Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

Engineered alkane-hydroxylating cytochrome $P450_{BM3}$ exhibiting native-like catalytic properties

Rudi Fasan, Mike M. Chen, Nathan C. Crook, Frances H. Arnold*

Department of Chemistry and Chemical Engineering, California Institute of Technology, 91125 Pasadena (CA), U.S.A.

Experimental Section

Plasmids and oligonucleotides. Plasmid pCWori[1] was used as cloning vector. Oligonucleotides were purchased from Invitrogen (Carlsbad, CA). The sequence of the primers used in this study are reported below:

#	Primer name	Sequence
1	L52I_for	5'-CGC GCT ACA TAT CAA GTC AGC-3'
2	L52I_rev	5'-GCT GAC TTG ATA TGT AGC GCG-3'
3	M145A_for	5'-GTA TCG GAA GAC GCG ACA CGT TTA ACG-3'
4	M145A_rev	5'-GTA TCG GAA GAC GCG ACA CGT TTA ACG-3'
5	V340M_for	5'-GAA GAT ACG ATG CTT GGA GGA G-3'
6	V340M_rev	5'-CTC CTC CAA GCA TCG TAT CTT C-3'
7	I366V_for	5'-CGT GAT AAA ACA GTT TGG GGA GAC G-3'
8	I366V_rev	5'-CGT CTC CCC AAA CTG TTT TAT CAC G-3'
9	E442K_for	5'-CGT TAA AAC CTA AAG GCT TTG TGG-3'
10	E442K_rev	5'-CCA CAA AGC CTT TAG GTT TTA ACG-3'
11	L324I_for	5'-CGA AGC GCT GCG CAT CTG GCC AAC TT-3'
12	L324I_rev	5'-AAG TTG GCC AGA TGC GCA GCG CTT CG-3'
13	S106R_for	5'-CTTACTTCCAAGGTTCAGTCAGCAGG-3'
14	S106R_rev	5'-CCT GCT GAC TGA ACC TTG GAA GTA AG-3'
15	BamHI_fwd	5'-CAC AGG AAA CAG GAT CCA TCG ATG CTT AGG-3'
16	SacI_rev	5'-CTA GGT GAA GGA ATA CCG CCA AGC GGA-3'
17	L437NNK_for	5'-CGA TAT TAA AGA AAC TNN KAC GTT AAA ACC-3'
18	L437NNK_rev	5'-GGT TTT AAC GTM NNA GTT TCT TTA ATA TCG-3'
19	T438NNK_for	5'-CGA TAT TAA AGA AAC TTT ANN KTT AAA ACC-3'
20	T438NNK_rev	5'-GGT TTT AAM NNT AAA GTT TCT TTA ATA TCG-3'
21	EcoRI_Rev	5'-CCG GGC TCA GAT CTG CTC ATG TTT GAC AGC-3'
22	L181NNK_for	5'-GGT CCG TGC ANN KGA TGA AGT AAT G-3'
23	L181NNK_rev	5'-CAT TAC TTC ATC MNN TGC ACG GAC C-3'
24	A82NNK_for	5'-CGT GAT TTT NNK GGA GAC GGG TTA-3'
25	A82NNK_rev	5'-TAA CCC GTC TCC MNN AAA ATC ACG-3'
26	A74NNK_for	5'-AAC TTA AGT CAA NNK CTT AAA TTC-3'
27	A74NNK_rev	5'-GAA TTT AAG MNN TTG ACT TAA GTT-3'
28	L75NNK_for	5'-GTC AAG CGN NK AAA TTC TTT CGT G-3'

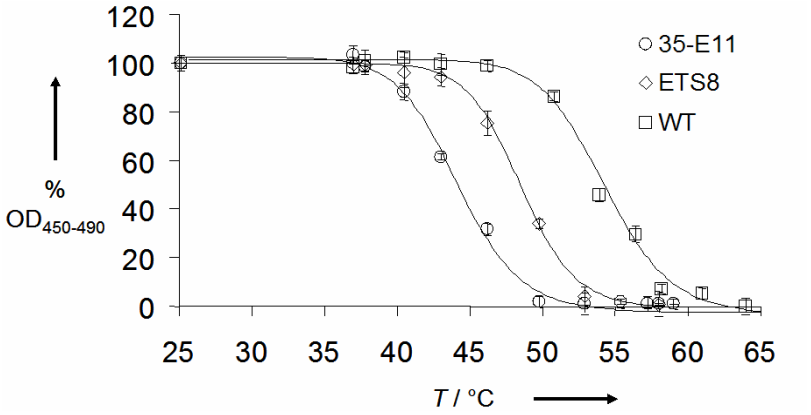
29	L75NNK_rev	5’-CAC GAA AGA ATT TMN NCG CTT GAC-3’
30	V78NNK_for	5’-GTC AAG CGC TTA AAT TCN NKC GTG ATT TT-3’
31	V78NNK_rev	5’-AAA ATC ACG MNN GAA TTT AAG CGC TTG AC-3’
32	A328NNK_for	5’-GGC CAA CTN NKC CTG CGT TTT CC-3’
33	A328NNK_rev	5’-GGA AAA CGC AGG MNN AGT TGG CC-3’
34	A184NNK_for	5’-GCA CTG GAT GAA NNK ATG AAC AAG-3’
35	A184NNK_rev	5’-CTT GTT CAT MNN TTC ATC CAG TGC-3’
36	L188NNK_for	5’-GAA CAA GNN KCA GCG AGC AAA TCC-3’
37	L188NNK_rev	5’-GGA TTT GCT CGC TGM NNC TTG TTC-3’
38	I401NNKfwd	5’-GCG TGC GTG TNN KGG TCA GCA G-3’
39	I401NNKrev	5’-CTG CTG ACC MNN ACA CGC ACG C-3’
40	T268NNKfwd	5’-GCG GGA CAC GAA NNK ACA AGT GGT C-3’
41	T268NNKrev	5’-GAC CAC TTG TMN NTT CGT GTC CCG C-3’
42	G265NNKfwd	5’-CAT TCT TAA TTG CGN NKC ACG AAA CAA CAA GTG-3’
43	G265NNKrev	5’-CAC TTG TTG TTT CGT GMN NCG CAA TTA AGA ATG-3’
44	A264NNKfwd	5’-CA TTC TTA ATT NNK GGA CAC GAA ACA ACA AGT G-3’
45	A264NNKrev	5’-CAC TTG TTG TTT CGT GTC CMN NAA TTA AGA ATG-3’
46	T260NNKfwd	5’-CAA ATT ATT NNK TTC TTA ATT GCG GGA C-3’
47	T260NNKrev	5’-GTC CCG CAA T TA AGA AMN NAA TAA TTT G-3’
48	L75NNKfwd	5’-GTC AAG CGN NKA AAT TTG TAC G-3’
49	L75NNKrev	5’-GTC CCG CAA TTA AGA AMN NAA TAA TTT G-3’
50	T88NNKfwd	5’-GAC GGG TTA TTT NNK AGC TGG ACG CAT G-3’
51	T88NNKrev	5’-GTC CCG CAA TTA AGA AMN NAA TAA TTT G-3’
52	F87NNKfwd	5’-GAC GGG TTA NNK ACA AGC TGG-3’
53	F87NNKrev	5’-CCA GCT TGT MNN TAA CCC GTC-3’
54	A82G_for	5’-CGT GAT TTT GGT GGA GAC GGG TTA-3’
55	A82G_rev	5’-TAA CCC GTC TCC ACC AAA ATC ACG-3’
56	FMN_for	5’-GCT GGT ACT TGG TAT GAT GCT-3’
57	FMN_rev	5’-CCA GAC GGA TTT GCT GTG AT-3’
58	FAD_for	5’-CGT GTA ACA GCA AGG TTC GG-3’
59	FAD_rev	5’-CTG CTC ATG TTT GAC AGC TTA TC-3’
60	G443NNK_for	5’-CGT TAA AAC CTG AAN NKT TTG TGG-3’
61	G443NNK_rev	5’-CCA CAA AMN NTT CAG GTT TTA ACG-3’
62	V445NNK_for	5’-CCT GAA GGC TTT NNK GTA AAA GCA -3’
63	V445NNK_rev	5’-TGC TTT TAC MNN AAA GCC TTC AGG-3’
64	T480NNK_for	5’-CGC TCA TAA TNN K CCG CTG CTT GTG-3’
65	T480NNK_rev	5’-CAC AAG CAG CGG MNN ATT ATG AGC G-3’
66	T515NNK_for	5’-CCG CAG GTC GCA NNK CTT GAT TCA C-3’
67	T515NNK_rev	5’-GTG AAT CAA GMN NTG CGA CCT GCG G-3’
68	P654NNK_for	5’-GCG GAT ATG NNK CTT GCG AAA ATG-3’
69	P654NNK_rev	5’-CAT TTT CGC AAG MNN CAT ATC CGC-3’
70	T664NNK_for	5’-GGT GCG TTT TCA NNK AAC GTC GTA GCA-3’
71	T664NNK_rev	5’-TGC TAC GAC GTT MNN TGA AAA CGC ACC-3’
72	D698NNK_for	5’-CAA GAA GGA NNK CAT TTA GGT G-3’
73	D698NNK_rev	5’-CAC CTA AAT GMN NTC CTT CTT G-3’
74	E1037NNK_for	5’-CAG CAG CTA GAA NNK AAA GGC CG -3’
75	E1037NNK_rev	5’-CGG CCT TTM NNT TCT AGC TGC TG-3’
76	BMfor_1504	5’-GCA GAT ATT GCA ATG AGC AAA GG-3’

77	BMrev1504	5’-CCT TTG CTC ATT GCA ATA TCT GC-3’
78	BMfor2315	5’-CGG TCT GCC CGC CGC ATA AAG-3’
79	BMrev2315	5’-CTT TAT GCG GCG GGC AGA CCG-3’

Enzyme library construction. HL1 library was prepared by introducing L52I, S106R, M145A, L324I, V340M, I366V, E442K substitutions into 35E11 singly and in combination by PCR overlap extension mutagenesis and SOEing. The previously described pCWori_35E11[2] served as template, BamHI_fwd and SacI_rev as megaprimers, and sequences 1 to 14 as mutagenizing primers. The amplified region (1.5 Kbp) was cloned in pCWori_35E11 using *BamH* I and *Sac* I restriction enzymes. Heme domain library HL2 was created by random mutagenesis of ETS8 heme domain residues 1-433 (*Taq* polymerase (Roche, Indianapolis, IN), BamHI_fwd and SacI_rev primers, and $MnCl_2$ from 50 to 300 μM). HL3 active-site libraries were prepared by site-saturation mutagenesis at positions 74, 75, 78, 82, 87, 88, 181, 184, 188, 260, 264, 265, 268, 328, 401, 437 and 438 using the 19A12 sequence as template. 437NNK and 437NNK sub-libraries were prepared using primers 17-20 and *BamH* I/*EcoR* I restriction enzymes (3.2 Kbp fragment). The remaining NNK libraries were created using primers 22-53 and *BamH* I/*Sac* I restriction enzymes. Heme recombination library HL4 was created by recombining mutations A74S, A74Q, V184S, V184T, V184A. HL5 heme saturation/recombination libraries were prepared by recombining mutations 74S/NNK, 82S/G, and 184S/T/NNK. RL1 and RL2 reductase domain libraries were prepared by random mutagenesis of 35E11 FMN domain (432-720) and FAD domain regions (724-1048). Error-prone PCR was carried out using *Taq* polymerase, pCWori_35E11 as template, FMN_for/FMN_rev and FAD_for/FAD_rev primer pairs, and $MnCl_2$ from 100 to 300 μM. The amplified regions were cloned in pCWori_35E11 using *Sac* I, *Nsi* I, and *EcoR* I restriction enzymes. RL3 and RL4 reductase domain libraries were created by site-saturation mutagenesis at positions 443, 445, 515, 580, 654, 664, 698, 1037 using 11-3 sequence as template, FMN_for and EcoRI_rev megaprimers, and primers 60 to 75. The amplified region (1.9 Kbp) was cloned in pCWori_11-3 using *Sac* I and *EcoR* I restriction enzymes. Library L9 was created by recombining G443A, P654K, T664G, D698G, and E1037G mutations and fusing the resulting FMN/FAD library to the 7-7 heme domain with *Sac* I and *EcoR* I restriction enzymes.

HTS and enzyme purification. High-throughput screening on dimethyl ether (DME) was carried out on 96-well plates as described.[1] Formaldehyde produced in the reaction was determined by addition of Purpald (168 mM in 2 M NaOH) and by monitoring absorbance at 550 nm. $P450_{BM3}$ variants were expressed and purified as described.[1] P450 concentration in the purified sample was measured in triplicate from the CO-binding difference spectra. Protein samples were aliquoted and stored at -80°C.

T_{50} determination. Samples of purified enzyme (~3 μM) were incubated for 10 minutes at different temperatures (from 20°C to 60°C) in a PCR thermocycler. After centrifugation, 160 μL protein solution were mixed with 40 μL 0.1 M sodium hydrosulfite on a microtiter plate. T_{50} values were calculated from heat-inactivation curves of CO-binding difference spectra (see below). Experiments were carried out in triplicate.



Oxidation rates, coupling efficiency, and total turnovers. Initial rates of propanol formation were determined at 25°C in 100 mM KPi pH 8.0 buffer (5 mL) using purified enzyme (100-400 nM) and propane-saturated buffer. Reactions were initiated by addition of 500 μM NADPH and stopped after 20 seconds by addition of 200 μL H_2SO_4 conc. Samples were analyzed by GC-ECD as described.[1] Initial rates of NADPH oxidation were measured by monitoring the decrease in OD_{340}. Rates were determined in triplicate and calculated over the first 20 seconds using extinction coefficient ε_{340} = 6210 M^{-1} cm^{-1}. Coupling values were calculated from the ratio of propanol formation rate to NADPH oxidation rate in the presence of propane. TTN reactions were carried out at 1 mL-scale using propane-saturated KPi buffer containing 20 nM purified protein and a cofactor regeneration system (10 U/mL isocitrate dehydrogenase, 7 mg/mL isocitrate, 0.15 mg/mL $NADP^+$). Vials were sealed and stirred for 24 hours at 4°C. Products were analyzed by GC using a Hewlett-Packard 5890 Series II Plus gas chromatograph**,** Supelco SPB-1 column (60 m x 0.52 mm x 0.5 μm film), 0.5 μL injection, FID detector, and the following separation program: 250 °C inlet, 275 °C detector, 80 °C oven for 2 min, 10°C/min gradient to 110°C, 25°C/min gradient to 275 °C, and 275 °C for 2.5 min. All measurements were performed at least in triplicate.

Whole-cell biotransformations. Biotransformations were carried out at 80 mL- and 300 mL-scale using temperature-controlled 100 mL- (Ochs-labor) and 1 L-fermenter (DasGip, 4x twin), respectively. Cells were resuspended in nitrogen-free modified M9 medium supplemented with 1% (w/v) glucose. A gas mixture of propane + air (or pure oxygen) was fed to the cells at an inlet gas flow rate of 5 and 10 L h^{-1}, respectively. In the 1 L-fermenter, pO_2 and pH were maintained at 100% and 7.2, respectively. The total amount of propanol produced in the reactions was determined summing the concentration of alcohol in the reactor and that in a bubbler connected to the reactor's gas outlet. At defined time intervals, 1 mL of cell suspension was removed, centrifuged, filtered, and subjected to GC analysis. P450 concentration in the cells was determined from CO-binding difference spectra on cell lysates obtained by sonication. Prior to bioconversions, freshly transformed cells were grown in modified M9 medium supplemented with 0.4% glucose and 1.5% yeast extract. At OD_{600} = 1.2, cells were induced with 0.25 mM IPTG and 0.25 mM δ-aminolevulininc acid and harvested after 10-12 hours. In addition to standard salts, modified M9 contained nutrients (calcium pantothenate, *p*-aminobenzoic acid, *p*-hydroxybenzoic acid, thiamine) and metals ($CoCl_2$, $CuSO_4$, $MnCl_2$, $ZnSO_4$) in trace amounts.

References

[1] M. W. Peters, P. Meinhold, A. Glieder, F. H. Arnold, *J. Am. Chem. Soc.* 2003, *125*, 13442-13450.

[2] P. Meinhold, M. W. Peters, M. M. Chen, K. Takahashi, F. H. Arnold, *ChemBiochem* 2005, *6*, 1765-1768.

[3] M. Wang, D. L. Roberts, R. Paschke, T. M. Shea, B. S. Masters, J. J. Kim, *Proc. Natl. Acad. Sci. U.S.A.* 1997, *94*, 8411-8416.

[4] A. J. Warman, O. Roitel, R. Neeli, H. M. Girvan, H. E. Seward, S. A. Murray, K. J. McLean, M. G. Joyce, H. Toogood, R. A. Holt, D. Leys, N. S. Scrutton, A. W. Munro, *Biochem. Soc. Trans.* 2005, *33*, 747-753.

[5] I. F. Sevrioukova, H. Li, H. Zhang, J. A. Peterson, T. L. Poulos, *Proc. Natl. Acad. Sci. U.S.A.* ***1999***, *96*, 1863-1868.

Table S1. Thermostabilization of 35E11. L52I, S106R, M145A, L324I, V340M, I366V, E442K mutations were introduced in variants ETS1 to ETS7. S106R and M145A mutations resulted in unfolded proteins (ETS2 and ETS7, respectively) and were not considered for further recombination.

Variant	L52I	L324I	V340M	I366V	E442K	T_{50} (°C)	ΔT_{50} (°C)
35E11						43.4 ± 0.6	**0**
ETS1	X					44.5 ± 0.3	**1.1**
ETS3		X				43.2 ± 0.1	**-0.3**
ETS4			X			46.0 ± 0.1	**2.6**
ETS5				X		47.1 ± 0.1	**3.7**
ETS6					X	45.0 ± 1.4	**1.6**
ETS8	X			X		48.5 ± 0.2	**5.1**
ETS9	X				X	46.8 ± 0.6	**3.4**
ETS10			X		X	44.2 ± 0.1	**0.8**
ETS11				X	X	46.6 ± 0.2	**3.2**
ETS12	X			X	X	misfolded	-
ETS13			X	X		misfolded	-
ETS14	X		X	X		46.1 ± 0.6	**2.7**
ETS15	X		X	X	X	47.8 ± 1.3	**4.4**
ETS16	X	X		X		45.9 ± 0.1	**2.5**
ETS17	X	X		X	X	47.2 ± 0.2	**3.8**
ETS18			X	X	X	45.5 ± 0.2	**2.1**
ETS19		X	X	X	X	45.6 ± 0.3	**2.2**
ETS20		X	X	X		44.6 ± 0.2	**1.2**
ETS21	X	X	X	X		45.8 ± 0.1	**2.4**
ETS22	X	X	X	X	X	47.0 ± 0.3	**3.6**

Figure S1. Map of the activity-enhancing reductase domain mutations on a homology model of $P450_{BM3}$ FAD-binding domain prepared on the basis of the rat cytochrome P450 reductase structure (PDB: 1AMO[3]). Structural similarity between the two is supported by a preview of the solved but not yet published structure of $P450_{BM3}$ FAD-binding domain.[4] Heme domain and FMN domain are represented as in PDB: 1BVY.[5] A 30-residue linker connects the C-terminus of the FMN-binding with the N-terminus of the FAD-binding domain (dotted line).