

Supporting Information

**Evolving P450_{pyr} Hydroxylase for Highly Enantioselective Hydroxylation at
Non-Activated Carbon Atom**

Son Quang Pham,^a Guillaume Pompidor,^b Ji Liu,^a Xiao-Dan Li^c and Zhi Li^{*a}

^aDepartment of Chemical and Biomolecular Engineering, National University of Singapore, 4 Engineering Drive 4,
Singapore 117576

^bSwiss Light Source, Paul Scherrer Institut, CH-5232 Villigen PSI, Switzerland

^cBiomolecular Research, Department of Biology and Chemistry, Paul Scherrer Institut, CH-5232 Villigen PSI,
Switzerland

E-mail: chelz@nus.edu.sg

List of content

1. Strains, growth media, and materials	3
2. Engineering of <i>E. coli</i> Top10-pTrc99A (his-tagged P450) recombinant expressing his-tagged P450pyr hydroxylase	4
3. Engineering of <i>E. coli</i> recombinant expressing Fdx200 and FdR1500	11
4. Purification of his-tagged P450pyr hydroxylase.....	12
5. Crystallization of his-tagged P450pyr hydroxylase.....	16
6. Directed evolution: mutant library creation	19
7. Directed evolution: enantioselectivity determination by MS assay and chiral HPLC analysis....	21
8. Reverse HPLC chromatogram	24
9. Chiral HPLC chromatogram	27
10. Substrate docking on protein structures by molecular dynamics (MD) simulation.....	28
11. Kinetic study of wild type P450pyr and P450 mutant I83H/M305Q/A77S	31
12. References.....	39

1 Strains, media, and materials

The strains and plasmids used or constructed for cloning and expression are listed in Table S1. Luria broth (LB), supplemented with antibiotics, were used throughout. All cultures were grown aerobically at 30 °C or 37 °C. Chemical or electro competent *E. coli* cell were use for plasmid transformation. To select *E. coli* transformants, ampicillin was used at 100 µg/ml, kanamycin was used at 50 µg/ml.

Table S1. Strains and plasmids used in this study

Strain or plasmid	Characteristic	Reference
<i>E. coli</i> Top10	Cloning strain	Invitrogen
<i>E. coli</i> BL21(DE3)	Cloning strain	Novagen
pTr99A	Cloning vector	Our group
pET21a(+)	Expressing vector	Novagen
pETDuet1	Expressing vector	Novagen
pRSFDuet1	Expressing vector	Novagen
pET28a	Expressing vector	Novagen
pCom8-PA7F200R1500	DNA Template	[1]
pRSFDuet1-P450	DNA Template	[2]
pETDuet1-FdxFdR	DNA Template	[2]

All chemical reagents and solvents were purchased from Acros, Sigma-Aldrich, Fluka, Fisher or Merk and used without further purification. *N*-benzylpyrrolidine **7** (97%, ACROS), (*R*)-1-benzylpyrrolidine-3-ol **8** (97% *ee*, Sigma-Aldrich), (*S*)-1-benzylpyrrolidine-3-ol **8** (99% *ee*, Sigma-Aldrich), POBr₃ (98.5%, Fluka), THF (99.9%, Sigma-Aldrich), Pentane (99.9%,

Sigma-Aldrich), CBr₄ (99%, Sigma-Aldrich), PPh₃ (99%, Sigma-Aldrich), CH₂Cl₂ (Fisher Scientific, AR grade), Pentane (99.9%, Sigma-Aldrich), LiAlD₄ (98% D, Sigma-Aldrich). Restriction enzymes, T4 DNA quick ligase kit, Taq DNA polymerase and dideoxynucleotides were obtained from New England Biolab. Miniprep kit, gel extraction kit and PCR purification kit were purchased from Qiagen. Oligonucleotides were synthesized by AIT biotech, Singapore.

Column chromatography was performed with Merck silica gel 60 (200–300 mesh) as column material. TLC analyses were carried out with pre-coated Silica Gel 60 F254 TLC-plate or Aluminum oxide and the compounds were visualized with UV light. FPLC system is AKTApurifier system. Ni-NTA column is HisPrep FF 16/10 column. SEC column is Superdex 200 10/300 GL. All were purchased from GE Healthcare. PCRs were carried out using a Mycycler Biorad.

2 Engineering of *E. coli* Top10-pTrc99A (his-tagged P450) recombinant expressing his-tagged P450pyr hydroxylase

2.1 PCR and gene cloning

P450pyr gene was amplified from pCom8-PA7F200R1500.^[1] To clone P450pyr hydroxylase with his-tag, special primers were designed as follow:

- Design primers (his-tag at C terminal)

- PF (Forward primer)

GGATTCCCATG.GAA.CAT.ACA.GGA.CAA.AGC.G

[NcoI site 3 bases upstream of the ATG start codon]

- PR1 (Reverse primer)

GGATTCCTGCAGCTA.ATG.ATG.ATG.ATG.ATG.ATG.TCC.TCC.TCC.CGC.GTG.GAC.G
CG.AAC.C

[PstI site downstream of the stop codon]

- Design primers (his-tag at N terminal)

- PF2 (Forward primer)

GGATTC~~CC~~ATG.GTT.CAT.CAT.CAT.CAT.CAT.CAT.AGC.AGC.GGC.GAA.CAT.ACA.G
GA.CAA.AGC.G

[NcoI site 3 bases upstream of the ATG start codon]

- PR (Reverse primer)

GGATTCCTGCAGCTA.CGC.GTG.GAC.GCG.AAC.C

[PstI site downstream of the stop codon]

- Design primer for cloning P450 with his-tag at N-terminal and co-express with Ferredoxin (Fdx)

- PF2 (Forward primer)

GGATTC~~CC~~ATG.GTT.CAT.CAT.CAT.CAT.CAT.CAT.AGC.AGC.GGC.GAA.CAT.ACA.G
GA.CAA.AGC.G

[NcoI site 3 bases upstream of the ATG start codon]

- PR3 (Reverse primer)

GGATTCCTGCAGTCA.ATG.CTG.CGC.GAG.AGG

[PstI site downstream of the stop codon]



Figure S1. Plasmid map of pTrc99A-hisP450pyr, pTrc99A-P450pyrhis and pTrc99A-hisP450pyr

From this cloning strategy, the clone with his-tag at N-terminal will have 10 more amino acids at N-terminal and the clone with his-tag at C-terminal will include 9 more amino acids at C-terminal.

For PCR amplification of P450_{pyr} gene with his-tag, the following program was used: 3 min at 94 °C; 25 cycles of 1 min at 94 °C, 1 min at 55 °C, and 1 min at 72 °C; 7 min at 72 °C; and 4 °C until further use.

PCR products were purified over 1% agarose gel in Tris-acetate-EDTA buffer. The plasmid pTr99A was chosen to clone P450_{pyr} since this plasmid previously succeeded to express active P450_{pyr}. pTr99A and PCR product were digested with 2 restrictive enzymes which restriction sites have been introduced in the PCR step. The digestion products were purified again over 1% agarose gel in Tris-acetate-EDTA buffer. The ligation was conducted using T4 quick ligase kit. The ligation solution was then transformed into *E. coli* Top10 competent cell.

2.2 Cell growth and expression of his-tagged P450_{pyr} hydroxylase

E. coli Top10-pTrc99A (his-tagged P450_{pyr}) was grown in LB medium containing 100 µg/ml ampicillin at 37 °C. When OD₆₀₀ reached 0.5–0.7, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM to induce the expression of his-tagged P450_{pyr}. During incubation for

further 5-10 hours at 25 °C, the cells were taken at selected time point from the culture and harvested by centrifugation (6000g for 10 min). The cells were then diluted to OD600 equal to 10, heated for 5 min at 95 °C, centrifuged at 13000 g for 10 min to get cell free extract (CFE). The CFE was mixed with SDS sample buffer at volume ratio 1: 1, heated for 5 min at 95 °C and run SDS-PAGE.

The gels were shown clearly the expression of a protein in between 47.5 to 62 kDa after induction.

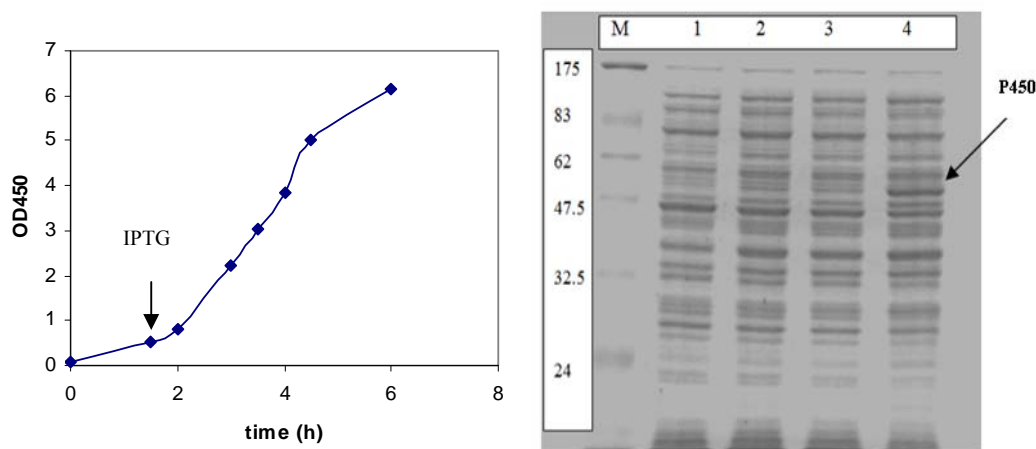


Figure S2. (left) Growth curve of Top10-pTrc99A-P450_{pyr} with his-tag at C terminal. (right) SDS-PAGE of CFEs of cells taken at different time points; lane 1: before induce, lane 2: after induce 1 h, lane 3: after induce 2 h, lane 4: after induce 4 h.

The size of the expressed protein is relative close to the real size of P450_{pyr} (47 kDa). This small difference was caused by the condition of preparing and running SDS-PAGE.

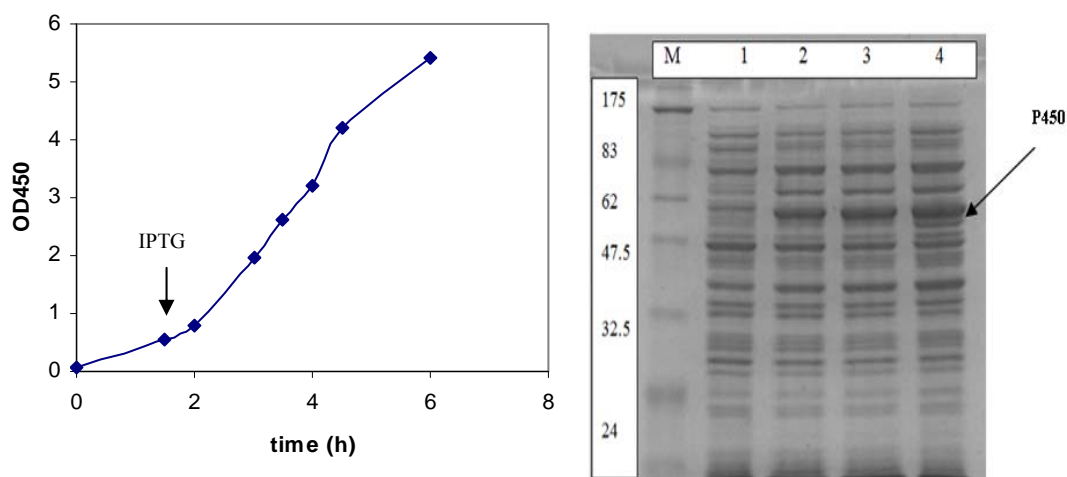


Figure S3. (left) Growth curve of Top10-pTrc99A-P450_{pyr} with his-tag at N-terminal. (right) SDS-PAGE of CFEs of cells taken at different time points; lane 1: before induce, lane 2: after induce 1 h, lane 3: after induce 2 h, lane 4: after induce 4 h.

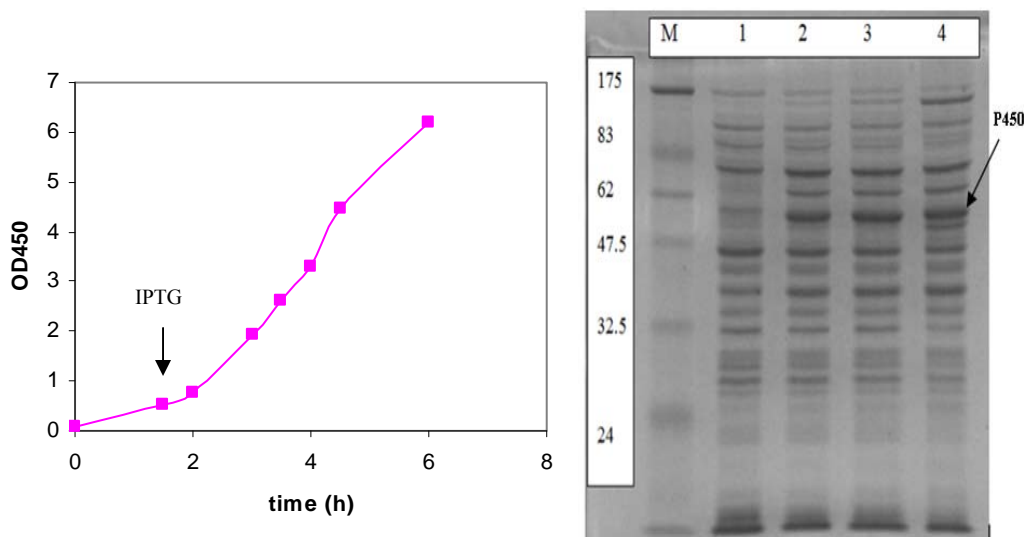
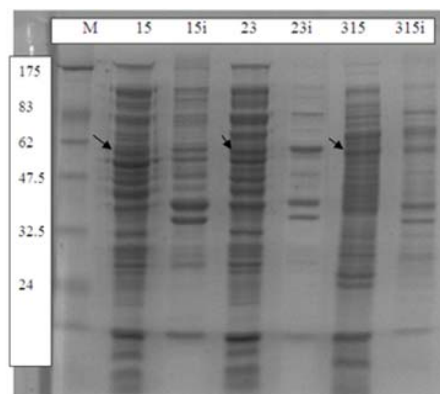


Figure S4. Growth curve and SDS pages time course. (left) Growth curve of Top10-pTrc99A-P450_{pyr}Fdx₂₀₀ with his-tag at N-terminal. (right) SDS-PAGE of CFEs of cells taken at different time points; lane 1: before induce, lane 2: after induce 1 h, lane 3: after induce 2 h, lane 4: after induce 4 h.

To check the solubility of the expression protein, the cells were broken by cell disruptor at 21 psi, 2 times and centrifuged to get the supernatant (soluble) and the pellet (insoluble). Both were used for SDS-PAGE.



15: supernatant of Top10-pTrc99A-P450 with his-tag at C terminal
23: supernatant of Top10-pTrc99A-P450 with his-tag at N terminal
315: supernatant of Top10-pTrc99A-P450 Fex200 with his-tag at N terminal
15i: insoluble part of Top10-pTrc99A-P450 with his-tag at C terminal
23i: insoluble part of Top10-pTrc99A-P450 with his-tag at N terminal
315i: insoluble part of Top10-pTrc99A-P450 Fex200 with his-tag at N terminal

Figure S5. SDS-PAGE of CFE and insoluble part of *E. coli* Top10 expression his-tagged P450pyr hydroxylase.

The gel showed that the expression band only appear in the soluble part of the total protein but not in the insoluble part. This mean the expressed protein is soluble.

2.3 CO difference spectra of cell free extract of *E. coli* Top10 expression his-tagged P450pyr

A special characteristic of P450 monooxygenase is that the reduce form can combine with carbon mono oxygen gas to give the maximum absorption at 450 nm. Cells were broken by Cell disruptor machine at 21psi, 2 times. The solution was then centrifuge at 23000 g for 50 min. The supernatant (15 – 20 g protein/l) was taken to measure CO difference spectral by mixing with few milligrams of reduce agent Dithiothreitol (DTT) then bubbling CO gas for 1 min and measured the absorption from 350 nm – 550 nm. All the CFEs showed CO difference at 450 nm. This confirmed the expression of P450 enzyme.

The active protein concentration was calculated by

$$c(\text{mM}) = \frac{A}{e(\text{cm}^{-1}\text{mM}^{-1}) \cdot b(\text{cm})} = \frac{A}{91(\text{cm}^{-1}\text{mM}^{-1}) \cdot 1(\text{cm})} \quad [3]$$

Base on molar extinction difference of $91 \text{ cm}^{-1}\text{mM}^{-1}$, the amount of active P450_{pyr} in the cell free extract was estimated and given in table S5.

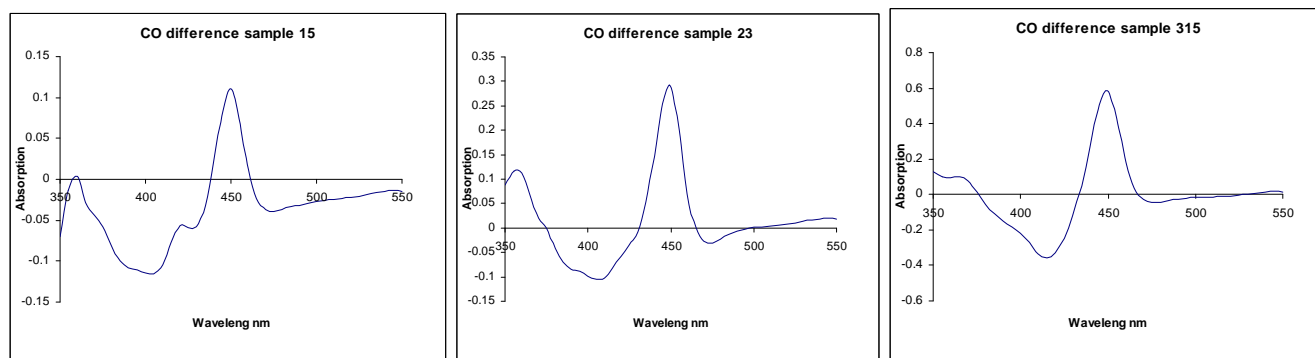


Figure S6. CO difference spectra

Table S2. Active P450 concentration calculated base on CO difference spectra

	23	315	15
Total protein (g/L)	16.1	18.6	20.1
CO difference (ΔA)	0.29	0.59	0.11
P450 (mM)	0.003	0.006	0.001
Active P450 (g/L)	0.15	0.3	0.06
Active P450 base on total soluble protein (%)	0.9	1.6	0.3

15:CFE of Top10-pTrc99A-P450 with his-tag at C terminal; 23:CFE of Top10-pTrc99A-P450 with his-tag at N terminal; 315:CFE of Top10-pTrc99A-P450 Fex200 with his-tag at N terminal

3 Engineering of *E. coli* recombinant expressing Fdx₂₀₀ and FdR₁₅₀₀

In order to test the activity of purified his-tag P450_{pyr}, an *E. coli* recombinant expressing Fdx₂₀₀ and FdR₁₅₀₀ was engineered. From the sequence of plasmid pCom8-PA7F200R1500, primers were designed to amplify DNA fragment containing genes of Ferredoxin and Ferredoxin reductase.

- Forward primer: HXN200-Fer-FW

GGATTCCCATATG.CCA.ACA.GTG.ACC.TAT.GTT.G

[NdeI site 3 bases upstream of the ATG start codon]

- Reverse primer: 1500-Red-RV1

GGATTCAAGCTT.TTA.GAG.GGA.GGT.TGG.G

[HindIII site downstream of the stop codon]

The restrictive enzymes NdeI and HindIII were used to digest the DNA fragment and plasmid pET21a. The ligation was conducted by T4 ligate. The cloning result was checked by same restrictive enzymes digestion and confirmed by DNA sequencing.

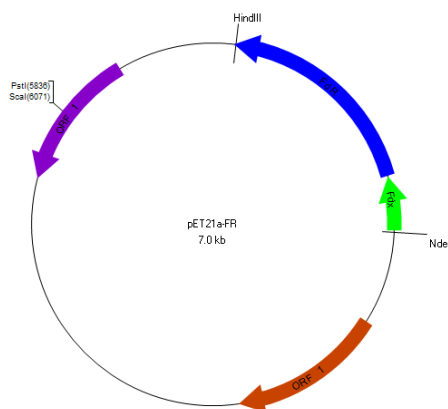


Figure S7. plasmid map of pET21a-FR

The recombinant *E. coli* BL21(DE3)-pET21a-FdxFdR was cultured in LB medium. When OD600 reached about 0.6, the cells were induced by adding IPTG to 0.5 mM. The cells were further grown for 9 h at 30 °C. Some cells were taken at different time points, and their CFE were used for SDS-PAGE. From SDS-PAGE a band in between 50 to 64 kDa was observed which is the similar size of FdR₁₅₀₀. The expression of Fdx₂₀₀ was not clear evidenced since the size of Fdx₂₀₀ is quite small and may overlap with other protein bands.

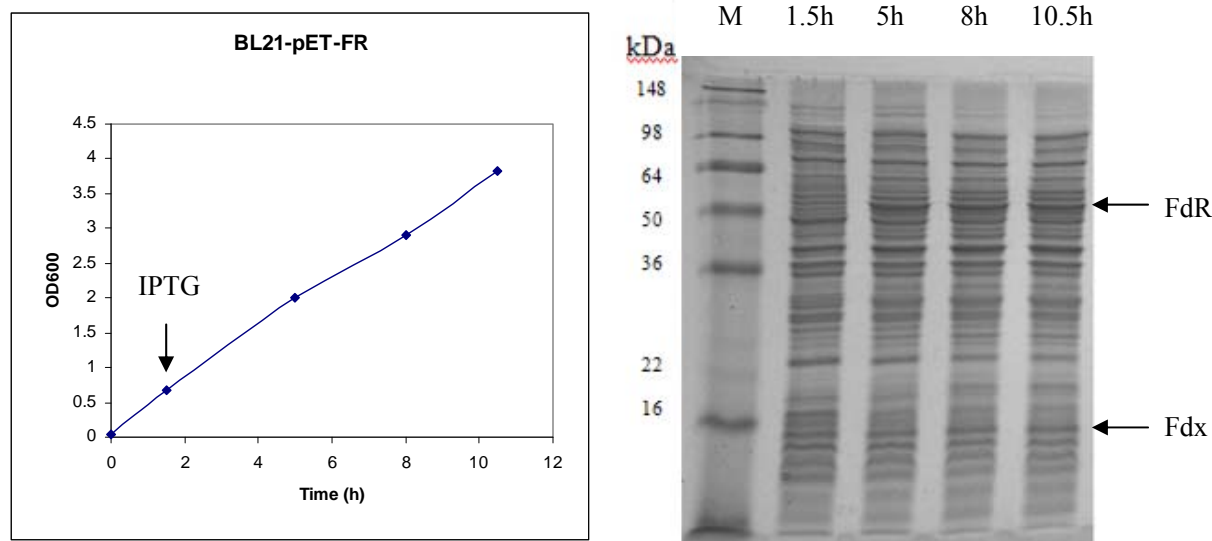


Figure S8. (left) Growth curve of *E. coli* BL21(DE3) pET21a-Fdx₂₀₀FdR₁₅₀₀. (right) SDS-PAGE of CFEs of cells taken at different time points; (M) protein marker

4. Purification of his-tagged P450_{pyr} hydroxylase

E. coli Top10-pTrc99A-hisP450_{pyr}Fdx₂₀₀ was grown and expressed in LB medium as described above. Cells were harvested at OD600 = 3 – 4, and cells pellet were collected after centrifugation and resuspended in buffer (50 mM NaH₂PO₄, 0.3 M NaCl, 10 mM imidazole, pH 8). Cells suspension were passed through cell disruptor at 21 psi, 2 times following by centrifugation (23000 g, 4 °C, and 50 min).

The purification was performed in FPLC with 2 chromatography columns. To monitor the purification progress, the solution was measured at 280 nm for total protein and 418 nm for P450 online. The temperature was kept at 4 °C for whole process. The cell free extract (50 ml) was loaded into pre-packed Ni-NTA column, prebalanced with 10mM imidazole buffer (100 ml) to eliminate non specific binding. The protein was washed by buffer with 15mM (100 ml) and then 30 mM (100 ml) imidazole. The target protein was finally eluted with buffer containing 250 mM imidazole (50 ml). The elution fraction containing his-tagged P450_{pyr} hydroxylase was concentrated by 30 kDa centrifuge filter. The concentrated protein was loaded in size exclusion chromatography column, eluted with 20 mM tris buffer (pH 7.5) containing 20 mM NaCl.

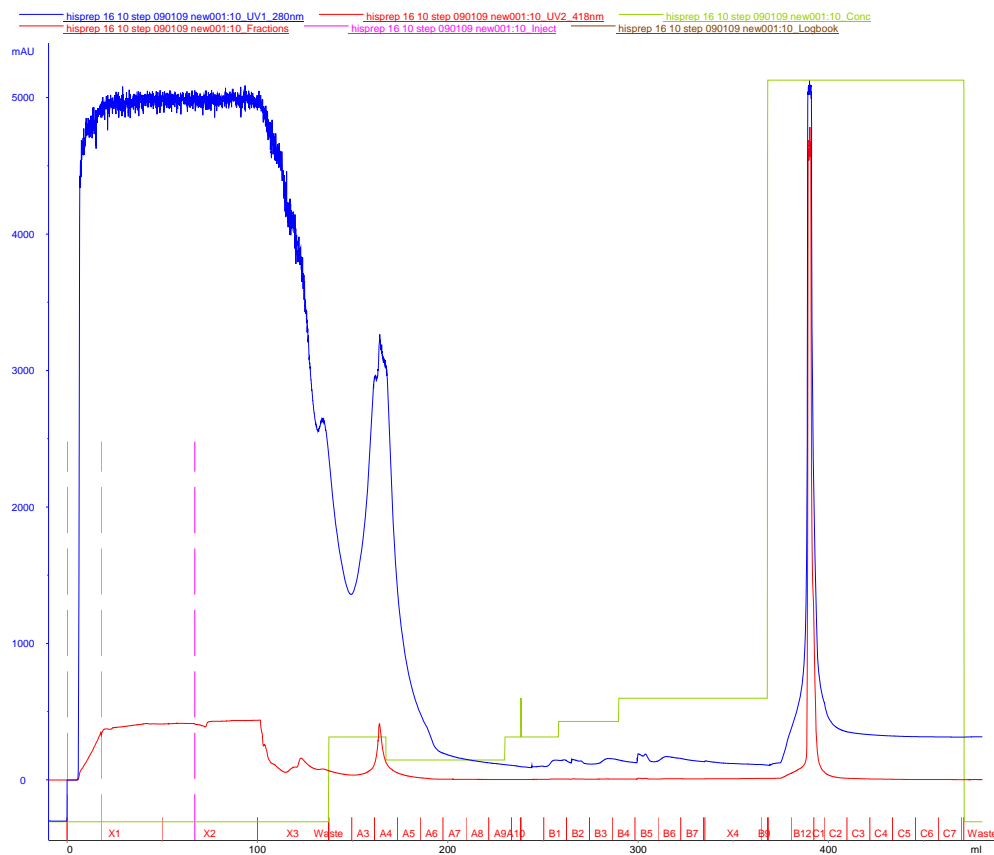


Figure S9. FPLC chromatogram of the purification of his-tagged P450_{pyr} on a Ni-NTA column

Blue line: UV absorption at 280 nm; Red line: UV absorption at 418 nm

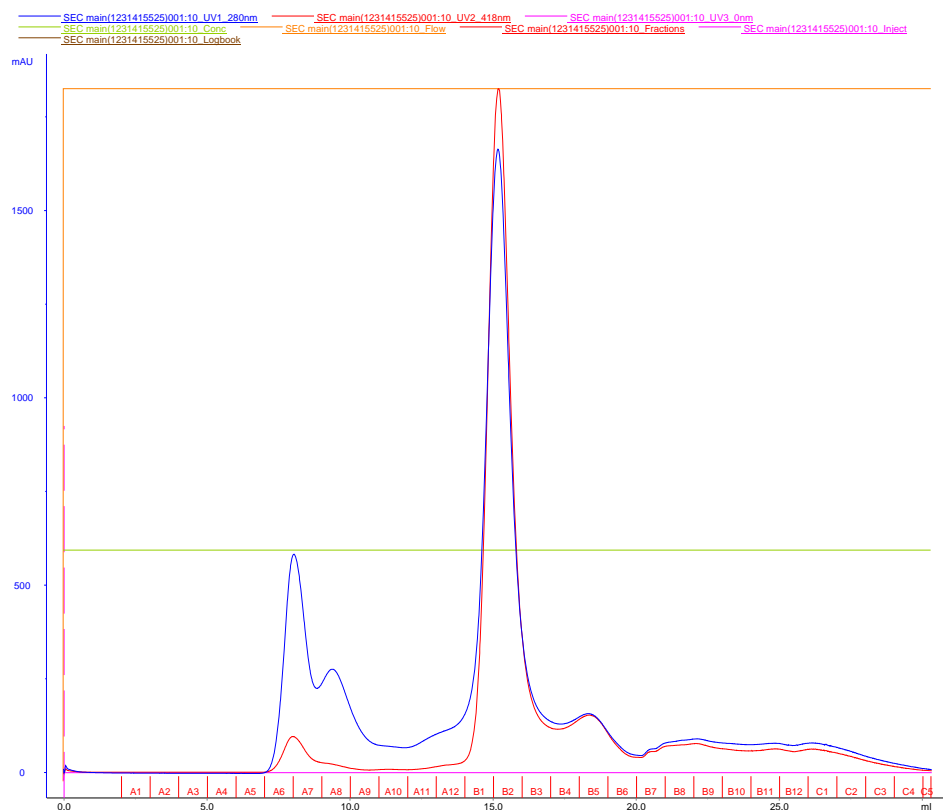


Figure S10. FPLC chromatogram of P450_{pyr} purification using SEC column

Blue line: absorption at 280 nm; Red line absorption at 418 nm

The activity of the his-tagged P450_{pyr} was determined by combining with CFE of *E. coli* BL21(DE3)-pET21a-FdxFdR, volume ratio 1: 1, to perform hydroxylation with *N*-benzyloxycarboxyl pyrrolidine with addition of NADH. The product formation was analyzed by reverse HPLC [Mobile Phase: 45% Acetonitrile - 55% 10 mM Phosphate buffer (pH7); 4 μ L injection Volume; Flow rate: 1 ml/min; Column: XBD C18]. Protein concentration was measured by Bradford method.

Table S3. Purification of his-tagged P450_{pyr} hydroxylase by FPLC

	Protein conc. [mg/ml]	Volume [ml]	Total protein [mg]	Specific		Yield [%]
				Total	Total	
				activity [U/g protein]	activity [mU]	
Cell free extract	17.1	45	769.5	1.5 ^a	1154	100
Elute fraction after Ni-NTA	2.8	12	33.6	15.6 ^a	524	45
Elute fraction after SEC	0.7	10	7	45.4 ^a	317	27
				[1595] ^b		

[a] The activity was calculated based on the biotransformation of *N*-benzyl oxycarboxyl pyrrolidine (2mM) with his-tagged P450_{pyr} hydroxylase and an equal volume of crude cell extract [14.9 g/l] of *E.coli* BL21(DE3) pET21a-FR expressing Ferredoxin and Ferredoxin reductase. [b] The activity was calculated based on the biotransformation of *N*-benzyl oxycarboxyl pyrrolidine (2 mM) with his-tagged purified P450_{pyr} hydroxylase [0.047 g/l], his-tagged purified Ferredoxin and his-tagged purified Ferredoxin reductase at a molar ratio of 1: 100: 1.35.

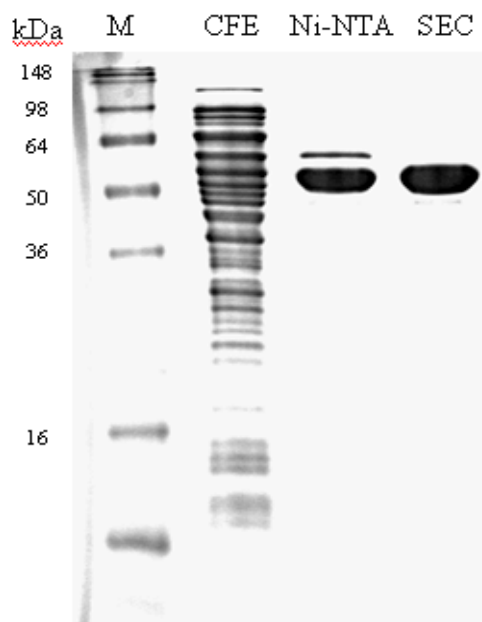


Figure S11. SDS-PAGE of purified his-tagged P450pyr hydroxylase after Ni-NTA and SEC column

5 Crystallization of his-tagged P450pyr hydroxylase

The purified his-tagged P450pyr hydroxylase was concentrated to 50 mg/mL for crystallization experiments. After initial screening, the best crystal was obtained from 1:1 mixtures of the protein solution and the mother liquor solution containing 100 mM Tris-HCl (pH 7.5), 20% PEG 4000, and 10% isopropanol after 5 weeks at 20 °C. One crystal was cryo-protected by soaking in a solution identical to the mother liquor supplemented with 25% glycerol and immediately frozen in the nitrogen stream at 100 K. Data from the flash-cooled crystal was collected at the beamline X10SA at Swiss Light Source. In order to maximize the anomalous signal of the iron ion of the heme for ab initio phasing, diffraction data were collected at the Fe K absorption edge ($\lambda = 1.73610 \text{ \AA}$) over a wedge of 180° and with an oscillation range of 0.25°. Friedel pairs were kept unmerged for phasing, but merged before refinement. The iron site was found using SHELX D,^[4] experimental phases were determined using SHELX E,^[4] and automatic building was carried out with ARP/wARP.^[5] The structure was

refined using Refmac^[6] and shown in Figure S12. The nearest structural homolog, found by structural alignment using DALI^[7], is the class I cytochrome P450-terp from *Pseudomonas* sp.^[8] (29 % homology, pdb ID 1CPT).

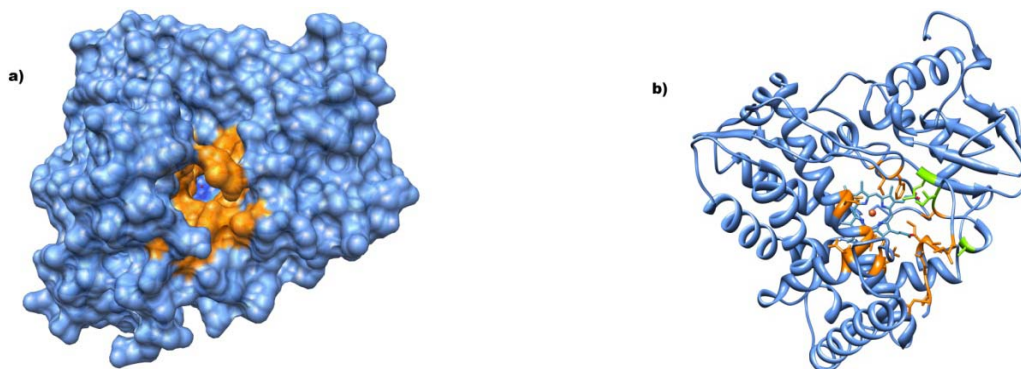


Figure S12. Structure of his-tagged P450pyr hydroxylase. a). The surface of the protein is depicted in blue. The surface corresponding to the selected mutations is shown in orange. The water molecule bound to the iron ion of the heme is depicted in red. b). The 3 residues corresponding to the triple mutations leading to the highest enantioselectivity are depicted in green. The other targeted residues are depicted in orange. Figures were prepared using Chimera.^[9]

Data-collection parameters, processing and refinement statistics are given in Table S4, and the X-ray structure was deposited as pdb ID 3RWL.

Table S4. Diffraction data collection parameters, processing statistics from *SCALA* from the CCP4 suite^[10] and refinement statistics. Values in parentheses refer to the highest resolution shell.

Space group	P3 ₁
Resolution range (Å)	50.0 - 2.00 (2.11- 2.00)
Unit cell parameters (Å)	a = b =112.60, c = 41.96 $\alpha = \beta =90^\circ, \gamma = 120^\circ$
N. of reflections	197203 (23980)
N. of unique reflections	38307 (5046)
R _{merge} (%)	7.1 (52.2)
I/ σ (I)	14.6 (2.9)
Completeness (%)	95.3 (85.8)
Multiplicity	5.1 (4.8)
R _{work} (%)	16.3
R _{free} (%)	19.4
Standard deviations on bond lengths (Å)	0.026
Standard deviations on bond angles (°)	2.07
Ramachadran plot: residues in (%):	
Favoured region	96.7
Allowed region	3.3
Disallowed region	0

6 Directed evolution: mutant library creation

The mutant library at a selected site was created using site saturation mutagenesis. Degenerated primers were designed at the specific amino acid residue with the sequence of NNK and 9 – 12 bases on the left and right. The NNK can encode for all 20 amino acid. PCR method was used to create the mutant plasmid base on template pRSFDuet1-P450pyr. The template was then eliminated using methylated restrictive enzyme DpnI. The mutated plasmid was transformed in *E.coli* BL21(DE3) – pETDuet1(Ferredoxin, Ferredoxin reductase). For each site, 32 distinct variant possibilities exist (32 possible codon substitutions). Hence, the screening of 188 transforms per site should be enough to cover more than 95% of all created variants.^[11]

Since 11 of the selected 20 sites were screened before, only 9 sites were used for the first round of evolution. In the first round, each of 9 key amino acid sites was mutated by saturation mutagenesis. The best mutant (I83H) was selected based on high-throughput *ee* assay and it was used as the template for the next round. In the second round, each of other 19 amino acid sites was mutated by saturation mutagenesis except the best amino acid site selected in round 1 (I83). Similarly the best mutant (I83H/M305Q) was then selected and used for the third round of evolution. In round 3, 5 best sites were used for evolution (A77, P99, N100, S182, and L302) which gave the best triple mutant I83H/M305Q/A77S.

Table S5. Primers for site-directed mutagenesis in the first round

I82	I82-F	GGA TAT GGC GGC NNK ATA ATC GAT GAC
	I82-R	GTC ATC GAT TAT MNN GCC GCC ATA TCC
I83	I83-F	GGC GGC ATC NNK ATC GAT GAC G
	I83-R	CGT CAT CGA TMN NGA TGC CGC C
M305	M305-F	CTT GCT CAT NNK CGC CGC ACG
	M305-R	CGT GCG GCG MNN ATG AGC AAG
D183	D183-F	GCT GGT CGN NKG TGA CAA CC
	D183-R	GGT TGT CAC MNN CGA CCA GC
S182	S182-F	CTT ACC CGC NNK TCG GAT GTG AC
	S182-R	GTC ACA TCC GAM NNG CGG GTA AG
A77	A77-F	CTC GTC CGA TNN KGG ATA TGG CG
	A77-R	CGC CAT ATC CMN NAT CGG ACG AG
A103	A103-F	CCA ATT TCA TCN NKA TGG ATC GGC C
	A103-R	GGC CGA TCC ATM NNG ATG AAA TTG G
P99	P99-F	GGA CTG GAT CTT NNK AAT TTC ATC GCG
	P99-R	CGC GAT GAA ATT MNN AAG ATC CAG TCC
L98	L98-F	GCG GAC TGG ATN NKC CCA ATT TC
	L98-R	GAA ATT GGG MNN ATC CAG TCC GC
F403	F403-F	CGT TCA AAT NNK GTG CGC GG
	F403-R	CCG CGC ACM NNA TTT GAA CG
L302	L302-F	GCA AAC ACC GNN KGC TCA TAT GC
	L302-R	GCA TAT GAG CMN NCG GTG TTT GC
V254	V254-F	CCT GAT CNN KGG CGG GAA C
	V254-R	GTT CCC GCC MNN GAT CAG G
D258	D258-F	CGG GAA CNN KAC CAC ACG
	D258-R	CGT GTG GTM NNG TTC CCG
L251	L251-F	GTA CTT NNK CTG ATC GTT GGC G
	L251-R	CGC CAA CGA TCA GMN NAA GTA C
T185	T185-F	GGA TGT GNN KAC CGC AGC
	T185-R	GCT GCG GTM NNC ACA TCC
T259	T259-F	CGG GAA CGA TNN KAC ACG C
	T259-R	GCG TGT MNN ATC GTT CCC G
G255	G255-F	CCT GAT CGT TNN KGG GAA CG
	G255-R	CGT TCC CMN NAA CGA TCA GG
N100	N100-F	GGA TCT TCC CNN KTT CAT CGC
	N100-R	GCG ATG AAM NNG GGA AGA TCC
I102	I102-F	CCC AAT TTC NNK GCG ATG GAT C
	I102-R	GAT CCA TCG CMN NGA AAT TGG G
T186	T186-F	GAT GTG ACA NNK GCA GCA CC

	T186-R	GGT GCT GCM NNT GTC ACA TC
--	--------	----------------------------

7 Directed evolution: enantioselectivity determination by MS assay and chiral HPLC analysis

For the evolution of each site, 192 selected colonies (4 controls: wild type P450pyr) were inoculated in 0.5ml TB medium (100 µg/ml Ampicilin, 50 µg/ml Kanamycin) in 2 96 deep well plates at 300rpm and 37 °C for 5.5 h, then induced with IPTG 0.5 mM, 5-Aminolevulinic acid hydrochloride (ALA) and FeCl₃ 0.5 mM in 0.5 ml TB, followed by shaking at 300 rpm and 30 °C for 12-14 h. The cells were divided equally to 2 plates. The plates were centrifuged at 5000 g and 4 °C for 10 min. After removing the supernatant in each well of the plates, the cells in each well were resuspended in 0.5 ml 50 mM phosphate buffer (pH 7.5) containing 2% glucose and 1 mM of either *R* or *S*-deuterated *N*-benzyl pyrrolidine respectively. The plates were then shaken at 300 rpm and 30 °C for 2 h; follow by centrifugation at 5000 g and 4 °C for 10 min. The supernatant was used for LC-MS analysis to obtain the ratio of signal [M]/[M+1]. The product *ee* was then calculated.^[12]

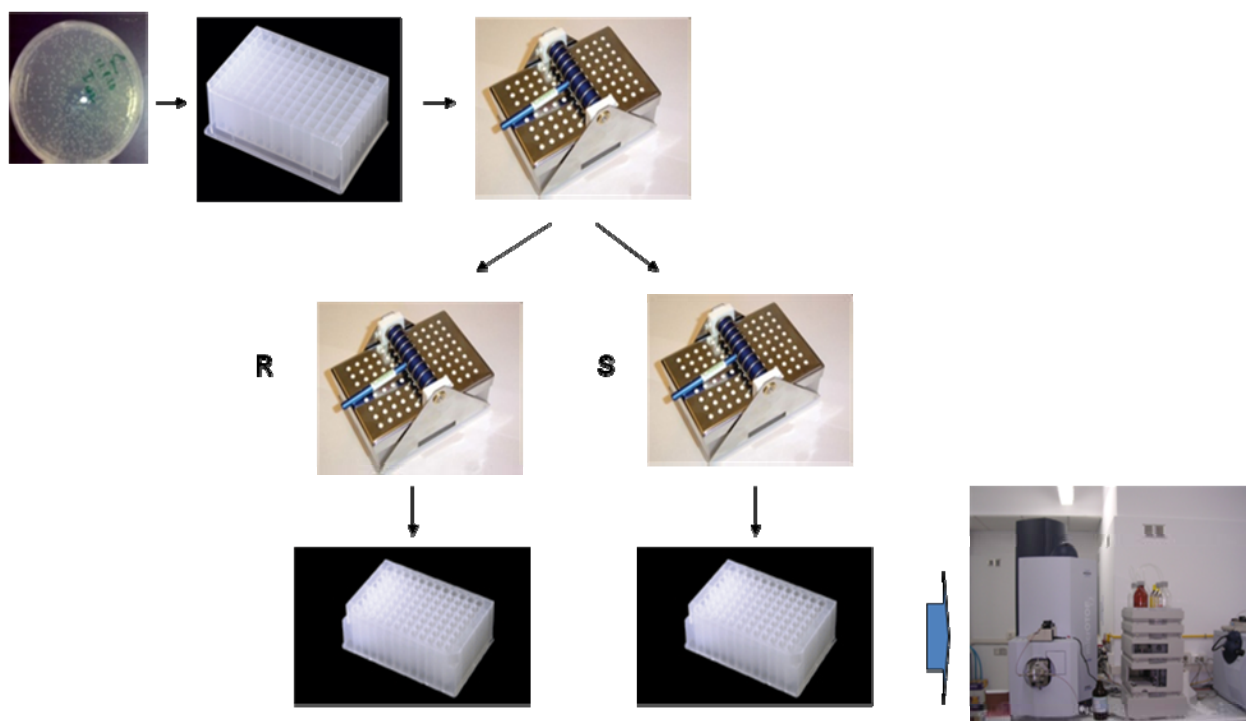


Figure S13. Screening procedure of the generated enzyme mutants

LC-MS analysis was performed by using Bruker UPLC-microTOF Q MS. Analytical conditions: LC: Water: Acetonitrile = 90:10, 1 min, 0.2 ml/min, sample injection 5 μ L, no column; MS: positive mode, 4 bar nebulizing gas, 4.9 kV spray voltage, drying gas flow of 8 L/min, drying gas temp. of 200 $^{\circ}$ C.

For the mutants which high *ee* determined by MS assay, their enantioselectivity were rechecked by chiral HPLC assay. The selected mutants were isolated and grown in shaking flask (50 ml). When the OD₆₀₀ reach 0.6 – 0.8, cells were induced with IPTG 0.5 mM and ALA 0.5 mM and further grown for 6 – 8 h at 25 $^{\circ}$ C. Cells were harvested at OD₆₀₀ from 6 – 9 by centrifugation at 5000 g and 4 $^{\circ}$ C and resuspended to 4 – 10 g/l cdw in 50 mM phosphate buffer (pH 7.5) for biotransformation with 10 mM *N*-benzyl pyrrolidine and 2% glucose at 30 $^{\circ}$ C and 300 rpm for 2 h. Cells were removed by centrifugation at 5000 g and 4 $^{\circ}$ C and the supernatant was analyzed by reverse HPLC [Eclipse Plus C18, 150 mm x 4.6 mm; eluent acetonitrile: 10 mM phosphate buffer (pH 7), 80: 20; flow rate: 1 ml/min; detection wavelength: 210 nm; retention time: *N*-benzyl-3-hydroxypyrrolidine (product), benzyl

alcohol (internal standard), N-benzyl pyrrolidine (substrate) = 4.9, 6.5, 8.4]. The supernatant was adjusted to pH 11-12 by addition of 5M KOH followed by extraction with ethyl acetate (3 times). The organic phase was separated and dried over Na₂SO₄, and the solvent was removed by evaporation. The pellet was dissolved in methanol and analyzed by chiral HPLC [Chiralcel OB-H (Daicel), 250 mm x 4.6 mm; eluent hexane:isopropanol, 98:2; flow rate: 1 mL/min; 4 μL injection volume; detection wavelength: 210 nm; t_R of (*R*), 15; t_R of (*S*), 23]. In the triple mutant (I83H/M305Q/A77S) case, the product was purified by column chromatography on aluminum oxide before analysis by chiral HPLC.

8 Reverse HPLC chromatogram

Mobile Phase : 80% Acetonitrile - 20% Phosphate buffer (pH 7.5) 10 mM

Injection Volume : 4 uL

Flow rate : 1 ml/min

Column : Eclipse Plus C18

Round 1 of evolution

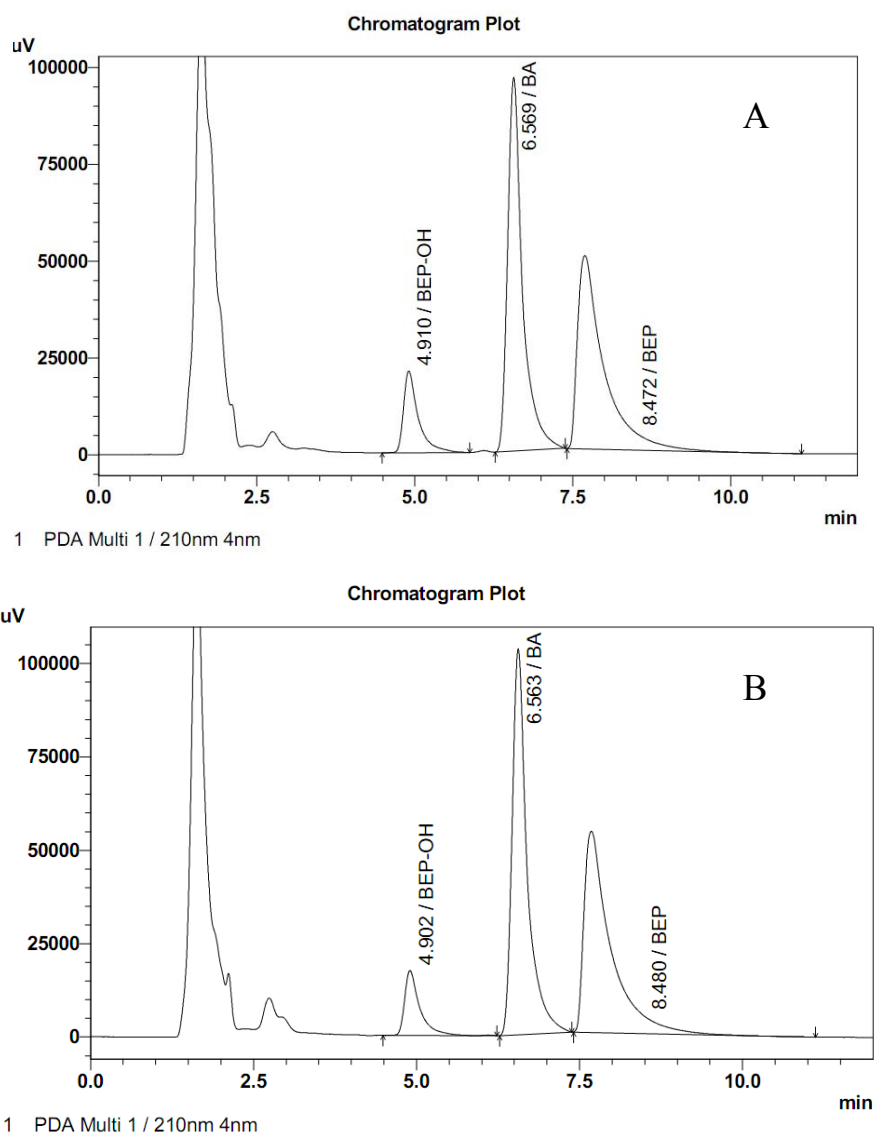


Figure S14. Reverse HPLC chromatogram of A: wild type P450pyr; B: P450pyr mutant I83H

Round 2 of evolution

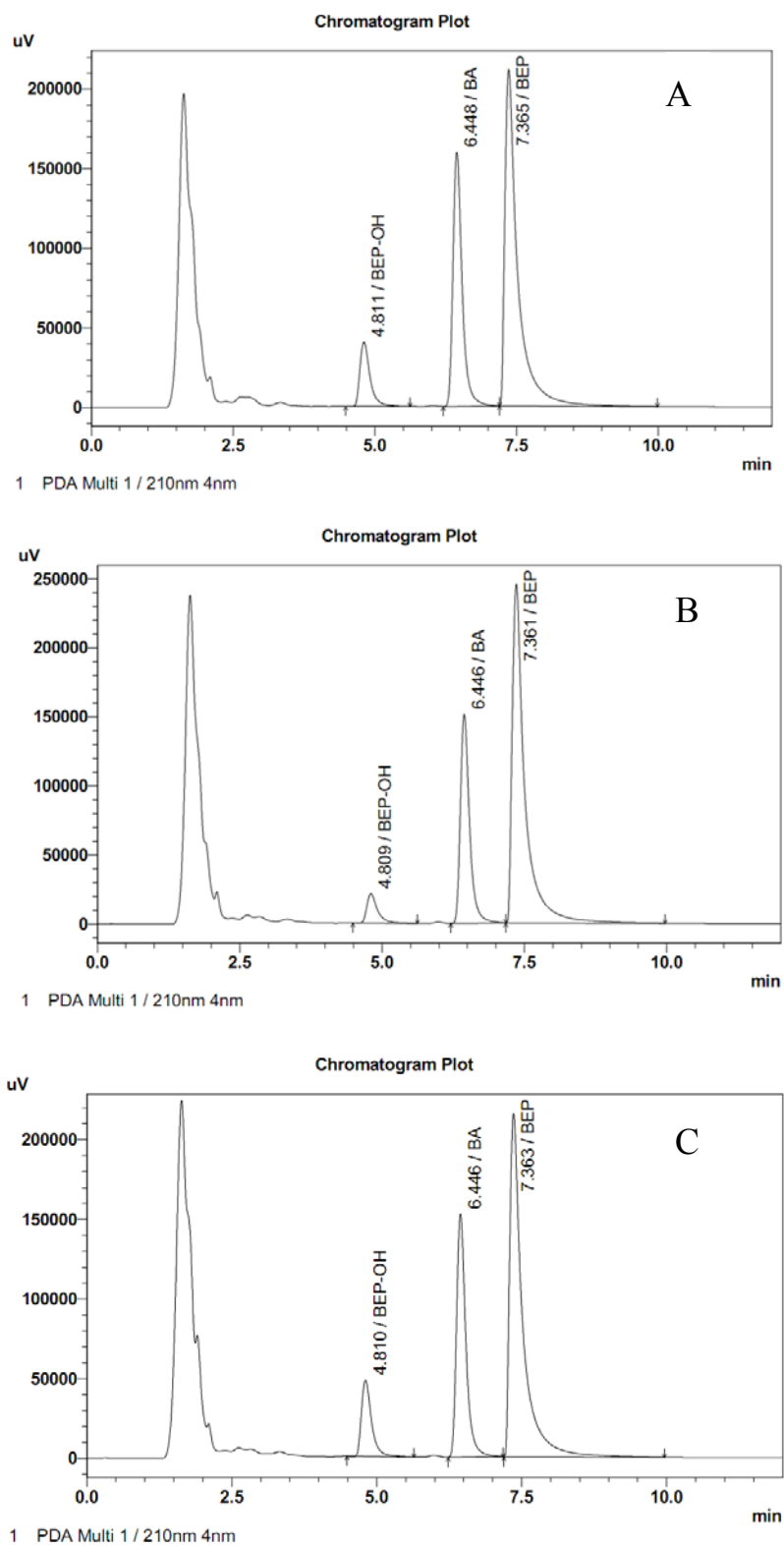


Figure S15. Reverse HPLC chromatogram of A: P405pyr mutant I83H; B: P405pyr mutant I83H/M305Q; C: P405pyr mutant I83H/A77S

Round 3 of evolution

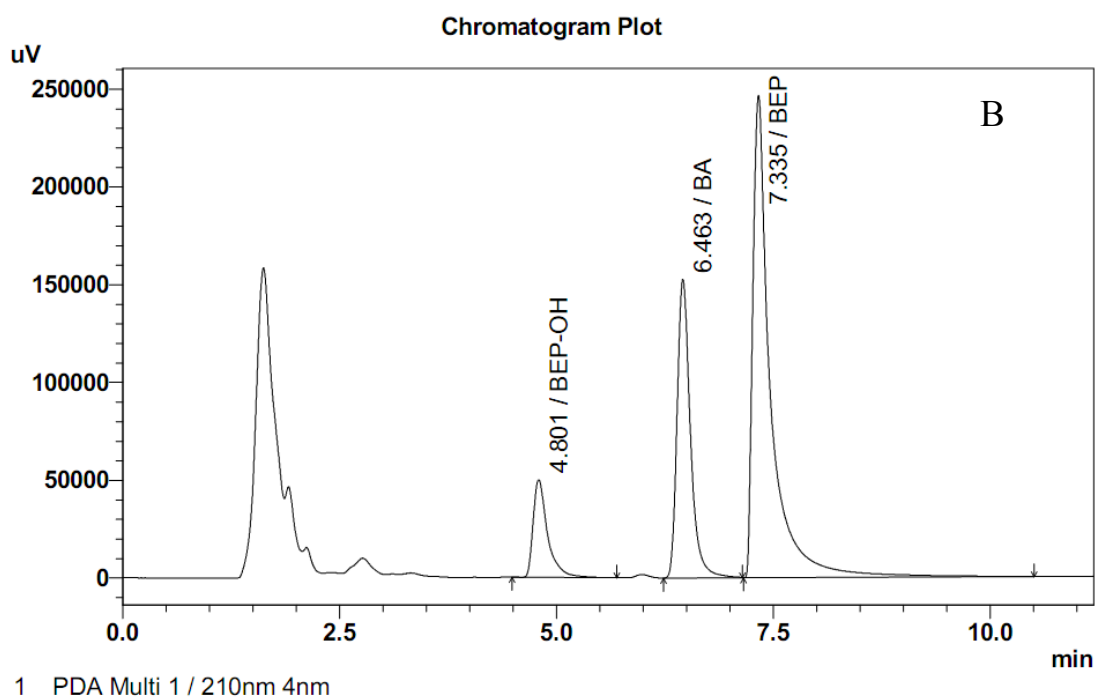
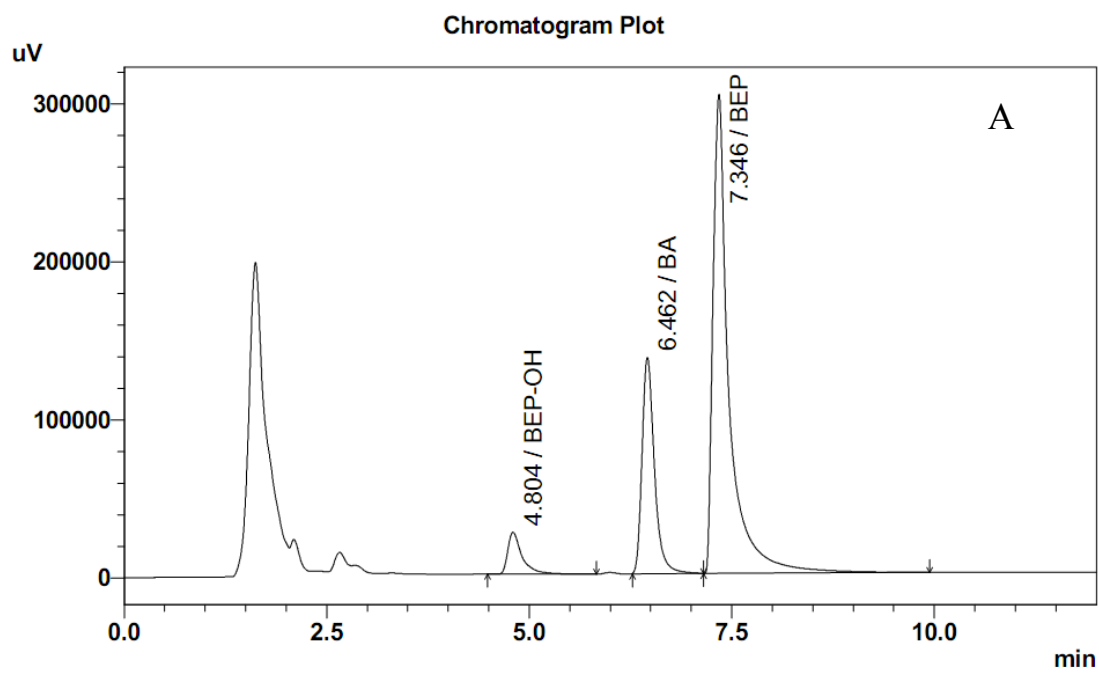


Figure S16. Reverse HPLC chromatogram of A: P405pyr mutant I83H/M305Q; B: P405pyr mutant I83H/M305Q/A77S

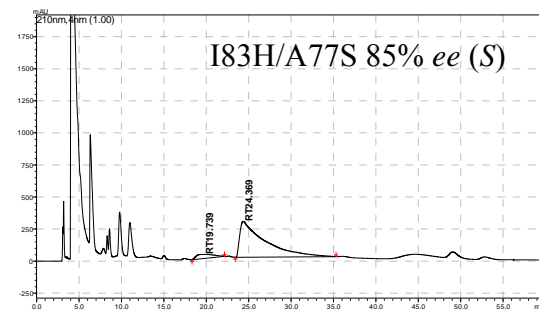
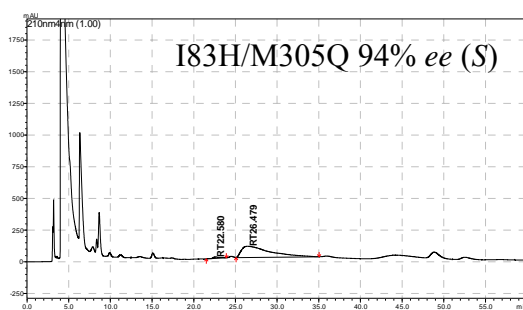
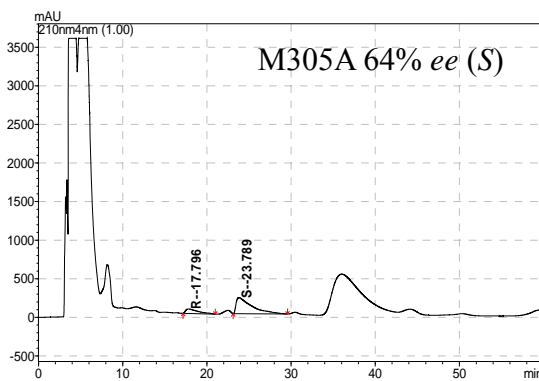
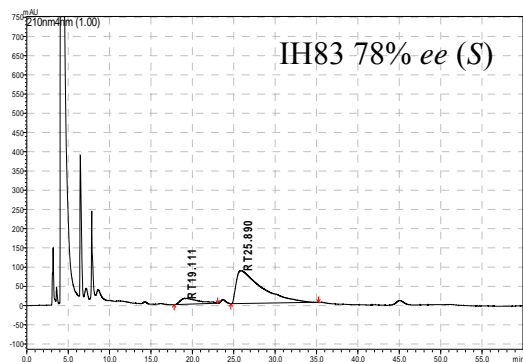
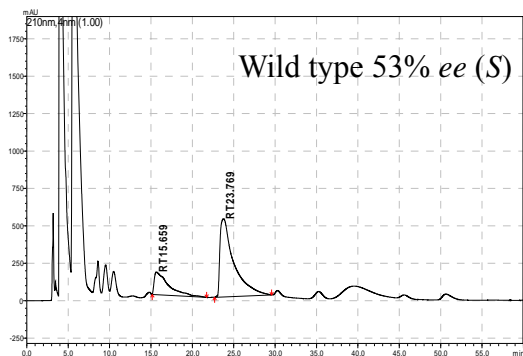
9 Chiral HPLC chromatogram

Mobile Phase : 98% n-hexane - 2% Isopropanol

Injection Volume : 1 μ L

Flow rate : 1 ml/min

Column : Chiralcel OB-H



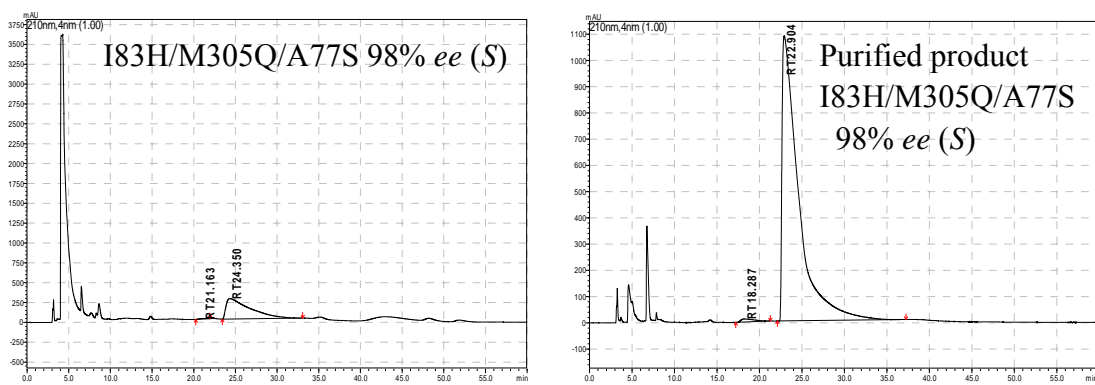


Figure S17. Chiral HPLC chromatogram of wild type P450pyr and best mutants generated in round 1, round 2 and round 3 respectively.

10 Substrate docking on protein structures by molecular dynamics (MD) simulation

The missing part of original X-ray structure of P450pyr was reconstructed by Modeller¹³ software and refined with PyRosetta.¹⁴ The structure of mutants was prepared with mutagenesis utility of PyMOL from the wild type structure. The protonation state of the individual residue was assigned by the value at neutral condition. The active site was set as the resting state. MD simulation was conducted with Gromacs 4.07 software¹⁵. The force field parameters were taken from the well-established Gromos53A6 force field¹⁶ and the literatures.^{17,18} Protein was solvated in a dodecahedral box containing SPC water and 100 mM of NaCl. During the whole modeling process, the electrostatic interaction was calculated by PME method, while Van de Waals interaction was computed with cut-off at 0.9nm. Bond length was constrained by LINCS method. The integration step was 2 fs.

The system was then minimized by steep descent method and equilibrated with several short restrained or unrestrained MD. The final production MD was conducted at isothermal-isobaric condition (NPT ensemble) at 300 K and 1 atm, which was coupled by Berendsen thermostat and Parrinello-Rahman barostat, respectively. The production MD covered of 10 ns of trajectory and 2500 structures were

recorded, from which the average structure was derived, and was adequately minimized and submitted to molecular docking.¹⁹

The docking was accomplished with Autodock Vina software.²⁰ The heme was set as ferryl-oxo-heme complex known as CpdI state.²¹ The final catalytically active pose was screened from the generated docking poses with established geometric criteria.^{21, 22}

As shown in the Figure S18, the productive geometry of substrate appeared within a shallow pocket containing Ile82, Leu102, Leu251, Val254 and Leu302. The benzyl group oriented towards the big loop (Ser75 to Asp105) in protein structure, and stacked with three hydrophobic amino acids: Ile82, Ile102 and Leu251. The pyrrolidine group posed towards the heme, and had its C3 atom located within 3.6 Å from the oxygen atom coordinated to the heme (heme-O). Two hydrophobic amino acids (Val254 and Leu302) were found in the vicinity of the pyrrolidine ring. The Leu302 weakly interacted to the pyrrolidine ring at the distance of 3.8 Å; while the side chain of Val254 oriented opposite to the pyrrolidine group and hardly impacted on the substrate's pose. Consequently, the substrate had been less restrained as there was sufficient space for the pyrrolidine group rotating to take other alternative positions.

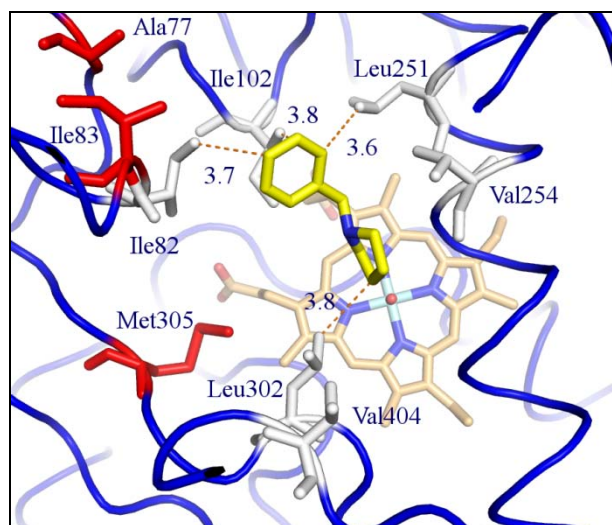
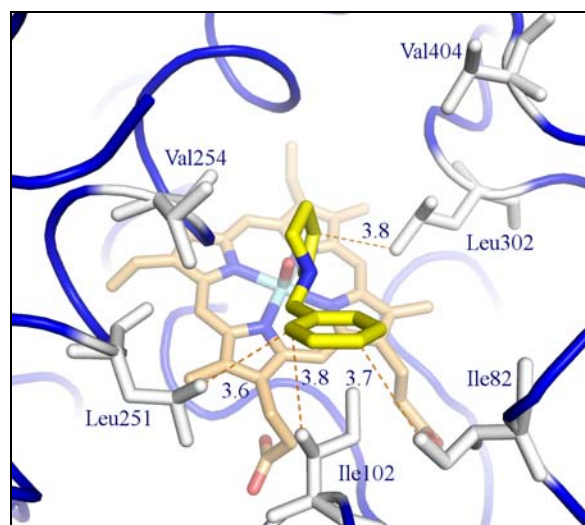


Figure S18. The side-view (left) and top-view (right) of substrate productive pose in wild type enzyme. The three mutation sites were illustrated as red stick. Hydrophobic amino acid residues were shown in white and labeled with the distance to the substrate.

In the triple mutant, the introduced Ser77 and His83 offered additional hydrogen bonds (Ser77OG-Gln89O, Ser77N-Gln89NE and His83NE2-Ile88N) within the big loop (Figure S19), which made the loop less flexible. The direct consequence of the reduced loop flexibility was that the C-terminal of the loop moved closer to the heme by approximately 1 Å. The Ile102 hence occupied the position for the substrate's benzyl group in the original substrate docking pose in the wild type enzyme (Figure S19). This unfavoured interaction flipped the substrate's docking pose over in the triple mutant.

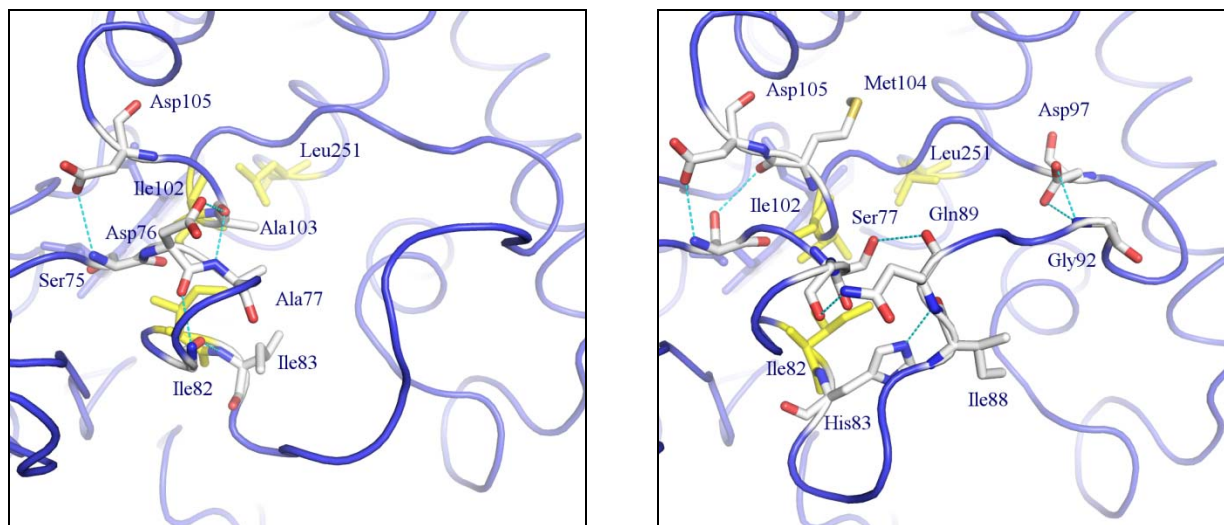


Figure S19. The loop (Ser75-Asp105) section of wild type of P450pyr (left) and triple mutant (right). The three hydrophobic amino acid residues interacts with substrate's benzyl group was shown in yellow sticks. Hydrogen bonds were represented as cyan dash line.

As shown in Figure S20, the substrate located within a narrow cleft consisting of Ile102, Val254, Leu302 and Val404, which may stabilize the substrate through hydrophobic interaction. The Ile82 and Leu251 that previously involved in substrate binding in the wild type enzyme did not take part in as long distance (>7 Å) was recorded. The C3 atom of pyrrolidine ring was 3.3 Å from the heme-O. Since the space within this binding site was very limited, the orientation of the substrate molecule was almost

constrained at the productive pose, and due to the relative geometric positions of Ile102 and Leu302, this geometric feature resulted in considerably high potential to form the *S*-hydroxylated product.

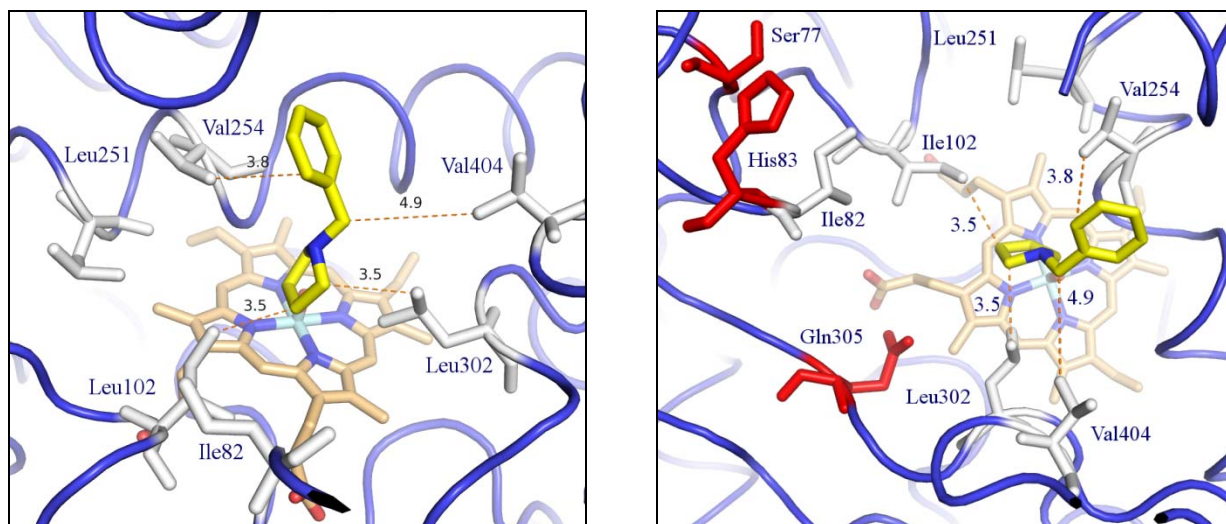


Figure S20. The side-view (left) and top-view (right) of substrate productive pose in the triple mutant. The three mutation sites were illustrated as red stick. Hydrophobic amino acid residues were shown in white and labeled with the distance to the substrate.

11 Kinetic study of wild type P450pyr and P450 mutant I83H/M305Q/A77S

11.1 Kinetic study of CFE of *E. coli* recombinant expressing P450pyr

E. coli BL21(DE3) pRSFDuet1-P450(wild type) pETDuet1-FR and *E. coli* BL21(DE3) pRSFDuet1-P450(mutant I83H/M305Q/A77S) pETDuet1-FR were inoculated overnight in LB medium as the seed culture. 1% seed were transferred in 50ml TB medium in 250ml shaking flask with baffles. Cells were grown at 37 °C to OD600 about 0.6 – 0.8 then induced by adding IPTG to final concentration of 0.25 mM (ALA and FeCl₃ were added to 0.5 mM). Cells were then grown at 25 °C for 5-7 hours to express protein. After harvesting cells were resuspended in 50 mM phosphate buffer (pH 7.5) and then broken by using cell disruptor at 21 psi, 2 times. Cell debris was removed by centrifugation (23000 rpm, 4 °C, and 50 min). The cell free extract was then use for biotransformation with CFE 4 g/l protein, *N*-benzyl pyrrolidine concentration 1 – 4 mM, NADH 2 mM, 30 °C, 300 rpm, 5 min.

The product (*N*-benzyl-3-hydroxypyrrolidine) concentration was analyzed by reverse HPLC [Eclipse Plus C18, 150 mm x 4.6 mm; eluent acetonitrile: 10 mM phosphate buffer (pH7), 80: 20; flow rate: 1 ml/min; detection wavelength: 210 nm]

The results are shown in figure S21 and table S6.

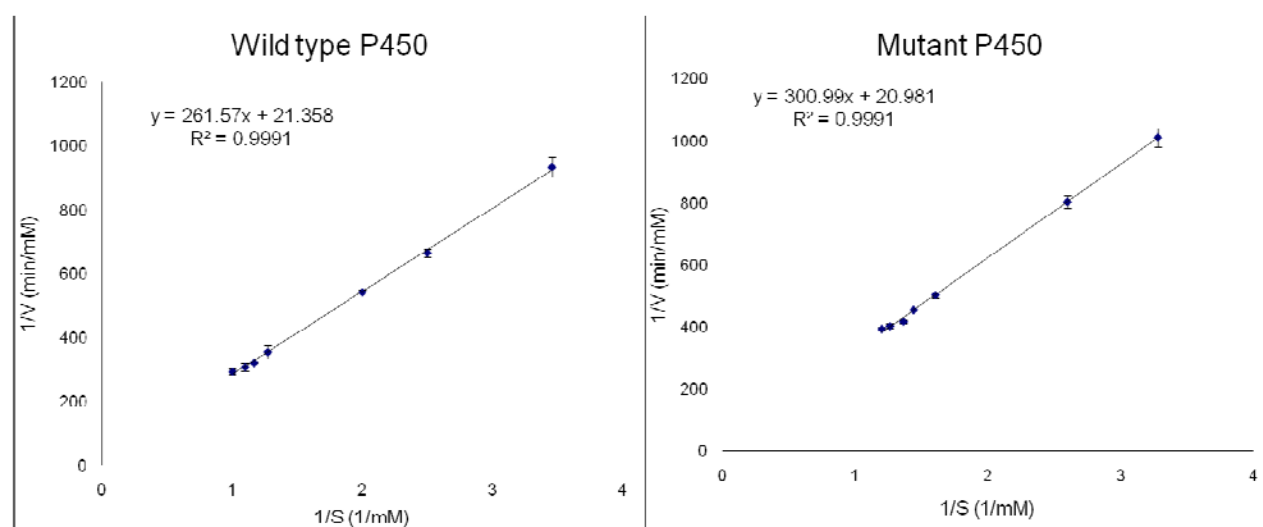


Figure S21. Lineweaver-Burk curve of kinetic data of cell free extract of the wild type and the best mutant of P450pyr.

Table S6. Kinetic data of cell free extracts of the wild type and the best mutant

	K_m	V_{max}	V_{max}/K_m
	(mM)	(μ M/min)	(μ M/min/mM)
CFE P450 Wild type	12.2	46.8	3.8
CFE P450 mutant I83H/M305Q/A77S	14.3	47.7	3.3

11.2 Kinetic study of purified his-tagged P450pyr and his-tagged P450pyr mutant I83H/M305Q/A77S

P450pyr mutant I83H/M305Q/A77S was cloned with his-tag in pRSFDuet1 at N-terminal using same

strategy described for wild type P450pyr. Ferredoxin and Ferredoxin Reductase were also cloned in pET21a to express his-tag at N-terminal.

- Design primers for Ferredoxin to express his-tag at N-terminal.

– (Forward primer) FdxpetF1

GGATTCCCATATG.CAT.CAT.CAT.CAT.CAT.CAT.GGC.AGC.AGC.CCA.ACA.GTG.ACC.TAT.
GTT.G

[NdeI site upstream of the start codon]

– (Reverse primer) FdxpetR1

GAATTCCCTCGAG.TCA.ATG.CTG.CGC.GAG.AGG.AAG

[XhoI site downstream of the stop codon]

- Design primers for Ferredoxin reductase to express his-tag at N-terminal.

– (Forward primer) FdRpetF1

GGATTCCCATATG.CAT.CAT.CAT.CAT.CAT.CAT.AGC.AGC.GGC.ATC.CAC.ACC.GGC.GTG
.ACC.G

[NdeI site upstream of the start codon]

– (Reverse primer) FdRpetR1

GAATTCCCTCGAG.TTA. GAG.GGA.GGT.TGG.GGA.CG T.TGC.G

[XhoI site downstream of the stop codon]

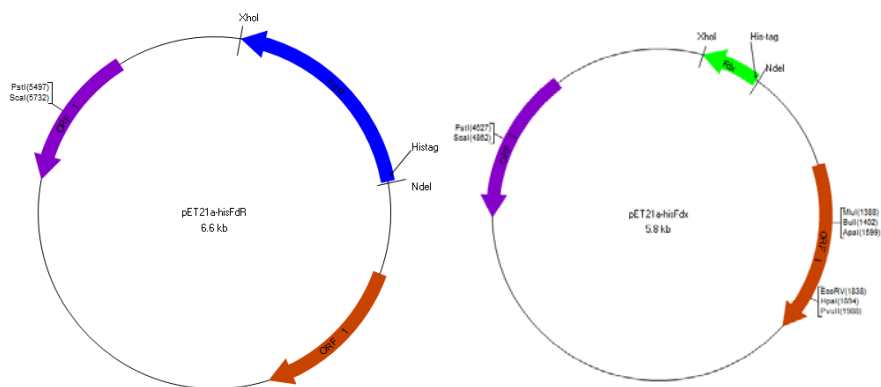


Figure S22. plasmid map of pET21a-hisFdR and pET21a-hisFdx.

Plasmids were then expressed in *E.coli* BL21(DE3) cells. Cells were grown in LB medium (50 ml) containing 100 µg/ml ampicilin at 37 °C. When OD600 reached 0.5–0.8, IPTG was added to a final concentration of 0.5 mM to induce the expression of his-tagged Ferredoxin or Ferredoxin reductase. During incubation for further 8.5 hours at 25 °C, the cells were taken at selected time point from the culture and harvested by centrifugation (6000 g for 10 min). The cells were then diluted to OD600 equal to 10, heated for 5 min at 95 °C, centrifuged at 13000g for 10min to get cell free extract (CFE). The CFE was mixed with SDS sample buffer at volume ratio 1: 1, heated for 5min at 95 °C and run SDS-PAGE. The expression of his-tagged Fdx was not clear which might the same reason as stated above. On the other hand, the expression of his-tagged FdR was proved by an observed band near 50 kDa.

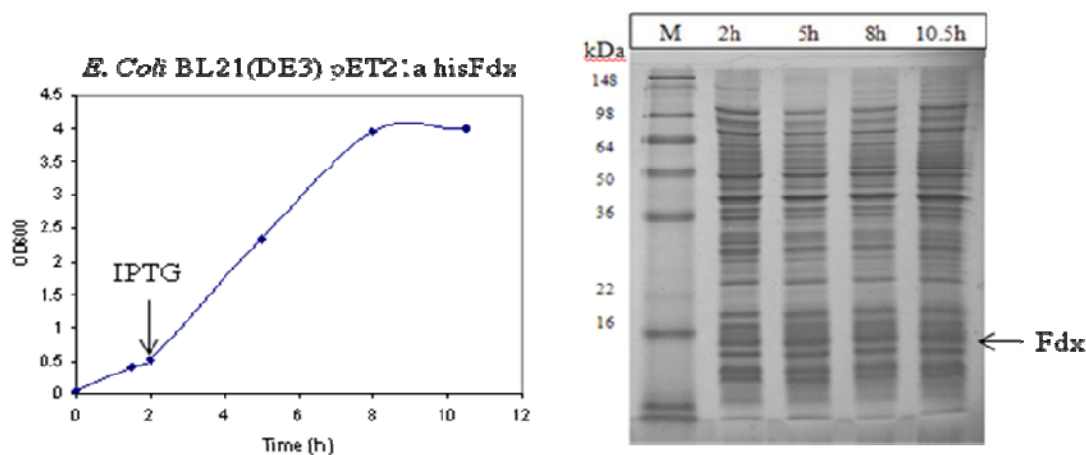


Figure S23. (left) Growth curve of *E. coli* BL21(DE3) pET21a hisFdx with his-tag at N-terminal. (right) SDS-PAGE of CFEs of cells taken at different time points.

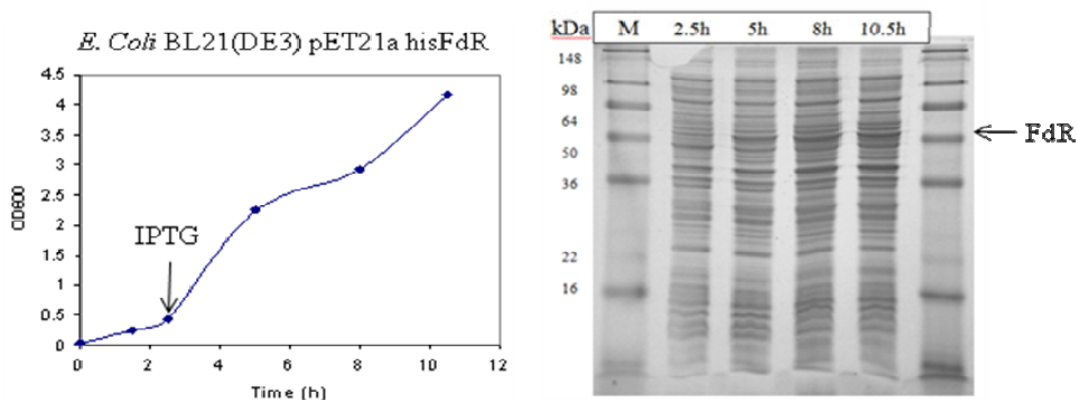


Figure S24. (left) Growth curve of *E. coli* BL21(DE3) pET21a hisFdR with his-tag at N-terminal. (right) SDS-PAGE of CFEs of cells taken at different time points.

The resulting recombinants were then inoculated overnight in LB medium as the seed culture. In each case, 1% seed were transferred in 400 ml TB medium in 2 L shaking flask with baffles. Cells were grown at 37 °C to OD600 about 0.6 – 0.8 then induced with 0.25 mM IPTG (in P450 expression, ALA and FeCl₃ were added to 0.5 mM). Cells were then grown at 22 °C for 12 – 14 hours to express the

protein. After harvesting, cells were resuspended in 50 mM NaH_2PO_4 buffer containing 0.3 M NaCl and 10 mM imidazole (pH 8). The cell suspensions were passed through a homogenizer (21 psi, 2 times). Cell debris was removed by centrifugation (23000 rpm, 4 °C, and 50 min). The CFEs were subjected to purification by FPLC with 2 chromatography columns Ni-NTA and SEC as described before.

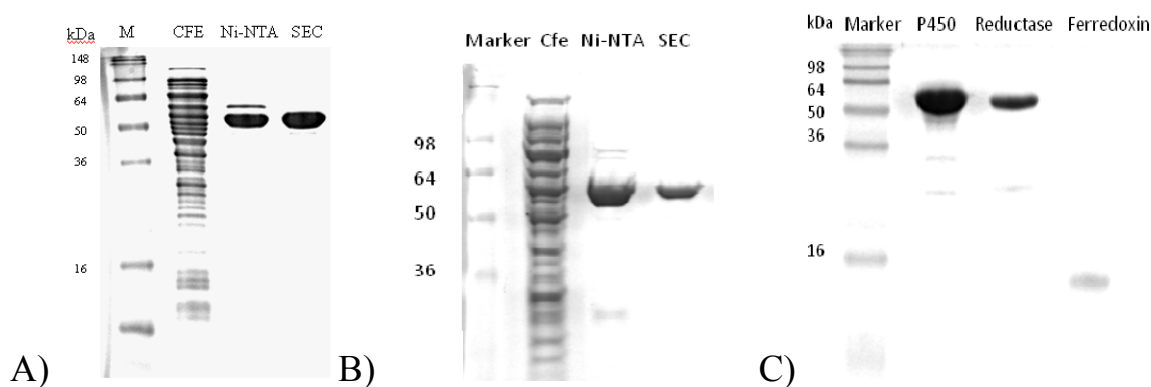


Figure S25. A) SDS-PAGE of purified his-tagged P450pyr. B) SDS-PAGE of purified his-tagged P450pyr mutant I83H/M305Q/A77S. C) SDS-PAGE of purified his-tagged P450pyr hydroxylase, his-tagged Ferredoxin and his-tagged Ferredoxin reductase. [CFE: cell free extract; Ni-NTA: elution fraction after Ni-NTA column; SEC: eluted fractions after SEC column]

To study the kinetic, the purified his-tagged P450pyr hydroxylase (wild type or best mutant) was mixed with purified his-tagged Ferredoxin and Ferredoxin reductase to a molar ratio of 1: 10: 1 in 50 mM phosphate buffer (pH 7.5). Biotransformation with this mixture purified enzymes was done as P450: Fdx: FdR = 5 uM: 50 uM: 5 uM, 0.4 – 5 mM *N*-benzyl pyrrolidine, 2 mM NADH, 30 °C, 300 rpm, 5 min.

The product (*N*-benzyl-3-hydroxypyrrolidine) concentration was determined by reverse HPLC [Eclipse Plus C18, 150 mmx4.6 mm; eluent acetonitrile: 10 mM phosphate buffer (pH 7), 80: 20; flow rate: 1ml/min; detection wavelength: 210 nm].

The results are shown in figure S26 and table S7.

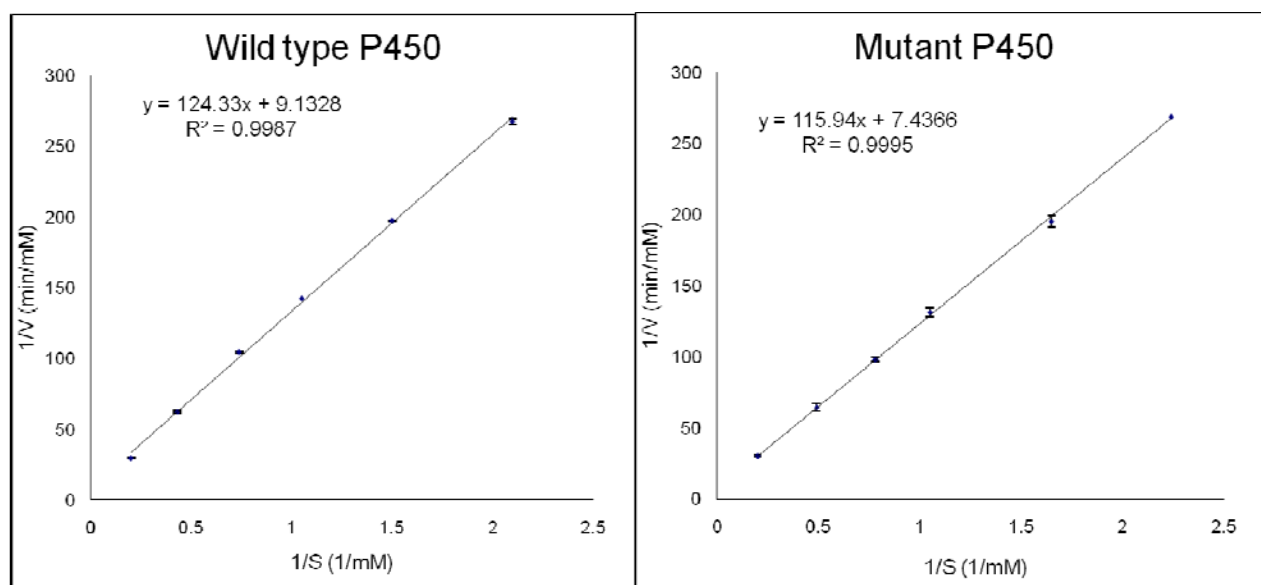


Figure S26. Lineweaver-Burk curve of kinetic data of purified his-tagged P450pyr and his-tagged P450pyr mutant I83H/M305Q/A77S

Table S7. Kinetic data of purified his-tagged P450pyr and his-tagged P450pyr mutant I83H/M305Q/A77S

	K_M	V_{max}	k_{cat}	k_{cat}/K_M
	(mM)	($\mu\text{M}/\text{min}$)	(1/min)	(1/min.mM)
His-tagged P450pyr	13.6	109.5	21.9	1.6
His-tagged P450pyr mutant I83H/M305Q/A77S	15.6	134.5	26.9	1.7

12 Reference

- [1] J. B. Beilen, E. G. Funhoff, A. V. Loon, A. Just, L. Kaysser, M. Bouza, R. Holtackers, M. Rothlisberger, Z. Li and B. Witholt, *Appl. Environ. Microbiol.*, 2006, **72**, 59.
- [2] W. L. Tang, Z. Li and H. Zhao, *Chem. Commun.*, 2010, **46**, 5461.
- [3] S. Omura and R. Sato, *J. Biol. Chem.*, 1964, **239**, 2379.
- [4] G.M. Sheldrick, *Acta Crystallogr. Sect. A*, 2008, **64**, 112.
- [5] R.J. Morris, A. Perrakis and V.S. Lamzin, *Methods Enzymol.*, 2003, **374**, 229.
- [6] G.N. Murshudov, A.A. Vagin and E.J. Dodson, *Acta Crystallogr. Sect. D*, 1997, **53**, 240.
- [7] L. Holm and P. Rosenström, *Nucleic Acids Res.*, 2010, **38**, 545.
- [8] C.A. Hasemann, K.G. Ravichandran, J.A. Peterson and J. Deisenhofer, *J. Mol. Biol.*, 1994, **236**, 1169.
- [9] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, G.M. Greenblatt, E.C. Meng and T.E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605.
- [10] Collaborative Computational Project Number 4, *Acta Crystallogr. Sect. D*, 1994, **50**, 760-763.
- [11] M. T. Reetz and J. D. Carballeira, *Nat. Protoc.*, 2007, **2**, 891.
- [12] Y. Chen, W. L. Tang, J. Mou and Z. Li, *Angew. Chem. Int. Ed.*, 2010, **49**, 5278.
- [13] A. Fiser and A. Sali, *Meth. Enzymol.*, 2003, **374**, 461.
- [14] S. Chaudhury, S. Lyskov and J. J. Gray, *Bioinformatics*, 2010, **26**, 689.
- [15] B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J Chem Theory Comput*, 2008, **4**, 435.
- [16] C. Oostenbrink, A. Villa, A. E. Mark and W. F. Van Gunsteren, *J Comput Chem*, 2004, **25**, 1656.
- [17] F. Autenrieth, E. Tajkhorshid, J. Baudry and Z. Luthey-Schulten, *J Comput Chem*, 2004, **25**, 1613.
- [18] A. Seifert, S. Tatzel, R. D. Schmid and J. Pleiss, *Proteins*, 2006, **64**, 147.

- [19] S. Kille, F.E. Zilly, J.P. Acevedo and M.T. Reetz, *Nat. Chem.*, 2011, **3**, 738.
- [20] O. Trott and A. J. Olson, *J Comput Chem*, 2010, **31**, 455.
- [21] I. G. Denisov, T. M. Makris, S. G. Sligar and I. Schlichting, *Chemical reviews*, 2005, **105**, 2253.
- [22] S. Shaik, D. Kumar, S. P. de Visser, A. Altun and W. Thiel, *Chemical reviews*, 2005, **105**, 2279.