

ARTICLE

DOI: 10.1038/s41467-018-06241-x

OPEN

Asymmetric assembly of high-value α -functionalized organic acids using a biocatalytic chiral-group-resetting process

Wei Song^{1,2}, Jin-Hui Wang^{1,2}, Jing Wu³, Jia Liu^{1,2}, Xiu-Lai Chen^{1,2} & Li-Ming Liu^{1,2,4}

The preparation of α -functionalized organic acids can be greatly simplified by adopting a protocol involving the catalytic assembly of achiral building blocks. However, the enzymatic assembly of small amino acids and aldehydes to form numerous α -functionalized organic acids is highly desired and remains a significant challenge. Herein, we report an artificially designed chiral-group-resetting biocatalytic process, which uses simple achiral glycine and aldehydes to synthesize stereodefined α -functionalized organic acids. This cascade biocatalysis comprises a basic module and three different extender modules and operates in a modular assembly manner. The engineered *Escherichia coli* catalysts, which contained different module(s), provide access to α -keto acids, α -hydroxy acids, and α -amino acids with excellent conversion and enantioselectivities. Therefore, this biocatalytic process provides an attractive strategy for the conversion of low-cost achiral starting materials to high-value α -functionalized organic acids.

¹State Key Laboratory of Food Science and Technology, Jiangnan University, Wuxi 214122, China. ²Key Laboratory of Industrial Biotechnology, Ministry of Education, Jiangnan University, Wuxi 214122, China. ³School of Pharmaceutical Sciences, Jiangnan University, Wuxi 214122, China. ⁴National Engineering Laboratory for Cereal Fermentation Technology, Jiangnan University, Wuxi 214122, China. Correspondence and requests for materials should be addressed to L.-M.L. (email: mingll@jiangnan.edu.cn)

The preparation of multifunctional chiral small molecules such as unnatural α -amino acids and α -hydroxy acids has attracted extensive attention from organic chemists because they play prominent roles in chemical synthesis, cosmetics, and pharmaceutical manufacturing^{1–5}. Therefore, a few elegant chemical and enzymatic synthesis methods have been elaborated to provide a sufficient number of chemicals to meet market demand^{6–9}. Among them, one particularly attractive synthesis route is the C–H functionalization-mediated asymmetric assembly of small building blocks^{10,11}.

However, modern chemical C–H functionalization methods heavily rely on precious and unsustainable transition metals (e.g., Pd, Rh, Ru, and Ir) and protected substrates^{12–14}. Hence, some researchers have turned to biocatalytic processes using some key enzymes such as threonine aldolases (TAs)^{9,15–17}, which could catalyze the aldol assembly of unprotected small amino acids (principally glycine) and aldehydes to produce small organic acids, providing a dramatically simplified method. Another representative example is pyruvate aldolase, which uses alanine and formaldehyde as starting materials for the biosynthesis of (*S*)- and (*R*)-2-amino-4-hydroxybutanoic acid⁹. Unfortunately, these methods only provide a limited range of products such as β -hydroxy- α -amino acids and *L/D*-homoserine^{9,18,19}. Therefore, the assembly of small amino acids and aldehydes to form numerous α -keto acids and α -hydroxy acids is highly desired and remains a significant challenge. Fortunately, artificially designed multi-enzyme cascades, which offer an alternative to related chemical methods and protein engineering^{3,11,17,20}, provide a useful tool for overcoming some challenging reactions^{21–26}, such as the oxy- and amino-functionalization of alkenes^{21,24,27,28}.

Here, we report an artificially designed chiral-group-resetting cascade biocatalysis protocol, which could transform unprotected achiral building blocks into prochiral α -keto acids and then to chiral α -hydroxy acids and α -amino acids. The platform operates in a modular assembly manner, enabling the tunable and predictable operation of cascade biocatalysis and retaining a high product diversity (including aliphatic, aromatic, heteroaromatic, and heterocyclic products) and modularity.

Results

Designing enzyme cascades for chiral group resetting. The designed modular cascade catalysis platform has two parts (Fig. 1a): (i) a basic module (BM), in which achiral aldehyde and glycine are assembled to prochiral α -keto acid (including the introduction and deletion of chiral $-\text{OH}/-\text{NH}_2$ groups), and (ii) extender modules (EMs), in which α -keto acid is converted into α -hydroxy/amino acid by reintroducing chiral $-\text{OH}/-\text{NH}_2$ groups (Fig. 1b). When a BM is coupled with EMs, for chiral $-\text{OH}/-\text{NH}_2$ groups, the whole process is presented as introduction–deletion–reintroduction, which can be seen as an $-\text{OH}/-\text{NH}_2$ -group resetting process (Fig. 2a, b). Using easily available glycine as a functionalized primer and nine aldehydes as extender units (Fig. 1c), nine prochiral α -keto acids, 18 chiral α -hydroxy acids, and 18 chiral α -amino acids (Supplementary Fig. 1) were obtained.

The BM starts with the TA-mediated asymmetric assembly (C–C bond formation while introducing chiral $-\text{OH}/-\text{NH}_2$ groups) of the achiral functionalized primer glycine (**2**) with aldehyde extender units (**1**) to produce β -hydroxy- α -amino acids (**3**) (Fig. 1b). The subsequent threonine deaminase (TD)-catalyzed atypical deamination (PLP-dependent α,β -elimination reaction²⁹, Supplementary Fig. 2) led to the elimination of the chiral group ($\text{C}_\alpha\text{-NH}_2$ and $\text{C}_\beta\text{-OH}$) of **3** while producing prochiral α -keto acids (**4**).

On the basis of the BM, EMs were designed for reintroducing chiral $-\text{OH}/-\text{NH}_2$ groups (Fig. 1b). Extender module 1 (EM1)

was designed for the *S*- or *R*-selective reduction of **4** to (*S*)- or (*R*)-hydroxy acids [(*S*)- or (*R*)-**5**] catalyzed by *L*- or *D*-hydroxyisocaproate dehydrogenase (*L*- or *D*-HicDH), coupled with NADH recycling by formate dehydrogenase (FDH). Extender module 2 (EM2) employed *L*-amino acid dehydrogenase (*L*-AADH) for converting **4** into (*S*)- α -amino acids [(*S*)-**6**] and FDH for NADH recycling. Since *D*-AADHs are not ubiquitous in nature, as opposed to *L*-selective AADHs, and natural *D*-AADHs are membrane-bound enzymes dependent on NADPH, which are much more expensive than NADH³⁰, we chose *D*-amino acid transaminase (*D*-AATA) for extender module 3 (EM3) for transforming **4** into (*R*)- α -amino acids [(*R*)-**6**] and glutamate dehydrogenase (GluDH), FDH, and glutamate racemase (GLR) for amino donor recycling.

The enzymes used in all of these modules were selected on the basis of literature reports on the abilities of specific enzymes and organisms to function in the presence of the required intermediates. For the further assembly of multiple modules and the optimization of enzyme expression in one *E. coli* strain, every module was cloned into four different but compatible plasmids (Supplementary Table 1)—pACYCDuet-1, pCDFDuet-1, pET-Duet-1, and pRSFDuet-1—to generate four recombinant plasmids (Supplementary Table 2)²⁸. The protein expressions from different plasmids harbored within the cells were induced with isopropyl- β -D-1-thiogalactopyranoside (IPTG), and an SDS-PAGE analysis confirmed the soluble expression of all of the enzymes used.

Production of prochiral α -keto acid. The validation of this system was initially focused on building a capable BM (Fig. 1b); therefore, the talented enzymes TA and TD were needed. First, *Pseudomonas aeruginosa* TA (*PaTA*) was selected owing to its high activity toward **1a** and **2**^{31,32}. Then, for TD screening, eight functionally known TDs were selected from the NCBI database via data mining³³ and expressed in *E. coli* BL21 (DE3). In order to obtain a complete conversion of **3** to **4**, the TD activities toward both *threo*- and *erythro*-*L*-phenylserine (*threo/erythro*-**3a**) were tested, and only TD from *Corynebacterium glutamicum* (*CgTD*) showed activities toward both configurations. Therefore, *PaTA* and *CgTD* were overexpressed and purified (Supplementary Fig. 3). To identify the potential bottleneck in the BM, the effects of the TA and TD concentrations on the initial reaction rate of **1a** to **4a** were examined in vitro. As shown in Fig. 3a, the initial reaction rate of **1a** to **4a** was slightly increased (7%) when increasing the TA concentration by twofold, while it was increased by 1.6-fold when doubling the concentration of *CgTD*. These results indicate that *CgTD* was the rate-limiting enzyme in the C–C bond formation and α,β -elimination cascade reaction. Furthermore, the activity of *CgTD* toward different substrates was investigated, and the results are shown in Fig. 3b. It was found that the activity of *CgTD* significantly decreases as the substrate size increased. In comparison with the case using **3i** as a substrate (a natural substrate of *CgTD*), the relative activity of *CgTD* toward bulky substrate **1a** decreased to 5% and exhibited no activity toward larger substrate **3e** (Fig. 3b). Therefore, to meet the needs of α -keto acid synthesis, it is imperative to improve the substrate specificity of *CgTD* toward bulky substrates by protein engineering strategies.

Then, we built a structural model of *CgTD* via homologous modeling by using *E. coli* TD (PDB ID: 1TDJ) as the template. To simulate the orientation of the external aldimine intermediate³⁴, ligands (**3a** and **3i**) were created with the nitrogen atom of **3** linked to the PLP cofactor (Fig. 3c, d). Docking these ligands into the cofactor-free structure yielded outputs with the PLP and substrates with approximately the same orientation, suggesting

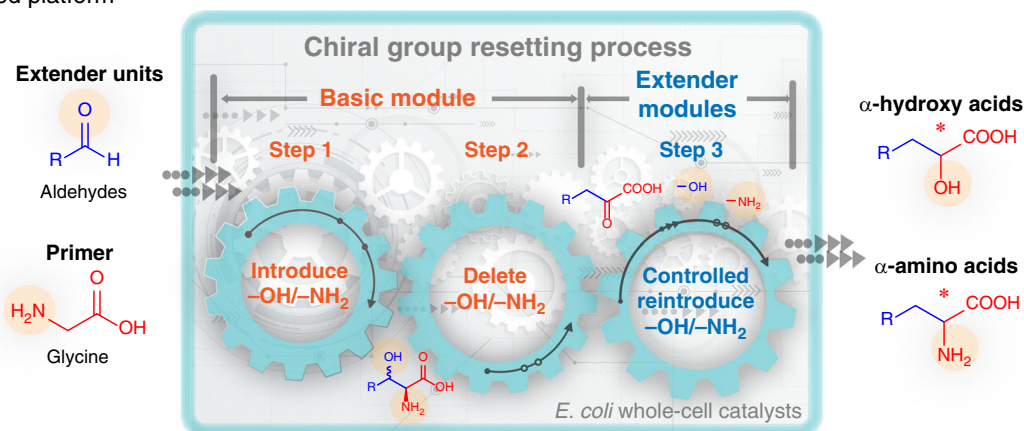
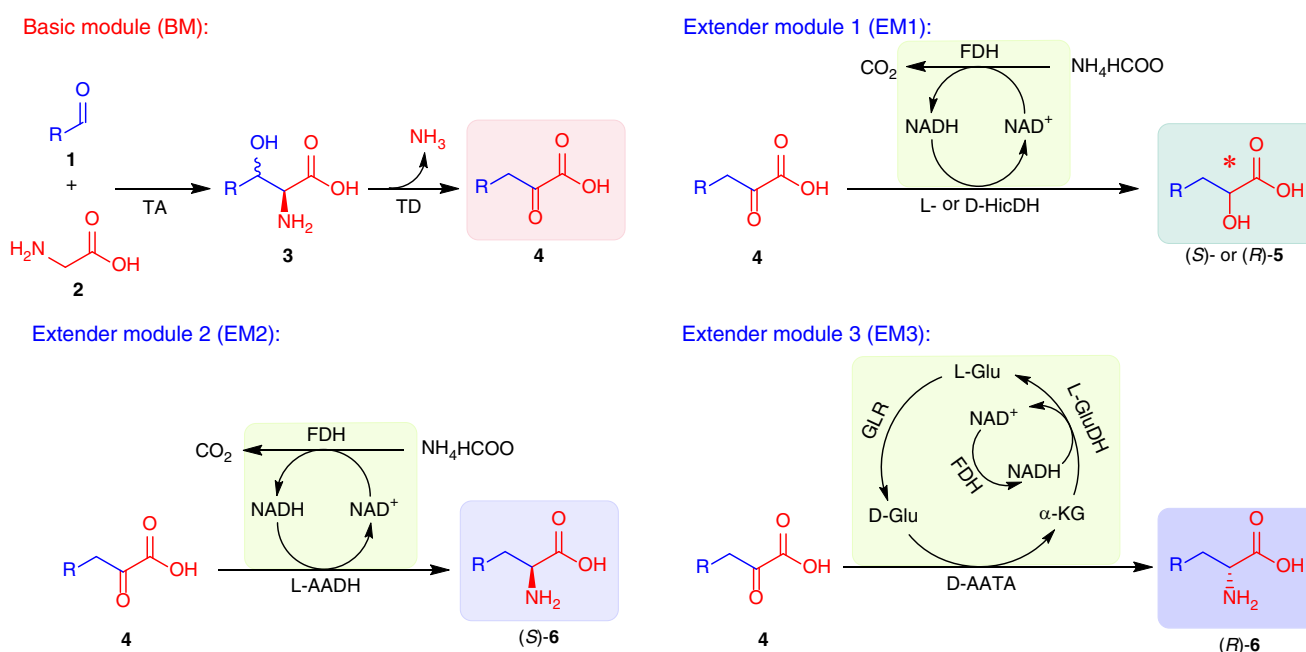
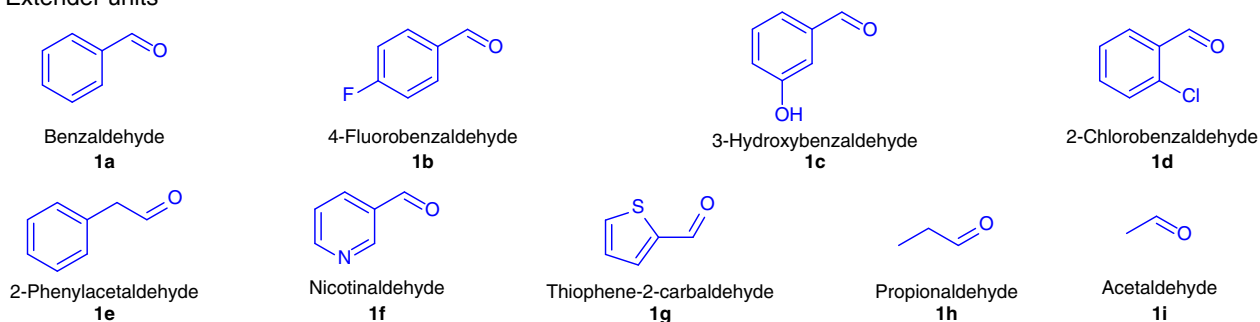
a Engineered platform**b** Modules demonstrated in this study**c** Extender units

Fig. 1 The synthesis of α -functionalized organic acids by proposed modular cascade biocatalysis platform. **a** The engineered modular cascade biocatalysis platform. **b** Four general basic enzyme modules. Basic module: threonine aldolase (TA) and threonine deaminase (TD) for condensation-deamination of aldehydes and glycine to α -keto acids; extender module 1: hydroxyisocaproate dehydrogenase (HicDH) and formate dehydrogenase (FDH) for reduction of α -keto acids to α -hydroxy acid; extender module 2: L-amino acid dehydrogenase (L-AAD) and FDH for reduction of α -keto acids to L- α -amino acids; extender module 3: D-amino acid transaminase (D-AATA), FDH, glutamate racemase (GLR), and glutamate dehydrogenase (GluDH) for transamination of α -keto acids to D- α -amino acids. Functionalized groups (α -position) and extender units (from aldehydes) are colored in red and blue, respectively. **c** Nine different extender units, including aromatic (**1a–e**), heteroaromatic (**1f**), heterocyclic (**1g**), and aliphatic (**1h–i**) aldehydes

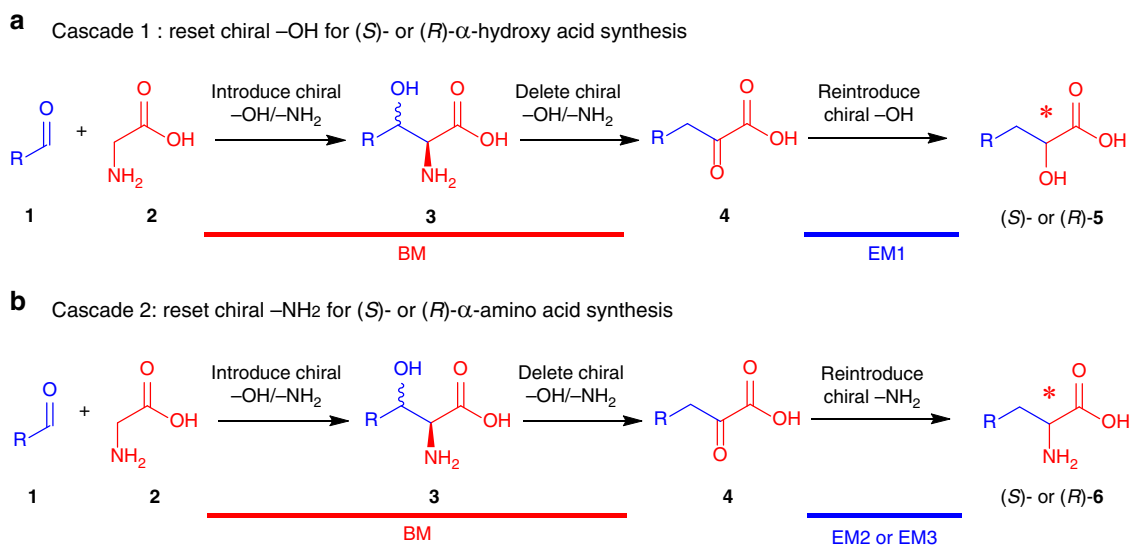


Fig. 2 The artificially designed chiral-group-resetting cascades. **a** Cascade 1: reset chiral –OH for stereodefined α -hydroxy acid synthesis via cascade BM with EM1. **b** Cascade 2: reset chiral –NH₂ for stereodefined α -amino acid synthesis via cascade BM with EM2 or EM3

that bulky substrates show the correct orientation. The surface of the binding pocket (Fig. 3e) also indicated that the volume of the catalytic cavity is sufficiently large to accept bulky substrates. On the other hand, by analyzing the substrate channel of CgTD, steric hindrance was identified as a potential bottleneck (a small entrance, Fig. 3f) for the bulky substrates. Two residues (Phe114 and Arg229) were identified as potential hot spots for broadening the substrate channel. Semisaturated mutation was used for protein engineering, and the highest initial reaction rate ($667 \mu\text{M min}^{-1}$) was observed for **1a** with the variant CgTD^{F114A,R229T} (with a larger substrate channel, Fig. 3g), thus representing a greater-than-18-fold increase when compared with that of CgTD_{WT} for **1a** ($37 \mu\text{M min}^{-1}$).

Four BM plasmids containing the PaTA and CgTD^{F114A,R229T} genes were transformed into the host strain that led to *E. coli* (OA01–04) (Supplementary Table 3). Among them, *E. coli* OA02 was found to be the best whole-cell catalyst in the biotransformation of **1a** into **4a** (Supplementary Fig. 4), which produced 380 mg l^{-1} **4a**. Furthermore, the reaction conditions for the conversion of **1a** to **4a** were optimized, including the pH, temperature, *E. coli* (OA02) concentration, and cosolvents (Supplementary Fig. 5). It was found that under the optimum conditions (pH 8.0, 25 °C, 10 g dcm l⁻¹, and 10% DMSO), the titer of **4a** was improved to 1494 mg l^{-1} with a conversion of 91%. However, low levels of unreacted substrate **1a** (2%) and intermediate **3a** (4%) still remained in the conversion system.

The synthetic potential of the BM was then tested with a range of aldehydes (**1b–i**), affording the corresponding α -keto acids (**4b–i**) at good conversions (68–91%) (Fig. 4). Overall, the conversion is higher than that in the case of PaTA alone (conversion < 80%, d.e. < 55%)³¹, which indicated that the BM could overcome the thermodynamically unfavorable equilibria of the TA reaction by coupling it with the irreversible TD reaction. However, even with the BM cascade system, it still cannot lead to complete conversion, which is possibly due to the reversible nature of the TA reaction and the relatively low activity of CgTD for the second reaction. For cascade processes, it is important that the second step is much faster than the first step to reach the theoretical conversion of 100%²⁸. Another reason is the metabolic background of *E. coli*, which could convert aldehydes into carboxylic acids or the corresponding alcohols and cause a loss of the starting material³⁵.

Production of α -hydroxy acid by reintroducing chiral –OH. To efficiently reintroduce chiral –OH for converting keto acids into α -hydroxy acids, the highly selective L- or D-HicDH, (L-HicDH obtained from *Lactobacillus confusus* DSM 201966 and D-HicDH obtained from *Lactobacillus paracasei* DSM 20008)^{36,37} was chosen and used to construct EM1. Then, four BM plasmids and eight EM1 plasmids were combinatorially combined and transformed into *E. coli* to engineer the catalysts for the conversion of glycine and aldehydes into α -hydroxy acids. Since plasmids with the same backbone are not compatible with each other, 24 different *E. coli* strains (OA05–28) were obtained (Supplementary Table 3), each coexpressing PaTA, CgTD, CbFDH, and highly selective L- or D-LDH (*E. coli* OA05–16 and *E. coli* OA17–28 provide S- and R-selective reducing activities, respectively). Next, the capacities of those 24 types of *E. coli* strains (OA05–28) on the transformation of **1a** into (S)- and (R)-**5a** were examined, and the results showed that *E. coli* (OA15) and *E. coli* (OA23) are the best whole-cell catalysts, which produced 1470 mg l^{-1} of (S)-**5a** (>99% e.e. and 89% conversion; Supplementary Fig. 6) and 1404 mg l^{-1} of (R)-**5a** (>99% e.e. and 85% conversion, Supplementary Fig. 7), respectively, in 36 h. The intermediate **4a** was found at relatively low levels (<1%); this indicated that EM1 has a high conversion efficiency.

To further explore the potential of the engineered cascade, aldehydes **1b–i** (10 mM) were biotransformed into α -hydroxy acids by using *E. coli* (OA15) and *E. coli* (OA23). As summarized in Table 1, (S)- and (R)-**5** were obtained with 55–92% conversion with excellent enantioselectivities (96–99%), demonstrating that the enzymatic cascade comprising BM and EM1 has a high efficiency and broad applicability.

Production of α -amino acid by reintroducing chiral –NH₂. In order to reintroduce the chiral –NH₂ group, L-AADH from *Bacillus badius*³⁸ was selected to construct EM2 for converting α -keto acids into (S)- α -amino acids. The combinatorial assembly of four BM plasmids and four EM2 plasmids led to 12 different *E. coli* whole-cell catalysts (OA29–40) (Supplementary Table 3), each coexpressing PaTA, CgTD, CbFDH, and highly selective L-AADH. Consequently, the 12 constructed *E. coli* strains were used for the transformation of **1a** into (S)-**6a**, and *E. coli* (OA34) showed the

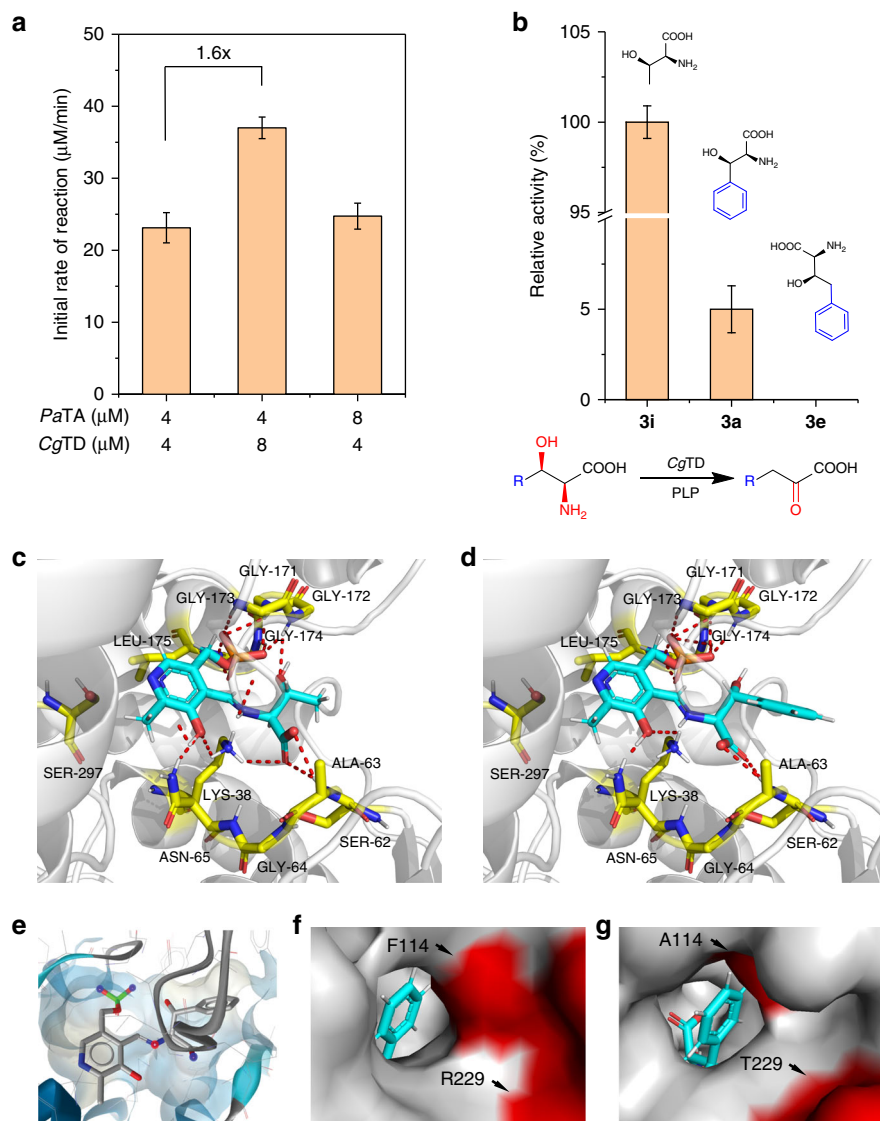


Fig. 3 Improvement of the substrate specificity of CgTD toward bulky substrates by protein engineering. **a** The effect of TA and TD concentrations on the initial reaction rate of **1a** to **4a** (reactions were performed in duplicate with PaTA and CgTD enzymes, with 2.5 mM aldehydes, 25 mM glycine, and 50 μM PLP). **b** CgTD activity toward different substrates. **c** Docking of the PLP/**3i** complex into the active site of CgTD. **d** Docking of the PLP/**3a** complex into the active site of CgTD. **e** The surface of binding pocket. **f** Substrate access channel of CgTD_{WT}. **g** Substrate access channel of CgTD^{F114A,R229T}. All enzymatic assays were performed in triplicate, and error bars indicate ±s.d.

best conversion capacities, which produced 1478 mg l⁻¹ (*S*)-**6a** (90% conversion and 95% e.e., Supplementary Fig. 16). Furthermore, *E. coli* (OA34) was used to catalyze different aldehydes **1b–i** (Table 2) to the corresponding (*S*)-amino acids, and excellent conversions of 65–93% (>89% e.e.) were detected.

Similarly, in order to produce (*R*)-α-amino acids, 12 different *E. coli* strains (OA41–52) were obtained, each coexpressing PaTA, CgTD, glucose dehydrogenase from *Bacillus megaterium* (BmGDH, for NADPH recycling), and D-AADH (a *meso*-2,6-D-diaminopimerate dehydrogenase mutant from *Symbiobacterium thermophilum*³⁹, StDAPDH^{H227V}). However, the best strain of *E. coli* (OA51) produced only 370 mg l⁻¹ (*R*)-**6a** (Supplementary Fig. 17). In order to improve the conversion of (*R*)-α-amino acids, we recruited the transamination route in EM3 (expressing *B. subtilis* D-AATA) (Fig. 1b) since D-AADHs are not ubiquitous in nature as opposed to L-selective AADHs (only *meso*-DAPDHs have been reported)³⁰. Thus, four BM plasmids and four EM3 plasmids were combinatorially assembled to obtain 12 *E. coli*

strains (OA53–64) (Supplementary Table 3). The 12 obtained types of *E. coli* strains were used for the conversion of **1a** into (*R*)-**6a**. Finally, the highest titer of (*R*)-**6a** (1396 mg l⁻¹ with 85% conversion and 99% e.e., Supplementary Fig. 18) was obtained with *E. coli* (OA59). The potential of *E. coli* (OA59) was examined with aldehydes **1b–i** (Table 2), and (*R*)-**6** was produced with conversions of 71–89%, except (*R*)-**6e** (32%) and (*R*)-**6g** (64%). Remarkably, all of the produced (*R*)-α-amino acids **6a–i** were in their enantiopure forms (e.e. >98%). In conclusion, the enzymatic cascade coupling BM with EM2/EM3 exhibited great potential for the synthesis of (*S*)/(*R*)-α-amino acids from renewable glycine and aldehydes.

Synthesis of α-functionalized organic acids on a 100-ml scale.

The synthesis of α-functionalized organic acids was scaled up to 100 ml to further evaluate the feasibility of the biocatalytic cascades. First, the tolerance of the key enzyme CgTD to **1a** and the

ratio of the catalyst to substrate **1a** were optimized for the scaled-up biotransformations, and the optimum conditions were identified to be 50 mM **1a** with catalyst-to-substrate ratios of 0.5–1:1 (g dcw l⁻¹ mM⁻¹) at 25 °C, pH 8.0, and 5% DMSO (Supplementary Fig. 27). Then, the preparative biotransformations of **1a**

(50 mM) with *E. coli* (OA02), *E. coli* (OA23), and *E. coli* (OA59) (30 g dcw l⁻¹) were performed. An additional 40-mM **1a** was added after 6 h when the remaining concentration of **1a** was lower than 10 mM. Subsequently, a total of 12.5 g l⁻¹ **4a**, 12.1 g l⁻¹ (*R*)-**5a** (99% e.e.), and 11.9 g l⁻¹ (*R*)-**6a** (99% e.e.) was produced at 85, 81, and 80% conversion, respectively, in 30–42 h. Further, the space-time yields (STYs) of **4a**, (*R*)-**5a**, and (*R*)-**6a** reached 10.0, 6.9, and 6.8 g l⁻¹ d⁻¹, respectively. Finally, **4a**, (*R*)-**5a**, and (*R*)-**6a** were isolated at 75, 69, and 61% yield by chromatography, extraction, and crystallization. Similarly, seven other valuable chemicals (**4i**, (*S*)-**5e**, (*S*)-**5i**, (*R*)-**5i**, (*S*)-**6e**, (*S*)-**6g**, and (*R*)-**6h**) with many applications (Supplementary Table 4) were successfully synthesized using the corresponding *E. coli* whole-cell catalysts and were isolated in yields of 45–78%.

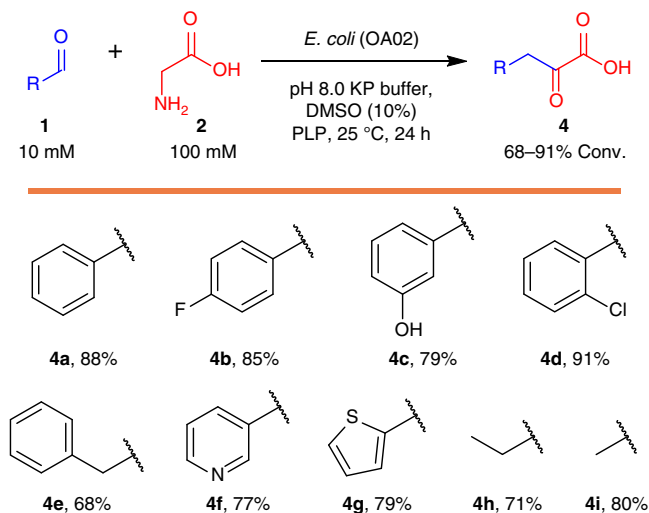


Fig. 4 Analytical biotransformation of glycine and aldehydes **1a–i** to α -keto acids **4a–i** with *E. coli* OA02. Reactions were performed in duplicate with resting cells *E. coli* (OA02) (10 g dcw l⁻¹) and **1a–i** (10 mM) in KP buffer (50 mM, pH 8.0, 100 μ M PLP, and 10% DMSO) at 200 rpm and 25 °C for 24 h. The values are averages of two experiments

Discussion

In this study, we report an artificially designed biocatalytic cascade, which comprises a BM and three different EMs, for the C–H functionalization-mediated asymmetric synthesis of α -functionalized organic acids. This biocatalytic cascade represents a chiral-group-resetting process that operates in a modular assembly manner, enabling the change of different catalyst modules and substrate units to facilitate the control of product species and configurations. Using this biocatalytic process with glycine as a primer and different aldehydes as extender units, 45 types of aliphatic, aromatic, heteroaromatic and heterocyclic products were obtained, including nine prochiral α -keto acids, 18 chiral α -hydroxy acids, and 18 chiral α -amino acids. Therefore, this biocatalytic process exhibited great potential for the production of a variety of α -functionalized organic acids.

Table 1 Analytical biotransformation of glycine and aldehydes to α -hydroxy acids by cascade catalysis

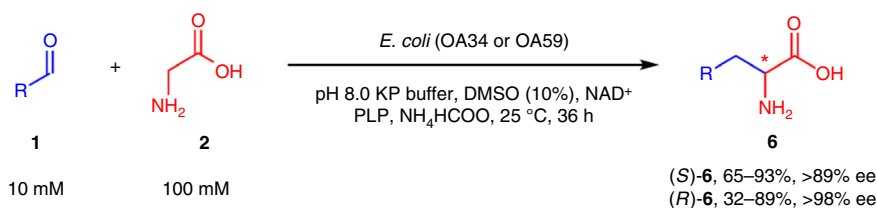
Reaction scheme showing the biotransformation of glycine (**2**) and aldehydes (**1**) to α -hydroxy acids (**5**) using *E. coli* (OA15 or OA23). The reaction conditions are pH 8.0 KP buffer, DMSO (10%), NAD⁺, 25 °C, 36 h, PLP, NH₄HCOO. The products are (*S*)-**5** (71–91%, >98% ee) and (*R*)-**5** (55–92%, >96% ee).

Substrate ^a	Product	Conv (%)	Yield ^b (mg l ⁻¹)	e.e. ^c (%)
1a	(<i>S</i>)- 5a	89	1470	99
	(<i>R</i>)- 5a	85	1404	99
1b	(<i>S</i>)- 5b	87	1593	99
	(<i>R</i>)- 5b	75	1374	99
1c	(<i>S</i>)- 5c	78	1413	99
	(<i>R</i>)- 5c	81	1467	98
1d	(<i>S</i>)- 5d	91	1816	99
	(<i>R</i>)- 5d	92	1836	99
1e	(<i>S</i>)- 5e	71	1272	98
	(<i>R</i>)- 5e	55	985	96
1f	(<i>S</i>)- 5f	78	1296	99
	(<i>R</i>)- 5f	70	1163	99
1g	(<i>S</i>)- 5g	71	1215	99
	(<i>R</i>)- 5g	78	1335	99
1h	(<i>S</i>)- 5h	87	1019	99
	(<i>R</i>)- 5h	83	972	99
1i	(<i>S</i>)- 5i	90	928	99
	(<i>R</i>)- 5i	80	825	99

^a Reactions were performed in duplicate with resting cells of *E. coli* (OA15 or OA23) (10 g dcw l⁻¹) and **1a–i** (10 mM) in 2 ml KP buffer (50 mM, pH 8.0, 100 μ M PLP, 1 mM NAD⁺, and 10% DMSO) at 200 rpm and 25 °C for 36 h

^b The conversion and yields were obtained after completion of the reactions and determined by HPLC analysis. The values are averages of two experiments

^c Enantiomeric excess (e.e.) was determined by chiral HPLC analysis (Supplementary Figs 8–15). The values are averages of two experiments

Table 2 Analytical biotransformation of glycine and aldehydes to α -amino acids by cascade biocatalysis

Substrate ^a	Product	Conv. (%)	Yield ^b (mg l ⁻¹)	e.e. ^c (%)
1a	(S)- 6a	90	1478	95
	(R)- 6a	85	1396	99
1b	(S)- 6b	80	1457	93
	(R)- 6b	84	1530	99
1c	(S)- 6c	75	1351	98
	(R)- 6c	71	1279	99
1d	(S)- 6d	93	1847	98
	(R)- 6d	89	1768	99
1e	(S)- 6e	65	1158	92
	(R)- 6e	32	570	98
1f	(S)- 6f	79	1189	96
	(R)- 6f	72	1028	99
1g	(S)- 6g	73	1243	89
	(R)- 6g	64	1089	99
1h	(S)- 6h	87	1010	97
	(R)- 6h	85	987	99
1i	(S)- 6i	91	929	98
	(R)- 6i	78	796	99

^a Reactions were performed in duplicate with resting cells of *E. coli* (OA34 or OA59) (10 g dcw l⁻¹) and **1a-i** (10 mM) in 2 ml KP buffer (50 mM, pH 8.0, 100 μ M PLP, 1 mM NAD⁺, and 10% DMSO) at 200 rpm and 25 °C for 36 h

^b The conversion and yields were obtained after completion of the reactions and determined by HPLC analysis. The values are averages of two experiments

^c Enantiomeric excess (e.e.) was determined by chiral HPLC analysis (Supplementary Figs 19–26). The values are averages of two experiments

The chiral group resetting process reported here provides a useful tool for the regulation of product configurations and species and helps to expand the application of unsatisfactory enzymes (such as TAs). During this process, the chiral information of each product was changed via three steps (Fig. 2): (i) introduce chiral –OH/–NH₂ groups (low selectivity), (ii) delete chiral –OH/–NH₂ groups, and (iii) reintroduce –OH/–NH₂ groups (with the desired configurations and species). With this process, both optically pure α -amino acids and α -hydroxy acids were finally obtained. Furthermore, this artificially designed biocatalytic process also overcame the thermodynamic (reversible reaction) and kinetic limitations (low diastereoselectivity for β -carbon) associated with the use of TAs^{31,40,41} by cascading with an irreversible TD reaction. From an organic synthesis perspective, for the regulation of product configurations and species, the chiral-group-resetting process is unique and advantageous over the majority of other enzymatic processes, e.g., asymmetric reduction (from achiral to *S/R*-NH₂/–OH)^{3,39}, stereoinversion (from *S*-NH₂/–OH to *R*-NH₂/–OH)^{3–5}, deracemization (from *rac*-NH₂/–OH to *R*-NH₂/–OH)^{3–5}, and type conversion (from *S*-NH₂ to *S/R*-OH)^{26,36,37}. In related chemical methods, the product configurations and species are the same as the substrate since the active chiral functional groups (e.g., –OH and –NH₂)^{10,11} were protected.

The achiral glycine and aldehydes, which are readily available and inexpensive, for the production of α -hydroxy acids and α -amino acids represent a sustainable alternative to other complementary methods. In contrast, in most enzymatic production processes of α -hydroxy acids and α -amino acids, the commonly

used substrates are keto acids^{3,39}, (*S*)-amino acids or their racemic mixtures^{3,26,36}, and (*S*)-hydroxy acids or their racemic mixtures^{4,5}; these are not readily available and much more expensive than the corresponding aldehydes. Regarding chemical methods, the protection and deprotection processes have the particular drawback of cost due to the protected substrates that are needed, such as *N*-phthalimido-protected *L*- α -alanine¹¹ and methyl or benzyl-protected (*S*)- α -lactic acid¹⁰. Thus, this biocatalytic process provides an attractive strategy for the conversion of low-cost achiral starting materials to high-value α -functionalized organic acids.

The biocatalytic process in this study is elegant and nicely exploits the power of cascades: (i) it provides a universal, efficient, and renewable catalyst to replace precious transition metals such as Pd^{10,11} in the modern chemical C–H functionalization-mediated asymmetric assembly of α -functionalized organic acids and (ii) when compared to other enzymatic methods^{3–5,9,26,28,36,39}, the universality and practicality of this biocatalytic cascade is more attractive owing to the more diverse product range that it supports, including liphatic, aromatic, heteroaromatic, and heterocyclic products from three classes (α -keto/hydroxy/amino acids). Furthermore, since enantiocomplementary *L*- and *D*-specific TAs were identified to accept nucleophilic substrates other than glycine (such as alanine, serine, and cysteine)^{42,43}, the expansion of this biocatalytic process could enable the synthesis of a vast array of α -functionalized products simply by combinatorically integrating different components such as other catalyst modules and substrate units.

Methods

Strains and plasmids. Wild-type *E. coli* BL21 DE3 was used for all molecular cloning and biocatalysis experiments. Plasmid based gene overexpression was achieved by cloning the desired gene(s) into a set of compatible plasmids pACYCDuet-1, pCDFDuet-1, pETDuet-1, and pRSFDuet-1 (Novagen, Darmstadt, Germany) (Supplementary Table 1) digested with appropriate restriction enzymes (see Supplementary Methods for genetic constructions). Further transformation of constructed plasmid(s) into *E. coli* strain gave an *E. coli* strain containing one or two designed modules for the desired transformations. All resulting plasmids and strains used in this study are listed in Supplementary Tables 2 and 3.

Docking simulations. The structure model of CgTD and its variants were built via SWISS-MODEL [<https://swissmodel.expasy.org/>] by using *E. coli* TD (PDB ID: 1TDJ [<https://www.rcsb.org/structure/1TDJ>]) as template. Docking simulations were performed using Autodock Vina and CgTD models. BioLuminate (Schrödinger) was used to draw ligands, and AutoDockTools was used to generate PDBQT files. To simulate the orientation of the external aldimine intermediate, ligands were created having the nitrogen atom linked to the PLP cofactor.

Enzyme mutagenesis. Site-directed mutagenesis was performed by PCR using mutagenic primers and plasmid pETduet-CgTD as template according to the manufacturer's instructions of QuickChange (Stratagene) (see Supplementary Methods for directed evolution experiments). The primers used in this study are listed in Supplementary Table 5. The *DpnI*-digested PCR product of 3 μ l was used to transform 80 μ l of *E. coli* JM109 chemically competent cells and colonies after transformation were incubated for DNA sequencing until all the designed mutants were obtained. Subsequently the plasmid of each mutant was extracted and transformed into *E. coli* BL21(DE3) for enzyme expression.

Biotransformation procedures. Analytic transformation of **1a-i** (10 mM) to **4a-i** were conducted with resting cells of *E. coli* (OA02) (10 g dcw l⁻¹) in 2 ml KP buffer (50 mM, pH 8.0, 100 μ M PLP, and 10% DMSO) at 200 rpm and 25 °C for 24 h. Analytic transformation of **1a-i** (10 mM) to (S)- and (R)-**5a-i** were conducted with resting cells of *E. coli* (OA15) and *E. coli* (OA23) (10 g dcw l⁻¹) in 2 ml KP buffer (50 mM, pH 8.0, 100 μ M PLP, and 10% DMSO) at 200 rpm and 25 °C for 36 h, respectively. Analytic transformation of **1a-i** (10 mM) to (S)- and (R)-**6a-i** were conducted with and resting cells of *E. coli* (OA33) and *E. coli* (OA71) (10 g dcw l⁻¹) in 2 ml KP buffer (50 mM, pH 8.0, 100 μ M PLP, and 10% DMSO) at 200 rpm and 25 °C for 36 h, respectively. At the end of the reaction, 100 μ l of supernatant was separated after centrifugation (13000 g, 2 min), diluted with MeCN/H₂O solution (900 μ l, 5:4) containing 0.1% TFA. The protein was removed by centrifugation and the solution was filtered through 0.22 μ m membrane filters. The resulting solution was analyzed by HPLC for quantifying the products under the conditions stated below. To determine the e.e. of **5a-i**, the sample (200 μ l) at the end of reaction was separated after centrifugation (13,000 g, 2 min), acidified with HCl (2.0 M) and extracted with EtOAc. After evaporation of ethyl acetate, the residue was dissolved in solvent (*n*-hexane: IPA = 95:5) for chiral HPLC analysis (see Supplementary Methods for analysis methods).

Preparative transformation of **1a** or **1i** (50 mM initial concentration and 40 mM addition at 6 h; total 90 mM) to **4a** or **4i** was conducted with resting cells of *E. coli* (OA02) (30 g dcw l⁻¹) in 100 ml KP buffer (50 mM, pH 8.0, 100 μ M PLP, and 5% DMSO) at 200 rpm and 25 °C for 30 h. Preparative transformation of **1a**, **1e**, or **1i** (50 mM initial concentration and 40 mM addition at 6 h; total 90 mM) to (S)-**5e**, (S)-**5i**, or (R)-**5a** was conducted with resting cells of *E. coli* (OA15) or *E. coli* (OA23) (30 g dcw l⁻¹) in 100 ml KP buffer (50 mM, pH 8.0, 100 μ M PLP, and 5% DMSO) at 200 rpm and 25 °C for 42 h, respectively. Preparative transformation of **1a**, **1e**, **1g**, or **1h** (50 mM initial concentration and 40 mM addition at 6 h; total 90 mM) to (S)-**6e**, (S)-**6g**, (R)-**6a**, or (R)-**6h** was conducted with and resting cells of *E. coli* (OA33) or *E. coli* (OA71) (30 g dcw l⁻¹) in 100 ml KP buffer (50 mM, pH 8.0, 100 μ M PLP, and 5% DMSO) at 200 rpm and 25 °C for 42 h, respectively. At the end of the reaction, the reaction mixtures were centrifuged at 6000 \times g for 30 min, and the supernatants was collected for product isolation (see Supplementary Methods for isolation protocols). The isolated products were further identified by nuclear magnetic resonance (NMR) and high resolution mass spectrometry (HRMS) analysis (see Supplementary Figs 28–47 and Supplementary Notes).

Data availability

The data supporting the findings of this study are available within the paper and its Supplementary Information files, or are available from the authors upon request.

Received: 28 April 2018 Accepted: 24 August 2018

Published online: 19 September 2018

References

- Gati, W. & Yamamoto, H. Second generation of aldol reaction. *Acc. Chem. Res.* **49**, 1757–1768 (2016).
- Szekrenyi, A. et al. Asymmetric assembly of aldose carbohydrates from formaldehyde and glycolaldehyde by tandem biocatalytic aldol reactions. *Nat. Chem.* **7**, 724–729 (2015).
- Parmeggiani, F. et al. Single-biocatalyst synthesis of enantiopure D-arylalanines exploiting an engineered D-amino acid dehydrogenase. *Adv. Synth. Catal.* **358**, 3298–3306 (2016).
- Adam, W. et al. Enzymatic resolution of chiral 2-hydroxy carboxylic acids by enantioselective oxidation with molecular oxygen catalyzed by the glycolate oxidase from spinach (*Spinacia oleracea*). *J. Org. Chem.* **62**, 7841–7843 (1997).
- Adam, W., Lazarus, M., Saha-Möller, C. R. & Schreier, P. Quantitative transformation of racemic 2-hydroxy acids into (R)-2-hydroxy acids by enantioselective oxidation with glycolate oxidase and subsequent reduction of 2-keto acids with D-lactate dehydrogenase. *Tetrahedron.: Asymmetry* **9**, 351–355 (1998).
- Furukawa, K., Inada, H., Shibuya, M. & Yamamoto, Y. Chemoselective conversion from α -hydroxy acids to α -keto acids enabled by nitroxyl-radical-catalyzed aerobic oxidation. *Org. Lett.* **18**, 4230–4233 (2016).
- Furukawa, K., Shibuya, M. & Yamamoto, Y. Chemoselective catalytic oxidation of 1,2-diols to α -hydroxy acids controlled by TEMPO-CIO₂ charge-transfer complex. *Org. Lett.* **17**, 2282–2285 (2015).
- Jiang, M., Jin, Y., Yang, H. & Fu, H. Visible-light photoredox synthesis of unnatural chiral α -amino acids. *Sci. Rep.* **6**, 26161 (2016).
- Hernandez, K. et al. Combining aldolases and transaminases for the synthesis of 2-amino-4-hydroxybutanoic acid. *ACS Catal.* **7**, 1707–1711 (2017).
- Chen, K., Li, X., Zhang, S. & Shi, B. Synthesis of chiral α -hydroxy acids via palladium-catalyzed C(sp³)-H alkylation of lactic acid. *Chem. Commun.* **52**, 1915–1918 (2016).
- He, J. et al. Ligand-controlled C(sp³)-H arylation and olefination in synthesis of unnatural chiral α -amino acids. *Science* **343**, 1216–1220 (2014).
- Baker, P. & Seah, S. Y. K. Rational approaches for engineering novel functionalities in carbon-carbon bond forming enzymes. *Comput. Struct. Biotech. J.* **2**, 1–10 (2012).
- Clapés, P. & Garrabou, X. Current trends in asymmetric synthesis with aldolases. *Adv. Synth. Catal.* **353**, 2263–2283 (2011).
- Hartwig, J. Evolution of C-H bond functionalization from methane to methodology. *J. Am. Chem. Soc.* **138**, 2–24 (2016).
- Steinreiber, J. et al. Threonine aldolases—an emerging tool for organic synthesis. *Tetrahedron* **63**, 918–926 (2007).
- Steinreiber, J. et al. Overcoming thermodynamic and kinetic limitations of aldolase-catalyzed reactions by applying multienzymatic dynamic kinetic asymmetric transformations. *Angew. Chem. Int. Ed.* **46**, 1624–1626 (2007).
- Hernandez, K. et al. Engineered L-serine hydroxymethyltransferase from *Streptococcus thermophilus* for the synthesis of α,α -dialkyl- α -amino acids. *Angew. Chem. Int. Ed.* **127**, 3056–3060 (2015).
- Seiple, I. B., Mercer, J. A. M., Sussman, R. J., Zhang, Z. & Myers, A. G. Stereoccontrolled synthesis of syn- β -Hydroxy- α -amino acids by direct aldolization of pseudoephedrine glycinamide. *Angew. Chem. Int. Ed.* **53**, 4642–4647 (2014).
- Trost, B. M. & Miede, F. Development of prophenol ligands for the diastereo- and enantioselective synthesis of β -hydroxy- α -amino esters. *J. Am. Chem. Soc.* **136**, 3016–3019 (2014).
- Prier, C. K., Zhang, R. K., Buller, A. R., Brinkmann-Chen, S. & Arnold, F. H. Enantioselective, intermolecular benzylic C-H amination catalysed by an engineered iron-haem enzyme. *Nat. Chem.* **9**, 629–634 (2017).
- Yu, H. et al. Bioamination of alkane with ammonium by an artificially designed multienzyme cascade. *Metab. Eng.* **47**, 164–189 (2018).
- Mutti, F. G., Knaus, T., Scrutton, N. S., Breuer, M. & Turner, N. J. Conversion of alcohols to enantiopure amines through dual-enzyme hydrogen-borrowing cascades. *Science* **349**, 1525–1529 (2015).
- Hepworth, L. J. et al. Enzyme cascades in whole cells for the synthesis of chiral cyclic amines. *ACS Catal.* **7**, 2920–2925 (2017).
- Li, A. et al. Whole-cell-catalyzed multiple regio- and stereoselective functionalizations in cascade reactions enabled by directed evolution. *Angew. Chem. Int. Ed.* **55**, 12026–12029 (2016).
- Schrittweiser, J. H., Velikogne, S., Hall, M. & Kroutil, W. Artificial biocatalytic linear cascades for preparation of organic molecules. *Chem. Rev.* **118**, 270–348 (2018).
- Yao, P. et al. Efficient biosynthesis of (R)- or (S)-2-hydroxybutyrate from L-threonine through a synthetic biology approach. *Adv. Synth. Catal.* **358**, 2923–2928 (2016).
- Both, P. et al. Whole-cell biocatalysts for stereoselective C-H amination reactions. *Angew. Chem. Int. Ed.* **55**, 1511–1513 (2016).

28. Wu, S. et al. Highly regio- and enantioselective multiple oxy- and amino-functionalizations of alkenes by modular cascade biocatalysis. *Nat. Commun.* **7**, 11917 (2016).
29. Parmeggiani, F., Weise, N. J., Ahmed, S. T. & Turner, N. J. Synthetic and therapeutic applications of ammonia-lyases and aminomutases. *Chem. Rev.* **118**, 73–118 (2017).
30. Xue, Y., Cao, C. & Zheng, Y. Enzymatic asymmetric synthesis of chiral amino acids. *Chem. Soc. Rev.* **47**, 1516–1561 (2018).
31. Franz, S. E. & Stewart, J. D. Threonine aldolases. *Adv. Appl. Microbiol.* **88**, 57–101 (2014).
32. Fesko, K. et al. Four types of threonine aldolases: similarities and differences in kinetics/thermodynamics. *J. Mol. Catal. B Enzym.* **52**, 19–26 (2008).
33. Xu, G., Yu, H., Zhang, X. & Xu, J. Access to optically active aryl halohydrins using a substrate-tolerant carbonyl reductase discovered from *Kluyveromyces thermotolerans*. *ACS Catal.* **2**, 2566–2571 (2012).
34. Toney, M. D. Controlling reaction specificity in pyridoxal phosphate enzymes. *Arch. Biochem. Biophys.* **433**, 279–287 (2011).
35. Bayer, T. et al. *In vivo* synthesis of polyhydroxylated compounds from a “hidden reservoir” of toxic aldehyde species. *ChemCatChem* **9**, 2919–2923 (2017).
36. Gourinchas, G. et al. A synthetic biology approach for the transformation of L- α -amino acids to the corresponding enantiopure (R)- or (S)- α -hydroxy acids. *Chem. Commun.* **51**, 2828–2831 (2015).
37. Busto, E., Richter, N., Grischek, B. & Kroutil, W. Biocontrolled formal inversion or retention of L- α -amino acids to enantiopure (R)- or (S)-hydroxyacids. *Chem. Eur. J.* **20**, 11225–11228 (2014).
38. Yamada, A., Dairi, T., Ohno, Y., Huang, X. L. & Asano, Y. Nucleotide sequencing of phenylalanine dehydrogenase gene from *Bacillus badius* IAM 11059. *Biosci., Biotechnol. Biochem.* **59**, 1994–1995 (1995).
39. Gao, X. et al. A newly determined member of the *meso*-diaminopimelate dehydrogenase family with a broad substrate spectrum. *Appl. Environ. Microbiol.* **83**, e00476–00417 (2017).
40. Fesko, K. Threonine aldolases: perspectives in engineering and screening the enzymes with enhanced substrate and stereo specificities. *Appl. Microbiol. Biotechnol.* **100**, 2579–2590 (2016).
41. Windle, C. L., Müller, M., Nelson, A. & Berry, A. Engineering aldolases as biocatalysts. *Curr. Opin. Chem. Biol.* **19**, 25 (2014).
42. Hernández, K., Szekrenyi, A. & Clapés, P. Nucleophile promiscuity of natural and engineered aldolases. *ChemBioChem.* **19**, 1–7 (2018).
43. Fesko, K., Strohmeier, G. A. & Breinbauer, R. Expanding the threonine aldolase toolbox for the asymmetric synthesis of tertiary α -amino acids. *Appl. Microbiol. Biotechnol.* **99**, 9651–9661 (2015).

Acknowledgements

This work is supported by the National First-class Discipline Program of Light Industry Technology and Engineering (LITE2018-08), the Key Technologies R & D Program of Jiangsu Province (BE2018623), and the National Natural Science Foundation of China (21878126).

Author contributions

W.S. and L.-M.L. conceived the work and designed the experiments. W.S. and J.-H.W. performed all the experiments. W.S., J.-H.W., J.W., X.-L.C., and L.-M.L. analyzed the results and drafted the manuscript. All authors read, edited and approved the final manuscript.

Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-018-06241-x>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018