to its redox partner, a mammalian-like diflavin reductase. The fusion of these two enzymatic activities makes soluble CYP102A1 an ideal model for mammalian, particularly human P450 enzymes (Munro et al., 2002). Recently, through rational design or directed evolution, wild-type CYP102A1 has been engineered to oxidize compounds showing little or no structural similarity to the natural fatty acid substrates of CYP102A1 (Yun et al., 2007; Kim et al., 2008; Kim et al., 2009).

Previously, a screening procedure was developed for the directed evolution of human CYP2A6 using indigo formation (Gillam et al., 1999). As *Escherichia coli* cultures produce indigo in the absence of supplemental indole, the rich expression media used in this procedure is likely to have provided sufficient indole as a substrate via tryptophan degradation. Indigo formation from indole oxidation was first used to screen CYP102A1 mutants by saturation mutagenesis at specific sites. A set of selected CYP102A1 mutants (F87V, F87V/L188Q, and F87V/L188Q/A74G) could efficiently hydroxylate indole, leading to the production of indigo and indirubin (Li et al., 2000). Recently, wild-type and mutant
forms of the bacterial CYP102A1 have been found to metabolize various drugs (Yun et al., 2007; Damsten et al., 2008). 7-Ethoxycoumarin, a substrate of CYP2E1, CYP2A6, and CYP1A2, was also found to be oxidized through reactions similar to those catalyzed by CYP102A1 mutants (Kim et al., 2008; Kim et al., 2005; Kim et al., 2006). Indole is a substrate for CYP2A6 (Wu et al., 2005). CYP2A6 mutants generated by random and site-directed mutagenesis were able to oxidize bulky substituted indole compounds, which are not substrates for the wild-type enzyme.

In the present study, we used random mutagenesis of wild-type CYP102A1 to generate mutants with catalytic activities similar to human P450 enzymes. Error-prone polymerase chain reaction (PCR) was carried out on the heme domain coding regions of the wild-type gene to generate a CYP102A1 DNA library. The library was transformed into *E. coli* for expression of the CYP102A1 mutants. A colorimetric colony-based method was adopted for primary screening of the mutants. When the P450 activities were measured at the whole-cell level, some of the blue colonies, but not the white colonies, showed apparent oxidation activity toward typical human P450 substrates, including coumarin, 7-ethoxycoumarin, phenacetin, and chlorzoxazone (Fig. 1). The enzymatic and physical properties of selected CYP102A1 mutants showing human P450 activity were further studied.
Materials and Methods

Chemicals. 7-Ethoxycoumarin, coumarin, 7-hydroxycoumarin (7-OH coumarin), 3-hydroxycoumarin (3-OH coumarin), phenacetin, acetaminophen, p-nitrophenol, and chlorzoxazone were purchased from Sigma-Aldrich (St. Louis, MO). 7-Ethoxyresorufin and 7-hydroxyresorufin were obtained from Invitrogen (Carlsbad, CA). All other chemicals and solvents used in this study were of analytical grade or high quality and were purchased from commercial suppliers and used without further purification.

Generation of the CYP102A1 libraries. Error-prone PCR was carried out on the heme domain coding regions (base pairs 428-888) of wild-type CYP102A1 to generate the first CYP102A1 DNA library. We constructed the pCW BM3-KX vector which includes KpnI and XhoI restriction sites by using silent mutagenesis. Oligonucleotide primers were used to introduce KpnI and XhoI restriction sites: KpnI forward, 5’-GAGCATATTGAGGTACCGGAAGAC-3’; XhoI reverse, 5’-TACTAGAACTCGAGCTGCTTCTTC-3’. The P450 heme domain region of pCW BM3-KX was randomly mutated using a GeneMorph® II random mutagenesis kit (Stratagene). The 50-µl reaction volume contained 50 pmol of each primer (forward primer 5’-GAGCATATTGAGGTACCGGAAGAC-3’ and reverse primer 5’-TACTAGAACTCGAGCTGCTTCTTC-3’), 70 ng of the template plasmid DNA (pCW BM3-KX), 2.5 U of Mutazyme® DNA polymerase, 0.2 µM of each dNTP and 10X Mutazyme® reaction buffer. The PCR reaction was initiated at 95°C for 2 min and run through 30 thermocycles of 95°C for 30 sec, 50°C for 30 sec, and 72°C for 1 min. After completion of the 30 cycles, the reaction medium was held at 72°C for 10 min. The amplified PCR library fragments were purified and cloned into the pCW BM3-KX vector using the KpnI and XhoI restriction sites. In the same way, a second library was constructed using the F162I/M237I
mutant, which had the highest 7-ethoxycoumarin oxidation activity among a set of mutants from the first screen as the template.

**Screening of the CYP102A1 libraries.** Randomized plasmid libraries were transformed into *E. coli* DH5αF’IQ. The *E. coli* cells expressing the CYP102A1 mutants were spread on Luria-Bertani (LB) broth agar expression plates, incubated at 37°C for 16 h and then stored at 4°C for approximately 2 weeks. LB agar expression plates contained 25 g of LB broth per liter (BD Biosciences), 15 g of Bacto agar per liter (BD Biosciences), and 100 µg of ampicillin per ml. No additives including indole, isopropyl-β-D-1-thiogalactopyranoside (IPTG), or δ-aminolevulinic acid (δ-ALA) were added to the plates. Approximately 10^3 blue-colored colonies were transferred into 96 deep well plates containing 0.3 ml of LB media and 100 µg/ml of ampicillin. The bacteria were initially cultured for 8 h at 37°C and 450 rpm (HT-MegaGrow incubator; Bioneer Co., Daejeon, Korea). Aliquots of the cell cultures (30 µl) were transferred to single wells of 96 deep well plates filled with 0.5 ml Terrific Broth expression media containing additives (10 µM IPTG, 0.5 mM δ-ALA) and ampicillin. These aliquots were cultured for 17 h at 30°C and 450 rpm. For primary screening, the whole cell activities of the mutants were measured using 7-ethoxycoumarin and coumarin as substrates. The cells were collected by centrifugation at 3,500 rpm for 10 min at 4°C and frozen at -20°C. Cell pellets were resuspended in 300 µl phosphate buffer (0.1 M, pH 8.0) containing 10 mM MgCl₂. After a 60-min incubation at 30°C and 400 rpm, 100 µM of the substrate was added, and the reactions were initiated by the addition of a reduced nicotinamide adenine dinucleotide phosphate (NADPH) generating system (final concentration of 0.5 mM NADP⁺, 10 mM glucose-6-phosphate, and 1.0 UI glucose-6-phosphate dehydrogenase). The plates were then incubated for 30 min at 37°C and 350 rpm. The reaction was terminated by the addition of 100 µl of 20% trichloroacetic acid (w/v), and the mixtures were centrifuged at
3000 rpm for 5 min at 4°C. Aliquots of the supernatant (50 µl) were transferred into a new black 96-well plate containing 150 µl Tris-HCl (pH 9.0), and the fluorescence (excitation at 355 nm and emission at 460 nm) of the mixture was measured using a microplate reader (Infinite M200, Tecan). Fifty mutants were selected for a second round of screening based on their high 7-ethoxycoumarin and coumarin oxidation activities.

**Protein expression and purification.** Wild-type and selected mutants of CYP102A1 were expressed in *E. coli* and purified as described (Kim et al., 2008). The CYP102A1 concentration was determined from the CO-difference spectra as described by Omura and Sato (1964) using \( \varepsilon = 91 \text{ mM/cm} \). For all of the wild-type and mutant enzymes, a typical culture yielded between 300 and 700 nM P450. The expression level of wild-type and mutant CYP102A1 was typically within the range of 1.0 to 2.0 nmol P450/mg cytosolic protein.

**P450 catalytic activity assays.** Several typical substrates of human P450s were used to examine the catalytic activities of the wild-type and the mutant forms of CYP102A1, as described below (Fig. 1 and 2). Typical steady-state reactions for 7-ethoxycoumarin oxidation, coumarin oxidation, phenacetin \( O \)-deethylation (PhOD), 7-ethoxyresorufin \( O \)-deethylation (EROD), chlorzoxazone 6-hydroxylation, and \( p \)-nitrophenol hydroxylation were composed of 50 pmol of CYP102A1 in 0.25 ml of 100 mM potassium phosphate buffer (pH 7.4) and the specified amount of substrate. To determine the turnover numbers of wild-type CYP102A1 and all of the tested mutants, we used substrate concentrations of 2.0 mM, 2.0 µM, 200 µM, and 800 µM for phenacetin, 7-ethoxyresorufin, chlorzoxazone, and \( p \)-nitrophenol, respectively.

**Oxidation of coumarin and 7-ethoxycoumarin –** Assays of 7-ethoxycoumarin and coumarin oxidation were performed using the wild-type CYP102A1 and its mutants. The reaction mixture contained 50 pmol of CYP102A1, 100 mM potassium phosphate buffer (pH
7.4), and 2 mM of the substrate in a total volume of 0.25 ml. Both substrate stocks (200 mM) were prepared in CH$_3$CN and diluted into the enzyme reaction such that the final organic solvent concentration was always less than 1% (v/v). The sample was pre-incubated for 10 min, and the reaction was initiated by the addition of the NADPH-generating system (0.5 mM NADP$^+$, 10 mM glucose 6-phosphate, and 1.0 IU glucose 6-phosphate dehydrogenase per ml) and continued in a water bath at 37°C. After a 30 min incubation, the reaction was stopped by the addition of 500 µl of cold CH$_2$Cl$_2$. Following centrifugation (3000 rpm for 10 min), 300 µl aliquots of the organic layer from each incubation were transferred to a clean glass tube, and the CH$_2$Cl$_2$ was removed under a N$_2$ stream. The metabolites of 7-ethoxycoumarin and coumarin were analyzed by high performance liquid chromatography (HPLC) and liquid chromatography-tandem mass spectrometry (LC/MS/MS), as described below.

**EROD** – Activities of EROD were measured using a fluorescence assay (Burke and Mayer, 1983). The reaction mixture contained 50 pmol of the CYP102A1 enzyme, 100 mM potassium phosphate buffer (pH 7.4), an NADPH-generating system, and 2 µM 7-ethoxyresorufin in a total volume of 0.25 ml. Pre-incubation with the reaction mixture was generally carried out for 5 min at 37°C. The reaction was initiated by the addition of the NADPH-generating system, incubated for 5 min at 37°C, and terminated with 0.5 ml of CH$_3$OH. Metabolites were measured using fluorescence and a resorufin standard (excitation at 535 nm and emission at 585 nm).

**PhOD** – The reaction mixture consisted of 50 pmol of CYP102A1, 100 mM potassium phosphate buffer (pH 7.4), an NADPH-generating system, and 2 mM phenacetin in a total volume of 0.25 ml. PhOD activity was determined by HPLC, as previously described (Yun et al., 2000). Incubations with the reaction mixture were performed for 20 min at 37°C, terminated with 50 µl of 17% HClO$_4$ and centrifuged ($10^3$ g for 10 min). Then 0.5 ml of a
A mixture of CHCl₃ and 2-propanol (6:4, v/v) was added to the supernatant to extract the products, followed by centrifugation (twice at 10³ g). The organic layers were combined, and the solvent was removed under N₂ gas. The products were analyzed by HPLC using a Gemini C₁₈ column (4.6×150 mm, 5 µm, Phenomenex) with a mobile phase of H₂O/CH₃OH/CH₃CO₂H (65:35:0.1, v/v/v; flow rate of 1.0 ml/min) and monitoring at A₂₅₄.

Chlorzoxazone 6-hydroxylation – The reaction mixture consisted of 50 pmol of CYP102A1, 100 mM potassium phosphate buffer (pH 7.4), an NADPH-generating system, and 200 µM chlorzoxazone in a total volume of 0.25 ml. Chlorzoxazone 6-hydroxylation activity was determined by HPLC, as previously described (Guengerich et al., 1991). Incubation with the reaction mixture was performed for 30 min at 37°C, terminated with 25 µl of 43% H₃PO₄ and 500 µl cold CH₂Cl₂, and centrifuged at 10³ g for 10 min. The organic layers were combined, and the solvent was removed under N₂ gas. The products were analyzed by HPLC using a Luna C₈ column (4.6×150 mm, 5 µm, Phenomenex) with a mobile phase of H₂O/CH₃CN/H₃PO₄ (72.5:27:0.5, v/v/v; flow rate of 1.25 ml/min) and monitoring at A₂₈₇.

p-Nitrophenol hydroxylation – The reaction mixture consisted of 50 pmol of CYP102A1, 100 mM potassium phosphate buffer (pH 7.4), an NADPH-generating system, and 800 µM p-nitrophenol in a total volume of 0.25 ml. p-Nitrophenol hydroxylation activity was determined by the spectrophotometric method previously described (Chang et al., 2006) with a minor modification. Incubation with the reaction mixture was performed for 5 min at 37°C, terminated with 100 µL of 20% trichloroacetic acid and centrifuged (10³ g for 10 min). After 0.1 ml of 2 M NaOH was added to the supernatant (0.2 ml), the absorbance at 535 nm was measured. The amount of product was determined using a p-nitrocatechol standard.

HPLC analysis of 7-ethoxycomarin and coumarin metabolites. The 7-
ethoxycoumarin metabolites were analyzed as previously described (Kim et al., 2008). Coumarin products were dissolved in 50% CH$_3$CN and analyzed by gradient HPLC using the Germini C$_{18}$ column. The mobile phase consisted of 10% CH$_3$CN containing 0.1% formic acid (buffer A) and 100% CH$_3$CN containing 0.1% formic acid (buffer B). The column was eluted at a flow rate of 1 ml/min by a gradient pump (LC-20AD; Shimadzu, Kyoto, Japan) with the following gradient: 0-2 min, isocratic with 30% buffer B; 2-8 min, 30%-35% buffer B gradient; 8-10 min, 35%-45% buffer B gradient; 10-15 min, isocratic with 45% buffer B; 15-17 min, 45%-30% buffer B gradient; 17-20 min, isocratic with 30% buffer B (all v/v). The absorbance at 330 nm was monitored.

**LC/MS/MS Analysis of Coumarin Metabolites.** LC/MS/MS analysis was performed on a Waters Synapt HDMS system (Waters, Milford, MA, USA) in the Korea Basic Science Institute facility (Gwangju, Korea). Analysis was performed in the ESI negative ion mode using an Acquity UPLC BEH C$_{18}$ column (1.7 µm; 2.1 X 100 mm). All analyses were performed using a gradient from buffer A (10% CH$_3$CN containing 0.1% formic acid) to buffer B (100% CH$_3$CN containing 0.1% formic acid). The following gradient program was used with a flow rate of 0.4 µl/min, and the sample was injected onto the column using the following A:B solvent mixtures: (70:30, v/v) for 0-1 min, (60:40, v/v) for 1-2 min, (50:50, v/v) for 2-2.5 min, (30:70, v/v) for 2.5-3.5 min, (10:90, v/v) for 3.5-4 min, (0:100, v/v) for 4.5-5.5 min, and (70:30, v/v) for 5.5-6 min. The temperature of the column was maintained at 30°C. ESI conditions were as follows: capillary 3.1 kV, cone voltage 40 V, source temperature of 80°C, desolvation temperature of 300°C, and desolvation gas flow of 450 l/h.

**Steady-state kinetic assays for 7-ethoxycoumarin and coumarin oxidation reactions.** To determine the kinetic parameters of the CYP102A1 mutants, we used 0.01 to 2 mM of 7-
ethoxycoumarin and coumarin. The products were analyzed by HPLC as described above. The kinetic parameters ($k_{cat}$ and $K_m$) were determined using nonlinear regression analysis with GraphPad PRISM software (GraphPad, San Diego, CA, USA). The data were fit to the standard Michaelis-Menten equation: $v = \frac{k_{cat}[E][S]}{([S] + K_m)}$. The equation is the standard Michaelis-Menten equation, where the velocity of the reaction is a function of the rate-limiting step in turnover ($k_{cat}$), the enzyme concentration ([E]), substrate concentration ([S]), and the Michaelis constant ($K_m$).

**Mutant modeling and automated ligand docking.** The 3-dimensional structure models of two mutants of CYP102A1, F162I/H236R and F162I/M185T/L188P/M237I, were constructed based on the X-ray structure of wild-type CYP102A1 (PDB code: 1BU7) (Sevrioukova et al., 1999) using MODELLER (Sali and Blundell, 1993) and were optimized using FoldX (Schymkowitz et al., 2005). The structure of the coumarin molecule was taken from the Cambridge Structural Database. The partial atomic charges were fit to the electrostatic potential computed by ab initio GAMESS calculations performed at the 6-31 G accuracy level (Schmidt et al., 1993).

Docking of coumarin was carried out using the flexible docking module based on the MMFF (Halgren, 1999) force field in the Discovery Studio environment (Accelrys Inc.). The active site residues around the heme pocket of the CYP102A1 mutants and coumarin were allowed to flex during the docking process. The binding site pockets for automated docking were assigned via the interaction of the CYP102A1 mutant with a set of 10,000 random coumarin conformers. The most energy-favorable coumarin conformer and its binding receptor were selected for further analysis. To better represent the overall strength of the hydrogen bonding between coumarin and the active site residues of the binding cavity in a docked complex, the hydrogen-bonding energy associated with each simulated H--O/H--N
distance was estimated using a module implemented in the AutoDock 4.0 program suite (Huey et al., 2007).

**Molecular dynamics simulation.** The molecular dynamics (MD) simulations were performed with the AMBER 10 program (Pearlman et al., 1995; Case et al., 2005). For each of the four dockings of coumarin with the CYP102A1 mutants (F162I/H236R and F162I/M185T/L188P/M237I) in vacuo MD trajectories were generated. The docking of coumarin to the CYP102A1 mutants (F162I/H236R and F162I/M185T/L188P/M237I) was initially performed according to two predicted topological binding sites (I or II) by several algorithms (Huang and Schroeder, 2008). The in vacuo MD runs were then performed at a constant volume and temperature with no periodic condition applied, as we were concerned with only studying the docking of the coumarin molecule to the mutated CYP102A1 enzymes. Each trajectory was equilibrated for 500 ps with a time step of 0.001 ps. Then a 2,500 ps run was performed, and the trajectory data were collected. The average conformational binding energy values were computed by averaging the total (non-polar + polar) energy values over the 2,500-ps run and then using this average to calculate the interaction energy values. A 10-Å cut-off for coulombic and long range forces were adopted in each simulation. A relative dielectric constant value of 1.0 was employed in all of the simulations.
Results and Discussion

Construction of the CYP102A1 libraries and screening the mutants with indigo formation. Error-prone PCR was used to construct randomized libraries and a colorimetric colony-based method was used in the screening of the mutants (Nakamura et al., 2001; Whitehouse et al., 2008; Zhang et al., 2009). First, error-prone PCR was used to generate a randomized library using wild-type CYP102A1 as the template. The heme domain coding region between the KpnI and XhoI restriction sites was randomly mutated. A colorimetric colony-based method was adopted for primary screening of the mutants. Approximately 300 blue colonies with indigo formation were transferred to 96 deep well plates and grown as previously described (Li et al., 2008) (Supplemental Data Fig. S1). As 7-ethoxycoumarin and coumarin are well recognized substrates of human P450 enzymes, whole cell oxidation activity of the mutants on the conversion of 7-ethoxycoumarin and coumarin to 7-hydroxycoumarin, a fluorescent product, was measured (Hirano et al., 2002). Colonies showing these catalytic activities were selected, and each mutated enzyme was expressed in E. coli for purification. The catalytic activities of the mutants on 7-ethoxycoumarin and coumarin were compared (Supplemental Data Fig. S2). The F162I/M237I mutant was found to have the highest activity toward both of the substrates and was selected as the template for the generation of a second library. The second screening was performed in the same way as the first screening. After error-prone PCR, approximately $10^3$ blue colonies were screened, and fluorescent-based whole cell assays were performed. The catalytic activities of the selected candidates (50 colonies) were confirmed by HPLC analysis. Through the iterative screenings, nine mutants with high activity for both 7-ethoxycoumarin and coumarin oxidation were finally chosen and subsequently characterized by nucleotide sequence
Oxidation of 7-ethoxycoumarin and coumarin by the CYP102A1 mutants. Wild-type CYP102A1, F162I/M237I, and nine other promising mutants were purified to measure the rate of 7-ethoxycoumarin oxidation at a fixed substrate concentration (2 mM). Two major metabolites, 7-hydroxycoumarin and 3-hydroxy, 7-ethoxycoumarin, were analyzed by HPLC (Kim et al., 2006). All of the mutants had apparent catalytic activity toward 7-ethoxycoumarin (Supplemental Data Fig. S3 A), while wild-type CYP102A1 did not possess any apparent catalytic activity for this substrate (<0.1 min⁻¹ for O-deethylation and 3-hydroxylation). The F162I/H236R and F162I/M185V/M237I mutants had high rates of O-deethylation (19.7 min⁻¹ and 17.8 min⁻¹, respectively), while the F162I/K187E/M237I and F162I/M185T/L188P/M237I mutants had high 3-hydroxylation activity (38.0 min⁻¹ and 41.6 min⁻¹, respectively). The F162I single mutant, and the L148I/F162I double mutant showed little effect on O-deethylation and 3-hydroxylation activity compared to the other mutants, and these mutants had low O-deethylation and 3-hydroxylation activity. Like the human P450 enzymes (Kim et al., 2006) and the recently reported mutants of CYP102A1 (Kim et al., 2008), all of the mutants examined in the present study showed a higher rate of 3-hydroxylation compared to their O-deethylation activity.

The catalytic activity of wild-type CYP102A1 and its mutants was also examined using coumarin, a typical P450 substrate. Coumarin is oxidized by the CYP102A1 mutants to produce two major metabolites (7-hydroxycoumarin and 3-hydroxycoumarin) and two unknown minor products (Fig. 1). The identities of the major metabolites and the substrate were verified by comparing the results of HPLC and LC/MS/MS with standard compounds (Supplemental Data Fig. S4). CYP2A6, the major enzyme for the hydroxylation reaction of coumarin in the human liver, catalyzes only the 7-hydroxylation reaction and not the 3-
hydroxylation reaction (Kim et al., 2005; Yun et al., 2005). However, unlike the CYP2A6 enzyme, all of the CYP102A1 mutants showed a preference for the 3-hydroxylation reaction over the 7-hydroxylation. 3-hydroxycoumarin is reported to be produced as a major metabolite of coumarin by CYP3A and CYP1A in humans, while CYP2A6 does not catalyze 3-hydroxycoumarin formation (Born et al., 2002). Our results obtained for the CYP102A1 mutants for coumarin 3-hydroxylation are similar to those for 7-ethoxycoumarin 3-hydroxylation (Supplemental Data Fig. S3). Wild-type CYP102A1 showed no catalytic activity for coumarin. The F162I/M185T/L188P/M237I mutant, which showed high activity for O-deethylation and 3-hydroxylation of 7-ethoxycoumarin, also had high activity for coumarin hydroxylation. The F162I and L148I/F162I mutants, which possessed low 7-ethoxycoumarin oxidation activity, also showed low coumarin oxidation activity.

Seven high-activity mutants were chosen and used to measure the kinetic parameters of the coumarin and 7-ethoxycoumarin oxidation reactions. The mutants displayed a more than 3-fold variation in $k_{cat}$ for both 7-hydroxylation and 3-hydroxylation reactions of coumarin. The F162I/K187E/M237I and F162I/M185T/L188P/M237I mutants were the most effective with 20 min$^{-1}$ and 4 min$^{-1}$ for 3-hydroxylation and 7-hydroxylation of coumarin, respectively (Table 1). For 7-ethoxycoumarin, the F162I/M185T/L188P/M237I and L148I/F162I mutants were the most effective with 41 min$^{-1}$ and 14 min$^{-1}$ for 3-hydroxylation and O-deethylation, respectively (Supplemental Data Table S1). The catalytic activity of the CYP102A1 mutants for coumarin 7-hydroxylation is comparable to that of CYP2A6 purified from a human liver (Yun et al., 1991).

Next, we determined the total turnover numbers (TTNs) of the CYP102A1 mutants for the coumarin and 7-ethoxycoumarin oxidation reactions. The overall range of 7-ethoxycoumarin O-deethylation was between 500 and 2,000. The overall range of 7-
ethoxycoumarin 3-hydroxylation was between 1,500 and 3,500. For 7-ethoxycoumarin oxidation, most of the mutants, except for F162I/E228K, showed an increase in product formation with increasing incubation time (Supplemental Data Fig. S5 A, B). In terms of the coumarin oxidation reactions, all of the tested mutants showed an increase in product formation with increased incubation time (Supplemental Data Fig. S5 C, D). In general, the CYP102A1 mutants seem to prefer the hydroxylation reaction over the O-deethylation reaction.

**Oxidation of other human P450 substrates is catalyzed by the CYP102A1 mutants.** The catalytic activities of the CYP102A1 mutant enzymes towards four substrates that are known to be metabolized by human P450 enzymes were investigated. First, the ability of wild-type CYP102A1 and a set of CYP102A1 mutants to catalyze human P450 substrates was measured at fixed substrate concentrations (phenacetin, 2.0 mM; 7-ethoxyresorufin, 2.0 µM; chlorozoxazone, 200 µM; p-nitrophenol, 800 µM; Fig. 2).

The CYP102A1 mutant enzymes converted phenacetin into one major metabolite. HPLC analysis identified the metabolite as acetaminophen (Fig. 1C) and LC-MS comparison (data not shown) with the authentic standard compound confirmed that the metabolite was acetaminophen. Known oxidation products of phenacetin that are catalyzed by human CYP1A2 are acetaminophen and acetol (Yun et al., 2000). Acetol was not formed by any of the CYP102A1 mutant enzymes. The turnover numbers for the entire set of 10 mutants for O-deethylation of phenacetin varied over a wide range. We found that wild-type CYP102A1 showed very low activity (0.40 min\(^{-1}\) for PhOD) under the test conditions (Fig. 2A), while the F162I/E228K mutant showed the highest activity (42 min\(^{-1}\)) among the mutants.

The EROD activity of the CYP102A1 mutants was also examined (Fig. 2B). Spectrofluorometric analysis identified the metabolite was resorufin and the identification...
was compared to an authentic resorufin standard compound (data not shown). Although the F162I/E228K mutant showed the highest EROD reaction rate (0.57 min⁻¹) among all of the mutants, its activity was relatively low compared to the other substrates tested. Interestingly, the F162I/F165L/M177T/M237I mutant did not show any apparent activity toward 7-ethoxyresorufin, but did show high hydroxylation activity toward chlorzoxazone (9.5 min⁻¹). All of the CYP102A1 mutants tested here could catalyze the hydroxylation reaction of chlorzoxazone and p-nitrophenol, which are typical human CYP2E1 substrates. The wild-type CYP102A1 did not possess any apparent activity toward these substrates (Fig. 2 C and D).

**Stability of the CYP102A1 mutants.** To estimate the stability of the mutants, we determined the thermal stability ($T_m$, °C) of wild-type CYP102A1 and the mutants (Supplemental Data Fig. S6). The $T_m$ value of wild-type CYP102A1 was determined to be 52.3 ± 0.22°C, while the $T_m$ value of the mutants varied. The F162I/M185T/L188P/M237I mutant, which had the highest 7-ethoxycoumarin and coumarin oxidation activities among the tested mutants, showed the lowest stability (47.5 ± 0.1°C). Recently, it was reported that the L188P mutation is associated with a destabilizing effect and causes a large conformational change in the enzyme during catalysis (Fasan et al., 2008). This conformational change appears to have a significant impact on the specificity and activity of the enzyme. Therefore, the mutant with the highest activity showed the lowest thermal stability. The mutants that contained substitutions of residues near L188 (F162I/K187E, F162I/K187E/M237I, and F162I/K187E/L188P/M237I) also showed a relatively low stability. On the other hand, the F162I/M237I and F162I/E228K mutants showed increased stability compared to the wild-type CYP102A1 (54.4 ± 0.4°C and 55.9 ± 0.3°C, respectively; Supplemental Data Fig. S6).
Models of the docking of coumarin to CYP102A1 and its mutants. The findings presented in the wet experiments above suggest that the structural basis for coumarin metabolism can be elucidated by the docking of coumarin to its probable binding site (either I or II) in the CYP102A1 mutants, F162I/H236R and F162I/M185T/L188P/M237I (Fig. 3B). Therefore, we docked coumarin to models of these two CYP102A1 mutants (Fig. 3). Several computational investigations by ligand docking and MD simulations were performed to study the relationship between the protein structure of CYP2A6 and the catalytic activity towards coumarin (Li et al., 2009; Lewis et al., 2006; Sansen et al., 2007). Previous studies assuredly gave close insight into the best pose orientation of coumarin docked in the desired orientation for hydroxylation in the topological aspect for the docking model complexes. The key residues involved in the formation of the active-site pocket of P450 can form a hydrogen bond with the carbonyl oxygen of coumarin which orientates the molecule for hydroxylation, as the relevant oxidizable carbon in coumarin lies directly above the heme iron in the model complex. The actual position of coumarin hydroxylation depends on the facility of approximation to the heme iron and the disposition of key hydrogen bond donor residues fairly close to the preferred sites of metabolism. As expected, coumarin binds to the active site of the CYP102A1 mutants with a similar pattern. All docked coumarin complexes have a similar orientation to the X-ray structure of the human microsomal CYP2A6 complexed with coumarin (PDB code: 1Z10) (Yano et al., 2005). Two threonine residues (T260 and T268) in the two CYP102A1 mutants form a hydrogen bond with the carbonyl oxygen of the coumarin ester group, which orientates the molecule for 7- and 3-hydroxylation. The relevant hydroxylation carbon atom in coumarin lies directly above the heme iron in the modeled complexes of the F162I/H236R and F162I/M185T/L188P/M237I mutants with coumarin. There also tends to be a $\pi-\pi$ stacking interaction between coumarin and F87 in the two
CYP102A1 mutants.

**Analysis of the binding complex of CYP102A1 with coumarin through MD stimulation.**

To more reliably explain why the two CYP102A1 mutants show a possibly higher catalytic efficiency against coumarin than wild-type CYP102A1, we performed molecular dynamics (MD) simulations on coumarin binding with the two mutants. A detailed analysis of the MD-simulated structures of the two CYP102A1 mutants in complex with coumarin revealed that T260 and T268 are the key residues in the structural change required for 7- and 3-hydroxylation, respectively. Coumarin binding to the active site of the CYP102A1 mutants caused considerable changes in the positions of the T260, A264, T268, and L437 residues compared to an adjacent region of the heme pocket in the wild-type CYP102A1. These changes followed hydrogen bond breakage between A364 N and T268 OG1, together with a \( \pi-\pi \) stacking interaction between the aromatic ring of coumarin and the side-chain of F87. Compared to the simulated complexes of F162I/H236R or F162I/M185T/L188P/M237I and coumarin, the average distances between the coumarin carbonyl oxygen and T260/T268 OG1s in the F162I/M185T/L188P/M237I mutant:coumarin complex are all shorter than in the F162I/H236R mutant:coumarin complex. Furthermore, coumarin binding to the F162I/M185T/L188P/M237I mutant is very similar to the binding to the F162I/H236R mutant, but the positional change of the A264/L437 residues and the spatial alteration of the non-polar binding cavity caused by the coumarin binding are significant only in the F162I/M185T/L188P/M237I mutant. Thus, these data suggest that coumarin binds more strongly to the F162I/M185T/L188P/M237I mutant than to the F162I/H236R mutant. The energy barrier for coumarin hydroxylation in the F162I/M185T/L188P/M237I mutant might be slightly lower than that in the F162I/H236R mutant. Since the mutated residues in the F162I/H236R and the F162I/M185T/L188P/M237I mutants stay outside of the active sites
and do not directly contact the coumarin molecule, these residues were initially expected to further increase or decrease the free space of the active site pocket so that coumarin binding would be easier for the CYP102A1 mutants than for wild-type CYP102A1.

The binding energy calculated via MD simulation suggests that the orientation of coumarin in binding site II for 3-hydroxylation is significantly more stable than the orientation of coumarin in binding site I for 7-hydroxylation. The stability of the coumarin orientation is due to a significant increase in the hydrophobic interaction between coumarin and the binding surface of the mutant enzyme. These results suggest that the coumarin binding structure in binding sites I and II of the F162I/M185T/L188P/M237I mutant should also be more stable than that of the F162I/H236R mutant (Supplemental Data Table S2). The average distance between the carbonyl oxygen and the T260 OG1 atom in the simulated coumarin complex, which orientates the molecule for 7-hydroxylation, was 3.57 Å and 2.52 Å for the F162I/H236R and the F162I/M185T/L188P/M237I mutants, respectively (Fig. 3 C and E). The average distances between the carbonyl oxygen, which serves to orientate the coumarin for 3-hydroxylation, and the OG1 atoms of T268 were 2.92 Å and 2.57 Å for the F162I/H236R and F162I/M185T/L188P/M237I mutants, respectively (Fig. 3 D and F).

A hydrogen bonding energy ($E_{\text{HB}}$) value can be evaluated for each hydrogen bond with the carbonyl oxygen of coumarin for each snapshot of the MD-simulated structure. The estimated total $E_{\text{HB}}$ of the MD-simulated hydrogen bonds between the carbonyl oxygen of coumarin and the heme pocket in each simulated complex structure is considered to be the average $E_{\text{HB}}$ value of all snapshots taken from the stable MD trajectory. The total hydrogen-bonding energies (-5.5 and -6.2 kcal/mol for the F162I/H236R and F162I/M185T/L188P/M237I mutants, respectively) for coumarin 3-hydroxylation estimated in this way are systematically higher (i.e., more negative) than the corresponding total.
hydrogen bonding energies (-4.6 and -4.9 kcal/mol for the F162I/H236R and F162I/M185T/L188P/M237I mutants, respectively) estimated in a way that is suitable for the catalysis of 7-hydroxylation. Moreover, the two sets of total $E_{HB}$ values are qualitatively consistent with each other in terms of the relative hydrogen-bonding strengths in the four simulated CYP102A1 complexes. In particular, the two sets of total $E_{HB}$ values consistently reveal that the overall strength of the hydrogen bonds between the carbonyl oxygen of coumarin and the heme pocket in the simulated coumarin binding complex structure of the F162I/M185T/L188P/M237I mutant is significantly stronger than that of the F162I/H236R mutant. Interestingly, hydrogen bonding between the substrate and the active site residue was suggested to be important for CYP2A6 catalyzed reactions (Wu et al., 2005). These results suggest a clear correlation between stabilization in binding site II and the catalytic efficiency of the F162I/M185T/L188P/M237I mutant for coumarin hydroxylation: the more stable the structure of binding site II, the lower the energy barrier, and the higher the catalytic efficiency.

**Analysis of the distance between the heme iron and oxidizable carbons of coumarin.**

Verifying the binding mode of coumarin that is most likely to bind the two CYP102A1 mutants tightly depends on the analysis of the distance between possible oxidizable carbons (C-3 or C-7) of coumarin and the heme iron with its ferric resting state in the two CYP102A1 mutants. The hydrogen bond interactions of the carbonyl oxygen from coumarin with the surrounding residues at the active site, as described above, are also very important for verifying which binding mode would be most likely to bind the two mutants tightly.

Since the interaction between oxidizable carbons and the Fe of heme plays an important role as coumarin binds to the F162I/H236R mutant, the distances from coumarin C-7 and C-3 to the Fe of heme were examined along all of the MD simulations (Supplemental Data Fig. S7 A and B). For the simulation in binding site I, the distance from coumarin C-7 to the Fe of
heme in the orientation of 7-hydroxylation stays below 3.7 Å and the average is 3.67 Å. For the simulation in binding site II, the distance from coumarin C-3 to the Fe of heme in the orientation of 3-hydroxylation is between 2.9 and 3.1 Å with an average of 3.01 Å. Therefore, the binding mode leading to 7-hydroxylation contributes much less to the binding of coumarin toward the F162I/H236R mutant than to that of 3-hydroxylation. The difference in the interaction energy of the coumarin complex with the F162I/H236R mutant at the two binding sites is -12.24 kcal/mol. These data indicate that the binding mode leading to 3-hydroxylation in binding site II is the most possible state when the 7-hydroxylation in binding site I binds with the F162I/H236R mutant (Supplemental Data Table S2).

The time dependence of the distances between oxidizable carbons of coumarin and the heme iron in the F162I/M185T/L188P/M237I mutant was also examined (Supplemental Data Figure S7 C and D). In binding site I, the distance from coumarin C-7 to the Fe of heme in the orientation of 7-hydroxylation is shorter than 3.0 Å and the average is 2.94 Å. In binding site II, distance from coumarin C-3 to the Fe of heme in the orientation of 3-hydroxylation is always shorter than 2.5 Å and the average is 2.45 Å. When the coumarin is docked in a reactive binding orientation leading to 3-hydroxylation, the interaction between coumarin C-3 and the Fe of heme would be much stronger than that between coumarin C-7 and the Fe of heme. As a comparison of binding interaction energy calculation in the complexes with the F162I/M185T/L188P/M237I mutant, the binding mode leading to 3-hydroxylation in the binding site II is favored at about 9.26 kcal/mol in contrast to 7-hydroxylation in binding site I (Supplemental Data Table S2).

The difference of interaction energies for the two CYP102A1 mutants at binding site I and binding site II average -15.48 and -12.50 kcal/mol, respectively (Supplemental Data Table S2). These data indicate that there is a stronger interaction between coumarin and the
F162I/M185T/L188P/M237I mutant. The interaction between oxidizable carbons of coumarin and the heme iron may contribute greatly to tight binding of coumarin toward the CYP102A1 mutants, as well as hydrogen bonding interaction with the surrounding residues at the active site.

In conclusion, we examined the possibility that CYP102A1 mutants with indole oxidation activity can have the catalytic activities of human P450 enzymes. Error-prone PCR was carried out on the heme domain coding region of the wild-type gene to generate a CYP102A1 DNA library. The library was transformed into *E. coli* to express the P450 mutants. A colorimetric colony-based method was used for primary screening of the mutants and the P450 activities were measured at the whole-cell level. Some of the blue colonies, but not the white colonies, showed apparent oxidation activity toward typical P450 substrates, including coumarin, 7-ethoxycoumarin, 7-ethoxyresorufin, phenacetin, chlorzoxazone, and *p*-nitrophenol. These results indicate that indigo formation provides a simple assay for identifying CYP102A1 mutants that have a greater potential for human P450 activity. Our computational findings suggest a correlation between the stabilization of the binding site and the catalytic efficiency of the CYP102A1 mutants towards coumarin: the more stable the structure in the binding site, the lower the energy barrier, and the higher the catalytic efficiency. Taken together, these data suggest that CYP102A1 mutants engineered by random mutagenesis can be developed as biocatalysts for industrial applications of human P450 activities.
References


Guengerich FP, Kim DH, and Iwasaaki M. (1991) Role of human cytochrome P-450 IIE1 in

DMD #30759


Kim KH, Isin EM, Yun CH, Kim DH and Guengerich FP (2006) Kinetic deuterium isotope effects for 7-alkoxycoumarin O-dealkylation reactions catalyzed by human cytochromes P450 and in liver microsomes - Rate-limiting C-H bond breaking in cytochrome P450


FOOTNOTES

This work was supported in part by the 21C Frontier Microbial Genomics and the Application Center Program of the Ministry of Education, Science and Technology of the Republic of Korea; the Korea Science and Engineering Foundation [Grant R01-2008-000-21072-02008]; and the Second Stage BK21 Project from the Ministry of Education, Science and Technology of the Republic of Korea.
Figure Legends

**Fig. 1.** HPLC analyses of the oxidation of human P450 substrates catalyzed by the CYP102A1 F162I/M185T/L188P/M237I mutant. The substrate and major product(s) are indicated. The substrates included in the reaction mixtures are as follows: (A) 7-ethoxycoumarin; (B) coumarin; (C) phenacetin; and (D) chlorzoxazone. Oxidation of the substrates were catalyzed by the F162I/M185T/L188P/M237I mutant (bold line), but not by wild-type CYP102A1 (dash line). Peaks were identified by comparing the retention times of the product with the reaction time of standard controls. Unidentified metabolites (marked with an asterisk) were also produced by the F162I/M185T/L188P/M237I mutant.

**Fig. 2.** Catalytic activity of the CYP102A1 mutants toward human P450 substrates. The following human P450 enzyme activities of the purified CYP102A1 mutant enzymes were characterized: (A) phenacetin O-deethylation; (B) 7-ethoxyresorufin O-deethylation; (C) chlorzoxazone 6-hydroxylation; (D) p-nitrophenol hydroxylation. Data are shown as the mean ± the standard error of the mean (SEM).

**Fig. 3.** Docking of coumarin to CYP102A1 mutant models and analysis of binding complex of the mutants with coumarin through molecular dynamics simulation. (A) The residues in red represent the residues mutated in the CYP102A1 F162I/H236R mutant and the residues in blue represent the residues mutated in the F162I/M185T/L188P/M237I mutant. (B) The two pockets for binding site I (for coumarin 7-hydroxylation) and binding site II (for coumarin 3-hydroxylation) are indicated by the cavity holes. The active sites of wild-type CYP102A1 (PDB code: 1BU7, gray) and its mutants (purple), F162I/H236R (C and D) and F162I/M185T/L188P/M237I (E and F), are shown. The green and salmon colors represent coumarin and heme, respectively. Although similar active site topologies have been obtained for wild-type CYP102A1 and the F162I/H236R and F162I/M185T/L188P/M237I mutants.
when the active sites are superimposed, the spatial orientation of the active sites of the two mutants is somewhat different from that of the wild-type. The average distances between the carbonyl oxygen and the T260 OG1 atom in the simulated coumarin complex, which orientates the molecule for 7-hydroxylation, were 3.57 Å and 2.52 Å for F162I/H236R (C) and F162I/M185T/L188P/M237I (E), respectively. The average distance between the carbonyl oxygen, which serves to orientate the coumarin for 3-hydroxylation, and the OG1 atoms of T268 was 2.92 Å and 2.57 Å for F162I/H236R (D) and F162I/M185T/L188P/M237I (F), respectively. For the simulation in binding site I, the average distance from coumarin C-7 to the Fe of heme in the orientation of 7-hydroxylation is 3.67 Å and 2.94 Å for F162I/H236R (C) and F162I/M185T/L188P/M237I (E), respectively. For the simulation in binding site II, the average distance from coumarin C-3 to the Fe of heme in the orientation of 3-hydroxylation is 3.01 Å and 2.45 Å for F162I/H236R (D) and F162I/M185T/L188P/M237I (F), respectively.
Table 1. Kinetic parameters of the CYP102A1 mutants for 7-hydroxylation and 3-hydroxylation of coumarin.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>7-OH coumarin</th>
<th>3-OH coumarin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k_{cat}$ (min$^{-1}$)</td>
<td>$K_m$ (mM)</td>
</tr>
<tr>
<td>F162I/M237I</td>
<td>1.2 ± 0.2</td>
<td>3.2 ± 0.8</td>
</tr>
<tr>
<td>L148I/F162I</td>
<td>2.2 ± 0.6</td>
<td>13 ± 1</td>
</tr>
<tr>
<td>F162I/K187E</td>
<td>1.2 ± 0.1</td>
<td>0.64 ± 0.07</td>
</tr>
<tr>
<td>F162I/E228K</td>
<td>1.4 ± 0.1</td>
<td>1.7 ± 0.3</td>
</tr>
<tr>
<td>F162I/H236R</td>
<td>1.9 ± 0.1</td>
<td>1.7 ± 0.2</td>
</tr>
<tr>
<td>F162I/K187E/M237I</td>
<td>2.5 ± 0.2</td>
<td>1.7 ± 0.2</td>
</tr>
<tr>
<td>F162I/M185T/L188P/M237I</td>
<td>4.0 ± 0.3</td>
<td>1.1 ± 0.2</td>
</tr>
</tbody>
</table>
FIGURE 1

A. 7-OH coumarin and 3-OH 7-ethoxycoumarin

B. 7-OH coumarin and 3-OH coumarin

C. Acetaminophen and phenacetin

D. Chlorozoxazone and 6-OH chlorozoxazone

$tr, \text{ min}$
FIGURE 2

(A) phenacetin

(B) ethoxyresorufin

(C) chlorzoxazone

(D) p-nitrophenol

nmol product/min/nmol P450

BM3 WT
F162I/M237I
L148I/F162I
F162I/K187E
F162I/E228K
F162I/H236R
F162I/M185V/M237I
F162I/K187E/M237I
F162I/F165L/M177T/M237I
F162I/M185T/L188P/M237I