Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

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

Shuke Wu

shukewu@mail.hzau.edu.cn

Huazhong Agricultural University https://orcid.org/0000-0003-0914-9277

Zhaoyang Qin

Huazhong Agricultural University

Yi Zhou

State Key Laboratory of Agricultural Microbiology, College of Life Science and Technology, Huazhong Agricultural University

Shenghao Peng

Huazhong Agricultural University

Wen Zheng

Huazhong Agricultural University

Zherui Xu

Huazhong Agricultural University

Yunpeng Bai

Northwest University

Matthias Höhne

TU Berlin

Thomas Ward

University of Basel https://orcid.org/0000-0001-8602-5468

Uwe Bornscheuer

University of Greifswald https://orcid.org/0000-0003-0685-2696

Article

Keywords:

Posted Date: October 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11382864/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: Yes there is potential Competing Interest. The authors, Z.Q., Y.Z., S.P., Z.X. and S.W., are inventors on a related Chinese patent (CN122382042A). All other authors declare no additional competing interests.

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

Zhaoyang Qin[1], Yi Zhou[1]*, Shenhao Peng[1], Wen Zheng[1], Zherui Xu[1], Yunpeng Bai[2], Matthias Höhne[3], Thomas R. Ward[4], 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

Abstract:

Photoenzymes have emerged as powerful biocatalysts for light-driven transformations, yet their engineering has been constrained by low-to-medium-throughput screening methods, and their catalytic efficiency for non-native reactions/substrates typically remains moderate. Here, we report the first ultrahigh-throughput screening (uHTS) platform for the directed evolution of a photoenzyme, achieved by engineering transcription factor-based biosensors coupled with fluorescence-activated cell sorting. Using the fatty acid photodecarboxylase CvFAP as a model, we performed three iterative rounds of uHTS combining multi-site saturation and random mutagenesis, and successfully obtained evolved variants (M9T and M9P) that exhibited >200-fold improvements in productivity towards two non-native aromatic substrates, with turnover numbers reaching up to 3,700. Molecular dynamics simulations revealed that the accumulated mutations synergistically optimized substrate access, pocket hydrophobicity, and catalytic distances. When integrated into enzyme cascades from

L-phenylalanine and L-tyrosine, these CvFAP variants enabled photobiocatalytic production of toluene (67.9 mM) and *p*-cresol (198 mM). Furthermore, by coupling with high-cell-density fermentation, we achieved direct conversion of glucose-derived crude amino acids to 4.3 g l^{-1} toluene and 21.4 g l^{-1} *p*-cresol without intermediate purification. Collectively, this work establishes uHTS as a powerful platform for reprogramming photoenzymes towards non-native substrates, and provides a sustainable route from renewable biomass to industrially important aromatic chemicals.

Light not only furnishes the living world with energy via photosynthesis, but also offers a distinctive means of directly driving biocatalysis[1-6]. In recent years, photoenzymes have been repurposed from natural enzymes/proteins by exploiting intrinsic photoactive cofactors or incorporating artificial photoactive cofactors. These repurposed photoenzymes catalyzed new-to-nature transformations with high conversions, typically after multiple rounds of directed evolution with low to medium throughput of 10^2–10^3 (e.g., single-site saturation mutagenesis (SSM) combined with microtiter plate screening). Such throughput limits constrain the search for highly active variants, leaving most repurposed photoenzymes with modest catalytic efficiency. In contrast, natural photoenzymes[7], although exceedingly rare, have been evolutionarily optimized for their native reactions. Notably, the fatty acid photodecarboxylase from *Chlorella variabilis* NC64A (CvFAP)[8-11] catalyzes the photodecarboxylation of long-chain fatty acids into alkanes with high catalytic efficiency (turn over numbers, TONs up to 8,000)[12]. Over several years of development, CvFAP has been engineered to convert non-native substrates for production of short-chain alkanes[13,14] and ethylbenzene[15], for kinetic resolution of racemic acids[16,17], and even for new-to-nature reactions[18-21]. However, these efforts have been relied on (semi-)rational design and screening of small mutant libraries, and the catalytic efficiency of CvFAP variants towards non-native substrates remains moderate (TONs < 1,000).

To break the catalytic-efficiency ceiling of photoenzymes for non-native substrates and to accelerate their engineering, ultrahigh-throughput screening (uHTS) methods are highly desirable[22-25], as they can rapidly survey vast sequence spaces (~10^5-10^8 variants) generated by multi-SSM, random mutagenesis, or recombination[26,27]. In particular, microfluidic assays employing cells or *in vitro* microcompartments, such as fluorescence-activated cell sorting (FACS) and fluorescence-activated droplet sorting (FADS), have been applied to many enzymes[22-25]. However, these methods typically required either fluorogenic proxy substrates or coupling of coproduct/cofactor formation (e.g., H_2O_2 or NADH) to a fluorescent signal. The former often uses substrates that diverge from the those in industrial application, whereas the latter is difficult for CvFAP, as it generates only CO_2 as the coproduct. To the best of our knowledge, uHTS has yet to be applied to the directed evolution of photoenzymes.

Building on the ability of CvFAP to defunctionalize biobased feedstocks, it holds potential to contribute toward a more sustainable society and future[28-30]. In this context, we have been interested in harnessing CvFAP for the bioproduction of commodity aromatics, specifically BTEX (benzene, toluene, ethylbenzene, and xylene)[31], which are currently derived mainly from petroleum refining (Fig. 1a). Among them, toluene, with a global annual production of 29 million tons[32], serves both as a major gasoline component and as a precursor to other aromatics. Another petrol-derived aromatics, *p*-cresol, is used in the production of antioxidants and plasticizers[33]. Although toluene and *p*-cresol are manufactured on an industrial scale, their natural biosynthetic pathways have only been elucidated recently. In anoxic microbial communities, the glycine radical enzyme PhdA/PhdB[34] decarboxylates phenylacetate to toluene, albeit at concentrations up to ~10 μM. Likewise, *p*-cresol is formed by the radical SAM enzyme ThiH[35] during thiamine biosynthesis in anaerobes, and by glycine radical enzymes, such as *p*-hydroxyphenylacetate decarboxylase[36] and arylacetate decarboxylase[37] in certain gut microbes. However, these radical enzymes are generally highly oxygen-sensitive and exhibit low catalytic efficiency[38], which limits their application in chemical synthesis. Thus far, the bioproduction of these

high-volume commodity aromatics remains an unmet challenge.

Herein, we report the engineering of transcription factor-based biosensors[39-42] and harnessing them as uHTS platforms for the directed evolution of the photoenzyme CvFAP (Fig. 1b). By constructing large mutant libraries through multi-site saturation mutagenesis and random mutagenesis, and subsequently screening them via FACS (Fig. 1c), we successfully reprogramed CvFAP to convert two non-native aromatic substrates, achieving >200-fold improvements in productivity and TONs up to 3,700 (Fig. 1d). Furthermore, the evolved CvFAP variants were integrated into enzyme cascades[43-45] and coupled with microbial fermentation, enabling the bioproduction of toluene and *p*-cresol at titers of 4.3 and 21.4 g l^{-1}, respectively (Fig. 1e). Collectively, this study not only demonstrates the power of uHTS in engineering photoenzymes, but also establishes photobiocatalytic cascades for the sustainable production of commodity chemicals from biobased feedstocks.

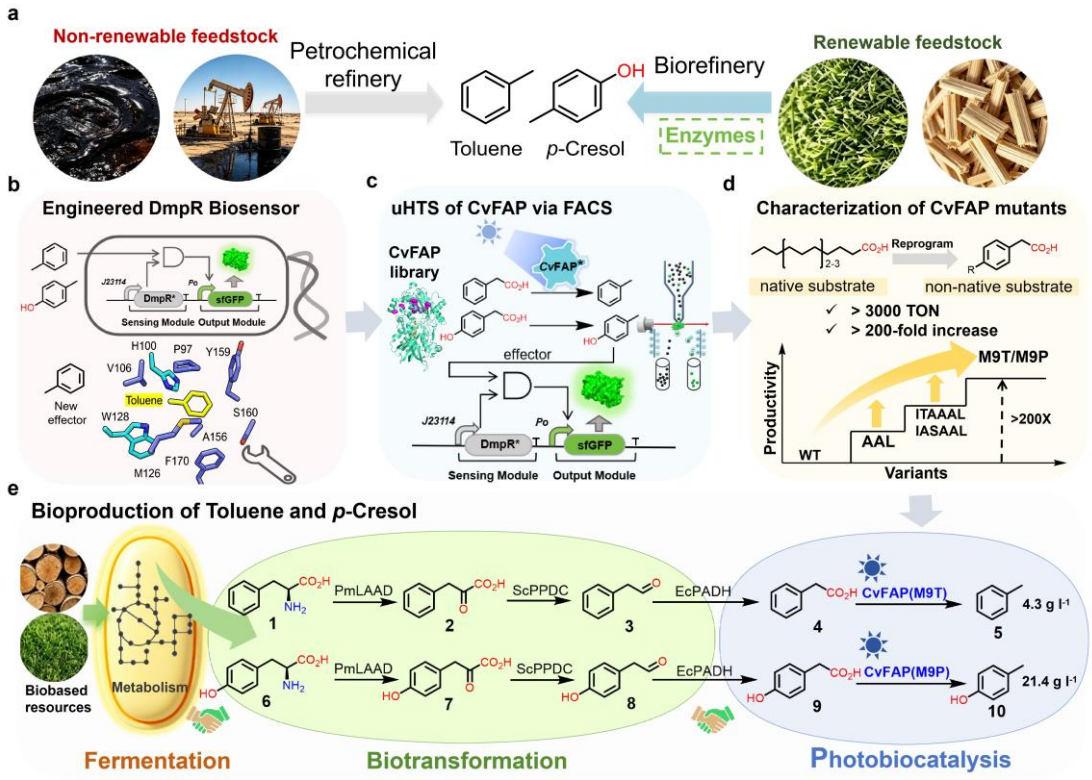


Fig.1 | Ultrahigh-throughput screening and evolution of CvFAP for photobiocatalytic production of toluene (5) and *p*-cresol (10). a, Existing and designed routes for production of toluene and *p*-cresol. **b**, Engineering of DmpR biosensors for selective detection of toluene and *p*-cresol. **c**, Evolution of CvFAP using uHTS technology. **d**, Characterization of CvFAP mutants toward non-native substrates. **e**, Coupled fermentation-cascade biotransformation-photobiocatalysis for

production of toluene and *p*-cresol from biobased resources (note that cofactors and by-products are not shown, see Fig. 6a for details). DmpR, a transcription factor activated by phenol; sfGFP, superfolder green fluorescent protein; J23114, a weak constitutive promoter; Po, DmpR-regulated promoter; PmLAAD, L-amino acid deaminase from *Proteus mirabilis*, ScPPDC, phenylpyruvate decarboxylase from *Saccharomyces cerevisiae*, EcPADH, phenylacetaldehyde dehydrogenase from *E. coli*.

Results

Construction and engineering of biosensors for toluene and *p*-cresol

To enable bioproduction of toluene (**5**) and *p*-cresol (**10**), we designed two enzyme cascades starting from L-phenylalanine (**1**) and L-tyrosine (**6**), respectively, each featuring a CvFAP-catalyzed decarboxylation of the corresponding intermediate, phenylacetate (**4**) and *p*-hydroxyphenylacetate (**9**), respectively (Fig. 1e). To reprogram CvFAP toward these non-native substrates, we employed product-responsive biosensors[39-42] coupled to GFP fluorescence as the uHTS readout (Fig. 2a). To detect **5** selectively, the xylene transcriptional activator XylR[46] was initially evaluated; however, it also responded to the substrate **4** (Supplementary Fig. S1), and attempts failed to eliminate this cross reactivity. Consequently, we shifted to detect **10** using the phenol-responsive transcriptional activator DmpR[47] (Fig. 2a,b). It substantially response to both phenol (Supplementary Fig. S2) and **10** (Fig. 2c), but not to **9** (Supplementary Fig. S3). To further enhance the sensitivity toward **10**, we introduced several previously reported beneficial mutations[48] toward phenol into DmpR. Among the variants tested, DmpR-M52I-Q71A displayed a lower background and a broader dynamic response range toward **10** (Fig. 2d, Supplementary Figs. S4 and S5, Table S1), while retaining no cross-response to substrate **9** (Fig. 2e). This variant was therefore selected as the biosensor for **10** in subsequent experiments.

Encouraged by these results and recent reports on engineering biosensors[49-52], we next aimed to engineer DmpR to detect **5** in the presence of **4**, while wild-type DmpR shows no response to neither **5** (Supplementary Fig. S6) or **4** (Supplementary Fig. S7). To facilitate the evolution of DmpR, we established a dual positive/negative selection

system employing sfGFP as the positive marker and SacB[53] as the negative selection marker (Fig. 2g). A toxicity assay confirmed effective counterselection by SacB in the presence of sucrose (Fig. 2f). Guided by the structure of DmpR (PDB: 6IY8)[54], we docked **10** into the ligand-binding domain (LBD) of DmpR (Fig. 2b). Residues H100 and W128 were found to be in close proximity to the phenol group of **10**, consistent with previous studies[54]. To engineered DmpR to **5** (no phenol group compared to **10**), H100 and W128 were subjected to simultaneous saturated mutagenesis to generate a DmpR mutant library. Cells harboring the DmpR library were first incubated with **5** and then subjected to FACS sorting to isolate mutants responsive to **5** based on GFP fluorescence (Fig. 2h). The sorted cells were incubated with sucrose without **5** to eliminate the cells with leaky expression. Through this positive/negative selection, two DmpR mutants, H100N and H100E, were identified to response to **5** but not to **4** (Fig. 2i,j, Supplementary Table S2). Between these two, DmpR-H100E exhibited a lower background signal (Supplementary Table S2), and was therefore selected as the biosensor for **5** in subsequent experiments.

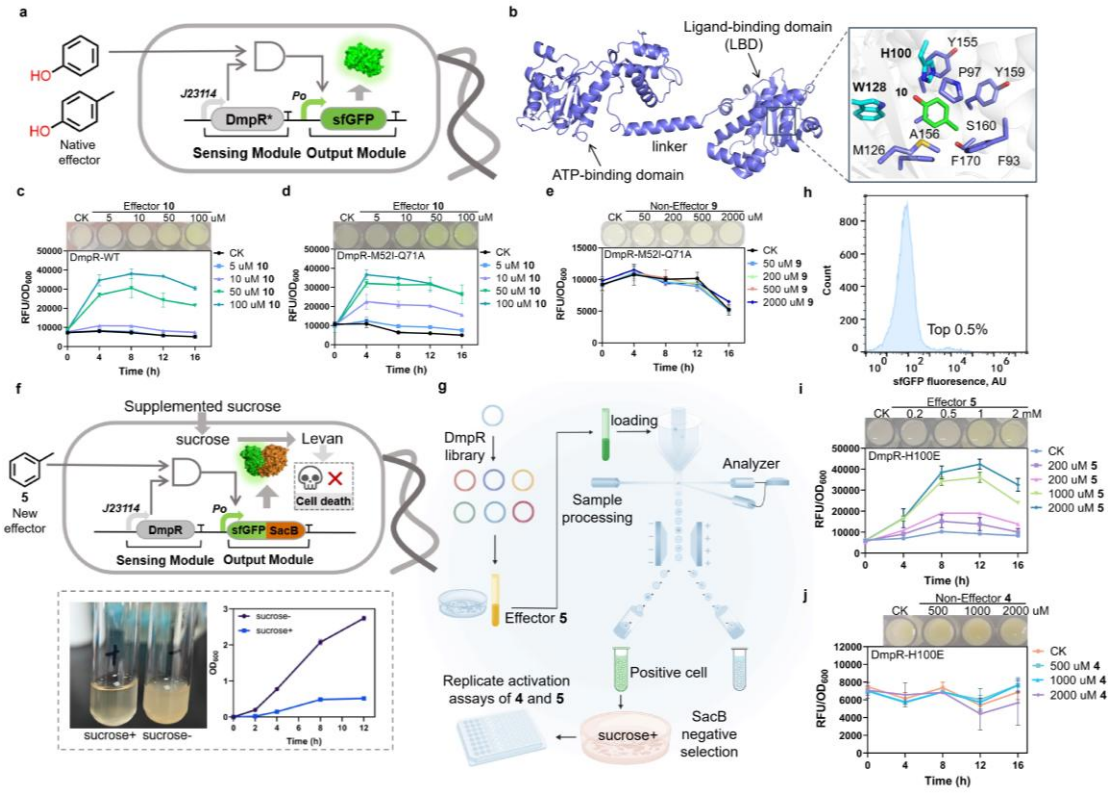


Fig. 2 | Construction and engineering of the *p*-cresol and toluene biosensors. a, Design of transcription factor DmpR-based *p*-cresol biosensor. **b**, Overall structure of

DmpR with **10** docked into the LBD of DmpR (**10**, green stick; key residues H100 and W128, cyan stick). **c**, Response of DmpR to **10** in different concentrations. **d**, Response of DmpR-M52I-Q71A to **10** in different concentrations. **e**, Response of DmpR-M52I-Q71A to **9** in different concentrations. **f**, Design of a dual positive/negative selection system with sfGFP (positive) and SacB (negative) selection markers. Toxicity assay of SacB was performed with phenol and with/without sucrose. **g**, FACS-based screening procedure for engineering DmpR to selectively response to **5**. **h**, Fluorescence distribution of DmpR mutation library (H100X-W128X) in FACS. **i**, Response of DmpR-H100E to **5** in different concentrations. **j**, Response of DmpR to **4** in different concentrations. Credit: photograph in **g**, biogdp (https://biogdp.com). Data in **c**, **d**, **e**, **f**, **i** and **j** are means ± s.d. of three technical replicates.

Design and construction of multi-site mutant libraries of CvFAP

Conventional CvFAP engineering, based on SSM or the focused rational iterative site-specific mutagenesis (FRISM)[55] strategy, samples only limited sequence space and has repeatedly identified very similar positions: residues 460 and 462 are critical for activity toward short-chain fatty acids[14], residues 462, 466, and 574 are key for improving novel functions in many evolution campaigns[16-21]. To expand the search beyond these semi-rational paradigms, we turned to the Saprot model[56], a protein language model that integrates structural information into sequence representations. Whole-enzyme fitness prediction suggested that mutations at residues 453, 454, 455, 460, 462, 466, 573, 574 and 575, could enhance CvFAP productivity (Fig. 3a). In parallel, molecular dynamics (MD) simulations highlighted that the loop spanning residues 453–462 plays a pivotal role in aromatic substrate conversion (Supplementary Fig. S10 and S11) and that residues 573–575 may form hydrogen bonds with the substrate carboxyl group. Guided by the convergence of Saprot predictions and MD simulations, we constructed three triple-site saturation mutagenesis libraries (Fig. 3b): Lib1 (460-462-466), combining established hotspots with position 466, revealed by both prior campaigns and Saprot; Lib2 (453-454-455), covering the loop segment around the Saprot-predicted residue 453; and Lib3 (573-574-575), targeting the MD-suggested substrate-anchoring region.

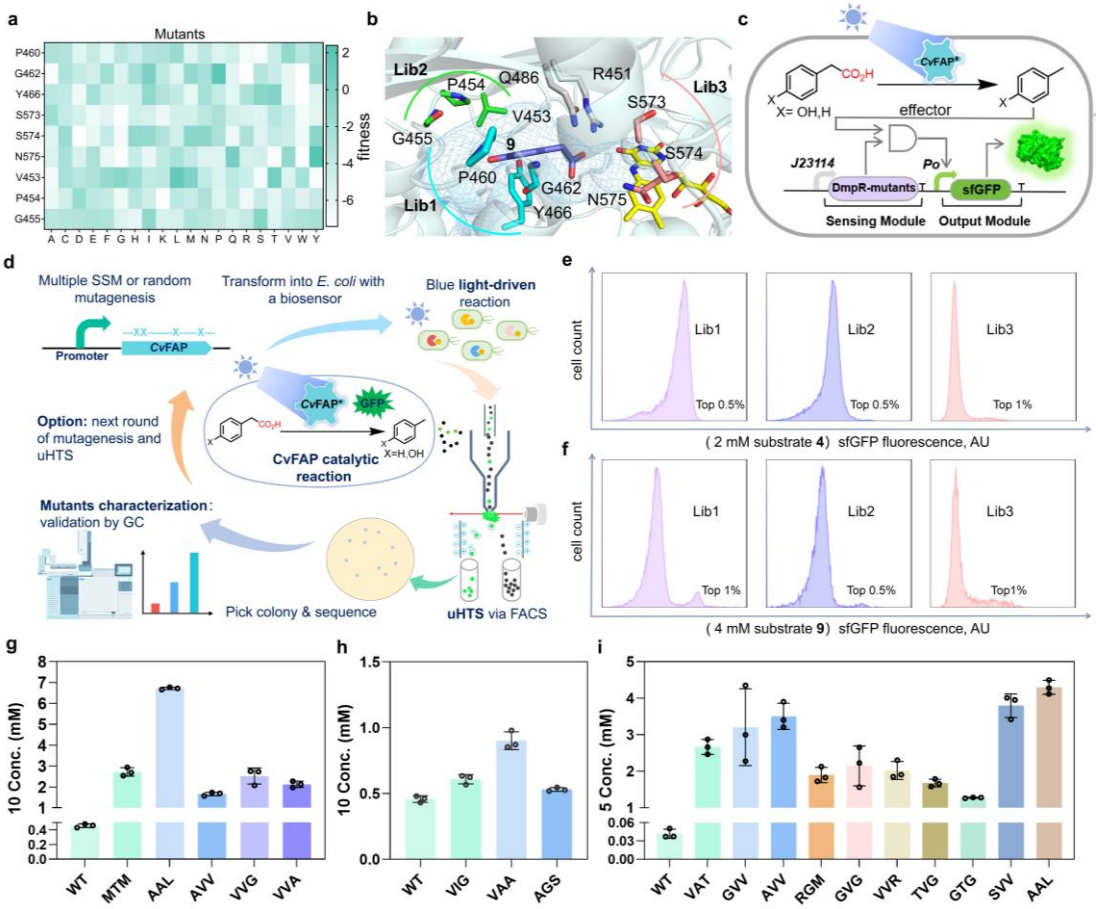


Fig. 3 | Directed evolution of CvFAP through uHTS via FACS (1st round). a, Fitness prediction of CvFAP single mutants by the Saprot model. **b**, Residues selected for constructing CvFAP libraries targeting **9** (Lib1-3 are in different color). **c**, Design of single cell system for photodecarboxylation and fluorescence generation. **d**, Schematic of biosensor-assisted uHTS screening for CvFAP evolution. **e**, Fluorescence distribution of CvFAP mutation libraries in the first round of uHTS toward **4**. Reaction conditions: 2 mM **4**, blue light 4 h. **f**, Fluorescence distribution of CvFAP mutation libraries in the first round of uHTS toward **9**. Reaction conditions: 4 mM **9**, blue light 4 h. **g**, Production of **10** by mutants isolated from Lib1 toward **9**. Mutants are named after the amino acids at 460-462-466. **h**, Production of **10** by mutants isolated from Lib2 toward **9**. Mutants are named after the amino acids at 453-454-455. **i**, Production of **5** by mutants isolated from Lib1 toward **4**. Mutants are named after the amino acids at 460-462-466. Data in **g**, **h**, and **i** are means ± s.d. of three technical replicates.

Ultrahigh-throughput screening of CvFAP via FACS

To evaluate the feasibility of biosensor-based uHTS of CvFAP, a pilot assay was performed using a previously reported CvFAP variant (P460A-G462I-Y466T, for ethylbenzene)[15] that exhibited improved productivity towards **4** and **9** (Supplementary Fig. S12). *E. coli* cells harboring the *p*-cresol biosensor were transformed with

plasmids encoding CvFAP wild type (WT) or P460A-G462I-Y466T mutant. The resulting cells were mixed and incubated with **9** under blue light illumination for 4 h, and subjected to FACS analysis and sorting based on GFP fluorescence. The improved variant was enriched and isolated (Supplementary Fig. S12), confirming the feasibility of the biosensor-based uHTS of CvFAP.

The three CvFAP mutant libraries were individually transformed into *E. coli* cells carrying either the toluene or the *p*-cresol biosensor. Following incubation with **4** or **9** and under blue-light illumination for 4 h, the cells were collected for FACS analysis (Fig. 3c,d). Populations exhibiting fluorescence intensity within the top 0.5-1% range were sorted (Fig. 3e,f, Supplementary Fig. S13) for subsequent verification by photobiocatalysis reaction and gas chromatography analysis. In the first round of uHTS with **9**, multiple mutants with substantially improved productivity were obtained. Among them, the triple mutant AAL (P460A-G462A-Y466L) from Lib1 exhibited the best performance, with a 15-fold increase in productivity relative to WT (Fig. 3g, Table 1). Several improved mutants were also identified from Lib2 (Fig. 3h), albeit with lower productivity than the AAL mutant. No significantly improved mutants were recovered from Lib3. When screening with substrate **4**, the same AAL mutant from Lib1 again performed best, showing a ~100-fold increase in productivity, while several other Lib1 mutants (e.g., VAT, GVV, and SVV) exhibited over ~60-fold improvements (Fig. 3i). No mutants with markedly improved productivity were obtained from Lib2 or Lib3.

To further improve the productivity, a second round of mutagenesis-uHTS was conducted using the AAL mutant as the template. Given that several improved mutants had been recovered from Lib2 in the first round, another triple-site saturation mutagenesis library (Lib4, Fig. 4a) targeting residues 453-454-455 based on the AAL mutant was constructed. To increase selection stringency, cells containing Lib4 and one of biosensors were incubated with a lower concentration of **4** or **9** under blue light illumination for a shorter time, and then screened by FACS (Fig. 4b,c). Screening against **9** yielded three mutants (IGTAAL, IATAAL, and ITAAAL, Table 1), all of which fully converted 10 mM **9** within 2 h (Fig. 4d). At higher substrate loadings,

ITAAAL retained outstanding activity, achieving 100%, 68%, 72% conversion at 20, 50, 100 mM **9**, respectively (Fig. 4e, Supplementary Fig. S14). Similarly, cells containing Lib4 were screened toward **4**, yielding two mutants (IASAAL and ITAAAL, Fig. 4f). After a 2 h reaction, the IASAAL mutant exhibited excellent productivity, achieving near-full conversion of 10 mM **4**, corresponding to a 243-fold improvement over WT. At 50 mM **4**, **5** accumulated to 22.5 mM (45% conversion). Given the high volatility of **5**, several biphasic reaction systems were evaluated. The system using *n*-hexadecane performed the best, affording 81% conversion of 50 mM **4** (Fig. 4g, h).

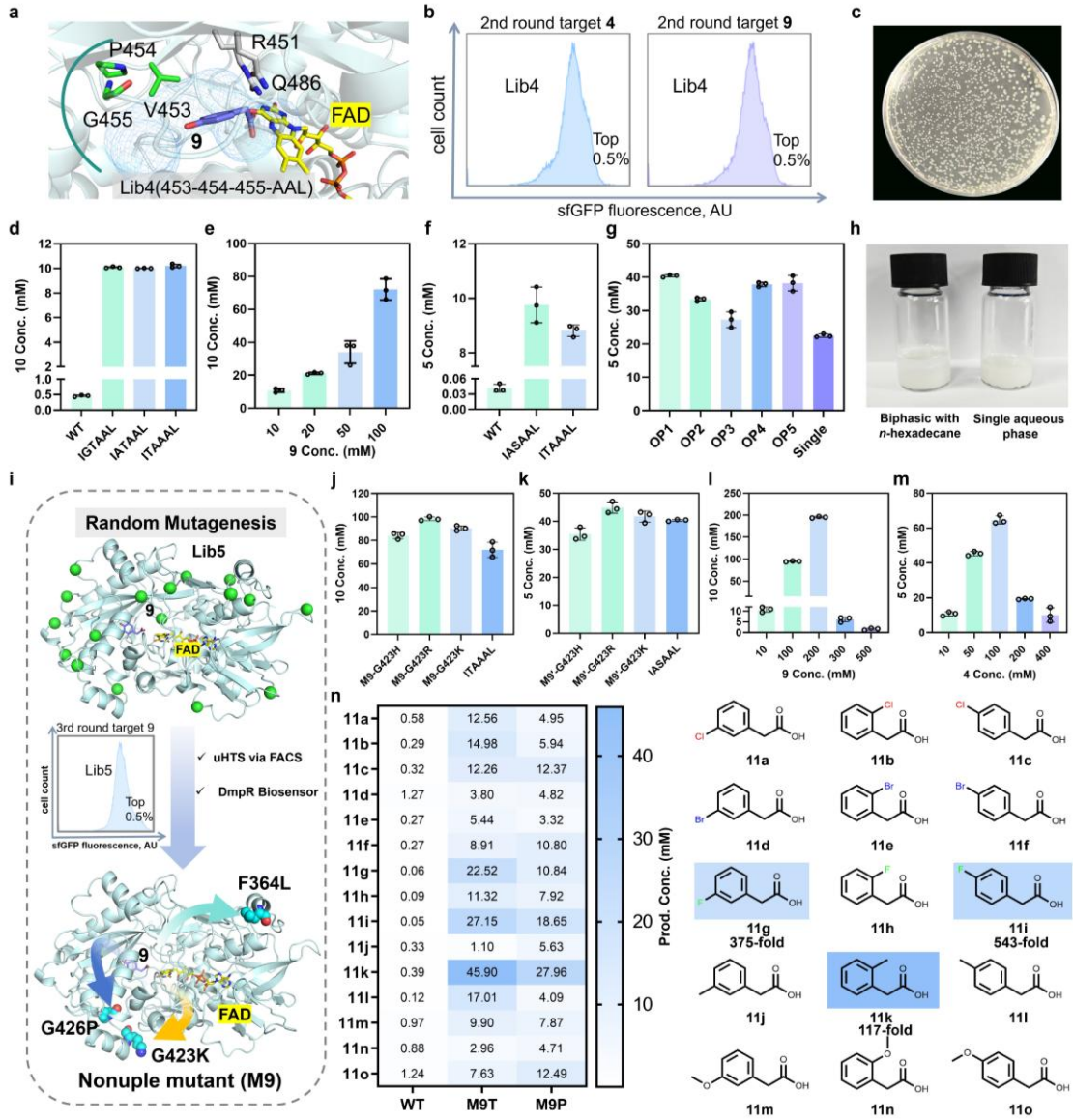


Fig.4 | Directed evolution of CvFAP through uHTS via FACS (2nd and 3rd round). **a**, Residues selected for Lib4 construction using variant AAL as the template. **b**,

Fluorescence distribution of CvFAP mutation libraries in the second round of uHTS targeting **4** and **9**. Reaction conditions: 1 mM **4** or 2 mM **9**, blue light 2 h. **c**, Cells from the second round of FACS sorting were plated on agar plate for recovery. **d**, Production of **10** by the second-round mutants toward **9**. **e**, Production of **10** from **9** in different concentrations by mutant ITAAAL. **f**, Production of **5** by second-round mutants toward **4**. **g**, Production of **5** from 50 mM **4** by mutant IASAAL in single phase or biphase systems with different organic solvents (OP1, *n*-hexadecane; OP2, dodecane; OP3, ethyl caprate; OP4, ethyl laurate; OP5, ethyl tetradecanoate). **h**, Photograph of biphasic system (*n*-hexadecane) and single aqueous phase system. **i**, Third round of evolution of CvFAP by random mutagenesis of ITAAAL and uHTS targeting **9**. **j**, Production of **10** from 100 mM **9** by mutant M9 with different amino acids at G423. **k**, Production of **5** from 50 mM **4** in biphasic system by mutant M9' with different amino acid at G423. **l**, Production of **10** from **9** in different concentrations by M9P. **m**, Production of **5** from **4** in different concentrations in biphasic system by M9T. **n**, Substrate scope of CvFAP WT and mutants M9T and M9P. Reaction conditions: biphasic system with *n*-hexadecane, 20 g l^{-1} whole-cell catalyst, 50 mM substrate, 22 °C, 5 h. Product concentrations in biphasic systems of **g**, **k**, **m**, and **n**, are normalized to the volume of aqueous phase. Data in **d**, **e**, **f**, **g**, **j**, **k**, **l**, **m**, and **n** are means ± s.d. of three technical replicates.

Table 1. Summary of key CvFAP mutants and their performance in producing 5 or 10.

CvFAP Mutant	Specific Mutations	Conc. of **5** or **10** Produced
AAL	P460A-G462A-Y466L	4.3 mM **5** from 10 mM **4** 6.7 mM **10** from 10 mM **9**
VAT	P460V-G462A-Y466T	2.7 mM **5** from 10 mM **4**
GVV	P460G-G462V-Y466V	3.2 mM **5** from 10 mM **4**
SVV	P460S-G462V-Y466V	3.8 mM **5** from 10 mM **4**
IGTAAL	**V453I-P454G-G455T**-P460A-G462A-Y466L	64.0 mM **10** from 100 mM **9**
ITAAAL	**V453I-P454T-G455A**-P460A-G462A-Y466L	72.1 mM **10** from 100 mM **9**
IATAAL	**V453I-P454A-G455T**-P460A-G462A-Y466L	56.4 mM **10** from 100 mM **9**
IASAAL	**V453I-P454A-G455S**-P460A-G462A-Y466L	40.4 mM **5** from 50 mM **4**
M9	**F364L-G423K-G426P**-V453I-P454T-G455A-P460A-G462A-Y466L	90.4 mM **10** from 100 mM **9**
M9'	**F364L-G423K-G426P**-V453I-P454A-G455S-P460A-G462A-Y466L	41.7 mM **5** from 50 mM **4**
M9P	F364L-**G423R**-G426P-V453I-P454T-G455A-P460A-G462A-Y466L	195.5 mM **10** from 200 mM **9**

M9T	F364L-**G423R**-G426P-V453I-P454A-G455S-P460A-G462A-Y466L	64.7 mM **5** from 100 mM **4**

Random mutagenesis of CvFAP and substrate scope study

Random mutagenesis remains a powerful complement to (semi-)rational and AI-guided engineering, as it can uncover beneficial mutations that design strategies miss[26,27]. Leveraging the established uHTS platform, we therefore subjected ITAAAL, the best variant toward **9** in the second round, to error-prone PCR (Lib5, Fig. 4i). Cells containing Lib5 and *p*-cresol biosensor were incubated with **9** under blue light illumination and then subjected to the third round of uHTS by FACS. We recovered a nonuple mutant M9 (F364L-G423K-G426P-ITAAAL, Table 1) that achieved 90.5% conversion of 100 mM **9** (Fig. 4j). The newly identified triple mutations (F364L-G423K-G426P) were combined with the IASAAL mutant (best for **4** in the second round) to generate a combinatorial variant M9' (F364L-G423K-G426P-IASAAL, Table 1) for substrate **4**. This nonuple mutant achieved 83.4% conversion of 50 mM **4** (Fig. 4k). Given that random mutagenesis may not deliver the optimal substitutions at each newly identified hit sites (F364, G423, and G426), we performed a conservation analysis of the CvFAP enzyme family (Supplementary Fig. S15 and S16). This analysis revealed that at position 423, basic amino acids appeared with the second-highest frequently, while proline was predominant at position 426. Therefore, G423 in M9 or M9' was mutated to two other basic amino acids, arginine (R) and histidine (H) and evaluated for conversion of **9** or **4** (Fig. 4j,k). Gratifyingly, the nonuple mutant F364L-G423R-G426P-ITAAAL (designated M9P, Table 1) achieved 98.5% conversion of 100 mM **9**, and F364L-G423R-G426P-IASAAL (designated M9T, Table 1) achieved 89.8% conversion of 50 mM **4**. At higher substrate loadings, M9P reached 97.8% conversion of 200 mM **9**, yielding **10** at 21.1 g l^{-1} (Fig. 4l, TON = 3,748), and M9T achieved 64.7% conversion of 100 mM **4**, producing **5** at 6.0 g l^{-1} (Fig. 4m, TON = 1270). Substrate inhibition was observed at elevated concentrations (300 mM for **9** or 200 mM for **4**).

The substrate scope of the M9T and M9P was further characterized using a panel of aromatic acids **11a**-**11o**. Both mutants exhibited markedly enhanced productivity relative to WT (3 to 500-fold) across all substrates tested (Fig. 4n). Notably, *o*-methylphenylacetic acid (**11k**) was converted to *o*-xylene (45.9 mM) with 91.8% conversion, and *p*-fluorophenylacetic acid (**11i**) reached 54.4% conversion to afford *p*-fluorotoluene (27.2 mM). These results demonstrate that both CvFAP mutants serve as efficient photoenzymes for the decarboxylation of non-native aromatic substrates. In summary, through three rounds of uHTS, the fatty acid photodecarboxylase CvFAP was successfully reprogramed to efficiently convert non-native aromatic acids, demonstrating the power of uHTS for photoenzymes engineering.

Computational insights from molecular dynamics simulations

The structures of CvFAP-M9P and M9T were predicted by using SWISS-MODEL. Relative to WT, the substrate-binding pockets of both mutants were smaller by 11-15 $Å^3$ and more hydrophobic, with pocket hydropathy decreasing from −0.37 in WT to −0.47 (M9T) and −0.51 (M9P), respectively (Fig. 5a-c). The reduced volume and the enhanced hydrophobicity could facilitate better accommodation of the aromatic substrates. Furthermore, beneficial mutations were found to be predominantly localized in the loop region spanning residues 453-462: substitutions of two proline residues at the loop termini (P454A/T and P460A) increase its flexibility (Fig. 5d, Supplementary Fig. S17), whereas mutated to polar amino acids (P454T and G455S) within the loop introduced new hydrogen bonds (Fig. 5e,f, Supplementary Fig. S18 and S19). These conformational changes indicated that the loop modulates substrate binding, prompting us to perform a dynamic analysis of the substrate-enzyme complex.

According to the catalytic mechanism of CvFAP, the flavin adenine dinucleotide (FAD) cofactor absorbs photon energy and abstracts one electron from the carboxyl group of the substrate[9], underscoring the importance of the substrate-FAD distance[18]. During 100-ns MD simulations, we calculated three distances from the two carboxyl oxygens and the carboxyl carbon of the substrate to the FAD-N_5 atom (Fig. 5g). These

distances in both M9T and M9P were significantly shorter than those in WT (Fig. 5h,i), which is expected to accelerate electron transfer. The binding energies between substrates and enzymes were further evaluated: both M9T and M9P presented lower binding energies, -37.8 (M9T) vs -25.68 (WT) kcal mol^{-1} binding to **4**, and -49.20 (M9P) vs -24.22 (WT) kcal mol^{-1} binding to **9**, in 100-ns simulations (Supplementary Fig. S20), indicating strengthened substrate-enzyme interactions.

Beyond the mutations in the substrate-binding pocket, beneficial mutations identified via random mutagenesis were further analyzed. G423R in both mutants is located at the substrate entrance (Fig. 5j, Supplementary Fig. S21 and S22). Combined structural analysis and steered molecular dynamics (SMD) simulations demonstrated that it facilitates substrates entry through electrostatic attraction, as reflected by a lower free energy for the substrate entry to the CvFAP channel (Supplementary Fig. S23). G426P in both mutants is lining the substrate channel. Analysis of the loop harboring G426P revealed a lower RMSF value for this region (Supplementary Fig. S24). By stabilizing the conformational integrity of the entrance region, G426P may indirectly optimize and stabilize the geometric configuration of the substrate channel. CaverDock analysis further showed that the substrate translocation within the channel of the mutant enzymes is associated with reduced binding energy (Fig. 5k and 5l). Collectively, these MD simulations and binding energy analyses indicate that the combined mutations in M9P and M9T synergistically optimize substrate access, pocket geometry, and catalytic distance, thereby underpinning the substantial gains in catalytic efficiency toward aromatic substrates.

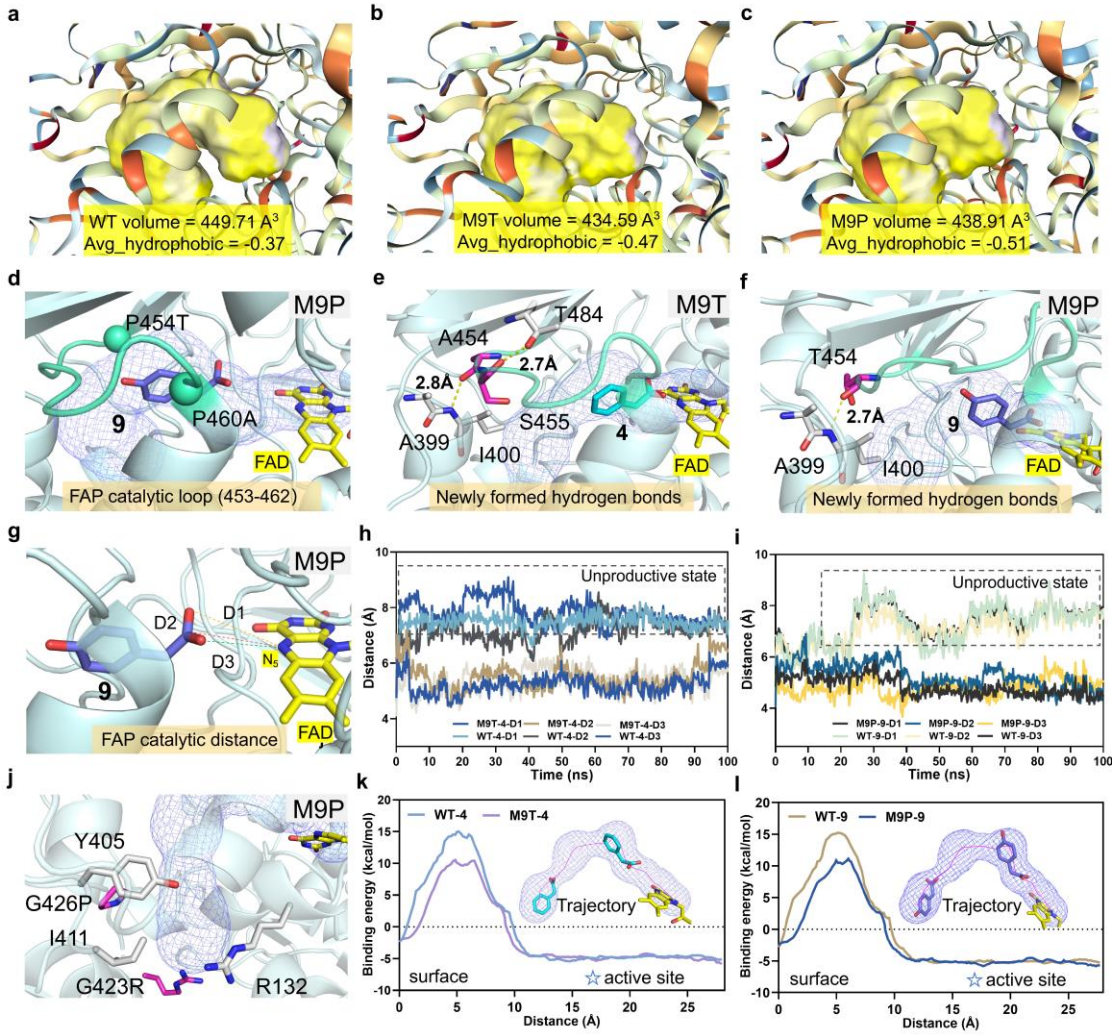


Fig.5 | Computational insights into the enhanced catalytic efficiency of CvFAP. **a-c**, Catalytic pocket volumes and hydropathy of CvFAP variants WT (**a**), M9T (**b**) and M9P (**c**). **d**, The loop region spanning residue 453-462 of M9P mutant (loop region, light green cartoon; **9**, purple stick model). **e-f**, Newly formed hydrogen bonds (yellow dashed lines) in the mutant M9T (**e**) and M9P (**f**). **g**, Three key catalytic distances between the substrate and CvFAP. D1, carboxyl O_1-FAD-N_5; D2, carboxyl C-FAD-N_5; D3, carboxyl O_2-FAD-N_5. **h**, Changes of catalytic distances between **4** and CvFAP WT or M9T during 100-ns MD simulations. **i**, Changes of catalytic distances between **9** and CvFAP WT or M9P during 100-ns MD simulations. **j**, Location of G423R and G426P near the substrate entry tunnel of CvFAP. **k**, Changes of binding energy during translocation of **4** in the channels of CvFAP WT and M9T. **l**, Changes of binding energy during translocation of **9** in the channels of CvFAP WT and M9P.

Bioproduction of toluene and *p*-cresol via enzyme cascades and fermentation

To enable the bioproduction of toluene (**5**) and *p*-cresol (**10**), we designed an enzyme cascade starting from L-phenylalanine (**1**) or L-tyrosine (**6**) (Fig. 6a).

L-Phenylalanine (**1**) is firstly converted to phenylpyruvate (**2**) by L-amino acid deaminase from *Proteus mirabilis* (PmLAAD)[57]; **2** is then transformed to phenylacetaldehyde (**3**) by the phenylpyruvate decarboxylase from *Saccharomyces cerevisiae* (ScPPDC)[58]; next, **3** is oxidized to phenylacetate (**4**) by phenylacetaldehyde dehydrogenase from *Escherichia coli* (EcPADH)[59]. The same enzyme cascade was applied to convert L-tyrosine (**6**) to *p*-hydroxyphenylacetate (**9**). In the final step, **4** or **9** are decarboxylated by CvFAP to yield **5** or **10**, respectively.

The genes encoding PmLAAD, ScPPDC, and EcPADH were assembled into an artificial operon in the plasmid pRSFduet-1. *E. coil* cells harboring this plasmid and co-expressing the three enzymes, were employed as whole-cell catalysts for cascade biotransformation of **1** to **4**. At 20 g l^{-1} cell loading, 10-50 mM of **1** was fully converted to **4** within 2-10 h; however, only 62.2% of 100 mM **1** was converted to **4** after 12 h (Fig. 6b, Supplementary Fig. S25). To address this, an NAD^+ recycling system (via co-expression of the NADH oxidase LbNOX[60]) was introduced, enabling full conversion of 100 mM **1** to **4**. The same whole-cell catalysts performed excellently on **6** without co-expressing LbNOX: 10-50 mM **6** was fully converted to **9** within 2-6 h, and 97.1% conversion of 100 mM **6** was achieved after 12 h (Fig. 6f). Furthermore, fed-batch addition of substrate **1** or **6** enabled the production of 198.5 mM titers of **4** or 199.9 mM titers of **9**, respectively (Fig. 6c,g, Supplementary Fig. S26).

After establishing optimal individual reaction conditions (Supplementary Fig. S27 and S28), we evaluated the integrated cascades in both one-cell and two-cell systems. Given that CvFAP catalyzes rapid decarboxylation (Fig. 6d,h) but it is unstable under prolonged illumination, the two-cell system was operated sequentially: *E. coli* cells co-expressing PmLAAD, ScPPDC, EcPADH (and LbNOX) firstly convert amino acids **1** or **5** into intermediates **4** or **9**, respectively; subsequently, CvFAP-M9T or M9P was added along with blue-light illumination to complete the reactions. This approach enabled, for the first time, high-titer bioproduction of **5** (67.9 mM, 6.3 g l^{-1}) and **10** (198.0 mM, 21.4 g l^{-1}) from natural amino acids via coupled cascade biotransformation and photobiocatalysis (Fig. 6e,i). In contrast, the one-cell

system, in which all enzymes are co-expressed and operate concurrently, showed lower production of 38.3 mM **5** and 89.7 mM **10** (Supplementary Fig. S29). This may attribute to inefficient expression of multiple enzymes, particularly CvFAP, in a single cell. Furthermore, insufficient substrate availability in concurrent mode may lead to light-induced damage to the CvFAP[61], thereby compromising its overall performance.

Finally, to demonstrate production from biobased glucose, high-cell-density fermentation was carried out in a 5-L bioreactor using the engineered strains described above for L-phenylalanine[62] and L-tyrosine[63] production. After two days of continuous fermentation and process, L-phenylalanine and L-tyrosine were obtained (Fig. 6j). Without further purification, these fermented crude amino acids were efficiently converted to toluene (4.3 g l^{-1}) and *p*-cresol (21.4 g l^{-1}) via coupled cascade biotransformation and photobiocatalysis (Fig. 6k).

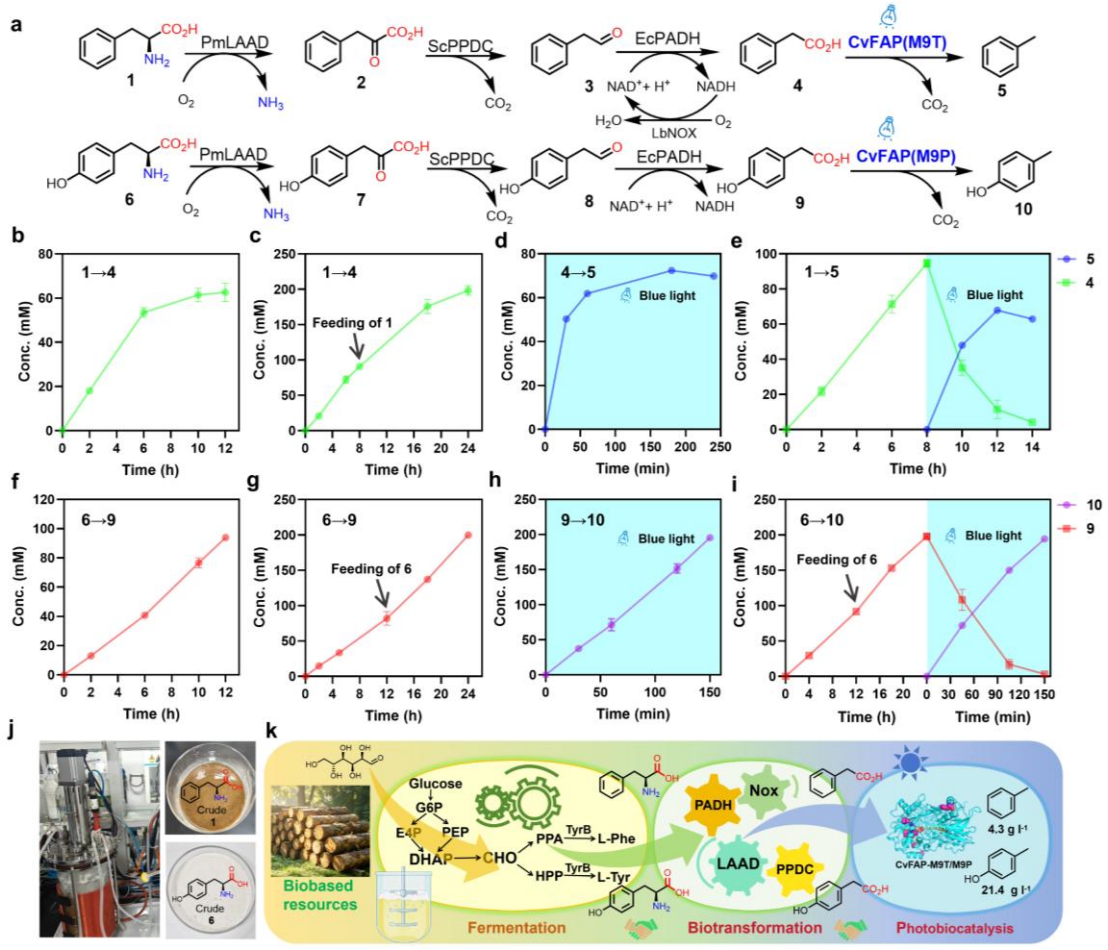


Fig.6 | Integrated multienzyme cascades and photobiocatalysis to produce toluene and *p*-cresol. a, Design of enzyme cascades to convert phenylalanine and

tyrosine to toluene and *p*-cresol, respectively. **b**, Cascade biotransformation of 100 mM **1** to **4** by *E. coli* co-expressing PmLAAD, ScPPDC, and EcPADH. **c**, Cascade biotransformation of 200 mM **1** to **4** by *E. coli* co-expressing PmLAAD, ScPPDC, EcPADH, and LbNOX. **d**, Photodecarboxylation of 100 mM **4** to **5** by *E. coli* expressing CvFAP-M9T in the biphasic system. **e**, Sequential biotransformation-photodecarboxylation of 100 mM **1** to **5** by *E. coli* co-expressing PmLAAD, ScPPDC, EcPADH, and LbNOX, and *E. coli* expressing CvFAP-M9T. **f**, Cascade biotransformation of 100 mM **6** to **9** by *E. coli* co-expressing PmLAAD, ScPPDC, and EcPADH. **g**, Cascade biotransformation of 200 mM **6** to **9** by *E. coli* co-expressing PmLAAD, ScPPDC, and EcPADH. **h**, Photodecarboxylation of conversion of 200 mM **9** to **10** by *E. coli* expressing CvFAP-M9P. **i**, Sequential biotransformation-photodecarboxylation of 200 mM **6** to **10** by *E. coli* co-expressing PmLAAD, ScPPDC, and EcPADH, and *E. coli* expressing CvFAP-M9P. **j**, Fermentative production of **1** and **6** in a 5-liter fermenter. **k**, Conversion of fermentative crude amino acids to produce **5** and **10** via coupled biotransformation and photobiocatalysis. Data in **b**, **c**, **d**, **e**, **f**, **g**, **h** and **i** are means ± s.d. of three technical replicates.

Conclusions

In this study, we established a comprehensive platform for the sustainable bioproduction of the important commodity aromatics, toluene and *p*-cresol, from biobased feedstocks, integrating ultrahigh-throughput screening, protein engineering, photobiocatalysis, and fermentation. Central to this work was the engineering of the fatty acid photodecarboxylase CvFAP to efficiently convert non-native aromatic substrates. To this end, we developed transcription factor DmpR-based biosensors for selective detection of toluene and *p*-cresol, and coupling them with FACS-based uHTS to screen large mutant libraries generated by multi-site saturation and random mutagenesis. Through three iterative rounds of screening, we isolated highly active CvFAP variants, notably M9T and M9P, with more than 200-fold improvements in productivity and TONs of up to 3,700 for *p*-cresol and 1,270 for toluene. MD simulations revealed that the accumulated mutations synergistically optimized substrate access, pocket volume and hydrophobicity, and catalytic distances, thereby enhancing substrate binding and electron transfer. These evolved photoenzymes were subsequently integrated into sequential two-cell cascades, together with upstream enzymatic conversion of L-phenylalanine and L-tyrosine, enabling photobiocatalytic production of toluene (67.9 mM) and *p*-cresol (198.0 mM) from amino acids.

Moreover, by coupling with high-cell-density fermentation of engineered strains, the cascades directly converted glucose-derived crude amino acids to toluene (4.3 g l^{-1}) and *p*-cresol (21.4 g l^{-1}) without intermediate purification. Collectively, this work not only showcases the power of uHTS in overcoming the catalytic efficiency ceiling of photoenzymes toward non-native substrates, but also provides a scalable and sustainable route from renewable biomass to industrially important aromatic chemicals, offering an alternative to conventional petroleum-based production. More broadly, the biosensor-enabled uHTS strategy demonstrated here should be transferable to the evolution of other photoenzymes, accelerating the development of photobiocatalytic transformations.

Methods

Materials and reagents

DNA fragments for plasmid assembly were amplified by polymerase chain reaction (PCR) using Phanta UniFi Super-Fidelity DNA Polymerase (P605, Vazyme). PCR primers (Supplementary Tables S3-S7) were synthesized at GeneCreate (Wuhan, China). The amplified fragments were purified using a DNA Gel Extraction Kit (BOLG901, BIOLOGY). Plasmids were constructed by seamless cloning using the ClonExpress Ultra One Step Cloning Kit V3 (Vazyme). Plasmids were extracted using a High-Pure Plasmid Mini-Prep Kit (BOLG905, BIOLOGY). The constructed plasmids were verified by Sanger sequencing at GeneCreate (Wuhan, China). Commercially available chemicals were purchased from suppliers including Sinopharm, Shanghai D&B Chemical, Adamas-beta, Energy Chemical, and FTSCI Science and Technology.

Characterization of whole-cell biosensors

The DmpR backbone plasmid was derived from pUC57-QS-DSF as described in a previous study[64]. To construct the biosensor, the LuxR/LuxI fragments were removed and the R0062 promoter driving superfolder GFP (sfGFP) was replaced with the native DmpR-responsive Po promoter, yielding the DmpR-sfGFP biosensor

plasmid. Recombinant *E. coli* BL21(DE3) strain harboring the biosensor plasmid was inoculated into a 20-ml reaction vial containing 4 ml of LB medium supplemented with 100 μg ml^{-1} ampicillin and cultured at 37 °C with shaking at 180 rpm for 12 h. The overnight culture was then diluted 100-fold into fresh LB medium with 100 μg ml^{-1} ampicillin and grown to an OD_{600} of ~0.8. Effectors (phenol, **5**, or **10** as DMSO stock solutions; **4** or **9** dissolved in KP buffer) were added to activate the biosensor, and the cultures were incubated at 30 °C for 16 h, with samples taken every 4 h for fluorescence measurement. At each time point, cells were harvested by centrifugation (10,000 × g, 6 min), washed twice with PBS buffer (100 mM, pH 7.0), and resuspended in PBS to an OD_{600} of 0.2–0.8. Aliquots (200 μl) were transferred to a 96-well microplate, and fluorescence (λ_{ex} = 485 nm; λ_{em} = 535 nm) was measured using a microplate reader (Spark, Tecan). Fluorescence values were normalized to the corresponding OD_{600}. Dose–response experiments were performed in 96-well microplates, and all measurements were carried out in three technical replicates.

Engineering of the toluene biosensor

DmpR mutant libraries targeting residues H100 and W128 were constructed by a two-step PCR method using degenerate primers (Supplementary Table S5). Library construction followed the procedures described in "Construction of CvFAP mutant libraries" below. To eliminate clones with leaky expression, the *sacB* gene encoding levansucrase was fused to sfGFP via a flexible (GGSGGGSGG) linker and used as the negative-selection marker. The DmpR mutant library plasmids were transformed into *E. coli* BL21(DE3) competent cells. The cells were inoculated into 4 ml of LB medium containing 100 μg ml^{-1} ampicillin in 20-ml reaction vials and cultured at 37 °C. When the OD_{600} reached ~0.8, **5** was added to a final concentration of 2 mM, followed by incubation at 30 °C to induce the biosensor response. After 10 h of incubation, the cells were harvested by centrifugation (10,000 × g, 6 min), washed twice with sterile PBS, and subjected to FACS (Sony MA900; see "Fluorescence-activated cell sorting" below). The top 0.5% of fluorescent cells were collected. The sorted cells were spread onto LB agar plates containing 100 g l^{-1} sucrose. After overnight growth, single colonies were picked and tested for activation

by **5** versus **4**. Clones showing strong activation by **5** but no activation by **4** were verified by Sanger sequencing and retained for further characterization.

Construction of CvFAP mutant libraries

The Saprot model was employed as described previously (SaprotHub_v2.ipynb)[56]. Prediction was performed on the Colab platform, using the CvFAP sequence lacking the signal peptide (residues 61-654) as input, and interpreted in conjunction with the CvFAP structure. For the multi-site saturation libraries (Lib1–Lib4), combinatorial mutagenic primers (Supplementary Table S6) were employed for simultaneous mutagenesis of three sites. The PCR product was purified and circularized by ligation with a kinase–ligase–*Dpn* I (KLD) enzyme mixture (22 °C, 12 h) in a 50-μl reaction containing 20 μl of PCR product, 5 μl of enzyme mixture and 25 μl of ligation buffer. The ligation products were purified and electroporated into *E. coli* DH5α; the cells were recovered in 500 μl of antibiotic-free LB medium (37 °C, 180 rpm, 2 h) and then transferred to 4 ml of LB medium containing 50 μg ml^{-1} kanamycin for overnight cultivation. An aliquot (100 μl) was spread on LB agar containing 50 μg ml^{-1} kanamycin to assess library coverage, and plasmids extracted from the remaining culture constituted the mutant library.

The random mutagenesis library (Lib5) was constructed from the pRSF-CvFAP plasmid using a random-mutagenesis kit (AG11809, Accurate); a template concentration of 0.5 ng μl^{-1} was used, maintaining a low mutation rate (0–4.5 mutations kb^{-1}) across the full-length CvFAP gene. The purified PCR products were assembled with the backbone vector by seamless cloning, and library construction followed the procedures described above.

Fluorescence-activated cell sorting

The CvFAP mutant libraries were electroporated into *E. coli* BL21(DE3) cells pre-harboring the corresponding biosensor (DmpR-H100E for screening toward **4**, or DmpR-M52I-Q71A for screening toward **9**). After recovery in antibiotic-free LB medium (500 μl, 37 °C, 180 rpm, 2 h), the cells were transferred to 4 ml of LB containing 100 μg ml^{-1} kanamycin and 100 μg ml^{-1} ampicillin and grown at 37 °C to an OD_{600} of ~0.8. CvFAP expression was induced with 0.1 mM isopropyl

β-D-thiogalactopyranoside (IPTG) and cultivation was continued at 30 °C for 5 h. The corresponding substrate **4** or **9** (1-4 mM) was then added, and the photocatalytic reaction was initiated by blue-light irradiation (450 nm) for 2-4 h. After the reaction, cells were collected and washed with PBS. Cells were analyzed and sorted on a Sony MA900 cell sorter. The flow rate was set to low; the FSC and SSC voltages were set to 10 and 11, respectively; the sample pressure was set to 4; and the event rate was maintained below 1,000 events s^{-1}. Cells exhibiting fluorescence in the top 0.5–1% of the population were sorted in purity mode into flow tubes. Aliquots of 50 ul, 100 ul and 200 ul were taken from the sorted cells in flow tubes and spread onto LB agar plates containing kanamycin (50 ug ml^{-1}), followed by overnight incubation at 37 °C. Single colonies grown on the plates were picked and transferred into a 24-well deep-well plate for subsequent validation of decarboxylation activity toward substrate **4** and substrate **9**.

Preparation of whole-cell catalysts

The recombinant strains were inoculated into 2 ml of LB medium containing 50 μg ml^{-1} kanamycin and cultured overnight (37 ℃, 200 rpm) as precultures. The precultures were transferred into 100 ml of TB medium containing 50 μg ml^{-1} kanamycin in 250-ml shake flasks and grown at 37 °C for ~4 h to an OD_{600} of ~0.8. After cooling at 4 ℃ for 15 min, protein expression was induced with 0.1 mM IPTG, and the cultures were incubated at 22 °C for 12 h (200 rpm). Cells were harvested by centrifugation (4,500 × g, 20 min), resuspended in KP buffer (0.2 M, pH 6.0) to afford the whole-cell catalysts, and stored at 4 °C until use. The CvFAP catalysts were protected from light during storage to prevent photodamage and irreversible enzyme inactivation and protein expression was verified by SDS-PAGE. For validation of initial screening results, the CvFAP mutants were cultured in 24-well deep-well plates under the same induction conditions.

Photocatalytic reaction by CvFAP

Photocatalytic reactions were performed in 5-ml vials with air-tight screw caps. To a 5-ml vial, whole-cell catalysts (final 10–20 g l^{-1}, dry cell weight), KP buffer (0.2 M, pH 6.0), and substrate stock solution in DMSO (final concentration 10–200 mM)

were added to form a 1-ml reaction system. The reactions were conducted at 22 °C under blue-light irradiation (450 nm) for 4 h. Unless otherwise stated, reactions were carried out in a single aqueous phase, and the product was extracted with 1 ml of methyl tert-butyl ether (MTBE) after the reaction. For toluene production, a biphasic system was applied owing to the high volatility of the product: *n*-hexadecane (1 ml) was added at a 1:1 (v/v) ratio to the aqueous phase, and the organic phase was collected after shaking for 3 min and centrifugation. During the biphasic system in 5-ml vials, the agitation speed was kept below 200 rpm to prevent emulsification of the aqueous and organic phases. In both cases, the organic phase (MTBE or *n*-hexadecane) was dried over anhydrous Na_2SO_4 and analyzed by GC, and product concentrations were normalized to the aqueous-phase volume.

GC analysis was performed on a gas chromatograph (Panna Instruments) equipped with an Agilent HP-5 column (30 m × 0.32 mm × 0.25 μm; FID; N_2 carrier gas) under the following temperature program: 90 °C (0.5 min), 20 °C min^{-1} to 200 °C, 80 °C min^{-1} to 320 °C (hold 1 min).

Construction of plasmids co-expressing PmLAAD, ScPPDC, and EcPADH

The CvFAP expression plasmid was described in a previous study[15]. The gene encoding PmLAAD was synthesized. The ScPPDC and EcPADH genes were amplified from the genomes of *Saccharomyces cerevisiae* and *Escherichia coli*, respectively. The three cascade genes were assembled as an artificial operon in pRSFDuet-1, and LbNOX was cloned into pACYCDuet-1 (Supplementary Table S4). The fragments were amplified with Phanta Flash Master Mix (Vazyme), purified, and assembled into the corresponding vectors with the ClonExpress Ultra One Step Cloning Kit V3 (Vazyme) following the manufacturer's instructions. The ligation products were transformed into *E. coli* DH5α, and the resulting plasmids were verified by Sanger sequencing.

Multi-enzyme cascade reaction

A two-cell system was employed for the cascade transformations. In the first stage, *E. coli* cells co-expressing PmLAAD, ScPPDC, and EcPADH (with or without LbNOX) converted **1** to **4** and **6** to **9**. To a 250-ml shaking flask, whole-cell catalysts

(final 10–20 g l^{-1}, dry cell weight), KP buffer (0.2 M, pH 6.0), and substrate (solid form, final concentration 10–100 mM) were added to form a 50-ml reaction system. The reaction mixture was incubated at 30 ℃ and 200 rpm. To achieve higher concentrations of **4** and **9**, additional feeding of substrate and whole-cell catalyst were performed: substrate **1** or **6** (solid form) was added at 12 h to an additional concentration of 100 mM, and whole-cell catalysts (using the reaction mixture to resuspend the cell pellet) were added at 12 h to an additional amount of 10 g l^{-1}. Samples were collected at regular intervals and analyzed by HPLC with an Agilent Poroshell 120 SB-C18 column (4.6 × 150 mm, 2.7 μm) and with a mobile phase of methanol/0.1% trifluoroacetic acid solution (50:50, v/v) at a flow rate of 1 ml min^{-1}. The column temperature was kept at 25 °C, and UV measurement was set to 210 nm and 280 nm. Upon completion, the supernatant was collected by centrifugation (10,000 × g, 10 min) and used to resuspend *E. coli* cells expressing the CvFAP variants for photocatalytic decarboxylation as described in the "Photocatalytic reaction by CvFAP" section.

Data availability

The data supporting the findings of this study are available within the article and its Supplementary Information. Data are also available from the corresponding authors upon request. Source data are provided with this paper.

References

1. Emmanuel, M. A. *et al.* Photobiocatalytic strategies for organic synthesis. *Chem. Rev.* **123**, 5459–5520 (2023).
2. Harrison, W., Huang, X. & Zhao, H. Photobiocatalysis for abiological transformations. *Acc. Chem. Res.* **55**, 1087–1096 (2022).
3. Fu, H. & Hyster T. K. From ground-state to excited-state activation modes: flavin-dependent "ene"-reductases catalyzed non-natural radical reactions. *Acc. Chem. Res.* **57**, 1087–1096 (2024).

4. Ma, Y. *et al.* Photobiocatalysis: more than just an interesting lab curiosity? *Chem. Catal.* **4**, 101077 (2024).

5. Yang, X. *et al*. Bridging chemistry and biology for light-driven new-to-nature enantioselective photoenzymatic catalysis. *Chem. Soc. Rev.* **54**, 5157–5188 (2025).

6. East, N., Polyzos, A. & Scott, C. Beyond flavoproteins: toward the industrialization of photobiocatalysis. *ACS Catal.* **15**, 20694−20716 (2025).

7. Taylor, A., Heyes, D. & Scrutton, N. S. Catalysis by Nature's photoenzymes. *Curr. Opin. Struct. Biol.* **77**, 102491 (2022).

8. Sorigué, D. *et al.* An algal photoenzyme converts fatty acids to hydrocarbons. *Science* **357**, 903–907 (2017).

9. Sorigué, D. *et al.* Mechanism and dynamics of fatty acid photodecarboxylase. *Science* **372**, eabd5687 (2021).

10. Xu, W., Sun, J., Chen, H. & Wu, Q. Tunable catalytic performance and new-to-nature reactions of fatty acid photodecarboxylase. *Acc. Chem. Res.* **59**, 382–396 (2026).

11. Bassiony, E., Margariti, A., Eser, B. E. & Kara, S. Advances and challenges in fatty acid photodecarboxylases: towards greener chemistry applications? *Green Chem.* **28**, 13103–13121 (2026).

12. Huijbers, M. M. E., Zhang, W., Tonin, F. & Hollmann, F. Light-driven enzymatic decarboxylation of fatty acids. *Angew. Chem. Int. Ed.* **57**, 13648–13651 (2018).

13. Amer, M. *et al.* Low carbon strategies for sustainable bio-alkane gas production and renewable energy. *Energy Environ. Sci.* **13**, 1818–1831 (2020).

14. Xu, W. *et al.* Rational design of fatty acid photodecarboxylase enables the efficient decarboxylation of medium- and short-chain fatty acids for the production of gasoline bio-alkanes. *Mol. Catal.* **524**, 112261 (2022).

15. Qin, Z. *et al.* Production of biobased ethylbenzene by cascade biocatalysis with an engineered photodecarboxylase. *Angew. Chem. Int. Ed.* **63**, e202314566 (2024).

16. Xu, J. *et al.* Light-driven kinetic resolution of α-functionalized carboxylic acids enabled by an engineered fatty acid photodecarboxylase. *Angew. Chem. Int. Ed.* **58**, 8474–8478 (2019).

17. Mou, K. *et al*. Stereodivergent protein engineering of fatty acid photodecarboxylase for

light-driven kinetic resolution of sec-alcohol oxalates. *Angew. Chem. Int. Ed.* **63**, e202318374 (2024).

18. Xu, J. *et al.* Light-driven decarboxylative deuteration enabled by a divergently engineered photodecarboxylase. *Nat. Commun.* **12**, 3983 (2021).
19. Ju, S. *et al.* Stereodivergent photobiocatalytic radical cyclization through the repurposing and directed evolution of fatty acid photodecarboxylases. *Nat. Chem.* **16**, 1339–1347 (2024).
20. Xu, W. *et al.* Catalytic promiscuity of fatty acid photodecarboxylase enables stereoselective synthesis of chiral α-tetralones. *Angew. Chem. Int. Ed.* **63**, e202412862 (2024).
21. Weissensteiner, F. *et al.* Establishing the fatty acid photodecarboxylase CvFAP as a platform for photobiocatalytic radical transformations. *ACS Catal.* **16**, 7758–7770 (2026).
22. Bozkurt, E. U., Ørsted, E. C., Volke D. C. & Nikel P. I. Accelerating enzyme discovery and engineering with high-throughput screening. *Nat. Prod. Rep.* **43**, 313–334 (2026).
23. Markel, U. *et al.* Advances in ultrahigh-throughput screening for directed enzyme evolution. *Chem. Soc. Rev.* **49**, 233–262 (2020).
24. Gantz, M., Neun, S., Medcalf, E. J., van Vliet, L. D. & Hollfelder, F. Ultrahigh-throughput enzyme engineering and discovery in in vitro compartments. *Chem. Rev.* **123**, 5571−5611 (2023).
25. Hecko, S. *et al.* Enlightening the path to protein engineering: chemoselective turn-on probes for high-throughput screening of enzymatic activity. *Chem. Rev.* **123**, 2832−2901 (2023).
26. Vidal, L. S., Isalan, M., Heap, J. T. & Ledesma-Amaro, R. A primer to directed evolution: current methodologies and future directions. *RSC Chem. Biol.* **4**, 271–291 (2023).
27. Wang, Y. *et al.* Directed evolution: methodologies and applications. *Chem. Rev.* **121**, 12384–12444 (2021).
28. Zimmerman, J. B., Anastas, P. T., Erythropel, H. C. & Leitner, W. Designing for a green chemistry future. *Science* **367**, 397–400 (2020).
29. Prather, K. L. J. Accelerating and expanding nature to address its greatest challenges. *Nat. Catal.* **3**, 181–183 (2020).

30. Intasian, P. *et al.* Enzymes, in vivo biocatalysis, and metabolic engineering for enabling a circular economy and sustainability. *Chem. Rev.* **121**, 10367−10451 (2021).

31. Wirt, S. A. & Prather, K. L. J. Metabolic pathway design for BTEX-based commodity chemicals: drop-ins and alternatives. *Curr. Opin. Biotech.* **100**, 103546 (2026).

32. Fabri, J., Graeser, U. & Simo, T. A. *Toluene. Ullmann's Encyclopedia of Industrial Chemistry* (Wiley-VCH, 2011).

33. Fiege, H. *Cresols and Xylenols. Ullmann's Encyclopedia of Industrial Chemistry* (Wiley-VCH, 2000).

34. Beller, H. R. *et al.* Discovery of enzymes for toluene synthesis from anoxic microbial communities. *Nat. Chem. Biol.* **14**, 451–457 (2018).

35. Kriek, M., Martins, F., Challand, M. R., Croft, A. & Roach, P. L. Thiamine biosynthesis in *Escherichia coli*: identification of the intermediate and by-product derived from tyrosine. *Angew. Chem. Int. Ed.* **46**, 9223–9226 (2007).

36. Yu, L., Blaser, M., Andrei, P. I., Pierik, A. J. & Selmer, T. 4-Hydroxyphenylacetate decarboxylases: properties of a novel subclass of glycyl radical enzyme systems. *Biochemistry* **45**, 9584–9592 (2006).

37. Lu, Q. *et al.* The glycyl radical enzyme arylacetate decarboxylase from *Olsenella scatoligenes*. *ACS Catal.* **11**, 5789−5794 (2021).

38. Rodrigues, A. V., Tantillo, D. J., Mukhopadhyay, A., Keasling, J. D. & Beller, H. R. Insight into the mechanism of phenylacetate decarboxylase (PhdB), a toluene-producing glycyl radical enzyme. *ChemBioChem*, **21**, 663–671 (2020).

39. Hossain, G. S., Saini, M., Miyake, R., Ling, H. & Chang, M.W. Genetic biosensor design for natural product biosynthesis in microorganisms. *Trends Biotechnol.* **38**, 797–810 (2020).

40. Bayer, T., Hänel, L., Husarcikova, J., Kunzendorf, A. & Bornscheuer, U. T. In vivo detection of low molecular weight platform chemicals and environmental contaminants by genetically encoded biosensors. *ACS Omega* **8**, 23227–23239 (2023).

41. Lu, M. *et al*. Transcription factor-based biosensor: a molecular-guided approach for advanced biofuel synthesis. *Biotechnol. Adv.* **72**, 108339 (2024).

42. Zhang, J., Zhang, D., Chan, S., Chang, M. W. & Poh, C. L. Emerging biosensor and

assay-enabled high-throughput screening solutions for enzyme and strain engineering. *Curr. Opin. Biotechnol.* **98**, 103439 (2026).

43. France, S. P., Hepworth, L. J., Turner, N. J. & Flitsch, S. L. Constructing biocatalytic cascades: in vitro and in vivo approaches to de novo multi-enzyme pathways. *ACS Catal.* **7**, 710–724 (2017).
44. Schrittwieser, J. H., Velikogne, S., Hall, M. & Kroutil, W. Artificial biocatalytic linear cascades for preparation of organic molecules. *Chem. Rev.* **118**, 270−348 (2018).
45. Benítez-Mateos, A. I., Roura Padrosa, D. & Paradisi, F. Multistep enzyme cascades as a route towards green and sustainable pharmaceutical syntheses. *Nat. Chem.* **14**, 489−499 (2022).
46. 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).
47. Shingler, V. & Moore, T. Sensing of aromatic compounds by the DmpR transcriptional activator of phenol-catabolizing *Pseudomonas* sp. strain CF600. *J. Bacteriol.* **176**, 1555–1560 (1994).
48. Kim, H. *et al.* Machine learning linked evolutionary biosensor array for highly sensitive and specific molecular identification. *Biosens. Bioelectron.* **170**, 112670 (2020).
49. Kunjapur, A. M. & Prather, K. L. Development of a vanillate biosensor for the vanillin biosynthesis pathway in *E. coli*. *ACS Synth. Biol.* **8**, 1958–1967 (2019).
50. d'Oelsnitz, S. *et al.* Using fungible biosensors to evolve improved alkaloid biosyntheses. *Nat. Chem. Biol.* **18**, 981–989 (2022).
51. Baiyoumy, A. *et al.* Directed evolution of a surface-displayed artificial allylic deallylase relying on a GFP reporter protein. *ACS Catal.*, **11**, 10705–10712 (2021).
52. d'Oelsnitz, S. *et al.* Using enantioselective biosensors to evolve asymmetric biocatalysts. *Nat. Chem. Biol.* (2026) https://doi.org/10.1038/s41589-026-02275-1
53. Recorbet, G. *et al.* Conditional suicide system of *Escherichia coli* released into soil that uses the *Bacillus subtilis* sacB gene. *Appl. Environ. Microbiol.* **59**, 1361–1366 (1993).
54. Park, K.-H. *et al.* Tetrameric architecture of an active phenol-bound form of the AAA+ transcriptional regulator DmpR. *Nat. Commun.* **11**, 2728 (2020).

55. Li, D., Wu, Q. & Reetz, M. T. Focused rational iterative site-specific mutagenesis (FRISM). *Methods Enzymol.* **643**, 225–242 (2020).

56. Su, J. *et al*. Democratizing protein language model training, sharing and collaboration. *Nat. Biotechnol.* (2025) https://doi.org/10.1038/s41587-025-02859-7

57. Molla, G., Melis, R. & Pollegioni, L. Breaking the mirror: L-amino acid deaminase, a novel stereoselective biocatalyst. *Biotechnol. Adv. 35*, 657–668 (2017).

58. Kneen, M. M. *et al.* Characterization of a thiamin diphosphate-dependent phenylpyruvate decarboxylase from *Saccharomyces cerevisiae*. *FEBS J.* **278**, 1842–1853 (2011).

59. Ferrández, A., Prieto, M. A., García, J. L. & Díaz, E. Molecular characterization of PadA, a phenylacetaldehyde dehydrogenase from *Escherichia coli*. *FEBS Lett*. **406**, 23–27 (1997).

60. Geueke, B., Riebel, B. & Hummel, W. NADH oxidase from Lactobacillus brevis: a new catalyst for the regeneration of NAD. *Enzyme Microb. Technol.* **32**, 205–211 (2003).

61. Wu, Y., Paul, C. E. & Hollmann, F. Stabilisation of the fatty acid decarboxylase from *chlorella variabilis* by caprylic acid. *ChemBioChem* **22**, 2420–2423 (2021).

62. 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).

63. 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).

64. He, J. *et al*. An engineered quorum-sensing-based whole-cell biosensor for active degradation of organophosphates. *Biosens. Bioelectron.* **206**, 114085 (2022).

Acknowledgements

We thank Zhi Li at National University of Singapore for helpful discussion of enzyme

cascades and providing the fermentation strain producing L-phenylalanine, Jian Xu and Shuyun Ju at Zhejiang University of Technology for helpful discussion of CvFAP, and Thomas Bayer at University of Bern and Nasreen Munshi at Nirma University for helpful discussion of biosensors. Gas chromatography-mass spectrometry data were acquired at the National Key Laboratory of Agricultural Microbiology Core Facility.

Author contributions

S.W. and U.T.B. conceived and initiated the project, and together with Y.Z., secured funding and supervised the work. Z.Q. engineered the biosensors, performed the uHTS of CvFAP, MD simulations, fermentation, cascade biotransformation, and photobiocatalysis. S.P. assisted the most of experiments. W.Z. conducted the SMD simulations and assisted in mechanistic interpretation. Z.X. assisted the fermentation and cascade biotransformation. Y.B. provided the DmpR plasmid and valuable suggestions on biosensors. M.H. provided valuable suggestions on CvFAP and photobiocatalysis. T.R.W. and U.T.B. provided valuable suggestions throughout the project. Z.Q. and S.W. wrote the manuscript, edited by Y.Z., T.R.W. and U.T.B., and approved by all authors.

Funding

We acknowledge the National Natural Science Foundation of China (No. 32100060 to Y.Z. and 22378155 to S.W.), Hubei Special Project for Science Development (No. 2024CSA060 to S.W.), Fundamental Research Funds for the Central Universities (No. 2662021SKQD001 to Y.Z. and 2662025PY011 to S.W.), the College of Life Science and Technology (LongYun Program 2.0 to S.W.), and Alexander von Humboldt-Stiftung (Humboldt Research Fellowship to S.W.).

Competing interests

The authors, Z.Q., Y.Z., S.P., Z.X. and S.W., are inventors on a related Chinese patent (CN122382042A). All other authors declare no additional competing interests.

Additional information

Supplementary information The online version contains supplementary material available at

Correspondence and requests for materials should be addressed to Yi Zhou, Uwe T. Bornscheuer, or Shuke Wu.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 20261015SICvFAPforTolueneandpCresol.pdf