Supplementary Information (SI) for Green Chemistry.
This journal is © The Royal Society of Chemistry 2025

# Supporting Information

## An artificial biocatalytic cascade for the synthesis of aryl L-*threo*-β-hydroxy-α-aminobutyric acids via carbon-chain extension strategy

Sisi Zhang,[‡a] Xi Chen,[‡b] Mengyao Chen,[a] Xiaohe Zhang,[a] Tianming Li,[a] Jianxiong Hao,[*a] Jinhui Feng,[b] Qiaqing Wu,[b] Dunming Zhu [b] and Qijia Chen[*a]

[1] College of Food Science and Biology, Hebei University of Science and Technology, Shijiazhuang 050018, China

[2] Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China

**Table of Contents**

1. Experimental Procedures................................................................S3-S8

2. Protein Sequences....................................................... ..............S9-S10

3. Supporting Figures......................................................................S11-S14

4. Supporting Table.........................................................................S15

5. References.....................................................................................S16

## Experimental Procedures

### Materials

The plasmids pET-21a, pET-28a, pET-32a, pACYCDuet-1, pCDFDuet-1, pETDuet-1, and pRSFDuet-1 were used to construct expression vectors. *E. coli* DH5α and *E. coli* BL21 (DE3) were used as cloning and expression hosts, respectively. Yeast alcohol dehydrogenase (ADH) and flavin mononucleotide (FMN) were obtained from Merck. Restriction enzymes and T4 DNA ligase were purchased from Thermo. ClonExpress II one step cloning kit was purchased from Vazyme. All other chemicals were purchased from Merck, Thermo, and Sinopharm Chemical.

### Construction of recombinant plasmids

TA from *Pseudomonas* sp. and PSDH from *P. xenovorans* were cloned as previously reported.[1] The gene sequences of PPDC from *A. Brasilense*, *P. fluorescens* and *S. cerevisiae*, TTA from *P. fluorescens*, ALDH from *P. terricola* and FR from *M. goodii* were codon-optimized, synthesized, and subcloned into pET-21a, pET-32a, or pRSFDuet-1.[1-5]

The PxPSDH gene was inserted into the *Nde* I and *Bgl* II digested pRSF-*ta*, resulting in the generation of the plasmid pRSF-*ta-psdh*. The fusion fragments containing RBS and PfPPDC were double-digested by *Bgl* II and *Xho* I, and then inserted into pRSF-*ta-psdh*. The resulting *ta-psdh-ppdc* gene was also sub-cloned to the other three plasmids, pACYCduet-1, pCDFduet-1, and pETduet-1, respectively. The PfTTA gene was inserted into a set of compatible plasmids between *Nco* I and *Bam*H

I. The fusion fragment containing PtALDH and thioredoxin protein, and MgFR were sequentially constructed on multiple cloning sites (MCS) I and MCS II by homologous recombination to give pACYC-*aldh-fr*, pCDF-*aldh-fr*, pET-*aldh-fr*, and pRSF-*aldh-fr*, respectively. Similarly, genetic engineering of recombinant plasmids bearing *ta-psdh* and *ppdc-tta* was also performed.

**Genetic engineering of e. coli strains co-expressing the multiple enzymes**

A set of compatible plasmids encoding the *ta-psdh-ppdc* gene was transformed into give *E. coli* (AAD 1-4). Each of the *tta* plasmids (pACYC-*tta*, pCDF-*tta*, pET-*tta*, and pRSF-*tta*) and each of the *aldh-fr* plasmids (pACYC-*aldh-fr*, pCDF-*aldh-fr*, pET-*aldh-fr*, and pRSF-*aldh-fr*) were co-transformed to give *E. coli* (TOR 1-12). Similarly, we also constructed *E. coli* (AD 1-4) co-expressing PsTA-PxPSDH, as well as *E. coli* (DTOR 1-12) co-expressing PfPPDC-PfTTA and PtALDH-MgFR.

**Cell growth and enzymatic activity assays**

The recombinant *E. coli* cells were grown overnight at 37 °C in LB medium supplemented with appropriate antibiotics (50 mg/L chloramphenicol, 50 mg/L streptomycin, 100 mg/L ampicillin and/or 50 mg/L kanamycin). One milliliter of seed culture was transferred to 50 mL of fresh LB containing appropriate antibiotics. After the $OD_{600}$ of the culture reached 0.6-0.8, a final concentration of 0.5 mM isopropyl-β-D-thio-galactopyranoside (IPTG) was added, and the mixture was incubated overnight at 25 °C. The cells were harvested by centrifugation (5000 ×*g*, 15 min).

L-TA activity was assessed using a mixture comprising 10 mM L-threonine, 0.1 mM PLP, 10 U ADH, 0.2 mM NADH and appropriate quantities of L-TA in Tris-HCl buffer (50 mM, pH 7.0). For L-TTA, the assay mixture contained 10 mM **1a**, 30 mM L-threonine, 0.1 mM PLP, 10 U ADH, 0.2 mM NADH and enzyme solution in Tris-HCl buffer (50 mM, pH 7.0). FR activity was assessed using a mixture comprising 0.3 mM FMN, 0.2 mM NADH and required amount of FR in Tris-HCl buffer (50 mM, pH 8.0). For these three enzymes, the reaction rates were measured by monitoring the decrease in absorbance at 340 nm ($\varepsilon$=6.22 $mM^{-1}$ $cm^{-1}$). ALDH activity was assessed in Tris-HCl buffer (50 mM, pH 9.0) containing 2 mM acetaldehyde, 1 mM $NAD^+$, 10 mM DTT, 100 mM KCl and enzyme solution. The reaction rate was measured by monitoring the increase in absorbance at 340 nm.

PSDH activity was evaluated in Tris-HCl buffer (50 mM, pH 7.0) containing 20 mM DL-β-phenylserine, 0.1 mM PLP and enzyme solution. PPDC activity was assessed in Tris-HCl buffer (50 mM, pH 7.5) containing 5 mM **3a**, 0.2 mM ThDP, 5 mM $MgCl_2$ and appropriate quantities of PPDC. The activities of PSDH and PPDC were spectrophotometrically determined according to the $FeCl_3$ and purpald colorimetric assays, respectively.[1]

**Biotransformation procedures**

Biotransformation of **1a** to **4a** was carried out at 30 °C for 6 h with 1 mL of Tris-HCl buffer (100 mM, pH 8.0) containing 10 mM **1a**, 100 mM glycine, 0.1 mM PLP, 0.5 mM ThDP, 5 mM $MgCl_2$, 10% (*v/v*) DMSO and 50 mg wet cell weight of *E. coli* (AAD). The concentrations of **1a**, **3a**, and **4a** were analyzed by HPLC.

Biotransformation of **4a** to **5a** was performed at 30 °C for 6 h with 1 mL of Tris-HCl buffer (100 mM, pH 8.0) containing 10 mM **4a**, 30 mM L-threonine, 0.1 mM PLP, 0.2 mM $NAD^+$, 5 mM FMN, 10% (*v/v*) DMSO, and 50 mg wet cell weight of *E. coli* (TOR). To determine the conversion and de value of the products, the reaction products were analysed by HPLC (Agilent 1260 Infinity) after derivatization with *o*-phthaldialdehyde/*N*-acetyl-cysteine (OPA/NAC) using a Zorbax Eclipse Plus C18 column (250 mm*4.6 mm, Agilent) at 30 °C with a tunable UV detector at 254 nm.[6]

Biotransformation of **1a** to **3a** was performed at 30 °C for 6 h with 1 mL of Tris-HCl buffer (100 mM, pH 8.0) containing 10 mM **1a**, 100 mM glycine, 0.1 mM PLP, 10% (*v/v*) DMSO and 50 mg wet cell weight of *E. coli* (AD). Biotransformation of **3a** to **5a** was performed at 30 °C for 6 h with 1 mL of Tris-HCl buffer (100 mM, pH 8.0) containing 10 mM **3a**, 30 mM L-threonine, 0.1 mM PLP, 0.2 mM $NAD^+$, 0.5 mM ThDP, 5 mM $MgCl_2$, 5 mM FMN, and 50 mg wet cell weight of *E. coli* (DTOR).

**Optimization of the reaction conditions**

The effect of the reaction temperature on *E. coli* (DTRO 11) was studied at various temperatures (25-37 °C). To investigate the effects of pH, the reactions were carried out in Tris-HCl (pH 7.0-8.5) and CHES-NaOH (pH 9.0). To explore the concentrations of the reaction components, various concentrations of ThDP (0-1 mM), $MgCl_2$ (0-10 mM), $NAD^+$ (0-1 mM), FMN (0-10 mM), and **3a** (10-50 mM) were added to the standard reaction mixtures. The optimal amount of whole-cell catalyst biomass (10-50 mg/mL) was investigated. All the experiments were performed in triplicate.

**Substrate scope**

Substrate scope was measured under optimal reaction conditions. The standard reaction mixtures consisted of 20 mM aldehyde **1**, 200 mM glycine, 0.1 mM PLP, 10% (*v/v*) DMSO, and 20 mg wet cell weight of *E. coli* (AD 4) mixed in Tris-HCl buffer (100 mM, pH 8.0) in a final volume of 1 mL. The reaction was carried out at 30 °C for 6 h and terminated by centrifugation (10,000 ×*g*, 1 min) followed by boiling (1 min) to remove the whole-cell catalyst. A mixture of 60 mM L-threonine, 0.2 mM ThDP, and 20 mg wet cell weight of *E. coli* (DTOR 11) was added to the supernatants of the aqueous-phase samples. The reaction was performed at 30 °C for 12 h. The resulting solution was used to analyze the conversion and de value.

**Scale-up synthesis of β-OH-homePhe 5a and β-OH-*p*-$NO_2$-homePhe 5p**

A solution of **1a** (106 mg) or **1p** (151 mg) in 5 mL of DMSO was added to 45 mL of Tris-HCl buffer (100 mM, pH 8.0) containing glycine (750 mg), PLP (1.2 mg) and wet cell weight of *E. coli* (AD 4) (1.0 g). When the aldehyde was completely consumed, the reaction was were terminated by centrifugation, followed by boiling to remove the whole-cell catalyst. And then, L-threonine (357 mg), ThDP (4.6 mg), and wet cell weight of *E. coli* (DTOR) (1.0 g) were subsequently added to the supernatants of the reaction samples.

At the end of the reaction, the mixture was centrifuged (10,000 ×*g*, 5 min) to remove cells. The supernatant was extracted with ethyl acetate, diluted with 450 mL MeOH and stored at 4°C overnight. The mixture was filtered to remove residual glycine and L-threonine, and concentrated under vacuum to 50 mL. The desired products, as

well as the remaining glycine and L-threonine, were isolated through an activated carbon adsorption method.[7] The products were further purified by flash chromatography on a silica gel column.

**Protein sequences**

PsTA from *Pseudomonas* sp.

Protein sequence:

MTDQSQQFASDNYSGICPEAWAAMEKANHGHERAYGDDQWTARAADHFR
KLFETDCEVFFAFNGTAANSLALSSLCQSYHSVICSETAHVETDECGAPEFFSN
GSKLLTARSEGGKLTPASIREVALKRQDIHYPKPRVVTITQATEVGSVYRPDE
LKAISATCKELGLNLHMDGARFSNACAFLGCTPAELTWKAGIDVLCFGGTKN
GMAVGEAILFFNRKLAEDFDYRCKQAGQLASKMRFLSAPWVGLLEDGAWL
RHAAHANHCAQLLSSLVADIPGVELMFPVEANGVFLQMSEPALEALRNKGW
RFYTFIGSGGARFMCSWDTEEARVRELAADIRAVMSA

PxPSDH from *Paraburkholderia xenovorans*

Protein sequence:

MSTAAPQHTDHTIDGEPIPTLDDIAAQHFALTPWVARTPVFDRLDFPSLEGTL
VNFKFELLQAGGSFKARGAFTNLLALDESQRSAGVTCVSGGNHAVAVAYAA
MRLGISAKVVLFRAANPARVALCRQYRADIVFAENIAEAFELVRRIEAEEGRY
FVHPFNGYRTVLGSATLGYEWATQTPDLEAVIVPIGGGGLAAGVATAMRLA
NPNVHLYGVEPEGADVMGKSFAANHTVRMGKMHGIADSLMSPHTEEYSYE
LCRRHIDQLVTVSDDQLRAAMLTLFGQLKLAVEPACAAATAALLGPLREQLQ
GKRVGVLLCGTNTDPVTFAAHIERARHSESLFPQ

PfPPDC from *Pseudomonas fluorescens*

Protein sequence:

MTNLPSTHINLTSEEVSLGDLVGRVLVESGIDDLFCIPGDFTMQLSRELLTTPG
LALRTMSHEYGTTLAALGYAVGKGVPGAVCFTYGVGVLNATNAIAQAYVE
RVPLLVFSGSPGTRERQAPLFLHHTIVDHQTQYRIMKEITVHQVCVTDPHQVL
EQLREAVALAVLHSRPVYIEIPRDLFQARVRYSPARRVPLEPSTRYSQAARQA
AELAYALVRKARDPVFVPGLDLKRRGLTDLAMRVCERLAMPWVATPMSRG
GIPVSHPNYRGIYAGPASPSRVTRELLAKCDVLMLIGEPNSDVNMGIASHIAK
GRLIHADDGKISVGRQHFNASTAEFLIAFSDVIHNAKTPLAPLTETQKDFIVPT
PASLYPSEESPLTPFDIINELNRHFVTQPDTQLVVDCGDVFFMSLGMFPADVLT
SPLYMSMGIAVPGAMGYQLGTGKRPIVLVGDGAFHMTGNELMRAAKFGLSP
IVIVLNNQRWASLSSDAADIALTEQMPMSFSAAGQFLQVQAFTATTGRELRQ
HLEEALNMDRPVLIDAQVDPSKRSYLCERFFDAVKGQQHLPKA

PfTTA from *Pseudomonas fluorescens*

Protein sequence:

MSNVKQQTAQIVDWLSSTLGKDHQYREDSLSLTANENYPSALVRLTSGSTAG
AFYHCSFPFEVPAGEWHFPEPGHMNAIADQVRDLGKTLIGAQAFDWRPNGGS
TAEQALMLAACKPGEGFVHFAHRDGGHFALESLAQKMGIEIFHLPVNPTSLLI
DVAKLDEMVRRNPHIRIVILDQSFKLRWQPLAEIRSVLPDSCTLTYDMSHDGG
LIMGGVFDSPLSCGADIVHGNTHKTIPGPQKGYIGFKSAQHPLLVDTSLWVCP
HLQSNCHAEQLPPMWVAFKEMELFGRDYAAQIVSNAKTLARHLHELGLDVT
GESFGFTQTHQVHFAVGDLQKALDLCVNSLHAGGIRSTNIEIPGKPGVHGIRL
GVQAMTRRGMKEKDFEVVARFIADLYFKKTEPAKVAQQIKEFLQAFPLAPLA
YSFDNYLDEELLAAVYQGAQR

PtALDH from *Pichia terricola*

Protein sequence:

MLRTATRTTFKSAQPSFMAAAAALRYYSHYPLSKSITLPNGKVYEQPTGLFIN
GEFVASQQHKTFEVINPSNEDEICHVYEARTEDVDAAVDAAYNAFHSEWSK
MDPSIRGEHLMKLAELMEKNKDTLAAIESMDNGKALFMAEIDVKLVINYLK
YCAGWADKLFGKVVDTGSNYFNYIKREPIGVCGQIIPWNFPLLMWSWKVGP
ALAAGNTIVLKTAESTPLSALYAAKLAKEAGIPDGVINIVSGFGKITGEAISTH
PKIKKLAFTGSTATGKHIMKAAAESNLKKVTLELGGKSPNIVFNDADIKKAV
NNLILGIFFNSGEVCCAGSRVFIQEGVYDQVLEEFKIAAEALKVGNPFEEGVFQ
GAQTSQQQLTKILGYVESGKDEGATLVTGGERLGDKGYFVKPTIFADVKPNM
KIYSEEIFGPFAVVTKFKTADEAIAMANDSEYGLAAGIHTTSLDTATYVANNL
EAGTVWINTYNDFHHNMPFGGFKQSGIGREMGEAAFENYTQWKTVRIAIND
GPQ

MgFR from *Mycobacterium goodii*

Protein sequence:

MSATDLSPTSLREAFGHFPSGVIAIAAEVDGTRVGLAASTFVPVSLEPPLVAFC
VQNSSTTWPKLKDLPSLGISVLGEAHDTAARTLAAKTGDRFAGLETESRDSG
AVFINGTSVWLESAIEQLVPAGDHTIVVLRVSDIVINEAVPPIVFHRSAFRKLG
A

## Supporting Figures

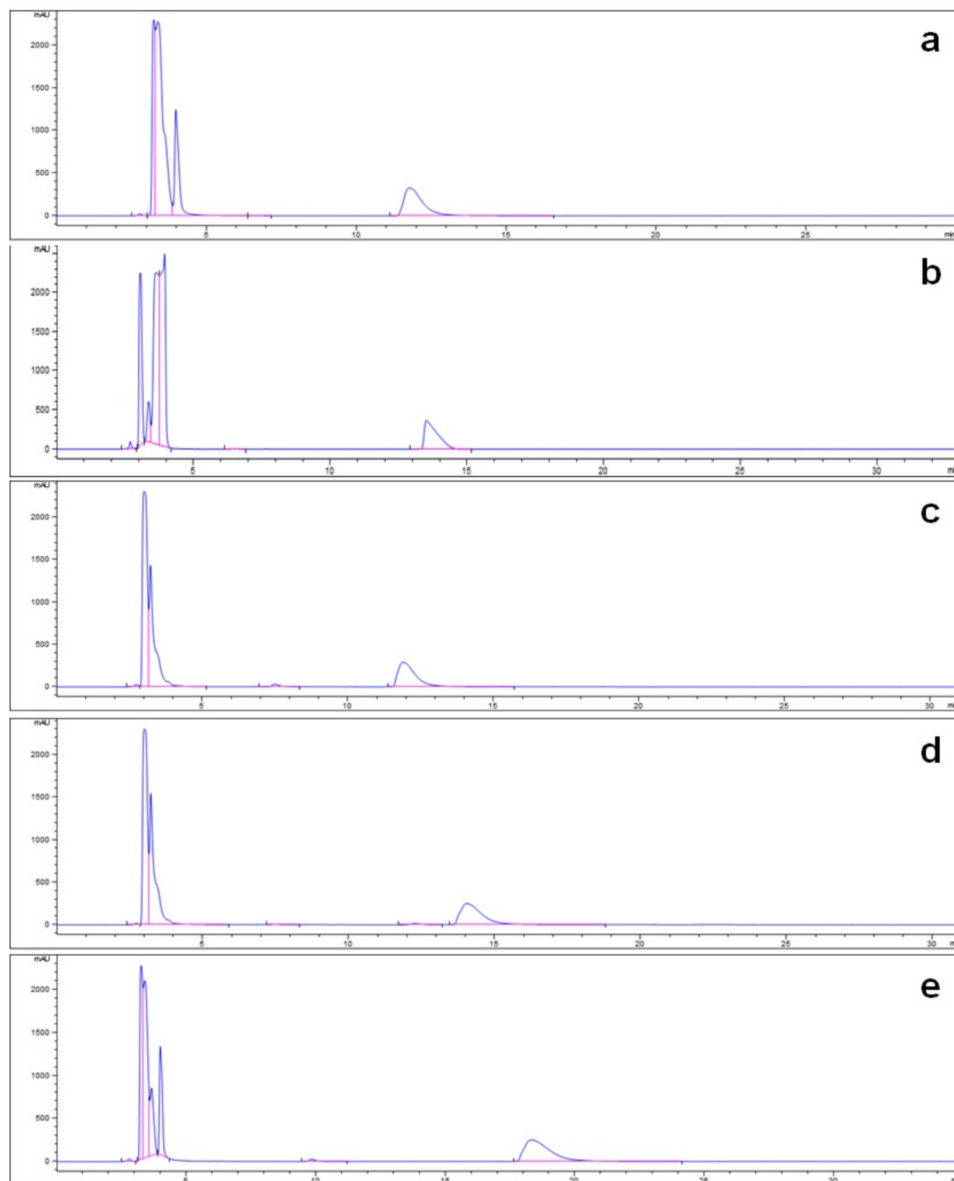



**Figure S1.** Determination of stereospecificity of PfTTA by HPLC chromatography after derivatization with OPA/NAC. The enzymatic catalytic products L-*threo*-β-hydroxy-homophenylalanine **5a** (a), L-*threo*-β-hydroxy-*p*-fluoro-homophenylalanine **5l** (b), L-*threo*-β-hydroxy-*p*-chloro-homophenylalanine **5m** (c), L-*threo*-β-hydroxy-*p*-bromo-homophenylalanine **5n** (d), and L-*threo*-β-hydroxy-*p*-nitro-homophenylalanine **5p** (e).

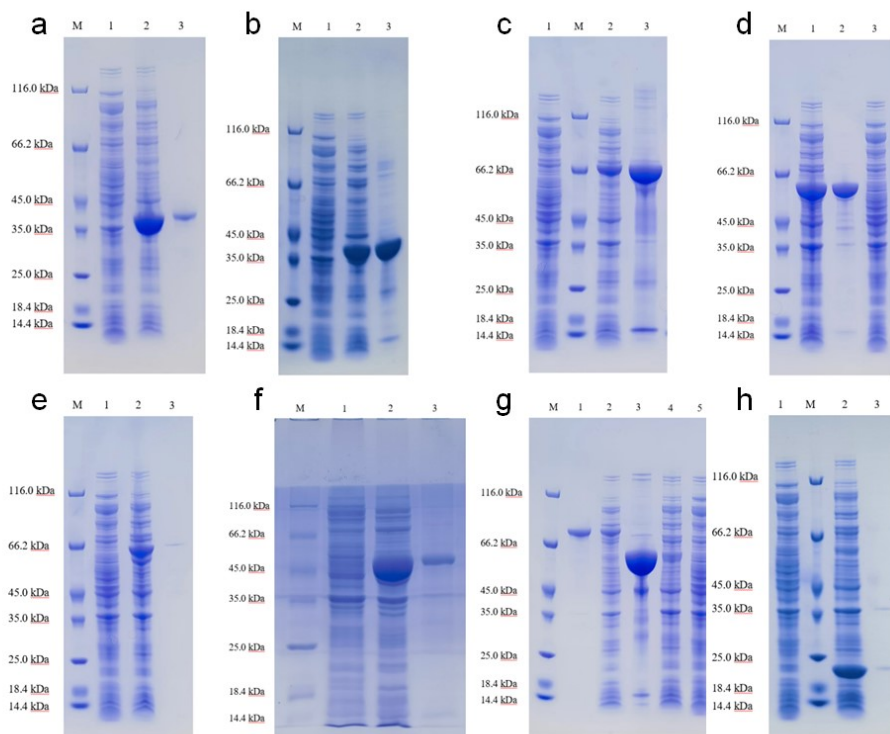



**Figure S2.** SDS-PAGE analysis of enzyme. (a) Lane 1: control check; Lane 2: supernatant of the lysate of PsTA; Lane 3, precipitate of the lysate of PsTA. (b) Lane 1: control check; Lane 2: supernatant of the lysate of PxPSDH; Lane 3, precipitate of the lysate of PxPSDH. (c) Lane 1: control check; Lane 2: supernatant of the lysate of ScPPDC; Lane 3, precipitate of the lysate of ScPPDC. (d) Lane 1: supernatant of the lysate of AbPPDC; Lane 2: precipitate of the lysate of AbPPDC; Lane 3: control check. (e) Lane 1: control check; Lane 2: supernatant of the lysate of PfPPDC; Lane 3, precipitate of the lysate of PfPPDC. (f) Lane 1: control check; Lane 2: supernatant of the lysate of PfTTA; Lane 3, precipitate of the lysate of PfTTA. (g) Lane 1, precipitate of the lysate of PtALDH-pET-32a; Lane 2: supernatant of the lysate of PtALDH-pET-32a; Lane 3, precipitate of the lysate of PtALDH-pET-28a; Lane 4: supernatant of the lysate of PtALDH-pET-28a; Lane 5, control check. (h) Lane 1: control check; Lane 2: supernatant of the lysate of MgFR; Lane 3, precipitate of the lysate of MgFR. Lane M: protein marker.

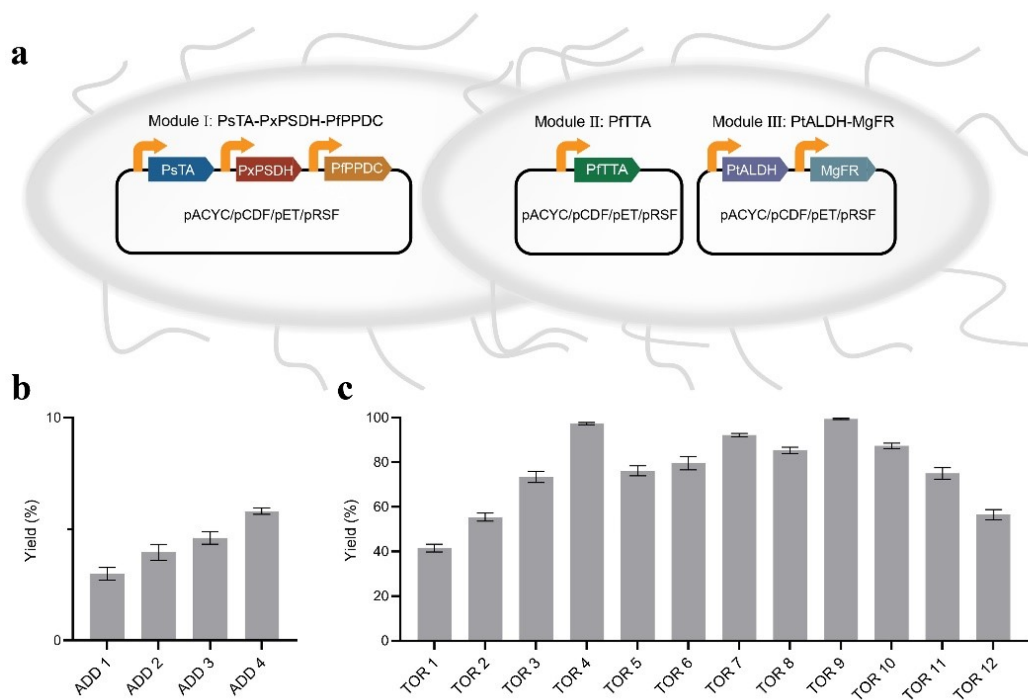
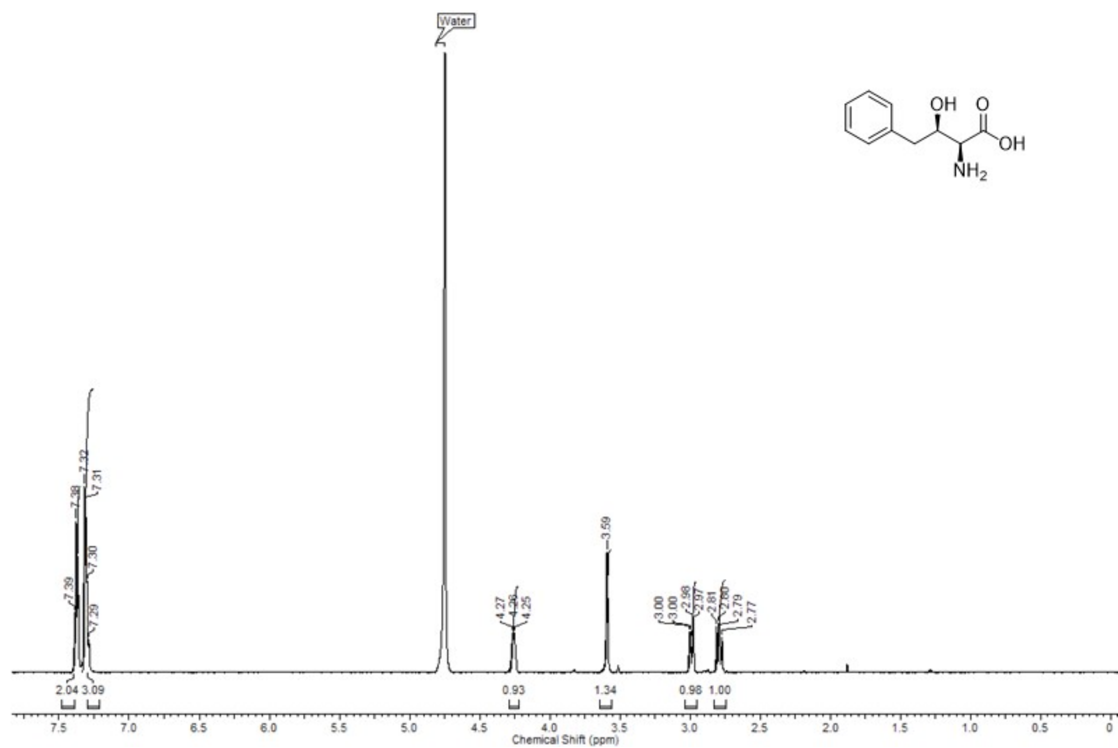



**Figure S3.** (a) Synthesis of **5a** by modular cascade biocatalysis using **4a** as an intermediate product. (b) Biotransformation of **1a** to **4a** catalyzed by *E. coli* (ADD 1-4). (c) Biotransformation of **4a** to **5a** catalyzed by *E. coli* (TOR 1-12).



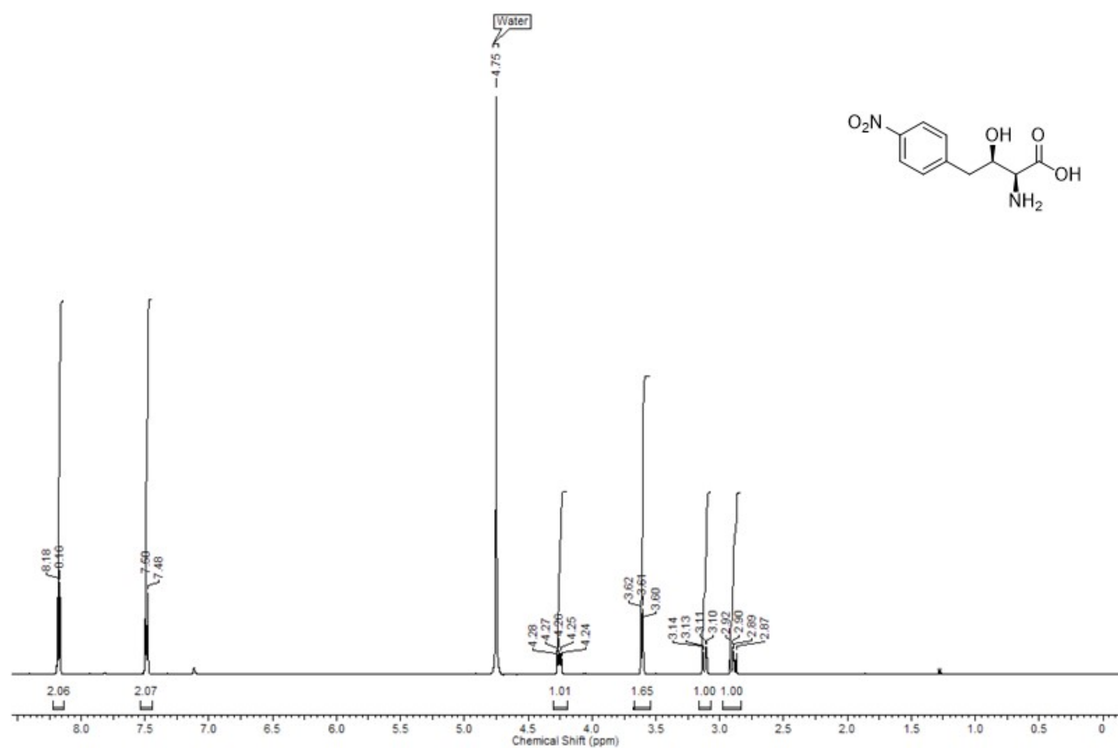




**Figure S4.** $^{1}$H spectrum of **5a** and **5p**.

## Supporting Tables

**Table S1** List of primers

| Name | Primers (5'-3') |
|---|---|
| PsTA-F | CATGCCATGGGCATGACCGACCAGTCTCAGCAGTTC |
| PsTA-R | CGGGATCCTTATCAAGCAGACATAACAGCACGGAT |
| PxPSDH-F | GGGTTTCATATGTCAACTGCCGCGCCGCAAC |
| PxPSDH-R | CCCTCGAGTTATTGAGGAAACAGTGACTCGCTATGCCGC |
| PfPPDC-F-1 | CATGCCATGGCGACCAACCTGCCGAGTACCCA |
| PfPPDC-R-1 | CGCGGATCCTGCTTTAGGCAGATGCTGCT |
| ScPPDC-F | CATGCCATGGGCATGGCACCTGTTACAATTGAAAAGTTC |
| ScPPDC-R | CGCGGATCCCTATTTTTTATTTCTTTTAAGTGCCGCTG |
| AbPPDC-F | CATGCCATGGGCATGAAACTGGCTGAAGCTCTGCTG |
| AbPPDC-R | CGCGGATCCTTCACGCGGAGCAGCGTG |
| PfTTA-F | CGCCATATGATGTCTAACGTTAAACAGCAGACCGC |
| PfTTA-R | CCGCTCGAGACGCTGAGCACCCTGGTAAACAG |
| PtALDH-F-1 | CATGCCATGGGCATGCTGCGTACCGCTACCCGTA |
| PtALDH-R-1 | CGCGGATCCCTGCGGACCGTCGTTGATAGC |
| MgFR-F | CATGCCATGGGCATGTCTGCTACCGACCTGTCTCCG |
| MgFR-R | CGCGGATCCAGCACCCAGTTTACGGAAAGCAGA |
| PtALDH-F-2 | TAAGAAGGAGATATACATATGCATATGAGCGATAAAATTATTCACCTG |
| PtALDH-R-2 | GGTTTCTTTACCAGACTCGAGCTGCGGACCGTCGTTGATA |
| PxPSDH-F-2 | TAAGAAGGAGATATACATATGTCAACTGCCGCGCCGCAA |
| PxPSDH-R-2 | CGATATCCAATTGAGATCTTTATTGAGGAAACAGTGACTCGC |
| PfPPDC-F-2 | CGCCATATGACCAACCTGCCGAGTACCCA |
| PfPPDC-R-2 | CCGCTCGAGTTATGCTTTAGGCAGATGCTGC |

## References


1. Q. Chen, X. Chen, J. Feng, Q. Wu, D. Zhu and Y. Ma, *ACS Catal.*, 2019, **9**, 4462-4469.
2. T. A. Scott, D. Heine, Z. Qin and B. Wilkinson, *Nat. Commun.*, 2017, 8.
3. J. E. Schaffer, M. R. Reck, N. K. Prasad and T. A. Wencewicz, *Nat. Chem. Biol.*, 2017, **13**, 737-744.
4. Z. Yao, C. Zhang, F. Lu, X. Bie and Z. Lu, *Appl. Microbiol. Biot.*, 2012, **93**, 1999-2009.
5. X. Chen, Y. Cui, J. Feng, Y. Wang, X. Liu, Q. Wu, D. Zhu and Y. Ma, *Adv. Synth. Catal.*, 2019, **361**, 2497-2504.
6. I. Molnar-Perl, *J. Chromatogr. A*, 2001, **913**, 283-302.
7. J. Yu, T. Wu, J. Liu, H. Zhang, and Q. Jiao. *Fine Chem.*, 2018, **35**, 65-70.