

Supplementary Information for

Ultrahigh-Throughput Evolution of a Photoenzyme for Efficient Bioproduction of Toluene and *p*-Cresol

Zhaoyang Qin¹, Yi Zhou^{1*}, Shenhao Peng¹, Wen Zheng¹, Zherui Xu¹, Yunpeng Bai²,
Matthias Höhne³, Thomas R. Ward⁴, Uwe T. Bornscheuer^{5*} & Shuke Wu^{1,5*}

¹ National Key Laboratory of Agricultural Microbiology, College of Life Science and Technology, Huazhong Agricultural University, Wuhan, China

² Shaanxi R&D Centre of Biomaterials and Fermentation Engineering, School of Chemical Engineering, Northwest University, Xi'an, China

³ Institute of Chemistry, Technical University of Berlin, Berlin, Germany

⁴ Department of Chemistry, University of Basel, Basel, Switzerland

⁵ Department of Biotechnology & Enzyme Catalysis, Institute of Biochemistry, University of Greifswald, Greifswald, Germany

E-mail: zhouyi1213@mail.hzau.edu.cn; uwe.bornscheuer@uni-greifswald.de;
shukewu@mail.hzau.edu.cn

Table of Contents

Supplementary Figures	S3
Supplementary Tables	S23
Supplementary Methods	S27
General information	S27
Construction of the XylR biosensor	S27
Preparation of competent cells containing the biosensor	S27
Transformation of the mutant library	S28
Pilot experiment for flow cytometric sorting.....	S28
Activity validation of FACS-sorted cells	S29
Optimization of the reaction pH	S29
Calculation of activity and TONs of CvFAP.....	S30
Fermentative production of 1 and 6	S30
Computational details.....	S31
Molecular docking simulations	S31
Molecular dynamics simulations	S32
Calculation of enzyme channels and catalytic pockets.....	S32
Docking and metadynamics simulations.....	S33
Key sequence information.....	S33
Supplementary References.....	S44

Supplementary Figures

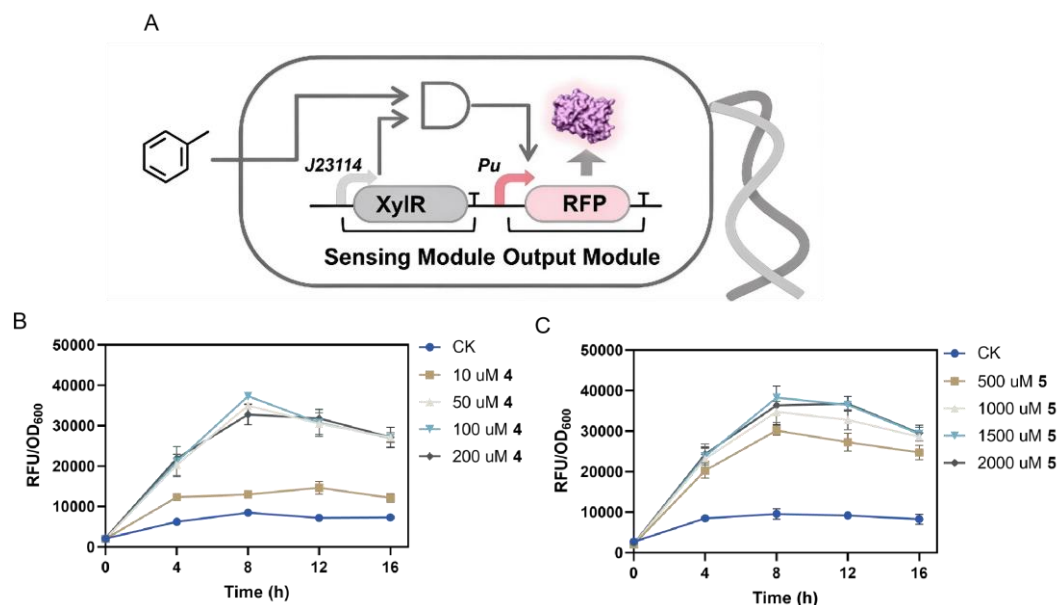


Figure S1. Construction and test of XylR-based biosensor. **A:** Design of the XylR-based biosensor. The reporting gene is the RFP (mcherry), constitutive promoter BBA_J23114 is used to drive the expression of the XylR transcription factor. **B:** Response of XylR to **4** in different concentrations (100 mM stock solution of **4** was prepared in KP buffer, 0.2 M, pH = 8.0). **C:** Response of XylR to **5** in different concentrations (100 mM stock solution of **5** was prepared in DMSO, after induction, the culture was transferred to a shaker at 30 °C to minimize the evaporation of **5**). Data in **B** and **C** are means $\pm$ s.d. of three technical replicates.

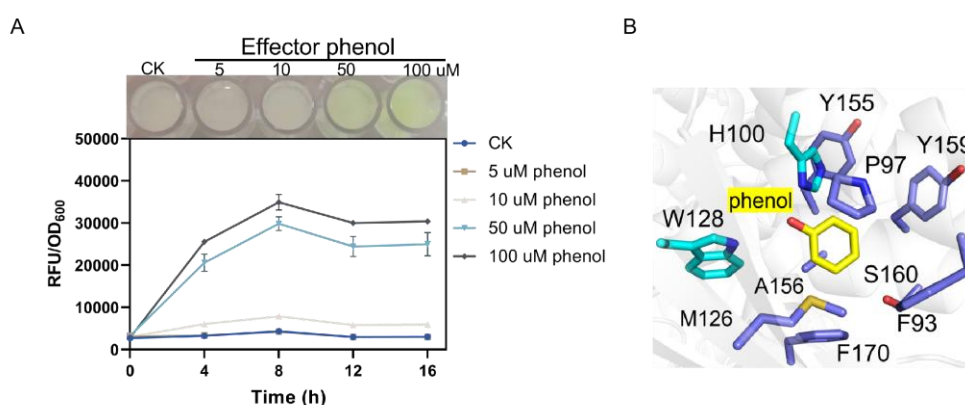


Figure S2. Study of DmpR-based biosensor to native effector phenol. **A:** Response of DmpR to phenol in different concentrations (100 mM stock solution of phenol was prepared in DMSO). **B:** Phenol is docked into the LBD domain of DmpR (phenol, yellow stick; key residues H100 and W128, cyan stick). Data in **A** are means $\pm$ s.d. of three technical replicates.

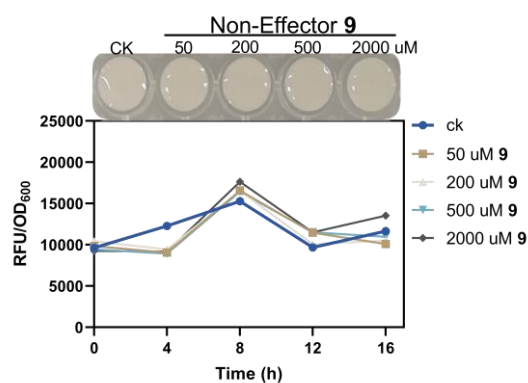


Figure S3. Study of DmpR-based biosensor toward 9. Response of DmpR to **9** in different concentrations (100 mM stock solution of **9** was prepared in DMSO).

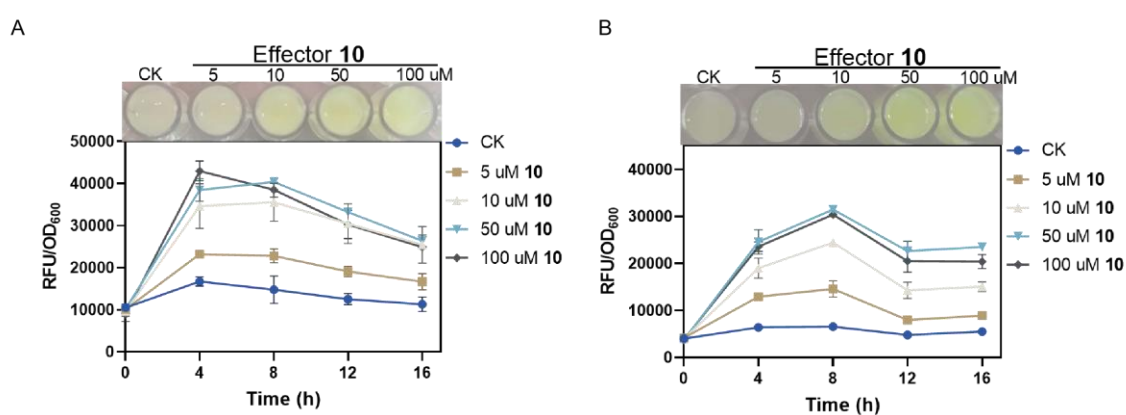


Figure S4. Study of other DmpR mutants toward 10. **A:** Response of DmpR-M52I to **10** in different concentrations. **B:** Response of DmpR-M52I-K188R to **10** in different concentrations. Data are means $\pm$ s.d. of three technical replicates.

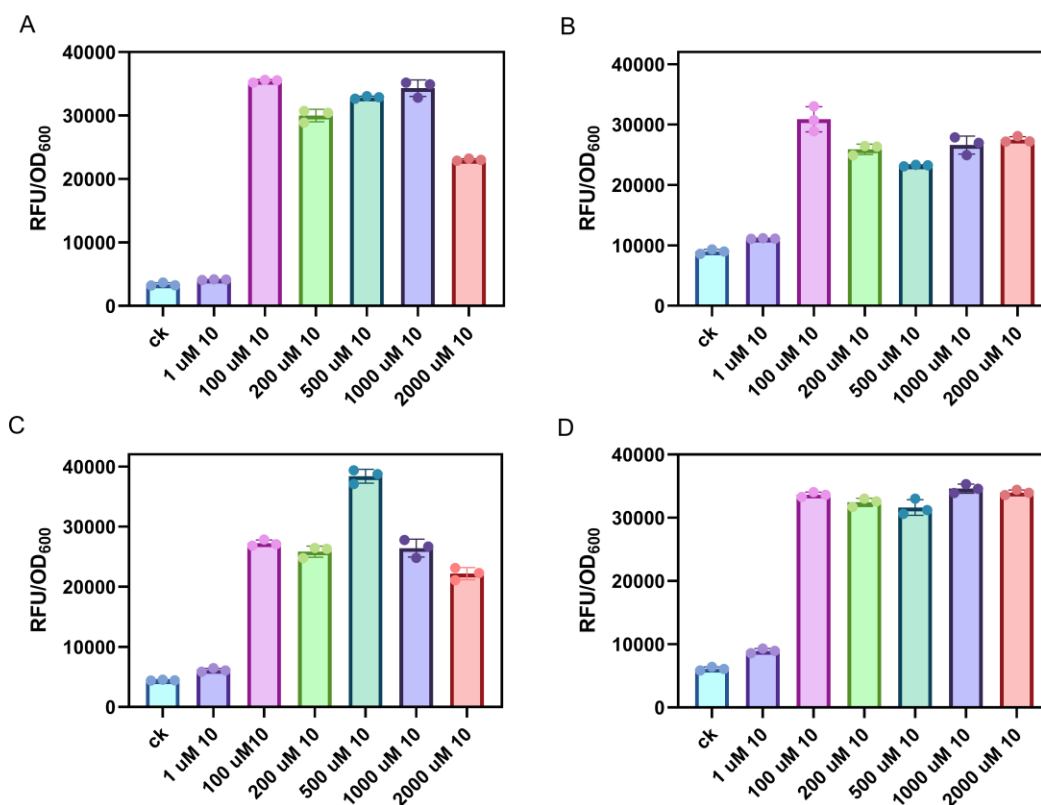


Figure S5. Operational range of the DmpR biosensors toward 10. **A:** Response range of DmpR-WT to varying concentrations of 10 (activation concentration of 10: 1–2000 μ M, 200 mM stock solution of 10, prepared in DMSO). **B:** Response range of DmpR-M52I to varying concentrations of 10. **C:** Response range of DmpR-M52I-K188R to varying concentrations of 10. **D:** Response range of DmpR-M52I-Q71A to varying concentrations of 10. The entire activation assay was performed in 25 mL reaction vials. When the OD₆₀₀ of *E. coli* culture reached 0.8, effector at different concentrations were added, followed by incubation at 30 °C. 100 mM stock solution of 10 was prepared in DMSO. Data are means $\pm$ s.d. of three technical replicates.

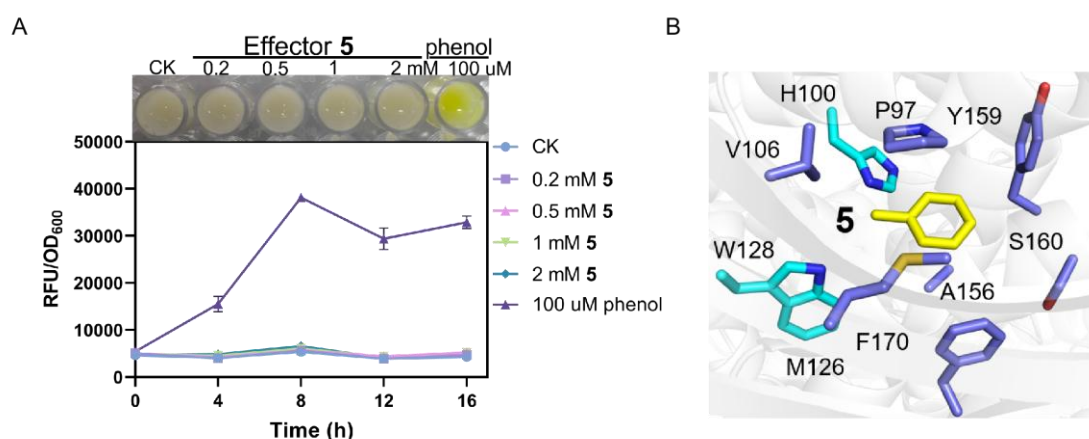


Figure S6. Study of DmpR-based biosensor toward 5. **A:** Response of DmpR to 5 in different concentrations (100 mM stock solution of 5 was prepared in DMSO, phenol was used as the control). **B:** 5 was docked into the LBD domain of DmpR (5, yellow sticks; key residues W128 and H100, cyan sticks). Data in **A** are means $\pm$ s.d. of three technical replicates.

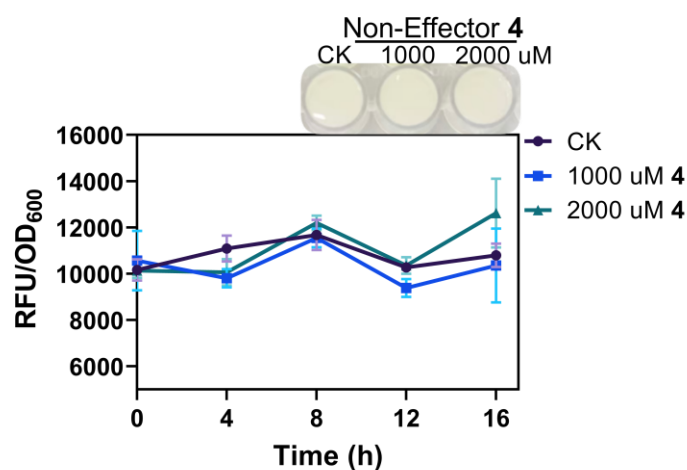


Figure S7. Study of DmpR-WT biosensor toward 4. Response of DmpR-WT to **4** in different concentrations (100 mM stock solution of **4** was prepared in DMSO).

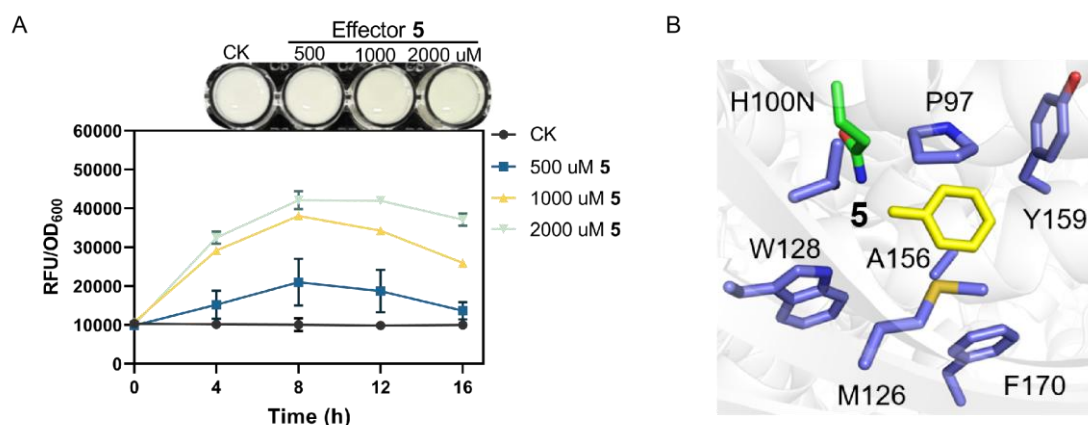


Figure S8. Study of DmpR-H100N biosensor toward 5. **A:** Response of DmpR-H100N to **5** in different concentrations (100 mM stock solution of **5** was prepared in DMSO). **B:** **5** was docked into the LBD domain of DmpR-H100N (the structure of DmpR-H100N was generated by homology modeling using DmpR-WT as the template; **5**, yellow sticks; key residue H100N, green sticks). Data in **A** are means $\pm$ s.d. of three technical replicates.

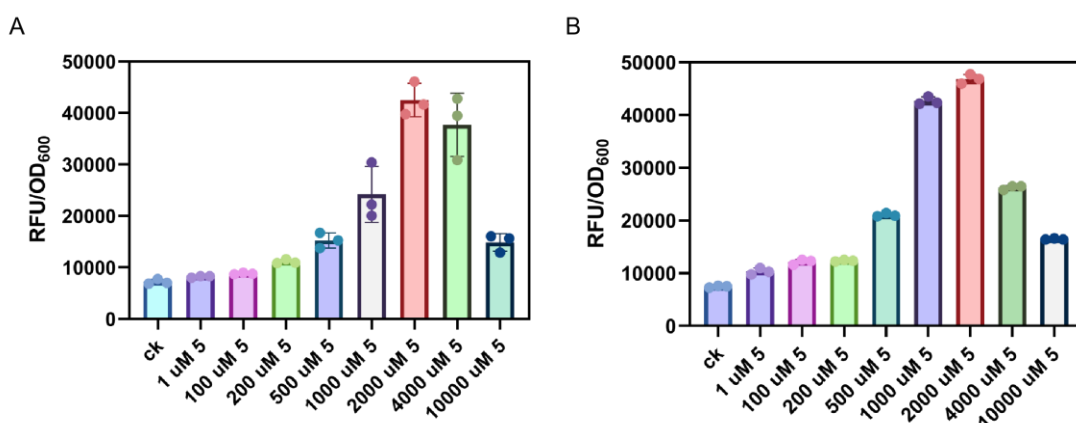


Figure S9. Operational range of the DmpR biosensors to 5. **A:** Response range of DmpR-H100N to varying concentrations of **5** (activation concentration of **5**: 1–10000 μ M, 200 mM stock solution of **5**, prepared in DMSO). **B:** Response range of DmpR-H100E to

varying concentrations of **5** (activation concentration of **5**: 1–10000 μM). The entire activation assay was performed in 25 mL reaction vials. When the OD₆₀₀ of *E. coli* culture reached 0.8, effector at different concentrations were added, followed by incubation at 30 °C. 100 mM stock solution of **5** was prepared in DMSO, Data are means $\pm$ s.d. of three technical replicates.

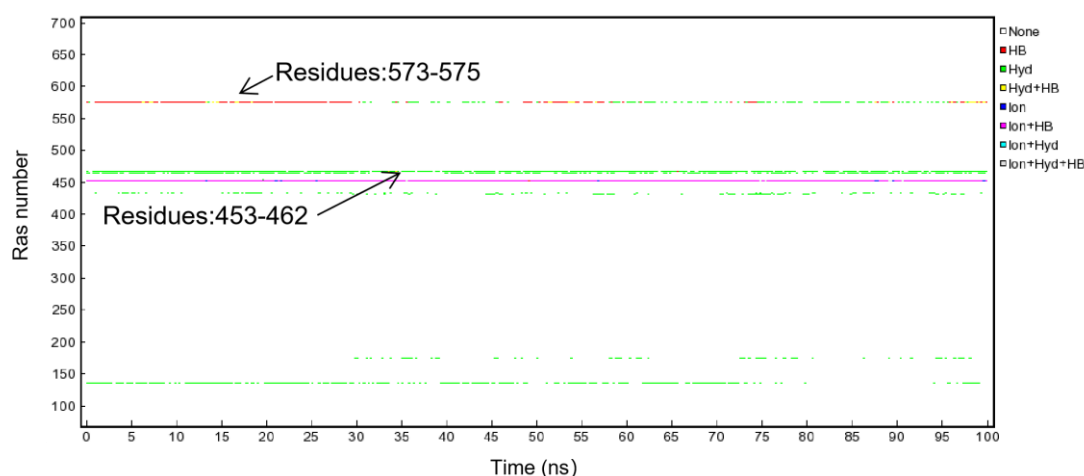


Figure S10. Per-residue ligand 4 interactions of the receptor CvFAP-WT. The plot shows the types of interactions by receptor residues with the ligand as a function of simulation time, three types of interactions are shown: Hydrogen bonds [red], hydrophobic [green] and ionic interactions [blue].

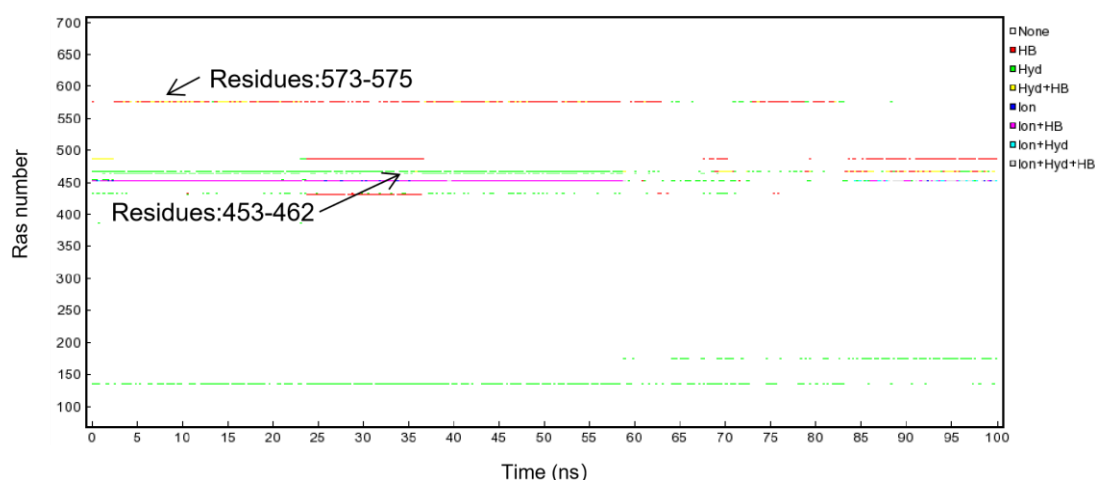


Figure S11. Per-residue ligand 9 interactions of the receptor CvFAP-WT. The plot shows the types of interactions by receptor residues with the ligand as a function of simulation time, three types of interactions are shown: Hydrogen bonds [red], hydrophobic [green] and ionic interactions [blue].

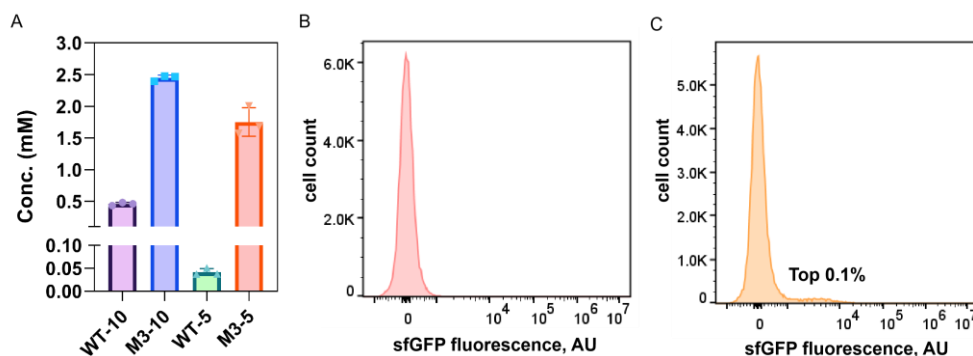


Figure S12. A pilot assay of uHTS of CvFAP. **A:** Production of **10** and **5** by CvFAP-WT and CvFAP-M3 (P460A-G462I-Y466T) catalyzed decarboxylation of 10 mM **9** and **4**, respectively. **B:** Fluorescence distribution of *E. coli* cells containing *p*-cresol sensor and CvFAP-WT after incubation with **9** (4 mM) under blue light illumination for 4 h. **C:** Fluorescence distribution of a mixture of *E. coli* cells containing *p*-cresol sensor and CvFAP-WT and *E. coli* cells containing *p*-cresol sensor and CvFAP-M3 (1000:1) after incubation with **9** (4 mM) under blue light illumination for 4 h. Data in **A** are means $\pm$ s.d. of three technical replicates.

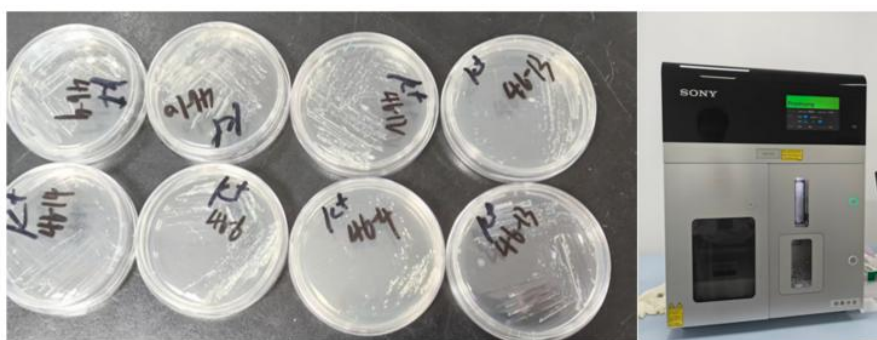


Figure S13. Isolation of cells collected in the first-round FACS. Photographs of cells on the plates after first round of sorting (for the further validation via single colony isolation, cell culture and expression, whole-cell photodecarboxylation and GC analysis) and the FACS equipment used.

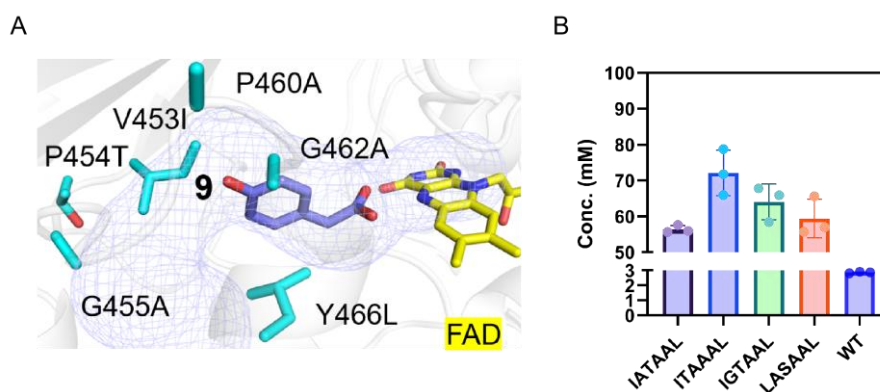


Figure S14. Study of the mutants obtained from 2nd round of evolution toward **9.** **A:** **9** was docked into the substrate-binding pocket of CvFAP mutant ITAAAL. **9**, purple stick models; mutated amino acids, cyan stick models. **B:** Production of **10** from 100 mM **9** by the second-round mutants. Reaction conditions: 0.2 M KB buffer pH = 6.0, 4 h, 22°C, 100 mM **9**, whole-

cell catalyst 20 g/L. Data in **B** are means $\pm$ s.d. of three technical replicates.

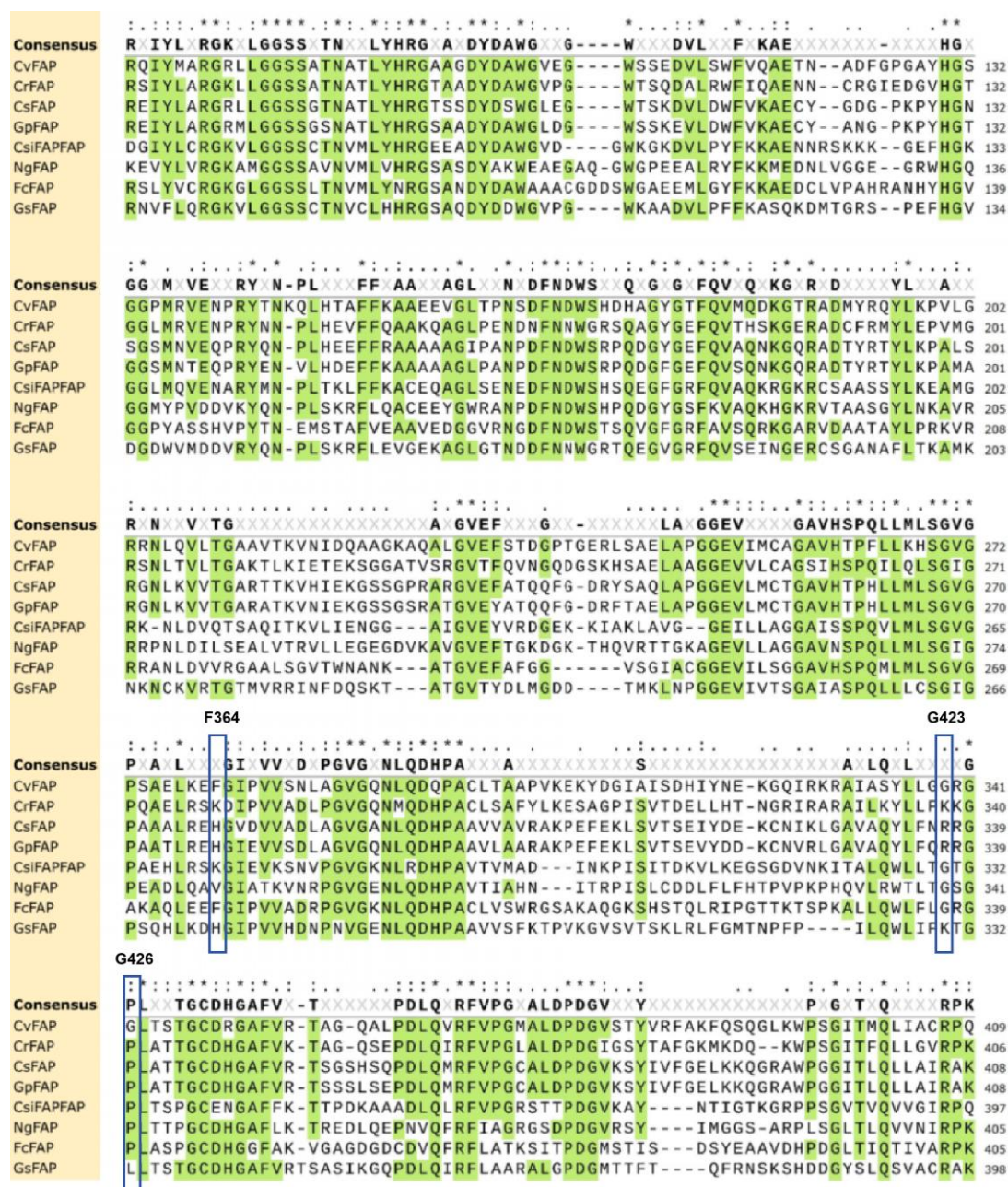


Figure S15. Multiple sequence alignment of several CvFAP homologues. CvFAP from *Chlorella variabilis* (A0A248QE08); CrFAP from *Chlamydomonas reinhardtii* (A8JHB7); CsFAP from *Coccomyxa subellipsoidea* strain C-169 (I0YJ13); GpFAP from *Gonium pectorale* (A0A150GC51); CsiFAP from *Chlamydomonas schloesseri* (A0A836BBQ2); NgFAP from *Nannochloropsis gaditana* (W7TN63); FcFAP from *Fragilariopsis cylindrus* CCMP1102 (A0A1E7FBP9); GsFAP from *Galdieria sulphuraria* (M2XJN3). The conservative amino acids 423 and 426 are marked with blue boxes. Multiple Sequence Alignment by CLUSTALW.

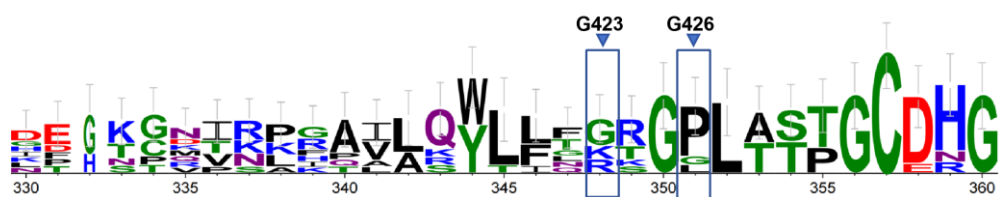


Figure S16. Sequence conservation analysis of residues G423 and G426 of CvFAP. Sequence conservation analysis of FAPs was performed using the sequences of different FAPs shown in the above figure. CvFAP from *Chlorella variabilis* (A0A248QE08); CrFAP from *Chlamydomonas reinhardtii* (A8JHB7); CsFAP from *Coccomyxa subellipsoidea* strain C-169 (I0YJ13); GpFAP from *Gonium pectorale* (A0A150GC51); CsiFAP from *Chlamydomonas schloesseri* (A0A836BBQ2); NgFAP from *Nannochloropsis gaditana* (W7TN63); FcFAP from *Fragilariopsis cylindrus* CCMP1102 (A0A1E7FBP9); GsFAP from *Galdieria sulphuraria* (M2XJN3), and the resulting figure was generated with MEGA7.

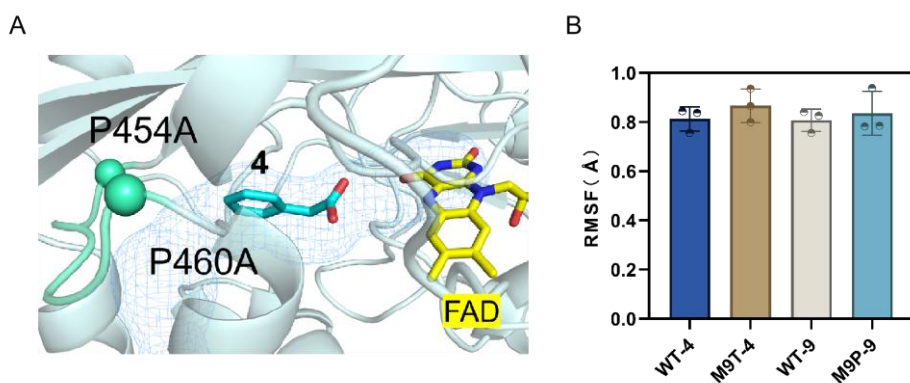


Figure S17. The loop region of residue 453-462 of M9T mutant and RMSF of the loop region. **A:** The loop region of residue 453-462 of the M9T mutant. loop 453-462, light green cartoon, the substrate channel, blue mesh model; **4**, cyan stick model; FAD, a yellow stick model. **B:** RMSF values for loop 453-462 in the WT and mutants M9T and M9P. Data in **B** are means $\pm$ s.d. of three independent MD simulations.

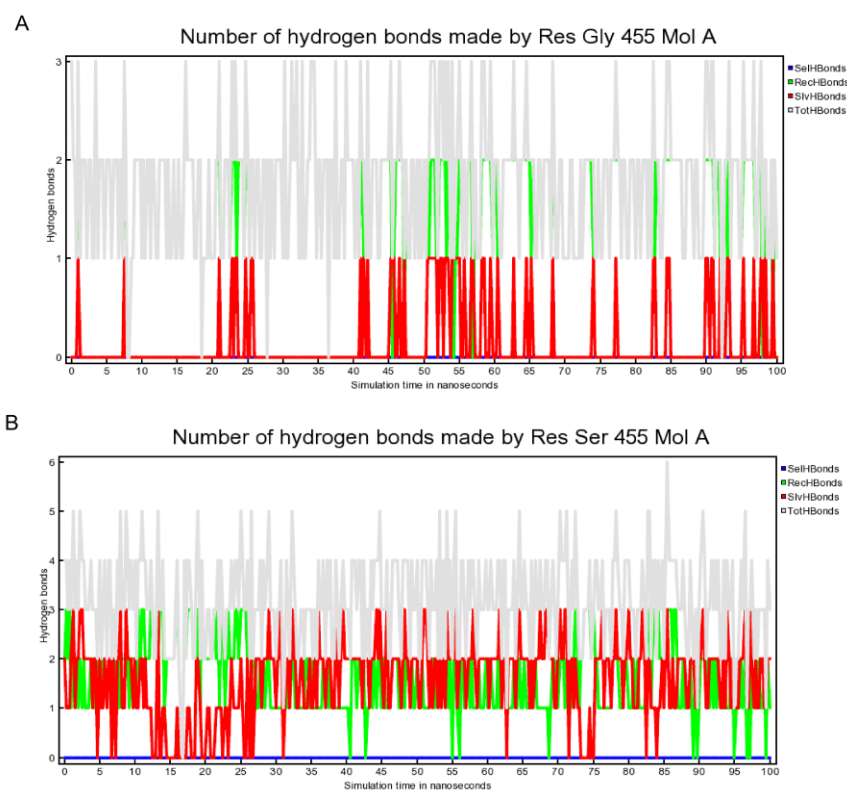


Figure S18. Calculation of hydrogen bond of G455 in CvFAP WT and S455 in M9T. A: Calculation of hydrogen bond of amino acid G455 in CvFAP-WT. **B:** Calculation of hydrogen bond of amino acid S455 in CvFAP-M9T. The plots show three kinds of hydrogen bonds: internal hydrogen bonds are in blue, hydrogen bonds with the rest of the solute are in green, those with the water are in red, and the sum of all three bonds are in grey.

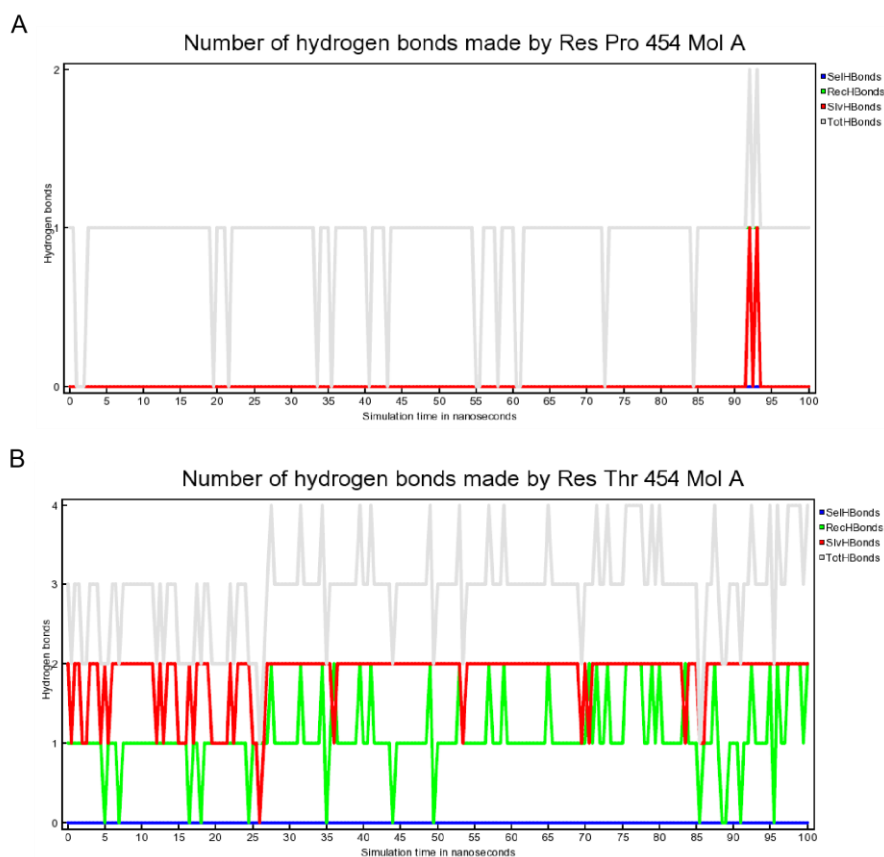


Figure S19. Calculation of hydrogen bond of P454 in CvFAP WT and T454 in M9T. A: Calculation of hydrogen bond of amino acid P454 in CvFAP-WT. **B:** Calculation of hydrogen bond of amino acid T454 in CvFAP-M9T. The plots show three kinds of hydrogen bonds: internal hydrogen bonds are in blue, hydrogen bonds with the rest of the solute are in green, those with the water are in red, and the sum of all three bonds are in grey.

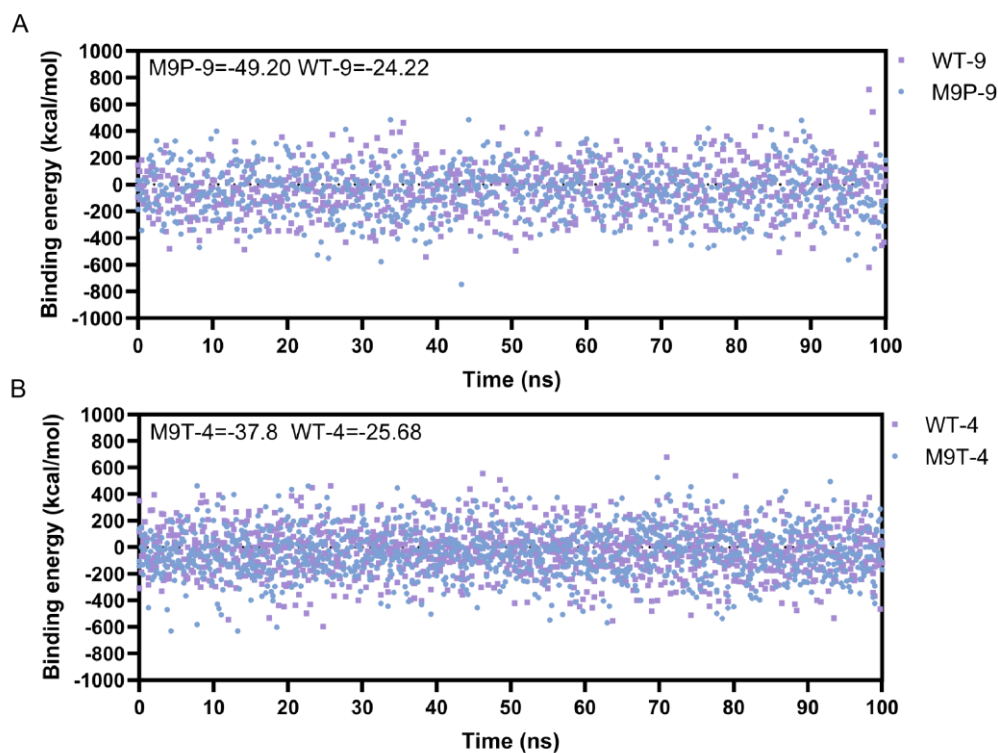


Figure S20. Calculation of the binding energy of substrate 4 and 9 within the catalytic pocket of CvFAP variants in 100-ns simulations. **A:** Binding energy of 9 within the catalytic pocket of CvFAP WT and M9P. **B:** Binding energy of 4 within the catalytic pocket of CvFAP WT and M9T.

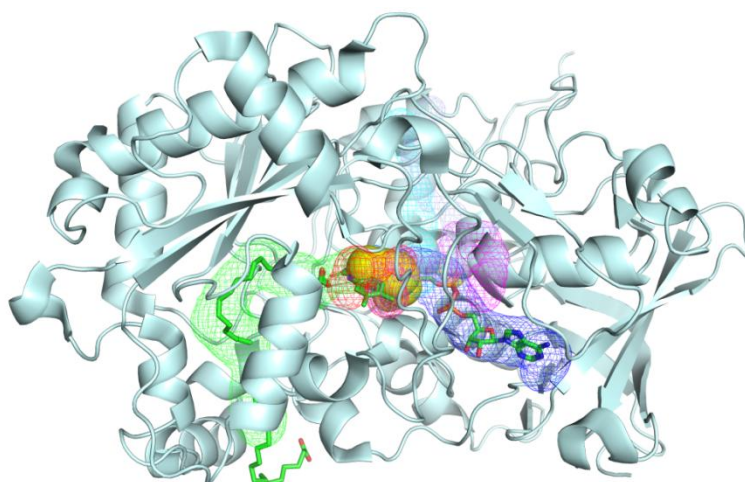


Figure S21. Potential substrate tunnels of CvFAP (PDB: 6YRU) were predicted using Caver. The major substrate tunnel is highlighted in green, and the native substrate stearic acid is in green stick model.

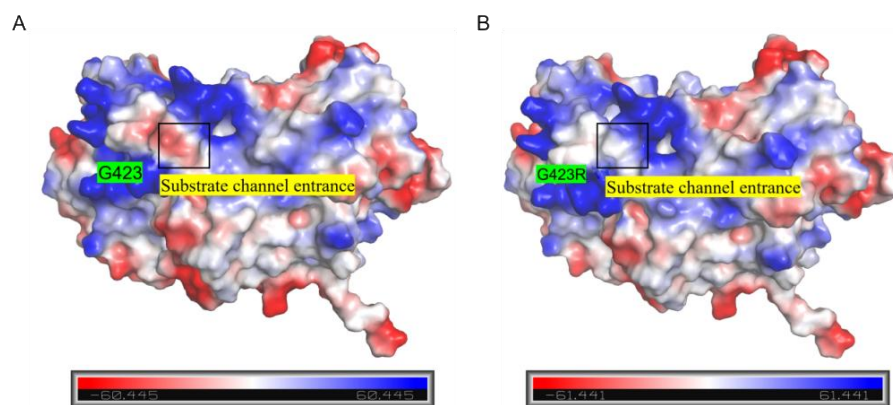


Figure S22. Electrostatic potential map of the surface of the CvFAP variants. A: Electrostatic potential map of the surface of the WT. **B:** Electrostatic potential map of the surface of the M9T. The black box indicates the entrance of the substrate channel. Negative charge is in red, and positive charge is in blue.

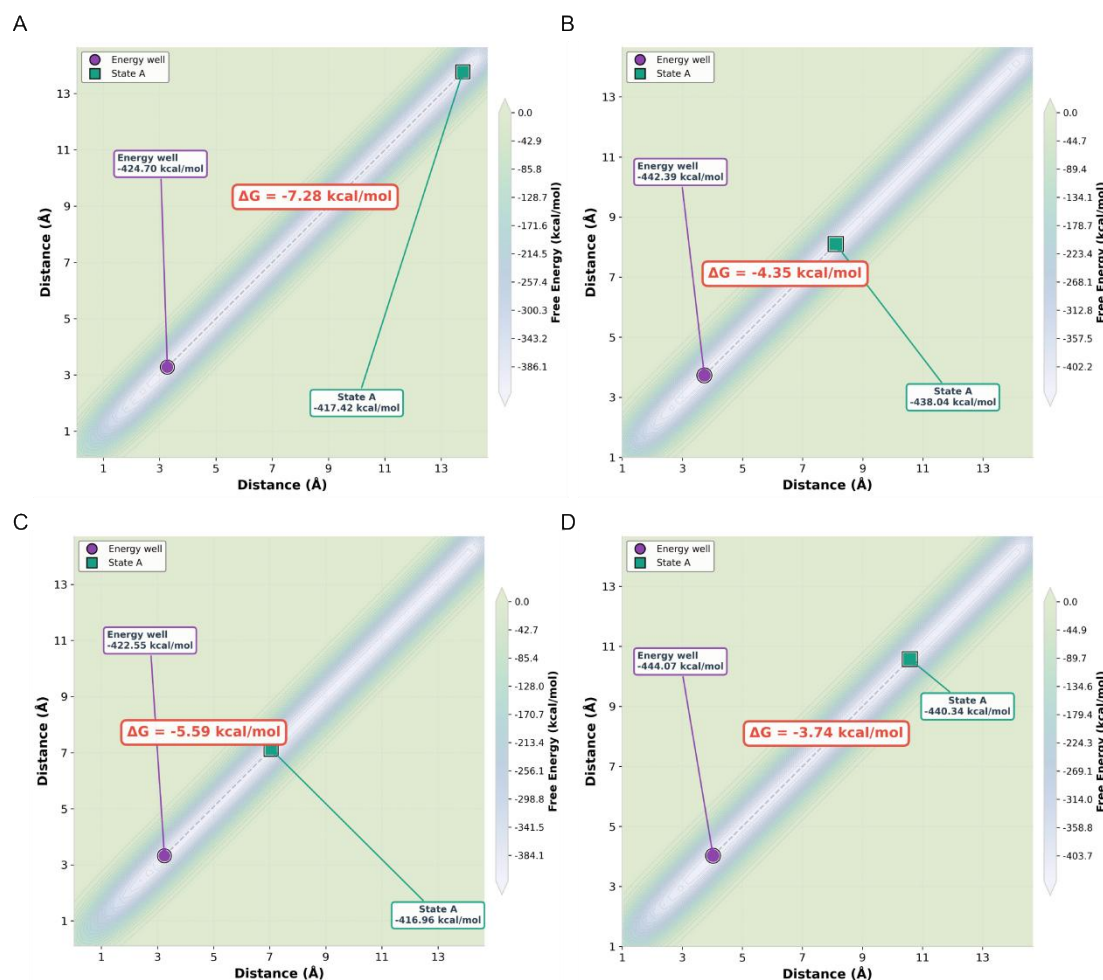


Figure S23. Energy changes during substrate entry into the tunnel calculated by SMD. A: calculated energy changes of substrate **9** from the mutant M9P to the enzyme channel entrance. **B:** Energy changes of substrate **9** from the WT to the enzyme channel entrance. **C:** Energy changes of substrate **4** from the M9T to the enzyme channel entrance. **D:** Energy changes of substrate **4** from the WT to the enzyme channel entrance. State A represents a fixed position where the substrate is docked away from the tunnel entrance (G423) of the enzyme. SMD

calculations were used to evaluate the energy change of the substrate from State A to the tunnel entrance position (G423).

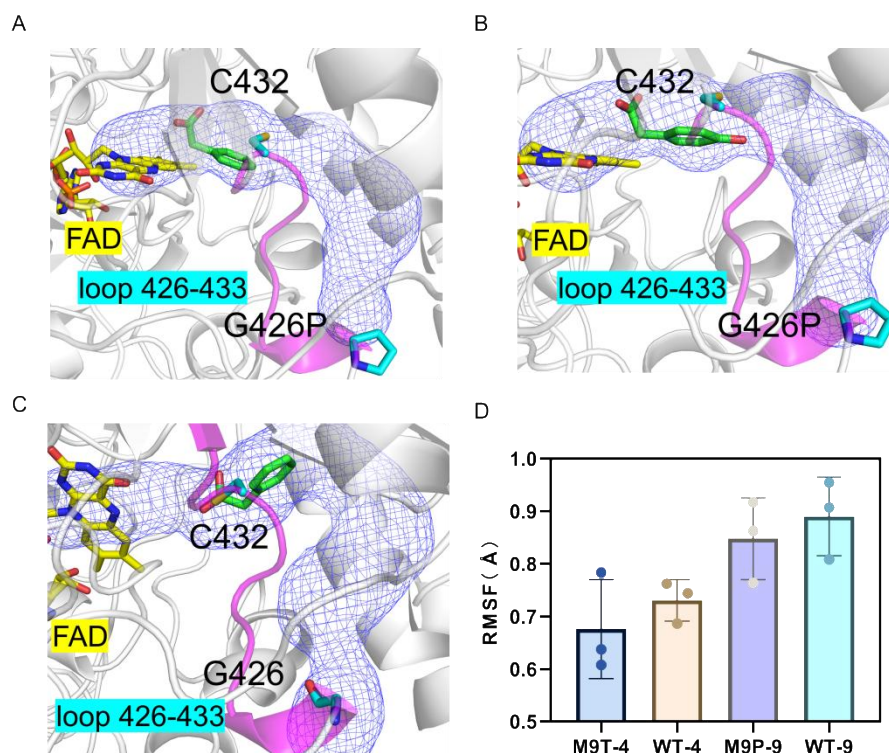


Figure S24. Analysis of the tunnel-forming loops in enzyme. **A:** Channel representation of the M9T mutant; loop 426–432 is shown as a purple cartoon model, the substrate channel is represented by a blue mesh model. **B:** Channel representation of the M9P mutant. **C:** Channel representation of the CvFAP-WT. **D:** Calculation of RMSF values for loop 426-432 in the mutants and wild type. Data in **D** are means $\pm$ s.d. of three independent MD simulations.

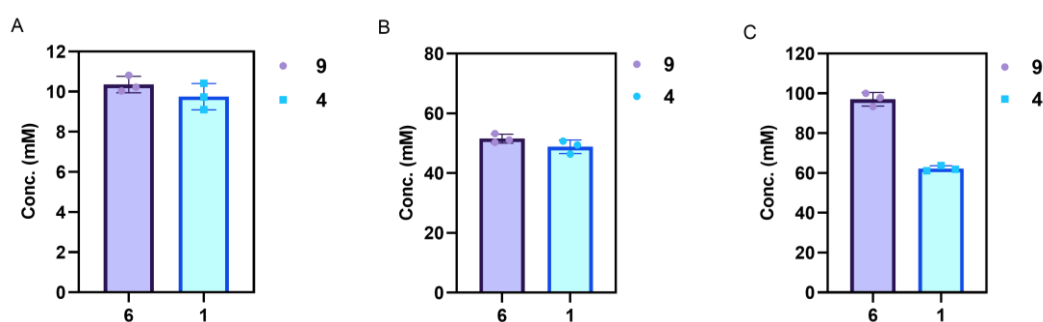


Figure S25. Production of 9 from 6 and 4 from 1 by *E. coli* co-expressing PmLAAD, ScPPDC, and EcPADH. **A:** Conversion rate of 10 mM 6 to 9 or 10 mM 1 to 4 via the three-enzyme cascade. **B:** Conversion rate of 50 mM 6 to 9 or 50 mM 1 to 4 via the three-enzyme cascade. **C:** Conversion rate of 100 mM 6 to 9 or 100 mM 1 to 4 via the three-enzyme cascade. Reaction conditions: 0.2 M KB buffer pH = 6.0, whole-cell catalysts 20 g l⁻¹, 30 °C, 2-24 h. Data are means $\pm$ s.d. of three technical replicates.

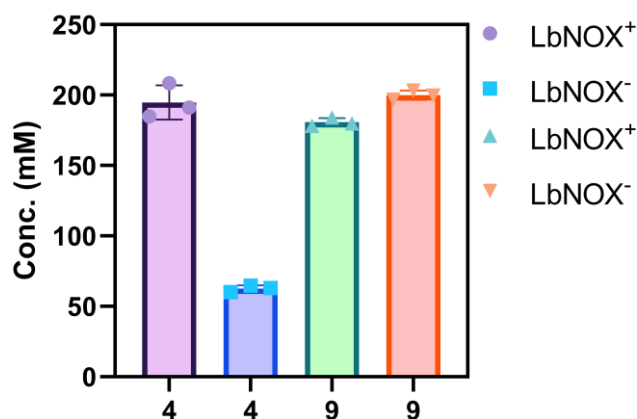


Figure S26. Production of 4 from 1 and 9 from 6 by *E. coli* co-expressing PmLAAD, ScPPDC, EcPADH, and with or without LbNOX. Reaction conditions: 100+100 mM substrate **1** or **6** (2nd feeding of 100 mM at 8 h or 12 h), 0.2 M KB buffer pH = 6.0, whole-cell catalysts 10+10 g l⁻¹ (2nd feeding of 10 g l⁻¹ at 8 h or 12 h), 30 °C, 24 h. Data are means $\pm$ s.d. of three technical replicates.

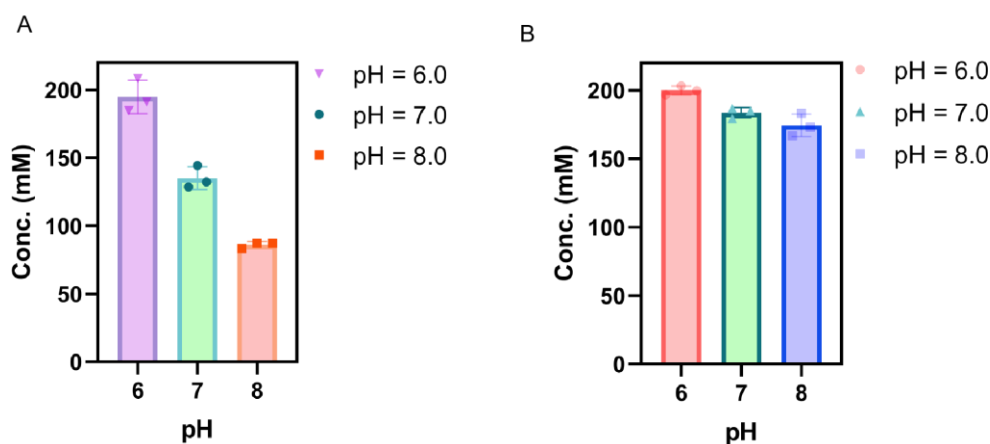


Figure S27. pH optimization of enzyme cascades for transformation of 1 and 6. A: The effects of different pH of transformation of **1** by enzyme cascade PmLAAD-ScPPDC-EcPADH-LbNOX. **B:** The effects of different pH of transformation of **6** by enzyme cascade PmLAAD-ScPPDC-EcPADH. Reaction conditions: 100+100 mM substrate **1** or **6** (2nd feeding of 100 mM at 8 h or 12 h), 0.2 M KB buffer pH = 6.0-8.0, whole-cell catalysts 10+10 g l⁻¹ (2nd feeding of 10 g l⁻¹ at 8 h or 12 h), 30 °C, 24 h. Data are means $\pm$ s.d. of three technical replicates.

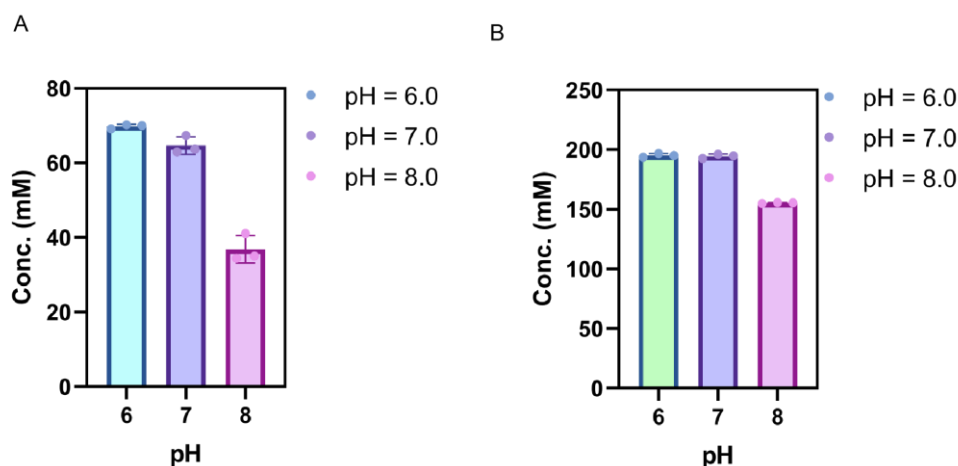


Figure S28. pH Optimization of decarboxylation by CvFAP M9T and M9P. **A:** Decarboxylation of 100 mM **4** by CvFAP M9T in KP buffer with different pH. **B:** Decarboxylation of 200 mM **9** by CvFAP M9T in KP buffer with different pH. Reaction conditions: 100 mM **4** or 200 mM **9**, 0.2 M KB buffer pH = 6.0-8.0, whole-cell catalysts 20 g l⁻¹, 22 °C, 4 h. Data are means $\pm$ s.d. of three technical replicates.

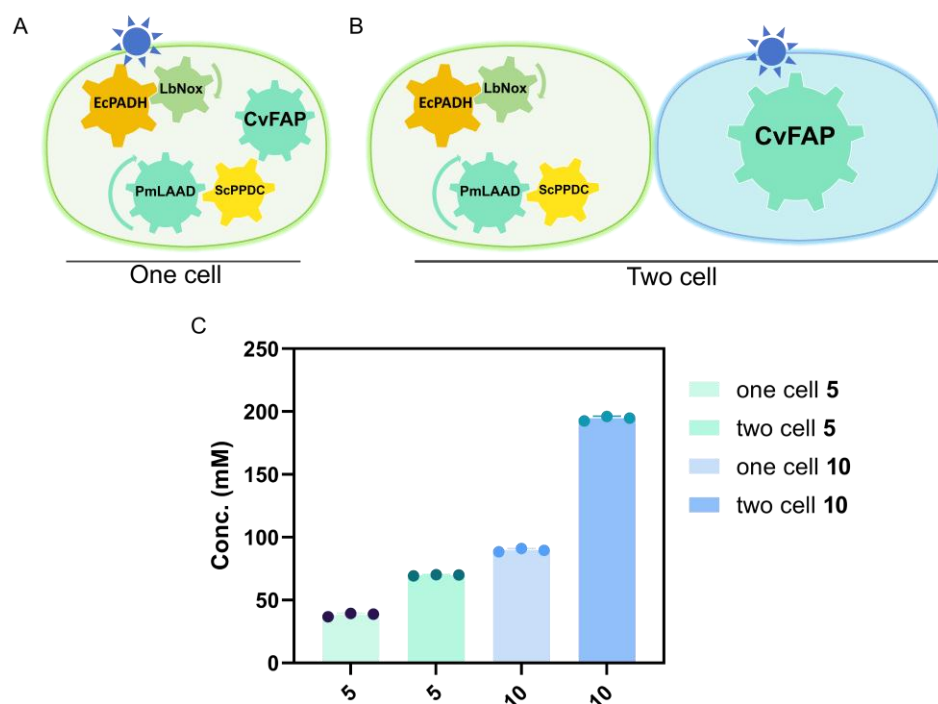


Figure S29. Comparison of one-cell and two-cell systems for cascade conversion of **1 to **5** or **6** to **10**.** **A:** one-cell system of a single *E. coli* strain harboring pRSFduet-CvFAP-M9T/M9P and pETduet-LAAD-PPDC-PADH-(LbNOX). **B:** Two-cell systems of two *E. coli* strains each harboring one plasmid of pRSFduet-CvFAP-M9T/M9P or pETduet-LAAD-PPDC-PADH-(LbNOX). **C:** Production of **5** or **10** by the one-cell and two-cell systems. To accumulate higher concentrations of **4** and **9**, substrate and whole-cell catalyst were fed in batches. One-cell reaction conditions: 100 mM **1** or 200 mM **6**, 0.2 M KB buffer pH = 6.0, whole-cell catalysts 20 g l⁻¹, 30 °C, 24 h, substrates **6** were added to the reaction system in two separate batches, 100+100 mM substrate **6** (2nd feeding of 100 mM at 12 h), the reaction was initially performed

under dark conditions, blue-light illumination was initiated 8 h after reaction onset and maintained until the reaction was complete. Two-cell reaction conditions are described in the “Multi-enzyme cascade reaction” section of the main text. Data are means $\pm$ s.d. of three technical replicates.

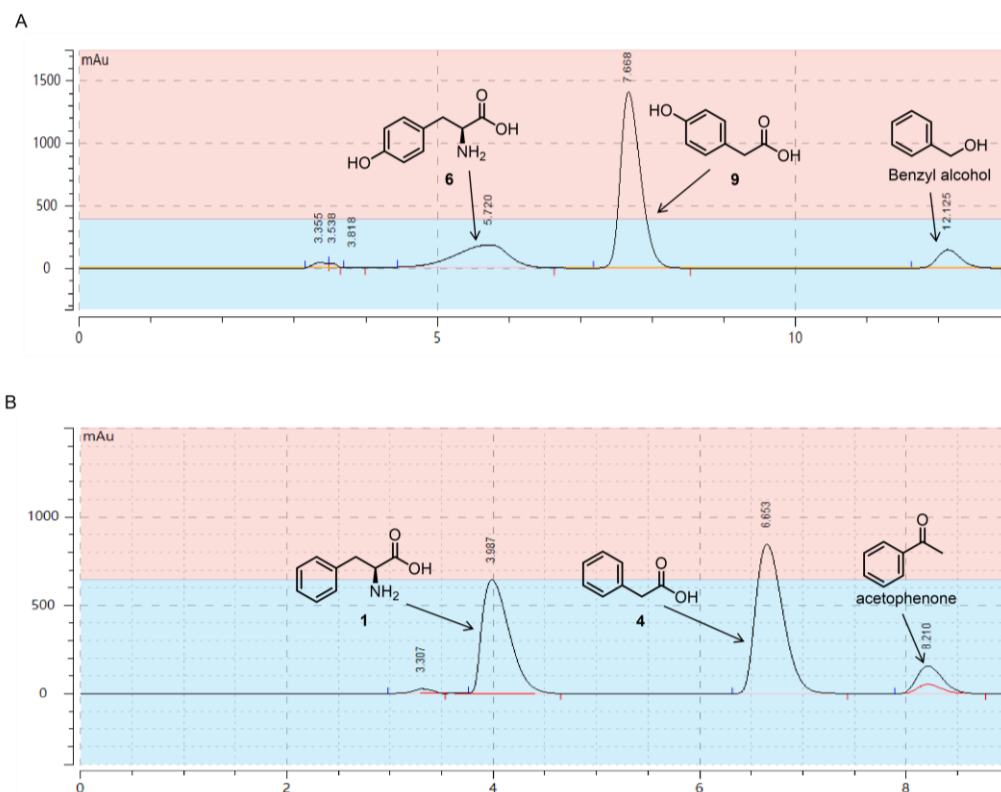
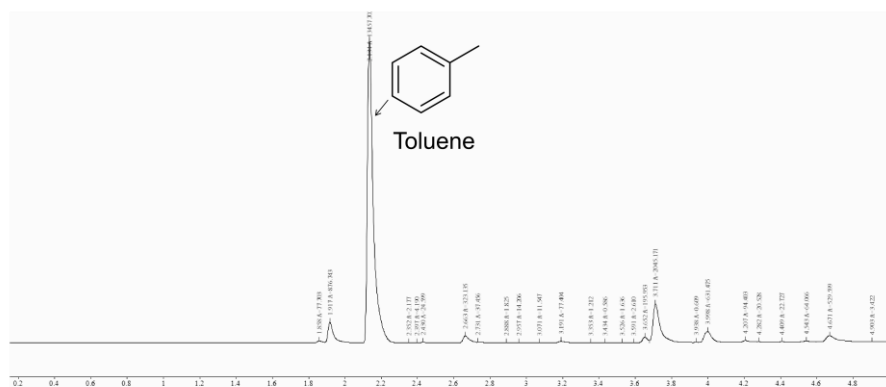


Figure S30. HPLC chromatograms of substrates and products of cascade biotransformation of 1 to 4 and 6 to 9. **A:** HPLC chromatogram of cascade biotransformation of **6** to **9**. Retention times: **6**, 5.270 min; **9**, 7.668 min; internal standard (benzyl alcohol), 12.125 min. **B:** HPLC chromatogram of cascade biotransformation of **1** to **4**. Retention times: **1**, 3.987 min; **4**, 6.653 min; internal standard (acetophenone), 8.210 min.

A



B

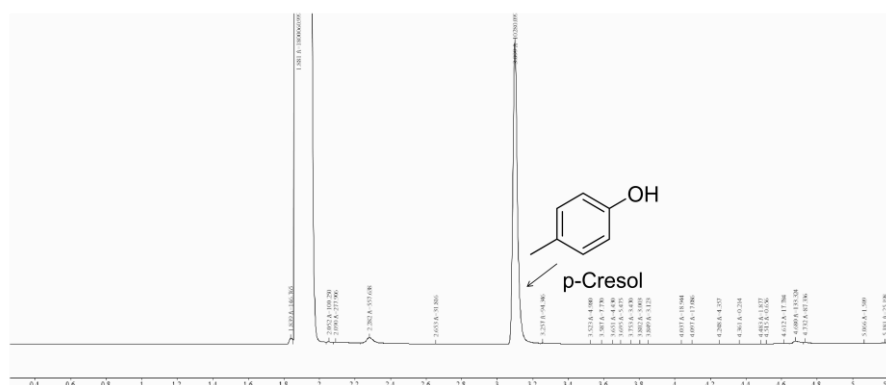


Figure S31. GC chromatograms of product toluene **5 and *p*-cresol **10**. A:** GC analysis of toluene **5** with a retention time of 2.134 min. **B:** GC analysis of *p*-cresol **10** with a retention time of 3.099 min.

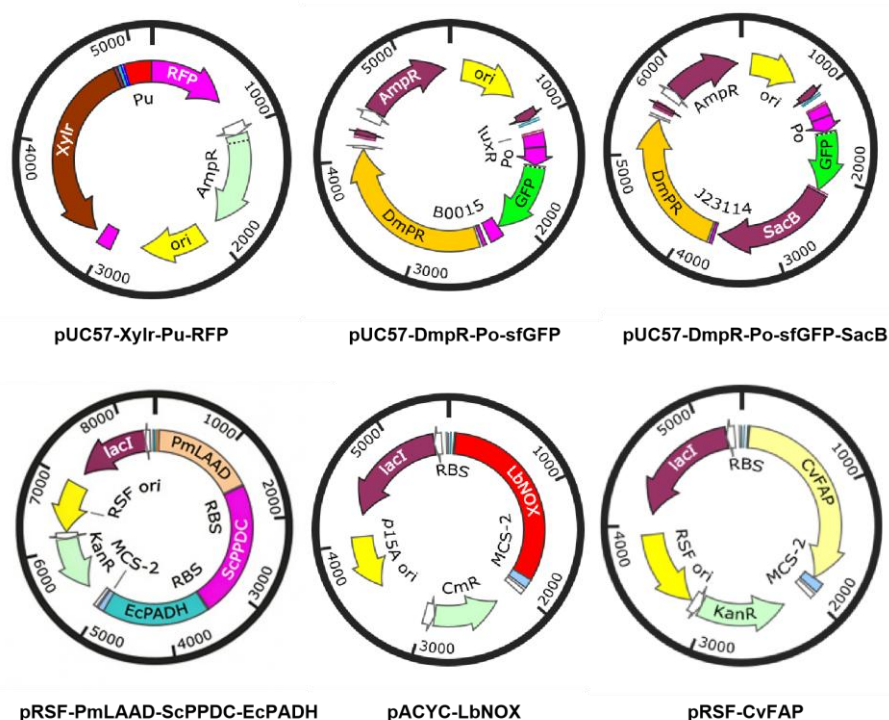


Figure S32. Key plasmid maps used in this study, including the biosensor plasmids, the multienzyme cascade plasmids, and the CvFAP plasmid.

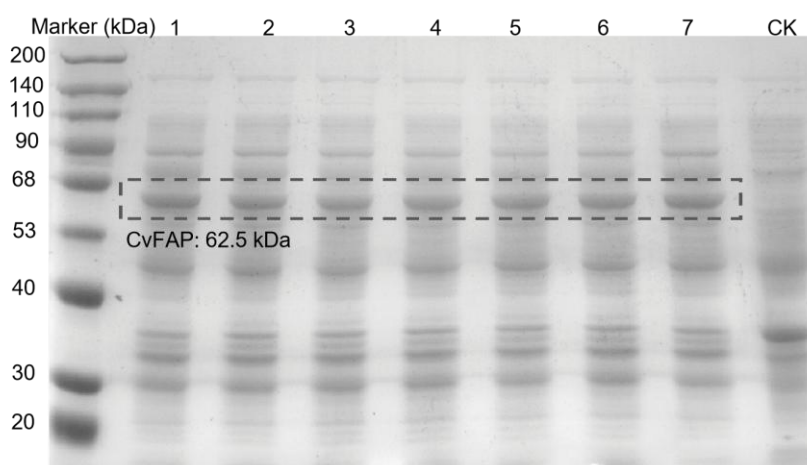


Figure S33. SDS-PAGE analysis of whole-cell protein of *E. coli* expressing CvFAP or its mutants. Lane 1: CvFAP-WT; Lane 2: CvFAP-AAL; Lane 3: CvFAP-VAT; Lane 4: CvFAP-ITAAAL; Lane 5: CvFAP-IASAAL; Lane 6: CvFAP-M9P; Lane 7: CvFAP-M9T; CK: *E. coli* cells containing the empty plasmid.

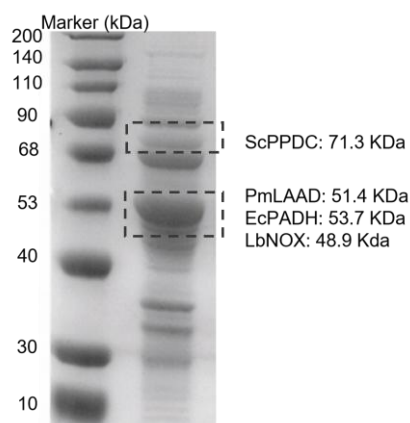


Figure S34. SDS-PAGE analysis of whole-cell protein of *E. coli* co-expressing PmLAAD, ScPPDC, EcPADH, LbNOX. Lanes 1: PmLAAD-ScPPDC-EcPADH-LbNOX (the upper larger band corresponds to ScPPDC, and the lower bands correspond to PmLAAD, EcPADH, and LbNOX). CK: *E. coli* cells containing the empty plasmid.

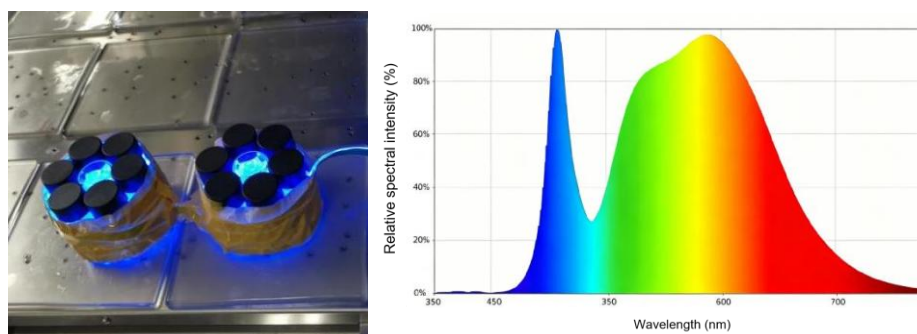


Figure S35. Photobiocatalytic set-up with blue LED strip and spectrum parameters of the LED strip. The shaking incubator was operated at 200 rpm and the temperature was maintained at 22 °C); Blue light strip parameters: PeakWave $\lambda_p=458.2$ nm, CenterWave $\lambda_c = 0.0$ nm, DeltaWave Amp = 24.0 nm, LED strip datasheet.

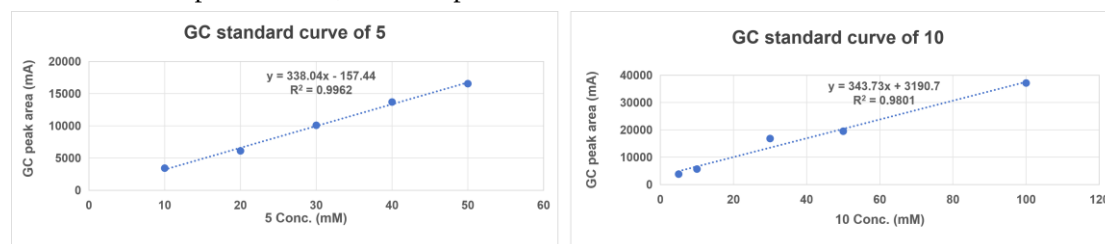


Figure S36. GC standard curves of 5 and 10.

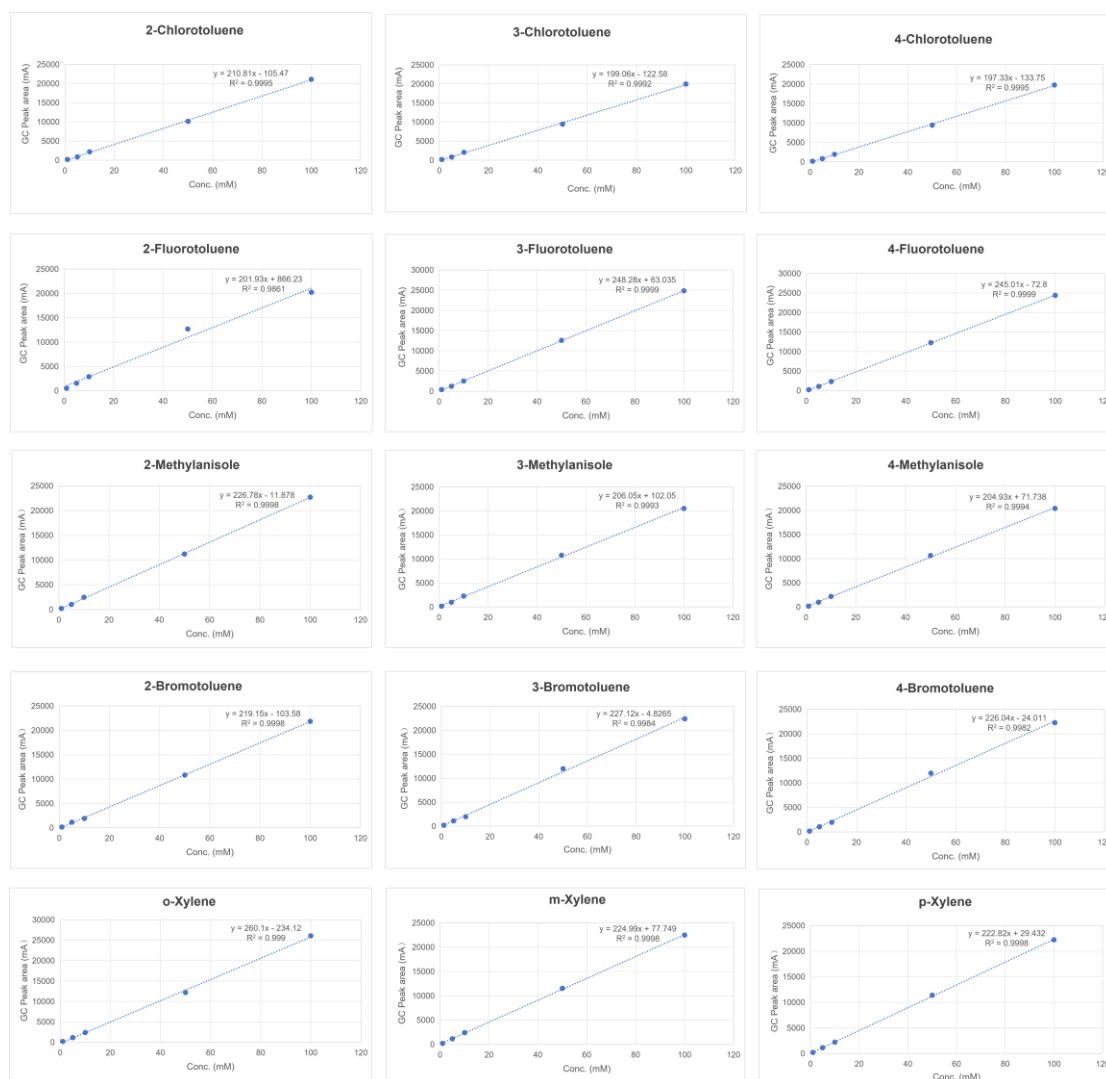


Figure S37. GC standard curves of toluene with different substituents.

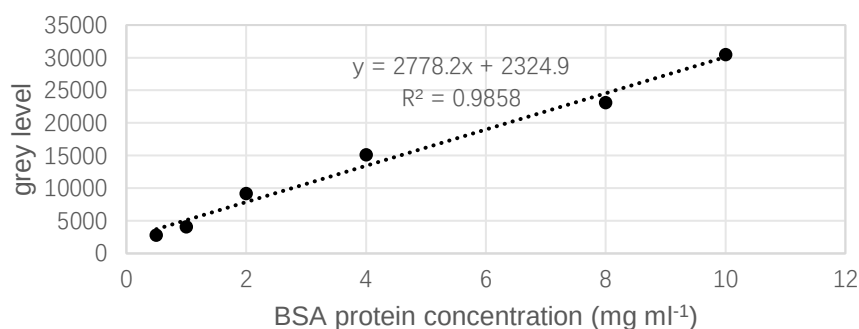


Figure S38. Standard curve of BSA on SDS-PAGE (Due to the special photosensitivity of CvFAP, the enzyme TONs were calculated by using the whole-cell catalyst. Different concentrations of BSA protein were prepared and the grayscale standard curves were made using SDS-PAGE and ImageJ).

Supplementary Tables

Supplementary Table S1. Performance parameters and standard deviations of the DmpR-based biosensor responsive to **10**.

TF	NFI _{max} ^a	Leaky expression NFI ₀ ^b	Operation range ^c	Dynamic range ^d
DmpR-WT	38059.43±2554.89	5799.66±709.36	0-2 mM	6.56-fold
DmpR-M52I	40391.00±477.07	12535.02±1347.07	0-2 mM	3.22-fold
DmpR-M52I-K188R	31464.87±278.27	6551.62±488.33	0-2 mM	4.80-fold
DmpR-M52I-Q71A	34967.31±899.50	6375.54±437.87	0-2 mM	5.48-fold

^a NFI_{max} represents the maximum fluorescence intensity of the biosensor upon effector addition. ^b NFI₀ represents the fluorescence intensity of the biosensor measured in the initial first measurement without effector addition. ^c Operation range represents the concentration range of effector to which the biosensor can respond. ^d Dynamic range represents the ratio between the NFI_{max} with effector and the NFI₀ without inducer.

Supplementary Table S2. Performance parameters and standard deviations of the DmpR-based biosensor responsive to **5**.

TF	NFI _{max} ^a	Leaky expression NFI ₀ ^b	Operation range ^c	Dynamic range ^d
DmpR-WT	No response	5799.66±709.36	No response	No response
DmpR-H100E	41222.06±696.67	9269.06±464.20	0-2 mM	4.44-fold
DmpR-H100N	42118.89±2315.68	11676.26±901.28	0-2 mM	3.61-fold

^a NFI_{max} represents the maximum fluorescence intensity of the biosensor upon effector addition. ^b NFI₀ represents the fluorescence intensity of the biosensor measured in the initial first measurement without effector addition. ^c Operation range represents the concentration range of effector to which the biosensor can respond. ^d Dynamic range represents the ratio between the NFI_{max} with effector and the NFI₀ without inducer.

Supplementary Table S3. Primers for construction of the XylR plasmid.

Name	Sequence
Xylr-Pu	F: GCTCACTCATTAGGCACCCAGGCTTATCGGCCCATTTGCTTTCA CAGATAAGCT R: GGTGTTGGCGGGTGTCTGGGGCTGGCTTAATTATTTATACAGTTCA TCCATA
RFP	F: ATGGTGTCCAAAGGCGAAGAAGACAA R: TTATTTATACAGTTCATCCATAC
Puc57	F: TTAAGCCAGCCCCGACACCCGCCAACACCC R: GCCTGGGGTGCCTAATGAGTGAGCTAA

Supplementary Table S4. Primers for construction of pRSFduet-PmLAAD-ScPPDC-EcPADH and pACYCduet-LbNOX.

Name	Sequence
pRSFduet/ pACYCduet	F: CTCGAGTCTGGTAAAGAAAC R: CATGGTATATCTCCTTATTAAAGTTAAA
ScPPDC	F: AGATCTTAAGGAGAAAATTTATGGCACCTGTTACAATTGAA R: TTTTAACTCCTTAGATATCTTATTTTTTATTTCTTTTAAAG
EcPADH	F: GATATCTAAGGAGTTAAAAAATGACAGAGCCGCATGTAGC R: TTTACCAGACTCGAGTTAATACCGTACACACACCG
PmLAAD	F: ACTTTAATAAGGAGATATACCATGAACATTAGTCGGCGCA R: AAATTTTCTCCTTAAGATCTTTACTTTTTGAAGCGGTCAAG
LbNOX	F: TAATAAGGAGATATACCATGAAAGTCACAGTTGTTGGT R: GTTTCTTTACCAGACTCGAGTTAAGCGTTAACTGATTGGGCA

Supplementary Table S5. Primers for construction of DmpR mutants and plasmids.

Name	Sequence
------	----------

Delete-LuxI/LuxR	F: TAAACGCAGAAAGGCCACCCGAAGGT R: CGCATGGTATGGATGAACTGTACAAA
Po	F: CGCAGGACATCAAGCAACGGCT R: TTAAGCATTGCTCAAGCGGCCTTGG
DmpR-H100nns	F: nnsTCACTCAAGGGTCTGGTCAAGGTCC R: CATCTGCGGCCAGCGAGGAA
DmpR-W128nns	F: nnsATCGACTCCTTCGAGGTG R: CTCCATCTCGGCATAGAAG
DmpR-M52I	F: attGTCAATACCCTGGGCATCGAA R: TTCCCGGCGAAAGCTGGCCAT
DmpR-Q71A	F: gcgTCCGGCCTGAAGGATGCCGA R: GTAACCATGGCGCAGGAACAAG
DmpR-188R	F: cgtCCGGCCGAAGAGTGGGACGA R: GCCAATGACCCGGCACTTGTCTG

Supplementary Table S6. Primers for construction of the mutant libraries of CvFAP.

Name	Sequence
Lib1	F: GACnnsGATnnsGTTAGCACCNnsGTTTCGTTTTGCCAAATTTTCAG R: CAGGGCCATACCCGGCACAAAGCGCACTT
Lib2	F: TTTnnsnnsnnsATGGCCCTGGACCCTGATGGC R: GCGCACTTGCAGATCCGGCAG
Lib3	F: CATnnsnnsnnsGCCATTACCGGTACATGTAAAATGG R: AATACTACGACGGATATATTCATCAATCT
Lib4	F: TTTnnsnnsnnsATGGCCCTGGACGCGGATGCG R: GCGCACTTGCAGATCCGGCAG (Same as Lib2)
FAP_plasmid Lib5	F: CTCGAGTCTGGTAAAGAAACCGC R: ATATCTCCTTATTAAAGTTAAAC
FAP-eqPCR Lib5	F: GTTTAACTTTAATAAGGAGATATACCATGGCAAGTGCAGTAGA AGAT R: CGCAGCAGCGGTTTCTTTACCAGACTCGAGTTATGCTGCCAC GGTAGCAGGGGCGG

Supplementary Table S7. The specific amino acid mutation sites of CvFAP variants obtained by FACS.

CvFAP-Mutants	CvFAP specific mutation sites
AAL	P460A-G462A-Y466L
MTM	P460M-G462T-Y466M
VVG	P460V-G462V-Y466G
VVA	P460V-G462V-Y466A
AVV	P460A-G462V-Y466V
VAT	P460V-G462A-Y466T
GVV	P460G-G462V-Y466V
SVV	P460S-G462V-Y466V
RGM	P460R-G462-Y466M
GVG	P460G-G462V-Y466G
VVR	P460V-G462V-Y466R
TVG	P460T-G462V-Y466G
GTG	P460G-G462T-Y466G
IASAAL	V453I-P454A-G455S-P460A-G462A-Y466L
IGTAAL	V453I-P454G-G455T-P460A-G462A-Y466L
ITAAAL	V453I-P454T-G455A-P460A-G462A-Y466L
IATAAL	V453I-P454A-G455T-P460A-G462A-Y466L
M9	F364L-G423K-G426P-V453I-P454T-G455A-P460A-G462A-Y466L
M9'	F364L-G423K-G426P-V453I-P454A-G455S-P460A-G462A-Y466L
M9P	F364L-G423R-G426P-V453I-P454T-G455A-P460A-G462A-Y466L
M9T	F364L-G423R-G426P-V453I-P454A-G455S-P460A-G462A-Y466L

Supplementary Methods

General information

All compounds in the reactions were analyzed by gas chromatography (GC, Panna Instruments) equipped with an HP-5 column (Agilent, 30 m × 0.320 mm × 0.25 μm, FID, N₂ as the carrier gas) or by high-performance liquid chromatography (HPLC, Unimicro) equipped with an Agilent SB-C18 column (Agilent, Poroshell 120, 4.6 × 150 mm, 2.7 μm). The power of the blue Flex LED strip is 12 W per meter × 5 meters. The distance from the LED light to the reaction sample is 1 cm.

Construction of the XylR biosensor

The *xylR*¹ gene from *Pseudomonas putida* (including its native promoter *Pu*) and the *mCherry* gene, encoding a red fluorescent protein (RFP), were synthesized by Synbio Technologies. First, the *XylR* gene was placed between the constitutive promoter BBa_J23114, the ribosome binding site (RBS) BBa_B0034, and the terminator BBa_B0014 by seamless cloning to form the sensing module. The promoter *Pu*, the RBS BBa_B0034, the *RFP* gene, and the terminator BBa_B0015 were then fused to construct the reporter module. The sensing and reporter modules were subsequently assembled into pUC57, and the resulting plasmid was transformed into *E. coli*, yielding the toluene-responsive biosensor plasmid (pUC57-XylR-Pu-RFP). The correct construction of the biosensor was confirmed by single-molecule nanopore sequencing (CoWin Biotech).

Preparation of competent cells containing the biosensor

A single colony of the biosensor strain was inoculated into 2 ml of LB medium (100 mg l⁻¹ ampicillin) and cultured overnight. This culture was used to inoculate 100 ml of fresh LB medium (100 mg l⁻¹ ampicillin), which was incubated at 37 °C and 220 rpm for 2–4 h. The optical density (OD₆₀₀) was measured every 30 min. When the OD₆₀₀ reached 0.4–0.6, cultivation was stopped. The culture was chilled on ice for 10 min. It was then transferred into pre-chilled 50 ml centrifuge tubes and centrifuged at 4 °C and 1,500 × g for 10 min, after which the supernatant was discarded. The cell pellet was gently resuspended in 10 ml of sterile ice-

cold H₂O until no visible clumps remained. The cells were centrifuged again under the same conditions (4 °C, 1,500 × g, 10 min), and the supernatant was discarded. This washing step with sterile H₂O was repeated once. Finally, the pellet was resuspended in 20 ml of sterile ice-cold 10% (v/v) glycerol. After centrifugation (4 °C, 1,500 × g, 15 min) and careful removal of the supernatant, the cells were resuspended in a minimal volume of sterile, ice-cold 10% (v/v) glycerol to concentrate them 100-fold (cells from 100 ml of culture were resuspended in 1 ml). The suspension was aliquoted (100 µl per tube) into 1.5 ml microcentrifuge tubes. The cells were flash-frozen in liquid nitrogen and stored at –80 °C for long-term preservation.

Transformation of the mutant library

The constructed mutant library was introduced into the prepared biosensor competent cells described above by electroporation using a MicroPulser electroporator (Bio-Rad). Electroporation was carried out using the bacterial electroporation program at a voltage of 4.0–5.5 kV. The electroporation cuvettes were pre-cooled on ice for 20 min before use. Then, 5 µl of the mutant library plasmid was added to the electrocompetent cells, and the mixture was incubated on ice for 10 min and transferred to a pre-chilled electroporation cuvette. After electroporation, 600 µl of LB medium without antibiotics was added to the cuvette. The mixture was then transferred to a centrifuge tube and recovered at 37 °C for 1 h. The constructed mutant library was used for subsequent experiments.

Pilot experiment for FACS sorting

Seed cultures of *E. coli* harboring *p*-cresol biosensor with either CvFAP-WT or CvFAP-P460A-G462I-Y466T were inoculated into 3 ml of LB medium (50 mg l^{–1} kanamycin and 100 mg l^{–1} ampicillin) in 20 ml tubes and incubated at 37 °C. When the OD₆₀₀ reached 1.0, 0.1 mM IPTG was added to induce expression of the CvFAP enzyme, and cultivation was continued at 30 °C for 5 h. Compound **9** was added as the substrate for CvFAP, and blue light (450 nm) was applied for 4 h. Subsequently, CvFAP-WT and CvFAP-P460A-G462I-Y466T were mixed at different ratios (e.g., 1,000:1) to generate mixed libraries. A negative control (no CvFAP) and positive controls (CvFAP-WT and CvFAP-P460A-G462I-Y466T) were included for analysis. A 100 µl aliquot of the mixed sample was diluted with 1 ml of sterile PBS buffer (sterilized

before sorting), washed twice, and then analyzed and sorted using FACS (Sony MA900 cell sorter). The excitation wavelength was 488 nm (FITC channel), and the emission wavelength was 530/30 nm. The system pressure was adjusted to achieve an analysis rate of 1,000 events per second. The top 0.1% fluorescent population was sorted, and the flow cytometry data were analyzed using FlowJo software. Samples collected in flow cytometry tubes after FACS sorting were spread onto LB agar plates (50 $\mu\text{g ml}^{-1}$ kanamycin) and incubated overnight at 37 °C. Single colonies were inoculated into test tubes containing 2 ml of LB medium supplemented with 50 $\mu\text{g ml}^{-1}$ kanamycin, followed by incubation at 37 °C for 12 h. The resulting bacterial cultures were sent to Sanger sequencing.

Activity validation of FACS-sorted cells

For validation of the CvFAP mutants, 24-well deep-well plates were used for cell culture. *E. coli* colonies grown on agar plates after FACS sorting were transferred into a 24-well deep-well plate containing 500 μl of LB medium supplemented with 50 $\mu\text{g ml}^{-1}$ kanamycin. After incubation at 37 °C for 10 h, 4.5 ml of terrific broth (TB) medium (containing 0.1 mM IPTG and 50 $\mu\text{g ml}^{-1}$ kanamycin) was added to each well, and the plates were transferred to a shaker at 22 °C to induce CvFAP mutant expression for 16 h. The *E. coli* cells were harvested by centrifugation (5,000 $\times$ g, 6 min, 4 °C), and the supernatant was discarded. The cell pellets were resuspended in potassium phosphate (KP) buffer (0.2 M, pH 6.0) and transferred to HPLC vials. The reaction system (500 μl) contained whole-cell catalyst (10 g l^{-1} wet cell weight) and 10 mM substrate **4** or substrate **9**. The reaction was carried out at 22 °C for 4 h in an incubator equipped with a blue-light strip (450 nm). After completion of the reaction, 500 μl of methyl *tert*-butyl ether (MTBE) was added to extract the products, and the concentrations of compounds **5** and **10** were determined by GC analysis. Product concentrations were normalized to the aqueous phase volume.

Optimization of the reaction pH

The CvFAP whole-cell catalyst was resuspended in KP buffers of varying pH (6.0–8.0), and reactions were carried out in 5 ml vials. Substrates **4** and **9** were prepared in DMSO. The final concentrations were 20 g l^{-1} for the whole-cell catalyst, 100 mM for substrate **4**, and 200

mM for substrate **9**. The reaction of substrate **4** was carried out in a biphasic system, with *n*-hexadecane added at an equal volume ratio. After the reaction, the mixture was vigorously shaken for 2 min and centrifuged. The upper *n*-hexadecane phase was collected for detection of product **5** after 4 h of reaction. After irradiation with blue light (450 nm) for 4 h, the reaction mixtures containing substrate **9** were extracted with methyl *tert*-butyl ether (MTBE) and analyzed by gas chromatography (GC). The same method was also used to optimize the pH for the LAAD-PPDC-PADH-LbNOX cascade. Amino acid substrates **1** and **6** were weighed as solid powders and added directly to the reaction vial. After the reaction, the reaction mixture was centrifuged, and the supernatant was collected for HPLC analysis.

Calculation of activity and TONs of CvFAP

Owing to the unusual photosensitivity of CvFAP, the apparent specific enzyme activity was calculated using *E. coli* whole-cell catalysts. First, bovine serum albumin (BSA) solutions at different concentrations were prepared, and grayscale standard curves were constructed using SDS-PAGE and ImageJ. The protein concentrations of CvFAP (WT) and variants (M9P and M9T) in cells were then calculated from the grayscale values of SDS-PAGE (Supplementary Fig. 32). The whole-cell reaction was performed in a 1 ml system as described above. The concentrations of substrates **4** and **9** were 100 mM and 200 mM, respectively, and the concentration of the whole-cell CvFAP catalyst was 20 g l⁻¹ wet cell weight. All reactions were performed in KP buffer at pH 6.0, and the reaction temperature was maintained at 22 °C under blue light (450 nm). After 4 h, products **5** and **10** were detected by GC. The conversions of substrates **4** and **9** to products **5** and **10** were used to calculate the apparent TON:

$$\text{TON} = (\text{total moles of product formed})/(\text{total moles of enzyme})$$

Fermentative production of **1** and **6**

The phenylalanine-fermenting strain² was inoculated into a 250 ml shake flask containing 100 ml of LB medium supplemented with 50 mg l⁻¹ streptomycin and grown overnight at 37 °C to serve as the seed culture for fermentation. The 100 ml seed culture was inoculated into 3 l of fermentation medium in a 5 l bioreactor. The strain was cultivated at 37 °C until the OD₆₀₀ reached approximately 10, at which point IPTG was added to a final concentration of 0.1 mM

to induce enzyme expression. The fermentation temperature was then set to 30 °C. The pH was maintained at approximately 6.8 using NH₄OH, and glucose (400 g l⁻¹) was supplied as the feed and added according to the dissolved oxygen signal in the bioreactor. Antifoam agent was added dropwise to suppress excessive foam generation. During the fed-batch fermentation, samples were collected periodically and analyzed by HPLC to quantify L-phenylalanine production. The fermentation medium for L-phenylalanine production in the bioreactor contained (per liter): 20 g glucose, 10 g (NH₄)₂SO₄, 5 g KH₂PO₄, 5 g MgSO₄, 5 g yeast extract, 15 mg FeSO₄, 15 mg MnSO₄, and 1 g betaine.

The engineered L-tyrosine-producing strain³ was inoculated into 100 ml of LB medium supplemented with 25 mg l⁻¹ streptomycin and 17 mg l⁻¹ chloramphenicol, then cultured overnight at 37 °C with shaking at 200 rpm as the preculture. Subsequently, 100 ml of the preculture was transferred into the 5 l bioreactor and incubated at 37 °C until the OD₆₀₀ reached 10 (approximately 12 h). At this stage, IPTG was added to induce expression of *aroG*^{fbr} and *tyrA*^{fbr}. The culture was further incubated at 30 °C with agitation at 500 rpm. The pH was maintained at approximately 6.8 using NH₄OH, and glucose (400 g l⁻¹) was used as the feed. The fermentation medium for L-tyrosine production contained (per liter): 20 g glucose, 10 g (NH₄)₂SO₄, 5 g KH₂PO₄, 5 g MgSO₄, 5 g yeast extract, 2 g tryptone, 15 mg FeSO₄, 1.5 g sodium citrate, 0.1 g thiamine HCl, and 1 ml of trace metal solution. The trace metal solution contained (per liter): 0.1 M HCl, 8.3 g FeCl₃, 0.84 g ZnCl₂, 0.13 g CuCl₂·2H₂O, 0.1 g CoCl₂·2H₂O, 0.1 g H₃BO₃, 0.1 g Na₂MoO₄, and 0.016 g MnCl₂. During the fed-batch fermentation, samples were collected periodically and analyzed by HPLC to quantify L-tyrosine production.

Computational details

Molecular docking simulations

Molecular docking simulations were performed using AutoDock⁴ (version 4.2.6). The modeling of CvFAP mutants was performed using SWISS-MODEL⁵⁻⁷. After removal of crystallographic waters and sodium counterions, CvFAP (PDB ID: 6YRU) and the mutants were used as the receptors, and substrates **4** and **9** were set as the ligands, with all rotatable bonds allowed to rotate freely. To include all possible binding

conformations, a $60 \times 60 \times 60$ grid box was used for the docking simulations. The Lamarckian genetic algorithm⁴ was applied, and each docking calculation consisted of 150 genetic algorithm runs. All other parameters were set to their default values. The conformation with the lowest binding energy was selected for further analysis.

Molecular dynamics simulations

MD simulations were performed using YASARA⁸ (version 23.12.24) with the AMBER14 force field. The CvFAP–ligand complex was immersed in a cubic simulation box filled with TIP3P water molecules, with a margin distance of 5 Å. After energy minimization (1010 steps), the simulation was run at pH 6.0 with a time step of 2.5 fs for a total of 100 ns, and 400 snapshots were extracted at 250 ps intervals for trajectory analysis. The CorrectDrift function was enabled to prevent the protein from drifting across periodic boundaries, and the trajectories were analyzed with the md_analysis module in YASARA.

Calculation of enzyme channels and catalytic pockets

Tunnel analysis was performed with CaverDock 1.26⁹, using the FAD-N5 atom as the starting point, and the binding energies of substrates **4** and **9** within the tunnels were calculated with the following parameters: minimum probe radius, 0.9 Å; shell depth, 4 Å; shell radius, 3 Å; clustering threshold, 3.5 Å; maximum distance, 3 Å; and desired radius, 5 Å.

The volume of the catalytic pocket of CvFAP and its mutants was calculated using the KVFinder-web server^{10,11} (<https://kvfinder-web.cnpem.br/>). Calculations were performed for wild-type CvFAP and the mutants. The pocket with ID KAI, corresponding to the catalytic cavity of CvFAP, was selected for subsequent analysis and visualization. Pocket hydropathy was calculated using KVFinder-web with the Eisenberg–Weiss hydropathy scale. For each cavity voxel, the hydropathy value of the side chain of the nearest interfacing amino acid was projected onto the cavity voxel. Average hydropathy (Avg Hyd.) of each pocket corresponds to the arithmetic mean of

hydropathy values over all cavity voxels; these values range from -1.42 (highly hydrophobic) to 2.6 (highly hydrophilic).

Docking and metadynamics simulations

The substrate **4** or **9** was docked using AutoDock, with the docking site positioned approximately $8 \text{ \AA}$ from residue G432 (wild type) or G432R (mutant). Subsequent energy minimization was carried out using Maestro 13.0, and protonation states were assigned at pH 7.0. A water box was constructed with a $10 \text{ \AA}$ buffer, employing the TIP3P water model. Sodium and chloride ions were added to neutralize the system to a net charge of zero. Metadynamics simulations were performed with Desmond V6.6¹², using the “specify center of mass” option. The distance was defined between group 1 (substrate) and group 2 (G432 or G432R). CV1 and CV2 were defined as the substrate and the G432 (G432R) residue, respectively. For CV1, the width was set to 0.5 and the wall potential to 15; for CV2, the width was set to 0.5 and the wall potential to 20. The Gaussian height was set to $0.03 \text{ kcal mol}^{-1}$. The simulation duration was 100 ns at a temperature of 303 K.

Key sequences information

>The optimized sequence of CvFAP

```
ATGGCAAGTGCCGTGGAAGATATTCGTAAAGTGCTGAGTGATAGTAGCAGTCCGGTTG
CCGGTCAGAAATATGATTATATTCTGGTGGGCGGTGGTACAGCCGCATGCGTGCTGGCA
AATCGTCTGAGTGCCGATGGTAGTAAACGTGTGCTGGTGTGCTGGAAGCAGGTCCGGATA
ATACCAGCCGCGATGTGAAAATTCCGGCAGCAATTACCCGTCTGTTTCGCAGCCCCGCTG
GATTGGAATCTGTTTAGCGAACTGCAAGAACAGCTGGCAGAACGCCAGATATATATGG
CCCGTGGTCGCCTGCTGGGCGGTAGCAGTGCCACCAATGCCACCCTGTATCATCGCGG
TGCAGCCGGTGACTATGATGCATGGGGCGTTGAAGGCTGGAGCAGTGAAGATGTTCTG
AGCTGGTTTGTTTCAGGCAGAAACCAATGCAGATTTTGGCCCCGGGTGCCTATCATGGCA
GTGGCGGTCCGATGCGCGTGGAATAATCCGCGTTATACCAATAAGCAGCTGCATACCGC
CTTTTCAAAGCCGCAGAAAGATTGGTCTGACCCCGAATAGCGATTTTAATGATTGGA
GCCATGATCATGCCGGTTATGGCACCTTTCAGGTTATGCAGGATAAAGGCACCCGCGCA
GATATGTATCGCCAGTATCTGAAACCGGTGCTGGGTGCGCCGAATCTGCAAGTGCTGAC
CGGCGCAGCCGTGACCAAAGTTAATATTGATCAGGCAGCAGGCAAAGCCCAGGCACT
GGGTGTGGAATTTTCTACCGATGGCCCCGACCGGCGAACGTCTGAGCGCCGAACCTGGCA
CCGGGCGGTGAAGTGATTATGTGCGCAGGCGCAGTTCATACCCCGTTTCTGCTGAAAC
```

ATAGCGGCGTTGGCCCGAGTGCAGAACTGAAAGAATTTGGCATTCCGGTGGTGAGCAA
TCTGGCCGGTGTGGCCAGAATCTGCAAGATCAGCCGGCCTGCCTGACCGCAGCACCG
GTAAAGAAAAATATGATGGCATTGCCATTAGTGATCATATCTATAATGAGAAAGGCCA
GATTCGCAAACGTGCCATTGCAAGTTATCTGCTGGGTGGTCGTGGTGGTCTGACCACT
ACCGGCTGTGATCGTGGCGCATTTGTGCGCACCGCAGGCCAGGCACTGCCGGATCTGC
AAGTGCCTTTGTGCCGGGTATGGCCCTGGACCCTGATGGCGTTAGCACCTATGTTTCGT
TTTGCCAAATTTAGAGCCAGGGTCTGAAATGGCCGAGTGGTATTACCATGCAGCTGAT
TGCATGTCGCCCCGAGAGTACCGGCAGTGTGGCCTGAAAAGCGCCGATCCGTTTGCC
CCGCCGAACTGAGCCCGGGTTATCTGACCGATAAAGATGGTGCAGATTTGGCAACCC
TGCCTAAAGGTATTCATTGGGCACGCGATGTGGCCCGTAGCAGTGGCCTGAGTGAATAT
CTGGATGGCGAACTGTTTCCGGGTAGCGGCGTGGTTAGTGATGATCAGATTGATGAATA
TATCCGTCGTAGTATTCATAGCAGCAATGCCATTACCGGTACATGTAAAATGGGCAATGC
AGGCGATAGCAGTAGTGTGGTTGATAATCAGCTGCGCGTTCATGGCGTTGAAGGTCTG
CGTGTGTGGATGCCAGCGTTGTGCCGAAAATTCCGGGCGGTGAGACCGGTGCCCCGG
TGGTTATGATTGCCGAACGCGCAGCAGCCCTGCTGACCGGTAAAGCCACCATTGGCGC
AAGCGCCGCCGCCCTGCTACCGTGGCAGCATAA

>CvFAP

MASAVEDIRKVLSDSSPVAGQKYDYILVGGGTAAACVLNRLSADGSKRVLVLEAGPDNT
SRDVKIPAAITRLFRSPLDWNLFSELQEQLAERQIYMARGRLLGGSSATNATLYHRGAAGD
YDAWGVEGWSSSEDLVSWFVQAETNADFGPGAYHSGGPMRVENPRYTQKQLHTAFFKA
AEEVGLTPNSDFNDWSHDHAGYGTQVMDKGTRADMYRQYLKPVLRNRLQVLTGA
AVTKVNIDQAAGKAQALGVEFSTDGPTGERLSAELAPGGEVIMCAGAVHTPFLKHSGV
GPSAELKEFGIPVVSNLAVGVQNLQDQPACLTAAPVKEKYDGLAISDHIYNEKGQIRKRAI
ASYLLGGRGGLTSTGCDRGAFVRTAGQALPDLQVRFVPGMALDPDGVSTYVRFQFQSQ
GLKWPSGITMQLIACRPQSTGSGVGLKSADPFAPPKLSPGYLTDKDGADLATLRKGIHWAR
DVARSSALSEYLDGELFPVSGVVSDDQIDEYIRRSIHSSNAITGTCKMGNAGDSSSVVDNQ
LRVHGVGLRVVDASVVPKIPGGQTGAPVVMIAERAAALLTGKATIGASAAAPAT

>The optimized sequence of PmLAAD

ATGAACATTAGTCGGCGCAAACCTGCTTTTAGGCGTAGGCGCAGCCGGTGTCTTGCTG
GCGGTGCTGCTTTGGTGCCGATGGTCCGCCGGGACGGCAAGTTCGTGGAGTCGAAGA
GTCGGGCGCTCTTCGTGGAATCCACCGAGGGGGCTTTACCGTCGGAGTCCGATGTCGT
GATCATCGGGGGTGGGATTCAAGGCATTATGACGGCAATTAATTTAGCGGAACGGGGC
ATGAGTGTTACCATCCTCGAAAAAGGTGAGGTAGCGGGCGAACAGTCAGGTGCGGGCA
TACTCTCAGATCATCTCCTATCAGACTAGTCCTGAAATCTTCCCTCTTCACCATTACGGC
AAGATCCTTTGGCGCGGCATGAATGAAAAAATTGGGGCGGACACCTCTTACCGTACAC
AGGGCCGCGTCGAGGCACTTGCCGATGAAAAAGCCCTCGACCGTGCACAGGAGTGGA
TCAAACTGCTAAAGAAACCGCGGGGTTTGACGTTCCATTGAATACTCGCATCATCAA
AGGGGAAGAGCTTTTGAACCGGTTAGTGGGTGCTCAAACACCTTGACGGTTCGCGGC
GTTTGAGGAGGATTCGGGTTCGTAGATCCGGAAACTGGCACCCCTACATTGGCCCGT
TACGCGAAGCAAATTGGCGTTAAGATTTACACGCATTGTGCTGTTGCGGGTATTGAAAC
GGCAGGTGGGAAAATTTCCGACGTCGTCACCGAGAAGGGGGCTATTCGCACGTCAAA
CGTCGTATTAGCGGGTGGGATTTGGTCCCGGTTATTTATGGGGAACATGGGCGTGGATT
TGCCGACGCTCAATGTTTACCTCAGCCAGCAGCGCGTTAGTGGGGTCCAGGCGCCCC

ACGTGGCAACGTGCACTTGCCTAACGGGATTCACTTTCGGGAACAAGCCGATGGCACG
TACGCTGTTGCTCCGCGTATCTTTACAAGTAGCATTGTGAAGGACTCCTTTCTCTTGGG
GCCTAAGTTTATGCATCTGCTGGGCGGCGGGGAGTTGCCATTAGAGTTCAGCATTGGTG
AGGATCTTTTCAATTCCCTCAAGATGCCAACGTCTTGGAATTAGACGAAAAGTCTCCT
TTTGAGCAGTATCGCATCGCTACAGCAACCCAAAACACAGAACACTTAGACGCAGTAT
TTCAGCGCATGAAGACAGAATTTCCGGTATTCGAGAAATCTCAGATTGTGGAACGGTG
GGGCGCTGTAGTTAGTCCAACCTTTCGATGAACTGCCTATTATCTCTGAGGTAAAGGAAT
ATCCTGGGCTCGTAATCAATACTGCGACGGTATGGGGTATGACAGAGGGCCCGGCTGC
GGGTGAAGTCACTGCCGACATTGTTACCGGGAAGAAGCCGGTCATCGACCCAACCCCT
TTCTCACTTGACCGCTTCAAAAAGTAA

>PmLAAD

MNISRRKLLLGVGAAGVLAGGAALVPMVRRDGFVESKSRALFVESTEGALPSESDVVII
GGGIQGIMTAINLAERGMSVTILEKGEVAGEQSGRAYSQIISYQTSPEIFPLHHYGKILWRG
MNEKIGADTSYRTQGRVEALADEKALDRAQEWIKTAKETAGFDVPLNTRIIKGEELSNRL
VGAQTPWTVAAFEEDSGSVDPETGTPTLARYAKQIGVKIYTHCAVRGIETAGGKISDVVTE
KGAIRTSNVVLGGIWSRLFMGNMGVDLPTLNVYLSQQRVSGVPGAPRGNVHLPNGIHF
REQADGTYAVAPRIFTSSIVKDSFLLGPKFMHLLGGGELPLEFSIGEDLFNSFKMPTSWKLD
EKSPFEQYRIATATQNTEHLDAVFQRMKTEFPVFEKSQIVERWGAVVSPTFDELPIISEVKE
YPGLVINTATVWGMTEGPAAGEVTADIVTGKKPVIDPTPFSLDRFKK

> The sequence of ScPPDC

ATGGCACCTGTTACAATTGAAAAGTTCGTAAATCAAGAAGAACGACACCTTGTTTCCA
ACCGATCAGCAACAATTCCCTTTGGTGAATACATATTTAAAAGATTGTTGTCCATCGATA
CGAAATCAGTTTTTCGGTGTTCCTGGTGACTTCAACTTATCTCTATTAGAATATCTCTATTC
ACCTAGTGTTGAATCAGCTGGCCTAAGATGGGTCGGCACGTGTAATGAACTGAACGCC
GCTTATGCGGCCGACGGATATTCCCGTTACTCTAATAAGATTGGCTGTTTAATAACCACG
TATGGCGTTGGTGAATTAAGCGCCTTGAACGGTATAGCCGGTTCGTTTCGCTGAAAATGT
CAAAGTTTTGCACATTGTTGGTGTGGCCAAGTCCATAGATTTCGCGTTCAAGTAACTTTA
GTGATCGGAACCTACATCATTTGGTCCCACAGCTACATGATTCAAATTTTAAAGGGCCA
AATCATAAAGTATATCATGATATGGTAAAAGATAGAGTCGCTTGCTCGGTAGCCTACTTG
GAGGATATTGAACTGCATGTGACCAAGTCGATAATGTTATCCGCGATATTTACAAGTAT
TCTAAACCTGGTTATATTTTGTTCCTGCAGATTTTGCGGATATGTCTGTTACATGTGATA
ATTTGGTTAATGTTCCACGTATATCTCAACAAGATTGTATAGTATAACCCTTCTGAAAACC
AATTGTCTGACATAATCAACAAGATTACTAGTTGGATATATTCCAGTAAAACACCTGCG
ATCCTTGGAGACGTACTGACTGATAGGTATGGTGTGAGTAACTTTTTGAACAAGCTTAT
CTGCAAAACTGGGATTTGGAATTTTCCACTGTTATGGGAAAATCTGTAATTGATGAGT
CAAACCCAACCTTATATGGGTCAATATAATGGTAAAGAAGGTTTAAAACAAGTCTATGAA
CATTTTGAACGTGTGCGACTTGGTCTTGCAATTTGGAGTCGACATCAATGAAATTAATAAT
GGGCATTATACTTTTACTTATAAACCAAATGCTAAAATCATTCAATTTTCATCCGAATTATA
TTCGCCTTGTGGACACTAGGCAGGGCAATGAGCAAATGTTCAAAGGAATCAATTTTGC
CCCTATTTTAAAAGAACTATACAAGCGCATTGACGTTTCTAACTTTCTTTGCAATATGA
TTCAAATGTAACCTCAATATACGAACGAAACAATGCGGTTAGAAGATCCTACCAATGGAC
AATCAAGCATTATTACACAAGTTCACCTACAAAAGACGATGCCTAAATTTTGAACCCCT
GGTGATGTTGTCGTTTGTGAAACAGGCTCTTTTCAATTCTCTGTTTCGTGATTTTCGATTT

CCTTCGCAGTTAAAATATATATCGCAAGGATTTTTCTTTCCATTGGCATGGCCCTTCCT
 GCCGCCCTAGGTGTTGGAATTGCCATGCAAGACCACTCAAACGCTCACATCAATGGTG
 GCAACGTAAAAGAGGACTATAAGCCAAGATTAATTTGTTTGAAGGTGACGGTGCAGC
 ACAGATGACAATCCAAGAAGTGAAGCACCATTCTGAAGTGCAATATCCACTAGAAGTT
 ATCATTTGGAACAATAACGGCTACACTATTGAAAGAGCCATCATGGGCCCTACCAGGTC
 GTATAACGACGTTATGTCTTGGAATGGACCAAATATTTGAAGCATTTCGGAGACTTCG
 ACGGAAAGTATACTAATAGCACTCTCATTCAATGTCCCTCTAAATTAGCACTGAAATTG
 GAGGAGCTTAAGAATTCAAACAAAAGAAGCGGGATAGAAGTTTGTAGAAGTCAAATTA
 GGCGAATTGGATTTCCCCGAACAGCTAAAGTGCATGGTTGAAGCAGCGGCACTTAAAA
 GAAATAAAAAATAA

>ScPPDC

MAPVTIEKFFVNQEERHLVSNRSATIPFGEYIFKRLLSIDTKSVFGVPGDFNLSLLEYLYSPSV
 ESAGLRWVGTCNELNAAYAADGYSRYSNKIGCLITTYGVGELSALNGIAGSFAENVKVLH
 IVGVAKSIDSRSSNFSNRNLHHLVPQLHDSNFKGPNHKVYHDMVKDRVACSVAYLEDIETA
 CDQVDNVIRDIYKYSKPGYIFVPADFADMSVTCDNLVNVPRISQQDCIVYPSENQLSDIINK
 ITSWIYSSKTPAILGDVLTDRYGVSNFLNKLICKTGIWNFSTVMGKSVIDESNPTYMGQYN
 GKEGLKQVYEHFELCDLVLFHGVVDINEINNGHYTFTYKPNAKIIQFHPNYIRLVDTRQNE
 QMFKGINFAPILKELYKRIDVSKLSLQYDSNVTQYTNETMRLEDPTNGQSSIITQVHLQKT
 MPKFLNPGDVVVCETGSFQFSVRDFAFPSQLKYISQGFFLSIGMALPAALGVGIAMQDHSN
 AHINGGNVKEDYKPRILIFEGDGAAQMTIQELSTILKCNIPLEVIIWNNNGYTIERAIMGPT
 RSYNDVMSWKWTKLFEAFGDFDGKYTNSTLIQCPSKLALKLEELKNSNKRSGIELLEVKL
 GELDFPEQLKCMVEAAALKRNKK

> The sequence of EcPADH

ATGACAGAGCCGCATGTAGCAGTATTAAGCCAGGTCCAACAGTTTCTCGATCGTCAAC
 ACGGTCTTTATATTGATGGTCGTCCTGGCCCCGCACAAAGTGAAAAACGGTTGGCGAT
 CTTTGATCCGGCCACCGGGCAAGAAATTGCGTCTACTGCTGATGCCAACGAAGCGGAT
 GTAGATAACGCAGTCATGTCTGCCTGGCGGGCCTTTGTCTCGCGTCGCTGGGCCGGGC
 GATTACCCGCAGAGCGTGAACGTATTCTGCTACGTTTTGCTGATCTGGTGGAGCAGCAC
 AGTGAGGAGCTGGCGCAACTGGAAACCCTGGAGCAAGGCAAGTCAATTGCCATTTC
 CGTGCTTTTGAAGTGGGCTGTACGCTGAACTGGATGCGTTATACCGCCGGGTAAACGA
 CCAAATCGCGGGTAAAACGCTGGACTTGTGATTCCCTTACCCAGGGGGCGCGTTA
 TCAGGCCTGGACGCGTAAAGAGCCGGTTGGCGTAGTGGCGGGAATTGTGCCATGGAA
 CTTTCCGTTGATGATTGGTATGTGGAAGGTGATGCCAGCACTGGCAGCAGGCTGTTCA
 ATCGTGATTAAGCCTTCGGAAACCACGCCACTGACGATGTTGCGCGTGGCGGAACTGG
 CCAGCGAGGCTGGTATCCCTGATGGCGTTTTTAATGTCGTCACCGGGTCAGGTGCTGTA
 TGCGGCGCGGCCCTGACGTCACATCCTCATGTTGCGAAAATCAGTTTTACCGGTTCAAC
 CGCGACGGGAAAAGGTATTGCCAGAAGTCTGCTGATCACTTAACGCGTGTAACGCTG
 GAACTGGGCGGTAAAAACCCGGCAATTGTATTAAGATGCTGATCCGCAATGGGTTAT
 TGAAGGCTTGATGACCGGAAGCTTCCTGAATCAAGGGCAAGTATGCGCCGCCAGTTTCG
 CGAATTTATATTGAAGCGCCGTTGTTTGACACGCTGGTTAGTGGATTTGAGCAGGCGGT
 AAAATCGTTGCAAGTGGGACCGGGGATGTCACCTGTTGCACAGATTAACCCTTTGGTT
 TCTCGTGCGCACTGCGACAAAGTGTGTTTCTCGACGATGCGCAGGCACAGCAAG
 CAGAGCTGATTCGCGGGTCGAATGGACCAGCCGGAGAGGGGTATTATGTTGCGCCAAC

GCTGGTGGTAAATCCCGATGCTAAATTGCGCTTAACTCGTGAAGAGGTGTTTGGTCCGG
 TGGTAAACCTGGTGCAGTAGCGGATGGAGAAGAGGCGTTACAACCTGGCAAACGACA
 CGGAATATGGCTTAACTGCCAGTGTCTGGACGCAAAATCTCTCCCAGGCTCTGGAATAT
 AGCGATCGCTTACAGGCAGGGACGGTGTGGGTAAACAGCCATACCTTAATTGACGCTA
 ACTTACCGTTTGGTGGGATGAAGCAGTCAGGAACGGGCCGTGATTTTGGCCCCGACTG
 GCTGGACGGTTGGTGTGAACTAAGTCGGTGTGTGTACGGTATTAA

>EcPADH

MTEPHVAVLSQVQQFLDRQHGLYIDGRPGAQSEKRLAIFDPATGQEIASTADANEADVND
 AVMSAWRAFVSRRWAGRLPAERERILLRFADLVEQHSEELAQLETLEQGKSIAISRAFEVG
 CTLNWMRYTAGLTTKIAGKTLDSLPLPQGARYQAWTRKEPVGVVAGIVPWNFPLMIGM
 WKVMPALAAAGCSIVIKPSETTPLTMLRVAELASEAGIPDGVFNVVTGSGAVCGAALTSHPH
 VAKISFTGSTATGKGIARTAADHLTRVTLELGKNPAIVLKDADPQWVIEGLMTGSFLNQG
 QVCAASSRIYIEAPLFDLTVSGFEQAVKSLQVPGMSPVAQINPLVSRAHCDKVCSFLDDA
 QAQQAELIRGSGNPAGEGYVAPTLVNPDAKLRLTREEVFGPVVNLVRVADGEEALQLA
 NDTEYGLTASVWTQNLSQLALEYSDRLQAGTVWVNSHTLIDANLPFGGMKQSGTGRDFGP
 DWLDGWCETKSVCVRY

> The sequence of LbNOX

ATGAAAGTCACAGTTGTTGGTTGTACACATGCCGGAACCTTTGCGATTAAACAAATCTT
 GGCCGAACACCCGGATGCCGAAGTGACCGTCTACGAACGTAACGATGTCATTTTCATT
 CTCTCTTGTGGAATCGCCCTTTACCTGGGGGGGAAAGTCGCTGACCCGCAAGGCCTCT
 TTTATTCAAGTCCTGAAGAACTCCAAAAATTAGGTGCCAATGTCCAAATGAATCACAAT
 GTTTTAGCAATCGATCCTGATCAAAAGACAGTGACCGTTGAGGACTTAACCAATCATGC
 ACAAACGACTGAGTCATACGATAAACTAGTCATGACGTCTGGTTCTTGGCCAATTGTTC
 CCAAGATTCCGGGCATCGATAGCGATCGCGTTAAGCTCTGCAAAAACCTGGGCACATGC
 GCAAGCTCTAATCGAAGATGCTAAAGAAGCCAAGCGGATTACCGTTATTGGTGCCGGC
 TATATTGGTGCTGAACTAGCAGAAGCCTACTCCACTACTGGTCATGACGTAACCTTAATT
 GATGCGATGGCCCGGGTGATGCCCAAGTACTTTGATGCTGATTTTACGGATGTCATTGA
 ACAAGATTATCGGGATCACGGTGTCCAACCTGCCTTAGGTGAAACGGTTGAAAGCTTT
 ACTGATAGTGCAACTGGGCTGACTATTAAGACTGATAAGAACAGCTATGAAACGGATCT
 CGCTATTTTATGCATTGGCTTTAGACCAAATACCGACCTGCTGAAAGGCAAAGTCGATA
 TGGCACCAAACGGCGCGATTATTACGGATGACTACATGCGTTCTTCTAACCCCTGATATT
 TCGCCGCTGGTGACAGTGCTGCTGTGCACTACAACCCAACCCATCAGAATGCTTACATT
 CCCTTAGCAACTAACGCGGTGCGTCAAGGTATCCTAGTCGGTAAAAACCTAGTTAAGC
 CGACCGTCAAGTATATGGGAACACAATCATCTTCTGGTTTGGCACTCTATGATCGGACG
 ATCGTCTCAACTGGTTTAAACGCTAGCAGCTGCAAAACAACAAGGGTTGAACGCTGAAC
 AAGTGATTGTTGAAGATAATTATCGCCCAGAGTTTATGCCGTCAACTGAACCCGTTTTG
 ATGTCATTAGTCTTTGACCCCGACACACACCGGATCTTAGGTGGTGCGTTAATGAGTAA
 ATACGACGTTTCACAATCGGCCAACACCCTTTCTGTTTGCATCCAAAACGAAAATACAA
 TTGATGACTTAGCGATGGTTGATATGCTCTTCCAACCTAACTTTGATCGACCATTCAACT
 ACCTAAACATCTTAGCGCAAGCTGCTCAGGCAAAGGTTGCCCAATCAGTTAACGCTTA
 A

>LbNOX

MKVTVVGCTHAGTFAIKQILAEHPDAEVTVYERNDVISFLSCGIALYLGGKVADPQGLFYSSPEELQKLGANVQMNHNVLIDPDQKTVTVEDLTNHAQTTESYDKLVM TSGSWPIVPKIPGIDSDRVKLCKNWAHAQALIEDAKEAKRITVIGAGYIGAELAEAYSTTGHDTVTLIDAMARVMPKYFDADFTDVIEQDYRDHGVQLALGETVESFTDSATGLTIKTDKNSYETDLAILCIGFRPNTDLLKGKVDMAPNGAIITDDYMRSSNPDIFAAGDSA AVHYNPTHQNAYIPLATNAVRQGILVGKNLVKPTVKYMGQTQSSSGLALYDRTIVSTGLTLAAAKQQGLNAEQVIVEDNYRPEFMPSTEPVLMSLVFDPDTHRILGGALMSKYDVSQSANTLSVCIQNENTIDDLAMVDMLFPQNFDRPFNYLNILAQAAQAKVAQSVNA

>The sequence of BsSacB

ATGAACATCAAAAAGTTTGCAAAACAAGCAACAGTATTAACCTTTACTACCGCACTGCTGGCAGGAGGCGCAACTCAAGCGTTTGCGAAAGAAACGAACCAAAAAGCCATATAAGGAAACATACGGCATTTCATATTACACGCCATGATATGCTGCAAATCCCTGAACAGCAA AAAAATGAAAAATATCAAGTTCCTGAATTCGATTTCGTCCACAATTA AAAAATATCTCTTCTGCAAAAGGCCTGGACGTTTGGGACAGCTGGCCATTACAAAACGCTGACGGCACTGTCGCAAACTATCACGGCTACCACATCGTCTTTGCATTAGCCGGAGATCCTAAAAATGCGGATGACACATCGATTTACATGTTCTATCAAAAAGTCGGCGAAACTTCTATTGACAGCTGGA AAAACGCTGGCCGCGTCTTTAAAGACAGCGACAAATTCGATGCAAATGATTCTATCCTA AAAGACCAAACACAAGAATGGTCAGGTTTCAGCCACATTTACATCTGACGGAAAAATCCGTTTATTCTACACTGATTTCTCCGGTAAACATTACGGCAAACAAACACTGACAACTGCA CAAGTTAACGTATCAGCATCAGACAGCTCTTTGAACATCAACGGTGTAGAGGATTATAA ATCAATCTTTGACGGTGACGGAAAAACGTATCAAAATGTACAGCAGTTCATCGATGAA GGCAACTACAGCTCAGGCGACAACCATACGCTGAGAGATCCTCACTACGTAGAAGATA AAGGCCACAAATACTTAGTATTTGAAGCAAACACTGGAAGTGAAGATGGCTACCAAGG CGAAGAATCTTTATTTAACAAGCATACTATGGCAAAGCACATCATTCTTCCGTCAAG AAAGTCAAAAACCTTCTGCAAAGCGATAAAAAACGCACGGCTGAGTTAGCAAACGGCG CTCTCGGTATGATTGAGCTAAACGATGATTACACACTGAAAAAAGTGATGAAACCGCT GATTGCATCTAACACAGTAACAGATGAAATTGAACGCGCGAACGTCTTTAAAATGAAC GGCAAATGGTACCTGTTCACTGACTCCCGCGGATCAAAAATGACGATTGACGGCATTACGTCTAACGATATTTACATGCTTGGTTATGTTTCTAATTCTTTAACTGGCCCATACAAGCCGCTGAACAAAACCTGGCCTTGTGTTAAAAATGGATCTTGATCCTAACGATGTAAACCTTTA CTTACTCACACTTCGCTGTACCTCAAGCGAAAGGAAACAATGTCGTGATTACAAGCTAT ATGACAAACAGAGGATTCTACGCAGACAAACAATCAACGTTTGCGCCAAGCTTCCTGC TGAACATCAAAGGCAAGAAAACATCTGTTGTCAAAGACAGCATCCTTGAACAAGGACAATTAACAGTTAACAATAA

>BsSacB

MNIKKFAKQATVLTFTTALLAGGATQAFAKETNPKPYKETYGISHITRHDMLQIPEQQKNEKYQVPEFDSSTIKNISSAKGLDVWDSWPLQNADGTVANYHGYHIVFALAGDPKNADDTSIYMFYQKVGETSIDSWKNAGRVFKDSKFDANDSILKDQTQEWSGSATFTSDGKIRLFYTFSGKHYGKQTLTTAQVNVASDSSLNINGVEDYKSIFDGDGKTYQNVQQFIDEGNYSSGDNHTLRDPHYVEDKGHKYLVFEANTGTEDGYQGEESLFNKAYYGKSTSFFRQESQKLLQSDKKRTAELANGALGMIELNDDYTLKKVMKPLIASNTVTDEIERANVFKMNGKWYLFDTDSRGSKMTIDGITSNDIYMLGYVSNSTGTPYKPLNKTGLVLKMDLDPNDVTFYTSHFVPPQAKGNNVVITSYMTNRGFIADKQSTFAPSFLNKGKKTSSVVKDSILEQQQLTVNK

>The optimized sequence of pUC57-DmpR-Po-sfGFP

TAAGCATTGCTCAAGCGGCCTTGGGCAATTGATCAAATGCTTAAAAAGTCTGCGCAAGCG
 CGGCTTAATTTGCTCGCTCCGATCATTCTAAAAATTAGAAACACATTGAAAAACATTACCT
 TGAAGTCTGTTTTAGACCTTGGCACAGCCGTTGCTTGATGTCCTGCGAAAGAGGAGAAAA
 CTAGAATGAGCAAAGGCGAAGAACTGTTTACGGGCGTCGTGCCGATTCTGGTGGAAGTGG
 ATGGCGATGTCAATGGTCACAAATTCAGCGTGCGCGGCGAAGGTGAAGGCGATGCGACCA
 ATGGTAAACTGACGCTGAAATTIATTTGACCACGGGTAAACTGCCGGTCCCGTGGCCGAC
 CCTGGTGACCACGCTGACGTATGGTGTCAGTGTTTCAGTCGTTACCCGGATCACATGAAA
 CGCCACGACTTTTTCAAATCCGCAATGCCGGAAGGTTATGTTCAAGAACGTACCATCTCATT
 TAAAGATGACGGCACCTACAAAACGCGCGCTGAAGTTAAATTCGAAGGTGATACCCTGGTC
 AACCGTATTGAACTGAAAGGCATCGATTTTAAAGAAGACGGTAATATTCTGGGCCATAAACT
 GGAATATAACTTCAATTCGCACAACGTGTACATCACCGCGGATAAACAGAAAAACGGTATC
 AAAGCCAACTTCAAAAATCCGCCATAATGTGGAAGATGGCAGCGTTCAACTGGCAGACCACT
 ATCAGCAAAACACCCCGATTGGTGATGGCCCGGTGCTGCTGCCGGACAATCATTACCTGAG
 CACGCAGTCTGTTCTGAGTAAAGATCCGAACGAAAAACGTGACCACATGGTCCTGCTGGA
 ATTTGTCACGGCGGCTGGTATCACGCACGGTATGGATGAACTGTATAAATAATAACCAGGCA
 TCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTTCGTTTTATCTGTTGTTGTGCG
 GTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAC
 TCGGATCCCGGGCCCGTCGAGGCCTTTCTGCGTTTATACTCGAGTTTATGGCTAGCTCAGTC
 CTAGGTACAATGCTAGCACTAGAAAAAGAGGAGAAAACTAGAATGCCGATCAAGTACAAGC
 CTGAAATCCAGCACTCCGATTTCAGGACCTGACCAACCTGATCCACTTCCAGAGCATGGA
 AGGCAAGATCTGGCTTGGCGAACAGCGCATGCTGTTGCTTCAGTTTTTCAGCGATGGCCAGC
 TTTTCGCCGGGAAATGGTCAATACCCTGGGCATCGAACGCGCCAAGGGCTTGTTCTGCGCC
 ATGGTTACCAGTCCGGCCTGAAGGATGCCGAACCTGGCCAGGAAGCTGAGACCGAATGCCA
 GCGAAGTCGGCATGTTCTCGCTGGGCGCAGATGTATTCACTCAAGGGTCTGGTCAAGGT
 CCGCCCCACCGAGCTCGATATCGACAAGGAATACGGGCGCTTCTATGCCGAGATGGAGTGG
 ATCGACTCCTTCGAGGTGAAAATCTGCCAGACCGACCTGGGGCAGATGCAAGACCCGGTG
 TGCTGGACTCTGCTCGGCTACGCCTGCGCCTATTCTCGGCGTTTCATGGGCCGGGAAATCAT
 CTTCAAGGAAGTCAGCTGCGGCGGCTGCGGCGGCGACAAGTGCCGGGTTCATTGGCAAGCC
 GGCCGAAGAGTGGGACGACGTTGCCAGCTTCAAACAGTATTTCAAGAACGACCCCATCATC
 GAGGAACTCTACGAGTTGCAATCGCAACTGTTGTCGCTGCGTACCAACCTCGACAAACAGG
 AAGGCCAGTACTACGGCATCGGTGACACCCCGGCCTACCAGACCGTGCGCAATATGATGGA
 CAAGGCCGCACAGGGCAAAGTCTCGGTGCTGCTGCTTGGCGAGACCGGGGTGCGCAAGG
 AGGTCATCGCGCGTAGCGTGACCTGCGCAGCAAACGCGCCGCCGAGCCCTTTGTGCGCG
 TGAAGTGTGCGGCGATCCCGCCGACCTGATCGAGTCCGAATTGTTGCGCGTGGAAGGAG
 GCGCCTTACCGGCGCCTCCAGTCACGCATGGGCGGCTTCGAGCGGGCCGACAAGGGCA
 CCATCTTCCTTGACGAGGTGATCGAACTCAGCCCGCGCGCTCAGGCCAGTCTGCTGCGCGT
 GCTGCAAGAAGGCGAGCTGGAGCGAGTTGGCGACAACCGCACGCGCAAGATCGACGTAA
 GGGTTATCGCCGCCACCCACGAGGACCTGGCCGAAGCGGTCAAGGCCGGGCGTTTTTCGCG
 CCGACCTGTACTACCGACTGAACGTTTTCCCGGTGGCGATCCCGGCGTTGCGCGAACGCCG
 CGAGGACATTCCACTGCTGGTTGAGCACTTCCTTCAGCGCTTCCACCAGGAGTACGGCAAG
 AGAACCTCGGCCTTTCAGGCAAAGCCCTGGAGGCCCTGCCTGCATTACAGTTGGCCGGGC
 AATATCCGTGAGCTGGAGAACGTATCGAGCGCGGCATCATCCTACCGATCCGAACGAAA

GCATCAGCGTGCAGGCGCTGTTCTACGGGCGCCGGAAGAGCCGCAGACCGCCAGCGAGC
GGGTGTCGTTGGACGGCGTGCTGATTCAGCCAGGCAATGGCCAGGGCAGTTGGATCAGCC
AGTTGTTGAGCAGCGGCCTGAGCCTCGACGAGATCGAGGAAAGCCTGATGCGCGAAGCCA
TGCAACAGGCCAACCAAAACGTCTCCGGTGCCGCGCGCTTGCTCGGCCTAAGCCGACCGG
CACTGGCCTATCGGCTGAAGAAAATCGGCATCGAAGGCTAG

The DmpR transcription factor is shown in yellow. The downstream promoter Po is in red. The reporter gene sfGFP is in green.

>The optimized sequence of pUC57-XylR-Pu-RFP

CCCGGGAAAGCGCGATGAACCTTTTTATCGCTGCCTTGATCAAATCGACAGGTGGTTATGC
GCGATTGATGATTTGCTCAAATACAGCCAGCGTGCTGTAGATTTTCTCTACATCCCCCCTTT
CTTTTTACAAAGAAAATCAATAATTTAGATGAAATAAGGGGATCGGTATAAGCAATGGCAT
GGCGGTTGCTAGCTATACGAGACTTAAAAATAAAAAATAGTGGTGACCCTTCAAAAGAGGAGA
AAACTAGAATGGTGTCCAAAGGCGAAGAAGACAATATGGCGATTATCAAAGAATTTATGCG
TTTCAAAGTTCACATGGAAGGTAGCGTCAATGGCCATGAATTTGAAATTGAAGGCGAAGGT
GAAGGCCGTCCGTATGAAGGTACCCAGACGGCAAACTGAAAGTCACCAAAGGCGGTCCG
CTGCCGTTTGCTTGGGATATTCTGTCACCGCAATTCATGTATGGTTCGAAAGCGTACGTTAA
ACACCCGGCCGATATCCCGGACTACCTGAACTGAGCTTTCCGGAAGGCTTCAAATGGGAA
CGTGTTATGAACTTCGAAGATGGCGGTGTGGTTACCGTCACGCAGGATAGCTCTCTGCAAG
ACGGTGAATTTATCTACAAAGTGAACTGCGCGGTACCAATTTCCCGTCTGATGGCCCCGTT
ATGCAGAAGAAAACGATGGGCTGGGAAGCGAGTTCCGAACGTATGTACCCGGAAGACGGT
GCCCTGAAAGGCGAAATCAAACAGCGCCTGAACTGAAAGATGGCGGTCAATTATGACGCA
GAAGTGAAGAACACGTACAAAGCTAAAAAACCGGTCCAACCTGCCGGGCGCATACAACGTG
AACATCAAACCTGGATATCACCTCCCACAACGAAGACTACACGATCGTTGAACAATATGAAC
GCGCAGAAGGTCGTCACCTCAACGGGTGGTATGGATGAACTGTATAAATAACGTAAAGCGGT
TAGCAACTAACCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTT
TTATCTGTTGTTTGTGCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGG
CCTTTCTGCGTTTATACTCGGATCCCGGGCCCGTCGAGGCCTTTCTGCGTTTATACTCGAGTT
TATGGCTAGCTCAGTCCTAGGTACAATGCTAGCACTAGAAAAGAGGAGAAAACCTAGAATGC
CGATCAAGTACAAGCCTGAAATTAAATGTCGCTTACATACAAACCCAAGATGCAGCATGAG
GATATGCAAGACCTTAGCAGCCAGATCCGTTTTCGTTGCCGCCGAAGGCAAGATCTGGTTGG
GAGAGCAGCGCATGCTCGTAATGCAGCTATCTACGCTGGCCAGCTTCCGTCGCGAAATTATC
AGCTTGATCGGCGTCGAGCGGGCCAAGGGTTTCTTCCTGCGGTTGGGCTATCAGTCCGGCC
TGATGGATGCCGAGCTGGCACGCAAGCTGCGGCCGCCATGCGCGAGGAGGAGGTGTTCC
TGGCTGGGCCTCAATTGTATGCGCTCAAGGGGATGGTCAAAGTACGCTTGCTGACAATGGA
TATCGCCATCCGGGACGGACGTTTCAACGTGGAGGCCGAGTGGATTGATTCCTTTGAAGTG
GATATCTGCCGAACCTGAGCTGGGCCTGATGAATGAGCCCGTCTGCTGGACGGTGCTAGGCT
ATGCTAGCGGCTATGGTTTCGGCATTATGGGCCGAGAAATCATTTTCCAGGAACTAGCTGT
CGCGGGTGCGGTGACGATAAATGCCTTATCGTCGGCAAGACCGCAGAAGAGTGGGGCGAT
GTCAGCAGTTTCAAGCCTACTTCAAAAGCGACCCGATCGTAGACGAGCGCTACGAGCTGC
AGACCCAGGTTGCCAACCTGCGCAACCGCCTGAAGCAGTACGATGGGCAGTATTACGGCAT
TGGCCATTCGCCAGCCTACAAGCGCATCTGTGAGACCATCGACAAGGCTGCACGCGGCAG
GGTTTCGGTCTGCTACTGGGTGAGACTGGGGTGGGCAAGGAGGTAATCGCGCGCAGCGT
GCATTTGCGCAGTGAGCGCGCAGAGCAACCCTTCGTCGCGGTGAACTGTGCGGCAATTCCG

CCGGATCTGATCGAGTCGGAACCTGTTTGGTGTGCGATAAGGGGCGCCTATACGGGCGCGGTCA
 ATGCACGCGCTGGACGTTTTGAACGGGCCAACGGCGGCACCATCTTTCTTGATGAGGTGAT
 CGAATTGACGCCGAGGGCCCAGGCCACCTGCTACGGGTATTGCAGGAAGGAGAGCTAGA
 GCGGGTCGGCGGCGACCGCACGCGAAAGGTCGACGTGAGGTTAATCACCGCAACAAACG
 AGAACCTGGAAGAGGCGGTCAAGATGGGGCGCTTTCGCGCAGACCTGTTCTTTCGGCTGA
 ATGTTTTTCCCGTGCATATCCCGCCGTTGCGCGAGCGCGTGGAAGATATCCCGCTGCTGGTC
 GAGCATTTTCTTAGAAGGCACCATAAGGAATACGGTAAGAAGACTCTTGGCCTGTCTGATC
 GAGCGATGGAGGCCTGCCTCCACTACCAATGGCCAGGCAATATCCGCGAGCTGGAGAACGC
 CCTTGAGCGCGGGGTGATTCTTACCGAGAGCAACGAAAGCATCAATGTCGAGTCGCTGTTT
 CCGGGGTTGGCGACGGCTACCGAAGGCGACAGGCTATCGAGCGAGGGCCGGTTGGAGGAG
 GAGTCCGGTGACAGTTGGTTTAGGCAAATTATCGACCAGGGCGTCAGCCTCGAAGATCTCG
 AAGCGGGTTTAATGCGCACGGCCATGGACCGTTGTGGGCAGAATATCTCACAGGCGGCGCG
 GTTGCTGGGATTGACCCGCCCGCAATGGCCTATCGACTTAAGAAGCTTGACCCAGCTTAT
 CTGTGAAAGCAATGGGCCGATAA

The XylR transcription factor is shown in yellow. The downstream promoter Pu is in red. The reporter gene RFP is in purple.

>The optimized sequence of pUC57-DmpR-Po-sfGFP-SacB

TAAGCATTGCTCAAGCGGCCTTGGGCAATTGATCAAATGCTTAAAAAGTCTGCGC
 AAGCGCGGCTTAATTTGCTCGCTCCGATCATTCTAAAAATTAGAAACACATTGAA
 AAACATTACCTTGAAGTCTGTTTTAGACCTTGGCACAGCCGTTGCTTGATGTCTT
 CCGAAAGAGGAGAAAACTAGAATGAGCAAAGGCGAAGAACTGTTTACGGGCGTC
 GTGCCGATTCTGGTGGAACTGGATGGCGATGTCAATGGTCACAAATTCAGCGTGCG
 CGGCGAAGGTGAAGGCGATGCGACCAATGGTAAACTGACGCTGAAATTTATTTGC
 ACCACGGGTAAACTGCCGGTCCCGTGGCCGACCCTGGTGACCACGCTGACGTATG
 GTGTCCAGTGTTCAGTCGTTACCCGGATCACATGAAACGCCACGACTTTTTCAA
 TCCGCAATGCCGGAAGGTTATGTTCAAGAACGTACCATCTCATTAAAGATGACGG
 CACCTACAAAACGCGCGCTGAAGTTAAATTCGAAGGTGATACCCTGGTCAACCGT
 ATTGAACTGAAAGGCATCGATTTTAAAGAAGACGGTAATATTCTGGGCCATAAACT
 GGAATATAACTTCAATTCGCACAACGTGTACATCACCGCGGATAAACAGAAAAACG
 GTATCAAAGCCAACCTTCAAATCCGCCATAATGTGGAAGATGGCAGCGTTCAACTG
 GCAGACCACTATCAGCAAAACACCCCGATTGGTGATGGCCCGGTGCTGCTGCCGG
 ACAATCATTACCTGAGCACGCAGTCTGTTCTGAGTAAAGATCCGAACGAAAAACG
 TGACCACATGGTCCTGCTGGAATTTGTCACGGCGGCTGGTATCACGCACGGTATGG
 ATGAACTGTATAAAGGCGGCAGCGGTGGTGGTAGCGGCGGCATGAACATCAAAAA
 GTTTGCAAAACAAGCAACAGTATTAACCTTTACTACCGCACTGCTGGCAGGAGGC
 GCAACTCAAGCGTTTGCGAAAGAAACGAACCAAAAGCCATATAAGGAAACATACG
 GCATTTCCCATATTACACGCCATGATATGCTGCAAATCCCTGAACAGCAAAAAAAT
 GAAAAATATCAAGTTCCTGAATTCGATTCGTCCACAATTAATAATATCTTCTGCA
 AAAGGCCTGGACGTTTGGGACAGCTGGCCATTACAAAACGCTGACGGCACTGTGCG
 CAACTATCACGGCTACCACATCGTCTTTGCATTAGCCGGAGATCCTAAAAATGCG
 GATGACACATCGATTACATGTTCTATCAAAAAGTCGGCGAAACTTCTATTGACAG
 CTGGAAAAACGCTGGCCGCGTCTTTAAAGACAGCGACAAATTCGATGCAAATGAT
 TCTATCCTAAAAGACCAAAACACAAGAATGGTCAGGTTTCAGCCACATTTACATCTGA

CGGAAAAATCCGTTTATTCTACACTGATTTCTCCGGTAAACATTACGGCAAACAAA
 CACTGACAACCTGCACAAGTTAACGTATCAGCATCAGACAGCTCTTTGAACATCAAC
 GGTGTAGAGGATTATAAATCAATCTTTGACGGTGACGGAAAAACGTATCAAAATGT
 ACAGCAGTTCATCGATGAAGGCAACTACAGCTCAGGCGACAACCATACGCTGAGA
 GATCCTCACTACGTAGAAGATAAAGGCCACAAATACTTAGTATTTGAAGCAAACAC
 TGGAACCTGAAGATGGCTACCAAGGCGAAGAATCTTTATTAAACAAAGCATACTATG
 GCAAAAGCACATCATTCTTCCGTCAAGAAAGTCAAAAACCTTCTGCAAAGCGATAA
 AAAACGCACGGCTGAGTTAGCAAACGGCGCTCTCGGTATGATTGAGCTAAACGAT
 GATTACACACTGAAAAAAGTGATGAAACCGCTGATTGCATCTAACACAGTAACAG
 ATGAAATTGAACGCGCGAACGTCTTTAAATGAACGGCAAATGGTACCTGTTCACT
 GACTCCCGCGGATCAAAAATGACGATTGACGGCATTACGTCTAACGATATTTACAT
 GCTTGGTTATGTTTCTAATTCTTTAACTGGCCCATACAAGCCGCTGAACAAAACCTG
 GCCTTGTGTTAAAAATGGATCTTGATCCTAACGATGTAACCTTTACTTACTCACACT
 TCGCTGTACCTCAAGCGAAAGGAAACAATGTCGTGATTACAAGCTATATGACAAAC
 AGAGGATTCTACGCAGACAAACAATCAACGTTTGCGCCAAGCTTCCTGCTGAACA
 TCAAAGGCAAGAAAACATCTGTTGTCAAAGACAGCATCCTTGAACAAGGACAATT
 AACAGTTAACAAATAAATCTCGGATCCCGGGCCCGTCGAGGCCTTTCTGCGTTTATAC
 TCGAGTTTATGGCTAGCTCAGTCCTAGGTACAATGCTAGCACTAGAAAAGAGGAGA
 AAACCTAGAATGCCGATCAAGTACAAGCCTGAAATCCAGCACTCCGATTTCAAGGA
 CCTGACCAACCTGATCCACTTCCAGAGCATGGAAGGCAAGATCTGGCTTGGCGAA
 CAGCGCATGCTGTTGCTTCAGTTTTTCAGCGATGGCCAGCTTTCGCCGGGAAATGGT
 CAATACCCTGGGCATCGAACGCGCCAAGGGCTTGTTCTGCGCCATGGTTACCAGT
 CCGGCCTGAAGGATGCCGAACCTGGCCAGGAAGCTGAGACCGAATGCCAGCGAAG
 TCGGCATGTTCTCGCTGGGCGCAGATGTATTCACTCAAGGGTCTGGTCAAGGTC
 CGCCCCACCGAGCTCGATATCGACAAGGAATACGGGCGCTTCTATGCCGAGATGGA
 GTGGATCGACTCCTTCGAGGTGAAAATCTGCCAGACCGACCTGGGGCAGATGCAA
 GACCCGGTGTGCTGGACTCTGCTCGGCTACGCCTGCGCCTATTCTCGGCGTTTAT
 GGGCCGGGAAATCATCTTCAAGGAAGTCAGCTGCGGCGGCTGCGGCGGCGACAA
 GTGCCGGGTCAATTGGCAAGCCGGCCGAAGAGTGGGACGACGTTGCCAGCTTCAA
 ACAGTATTTCAAGAACGACCCCATCATCGAGGAACCTCTACGAGTTGCAATCGCAAC
 TGTGTGCTGCTGCGTACCAACCTCGACAAACAGGAAGGCCAGTACTACGGCATCGG
 TCAGACCCCGGCCTACCAGACCGTGCGCAATATGATGGACAAGGCCGCACAGGGC
 AAAGTCTCGGTGCTGCTGCTTGCGGAGACCGGGGTCGGCAAGGAGGTATCGCGC
 GTAGCGTGCACCTGCGCAGCAAACGCGCCGCCGAGCCCTTTGTGCGGGTGAACCTG
 TGCGGCGATCCCGCCGGACCTGATCGAGTCCGAATTGTTTCGGCGTGAAAAAGGC
 GCCTTCACCGGCGCCTCCCAGTCACGCATGGGCGGCTTCGAGCGGGCCGACAAGG
 GCACCATCTTCCTTGACGAGGTGATCGAACTCAGCCCGCGCGCTCAGGCCAGTCT
 GCTGCGCGTGCTGCAAGAAGGCGAGCTGGAGCGAGTTGGCGACAACCGCACGCG
 CAAGATCGACGTAAGGGTTATCGCCGCCACCCACGAGGACCTGGCCGAAGCGGTC
 AAGGCCGGGCGTTTTTCGCGCCGACCTGTACTACCGACTGAACGTTTTCCCGGTGG
 CGATCCCGGCGTTGCGCGAACGCCGCGAGGACATTCCACTGCTGGTTGAGCACTT
 CCTTCAGCGCTTCCACCAGGAGTACGGCAAGAGAACCCTCGGCCTTTCAGGCAAA
 GCCCTGGAGGCCTGCCTGCATTACAGTTGGCCGGGCAATATCCGTGAGCTGGAGA
 ACGTCATCGAGCGCGGCATCATCCTCACCGATCCGAACGAAAGCATCAGCGTGCA

GGCGCTGTTCTACGGGCGCCGGAAGAGCCGCAGACCGCCAGCGAGCGGGTGTC
GTTGGACGGCGTGCTGATTCAGCCAGGCAATGGCCAGGGCAGTTGGATCAGCCAG
TTGTTGAGCAGCGGCCTGAGCCTCGACGAGATCGAGGAAAGCCTGATGCGCGAA
GCCATGCAACAGGCCAACCAAAACGTCTCCGGTGCCGCGCGCTTGCTCGGCCTAA
GCCGACCGGCACTGGCCTATCGGCTGAAGAAAATCGGCATCGAAGGCTAG

The DmpR transcription factor is shown in yellow. The downstream promoter Po is in red. The reporter gene sfGFP is in green. The negative selection marker SacB gene is in gray.

Supplementary References

1. Fernández, S., De Lorenzo, V. & Pérez-Martin, J. Activation of the transcriptional regulator XylR of *Pseudomonas putida* by release of repression between functional domains. *Mol. Microbiol.* **16**, 205–213 (1995).
2. Sekar, B. S., Lukito, B. R. & Li, Z. Production of natural 2-phenylethanol from glucose or glycerol with coupled *Escherichia coli* strains expressing l-phenylalanine biosynthesis pathway and artificial biocascades. *ACS Sustain. Chem. Eng.* **7**, 12231–12239 (2019).
3. Xu, Z. *et al.* Biosynthesis of 4-ethylphenol via an enzyme cascade with an engineered fatty acid photodecarboxylase. *J. Agric. Food Chem.* **74**, 22942–22951 (2026).
4. Morris, G. M. *et al.* AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2009).
5. Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
6. Bienert, S. *et al.* The SWISS-MODEL Repository—new features and functionality. *Nucleic Acids Res.* **45**, D313–D319 (2017).
7. Studer, G. *et al.* QMEANDisCo—distance constraints applied on model quality estimation. *Bioinformatics* **36**, 1765–1771 (2020).
8. Land, H. & Humble, M. S. YASARA: A Tool to Obtain Structural Guidance in Biocatalytic Investigations. in *Protein Engineering* (eds Bornscheuer, U. T. & Höhne, M.) vol. 1685 43–67 (Springer New York, New York, NY, 2018).
9. Marques, S. M. *et al.* Caver Web 2.0: analysis of tunnels and ligand transport in dynamic ensembles of proteins. *Nucleic Acids Res.* **53**, W132–W142 (2025).
10. Guerra, J. V. D. S. *et al.* ParKVFinder: A thread-level parallel approach in biomolecular cavity detection. *SoftwareX* **12**, 100606 (2020).
11. Guerra, J. V. S., Ribeiro-Filho, H. V., Pereira, J. G. C. & Lopes-de-Oliveira, P. S. KVFinder-web: a web-based application for detecting and characterizing biomolecular cavities. *Nucleic Acids Res.* **51**, W289–W297 (2023).
12. Bowers, K. J. *et al.* Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. in *ACM/IEEE SC 2006 Conference (SC'06)* 43–43 (IEEE, Tampa, FL, 2006).