Electronic Supplementary Material (ESI) for Catalysis Science & Technology.

Electronic Supplementary Information

Modulating stereoselectivity and catalytic efficiency of carbenoid reactions catalysed by Self-sufficient P450s

Binhao Wang, Cuiping You, Guochao Xu,* and Ye Ni,*

Key laboratory of industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi214122, Jiangsu, China. Phone number: (+86) 0510-85918206.

*Corresponding authors: yni@jiangnan.edu.cn (Y Ni); guochaoxu@jiangnan.edu.cn (Guochao Xu).

1. Materials and Methods

1.1 Materials

The racemic standard of ethyl 2-phenylcyclopropane carboxylate and styrene were obtained from MACLIN. Ethyl diazoacetate was provided by Damas-beta. Other required reagents, such as ethanol, acetonitrile, N-hexane, etc., were purchased from Sigma Aldrich.

1.2 Strain and amino acid sequence

E. coli BL21 (DE3) was employed for the expression of recombinant proteins, six self-sufficient P450s include: $P450_{TT}$ (Genbank number: WP_055423153.1), $P450_{JT}$ (Genbank number: WP_104007096.1), CYP505X (Genbank number: XP_754698), CYP152K6 (GenBank number: 6FYJ_A), CviUPO (GenBank number: 7ZCL_B), and CYP505A30 (GenBank number: XP_003663647.1). Above nucleotide sequences were synthesized by Tianlin (Wuxi) Technology Co., LTD. and subsequently cloned into pET 28a (+) vector. For the positive control experiments, the nucleotide sequences of $P450_{BM3}$, myoglobin, and cytochrome C were synthesized according to the references.[1-3]

1.3 Medium

Luria-Bertani (LB) medium ($g \cdot L^{-1}$): peptone 10.0, yeast extract 5.0, sodium chloride 10.0, pH 7.2, solid medium with 2% AGAR.

Super Broth (TB) medium ($g \cdot L^{-1}$): yeast extract 24.0, peptone 12.0, K2HPO3 16.4, KH2PO3 2.3, glycerol 5.0.

1.4 Mutagenesis of substitution of the key axial residues of P450s

Substitution of key axial residues in P450s was performed through whole plasmid PCR mutagenesis, using six P450 genes as templates. The primers required for introducing the critical residue changes were exclusively synthesized by Sangon Biotech (Shanghai) Co., LTD. **(Table S1)**. PCR amplification system contained complementary primers and used KOD-Plus-Neo DNA polymerase to amplification sequence. PCR program involved an initial pre-denaturation at 94°C for 4 min, followed by amplification for 25 cycles with denaturation at 98°C for 10 s, annealing at 55°C for 15 s and elongation at 72°C for 5 min, and then further elongation at 72°C for 10 min. The plasmids verified by nucleic acid glue and digested with 1 μL of DpnI for 0.5 h at 37°C to remove the methylated templates. Then 10 μL of the digestion mixture were transformed into *E. coli* BL21 (DE3). Single

colonies were picked up and cultivated in LB medium (50 μg/mL Kan). The plasmids of mutations were confirmed by sequencing.

Table S1. List of primers

Primer	Sequence
$P450_{TT}$-C385H-1	GCCGAGGTGCTGATGGGCGCCGTAACC
$P450_{TT}$ -C385H-2	CATCAGCACCTCGGCCGTAACCTCGG
$P450_{TT}$ -C385S-1	CATCAGAGCCTCGGCCGTAACCTCGG
$P450_{TT}$ -C385S-2	GCCGAGGCTCTGATGGGCGCCGTAACC
$P450_{JT}$-C377S-1	CATCAGAGCATGGGCAAAAACATCGGC
$P450_{JT}$ -C377S-2	GCCCATGCTCTGATGGGCACCATAGCCG
$P450_{JT}$ -C377H-1	CATCAGCACATGGGCAAAAACATCGGC
$P450_{JT}$ -C377H-2	GCCCATGTGCTGATGGGCACCATAGCC
CYP505A30-C411S-1	CGTGCGAGCATCGGGCGTCCTTTCGCG
CYP505A30-C411S-2	CCCGATGCTCGCACGCATACCAGTACCG
CYP505A30-C411H-1	CGTGCGCACATCGGGCGTCCTTTCGCG
CYP505A30-C411H-2	CCCGATGTGCGCACGCATACCAGTACC
CYP505X-C414H-1	CGCGCACACATTGGTCGTCCGTTTGCG
CYP505X-C414H-2	ACCAATGTGTGCGCGCATACCGTTGCC
CYP505X-C414S-1	CGCGCAAGCATTGGTCGTCCGTTTGCG
CYP505X-C414S-2	ACCAATGCTTGCGCGCATACCGTTGCC
CYP152K6-C365S-1	CACCGCAGCGCTGGCGAATGGGTTACC
CYP152K6-C365S-2	GCCAGCGCTGCGGTGGCCCATGAAGTA
CYP152K6-C365H-1	CACCGCCACGCTGGCGAATGGGTTACC
CYP152K6-C365H-2	GCCAGCGTGGCGGTGGCCCATGAAGTA
CviUPO-C20S-1	GCTCCGAGCCCGGCTATGAACAGCCTG
CviUPO-C20S-2	AGCCGGGCTCGGAGCGCGAAATTCACCC
CviUPO-C20H-1	GCTCCGCACCCGGCTATGAACAGCCTG
CviUPO-C20H-2	AGCCGGGTGCGGAGCGCGAAATTCACC

1.5 Protein expression and purification

Recombinant *E. coli* BL21 (DE3) including six P450s and variants were cultivated at 37°C and 180 rpm in LB medium supplemented with 1% glucose. The cultures were grown until OD_{600} reached 0.6~0.8. Subsequently, they were induced with 0.2 mM isopropyl β-D-thiogalactopyranoside (IPTG) for gene expression. 2 mM $FeSO_4$ and 1 mM 5-aminolevulinic acid hydrochloride (ALA) were added for synthesis of iron porphyrin in the catalytic center of P450s. And further culture was conducted at 16 °C, 120 rpm for 24 h. After culture, the culture medium was removed by centrifugation at 4000 rpm for 10 min. The induced cells were harvested by centrifugation (4°C, 8,000×g for 10 min) to be freeze-dried. When protein purification was required, the cells were resuspended in phosphate buffer A (25 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole, 5 mM β-mercaptoethanol, pH 7.4), then disrupted with ultra-sonication (work for 2 s, pause for 3 s, 150 W). The resulting supernatant was loaded onto a His-Trap HP nickel affinity column pre-equilibrated with buffer A using AKTA Avant System (GE Healthcare, USA), P450s were eluted off by a 20-300 mM imidazole gradient in buffer B (25 mM Tris-HCl, 300 mM NaCl, 300 mM imidazole, 5 mM β-mercaptoethanol, pH 7.4). The collected proteins were verified by SDS-PAGE analysis and then concentrated by an ultrafiltration tube. The target pure enzyme is subjected to the next experiment.

1.6 Small-scale of carbenoid model reactions

The reactions were initiated by adding 10 μL of styrene (1.2 M stock solution in methanol), followed by the addition of 10 μL of EDA (0.4 M stock solution in methanol) using a syringe under anaerobic conditions. The reaction mixture was stirred for 12 hours at 16℃. Upon completion of the reaction, it was terminated by adding an equal volume of 3 M HCl. The resulting solution was then centrifuged at 12,000 rpm for 10 min to separate the phases. After centrifugation for 5 min, the solution was filtered using a 0.22 μm organic filter membrane. Subsequently, the analysis was performed using an ACQUITY UPC2 system by Waters Corporation, with a Trefoil CEL2 Supercritical Column (3.0 × 150 mm, 2.5 μm) employed. The mobile phase A consisted of a mixture of ethanol and acetonitrile. UV detection was carried out at a wavelength of 210 nm, with the column temperature maintained at 30 °C. The flow rate was set at 1.2 mL/min, and an injection volume of 3 μL was used. It was recommended to conduct at least two parallel experiments for each group of reactions. The enzymatic cyclopropane reaction system consists of crude enzymes (**Table S2**) and purified enzymes (**Table S3**).

Table S2. Reaction system for crude enzymes

Ingredients	Volume (μ L)
Crude enzyme	3444
Styrene（1.2 M）	33
EDA（400 mM）	50
NaS_2O_4	33

Table S3. Reaction system for purified enzymes

Ingredients	Volume（ μ L）
Purified Enzyme	10 μM
Styrene（1.2 M）	33
EDA（400 mM）	50
NaS_2O_4	33
Tris·HCl buffer（100 mM pH=8.0）	Filled up to 2000

1.7 Phylogenetic tree analysis, Sequence analysis, Moleclular modeling and Simulation

Phylogenetic tree was constructed using is the maximum likelihood (ML) approach, and the software used is RAxML 8.2.12 and PhyloSuite.[4, 5] Using $P450_{BM3}$, myoglobin, and six self-sufficient P450s as templates, homologous sequences was searched on the NCBI website (https://www.ncbi.nlm.nih.gov/), incomplete and redundant sequences were removed, and the sequences based on sequence consistency were rename. Sequence alignment was performed using the multiple sequence alignment software MAFFT,[6] with specific parameters: FFT-NS-2 (default) strategy, Gap opening penalty, default: 1.53, Offset (works like gap extension penalty), default: 0.0. Before the tree constructed, it is essential to determine evolutionary model for the sequence using Model Finder.[7] Bayesian Information Criterion (BIC) was employed to identify the best-fitting substitution model among various evolutionary models: LG+G4. Finally, online phylogenetic tree editing tool ITOL v5 (https://itol.embl.de/) was employed to beautify the tree.[8]

Sequence alignments and conserved analysis were conducted using online tool Clustal

(https://www.genome.jp/tools-bin/clustalw) and NCBI service (https://www.ncbi.nlm.nih.gov/). To further enhance the obtained results, online website ENDscript/ESPrip was utilized for beautification.[9]

With the exception of certain protein structures like $P450_{TT}$ and $P450_{BM3}$, all molecular models were generated using AlphaFold 2. The analysis of interactions around the heme group was performed through Discovery Studio software. The enzyme-substrate complexes were further optimized using the CHARM force field in the Discovery Studio software package. The structures were visualized and generated using Pymol.

2 Standard curve of product (EPC) using ultra-performance convergence chromatography

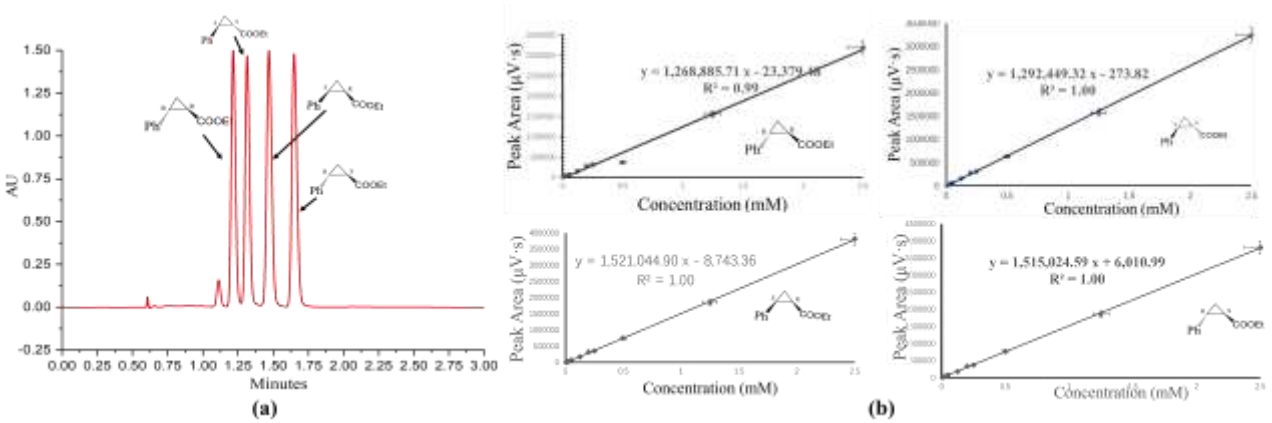


Figure S1. Analysis of ethyl 2-phenyl cyclopropane methane carboxylate (EPC) using ultra-performance convergence chromatography. (a) The ultra-performance convergence chromatography results for different configurations of EPC were as follows: (R, R)-EPC eluted at 1.216 min, (S, S)-EPC eluted at 1.315 min, (S, R)-EPC eluted at 1.470 min, and (R, S)-EPC eluted at 1.649 min. (b) Standard curve of different configurations of EPC analyzed by ultra-performance convergence chromatography.

4. Expression and purification of recombinant P450s

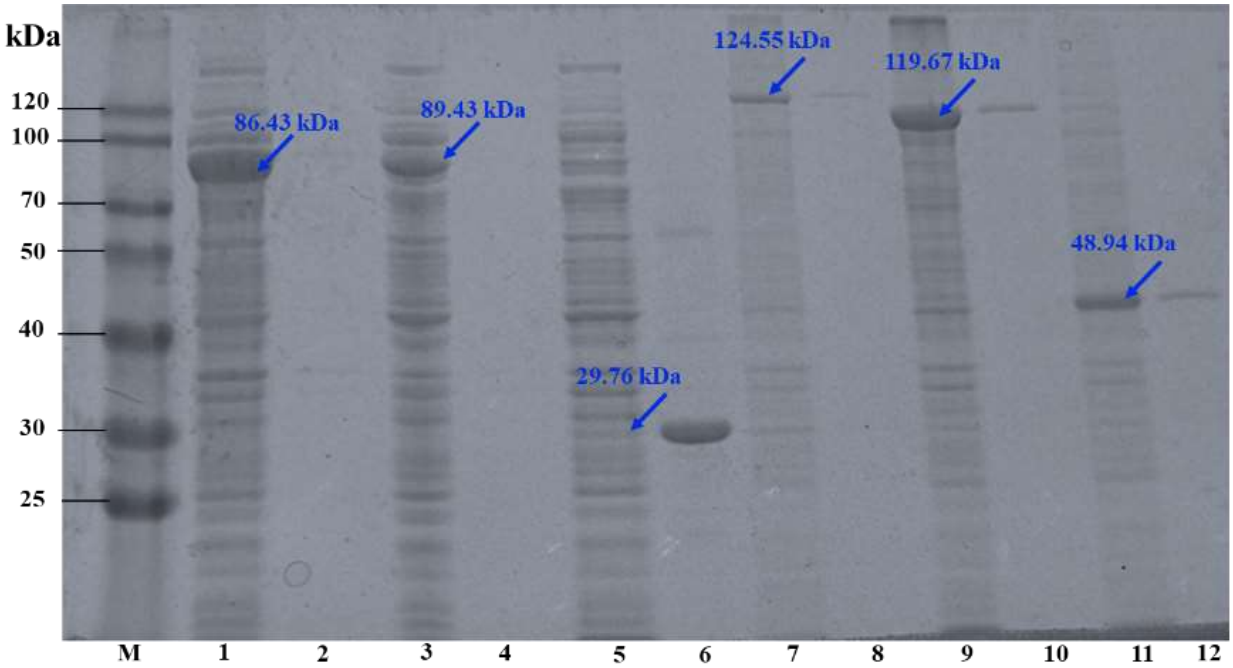


Figure S2. SDS-PAGE analysis of six recombinant P450s. The results showed that six different candidate genes for self-sufficient P450 skeleton enzymes were all soluble and in line with the expected protein size. M: protein Marker, lane 1~2: $P450_{TT}$ supernatant, precipitate; Lane 3~4: $P450_{JT}$ supernatant, precipitate; Lane 5~6: CviUPO supernatant, precipitate; Lane 7~8: CYP505X supernatant, precipitation; Lane 9~10: CYP505A30 supernatant, precipitate; Lane 11~12: CYP152K6 supernatant, precipitate

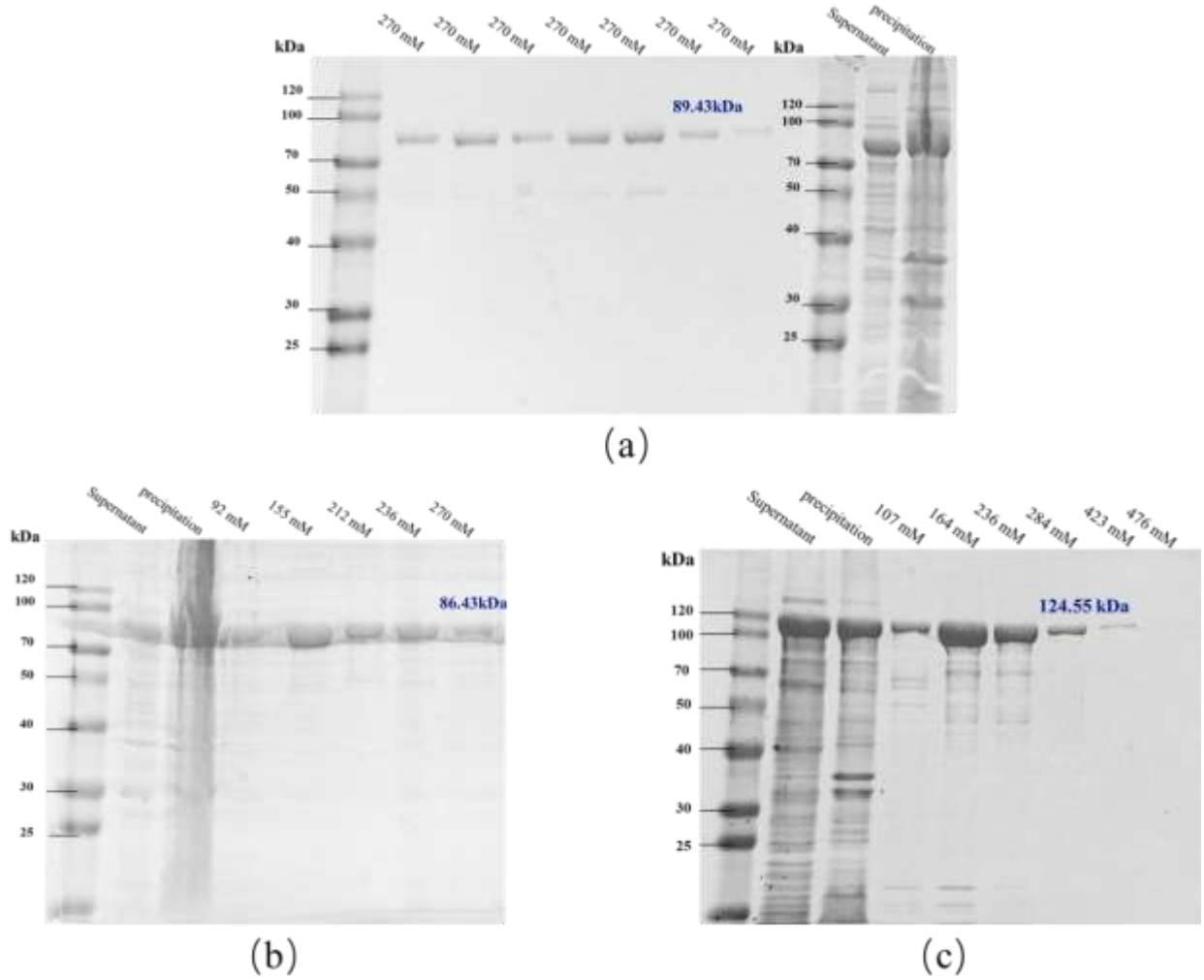


Fig S3. SDS-PAGE analysis of three purified self-sufficient P450 enzymes. (a) $P450_{JT}$-WT. (b) $P450_{TT}$-C385S. (c) CYP505X-WT.

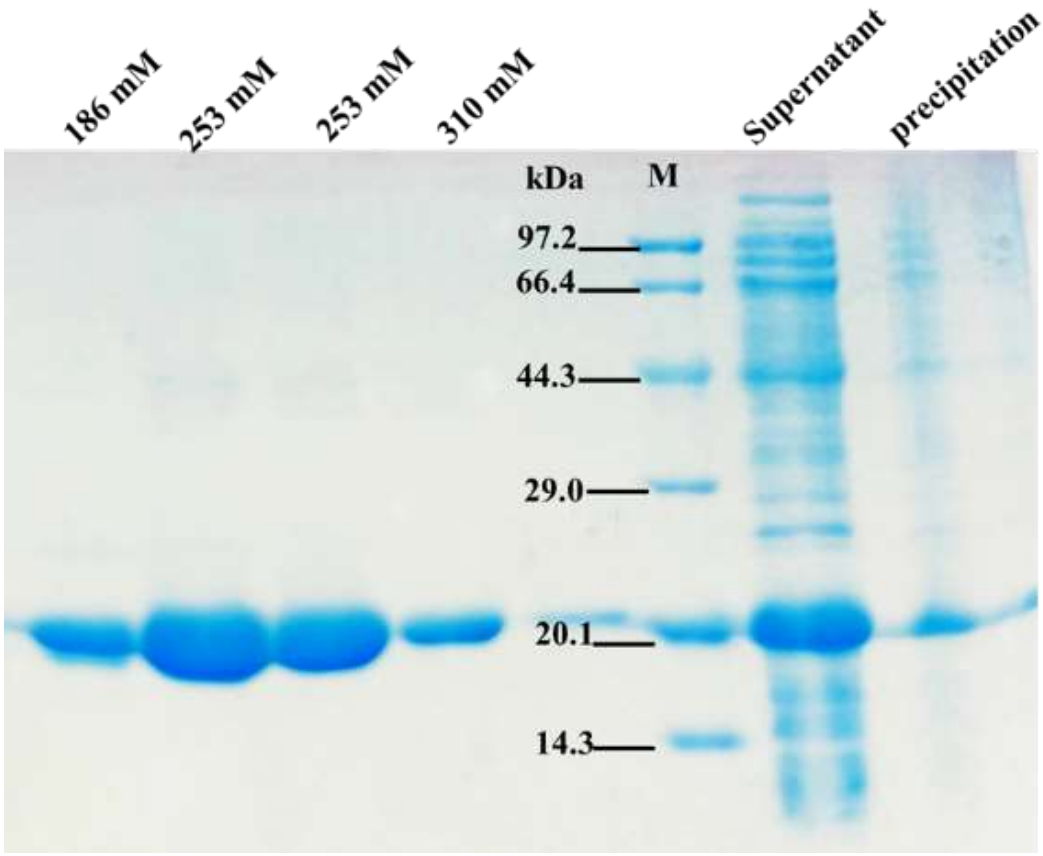


Fig S4. SDS-PAGE analysis of purified myoglobin

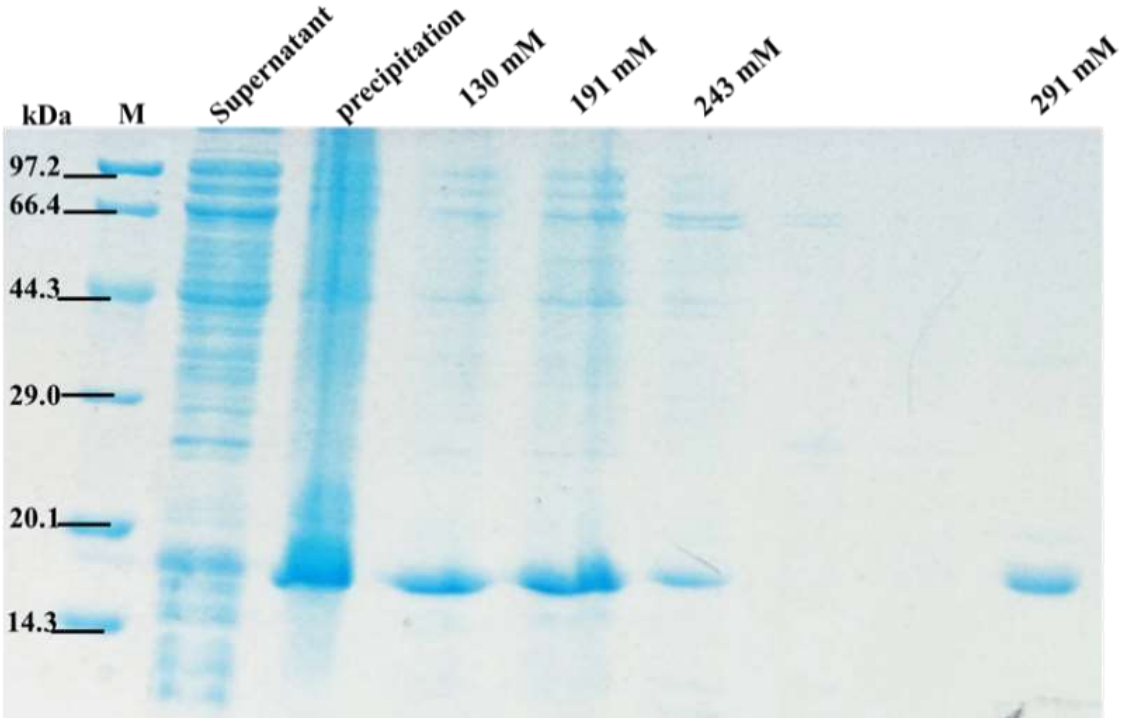


Fig S5. SDS-PAGE analysis of purified cytochrome C

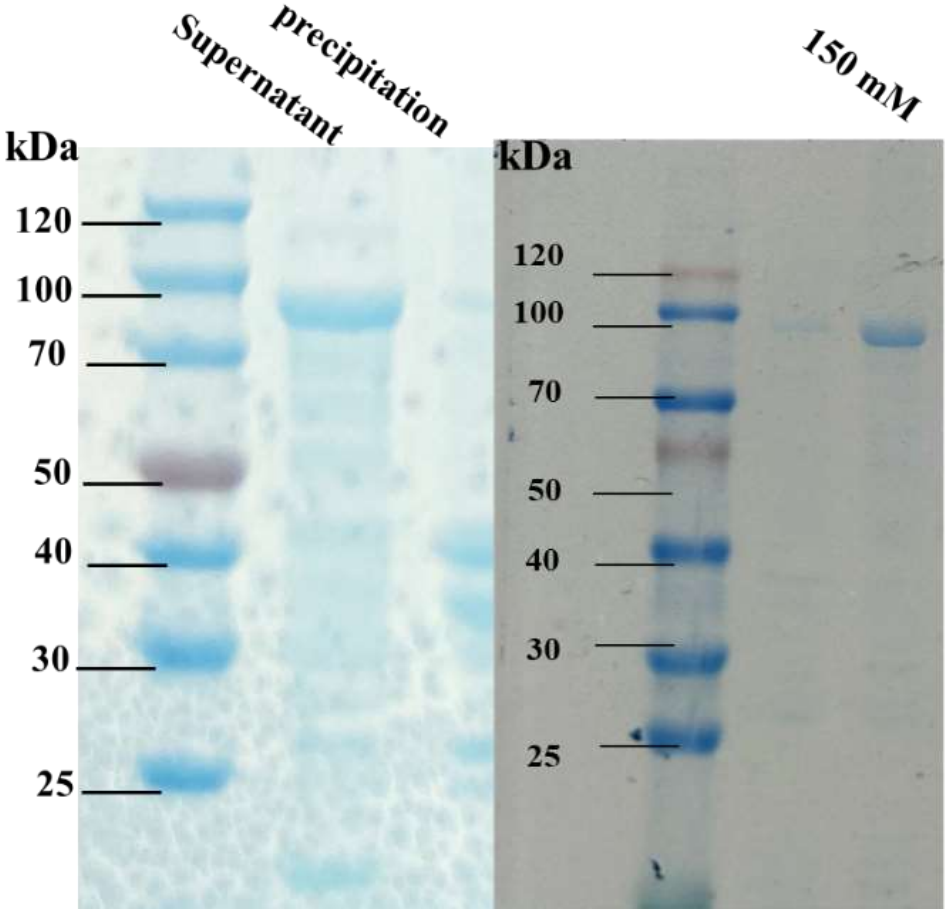


Fig S6. SDS-PAGE analysis of purified $P450_{BM3}$

5. Comparison of carbenoid model reactions catalyzed by different self-sufficient P450s

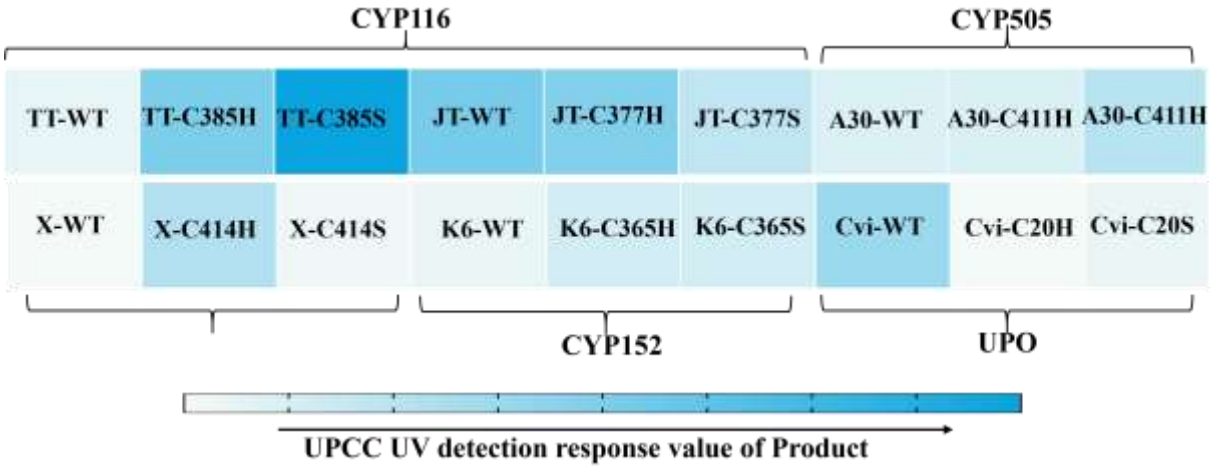


Fig S7. Fingerprint of carbenoid model reactions catalysed by different self-sufficient P450s. Experiments were performed using crude extract of recombinant P450s. Catalyst were normalized such that all lyophilized cells added to buffer in the same amount for ultrasonication. Product was determined by ultra-performance convergence chromatography.

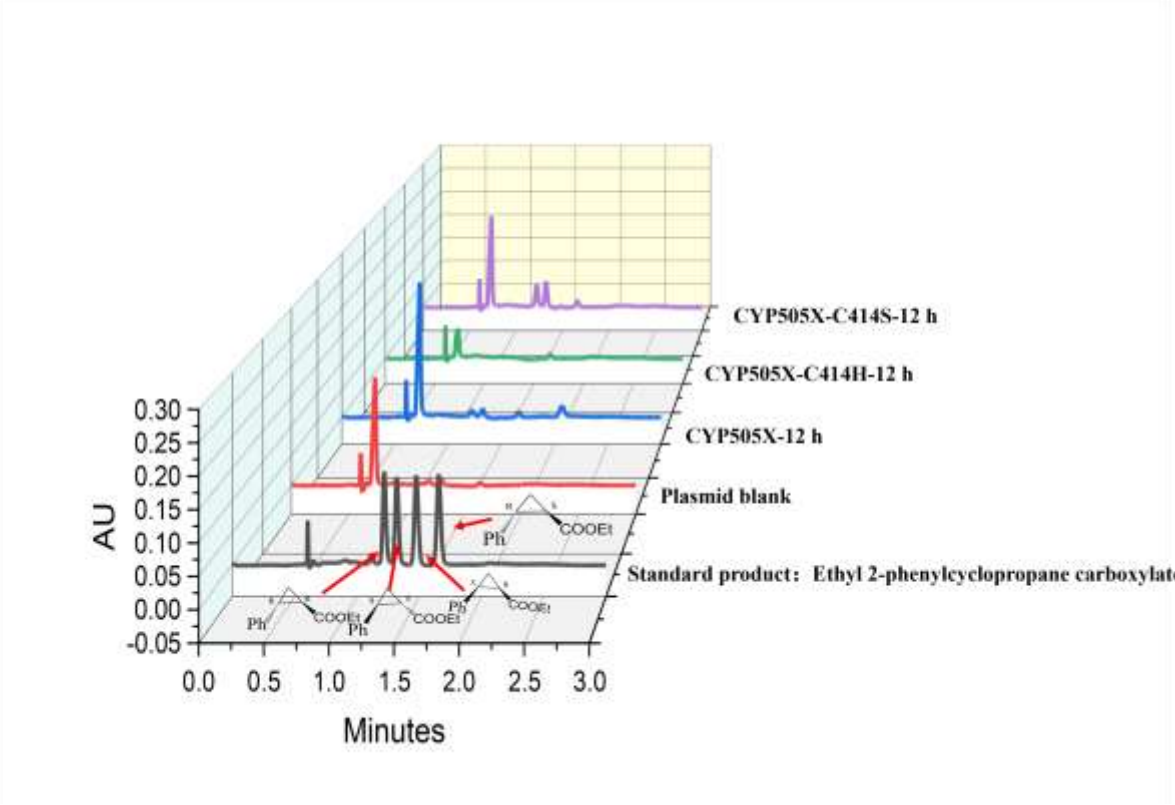


Fig S8. Ultra-performance convergence chromatography analysis of reactions catalyzed by crude enzymes of $P450_{TT}$ and its variants.

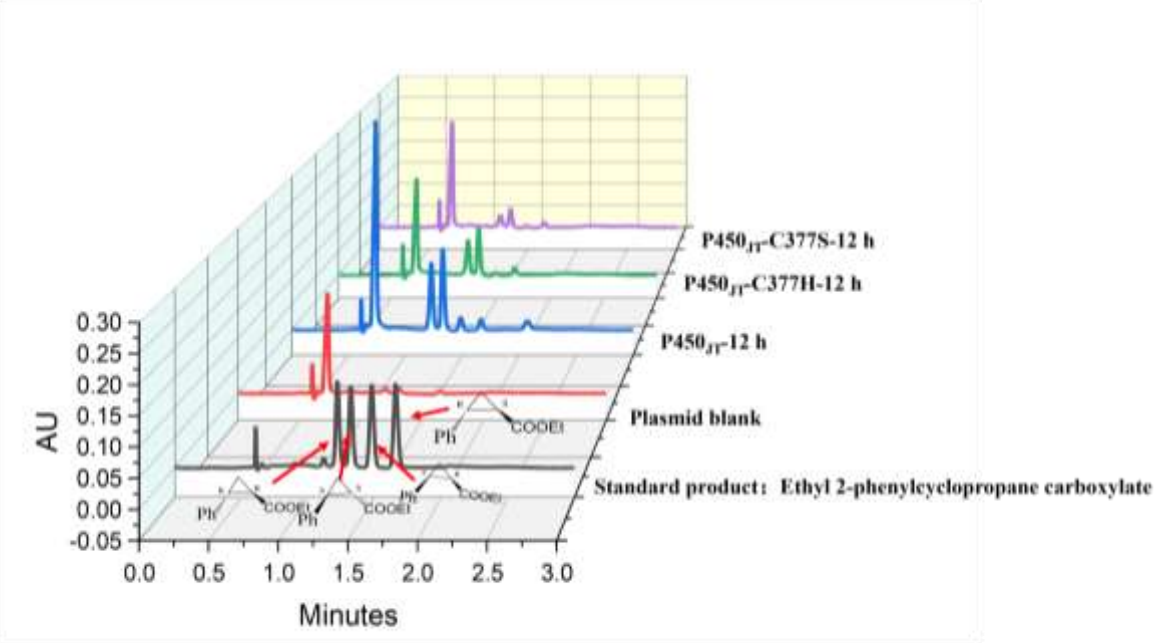
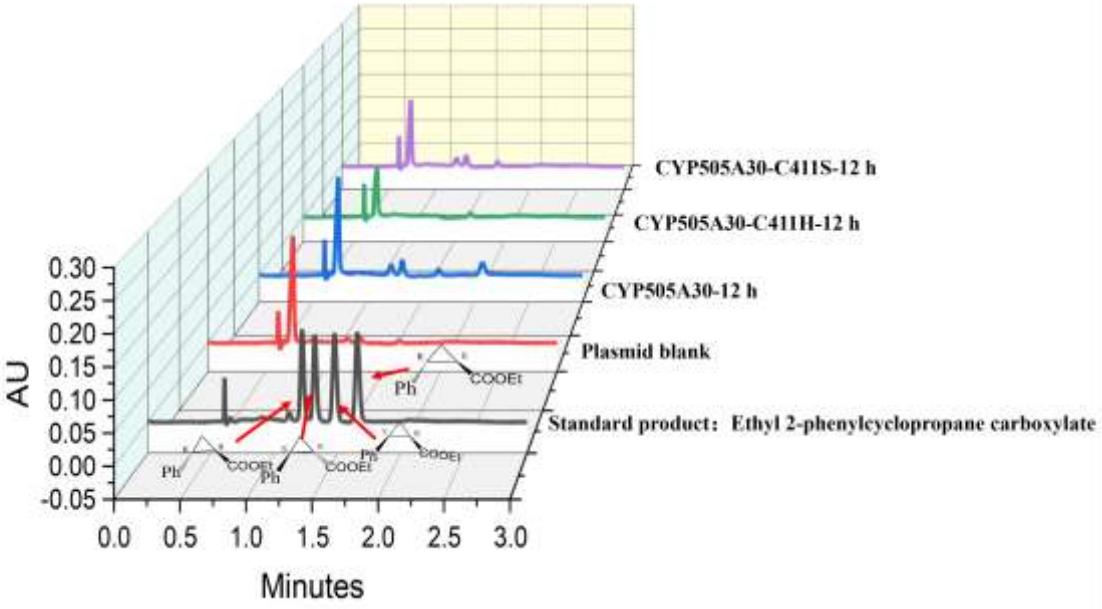


Fig S9. Ultra-performance convergence chromatography analysis of reactions catalyzed by crude enzymes of $P450_{JT}$ and its variants.



147

148 **Fig S10.** Ultra-performance convergence chromatography analysis of reactions catalyzed by crude

149 enzymes of CYP505A30 and its variants.

150

151

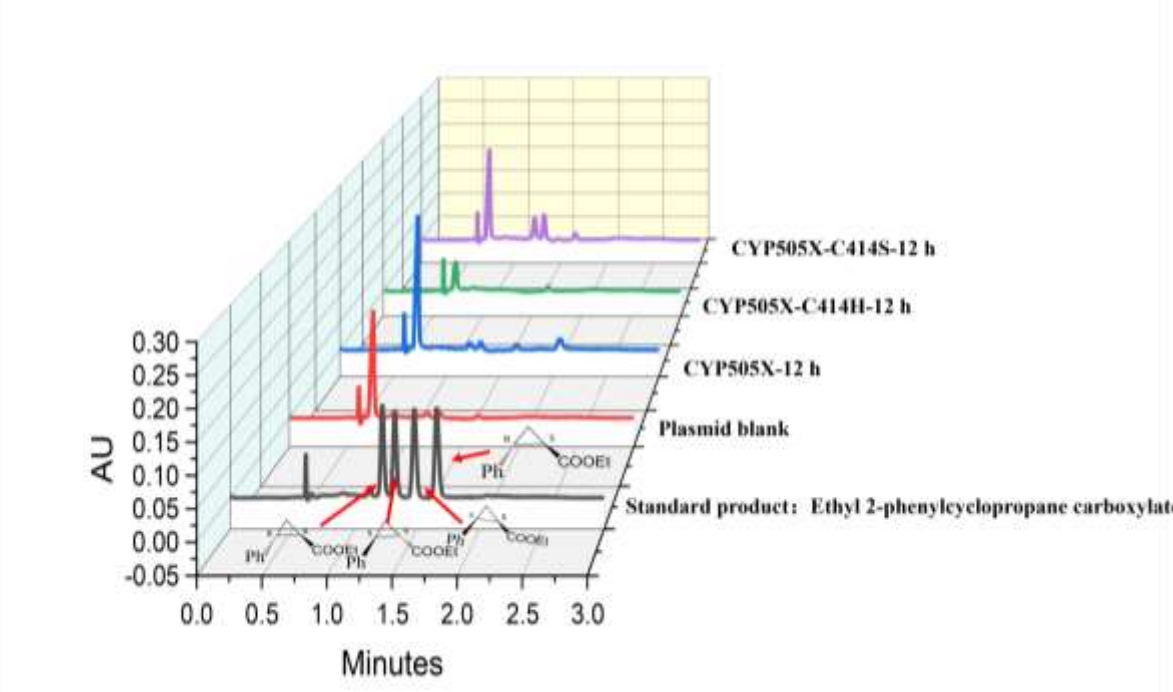


Fig S11. Ultra-performance convergence chromatography analysis of reactions catalyzed by crude enzymes of CYP505X and its variants.

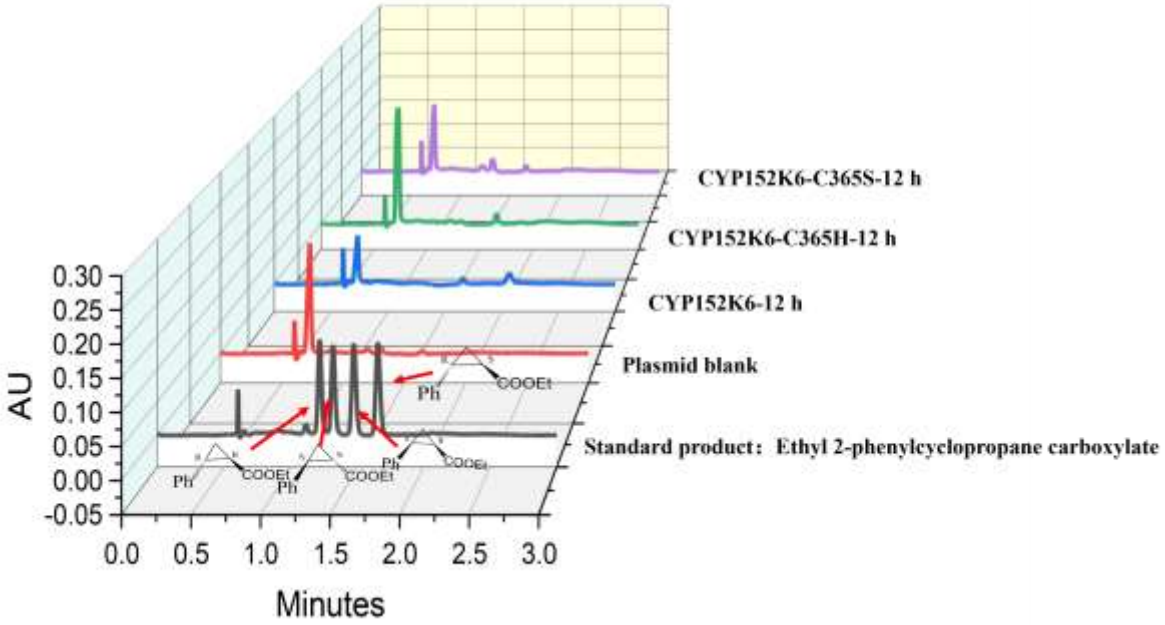


Fig S12. Ultra-performance convergence chromatography analysis of reactions catalyzed by crude enzymes of CYP152K6 and its variants.

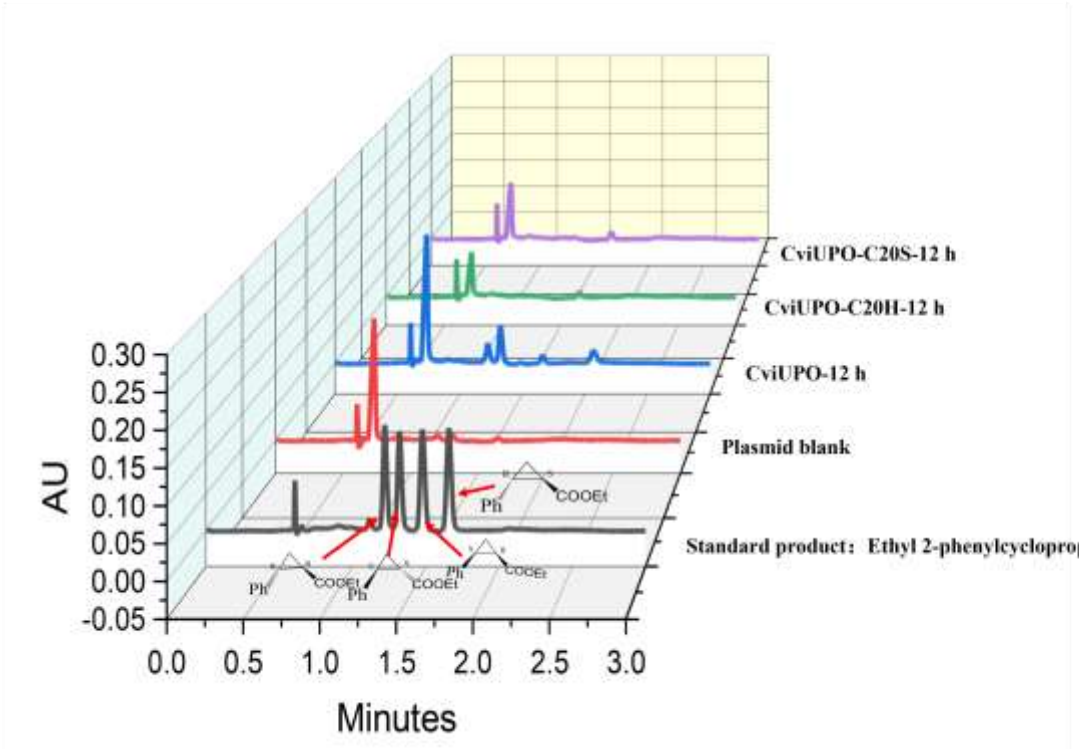


Fig S13. Ultra-performance convergence chromatography analysis of reactions catalyzed by crude enzymes of CviUPO and its variants.

6. Comparison of turnover frequencies (TOFs) and stereoselectivity of P450s

Table S4. Comparison of turnover frequencies (TOFs) and stereoselectivity of carbenoid model reaction catalyzed by different purified P450s

Purified Enzymes	TOF (min^{-1})	de_E (%)	ee_{cis} (%)	ee_{trans} (%)
$P450_{TT}$-WT	0.75	94.0	99.9	79.8
$P450_{TT}$-C285H	0.80	95.4	99.9	83.3
$P450_{TT}$-C285S	1.50	31.5	58.0	20.9
$P450_{JT}$-WT	1.20	50.7	68.0	33.4
$P450_{JT}$-C377H	0.45	84.0	99.9	56.7
$P450_{JT}$-C377S	0.70	83.3	93.5	51.9
CYP505X	0.95	91.6	99.9	75.8
P411-CIS	0.40	99.9	99.9	n.d.
P450-CIS	0.30	99.9	99.9	n.d.
Myoglobin	0.50	81.1	99.9	99.9
Cytochrome C	0.30	84.4	69.2	99.9

8. Residues interacted with the heme cofactors of different P450s

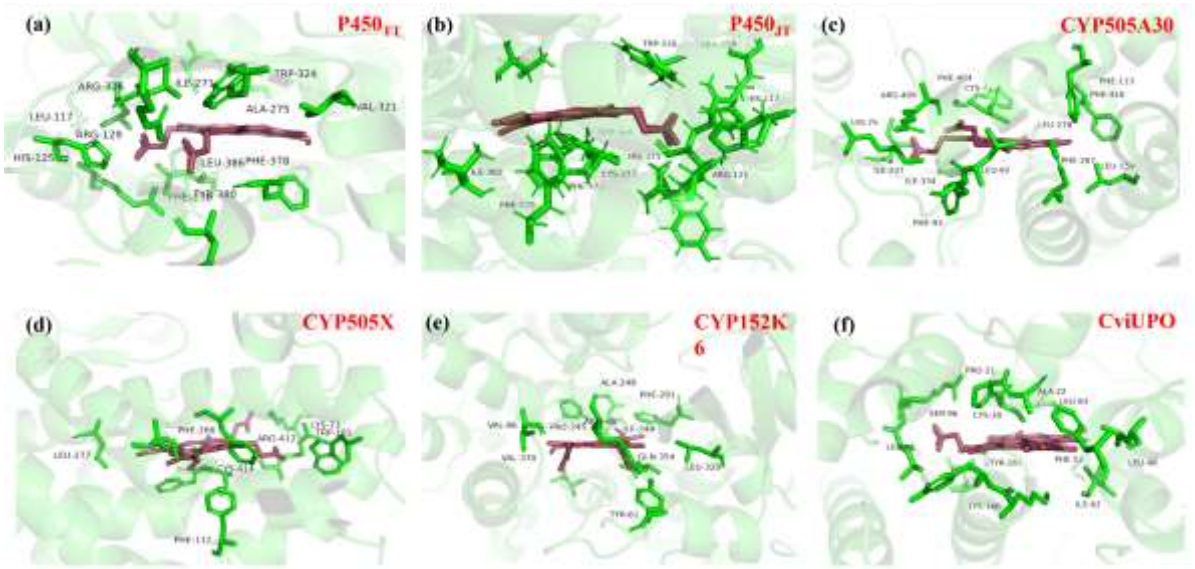


Fig S14. Residues interacted with the heme cofactors of different P450s. Molecular simulations were conducted on the heme pocket of the following proteins: (a) $P450_{TT}$ (PDB: 6GII), (b) $P450_{JT}$ (homology model built using AlphaFold 2), (c) CYP505A30 (homology model built using AlphaFold 2), (d) CYP505X (homology model built using AlphaFold 2), (e) CYP152K6 (homology model built using AlphaFold 2), and (f) CviUPO (homology model built using AlphaFold 2). The heme cofactor (represented in red) and the surrounding residues were visualized as sticks.

Reference

1. Z. J. Wang, H. Renata, N. E. Peck, C. C. Farwell, P. S. Coelho and F. H. Arnold, *Angew Chem Int Ed Engl*, 2014, **53**, 6810-6813.
2. T. Heel, J. A. McIntosh, S. C. Dodani, J. T. Meyerowitz and F. H. Arnold, *Chembiochem*, 2014, **15**, 2556-2562.
3. S. B. J. Kan, X. Huang, Y. Gumulya, K. Chen and F. H. Arnold, *Nature*, 2017, **552**, 132-136.
4. A. Stamatakis, P. Hoover and J. Rougemont, *Syst Biol*, 2008, **57**, 758-771.
5. D. Zhang, F. L. Gao, I. Jakovlic, H. Zou, J. Zhang, W. X. Li and G. T. Wang, *Mol Ecol Resour*, 2020, **20**, 348-355.
6. K. Katoh, J. Rozewicki and K. D. Yamada, *Brief Bioinform*, 2019, **20**, 1160-1166.
7. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler and L. S. Jermiin, *Nat Methods*, 2017, **14**, 587-+.
8. I. Letunic and P. Bork, *Nucleic Acids Res*, 2021, **49**, W293-W296.
9. X. Robert and P. Gouet, *Nucleic Acids Res*, 2014, **42**, W320-W324.