Supplementary Information

Inverting the Enantioselectivity of P450pyr Monooxygenase by Directed Evolution

Weng Lin Tang, Zhi Li, and Huimin Zhao**

Experimental Procedures

Materials. *Escherichia coli* BL21(DE3) and the expression plasmids were purchased from Novagen (Madison, WI). PCR grade dNTPs were obtained from Roche Applied Sciences (Indianapolis, IN). Restriction enzymes, Phusion High-Fidelity DNA polymerase, T4 DNA ligase and their corresponding buffers were purchased from New England Biolabs (NEB) (Beverly, MA). D-Glucose was purchased from ThermoFisher (Pittsburgh, PA). Ampicillin, kanamycin, isopropyl *β*-D-thiogalactopyranoside (IPTG), nitro blue tetrazolium (NBT), phenazine methosulfate (PMS), NAD^+, and $NADP^+$ were purchased from Sigma (St. Louis, MO). Other required salts and reagents were purchased from either Fisher or Sigma-Aldrich. The Ni-NTA agarose, QIAprep spin plasmid mini-prep kit, QIAEX II gel purification kit, and QIAquick PCR purification kit were purchased from Qiagen (Valencia, CA). Various oligonucleotide primers were obtained from Integrated DNA Technologies (Coralville, IA). SDS-PAGE gels, buffers and protein size markers were purchased from Bio-Rad (Hercules, CA).

Cloning, Overexpression, and the Purification of the BRD Alcohol Dehydrogenase. The genomic DNA from *Micrococcus luteus* was extracted using the Wizard Genomic DNA purification kit from Promega (Madison, WI). The BRD gene was then PCR amplified from the genomic DNA using the following primers: BRDfwd, 5'- TTA ACT ACT CAT ATG CGA CGG ATG ACG CTG CC -3' and BRD rev, 5'- TTA ACT ACT GGA TCC TTA CAG CAT TTC CAG TGG TCG CG -3'. The amplified gene was digested with *Nde*I and *Bam*HI and ligated into the pET15b vector. The first 41 codons were optimized[1] by digesting the pET15b-BRD using *Nde*I and *Xho*I and replacing it with the *Nde*I/*Xho*I digested PCR amplified fragment (5'-tta act CAT ATG CGT CGT ATG ACT TTA CCA TCT GGT GAA TCT ATT CCA GTT TTA GGT CAA GGT ACT TGG GGT TGG GGT GAA GAT CCA GGT CGT CGT GGT GAT GAA GTT GCT GCT TTA CAT GCT GGT CTC GAG tta act-3', *Nde*I and *Xho*I restriction sites are underlined, flanking bases are in small caps). This PCR fragment was synthesized using overlap extension PCR of the following three fragments: fragfwd1 5'-TTA ACT CAT ATG CGT CGT ATG ACT TTA CCA TCT GGT GAA TCT ATT CCA GTT TTA GGT CA-3', fragfwd2 5'-AA GAT CCA GGT CGT CGT GGT GAT GAA GTT GCT GCT TTA CAT GCT GGT CTC GAG TTA ACT-3', fragrev 5'-ACC ACG ACG ACC TGG ATC TTC ACC CCA CCC CAA GTA CCT TGA CCT AAA ACT GGA ATA GA-3'. The resulting pET15b-BRD* was electroporated into electrocompetent BL21(DE3) and then plated on LB plates containing 100 μg/mL ampicillin. The BL21(DE3) pET15b-BRD* was cultivated in 5 mL LB media overnight and then inoculated into 500 mL BRD Exp media[1] (1.5% (w/v) glycerol, 1.5% (w/v) Bacto tryptone, 0.4% (w/v) Bacto yeast extract, 0.2% (w/v) sodium chloride, 0.8% (w/v) potassium dihydrogen phosphate, 0.05% (w/v) magnesium sulfate heptahydrate, pH 6.0). The culture was induced at OD_{600}=0.6-0.8 and cultured at 23°C for 18 h at 250 rpm. The cells were harvested by centrifugation and lysed by French press. The BRD protein was purified using gravity flow columns containing 3 mL of Ni-NTA resin washed with 60 mL 10 mM imidazole buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0), 60 mL 20 mM imidazole buffer, 30 mL 30 mM imidazole buffer and 3 mL 50 mM imidazole buffer. Proteins were eluted with 100 mM imidazole buffer and the protein fractions were analyzed using the Bradford reagent (Bio-Rad). Proteins were concentrated a using Millipore Amicon Ultra-4 centrifugal filter device (MWCD 10 kDa) at 4,000 rpm at 4°C, washed three times with 50 mM potassium phosphate buffer (pH 7.5), and concentrated again. Concentrated protein was stored at -80°C in 15% glycerol.

Cloning, Overexpression and the Purification of the RDR Alcohol Dehydrogenase. The RDR gene[2] from *Devosia riboflavina* was codon optimized and synthesized by Genscript Corp (Piscataway, NJ). The gene was PCR amplified further using forward primer RDRfwd NdeI, TTA ACT ACT CAT ATG TCC CAG GAT TTT TCA GGC AAG GTC and reverse primer RDRrev XhoI, TTA ACT ACT CTC GAG CTA TTG GGC GAC GTA GCC GC. The amplified gene was digested with *Nde*I and *Xho*I and ligated into the pET28a vector. The resulting construct was then electroporated into electrocompetent BL21(DE3) and plated on LB plates containing 50 μg/mL kanamycin. The BL21(DE3) pET28a-RDR was cultured in

5 mL LB media overnight and was then inoculated into 500 mL TB media. The subsequent protein expression and purification steps are the same as that for the BRD alcohol dehydrogenase.

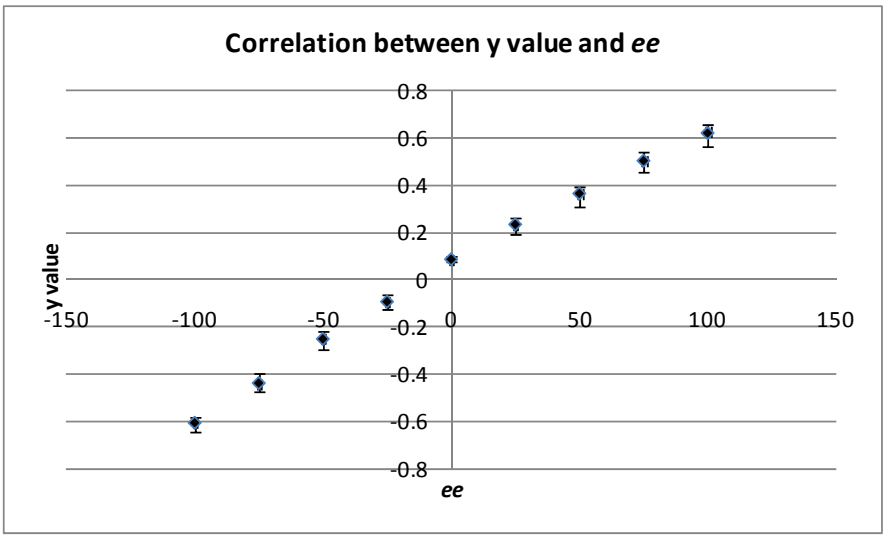

Figure S1. Graph showing the linear correlation between y value and *ee* where $y = \frac{OD_{580}^{BRD} - OD_{580}^{RDR}}{OD_{580}^{BRD} + OD_{580}^{RDR}}$. $OD_{580}{}^{BRD}$ is linearly correlated to the (*S*)-**2** concentration whereas $OD_{580}{}^{RDR}$ is linearly correlated to the (*R*)-**2** concentration.

Homology Modeling.[3] The following structures were downloaded from the Protein Data Bank (PDB) database (PDB accession code): CYP119 from *Sulfolobus solfactaricus* (1F4T)[4] and P450st from *Sulfolobus tokodaii* (1UE8).[5] These two structures were aligned using Insight II software (Insight II, version 2000; Accelrys Inc., San Diego, CA) based on its conserved structural regions to achieve the lowest root-mean-square (RMS) score. The amino acid sequence of P450pyr was manually aligned with the structural alignment such that the aligned sequences represented homologous structural regions. Using default parameters of the automated MODELER module within Insight II, with the alignment as input and using moderate refinement of the structure and loop regions, we obtained approximately 30 structural models. By visually inspecting the models for obvious flaws, comparing the scores from the Profiles 3-D function, and the ProStat inspection of ϕ and ψ angles, the best model was selected. In this model, the ϕ and ψ angles were determined to be 82.4% within their expected values, which was comparable to 86.8% and 83.3% for identical analysis of the template PDB structures 1F4T and 1UE8, respectively. A value of 79.9% correct self-compatibility of amino acids with the modeled structure was obtained when inspected by Profiles-3D (Insight II, default parameters). Therefore, the model seemed reasonably well constructed despite the low homology with the templates. Docking of substrate **1** onto the heme binding site of the best model was carried out using the Molecular Operating Environment (MOE; Chemical Computing Group Inc., Montreal, Canada). The whole structure was then subjected to energy minimization to relieve steric and torsional artifacts from the modeling and docking processes.

Cloning of an Active Recombinant *E. coli* BL21(DE3). Plasmid pCom8-PA7F200R1500[6] was used as a PCR template to amplify the coding region of cytochrome P450pyr, ferredoxin (Fdx) and ferredoxin reductase(FdR). The PCR amplified P450pyr gene was digested and ligated into the pRSFDuet vector. Similarly, the PCR amplified ferredoxin and ferredoxin reductase gene were digested and cloned into the MCS 1 and MCS 2 of the pETDuet vector respectively. The resulting plasmids from the ligations were separately transformed into chemical competent *E. coli* DH5α cells and plated on LB agar plates containing the necessary antibiotics. Finally, the purified pRSFDuet P450pyr and pETDuet Fdx plasmids were transformed into electrocompetent *E. coli* BL21(DE3) and plated on LB agar plates containing 50 µg/ml kanamycin and 100 µg/ml ampicillin.

Library Generation. Megaprimer PCR method was used to generate libraries where single residues were randomized to all 20 possible amino acids. In the first round of saturation mutagenesis, pRSFDuet P450pyr wild type plasmid was used as template. Briefly, three primers were used to generate an amplified gene library consisting of a saturation mutagenized residue: two primers flanking the P450pyr region (PF200fwd, 5'-TTA ACT ACT CCA TGG AAC ATA CAG GAC AAA GCG CGG-3', and P450rev KpnI, 5'-TTA ACT ACT GGT ACC CTA CGC GTG GAC GCG AAC-3') and a degenerate primer incorporating the residue to be mutated. The degenerate primer substituted the codon corresponding to the target residue with the sequence NNS and contained 9-10 additional bases on either side (5' and 3'). The NNS substitution allowed the incorporation of all 20 amino acids while limiting the codon possibilities to only 32. For each gene library with a randomized codon, 2 PCRs were performed. Firstly, the megaprimer containing the NNS substitution at the codon of interest was amplified. Each PCR contained (50 mL of final volume): 1× Phusion HF buffer,0.2 mM of each dNTP (Roche), 0.5 μM of appropriate degenerate primer, 0.5 μM of appropriate flanking primer (PF200fwd or P450rev KpnI), and 50 ng of template plasmid. PCRs were carried out on a MJ Research (Watertown, MA) PTC-200 thermocycler with cycle conditions of 98°C 30 s, (98°C 10 s, 58°C 30 s, and 72°C 15 s) × 25 cycles and 72°C 10 min. The megaprimer product was purified and eluted into 50 μL EB buffer using the QIAquick PCR purification kit. 10 uL of the megaprimer was then combined into a 100 μL PCR reaction which shares the same conditions to the first PCR step described above, except the primers used were replaced with the megaprimer and its corresponding flanking primer, and a different program was used: 98°C 30 s, (98°C 10 s, 58°C 30 s, and 72°C 45 s) × 30 cycles and 72°C 10 min. PCR product from this reaction was purified from a 1% agarose gel by using the QIAEX II gel purification kit. The insert was digested with *Nco*I and *Kpn*I and ligated into the pRSFDuet vector. This library was electroporated into *E. coli* BL21(DE3) competent cells containing the pETDuet Fdx FdR vector.

Library Screening. For identification of enantioselective P450pyr variants, a 96-well plate two-enzyme based colorimetric assay was developed based on a colorimetric nitro blue tetrazolium (NBT) - phenazine methosulfate (PMS)[3]. Library colonies were picked with sterile toothpicks and used to inoculate 500 μL of TB medium containing 50 μg/mL kanamycin and 100 μg/mL ampicillin in 96-deep well (2 mL) plates. The plates were shaken at 500 rpm, 37°C for 6 h, and protein expression was induced by adding 500 μL of TB medium containing 0.5 mM IPTG and 0.5 mM δ-ALA, followed by shaking at 500 rpm, 30°C for 14 h. The plates were centrifuged at 4000 rpm for 10 min at 4°C, and the cells were resuspended in 300 μL of 50 mM potassium phosphate buffer pH 7.0 containing 2 % (w/v) D-glucose and 10 mM of **1**. The plates were then shaken at 1100 rpm, 30°C for 5 h and were later centrifuged at 4000 rpm for 10 min at 4°C. An aliquot (80 μL) of the supernatant was transferred to two fresh plates, one containing ~5 μg BRD enzyme while the other plate contained ~5 μg of the RDR enzyme. The NBT assay solution (final concentration 2 mg/mL NBT, 0.1 mg/mL PMS, 0.5 mM NAD(P)H) was added to each plate to initiate the reaction. The plates were incubated at room temperature in darkness for 2.5 h and the absorbance at 580 nm was measured in a SpectraMax 340PC microplate reader (Molecular Devices, Sunnyvale, CA). Enantioselective mutants were identified by comparing the ratio of the absorbance on both plates against that of the parent enzyme.

Whole Cell Assay in Shaking Flasks. To confirm the positive mutants, they were streaked on LB plates containing 50 μg/mL kanamycin and 100 μg/mL ampicillin, and single colonies were picked into 5 mL of liquid LB medium containing the antibiotics. After shaking at 37°C overnight, 1 ml of the culture were inoculated into 10 mL of liquid TB medium containing the antibiotics, 0.5 mM IPTG and 0.5 mM δ-ALA. The culture was shaken at 250 rpm, 30°C for 6 h and then centrifuged at 4000 rpm for 5 min at 4°C. The cells were resuspended (to 10 g cell dry weight per liter) in 50 mM potassium phosphate buffer pH 7.0 containing 2 % (w/v) D-glucose and 10 mM of **1**, and further shaken at 250 rpm, 30°C for 4 h. After centrifuging at 4000 rpm, 80 μL aliquot of the supernatant was taken to be analyzed by using the Hypersil BDS-C18 (5 μm) column (15 mm × 4 mm) on the reverse-phase HPLC.[7] The chiral product in the remaining supernatant was extracted using ethyl acetate in a standard work up procedure and the chirality was analyzed by using the Chiralcel OB-H (250 mm × 4.6mm) column.[8] The *ee* of the positive mutants were verified in triplicates.

DNA Sequencing and Analysis. Plasmid DNA from selected mutants was isolated using QIAprep spin plasmid mini-prep kits. Sequencing reactions consisted of 100 to 200 ng of template DNA, 10 pmol of

each primer, sequencing buffer, and the BigDye reagent (Applied Biosystems, Foster City, CA). Reactions were carried out for 31cycles of 96°C for 20 s, 50°C for 10 s, and 60°C for 3.5 min in an MJ Research PTC-200 thermal cycler (Watertown, MA). Prepared samples were submitted to the Biotechnology Center at the University of Illinois for sequencing on an ABI PRISM 3700 sequencer (Applied Biosystems, Foster City, CA).

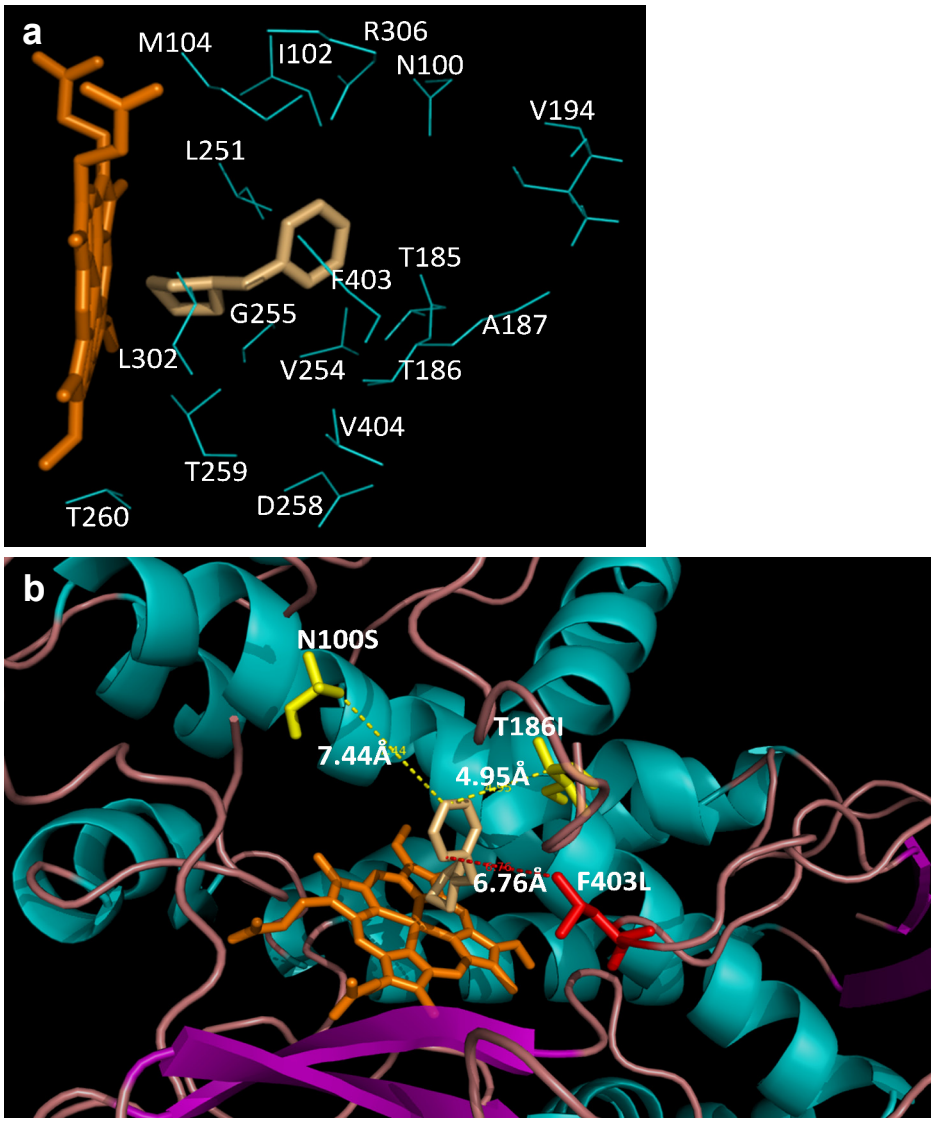

Figure S2. (a) Homology model showing the 17 residues that were identified within 5Å of the heme (orange)-docked substrate (light orange). (b) The mutation sites are shown and labeled (yellow for the (*R*)-mutant and red for the (*S*)-mutant), and the distances between the substrate **1** and the mutation site are shown in Å.

1. N. Kizaki, Y. Yasohara and J. Hasegawa, *US Patent*, 7,033,808 B2, 2006.
2. N. Y. Kizaki, Y.; Nagashima, N.; Hasegawa, J., *J. Mol. Cat. B-Enzymatic*, 2008, **51**, 73-80.
3. R. Woodyer, W. A. van der Donk and H. Zhao, *Biochemistry*, 2003, **42**, 11604-11614.
4. J. K. Yano, L. S. Koo, D. J. Schuller, H. Li, P. R. Ortiz de Montellano and T. L. Poulos, *J. Biol. Chem.*, 2000, **275**, 31086-31092.

5. Y. Oku, A. Ohtaki, S. Kamitori, N. Nakamura, M. Yohda, H. Ohno and Y. Kawarabayasi, *J. Inorg. Biochem.*, 2004, **98**, 1194-1199.
6. J. B. van Beilen, E. G. Funhoff, A. van Loon, A. Just, L. Kaysser, M. Bouza, R. Holtackers, M. Rothlisberger, Z. Li and B. Witholt, *Appl. Environ. Microbiol.*, 2006, **72**, 59-65.
7. Z. Li, H.-J. Feiten, D. Chang, W. A. Duetz, J. B. van Beilen and B. Witholt, *J. Org. Chem.*, 2001, **66**, 8424-8430.
8. Z. Li, H.-J. Feiten, J. B. van Beilen, W. Duetz and B. Witholt, *Tetrahedron: Asymmetry*, 1999, **10**, 1323-1333.