

Supporting Information

**Evolving P450_{pyr} Monooxygenase for Highly
Regioselective Terminal Hydroxylation of *n*-
butanol to 1,4-butanediol**

Yi Yang, Yu Tse Chi, Hui Hung Toh and Zhi Li*

Department of Chemical and Biomolecular Engineering, National University of Singapore, 4
Engineering Drive 4, Singapore 117585

E-mail: chelz@nus.edu.sg

Table of Contents

| | |
|--|----|
| 1. Chemicals | 3 |
| 2. Strains and biochemicals | 3 |
| 3. Identification of suitable amino acid residues of P450pyr for Iterative Saturation Mutagenesis (ISM) | 4 |
| 4. Generation of P450pyr monooxygenase mutant library | 5 |
| 5. Calibration of purpald assay for determination of formaldehyde..... | 7 |
| 6. General procedure of screening P450pyr mutants for terminal hydroxylation of alcohol by using surrogate substrate-based colorimetric HTS assay | 8 |
| 7. General procedure for biohydroxylation of <i>n</i> -butanol with <i>E.coli</i> expressing positive P450pyr mutants in shaking flask..... | 9 |
| 8. GC analysis of 1,4-butanediol | 10 |
| 9. MS analysis of the derivative of 1,4-butanediol made by BSTFA-TMCS method..... | 11 |
| 10. Molecular modelling of substrates docking in P450pyr and its mutants..... | 12 |
| 11. References | 15 |

1. Chemicals

Following chemicals were purchased from Sigma-Aldrich and used without further purification:

2-methoxyethanol ($\geq 99\%$), *n*-butanol ($\geq 99\%$), 1,2-butanediol ($\geq 98\%$), 1,3-butanediol ($\geq 99\%$), 1,4-butanediol ($\geq 99\%$), purpald ($\geq 99\%$), δ - aminolevulinic acid hydrochloride (ALA) ($\geq 97\%$), Isopropyl β -D-1-thiogalactopyranoside (IPTG) D-Glucose ($\geq 99.5\%$), dimethylformamide (DMF) ($\geq 99\%$), ethanol ($\geq 99\%$), ammonium sulphate ($\geq 99\%$), BSTFA (contains 1% TMCS, 99%), Sodium phosphate dibasic ($\geq 99\%$), and Sodium phosphate monobasic ($\geq 99\%$).

2. Strains and biochemicals

Escherichia coli BL21(DE3) strain was purchased from Novagen. Plasmids pRSFDuet-P450pyr and pETDuet-Fdx-FdR were obtained from our own laboratory collections. Restriction enzyme *Dpn* I, Deoxynucleotide (dNTP) Solution Mix, and Q5 Hot Start High-Fidelity DNA Polymerase were purchased from New England Biolabs (NEB). Tris-acetate-EDTA (TAE) buffer, DNA loading dye, and DNA ladder were purchased from Thermo Scientific. Oligonucleotides were synthesised by AIT biotech, Singapore. LB Broth, Bacto Yeast Extract, and Bacto Tryptone were purchased from Biomed Diagnostics. Antibiotic ampicillin and kanamycin were purchased from Sigma-Aldrich. QIAquick Gel Extraction Kit (Qiagen) and QIAprep spin plasmid miniprep Kit were purchased from Qiagen.

3. Identification of suitable amino acid residues of P450pyr for Iterative Saturation Mutagenesis (ISM)

To identify the suitable amino acid residues for iterative saturation mutagenesis (ISM), *n*-butanol **1** was docked into the x-ray structure of P450pyr with the catalytic heme modified as the Cpd I ferryl-oxo-complex.^{S1} The catalytically relevant active binding posture for *n*-butanol was obtained based on established geometric criteria for substrate-heme binding in P450 enzymes.^{S2} As shown in Figure S1, 22 amino acids residues were selected for ISM. 11 amino acids were located within ~ 6.5 Å of the docked substrate: (i) important hydrophobic residues: Ile82, Ile83, Ile102, Leu302, Met305, and Phe403; and (ii) residues on the heme-proximal helix: Leu251, Val254, Gly255, Asp258, and Thr259. 4 amino acids located along another helix (Ser182, Asp183, Thr185 and Thr186) were also chosen due to their relatively abrupt position in the substrate access channel. 7 important amino acids from the ‘big loop’ (Ala77, Ile88, Gln89, Leu98, Pro99, Asn100, and Ala103) were also chosen for ISM, since in our previous study an improvement in P450pyr enantioselectivity was attributed to a reshaped binding pocket *via* inward movement of this ‘big loop’ (Ser74-Asp105).^{S1}

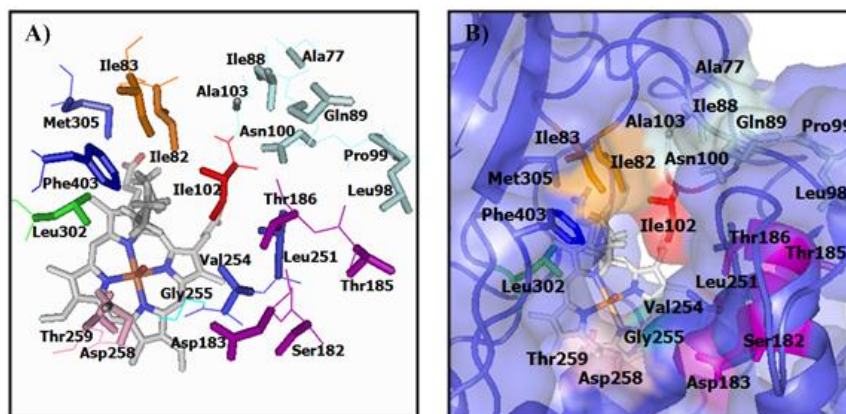


Figure S1. (A) Spatial overview of 22 selected amino acid residues for ISM, in relation to the active binding posture of *n*-butanol (grey cartoon stick with hydrogen atoms) in P450pyr. (B) Surface presentation of the same overview with additional structural motifs for a more aesthetic appreciation of the enzyme-substrate binding posture.

4. Generation of P450pyr monooxygenase mutant library

To generate P450pyr monooxygenase mutant library, PCR was carried out on each of the selected amino acid site for ISM using the designed primers shown in Table S1. Each PCR reaction tube contained the following mixture: 2.5 μ L forward and reverse primers of a particular target site, 10 μ L Q5 reaction buffer, 1 μ L 10 mM dNTP mixture, 0.5 μ L (10 ng) template DNA, 0.5 μ L Q5 Hot Start High-Fidelity DNA polymerase and 33 μ L of nuclease-free H₂O. PCR amplification was carried out on Bio-Rad S1000 Thermal Cycler using the following thermal cycling program: 98 °C for 30 s, 25 cycles [98 °C for 10 s, annealing temperatures for each pair of primers (see Table S1 T_m*) for 30 s, 72 °C for 3 min], and 72 °C for 5 min.

Table S1. Primer sequences used for site-directed mutagenesis

| Amino acid | | Primer Sequence (5' to 3') ^a | T _m * (°C) ^b |
|------------|--------|---|------------------------------------|
| A77 | A77-F | C TCG TCC GAT NNK GGA TAT GGC G | 69 |
| | A77-R | C GCC ATA TCC MNN ATC GGA CGA G | |
| I82 | I82-F | GGA TAT GGC GGC NNK ATA ATC GAT GAC | 69 |
| | I82-R | GTC ATC GAT TAT MNN GCC GCC ATA TCC | |
| I83 | I83-F | GGC GGC ATC NNK ATC GAT GAC G | 70 |
| | I83-R | C GTC ATC GAT MNN GAT GCC GCC | |
| I88 | I88-F | C GAT GAC GGC NNK CAA AAA GG | 65 |
| | I88-R | CC TTT TTG MNN GCC GTC ATC G | |
| Q89 | Q89-F | GAC GGC ATT NNK AAA GGT GGC G | 68 |
| | Q89-R | C GCC ACC TTT MNN AAT GCC GTC | |
| L98 | L98-F | GC GGA CTG GAT NNK CCC AAT TTC | 67 |
| | L98-R | GAA ATT GGG MNN ATC CAG TCC GC | |
| P99 | P99-F | GGA CTG GAT CTT NNK AAT TTC ATC GCG | 67 |
| | P99-R | CGC GAT GAA ATT MNN AAG ATC CAG TCC | |
| N100 | N100-F | G GAT CTT CCC NNK TTC ATC GC | 63 |
| | N100-R | GC GAT GAA MNN GGG AAG ATC C | |
| I102 | I102-F | CCC AAT TTC NNK GCG ATG GAT C | 65 |
| | I102-R | G ATC CAT CGC MNN GAA ATT GGG | |
| A103 | A103-F | CC AAT TTC ATC NNK ATG GAT CGG CC | 68 |
| | A103-R | GG CCG ATC CAT MNN GAT GAA ATT GG | |
| S182 | S182-F | CTT ACC CGC NNK TCG GAT GTG AC | 69 |
| | S182-R | GT CAC ATC CGA MNN GCG GGT AAG | |
| D183 | D183-F | GC TGG TCG NNK GTG ACA ACC | 67 |
| | D183-R | GGT TGT CAC MNN CGA CCA GC | |
| T185 | T185-F | G GAT GTG NNK ACC GCA GC | 64 |
| | T185-R | GC TGC GGT MNN CAC ATC C | |
| T186 | T186-F | GAT GTG ACA NNK GCA GCA CC | 64 |
| | T186-R | GG TGC TGC MNN TGT CAC ATC | |

| | | | |
|------|------------------|--|----|
| L251 | L251-F L251-R | GTA CTT NNK CTG ATC GTT GGC G C GCC AAC GAT CAG MNN AAG TAC | 64 |
| V254 | V254-F V254-R | C CTG ATC NNK GGC GGG AAC GTT CCC GCC MNN GAT CAG G | 67 |
| G255 | G255-F G255-R | C CTG ATC GTT NNK GGG AAC G C GTT CCC MNN AAC GAT CAG G | 64 |
| D258 | D258-F D258-R | C GGG AAC NNK ACC ACA CG CG TGT GGT MNN GTT CCC G | 64 |
| T259 | T259-F T259-R | C GGG AAC GAT NNK ACA CGC GCG TGT MNN ATC GTT CCC G | 64 |
| L302 | L302-F L302-R | G CAA ACA CCG NNK GCT CAT ATG C G CAT ATG AGC MNN CGG TGT TTG C | 68 |
| M305 | M305-F M305-R | CTT GCT CAT NNK CGC CGC ACG CGT GCG GCG MNN ATG AGC AAG | 71 |
| F403 | F403-F F403-R | CGT TCA AAT NNK GTG CGC GG CC GCG CAC MNN ATT TGA ACG | 66 |

^a N represents any possible bases and K represents either guanine (G) or thymine (T)

^b Tm* is the value of optimal annealing temperatures calculated with NEB Tm Calculator <https://www.neb.com/tools-and-resources/interactive-tools/tm-calculator>

PCR products were subjected to DNA gel electrophoresis and the location of DNA in gel was detected using gel imaging system. Figure S2 showed an example of PCR products at 4 different sites, and the PCR product bands as well as some primer dimers can be seen clearly in this image. The product band was excised from the gel and then purified using QIAquick Gel Extraction Kit (Qiagen). The purified DNA was digested with methylated restrictive enzyme *Dpn I* at 37°C overnight.

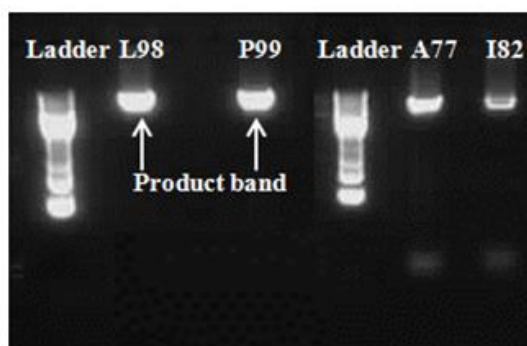


Figure S2. Image of gel electrophoresis after PCR amplification of site L98, P99, A77 and I82, respectively.

P450pyr mutant library was created by transferring the digested DNA into competent *E. coli* BL21(DE3) cells containing the pETDuet-Fdx-FdR vector. According to the ISM theoretical

calculations, a total of 94 mutant clones are required for each amino acid residue to have 95% coverage of all 20 possible amino acids.^{S3} Therefore, 188 clones for each site were taken to two 96-deep well plates in our experiment to ensure the coverage of > 95% for all possible 20 amino acids.

5. Calibration of purpald assay for determination of formaldehyde

A series of standard formaldehyde solutions was prepared, ranging from 0.05 mM to 1 mM. 100 μ L aliquot from each concentration were transferred to a 96-well plate. 50 μ L purpald solution containing 100 mM purpald and 1 M NaOH were added. After incubating at room temperature for 10 min in the presence of air, the reaction was fully completed and the absorbance of intense purple product in each well was measured at 550 nm using a spectrophotometer.

As shown in Fig. S3, the purpald assay produced excellent sensitivity and linearity toward formaldehyde. After plotting the absorbance value of the purple product versus the concentration of formaldehyde at each standard concentration, a calibration curve was set up. Using this calibration curve, we can quickly estimate the concentration of formaldehyde based on UV absorption.

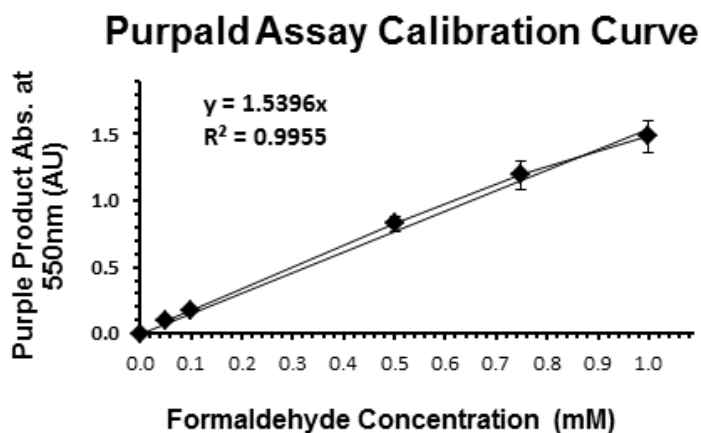


Figure S3. Calibration curve of purpald assay for determination of formaldehyde concentration.

6. General procedure of screening P450pyr mutants for terminal hydroxylation of alcohol by using surrogate substrate-based colorimetric HTS assay

Multiple single colony of *E. coli* (P450pyr mutant) generated in Section 4 were picked from LB agar plate and inoculated into each well of 96-deep well plate with 600 μ L TB medium containing 50 μ g/mL kanamycin and 100 μ g/mL ampicillin. The deep-well plate was shaken at 900 rpm and 37 $^{\circ}$ C for 8 h. 100 μ L cell cultures were taken out and mixed with 100 μ L 50% glycerol to prepare cell stock which was stored in -80 $^{\circ}$ C refrigerator. After that, 800 μ L TB medium containing 0.5 mM IPTG, 0.5 mM ALA, 50 μ g/mL kanamycin and 100 μ g/mL ampicillin were added to the remaining cell cultures in each well to induce the P450pyr protein expression. After expressing at 22 $^{\circ}$ C for 8 h, *E. coli* cells were harvested by centrifugation at 3220 g and room temperature for 15 min. 500 μ L 100 mM potassium phosphate buffer (pH 8.0) containing 2 % (w/v) D-glucose and 5 mM 2-methoxyethanol **3** were added to the cell pellets in each well. The biotransformation was performed at 900 rpm and 30 $^{\circ}$ C for 4 h and stopped by centrifugation at 3220 g for 10 min. 100 μ L aliquot from each well were transferred to a microliter plate that contained 50 μ L aqueous solution of 1 M NaOH and 100 mM purpald. After incubating at room temperature for 10 min, the absorbance of each well was measured at 550 nm using spectrophotometer (Thermo Scientific Multiskan Go).^{S4} Mutants with hydroxylation activity towards the surrogate substrate 2-methoxyethanol **3** will produce formaldehyde **6**, which react with purpald **7** to give purple color compound **8**. The principle of this surrogate substrate-based colorimetric HTS assay was shown in Figure S4A. The general procedure of this assay was shown in Figure S4B. The product with purple color can be directly checked by naked eye. A representative example was shown in Figure S5.

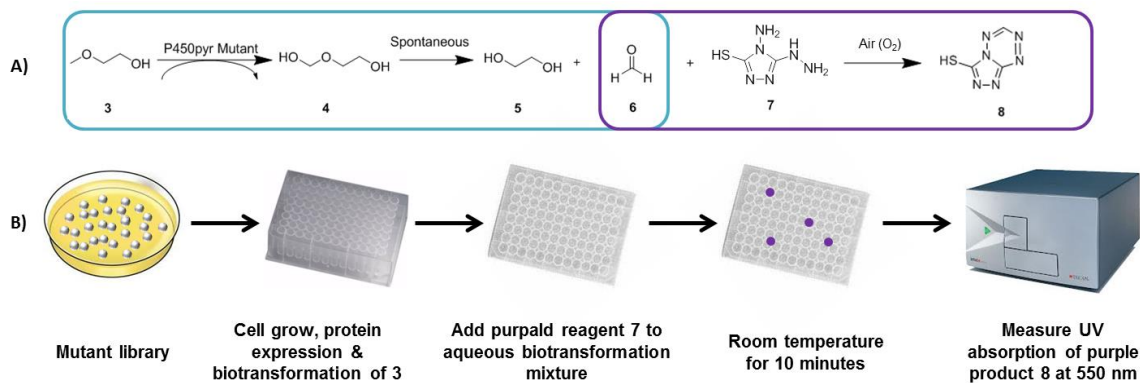


Figure S4. A) Principle of surrogate substrate-based colorimetric HTS assay B) Procedure of screening *E. coli* (P450pyr mutants) for the terminal hydroxylation of *n*-butanol **1** by using surrogate substrate-based colorimetric HTS assay.

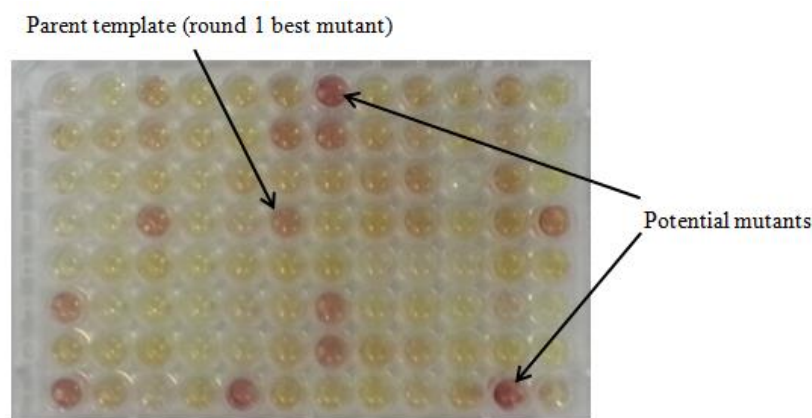


Figure S5. A representative example of 96-well plate after using high-throughput colorimetric assay. The stronger purple color indicating higher concentration of the purple product **8**.

7. General procedure for biohydroxylation of *n*-butanol with *E. coli* expressing positive P450pyr mutants in shaking flask

E. coli strains containing positive P450pyr mutants were inoculated into 50 mL TB containing 50 $\mu\text{g/mL}$ kanamycin and 100 $\mu\text{g/mL}$ ampicillin in multiple shaking flasks, respectively. The mixtures were shaken at 250 rpm and 37 $^{\circ}\text{C}$ until OD_{600} researched 0.8. The proteins expression

were induced by adding IPTG and ALA to a final concentration of 0.5 mM, respectively, and the mixture was shaken for additional 6 h at 250 rpm and 22 °C.

The cell pellets were collected by centrifugation at 3220 g and room temperature for 10 min. The cell were resuspended to a density of 24 g cdw L⁻¹ in 10 mL 100 mM potassium phosphate buffer (pH 8.0) containing 2 % (w/v) D-glucose and 5 mM of *n*-butanol **1**. Biotransformation was performed at 300 rpm and 30 °C for 4 h, and the formation of 1,4-butanediol **2** was analyzed by GC.

8. GC analysis of 1,4-butanediol

GC analysis was performed by using Agilent GC 6890 with Agilent HP-INNOWAX column (25 m × 0.32 mm) in splitless injection mode with inlet temperature of 280 °C and detector temperature of 220 °C. Temperature program: 60 °C for 3 min, increased to 180 °C at 30 °C/min, and further increased to 220 °C at 20 °C/min. Retention times: 6.86 min for 2-hexnaol **9** (internal standard), 8.58 min for 1,2-butanediol, 8.93 min for 1,3-butanediol, and 9.47 min for 1,4-butanediol **2**.

To analysis the biotransformation mixture, the supernatant was first concentrated at room temperature with Eppendorf Concentrator Plus. 32 % (w/w) ethanol and 16 % (w/w) ammonium sulfate were then added to the concentrated supernatant to form two-liquid phases to extract 1,4-butanediol **2** into ethanol phase.^{S5} 100 µL ethanol phase were taken and used for GC analysis.

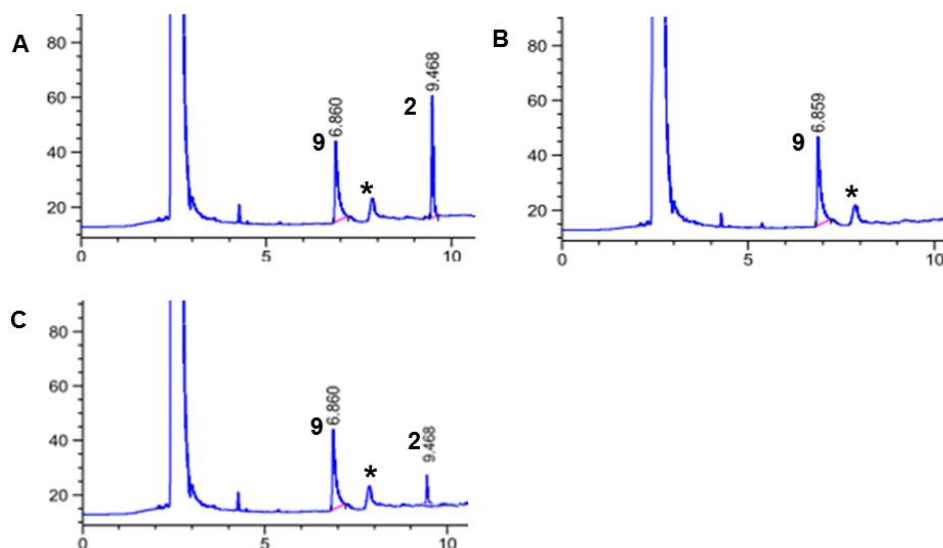


Figure S6. GC chromatograms of the biotransformation mixture from the hydroxylation of *n*-butanol **1** with *E. coli* strain expressing P450pyrI83M/I82T. (A) *E. coli* cell was induced with IPTG, (B) *E. coli* cell was not induced and (C) *E. coli* cell was not induced, and 0.05mM 1,4-butanediol was added to the biotransformation mixture. 2-hexanol **9** was used as internal standard; **2** was the bioproduct 1,4-butanediol; * unrelated cell metabolite.

9. MS analysis of the derivative of 1,4-butanediol made by BSTFA-TMCS method

To further confirm the identity of the biotransformation product is 1,4-butanediol, the mixture from biohydroxylation of *n*-butanol **1** was derivatised by BSTFA-TMCS method.^{S6} 1mL biotransformation supernatant was dried at 45 °C with Eppendorf Concentrator Plus, followed by the addition of 20 μ L DMF and 100 μ L BSTFA-TMCS reagent. The derivatisation reaction was conducted at 70 °C for 30 min. The derivatised mixtures were centrifuged at 20376 g for 5 min, and finally 60 μ L sample was used for GC-MS analysis. Similarly, the standard 1,4-butanediol **2** was derivatised by using the same procedure.

GC-MS analysis was performed using Agilent GC-MSD system 6890-5973A with GC-MSD HP-5 MS column in splitless injection mode with inlet temperature of 250 °C and the MS interface transfer line temperature of 270 °C. Temperature program: 80 °C for 1.5 min, increased to 140 °C at 10 °C/min, hold at 140 °C for 3 min, increased to 300 °C at 100 °C/min and hold at 300 °C for 5 min. Retention time for the derivatised **2** (1,4-butanediol bis(trimethylsilyl)ether) is 6.25 min. The MS of the peak at 6.25 min in GC chromatogram was shown in Fig. S7.

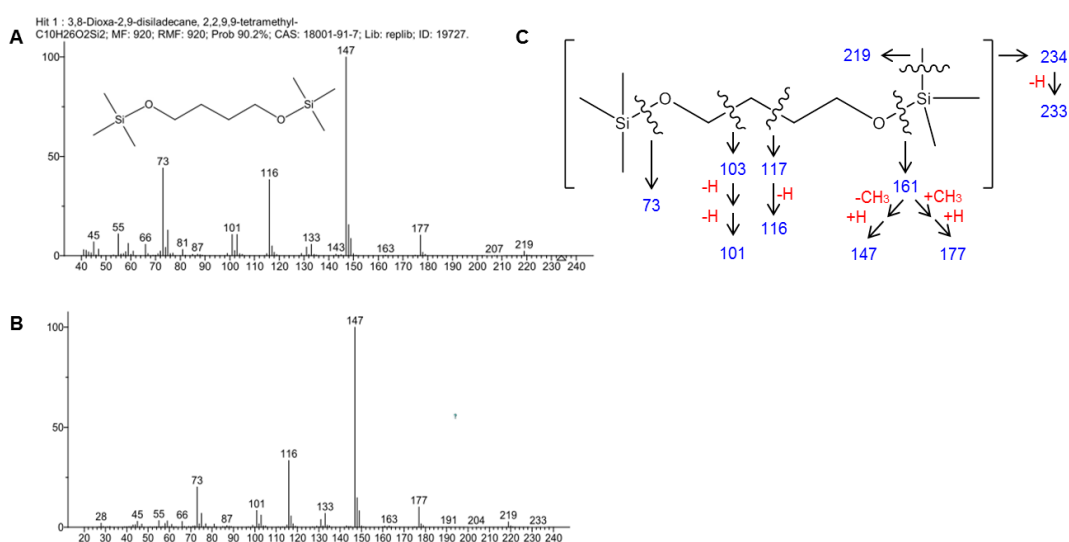


Figure S7. MS analysis of peak at 6.25 min in the GC chromatogram for BSTFA-TMCS derivative of 1,4-butanediol. (A) derivative from standard 1,4-butanediol; (B) derivative from biotransformation product with induced cell of *E. coli* (P450pyrI83M). (C) Proposed fragmentation of the derivatised 1,4-butanediol.

10. Molecular modelling of substrates docking in P450pyr and its mutants

The structure of P450pyr was obtained based on x-ray structure and enhanced using MODELLER^{S7} via CHIMERA^{S8} interface to supplant seven non-critical missing residues via the high precision Discrete Optimized Protein Energy score (DOPE-HR) algorithm.^{S9} Amino

acid residue mutations were performed *via* ACCELRY'S Discovery Studio.^{S10} Ramachandran plots for P450pyr, P450pyr I83M and P450pyrI83M/I82T models were generated before and after molecular dynamics (MD) simulations, and all revealed > 90 % of residues within allowed regions.

MD simulations of wild type and mutant enzymes were all conducted with GROMACS 4.5^{S11} using GROMOS 53a6 force field at a 0.2 ps recording interval in a water system with 150 mM NaCl. Long-range electrostatic interactions were modelled *via* Particle Mesh Ewald (PME) method, and short range van der Waals interactions were localized at 10 Å. Bond lengths were constrained *via* LINCS algorithm.^{S12} Energy minimization was performed with Steepest descent algorithm, in which V-rescale thermostat and Parrinello-Rahman barostat were used to stabilize temperature at 300 K and pressure at 1 atm, respectively. 5 ns production runs coupled to NPT were conducted, and the recorded enzyme conformers were clustered using Gromos algorithm with RMSD cutoff at 0.18 nm.^{S13} The main cluster centroids were docked with *n*-butanol using Autodock VINA,^{S14} where the catalytic heme was modified as the Cpd I ferryl-oxo-complex state.^{S15} Docking results were evaluated via binding energy scores with docked clustering at 1 Å non-fitted RMSD cutoff. Catalytically relevant active binding postures for terminal hydroxylation of *n*-butanol required the terminal substrate-C to be at closest approach to heme-O atom at a distance of < 6 Å, and also with Fe-O-H angles in the optimal range of 110° to 160° as part of established geometric criteria for substrate-heme binding in P450 enzymes.^{S2, S16}

Figures S8 showed the MD simulated enzyme models for P450yr, P450pyrI83M, and P450pyrI83M/I82T, respectively. They depict the spatial relation between the two mutated residues Ile83 and Ile82, and their nearby residues Ile102, Leu302, Thr259, Phe403, and Ser75. For P450pyr, the substrate *n*-butanol **1** was well docked in the binding pocket between the two residues Ile102 and Leu302. The substrate -OH group was at the closest approach to heme-O (2.0

Å). Such a posture was clearly inadequate for any hydroxylation. Therefore, P450pyr don't have hydroxylation activity for *n*-butanol.

As shown in Fig. 8B, the single mutation I83M reshaped the binding pocket *via* a "double domino effect": Domino effect 1: I83M introduced a longer residue Met83 with increased hydrophobic contact. Therefore, the neighbour residue Phe403 was reoriented slightly from the vicinity of Met83, with a distance of 6.7 Å for Met83-Phe403 in comparison with a 4 Å Ile83-Phe403 distance in P450pyr. This allowed more movement for nearby core residue Leu302, where the Leu302-Phe403 distance increased from 3.8 Å to 5.2 Å in P450pyrI83M. Domino effect 2: The increased hydrophobic contact of the longer Met83 residue influenced the neighbour residue Ile82, giving a 6.5 Å Met83-Ile82 distance which was bigger than the 5.4 Å Ile83-Ile82 distance in P450pyr. This shift affected nearby residue Ile102 so that it was slightly closer to Ile82 with a distance of 2.9 Å (compared to the 4.3 Å in P450pyr, Fig. S8A). Overall, the perturbations experienced by Phe403 and Ile82 made Ile102 and Leu302 more closely located, so that the Ile102-Leu302 closest distance reduced from 7.4 Å in P450pyr to 5.7 Å in P450pyrI83M. Thus the substrate changed its orientation, in which the -OH group oriented toward the hydrophilic residue Thr259 and substrate terminal-C closer to heme-O with a distance of 3.7 Å. This explains the occurrence of terminal hydroxylation activity. The distances of other C-atoms to the hem-O were 4.0 Å, 4.4 Å, and 5.0 Å, respectively (Fig. S8B). Therefore, the regioselectivity to terminal hydroxylation is higher.

For P450pyrI83M/I82T, I82T mutation gave a hydrophilic Thr82 residue which disrupted the hydrophobic interaction between Met83 and Ile102 (Figure S8C). This changed the nearby hydrophobic residue Ser75 to encroach upon the binding pocket. The structural integrity of the binding pocket surface at the Ile102-Phe403-Met83 was still maintained, with Leu302-Phe403 distance of 5.4 Å and Phe403-Met83 distance of 6.9 Å. For the binding pocket size between Ile102 and Leu302, their closest distance was expanded to 8.8 Å (Fig. S8C). The substrate docked

well in this binding pocket, with the -OH group well oriented toward residue Ser75 with a distance of 3.3 Å due to possible hydrophilic interaction. Here, the substrate terminal-C was the closest one to heme-O of distance 3.4 Å. This distance accounts for the terminal hydroxylation activity. Compared to 3.7 Å terminal-C to heme-O distance in P450pyrI83M, the shorter distance in P450pyrI83M/I82T would give a higher terminal hydroxylation activity. The distances of other C-atoms to the hem-O were: 4.8 Å, 5.4 Å and 6.8 Å, respectively (Fig. S8C). This explains the excellent regioselectivity for the terminal hydroxylation.

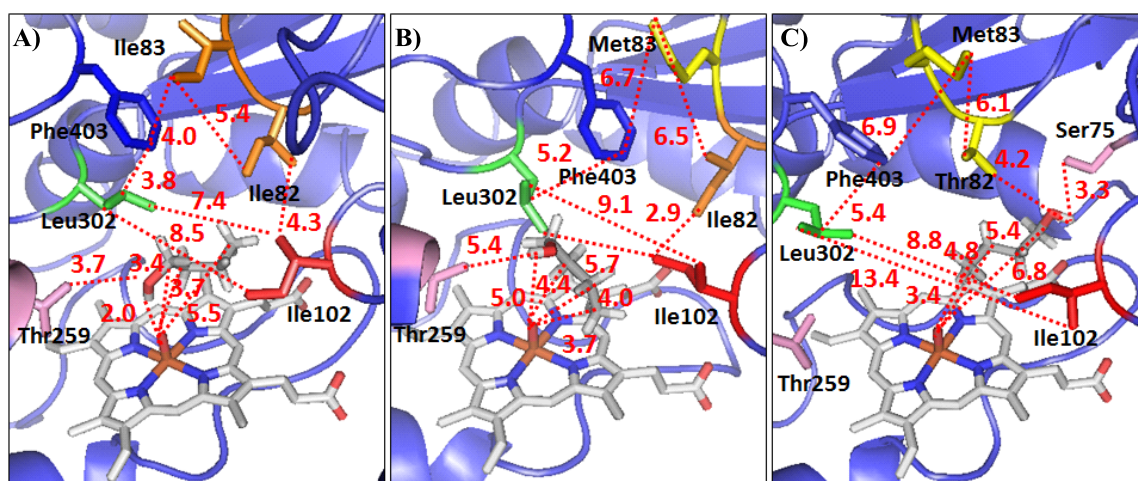


Figure S8. Positional relation of the key residues Ile102, Thr259, Leu302, Phe403, Ser75 and mutated residues (82, 83) in enzyme-*n*-butanol binding pose. (A) P450pyr (B) P450pyrI83M (C) P450pyrI83M/I82T. Mutated residues are shown in yellow. Distances (in angstrom) are denoted by red dashed lines.

11. References

- S1. S. Q. Pham, G. Pompidor, J. Liu, X. D. Li and Z. Li, *Chem. Commun.*, 2012, **48**, 4618.
- S2. S. Shaik, S. Cohen, Y. Wang, H. Chen, D. Kumar and W. Thiel, *Chem. Rev.*, 2010, **110**, 949.
- S3. M. T. Reetz and J. D. Carballeira, *Nat. Protoc.*, 2007, **2**, 891.

- S4. G. Avigad, *Anal. Biochem.*, 1983, **134**, 499.
- S5. Z. Li, H. Teng and Z. Xiu, *Process. Biochem.*, 2010, **45**, 731.
- S6. H. Yim, R. Haselbeck, W. Niu, C. Pujol-Baxley, A. Burgard, J. Boldt, J. Khandurina, J. D. Trawick, R. E. Osterhout, R. Stephen, J. Estadilla, S. Teisan, H. B. Schreyer, S. Andrae, T. H. Yang, S. Y. Lee, M. J. Burk and S. Van Dien, *Nat. Chem. Biol.*, 2011, **7**, 445.
- S7. A. Sali and T. Blundell, *J. Mol. Biol.* 1993, **5**, 234.
- S8. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605.
- S9. M.-Y. Shen and A. Sali, *Prot. Sci.*, 2006, **15**, 2507.
- S10. Accelrys Software Inc., *Discovery Studio Visualizer, Release 3.5*, San Diego: Accelrys Software Inc., 2012
- S11. H. J. C. Berendsen, D. Vandespoel and R. Vandrunen, *Comput. Phys. Commun.*, 1995, **91**, 43.
- S12. B. Hess, H. Bekker, H. J. C. Berendsen and J. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463.
- S13. X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren and A. E. Mark, *Angew. Chem. Int. Edit.*, 1999, **38**, 236.
- S14. (a) G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.*, 2009, **30**, 2785; (b) O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455.
- S15. I. G. Denisov, T. M. Makris, S. G. Sligar and I. Schlichting, *Chem. Rev.*, 2005, **105**, 2253.
- S16. Y. Zhang, P. Morisetti, J. Kim, L. Smith and H. Lin, *Theor. Chem. Acc.*, 2008, **121**, 313.