Supporting Information

Stereoselective decarboxylation of *β*-functionalized carboxylic acids enabled by engineered fatty acid photodecarboxylase

Xinping Zhang, Jieyu Zhou, Jie Zhang, Xiangyuan Gu, Huiru Wang, Shunyu Zhao, Feifan Luo, Ye Ni**

State Key Laboratory of Food Science and Resources, Jiangnan University, Wuxi 214122, Jiangsu, China

Key laboratory of industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi 214122, Jiangsu, China

Key Laboratory of Industrial Synthetic Biology of Jiangsu Province, Jiangnan University, Wuxi 214122, China

*Corresponding authors

Dr. Jieyu Zhou

E-mail address: zhoujieyu@jiangnan.edu.cn

ORCID: 0000-0003-3292-0314

Prof. Ye Ni

E-mail address: yni@jiangnan.edu.cn

ORCID: 0000-0003-4887-7517

Contents

Supplementary Tables:

Table S1. List of abbreviations.

Abbreviations	Name
*Cv*FAP	Fatty acid photodecarboxylase from *Chlorella variabilis*
FAD	Flavin Adenine Dinucleotide
A	Alanine (Ala)
F	Phenylalanine (Phe)
I	Isoleucine (Ile)
S	Serine (Ser)
H	Histidine (His)
E	Glutamic acid (Glu)
D	Aspartic acid (Asp)
T	Threonine (Thr)
Y	Tyrosine (Tyr)
N	Asparagine (Asn)
R	Arginine (Arg)
K	Lysine (Lys)
Q	Glutamine (Gln)
W	Tryptophan (Trp)
G	Glycine (Gly)
P	Proline (Pro)
M	Methionine (Met)
C	Cysteine (Cys)
L	Leucine (Leu)
V	Valine (Val)
MD simulations	Molecular Dynamics simulations
GC-MS	Gas Chromatography-Mass Spectrometry
FRISM	Focused Rational Iterative Site-specific Mutagenesis
SSM	site-saturation mutagenesis
ROS	reactive oxygen species
1-NITC	1-naphthyl isothiocyanate
HPLC-UV	High Performance Liquid Chromatography-Ultraviolet
DMF	dimethylformamide

Table S2. Preliminary screening of N575 mutants for decarboxylation of *β*-hydroxydecanoic acid.

Enzyme	2-Nonanol (mM)
WT	1.136
N575A	ND
N575F	1.448
N575I	ND
N575S	ND
N575H	ND
N575E	ND
N575D	ND
N575R	ND
N575Q	0.471
N575K	ND
N575Y	0.594
N575P	ND
N575T	ND
N575G	ND
N575W	ND
N575M	ND
N575C	ND
N575L	ND
N575V	ND

ND, not detected.

Table S3. Statistical significance analysis of all variants relative to WT (Dunnett's test).

Dunnett's multiple comparisons test	Significant	Summary	Adjusted P Value
WT vs. V463K (K)	No	ns	0.1572
WT vs. V463D (D)	No	ns	0.1398
WT vs. V463E (E)	Yes	*	0.0355
WT vs. V463Q (Q)	Yes	*	0.0398
WT vs. V463N (N)	Yes	****	<0.0001
WT vs. V463H (H)	Yes	****	<0.0001
WT vs. V463Y (Y)	Yes	****	<0.0001
WT vs. V463W (W)	Yes	****	<0.0001
WT vs. V463S (S)	Yes	****	<0.0001
WT vs. V463T (T)	Yes	****	<0.0001
WT vs. V463G (G)	Yes	*	0.0440
WT vs. V463P (P)	Yes	*	0.0108
WT vs. V463A (A)	Yes	****	<0.0001
WT vs. V463M (M)	Yes	****	<0.0001
WT vs. V463C (C)	Yes	****	<0.0001
WT vs. V463F (F)	No	ns	0.2671
WT vs. V463L (L)	No	ns	0.8559
WT vs. V463R (R)	No	ns	0.9990
WT vs. V463I (I)	Yes	*	0.0394
WT vs. S574A/V463A	Yes	****	<0.0001
WT vs. N575F/V463A	Yes	****	<0.0001
WT vs. Y466A/V463A	Yes	****	<0.0001
WT vs. S574A/V463T	Yes	****	<0.0001
WT vs. N575F/V463T	Yes	****	<0.0001
WT vs. Y466A/V463T	Yes	****	<0.0001
WT vs. S574A/V463S	Yes	****	<0.0001
WT vs. N575F/V463S	Yes	****	<0.0001
WT vs. Y466A/V463S	Yes	**	0.0015
WT vs. Y466A/S574A	Yes	****	<0.0001
WT vs. N575F/S574A	Yes	**	0.0013
WT vs. Y466A/N575F	No	ns	0.9957
WT vs. Y466A/N575F/V463T	No	ns	>0.9999
WT vs. Y466A/S574A/V463A	Yes	**	0.0064
WT vs. Y466A/N575F/V463S	Yes	****	<0.0001
WT vs. N575F/S574A/V463S	Yes	****	<0.0001
WT vs. N575F/S574A/V463T	Yes	*	0.0475
WT vs. N575F/S574A/V463A	No	ns	0.9993
WT vs. Y466A/N575F/V463A	No	ns	0.7517
WT vs. Y466A/N575F/S574A	Yes	**	0.0037

Table S4. Binding free energies between substrate 1a and WT *Cv*FAP, V463A and V463T (from three parallel replicates).

Enzyme	Binding free energy (kJ/mol)
WT	-440.4 ± 36.0
V463T	-640.8 ± 68.0
V463A	-582.7 ± 15.2

Table S5. Primer sequences used for site-saturation mutagenesis of residue V463.

Mutation	Sequence
V463A forward	GATCCGGATGGCGCTAGCACCTATGTG
V463A reverse	CACATAGGTGCTAGCGCCATCCGGATC
V463F forward	GATCCGGATGGCTTTAGCACCTATGTG
V463F reverse	CACATAGGTGCTAAAGCCATCCGGATC
V463I forward	GATCCGGATGGCATTAGCACCTATG
V463I reverse	CACATAGGTGCTAATGCCATCCGGA
V463E forward	GATCCGGATGGCGAAAGCACCTATG
V463E reverse	CACATAGGTGCTTTCGCCATCCGGA
V463H forward	GATCCGGATGGCCATAGCACCTATG
V463H reverse	CACATAGGTGCTATGGCCATCCGGA
V463S forward	GATCCGGATGGCTCTAGCACCTATG
V463S reverse	CACATAGGTGCTAGAGCCATCCGGAT
V463L forward	GATCCGGATGGCCTTAGCACCTATGT
V463L reverse	CACATAGGTGCTAAGGCCATCCGGA
V463I forward	GATCCGGATGGCATTAGCACCTATGT
V463I reverse	CACATAGGTGCTAATGCCATCCGGA
V463P forward	GATCCGGATGGCCCTAGCACCTATGT
V463P reverse	CACATAGGTGCTAGGGCCATCCGGA
V463Y forward	GATCCGGATGGCTATAGCACCTATGT
V463Y reverse	CACATAGGTGCTATAGCCATCCGGA
V463W forward	GATCCGGATGGCTGGAGCACCTATGT
V463W reverse	CACATAGGTGCTCCAGCCATCCGGA
V463T forward	GATCCGGATGGCACTAGCACCTATGT
V463T reverse	CACATAGGTGCTAGTGCCATCCGGA
V463C forward	GATCCGGATGGCTGTAGCACCTATGT
V463C reverse	CACATAGGTGCTACAGCCATCCGGA
V463N forward	GATCCGGATGGCAATAGCACCTATGT
V463N reverse	CACATAGGTGCTATTGCCATCCGGA
V463Q forward	GATCCGGATGGCCAAAGCACCTATGT
V463Q reverse	CACATAGGTGCTTTGGCCATCCGGA
V463K forward	GATCCGGATGGCAAAAGCACCTATGT
V463K reverse	CACATAGGTGCTTTTGCCATCCGGA
V463R forward	GATCCGGATGGCCGTAGCACCTATGT
V463R reverse	CACATAGGTGCTACGGCCATCCGGA
V463G forward	GATCCGGATGGCGGTAGCACCTATGT
V463G reverse	CACATAGGTGCTACCGCCATCCGG
V463M forward	GATCCGGATGGCATGAGCACCTATGT
V463M reverse	CACATAGGTGCTCATGCCATCCGG
V463D forward	GATCCGGATGGCGATAGCACCTATGT
V463D reverse	CACATAGGTGCTATCGCCATCCGG

Supplementary Figures:

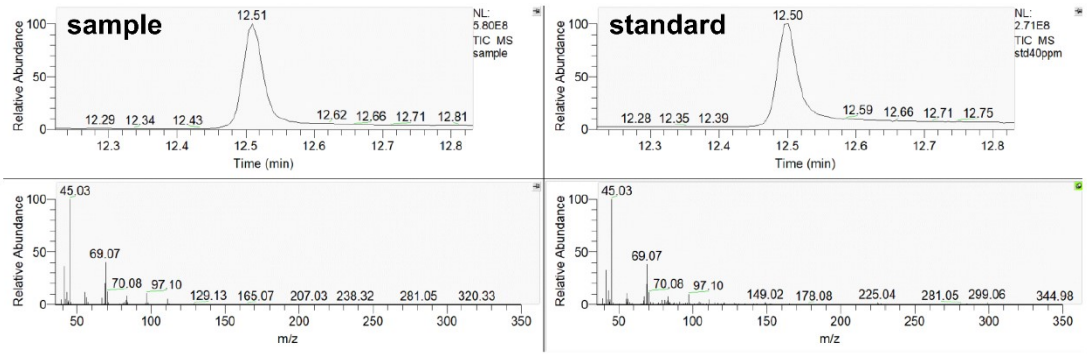


Figure S1. GC-MS analysis of enzymatic decarboxylation product 2-nonanol. (Left: mass spectrum of enzymatic product by WT; Right: mass spectrum of 2-nonanol reference standard.)

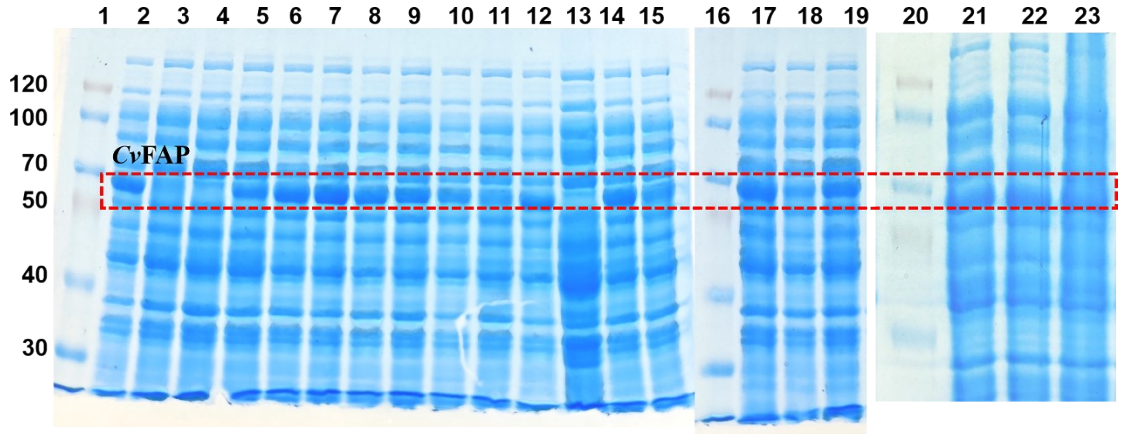


Figure S2. SDS-PAGE analysis for the expression of WT and mutants of *Cv*FAP. Lane 1, 16 and 20: protein markers. Lane 1-15: crude enzyme of WT *Cv*FAP, V463I, V463L, V463F, V463C, V463M, V463A, V463P, V463G, V463T, V463S, V463W, V463Y, V463H. Lane 17-19: crude enzyme of V463N, V463Q, V463E. Lane 21-23: crude enzyme of V463D, V463K, V463R.

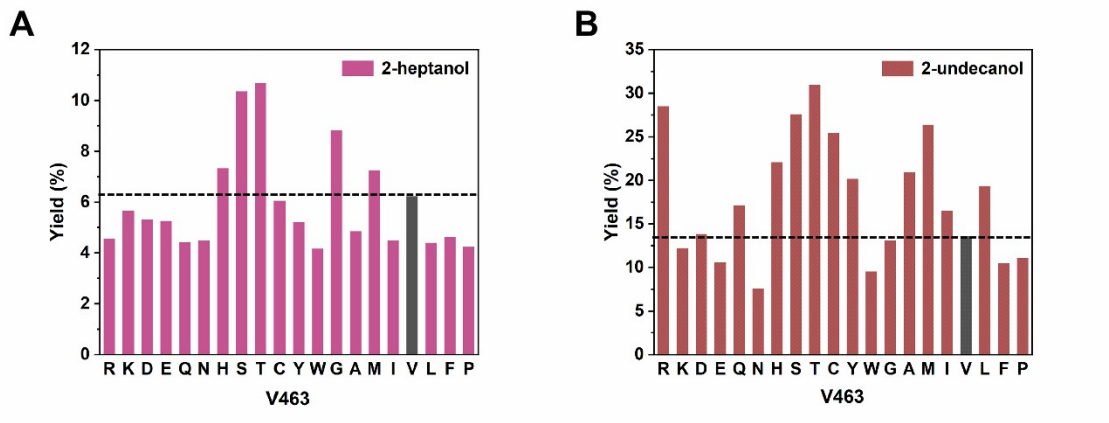


Figure S3. Comparison of decarboxylation yield of *β*-hydroxycarboxylic acid substrates by V463 variants. A) *β*-hydroxyoctanoic acid. B) *β*- hydroxydodecanoic acid. (the single-letter codes denote the corresponding amino acids)

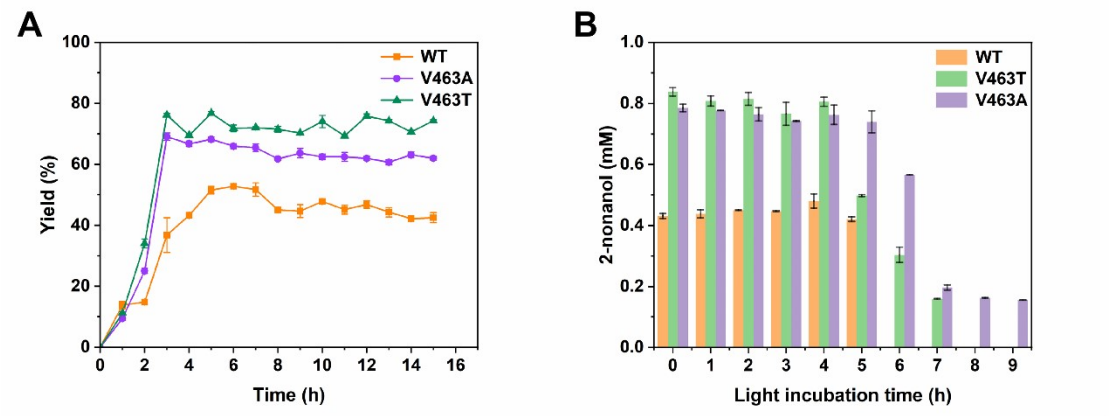


Figure S4. Evaluation of photostability of WT, V463A and V463T. A) Time course of decarboxylation. B) Photostability.

Reaction conditions: *β*-hydroxydecanoic acid (substrate final concentration: 10 mM), crude enzyme solution (0.5 g wet cell in 10 mL pH 7.5 Tris-HCl), irradiation of blue LEDs at 20 °C. Samples (500 μL) were collected hourly for GC analysis.

Reaction conditions of photostability: crude enzyme (0.5 g wet cell in 10 mL pH 7.5 Tris-HCl) exposed to the blue LEDs at 20 °C, *β*-hydroxydecanoic acid (final concentration: 1 mM). Samples (1 mL) were taken every hour.

GC data and Gas chromatography-mass spectrometry (GC-MS) analysis of light-driven production of chiral alcohol.

2c, chiral GC (Agilent CYCLOSIL-B, TR = 14.9 min, TS = 15.3 min, temperature conditions: initial temperature 40 ℃, holding 10 min, then 2 ℃/min to 80 ℃, then 30 ℃/min to 210 ℃, holding 5 min).

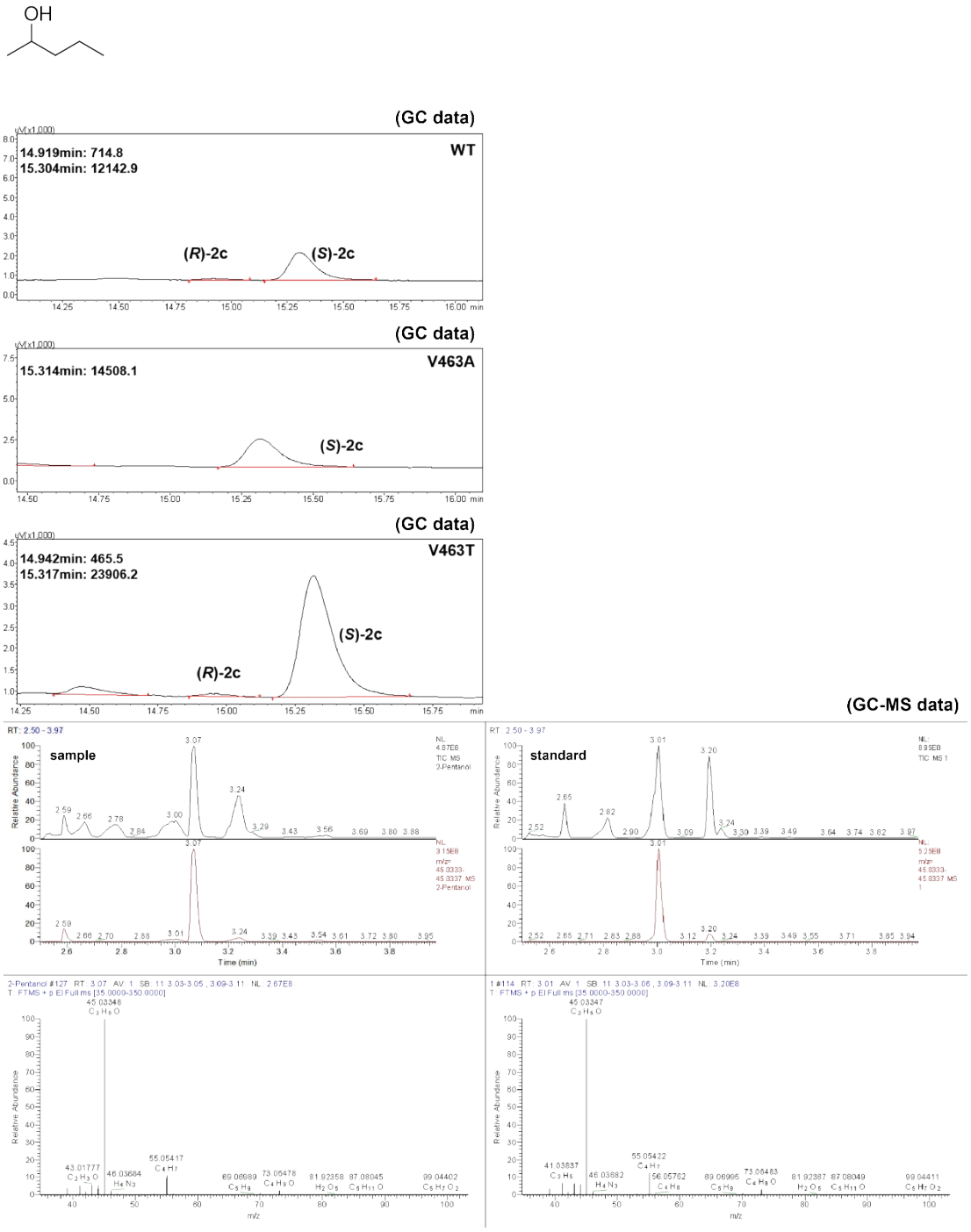


Figure S5. GC data and GC-MS data for the light-driven chiral production 2c.

3c, chiral GC (Agilent CYCLOSIL-B, TR =32.82 min, TS =33.0 min, temperature conditions: initial temperature 40 ℃, holding 10 min, then 2 ℃/min to 100 ℃, then 30 ℃/min to 220 ℃, holding 5 min).

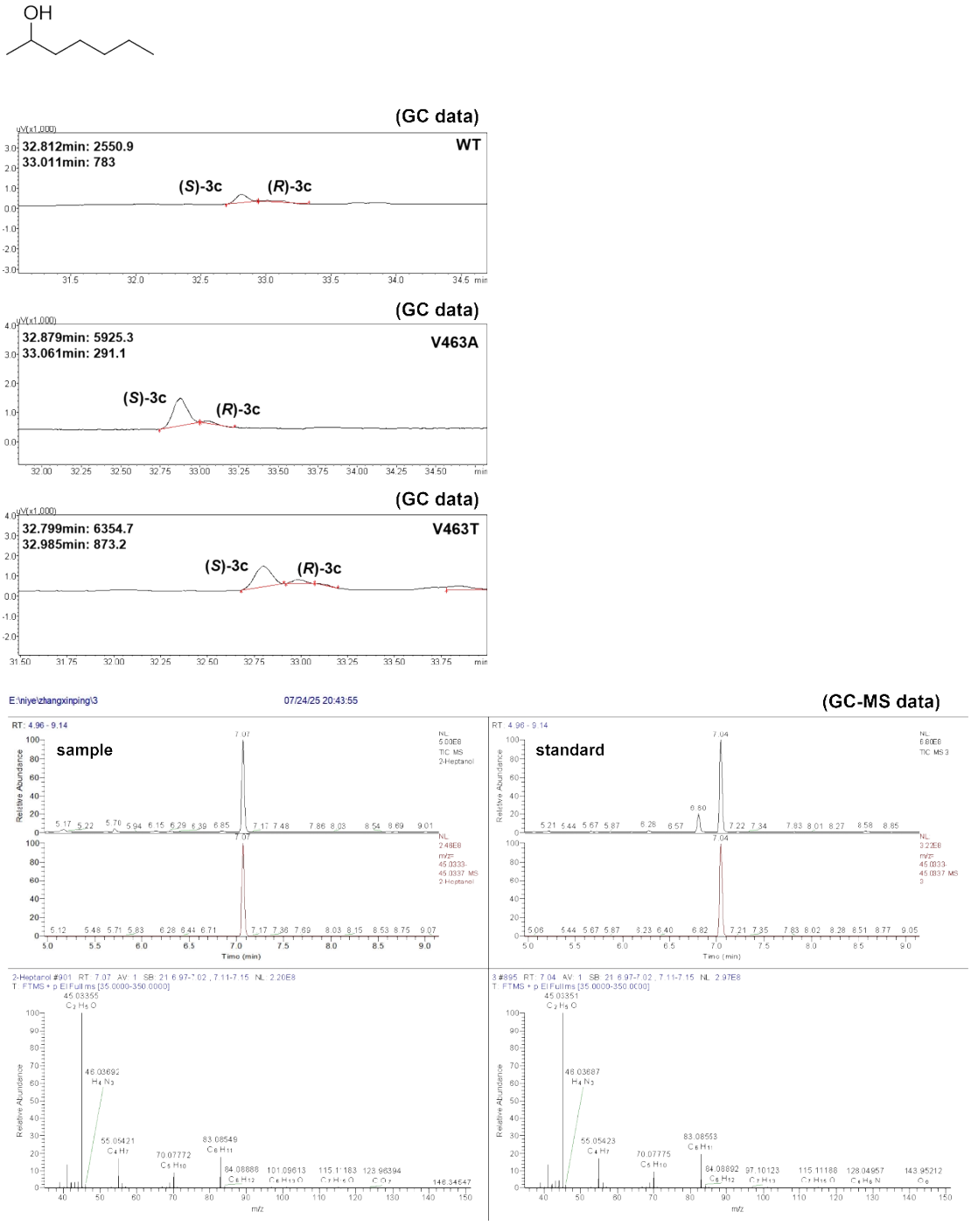
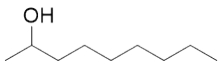


Figure S6. GC data and GC-MS data for the light-driven chiral production 3c.

1c, chiral GC (Agilent CP7502, TR = 31.2 min, TS = 31.6 min,[1] temperature conditions: initial temperature 80 ℃, holding 20 min, then 1 ℃/min to 120 ℃, then 30 ℃/min to 180 ℃, holding 3 min).



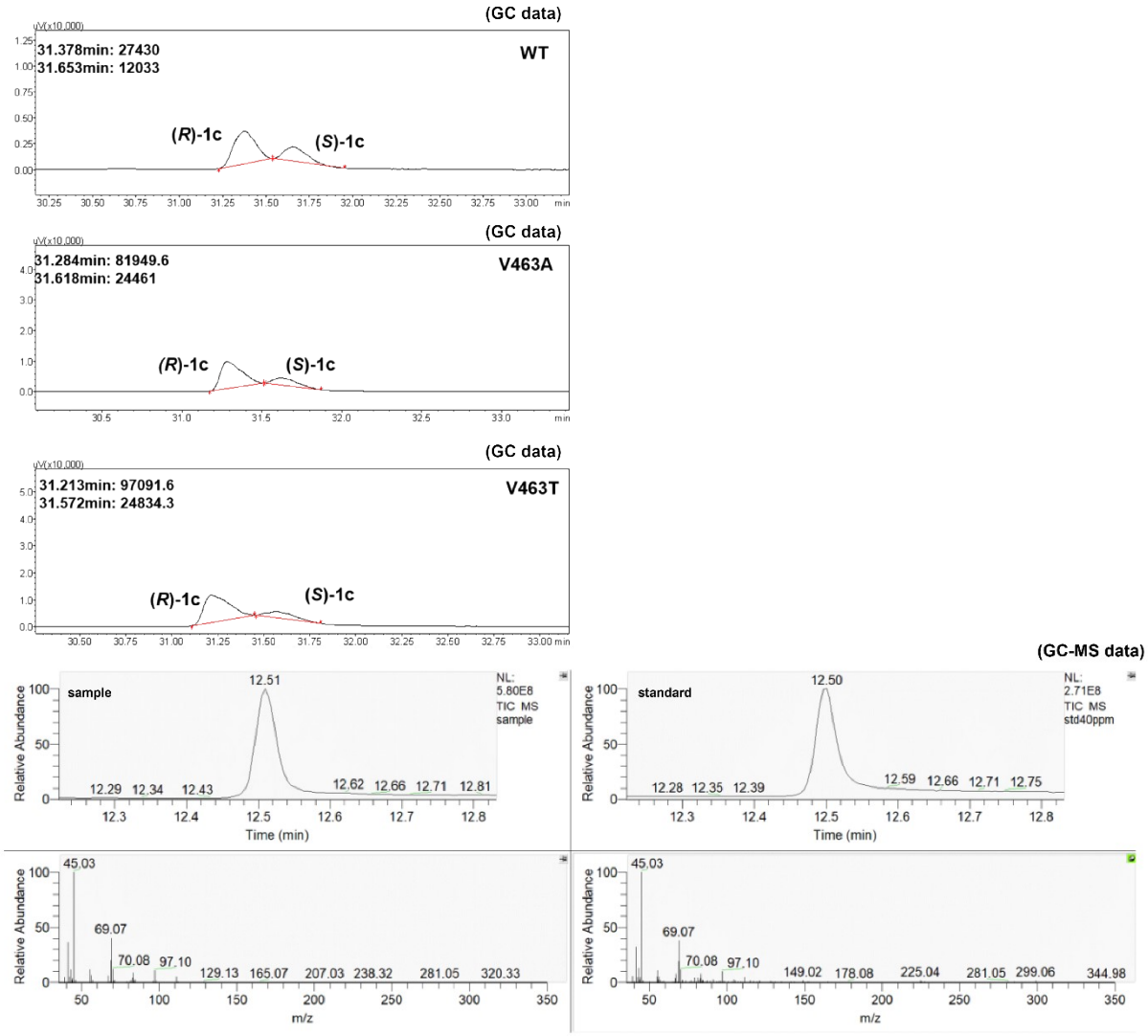


Figure S7. GC data and GC-MS data for the light-driven chiral production 1c.

4c, chiral GC (Agilent CP7502, TR = 41.1 min, TS = 41.5 min, temperature conditions: initial temperature 100 ℃, holding 10 min, then 0.5 ℃/min to 150 ℃, then 30 ℃/min to 190 ℃, holding 3 min).

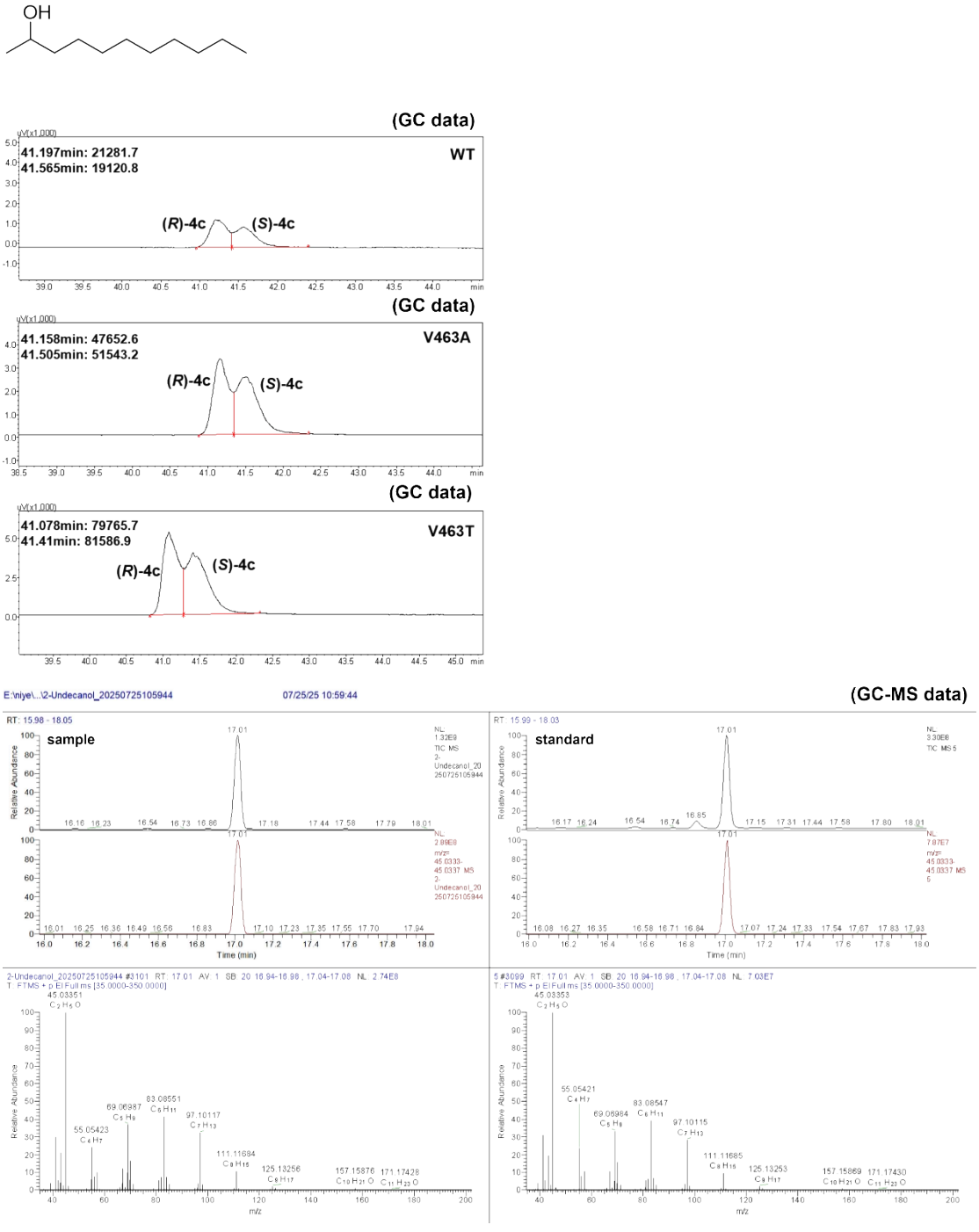


Figure S8. GC data and GC-MS data for the light-driven chiral production 4c.

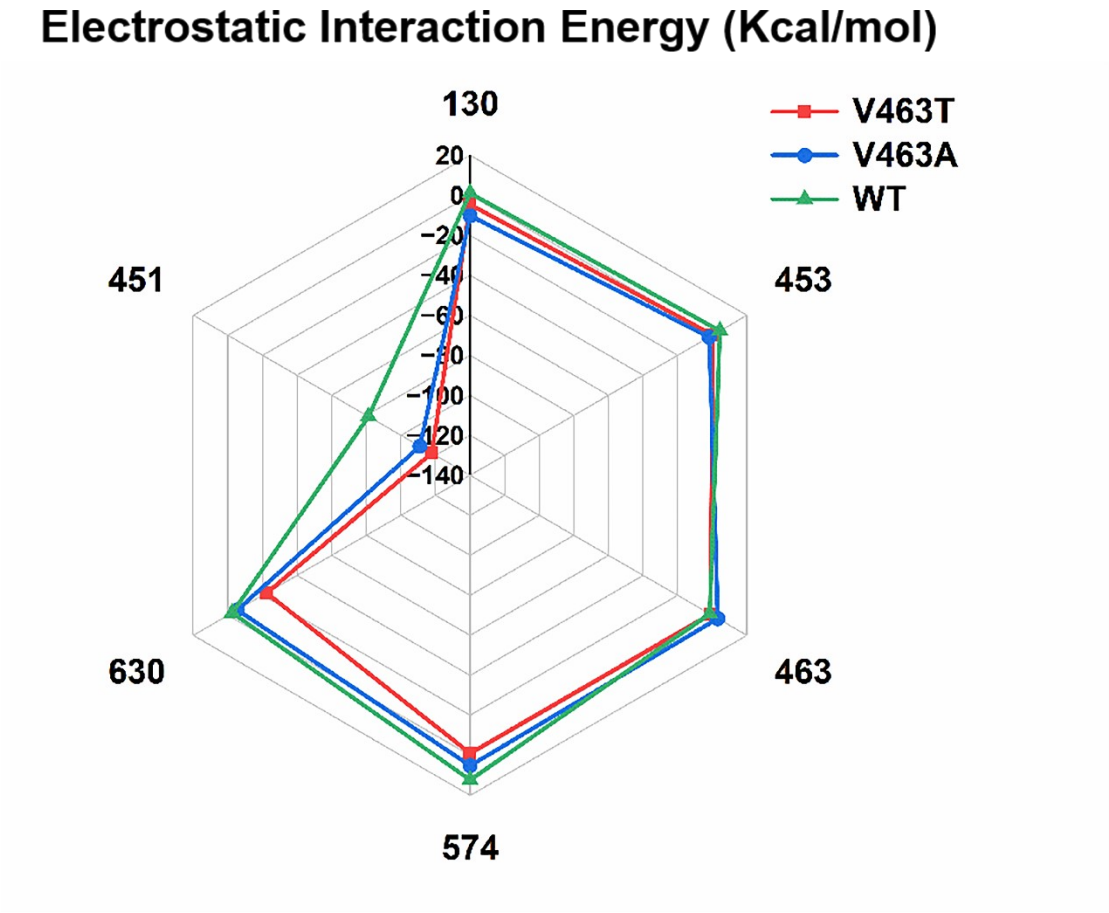
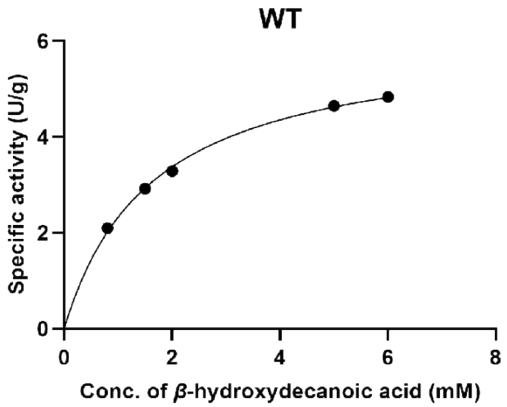


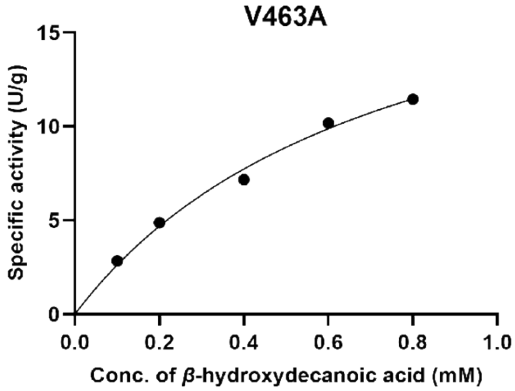
Figure S9. Electrostatic interaction energy (Kcal/mol) of substrate *β*-hydroxydecanoic acid (1a) with key residues of WT *Cv*FAP, V463A and V463T.

A



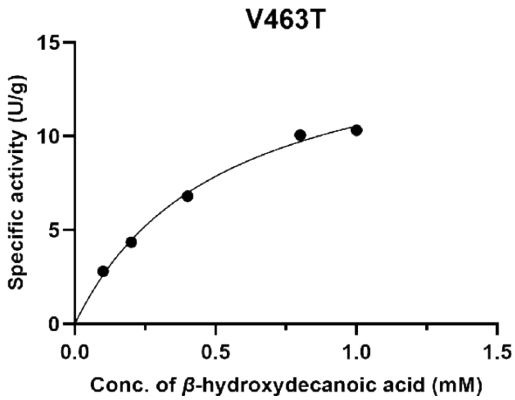
Best-fit values	
V_{max}	6.135
K_m	1.637
Std.Error	
V_{max}	0.109
K_m	0.081
R^2	0.9949

B



Best-fit values	
V_{max}	22.21
K_m	0.749
Std.Error	
V_{max}	1.945
K_m	0.114
R^2	0.9887

C



Best-fit values	
V_{max}	15.97
K_m	0.515
Std.Error	
V_{max}	0.608
K_m	0.043
R^2	0.9941

Figure S10. Kinetic measurement of WT *Cv*FAP, V463A and V463T with *β*-hydroxydecanoic acid (1a). A) WT. B) V463A. C) V463T. The data were fitted using nonlinear regression to the Michaelis-Menten equation in GraphPad Prism 8.

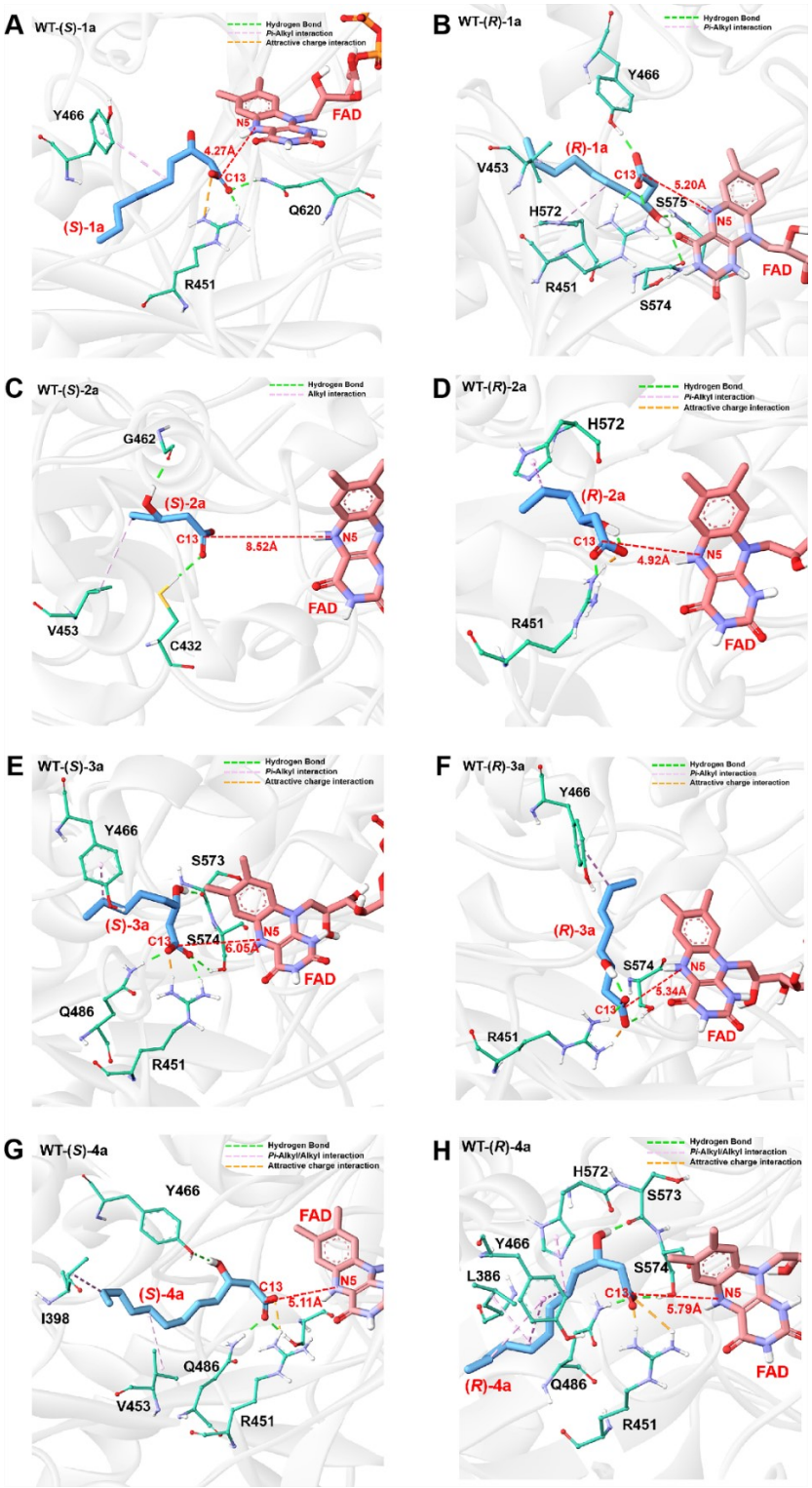


Figure S11. Representative cluster structures of WT with (*S*)/(*R*)-1a/2a/3a/4a. A) (*S*)-1a. B) (*R*)-1a. C) (*S*)-2a. D) (*R*)-2a. E) (*S*)-3a. F) (*R*)-3a. G) (*S*)-4a. H) (*R*)-4a.

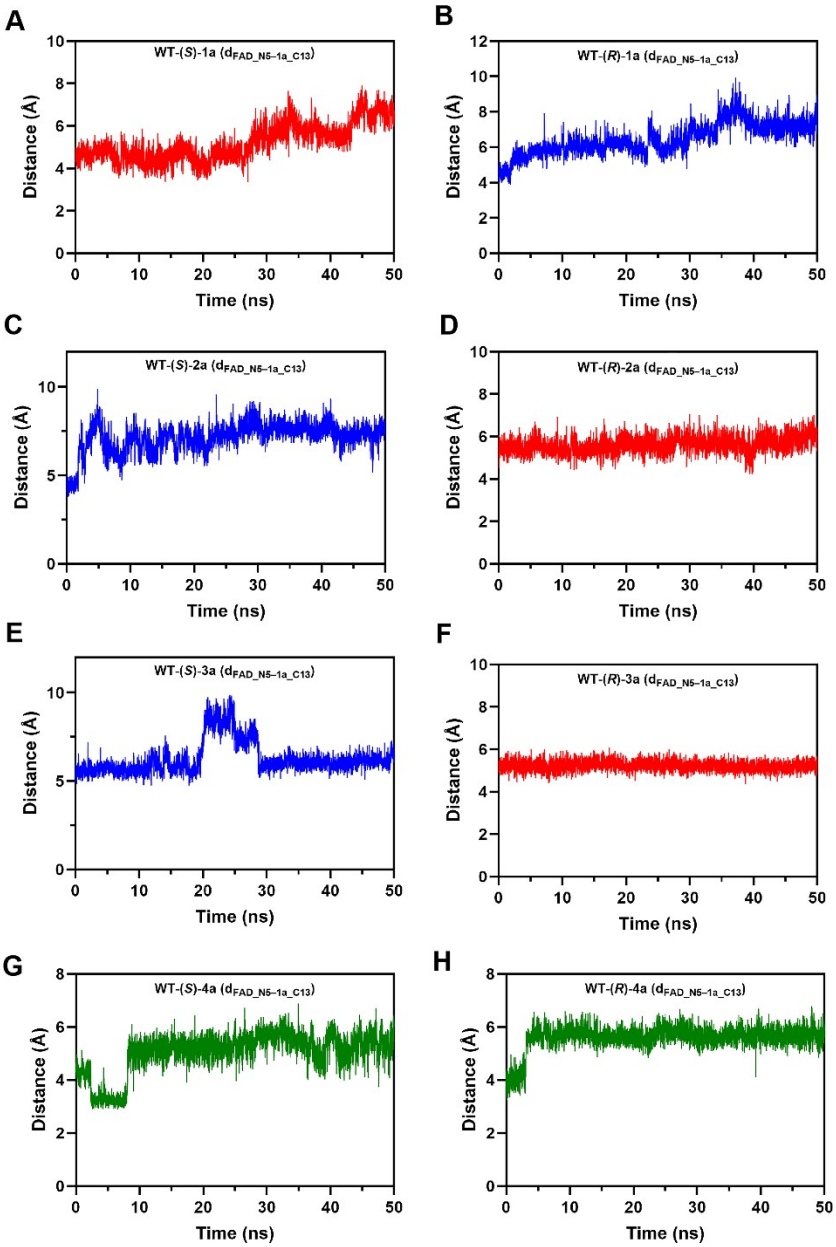


Figure S12. Molecular dynamics simulations of WT with (*S*)/(*R*)-1a/2a/3a/4a. A) (*S*)-1a. B) (*R*)-1a. C) (*S*)-2a. D) (*R*)-2a. E) (*S*)-3a. F) (*R*)-3a. G) (*S*)-4a. H) (*R*)-4a.

Here, the red solid line represents the preferred conformation of WT, the blue dashed line represents the disfavored conformation, and the green line indicates conformations with no significant selectivity.

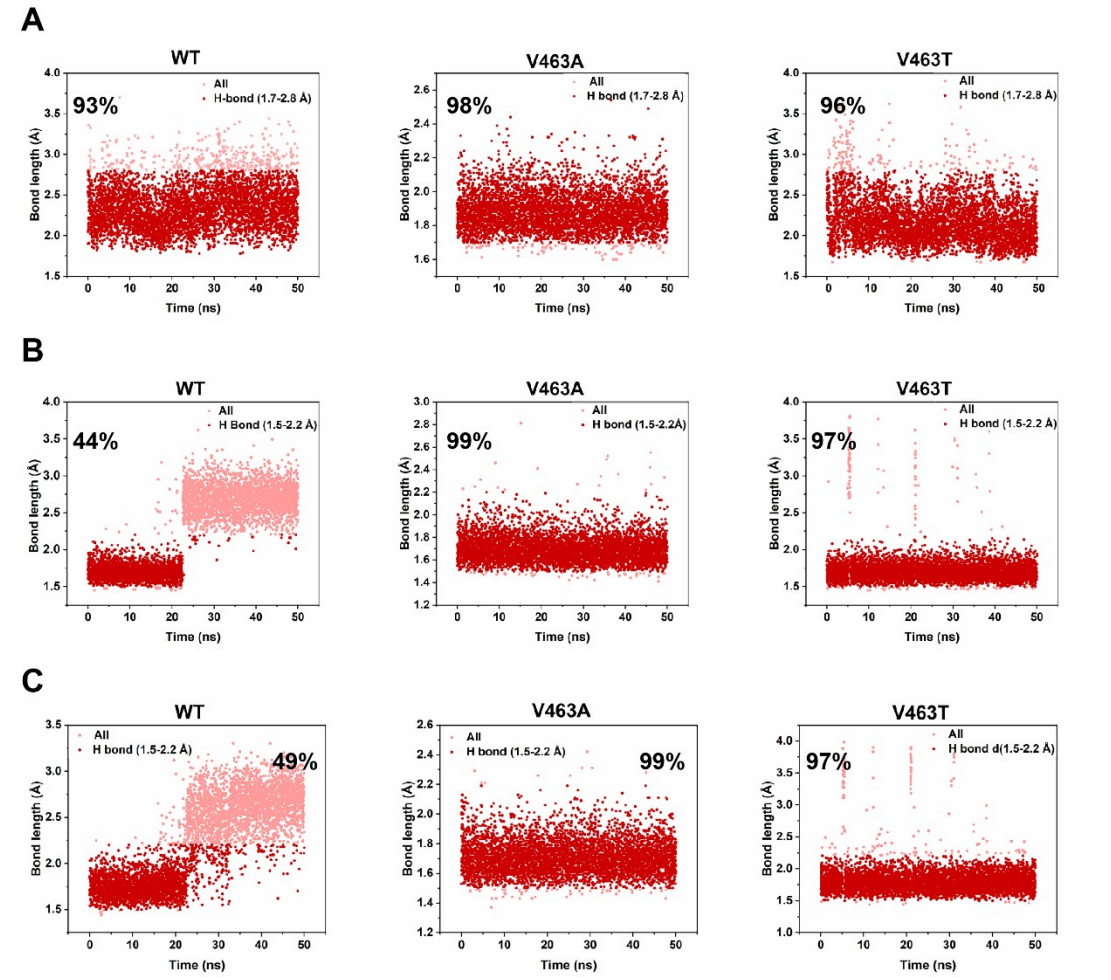


Figure S13. Statistics of key hydrogen bonds for FAD anchoring in WT *Cv*FAP, V463A, and V463T variants from molecular dynamics simulations. A) Statistics of H-bond$^{FAD_O2\text{-}FAP}$. B) Statistics of H-bond$^{FAD_H5\text{-}FAP}$. C) Statistics of H-bond$^{FAD_H7\text{-}FAP}$. Here, H-bond$^{FAD_O2\text{-}FAP}$ means H-bond formation between FAD_O2 (acceptor) and FAP_NH2 (donor), the probability was calculated for D-A distances within the canonical range of 1.7-2.8 Å, H-bond$^{FAD_H5\text{-}FAP}$ means H-bond formation between FAD_H5 (donor) and FAP_O784 (acceptor), the probability was calculated for D-A distances within the canonical range of 1.5-2.2 Å, H-bond$^{FAD_H7\text{-}FAP}$ means H-bond

formation between FAD_H7 (donor) and FAP_O785 (acceptor), the probability was calculated for D-A distances within the canonical range of 1.5-2.2 Å.

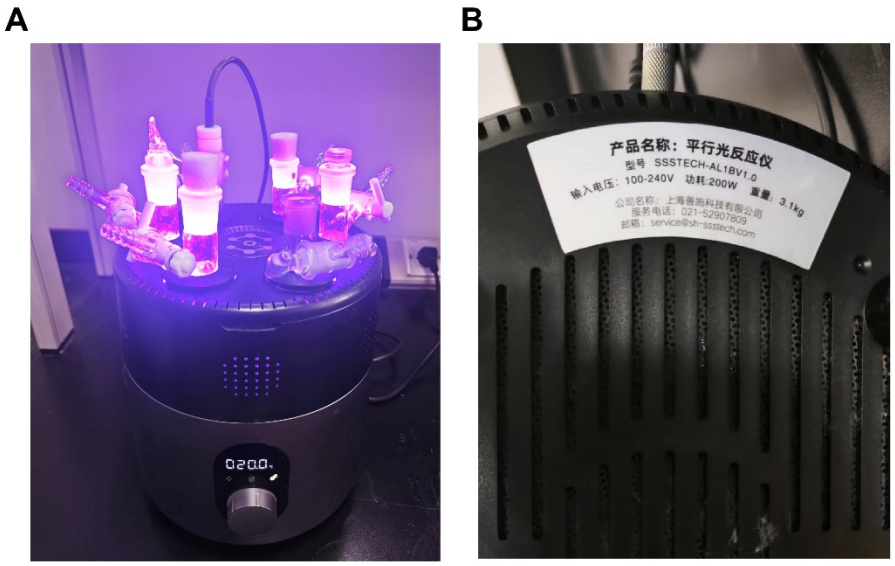


Figure S14. Photoreactor used in this study. A) Photoreactor. B) Manufacturer information: Shanshi Technology Co., Ltd. (Shanghai, China); model: SSSTECH-AL1BV1.0.

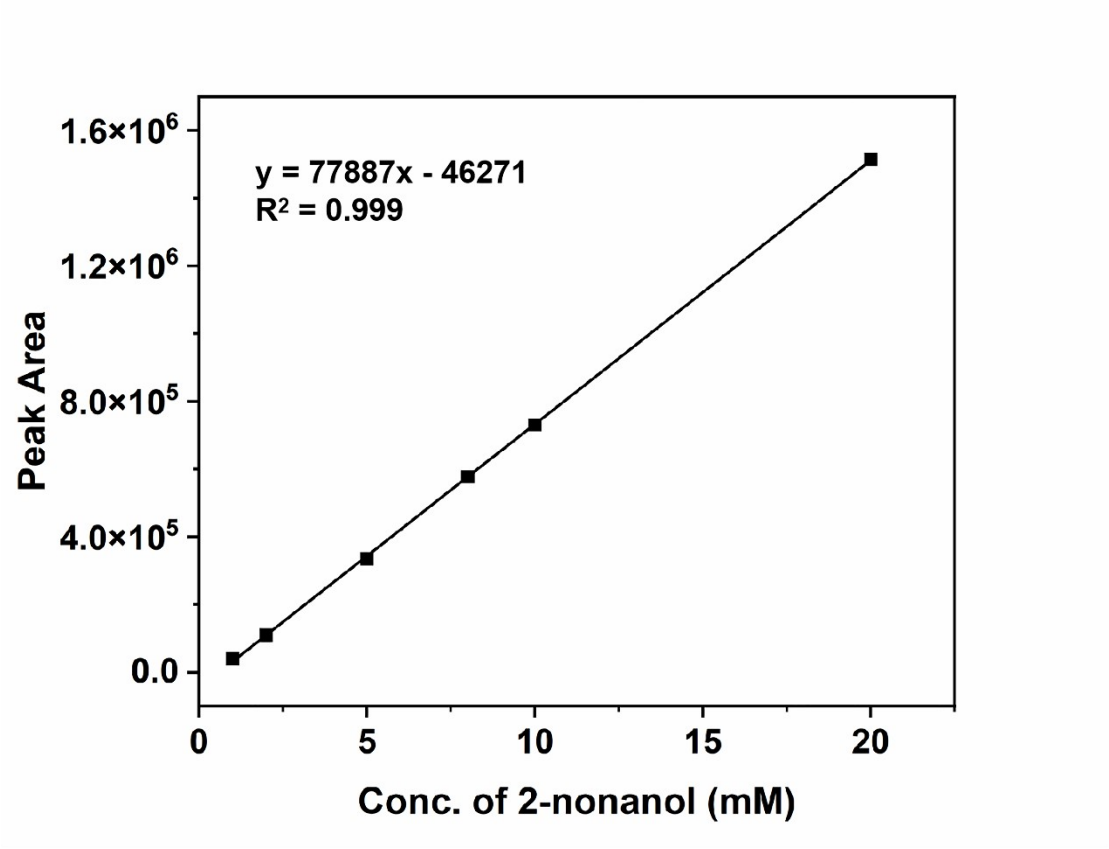


Figure S15. Standard curve for the quantification of 2-nonanol (1c). (Agilent HP-5, temperature conditions: initial temperature 40 ℃, holding 1 min, then 10 ℃/min to 120 ℃, then 30 ℃/min to 250 ℃, holding 5 min).

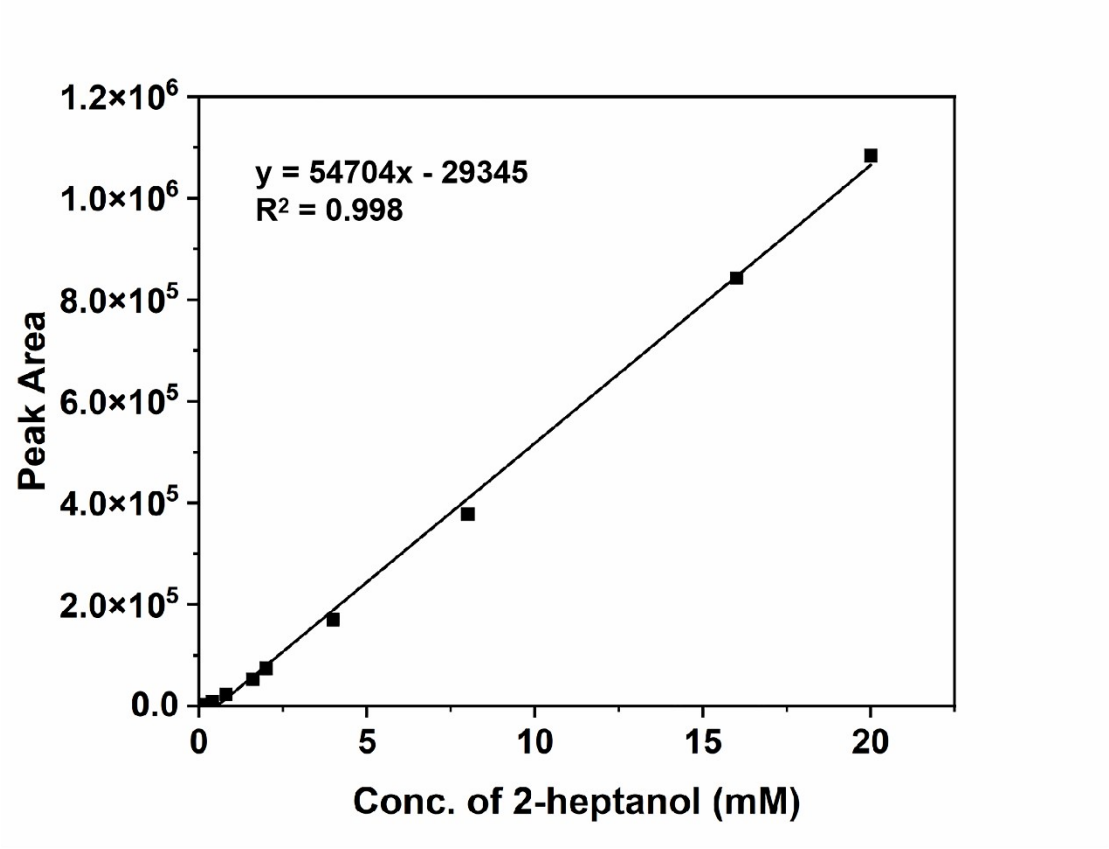


Figure S16. Standard curve for the quantification of 2-heptanol (3c). (Agilent HP-5, temperature conditions: initial temperature 40 ℃, holding 1 min, then 10 ℃/min to 120 ℃, then 30 ℃/min to 250 ℃, holding 5 min).

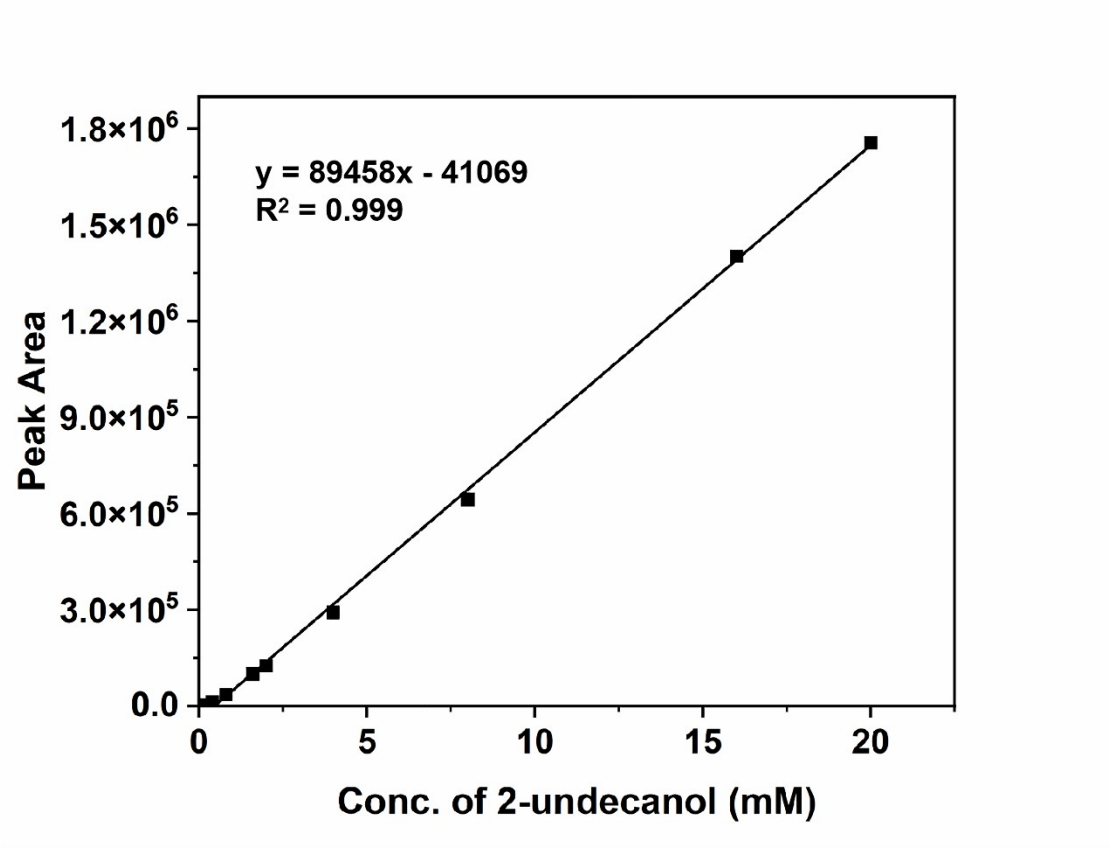


Figure S17. Standard curve for the quantification of 2-undecanol (4c). (Agilent HP-5, temperature conditions: initial temperature 40 ℃, holding 1 min, then 10 ℃/min to 120 ℃, then 30 ℃/min to 250 ℃, holding 5 min).

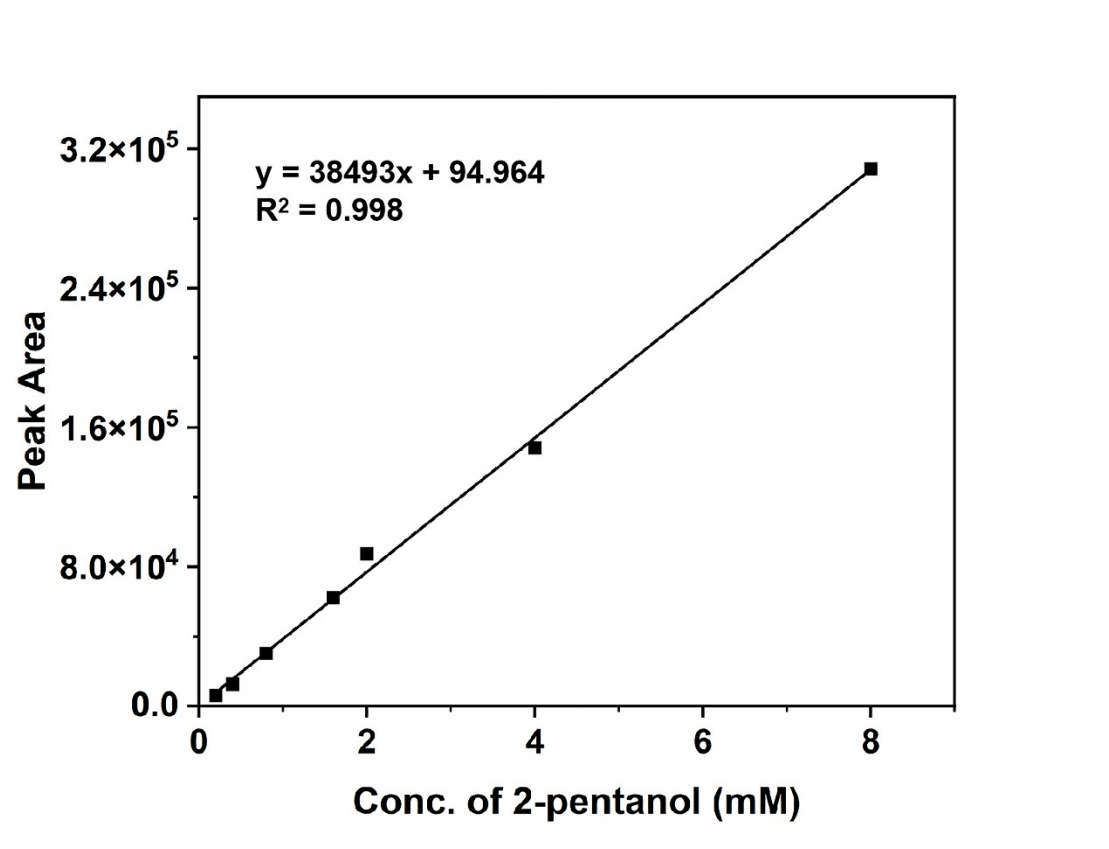


Figure S18. Standard curve for the quantification of 2-pentanol (2c). (Agilent CYCLOSIL-B, temperature conditions: initial temperature 40 ℃, holding 10 min, then 2 ℃/min to 80 ℃, then 30 ℃/min to 210 ℃, holding 5 min).

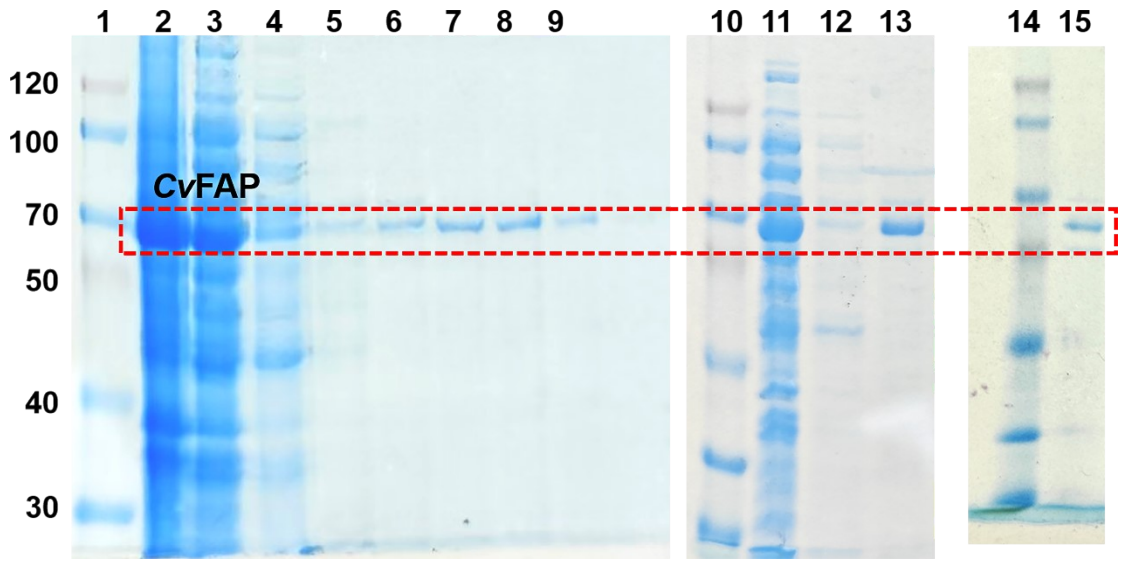


Figure S19. SDS-PAGE analysis of purified WT and variants enzymes of *Cv*FAP. Lane 1, 10 and 24: protein markers. Lane 2-9: purified enzyme of WT, 2: filtered cell extract of WT crude enzyme, 3: flow-through of WT, 4: 50 mM imidazole eluate, 5: 100 mM imidazole eluate, 6: 150 mM imidazole eluate, 7: 200 mM imidazole eluate, 8: 250 mM imidazole eluate, 9: 300 mM imidazole eluate. Lane 11-13: purified enzyme of V463A, 11: flow-through of V463A, 12: 50 mM imidazole eluate, 13: 300 mM imidazole eluate. 15: 300 mM imidazole eluate of V463T.

Experimental Section

Materials.

All the reagents were purchased from commercial suppliers (*β*-hydroxydecanoic acid, *β*-hydroxycaproic acid, *β*-hydroxyoctanoic acid and *β*-hydroxydodecanoic acid were purchased from Macklin, other reagents were purchased from Adamas).

Strains and chemicals.

The sequence of *Cv*FAP was commercially synthesized (GENEWIZ, Wuxi, China). All plasmids were from our laboratory and expressed in Escherichia. coli BL21 (DE3). Primers used in polymerase chain reaction (PCR) in this work are listed in Table S5. Reagents for PCR were obtained from Takara Biotechnology (Otsu, Japan). Prime Star Max Polymerase used in QuikChange site-directed mutagenesis method was obtained from Takara Biotechnology (Otsu, Japan). *E. coli* strains were grown in Luria-Bertani (LB) medium at 37 °C. Antibiotics were added as kanamycin (50 μg/mL).

Mutant Construction.

Site-directed mutagenesis was performed via whole-plasmid amplification using PrimeSTAR DNA Polymerase (Takara, Japan). The PCR mixture (20 μL) contained 10 μL of enzyme premix, 0.5 μM each of forward and reverse primers (1 μL each) and 40 ng of plasmid template. The amplification protocol was as follows: initial denaturation at 98 °C for 10 min; 25 cycles of denaturation at 98 °C for 30 s, annealing at a temperature determined by the primer Tm for 30 s, and extension at 72 °C for 1 min 20

s; followed by a final extension at 72 °C for 10 min, and then cooling to 16 °C for 10 min. After amplification, the reaction mixture was digested with DpnI restriction enzyme (Takara, Japan) to remove the methylated template DNA. The resulting amplified plasmid was transformed into competent *E. coli* cells, which were then spread on LB agar plates containing kanamycin (50 mg/L). The plates were incubated at 37 °C for 12 h. Single colonies were picked and the mutation sites were verified by DNA sequencing. Successfully mutated strains were stored for further use.

Protein Expression.

The strain pET28b-*Cv*FAP (BL21) provided by the company was streaked onto LB agar plates containing kanamycin at a final concentration of 50 mg/L. The plates were incubated overnight (12 h) in a 37 °C incubator. After 12 h of incubation, single colonies were picked from the plates, with 3 biological replicates selected. Each replicate was inoculated into a culture tube containing 4 mL of LB medium supplemented with kanamycin (final concentration 50 mg/L) to prepare seed cultures of the three strains. Following 12 h of seed culture, each seed culture was inoculated at a 2% (v/v) ratio into 100 mL of LB medium containing kanamycin (final concentration 50 mg/L). The inoculated bacterial cultures were incubated in a 37 °C shaking incubator until the optical density at 600 nm (OD_{600}) reached 0.6. Isopropyl *β*-D-1-thiogalactopyranoside (IPTG) was then added to the cultures at a final concentration of 0.1 mM (using a 200 mM stock solution), and the cultures were transferred to a 17 °C

shaking incubator for induction over 18 h. After induction, 40 mL of the bacterial culture was centrifuged at 8000 rpm for 10 min to remove the supernatant. The bacterial pellet was resuspended in 10 mL of Tris-HCl buffer. The resuspended bacterial suspension was subjected to cell disruption using an ultrasonic homogenizer under the following conditions: power 360 W, with a total disruption duration of 10 min. The disrupted suspension was centrifuged at 8000 rpm for 30 min at 4 °C, and the supernatant was collected for subsequent experiments (Fig. S1).

Calculation of Binding Free Energies.

The binding free energy was estimated using the MM-PBSA/GBSA approach based on molecular dynamics (MD) trajectories. All calculations were performed using GROMACS 2023 with the Intel 2021 compiler and IntelMPI. The analysis procedure is as follows: First, an index file was generated to define the protein (residues 1–586), the ligand (residue 587), and the complex. Subsequently, topology files and corresponding trajectory files were extracted for the ligand (state B), the protein (state A), and the complex (state AB) separately. Potential energies were then computed for each state by rerunning the trajectories with the mdrun module. Finally, the relevant energy terms (including Coulombic, Lennard-Jones, and potential energies) were extracted for the ligand, protein, and complex using the energy tool. The binding free energy was derived by combining the single-trajectory energies of the three states.

Statistical Analysis

All experiments were performed in triplicate (n = 3 independent replicates), and results are presented as mean ± standard deviation (SD). Statistical significance analysis for the data was carried out using GraphPad Prism (version 8.0.1; GraphPad Software, San Diego, CA, USA). One-way analysis of variance (ANOVA) was conducted, followed by Dunnett's multiple comparisons test, in which the mean value of each variant group was compared with that of the wild-type (WT) control group. P-value < 0.05 was considered statistically significant. The corresponding significance levels are indicated in the figures as follows: * for $P < 0.05$, ** for $P < 0.01$, *** for $P < 0.001$, and **** for $P < 0.0001$; "ns" denotes not significant ($P \geq 0.05$).

Reference

(1) Pianaro, A. M., C.; Kerr, W. E. ; Singer, R. B.; Patricio, E. F. L. R. A. ; Marsaioli, A. J. Stingless Bees: Chemical Differences and Potential Functions in Nannotrigona testaceicornis and Plebeia droryana Malesand Workers. *J. Chem. Ecol.* **2009**, *35*, 1117–1128.