

# Direct asymmetric synthesis of $\beta$ -branched aromatic $\alpha$ -amino acids using engineered phenylalanine ammonia lyases

Received: 14 March 2024

Accepted: 16 September 2024

Published online: 26 September 2024

Check for updates

Chenghai Sun<sup>1</sup>✉, Gen Lu<sup>2</sup>, Baoming Chen<sup>2</sup>, Guangjun Li<sup>2</sup>, Ya Wu<sup>2</sup>, Yannik Brack<sup>1</sup>, Dong Yi<sup>3</sup>, Yu-Fei Ao<sup>1</sup>, Shuke Wu<sup>4</sup>, Ren Wei<sup>1</sup>, Yuhui Sun<sup>2</sup>, Guifa Zhai<sup>2</sup>✉ & Uwe T. Bornscheuer<sup>1</sup>✉

$\beta$ -Branched aromatic  $\alpha$ -amino acids are valuable building blocks in natural products and pharmaceutically active compounds. However, their chemical or enzymatic synthesis is challenging due to the presence of two stereocenters. We design phenylalanine ammonia lyases (PAL) variants for the direct asymmetric synthesis of  $\beta$ -branched aromatic  $\alpha$ -amino acids. Based on extensive computational analyses, we unravel the enigma behind PAL's inability to accept  $\beta$ -methyl cinnamic acid ( $\beta$ -MeCA) as substrate and achieve the synthesis of the corresponding amino acids of  $\beta$ -MeCA and analogs using a double (PcPAL-L256V-I460V) and a triple mutant (PcPAL-F137V-L256V-I460V). The reactions are scaled-up using an optimized *E. coli* based whole-cell biotransformation system to produce ten  $\beta$ -branched phenylalanine analogs with high diastereoselectivity ( $dr > 20:1$ ) and enantioselectivity ( $ee > 99.5\%$ ) in yields ranging from 41–71%. Moreover, we decipher the mechanism of PcPAL-L256V-I460V for the acceptance of  $\beta$ -MeCA and converting it with excellent stereoselectivity by computational simulations. Thus, this study offers an efficient method for synthesizing  $\beta$ -branched aromatic  $\alpha$ -amino acids.

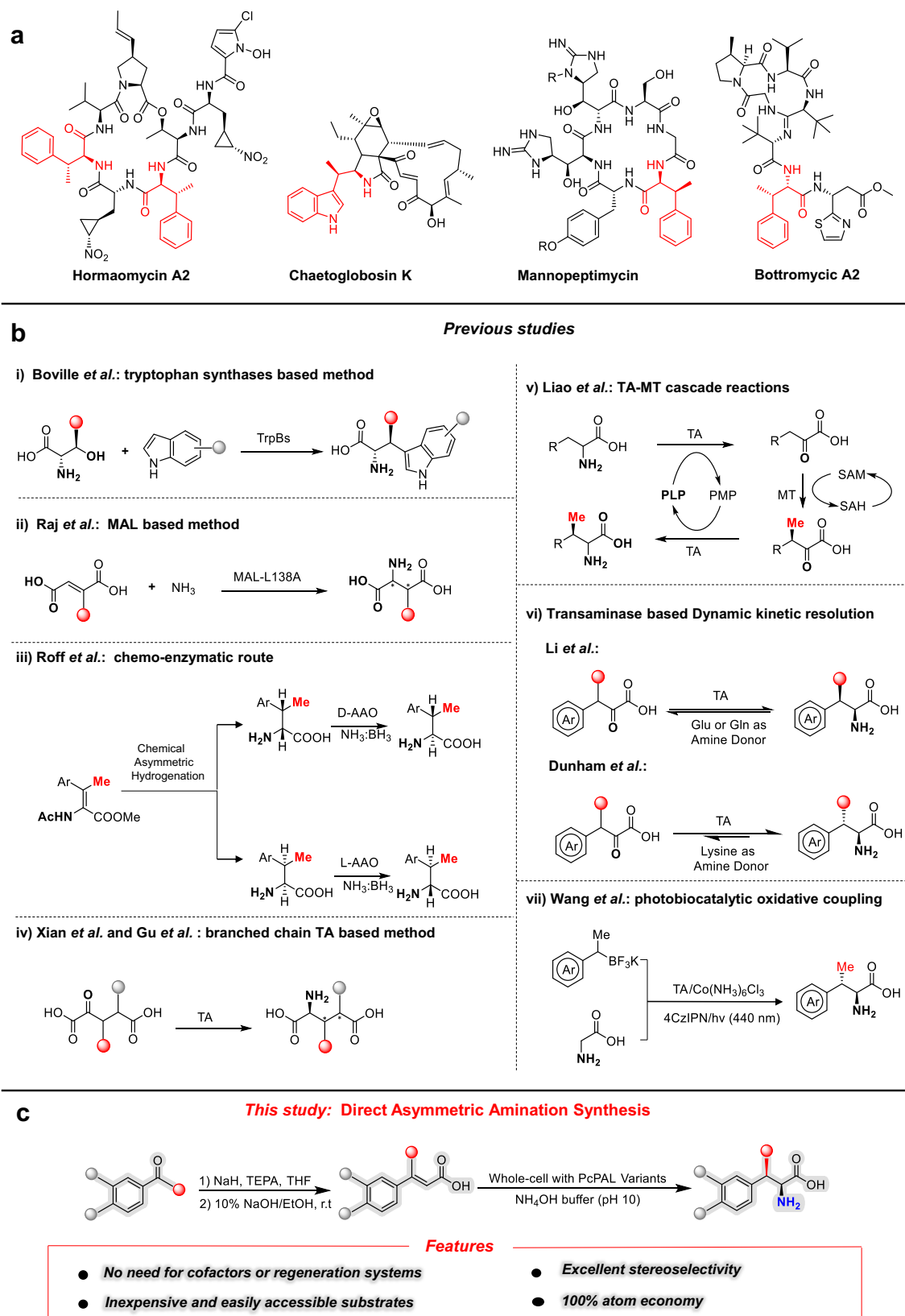
$\beta$ -branched aromatic  $\alpha$ -amino acids are an important class of non-proteinogenic amino acids, which have been found in various natural products and bioactive compounds (Fig. 1a)<sup>1–6</sup>. Although several studies have demonstrated that adding  $\beta$ -substituents can significantly enhance the biological activity of the corresponding parent molecules<sup>7–10</sup>, the incorporation of these substituents is still challenging due to the presence of two consecutive chiral centers in  $\beta$ -branched  $\alpha$ -amino acids. To date, only two chemical approaches have been developed to synthesize  $\beta$ -branched  $\alpha$ -amino acids. One approach involved the asymmetric hydrogenation of  $\beta,\beta$ -disubstituted dihydroamino acids<sup>11</sup>, while the other one was direct asymmetric C-H activation using aliphatic quinolyl carboxamide analogs as starting

materials<sup>12,13</sup>. However, the limitations of these chemical asymmetric synthesis constrained their widespread adoption in industry, especially because of the excessive usage of rare transition-metal catalysts, the requirement of pre-installed directing groups, and the suboptimal diastereoselectivity.

In comparison with chemical synthesis, biocatalysis has played an increasingly pivotal role in the synthesis of chiral compounds, owing to its mild reaction conditions and exceptional stereoselectivity<sup>14–17</sup>. In recent years, several enzymatic processes have been developed to produce  $\beta$ -branched  $\alpha$ -amino acids, including: (i) engineering of tryptophan synthases (TrpBs) to generate  $\beta$ -alkyl tryptophan analogs<sup>18</sup>; (ii) engineering of methylaspartate ammonia lyases (MALs) for the

<sup>1</sup>Department of Biotechnology and Enzyme Catalysis, Institute of Biochemistry, University of Greifswald, Greifswald, Germany. <sup>2</sup>School of Pharmacy, Tongji Medical College of Huazhong University of Science and Technology, Hubei Key Laboratory of Natural Medicinal Chemistry and Resource Evaluation, Wuhan, China. <sup>3</sup>Research Center for Systems Biosynthesis, China State Institute of Pharmaceutical Industry, National Key Laboratory of Lead Druggability Research, Shanghai, China. <sup>4</sup>College of Life Science and Technology, Huazhong Agriculture University, Wuhan, China.

✉ e-mail: [chenghai.sun23@outlook.com](mailto:chenghai.sun23@outlook.com); [gfzhai@hust.edu.cn](mailto:gfzhai@hust.edu.cn); [uwe.bornscheuer@uni-greifswald.de](mailto:uwe.bornscheuer@uni-greifswald.de)



**Fig. 1 | Representative natural products and strategies for synthesizing  $\beta$ -branched aromatic amino acids.** **a** Examples of bioactive compounds containing  $\beta$ -branched aromatic amino acids. **b** Reported<sup>18–26</sup> strategies for the biocatalytic synthesis of  $\beta$ -branched aromatic amino acids. **c** Direct asymmetric amination synthesis method developed in this study. MAL methylaspartate ammonia lyase,

AAO amino acid oxidase, TA transaminase, PLP pyridoxal-5'-phosphate, PMP pyridoxamine-5'-phosphate, MT methyltransferase, SAM S-Adenosylmethionine, SAH S-adenosylhomocysteine, THF Tetrahydrofuran, TEPA Triethylphosphonoacetate, PcPAL Phenylalanine ammonia lyase from *Petroselinum crispum*.

synthesis of  $\beta$ -branched aspartate derivatives<sup>19</sup>; (iii) chemo-enzymatic approaches employing amino acid oxidases (AAOs) to obtain enantiomerically pure  $\beta$ -methylated  $\alpha$ -amino acids<sup>20</sup>; (iv) utilization of transaminases (TAs) for the production of branched glutamate analogs<sup>21, 22</sup>; (v) integration of cascade reactions involving TAs and methyltransferases to produce  $\beta$ -methylated  $\alpha$ -amino acids<sup>23</sup>; (vi) implementing dynamic kinetic resolution techniques based on TAs<sup>24, 25</sup> and (vii) photobiocatalytic oxidative coupling method to prepare  $\beta$ -methylphenylalanines<sup>26</sup> (Fig. 1b). Despite this progress propelled by the above concepts, some of them are hampered by the need of (light-sensitive) pyridoxal-5'-phosphate (PLP) or *S*-adenosylmethionine (SAM) requiring a complex cofactor regeneration system. Although no co-factors are required for the reactions catalyzed by MALs, these enzymes often exhibit inadequate diastereoselectivity<sup>19</sup>.

Compared to the above biosynthetic reactions, a direct asymmetric amination of  $\beta$ -branched cinnamic acid (CA), and its analogs in an anti-Michael regioselectivity way, is much more attractive to synthesize  $\beta$ -branched aromatic  $\alpha$ -amino acids. In principle, phenylalanine ammonia lyases (PALs, EC 4.3.1.24) possess the unique capability to catalyze non-oxidative and reversible amination reactions using CA-like substrates<sup>27–30</sup>. Thus, they stand out as promising enzymes for such anti-Michael amination reactions. Besides, PALs can intrinsically generate the catalytic prosthetic group 3,5-dihydro-5-methylidene-4H-imidazol-4-one (MIO) through autocatalysis, eliminating the need for cofactors and cofactor regeneration systems. These features collectively render PALs particularly valuable in industrial applications for the synthesis of phenylalanine analogs, which has been evidenced by several leading companies, including Pfizer (USA)<sup>31</sup>, DSM (Netherlands)<sup>32</sup>, Novartis (Switzerland)<sup>33</sup>, and Johnson Matthey (UK)<sup>34</sup>.

In previous reports, PALs from *Rhodotorula glutinis* (RgPAL)<sup>35–37</sup>, *Rhodospiridium toruloides* (RtPAL)<sup>38</sup>, *Petroselinum crispum* (PcPAL)<sup>39</sup>, and *Anabaena variabilis* (AvPAL)<sup>40, 41</sup> have been well studied, especially including the engineering of their hydrophobic binding pockets to expand their substrate scope<sup>42–47</sup>. However, in these investigations, only CA analogs with single or multiple substituents exclusively present at the phenyl ring could be converted, including heterocyclic systems, bulkier aryl systems, or naphthalene rings<sup>36, 43, 48</sup>. An early study in 1994 demonstrated that  $\beta$ -deuterated phenylalanine derivatives could be converted in PAL-mediated deamination reactions<sup>49</sup>. Subsequently, Turner et al. achieved the amination reaction using AvPAL with  $\beta$ -deuterated cinnamic acid<sup>50</sup>. Although this reaction resulted in a low enantiomeric excess (ee) value, it provided evidence that  $\beta$ -deuterated substrates are accepted by PALs. However, for other substitutions apart from deuterium, such as fluorine and methyl substitutions, no conversion was observed with eukaryotic RgPAL<sup>36</sup>, bacterial AvPAL<sup>24</sup>, or the corresponding amino acid mutase (PAM)<sup>51</sup>. These results underscore the substantial challenges associated with utilizing PALs for the amination of CA analogs bearing  $\beta$ -substituents such as alkyl.

Here, we report a computational analysis based rational design approach to decipher the underlying reason for the inability of PALs to accept  $\beta$ -methyl-cinnamic acid ( $\beta$ -MeCA) as substrate. This leads to the creation of two efficient PAL mutants (PcPAL-L256V-I460V and PcPAL-F137V-L256V-I460V) which catalyze the direct asymmetric amination of various  $\beta$ -MeCA analogs with high diastereoselectivity (dr > 20:1) and enantioselectivity (ee > 99.5%) (Fig. 1c). This work expands not only the substrate scope of PALs, but especially provides a direct reaction pathway for the synthesis of  $\beta$ -branched aromatic  $\alpha$ -amino acids.

## Results

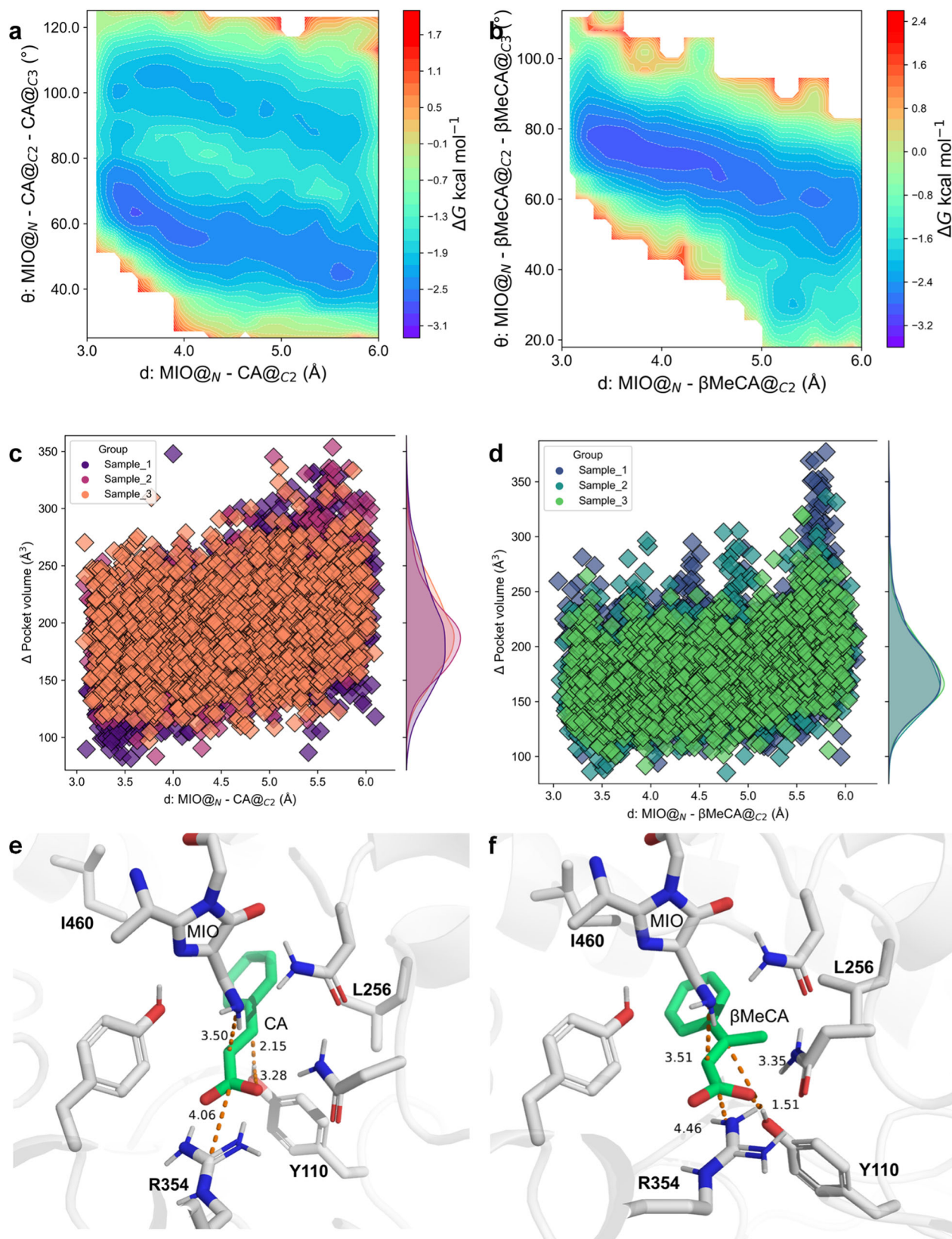
### In silico analyses of PcPAL towards $\beta$ -MeCA

Currently, the E1cb (via a carbanion intermediate) and the E2 (concerted) mechanisms are widely accepted for PcPAL-mediated deamination reactions<sup>27</sup>. Consequently, the amination reactions catalyzed by PcPAL are likely to proceed through the reverse E1cb and E2 mechanisms (Fig. S1). Despite an ongoing debate between these

mechanisms, both of them involve two critical steps in the amination reactions: amination and protonation. The key residues involved in these processes include the MIO-NH<sub>2</sub> adduct, Tyr110, and R354. Among them, the MIO-NH<sub>2</sub> adduct is responsible for amidating the  $\alpha$ C of the substrate, Tyr110 is responsible for protonating the  $\beta$ C, and R354 forms strong hydrogen bonds with the carboxyl group of the substrate to position the substrate. Based on these catalytic processes, we conducted a series of in silico analyses to elucidate the underlying reasons for the inability of PALs to recognize  $\beta$ -substituted substrates. Firstly, CA and  $\beta$ -MeCA were individually docked into PcPAL (PDB:6HQF)<sup>52</sup> (Fig. S2a), and based on the docking results, we further performed conventional molecular dynamics (cMD) simulations. The results showed that in the ( $\beta$ -MeCA)/(PcPAL-WT) complex, the distance variation range between the N atom of MIO and the  $\alpha$ C of  $\beta$ -MeCA was narrower than that in the (CA)/(PcPAL-WT) complex (Figs. S3a, b). Besides, by examining the hydrogen bond interactions between the substrates and amino acids within the active pocket (Source Data 3), it was found that in the ( $\beta$ -MeCA)/(PcPAL-WT) complex, the carboxyl group of  $\beta$ -MeCA forms a high-frequency hydrogen bond with Tyr110, appearing in 99% of the recorded 5000 snapshots (Source Data 3). These evidences suggest that  $\beta$ -MeCA is more stable within the pocket. Although the reason for the inactivity of the ( $\beta$ -MeCA)/(PcPAL-WT) complex remains unclear, based on the above observations, we reasonably speculate that the inactivity is likely due to the perturbation of the substrate positioning dynamics<sup>53, 54</sup>.

To validate our hypothesis, we further carried out umbrella sampling (US) simulations to determine the optimal amination distances (distance 1 in Fig. S4) when CA and  $\beta$ -MeCA have the lowest energy conformations in PcPAL-WT. For substrate CA, the optimal amination distance ranged from 3.5 to 3.9 Å when it had the lowest energy conformations in PcPAL-WT, whereas the corresponding distance ranged from 3.6 to 3.7 Å for  $\beta$ -MeCA (Fig. S2b). Besides these different distance ranges detected in the presence of CA and  $\beta$ -MeCA, we also observed differences related to the angle between MIO and the double bond of the substrates, which is important for the amination process. Specifically, we observed two angles in the (CA)/(PcPAL-WT) complex (70° and 109°, Fig. 2a). However, in the ( $\beta$ -MeCA)/(PcPAL-WT) complex, only one angle was detected around 80° (Fig. 2b). According to the Bürgi-Dunitz angle theory<sup>55</sup>, the N atom of MIO should attack from the direction of  $105 \pm 5^\circ$  relative to the C=C bond to achieve effective orbital overlap for C-N bond formation. This missing bond-forming angle, along with the relatively narrow optimal amination distance range detected in the ( $\beta$ -MeCA)/(PcPAL-WT) complex, suggests that  $\beta$ -MeCA is less flexible and more stable than CA. Further, we compared the pocket volumes in the presence of both CA and  $\beta$ -MeCA. The results showed that the mean pocket volumes of the (CA)/(PcPAL-WT) complex and the ( $\beta$ -MeCA)/(PcPAL-WT) complex were 192.8 Å<sup>3</sup> and 177.4 Å<sup>3</sup>, respectively (Fig. 2c, d). The smaller pocket volume within the ( $\beta$ -MeCA)/(PcPAL-WT) complex indicates that  $\beta$ -MeCA was confined to a very narrow cavity, which restricts its dynamic changes. Overall, all the US simulation results observed in the ( $\beta$ -MeCA)/(PcPAL-WT) complex indicated that the presence of  $\beta$ -methyl substitution may disrupt the normal substrate positioning dynamic.

Following, Quantum Mechanics/Molecular Mechanics (QM/MM) simulations were employed to further optimize the conformations of PcPAL-WT and the substrate complexes. The QM/MM simulation results showed that, compared to the CA substrate (Fig. 2e) a strong hydrogen bond was detected between the hydroxyl hydrogen of Tyr110 and the carboxyl group of  $\beta$ -MeCA (1.51 Å, Fig. 2f). The formation of this strong hydrogen bond is thermodynamically unfavorable in the enzymatic reaction of PcPAL-WT, not only restricting the flexibility of  $\beta$ -MeCA in the active pocket but also significantly impeding the proton transfer process from Tyr110 to the  $\beta$ C-carbon atom.



**Fig. 2 | In silico analyses for the substrate recognition mechanism of PcPAL.** **a, b** Gibbs energy diagram of the cMD simulations projected onto the distance from the amino N of MIO to the  $\alpha$ C of the substrates (X-axis) and the angle between MIO and the double bond of the substrates (Y-axis) within the (CA)/(PcPAL-WT) complex and the ( $\beta$ -MeCA)/(PcPAL-WT) complex, respectively. **c, d** the pocket volumes detected during the cMD simulations of (CA)/(PcPAL-WT)

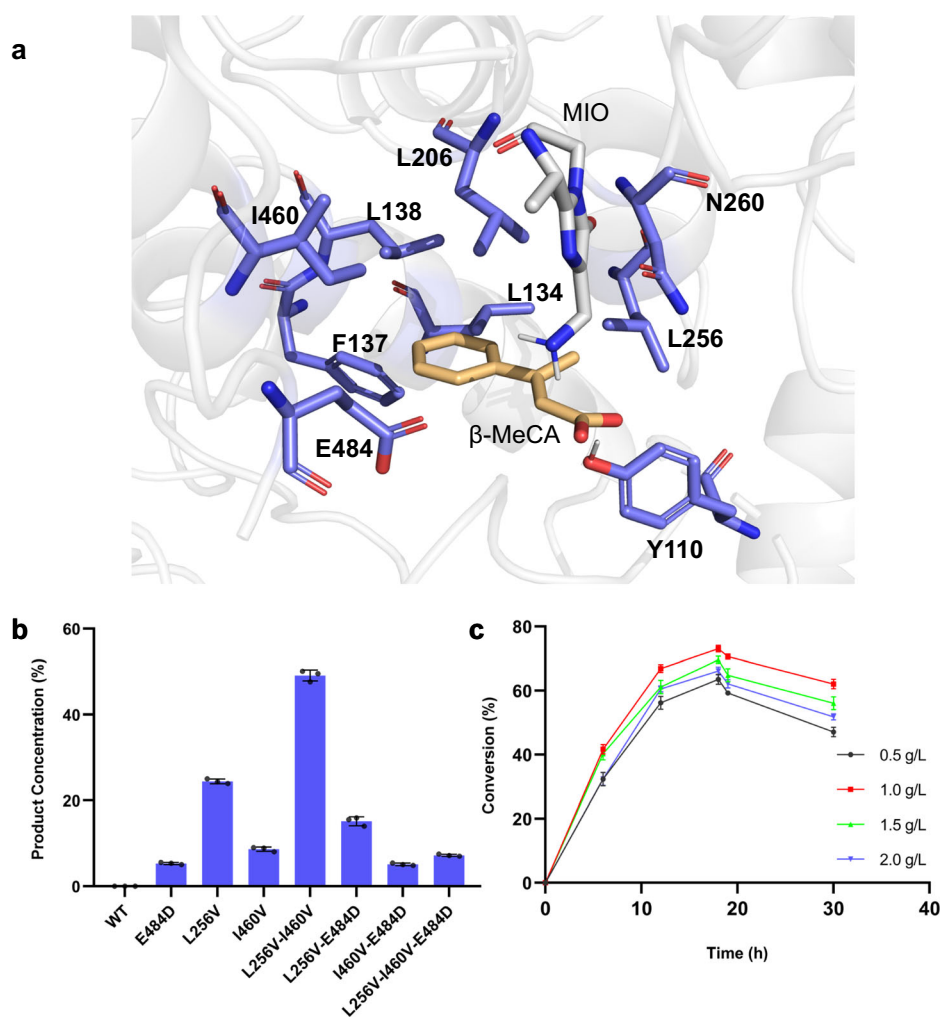
complex and the ( $\beta$ -MeCA)/(PcPAL-WT) complex, respectively. **e, f** Optimized pre-reaction states of CA and  $\beta$ -MeCA in the PcPAL-WT using the Quantum Mechanics/Molecular Mechanics (QM/MM) methodology, respectively. Source data are provided in Source Data 3 to Source Data 5. CA: cinnamic acid;  $\beta$ -MeCA:  $\beta$ -methyl cinnamic acid; MIO: 3,5-dihydro-5-methylidene-4H-imidazol-4-one.

In summary, our computational simulations indicated that the incapacity of PcPAL-WT to aminate  $\beta$ -MeCA raised from the perturbation of  $\beta$ -MeCA's positioning dynamic within the active pocket, which is primarily due to the strong hydrogen bond formation between the carboxyl group of  $\beta$ -MeCA and Tyr110. Consequently, we hypothesized that alleviating the stringent active pocket of PcPAL might help restore the correct positioning dynamic of  $\beta$ -MeCA analogs to achieve the asymmetric amination for the synthesis of  $\beta$ -branched amino acids. Besides, our computational results indicated that to induce reactive conformations, it is essential to engineer not only the residues around the  $\beta$ -methyl group of  $\beta$ -MeCA, but also those around its phenyl ring. This is crucial as both sets of residues were essential to form intricate interactions with  $\beta$ -MeCA and influence its conformation in the active pocket.

### Rational engineering of PcPAL for the amination of $\beta$ -MeCA

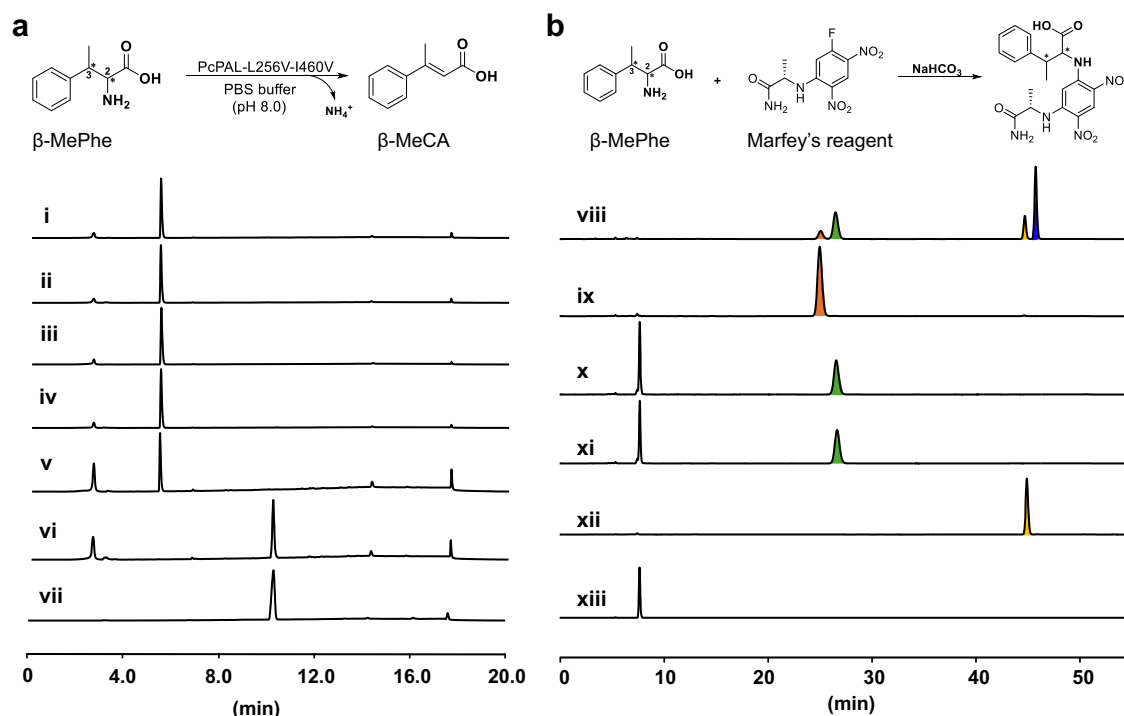
As the above computational results provided us with understandings for the rational engineering of PcPAL, we identified eight key residues around the substrate binding pocket (L134, F137, L138, I460, E484, L256, N260, and L206), which could affect the correct conformation of  $\beta$ -MeCA (Fig. 3a). Among them, L256 and N260 are

located near the  $\beta$ -Me group, while the others are around the phenyl ring of  $\beta$ -MeCA. Subsequently, each of these residues was individually substituted with three smaller residues (A, G, and V). Additionally, we also investigated three other mutants, previously established in our laboratory (E484D, F137V-I460V<sup>43</sup>, and F137A-I460A<sup>43</sup>) to explore their potential. Notably, the latter two mutants, developed by the Bencze's group<sup>43</sup>, were investigated for their ability to catalyze the amination of bulky substrates, such as styrylacrylate and biarylalanines. After recombinant expression in *E. coli* BL21 (DE3) and purification of these 27 mutants, an in vitro screening against  $\beta$ -MeCA was performed. Remarkably, after reacting with three mutants, i.e., L256V, I460V, and E484D, the product concentrations in the reaction system achieved 24.3%, 8.7%, and 5.3%, respectively. Interestingly, L256V and I460V mutants were well studied previously for synthesizing various  $\alpha$ -phenylalanine analogs<sup>43–45,56,57</sup>, yet their ability to convert  $\beta$ -substituted CA analogs was not explored yet. Through analyzing the conformation of  $\beta$ -MeCA within the docking result, we found that L256 is positioned in close proximity to the  $\beta$ -methyl group of the substrate (Fig. 3a). After substitution with Val, more space becomes available to accommodate the  $\beta$ -methyl group. This is a



**Fig. 3 | Results of engineering and optimizing whole-cell catalysis conditions for PcPAL in  $\beta$ -MeCA amination.** **a** Key bulky residues around the substrate  $\beta$ -MeCA modeled in PcPAL-WT. The active site illustrations are based on the crystal structure of PcPAL (Protein Data Bank ID: 6HQF). The substrate and MIO group are highlighted in orange and gray, respectively. **b** Enzyme activity of PcPAL mutants towards  $\beta$ -MeCA. Reaction conditions: 100  $\mu$ g purified enzyme, 1 mM  $\beta$ -MeCA, total 200  $\mu$ l in  $\text{NH}_4\text{OH}$  buffer (pH 10.0), 30  $^\circ\text{C}$ , 450 rpm for 24 h. After that the reaction

was stopped by adding HCl and centrifuged for further HPLC analysis. **c** Time curves with different substrate concentrations. Three parallel samples were collected from each reaction for HPLC analysis. Product concentrations were calculated using a standard curve and the peak area of the product. Error bars (SD) were derived from triplicate experiments and created using GraphPad Prism 8. Data are presented as mean values  $\pm$  SD. Source data are provided in Source Data 1.  $\beta$ -MeCA:  $\beta$ -methyl cinnamic acid; MIO: 3,5-dihydro-5-methylidene-4H-imidazol-4-one.



**Fig. 4 | HPLC data for determining the absolute configuration of the amination product.** **a** In vitro deamination assays with  $\beta\text{-MePhe}$  of different absolute configurations using the PcPAL-L256V-I460V mutant. Reaction conditions: 100  $\mu\text{g}$  purified enzyme, 1 mM various  $\beta\text{-MePhe}$ , total 500  $\mu\text{L}$  in PBS buffer (pH 8.0), 30  $^\circ\text{C}$ , 450 rpm for 24 h. i) (2*S*, 3*S*)- $\beta\text{-MePhe}$ ; ii) deamination reaction of (2*S*, 3*S*)- $\beta\text{-MePhe}$ ; iii) (2*R*, 3*R*)- $\beta\text{-MePhe}$ ; iv) deamination reaction of (2*R*, 3*R*)- $\beta\text{-MePhe}$ ; v) (2*S*, 3*R*)- $\beta\text{-MePhe}$ ; vi) deamination reaction of (2*S*, 3*R*)- $\beta\text{-MePhe}$ ; vii)  $\beta\text{-MeCA}$  standard.

**b** Derivatization results with  $\beta\text{-MePhe}$  of different configurations and the in vitro

reaction product formed by using Marfey's reagent. viii) standard mixture of  $\beta\text{-MePhe}$  derivatized with Marfey's reagent; ix) (2*S*, 3*S*)- $\beta\text{-MePhe}$  derivatized with Marfey's reagent; x) (2*S*, 3*R*)- $\beta\text{-MePhe}$  derivatized with Marfey's reagent; xi) derivatization of the in vitro amination reaction product produced by the PcPAL-L256V-I460V mutant; xii) (2*R*, 3*R*)- $\beta\text{-MePhe}$  derivatized with Marfey's reagent; xiii) blank containing solely Marfey's reagent and enzyme.  $\beta\text{-MePhe}$   $\beta\text{-methyl phenylalanine}$ ,  $\beta\text{-MeCA}$   $\beta\text{-methyl cinnamic acid}$ .

plausible explanation for the significantly higher activity exhibited by the L256V mutant. I460 and E484 are located around the phenyl group of  $\beta\text{-MeCA}$ . After turning them into smaller residues, we thought that the substrate might undergo spatial displacement within the active pocket. This should lead to the generation of additional conformations, among which favorable ones for the reaction may be present, thereby demonstrating reaction activity. To find much better performing variants, we further combined these promising mutants to generate four additional mutants (L256V-I460V, L256V-E484D, I460V-E484D, and L256V-I460V-E484D). Among them, the L256V-I460V mutant showed the best amination activity with a product concentration in the reaction system of 49.1% (Fig. 3b). Contrary to our initial expectations, the double mutant L256V-I460V showed much better activity than the triple mutant L256V-I460V-E484D (product concentrations in the reaction system: 7.2%), indicating that residues at positions 256 and 460 were the most crucial mutation sites. Additionally, it is intriguing to note that when the E484D mutation was combined with other mutations, it consistently led to a decrease in the activity of the original mutants, which indicated that the multi-point mutation combined with E484D seems to be unfavorable to ensure a correct binding position of  $\beta\text{-MeCA}$  in the active site pocket.

To further evaluate the catalytic efficiency of the PcPAL-L256V-I460V mutant, we measured its kinetic constants in the deamination reaction using (2*S*, 3*R*)- $\beta\text{-MePhe}$  as substrate (Fig. S6). The results revealed a relatively high  $K_m$  value ( $K_m = 0.7304 \pm 0.1195$  mM), which indicated the low affinity of (2*S*, 3*R*)- $\beta\text{-MePhe}$  to the mutant. This weak affinity allows the product to leave the catalytic pocket more easily during the amination reaction, thereby facilitating the amination reaction.

### Determination of the absolute configuration of the amination product

Encouraged by the above findings, we elucidated the absolute configuration of the amination products by employing two distinct methods. In the first method, by utilizing the three  $\beta\text{-methyl phenylalanine}$  ( $\beta\text{-MePhe}$ ) standards with 2*S*,3*S*-, 2*S*,3*R*-, 2*R*,3*R*- configurations, the in vitro deamination reactions were conducted using the most efficient variant (PcPAL-L256V-I460V). High-performance liquid chromatography (HPLC) analysis demonstrated that only (2*S*,3*R*)- $\beta\text{-MePhe}$  was deaminated with full (~100%) conversion (Fig. 4a). Given the reversible nature of PAL-mediated reactions, we deduced that the stereo configuration of the amination product should hence be 2*S*,3*R*. To further validate our hypothesis, we conducted the second approach, in which Marfey's reagent was used to derivatize the above three  $\beta\text{-MePhe}$  standards as well as the crude amination reaction solution which was formed by the PcPAL-L256V-I460V mutant. As (2*R*, 3*S*)- $\beta\text{-MePhe}$  was not commercially available, the  $\beta\text{-MePhe}$  mixture standard which contains all four configurations was chosen for the derivatization reaction. After analyzing this by LC-HRMS, we detected the derivatization product of (2*R*, 3*S*)- $\beta\text{-MePhe}$  (blue peak in Fig. 4b viii), confirming the presence of the (2*R*, 3*S*)-typed substrate in the standard mixture. Moreover, we found that when the crude amination reaction solution was reacted with Marfey's reagent, a peak with the same retention time was observed as the one in the (2*S*, 3*R*)-type standard (Fig. 4b and Fig. S7). Hence, the results from both experiments clearly confirmed that the absolute configuration of the amination product was 2*S*, 3*R*. Furthermore, the amination reaction catalyzed by the PcPAL-L256V-I460V mutant proceeded with excellent diastereoselectivity ( $dr > 20:1$ ) and enantioselectivity ( $ee > 99.5\%$ ).

## Development of a whole-cell biotransformation process

To enhance the practicality of this direct asymmetric amination reaction for synthesizing  $\beta$ -branched aromatic amino acids, we developed a whole-cell biotransformation (Fig. 1c)<sup>44,56</sup>. Therefore, the gene of the PcPAL-L256V-I460V mutant was introduced into the strain *E. coli* BL21 (DE3)<sub>Δ</sub>tyrB, which lacks the aromatic amino acid transaminase gene *tyrB* and thus avoids potential transamination of the generated product within the cells<sup>58</sup>. After expressing the PcPAL-L256V-I460V mutant at 28 °C for 20 h, the cells were harvested and then resuspended in ammonia buffer for the biotransformation. To make this process efficient, we optimized different conditions by using  $\beta$ -MeCA as substrate and a commonly used 5 M ammonium hydroxide buffer (NH<sub>4</sub>OH buffer, pH 10) as the source of ammonia<sup>48,50</sup>. Considering the possible membrane permeability issues, clarified lysate of the harvested cells was used for transformation. However, the results showed that employing whole-cells containing the PcPAL-L256V-I460V gave significantly higher conversions than using clarified lysate, presumably due to higher enzyme stability *in vivo* under the same catalytic conditions (Fig. S8). Moreover, by resuspending the cultivated cells in NH<sub>4</sub>OH buffer to reach an OD<sub>600</sub> of 30, four distinct substrate concentrations (0.5, 1.0, 1.5, and 2.0 g/L) were systematically investigated. The maximum product concentrations in the reaction system (71%) was achieved at a substrate concentration of 1.0 g/L after 18 h reaction time. Longer reaction times lead to a gradual decline indicating a potential shift in the reaction equilibrium towards the reverse deamination reaction (Fig. 3c). Optimal reaction conditions were thus identified using 5 M NH<sub>4</sub>OH buffer, 1 g/L substrate concentration, and 18 h reaction time.

## Exploration of the substrate scope

Given the excellent stereoselectivity demonstrated by PcPAL-L256V-I460V, as well as the simple and practical whole-cell biotransformation method, we next explored the substrate scope of this amination reaction. 15  $\beta$ -MeCA analogs with different substituents at the phenyl ring were firstly synthesized using a straightforward chemical method (Fig. S9)<sup>59</sup>. Considering that the amino acids around the phenyl ring of substrates in the active pocket of PcPAL, such as L134, F138, and F138, could also potentially influence the conformations of substrates, we conducted screening assays on the synthesized substrates using our mutant library. The HPLC results showed that six substrates (**1**, **2**, **4** to **8**) could be converted by the PcPAL-L256V-I460V mutant. Despite substrate **3**, **9**, and **10** could be converted by both PcPAL-F137V-I460V and PcPAL-L256V mutants, the combination mutant PcPAL-F137V-L256V-I460V showed higher conversion for these three substrates (Fig. S10), highlighting the crucial role of the L256 position in enhancing reaction activity.

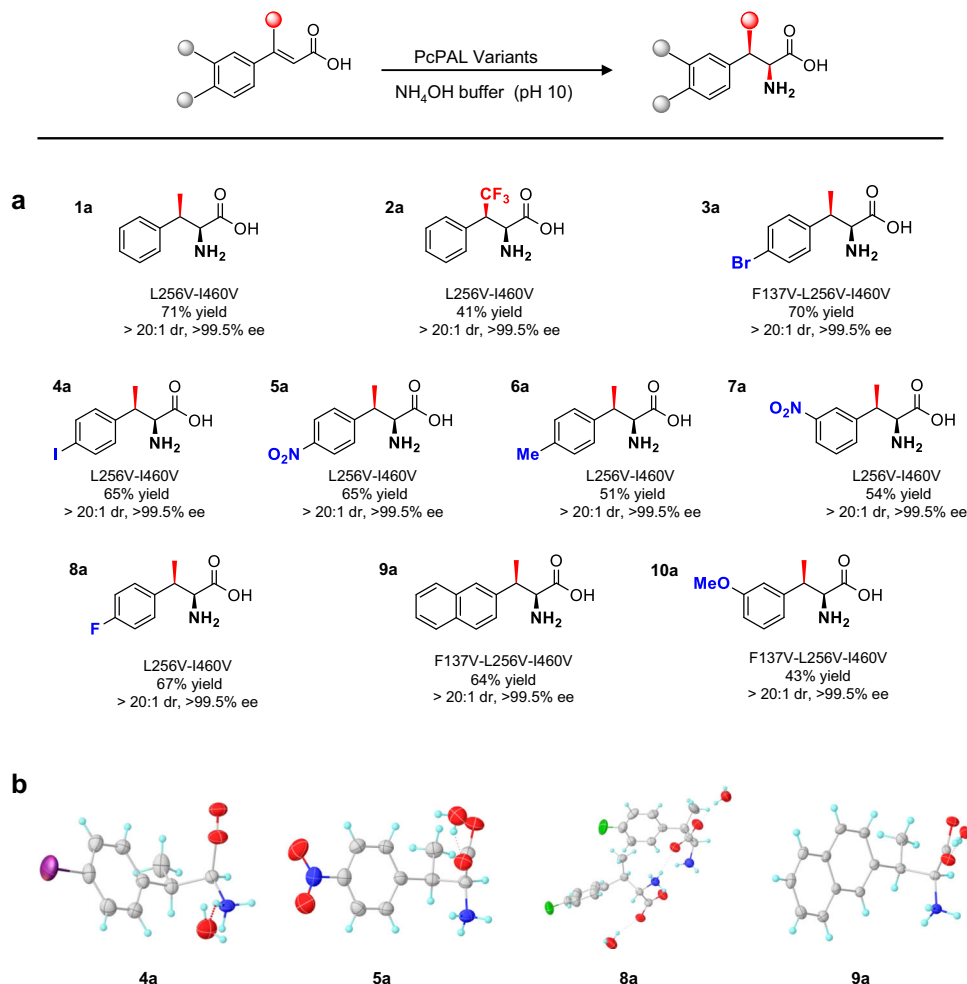
Next, we scaled up the reactions using the *E. coli* based whole-cell biotransformation method by using 30 mg substrates and obtained the ten  $\beta$ -substituted amino acid products with yields ranging from 41% to 71% (Fig. 5). In previous reports, the absolute configuration of  $\beta$ -substituted amino acid products were typically determined by referencing to previously reported NMR and specific rotation data<sup>24</sup>. For the newly synthesized  $\beta$ -substituted amino acid products reported here, we assigned the absolute configurations for four products (**4a**, **5a**, **8a**, and **9a**, Fig. 5b, Table S2 to S5) by X-ray crystallography and this unequivocally confirmed that the absolute configuration of all four crystallized products were 2*S*,3*R*, in accordance with the absolute configuration determined for  $\beta$ -MePhe **1a** (Fig. 4). In addition, a comprehensive analysis of all the products was carried out, including NMR and specific rotation data analysis. Remarkably, all these analytical data consistently confirmed that the configurations of the produced  $\beta$ -branched phenylalanine analogs were 2*S*, 3*R*. All chiral reaction products were obtained with high diastereoselectivity (*dr* > 20:1) and optical purity (*ee* > 99.5%) (Fig. 5).

Next, by comparing the yields of these productive reactions, we identified three factors that could affect the amination reaction efficiency. Firstly, the position of substitutions plays a significant role. When the electronegativity of the substituents was the same, substrates with *para* substitutions (e.g., **3a**, **4a**, **5a**, and **8a**) exhibited higher reactivity compared to those with *meta* substitutions (e.g., **p7**). Notably, substrates with *ortho* substitutions, such as **11** and **12** (Fig. S9), were not converted. Secondly, the electronegativity of substituents could also influence the outcomes. A comparison of yields between products **3a–5a** and **6a**, as well as **7a** and **10a**, indicates that electron-withdrawing groups could benefit the conversion. Thirdly, the size of the substitutions at the  $\beta$ -position also plays a crucial role. Although mutants could accommodate the CF<sub>3</sub> group at the  $\beta$ -position, which is similar in size to the methyl group, the yield of **2a** remained relatively low (41 %). The introduction of this strong electron-withdrawing CF<sub>3</sub> substituent should reduce the nucleophilicity of the  $\alpha$  carbon, thus resulting in a slow amination process. We thought that this accounts for the diminished activity observed for the CF<sub>3</sub>-substituted substrate. However, the ability of mutants to recognize the CF<sub>3</sub>-substituted substrates strongly implies that steric hindrance also plays a central role in impeding the reaction. Conversely, when larger substitutions like ethyl, propyl, and isopropyl were present at the  $\beta$ -position (**13** to **15**, Fig. S9), no reaction was observed. This observation further underscores the high sensitivity of PcPALs' active sites to steric effects at the  $\beta$ -position of the substrates.

## Elucidating the molecular basis of PcPAL-L256V-I460V for converting $\beta$ -MeCA

To reveal the molecular basis for accepting  $\beta$ -MeCA, as well as the excellent stereoselectivity during the PcPAL-L256V-I460V mediated amination reactions, the protein structure of this mutant was predicted using AlphaFold 2 and docked with  $\beta$ -MeCA into its active pocket. Subsequently, a series of computational simulations were conducted based on the docking results. Interestingly, similar simulation results were observed in the ( $\beta$ -MeCA)/(PcPAL-L256V-I460V) complex compared to the (CA)/(PcPAL-WT) complex. Specifically, in both complex systems, angles conforming to the Bürgi-Dunitz angle theory were detected, and both angles were around 109° (Figs. 6a and 2a). In addition, in the ( $\beta$ -MeCA)/(PcPAL-L256V-I460V) complex, the mean pocket volume of  $\beta$ -MeCA was around 207.9 Å<sup>3</sup>, much larger than that in the ( $\beta$ -MeCA)/(PcPAL-WT) complex (177.4 Å<sup>3</sup>, Figs. 6b and 2d), meaning that  $\beta$ -MeCA was more flexible in the mutant. Furthermore, the optimized conformation of the ( $\beta$ -MeCA)/(PcPAL-L256V-I460V) complex was determined through QM/MM simulations, revealing a distance of 3.28 Å between the hydroxyl hydrogen of Tyr110 and the carboxyl group of  $\beta$ -MeCA. This distance is greater than that from the hydroxyl hydrogen of Tyr110 to  $\beta$ C (2.11 Å) (Fig. 6c), with a corresponding hydrogen bond angle of 118.63° (Table S6, Source Data 5), indicating that the strong hydrogen bond interaction present in the ( $\beta$ -MeCA)/(PcPAL-WT) complex (Fig. 2f) does not exist in the mutant. Overall, the simulation results shown in Fig. 6a–c demonstrate that the conformation of  $\beta$ -MeCA in the mutant was almost same to that of CA in PcPAL-WT. Considering the conformation of CA in the PcPAL-WT, this could be used as a reference for assessing the potential occurrence of reactions when  $\beta$ -MeCA is used as a substrate. Indeed, the simulation results indicated that the positioning dynamic of  $\beta$ -MeCA was relatively flexible within PcPAL-L256V-I460V mutant, which could facilitate the amination reaction process.

To gain a deeper understanding of why  $\beta$ -MeCA showed different positioning dynamics within the PcPAL-WT and the PcPAL-L256V-I460V mutants, we further analyzed another two distances (distance 3: from  $\beta$ C of  $\beta$ -MeCA to amino acid 256; distance 4: from the benzene ring center of substrate to amino acid 460 in Fig. S4). In the complexes ( $\beta$ -MeCA)/(PcPAL-L256V-I460V) and (CA)/(PcPAL-WT), two distinct



**Fig. 5 | Substrate scope exploration and X-ray crystal structures of the amino acids formed.** **a** Structures of the produced  $\beta$ -branched aromatic  $\alpha$ -amino acids. Yields refer to the isolated product after purification. **b** X-ray crystal structures of

products **4a**, **5a**, **8a** and **9a**. Different colors represent different atoms: gray for carbon atoms, cyan for hydrogen atoms, red for oxygen atoms, blue for nitrogen atoms, and green for fluorine atoms. Source data are provided in Source Data 5.

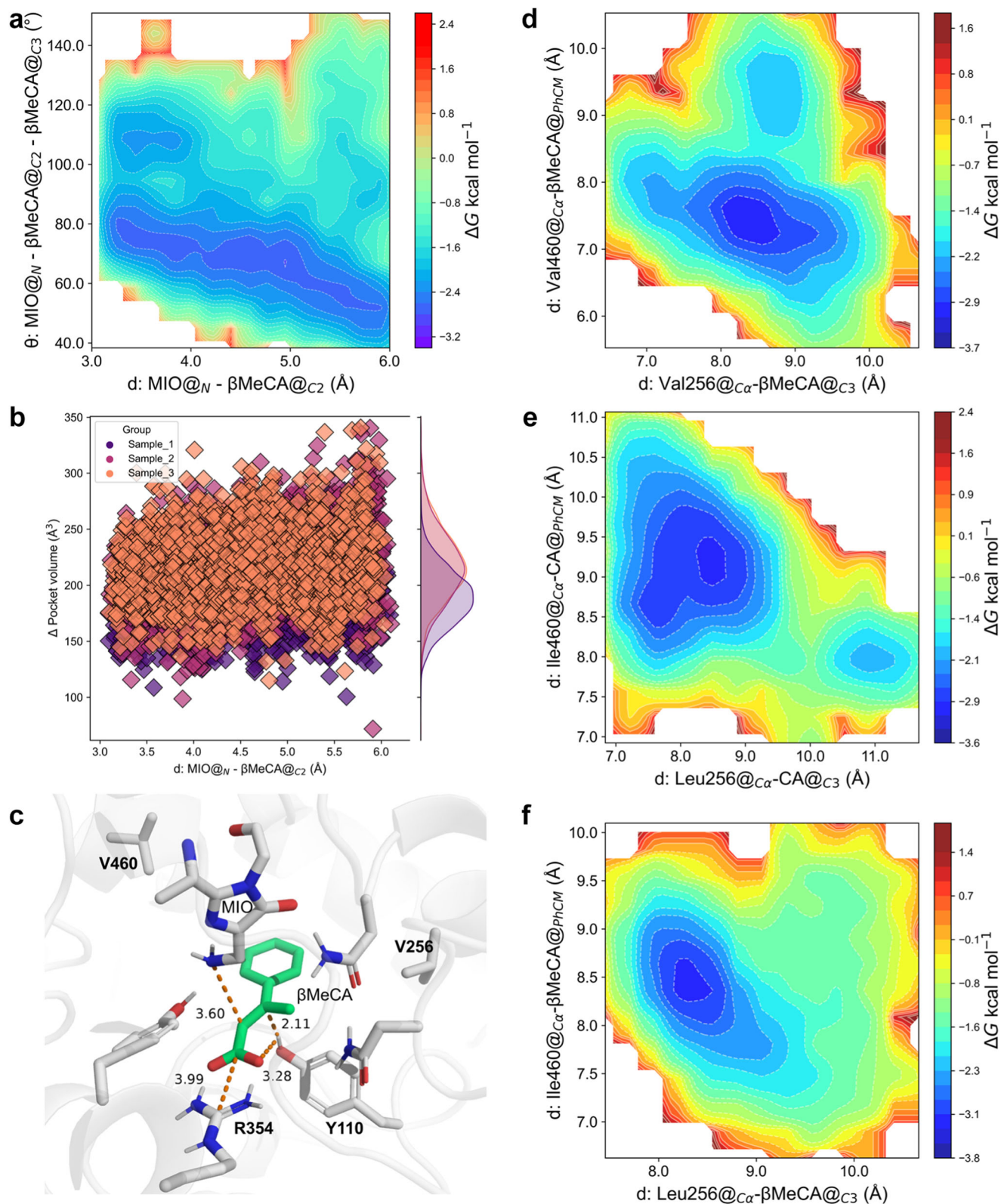
low-energy regions can be observed, each characterized by a broader primary energy well separated by a lower energy barrier (-1.2 kcal/mol), indicating that the substrate can readily traverse this barrier to adjust its conformation (Fig. 6d, e). Within the main low-energy regions, in the ( $\beta$ -MeCA)/(PcPAL-L256V-I460V) complex, distance 3 ranges from 7.9 to 8.6 Å, and distance 4 ranges from 8.0 to 8.9 Å (Fig. 6d). Similarly, in the (CA)/(PcPAL-WT) complex, distance 3 ranges from 7.4 to 8.9 Å, and distance 4 ranges from 8.3 to 9.8 Å (Fig. 6e). However, in the ( $\beta$ -MeCA)/(PcPAL-WT) complex, only one low-energy region is observed, with distance 3 ranging from 7.9 to 8.6 Å, and distance 4 ranging from 8.0 to 8.9 Å (Fig. 6f). These results indicate that compared to the ( $\beta$ -MeCA)/(PcPAL-WT) complex,  $\beta$ -MeCA in (PcPAL-L256V-I460V) exhibits a greater range of motion and a broader conformational space. Considering the wider angular range between MIO's N atom and the double bond in ( $\beta$ -MeCA)/(PcPAL-L256V-I460V) and (CA)/(PcPAL-WT) complexes, these findings support our hypothesis that substrate positioning dynamics influence the occurrence of the reaction.

Additionally, the unique and precise distribution of amino acids in the PcPAL's substrate binding pocket determines the stereoselectivity of the asymmetric amination reaction. The amino acids around the benzene ring are mainly hydrophobic, while the ones around the carboxyl group are more hydrophilic (Fig. 6c). Guided by the unique distribution of amino acids within the pocket, the  $\beta$ -methyl group of  $\beta$ -MeCA demonstrates a preferential

orientation towards the hydrophobic region. This will be advantageous for the enzyme's adaptation and stereoselectivity towards  $\beta$ -MeCA analogs. Importantly, due to the fact that the critical MIO group is positioned on the *si-si* face of the substrate plane, the chiral center formed after the amino group attacking  $\alpha$ C can only be 2S. Additionally, the crucial Tyr110 is located on the *Re-Re* face of the substrate plane. In such case, when the  $\beta$ C is protonated, the chiral center formed can only be 3R. Overall, the distinctive amino acid distribution provides essential molecular foundations for correctly accommodating  $\beta$ -MeCA and facilitating a highly stereoselective asymmetric amination reaction.

## Discussion

In summary, through computational analysis and a rational protein engineering approach, we successfully elucidated and addressed the long-standing challenge of PAL's inability to aminate  $\beta$ -substituted CA analogs. Our method for the direct asymmetric amination of  $\beta$ -substituted CA analogs to generate  $\beta$ -branched aromatic  $\alpha$ -amino acids possesses several advantages, such as readily available substrates, excellent atom economy, no need for cofactors, and high regio- and stereo selectivity. These results represent a significant advancement in the field of PALs, contributing not only to the expansion of their applications but also providing an approach for the synthesis of  $\beta$ -branched aromatic  $\alpha$ -amino acids.



**Fig. 6 | Molecular basis explaining the amination of the  $\beta$ -MeCA substrate by the mutants.** **a** Gibbs energy diagram of the cMD simulations projected onto the distance from the amino N of MIO to the  $\alpha$ C of the substrates and the angle between MIO and the double bond of the substrates within the ( $\beta$ -MeCA)/(PcPAL-L256V-I460V) mutant. **b** the pocket volume detected during the US simulations of the ( $\beta$ -MeCA)/(PcPAL-L256V-I460V) complex. **c** Optimized pre-reaction states of  $\beta$ -MeCA in the PcPAL-L256V-I460V mutant using the Quantum Mechanics/Molecular

Mechanics (QM/MM) methodology. **d**, **e**, **f** Gibbs energy diagram of the cMD simulations projected at the distance from  $\beta$ C of  $\beta$ -MeCA to amino acid 256 and the benzene ring center and amino acid 460 within the ( $\beta$ -MeCA)/(PcPAL-L256V-I460V) complex, (CA)/(PcPAL-WT) complex, and ( $\beta$ -MeCA)/(PcPAL-WT) complex respectively. Source data are provided in Source Data 3 to Source Data 5. CA cinnamic acid,  $\beta$ -MeCA  $\beta$ -methyl cinnamic acid, MIO 3,5-dihydro-5-methylidene-4H-imidazol-4-one.

## Methods

### Construction of PcPAL mutant libraries

The plasmid PcPAL-pETs<sup>58</sup> was used as template and site-directed mutagenesis using the QuikChange<sup>®</sup> method was used for the creation of mutants. Primers were designed using the Agilent (USA) web tool QuikChange<sup>®</sup> Primer Design and provided in Table S1. The PCR protocol employed the Pfu Plus polymerase from Roboklon (Germany) using the following steps: initial denaturation at 98 °C for 30 s, followed by 25 cycles of denaturation at 98 °C for 10 s, annealing at 68 °C for 10 s, and extension at 72 °C for 3.5 min. A final extension step was performed at 72 °C for 10 min. The PCR products were then digested using DpnI and used for the transformation of *E. coli* (Top 10) chemical competent cells through heat shock. The introduced mutations were confirmed by sequencing.

### Cloning, expression and purification of PcPAL and associated proteins

The plasmids harboring the genes encoding the PcPAL variants were inserted into the *E. coli* BL21(DE3) strain by heat shock transformation. The transformed cells were incubated overnight in a culture tube containing 4 mL LB medium supplemented with kanamycin (40 mg/mL) at 37 °C with constant shaking at 180 rpm. The overnight cultures were then transferred to a 500 mL flask containing 100 mL of TB medium supplemented with kanamycin (50 mg/mL) and incubated at 37 °C with constant shaking at 180 rpm. Once the cultures reached an OD<sub>600</sub> of 0.6, the temperature was lowered to 28 °C, and the cultures were incubated for an additional 20 h with constant shaking at 180 rpm. Cells were centrifuged at 10,000 × *g* for 30 min, and the cell pellet was resuspended in 20 mL lysis buffer (25 mM HEPES, pH 7.5, 300 mM NaCl, 5 mM imidazole, 10% glycerol) and lysed using ultrasonication. The insoluble debris was removed by centrifugation at 10,000 × *g* and 4 °C for 1 h. The protein supernatant was incubated with 1 mL of Ni-NTA Sepharose resin for 0.5 h with slow and constant rotation at 4 °C. The protein-resin mixtures were then loaded onto a gravity flow column, and the proteins were eluted using increasing concentrations of imidazole (25 mM, 50 mM, 300 mM) in Buffer A (25 mM HEPES, pH = 7.5, 300 mM NaCl, 10% glycerol). The purified proteins were further subjected to desalting using PD-10 desalting columns and buffer B (25 mM HEPES, pH = 7.5, 50 mM NaCl, 10% glycerol), followed by concentration using an Amicon Ultra-4 (GE Healthcare) centrifugal filter. The purity of the proteins was assessed using 12% acrylamide SDS-PAGE (Fig. S11).

### General method for screening in vitro amination reactions

A purified PcPAL variant (100 μg) was added to a solution of the substrate (1 mM, 500 μL) in NH<sub>4</sub>OH buffer (5 M, pH = 10, adjust by H<sub>2</sub>SO<sub>4</sub>). The mixture was incubated at 30 °C, 500 rpm for 24 h. After that, the reaction was stopped with 500 μL 10% H<sub>2</sub>SO<sub>4</sub> (v/v), thoroughly shaken and centrifuged (12,000 × *g*, 5 min). The supernatant was transferred to a filter vial and used directly for UPLC analysis. UPLC analysis of reactions were performed on a Luna<sup>®</sup> Omega column (C18, 5 μm, 150 × 4.6 mm, Phenomenex) at a flow rate of 1 mL min<sup>-1</sup> and detected at 210 nm over a 10 min gradient program with water (eluent A) and acetonitrile (eluent B): *T* = 0 min, 5% B; *T* = 4 min, 60% B; *T* = 6 min, 5% B; *T* = 10 min, 5% B.

### Biotransformations and product purification

Recombinant strains were inoculated in TB medium (2 L) by growth to an OD<sub>600</sub> of 0.8–1.0 at 37 °C. After expression at 28 °C and 220 rpm for 20 h, the cultures were harvested by centrifugation at 4000 × *g* at 4 °C and washed twice with ddH<sub>2</sub>O. After that, the cells were resuspended in 30 mL 5 M NH<sub>4</sub>OH buffer (pH = 10) to reach an OD<sub>600</sub> of 30, into which different substrates (1 g/L) and 1% Triton X-100 were added. After 18 h incubation at 28 °C, the reaction mixture was acidified slowly with aq. H<sub>2</sub>SO<sub>4</sub> (50% v/v) at 0 °C, centrifuged (8000 *g*, 10 min, 4 °C) and applied

to a column containing washed Dowex 50 W resin for desalination. The resin was washed with water until the eluate was neutral, then the product was eluted with aq. NH<sub>4</sub>OH (15% v/v). The fractions containing the product were pooled and concentrated by evaporating in a centrifugal evaporator. After that, the crude products were further separated by semi-preparative HPLC (SHIMADZU, SPD-20A 230 V) equipped with a Phenomenex Luna C18 reverse phase column (5 μm, 250 × 10 mm). The purification process involved isocratic elution with a mixture of 12% acetonitrile and water containing 0.1% formic acid for a duration of 30 min.

### NMR spectroscopy

The NMR spectra were recorded on a Bruker Avance III spectrometer at a <sup>1</sup>H frequency of 300 MHz or 400 MHz or 600 MHz. Samples (varying from 1 to 15 mg) were dissolved in 400 μL MeOH-*d*<sub>4</sub> or D<sub>2</sub>O (Cambridge Isotope) and all spectra were recorded at 25 °C (298 K). NMR experiments were processed with the Bruker Topspin program and analyzed with the MestReNova software (Version 6.1.0-6224).

### X-ray crystallographic analysis

Compounds **4a**, **5a**, **8a**, and **9a** were obtained as colorless crystals by recrystallization from a water/methanol system using the vapor diffusion method. The single-crystal X-ray diffraction data were collected using the XtaLAB Synergy Custom and XtaLAB PRO MM007HF diffractometers with Cu-Kα radiation (λ = 1.542 Å). The structures were solved by the intrinsic phasing solution method. The models were refined with ShelXL-2018/3 using least-squares minimization. The structures were visualized by Olex2 (version 1.5).

### Preparation of the MIO-group containing tetramer structures of PcPAL and docking analysis

AlphaFold 2 was used to predict the tetrameric form 3D structures of PcPAL and its related mutants. The MIO structure within them was derived from a resolved crystal structure. We extracted the MIO-group from the crystal structure (PDBID: 6RGS) and re-edited it in PyMOL (version 1.8.0.2) to aminate the MIO-group. Then this non-standard amino acid was checked and capped, and structural optimization and vibrational frequency checks were conducted using ORCA 5.0.4<sup>60</sup> at the r2SCAN-3c<sup>61</sup> theory level, using the optimized conformation in CPCM (ε = 4) implicit solvent as the initial conformation for MIO to ensure the accuracy of its geometry and electronic structure. Subsequently, RESP2 charge fitting was carried out using Multiwfn 3.8\_dev<sup>62</sup>, and it was parameterized using the Amber/GAFF2 atom types through the Antechamber tool. Finally, the missing parameters were completed using parmchk2, and the results were visually inspected and subjected to dynamic testing. After that, the docking experiments were carried out on Autodock Tools (version 1.5.6), with the grid size restricted to a 10 × 10 × 10 Å box centered on the midpoint between the MIO residue and Tyr110 of the tyrosine loop. PyMOL (version 1.8.0.2) was used to visualize 3D structure analysis and comparison.

### Conventional molecular dynamics (cMD)

Protonation prediction for the protein was performed using H++<sup>63</sup>, and the initial structure underwent 2,500 steps of steepest descent and 2,500 steps of conjugate gradient optimization using the Amber22 program<sup>64</sup>. The protein residues were parameterized using the AMBER14SB force field. The system was solvated using a periodic dodecahedron with a 10 Å buffer between solute atoms and the box edges, and appropriate amounts of Na<sup>+</sup> and Cl<sup>-</sup> ions were added to neutralize the system. Subsequently, a 2 ns NVT heating simulation was conducted to equilibrate the molecular system. The NVT simulation employed the Langevin thermostat to gradually raise the temperature to 300 K, with hydrogen bonds constrained using the SHAKE algorithm and a time step of 2.0 fs. This was followed by a 5 ns NPT simulation to further relax the system, also using the Langevin thermostat

to maintain the temperature at 300 K, with hydrogen bonds constrained using SHAKE and a time step of 2.0 fs. Long-range electrostatics were calculated using the particle mesh Ewald method with a non-bonded interaction cutoff of 10 Å. The MD simulations were run for 100 ns with a 2 fs time step, saving coordinates every 2 ps to the trajectory file. The simulation trajectories were analyzed using the cpptraj tool.

### Umbrella sampling simulations

The initial coordinates for umbrella sampling were derived from equilibrated conformations obtained during the MD phase, followed by three independent 10 ns NPT simulations to serve as the starting points for the three parallel sets of umbrella sampling. Umbrella sampling (US) utilizes a reaction coordinate describing the distance between two heavy atoms involved in the nucleophilic addition reaction (the N atom of MIO to the  $\alpha$ C of CA or  $\beta$ -MeCA), aiming to reveal differences between the WT and mutant in the pre-reactive state. We employed 38 simulation windows with a step size of 0.75 Å between windows. The system was constrained to other values by introducing an artificial harmonic bias potential with a constraint force of 50 kcal/mol. Each sampling window underwent 5,000 steps of energy minimization and 50 ps of pre-equilibration, followed by a 2 ns NPT simulation. Throughout the umbrella sampling process, a time step of 1.0 fs was used, and other simulation parameters were consistent with pre-equilibration. Three repetitions were carried out starting from different equilibrated points. Weighted Histogram Analysis Method<sup>65</sup> was employed to compute the potential of mean force (PMF).

### Analysis of the USMD results

Cpptraj and PLUMED<sup>66</sup> were used to analyze the umbrella sampling trajectories to extract atomic distance and angle information. The protein pocket volumes and substrate volumes during umbrella sampling were calculated using Fpocket (<https://github.com/Disngine/fpocket>, Accessed on June 2, 2024) and MoloVol<sup>67</sup>, respectively. For protein pocket volume calculation, the `min_alpha_size` was set to 2 Å, and the `max_alpha_size` was set to the default value. For molecular volume calculation, the `rprobe` was set to 1.2 Å, and the grid was set to 0.1 Å. The volume difference calculation formula is as follows:

$$\Delta V = V_{\text{protein}} - V_{\text{substrate}}$$

After obtaining the results of the umbrella sampling, the coordinates of the corresponding window at the lowest energy reaction coordinate were read in. A 2 ns classical molecular dynamics (cMD) simulation was performed to sample the pre-reaction state. The structures from the pre-reaction state sampling were clustered using cpptraj, and potential errors in the clustering process were minimized using 2,500 steps of steepest descent and 1,000 steps of conjugate gradient minimization.

In the pre-reaction state sampling, we first read the coordinates corresponding to the lowest energy point from the PMF curve obtained through umbrella sampling. Then, we performed a 2 ns cMD simulation starting from these coordinates, recording snapshots every 1 ps. Finally, we used cpptraj to perform clustering analysis on these snapshots, setting a total of 5 conformational clusters.

### Quantum mechanics/molecular mechanics (QM/MM)

Based on clustering results, we optimized the conformations using QM/MM. This approach allows us to incorporate electronic correlation effects within the system at a higher quantum mechanics (QM) level, while describing interactions using molecular mechanics (MM) methods, aiming to achieve more accurate and reliable interaction descriptions. The QM/MM interface was executed using the ASH program (<https://github.com/RagnarB83/ash>, Accessed on June 10, 2024.). The QM region including the substrate, MIO, ARG354, TYR110,

TYR351, ASN260, and ASN384, was treated with the BLYP function<sup>68</sup>, def2-SVP basis set<sup>69</sup>, def2/J auxiliary basis set<sup>70</sup>, and D3BJ dispersion correction<sup>71</sup> level using ORCA 4.0.5. The MM region was described using OpenMM 8.0<sup>72</sup> and Amber ffi4SB force field. To handle electrostatic interactions between QM and MM layers, the electrostatic embedding technique was applied, and the link atom method resolved boundary issues between QM and MM layers. The active region is defined as within a 10 Å radius around the amino N of MIO. This method allows us to simulate the interactions of the wild-type and mutant in the pre-reaction state more precisely.

### Determination of kinetic constants

The kinetic constants were measured in triplicate at 30 °C using 96-well UV-microplates. The reactions were carried out in 50 mM PBS buffer (pH 8.0) and the mutant enzyme (PcPAL-L256V-I460V) was maintained at a constant concentration of 0.6  $\mu$ M, with eight different substrate concentrations ranging from 0.015625 to 1.5 mM. Kinetic measurements were performed using UV spectroscopy to monitor the production of  $\beta$ -MeCA over a 10 min period at a wavelength of 260 nm. The kinetic constants were then calculated from Michaelis-Menten plots via non-linear regression analysis.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

Crystallographic data for the structures reported in this article have been deposited at the Cambridge Crystallographic Data Centre, under deposition nos. CCDC 2293061 (**4a**), 2293062 (**8a**), 2293063 (**5a**) and 2313111 (**9a**). Copies of the data can be obtained free of charge via <https://www.ccdc.cam.ac.uk/structures/>. The crystal structure of PcPAL (PDB: 6HQF) is available at <https://www.rcsb.org/structure/6HQF>. All other characterization data and detailed experimental procedures are available in the supplementary information file. Source data are provided together with this publication.

### References

- Huang, Y. T. et al. In vitro characterization of enzymes involved in the synthesis of nonproteinogenic residue (2S, 3S)- $\beta$ -methylphenylalanine in glycopeptide antibiotic mannopeptimycin. *Chem-BioChem* **10**, 2480–2487 (2009).
- Andres, N., Wolf, H. & Zähler, H. Hormaomycin, a new peptide lactone antibiotic effective in inducing cytodifferentiation and antibiotic biosynthesis in some *Streptomyces* species. *Z. Naturforsch. C. J. Biosci.* **45**, 851–855 (1990).
- Höfer, I. et al. Insights into the biosynthesis of hormaomycin, an exceptionally complex bacterial signaling metabolite. *Chem. Biol.* **18**, 381–391 (2011).
- Shimamura, H. et al. Structure determination and total synthesis of bottromycin A2: a potent antibiotic against MRSA and VRE. *Angew. Chem. Int. Ed.* **121**, 932–935 (2009).
- Yamada, T. et al. Synthesis and evaluation of antibacterial activity of bottromycins. *J. Org. Chem.* **83**, 7135–7149 (2018).
- Cutler, H. G. et al. Chaetoglobosin K: a new plant growth inhibitor and toxin from *Diplodia macrospora*. *J. Agric. Food Chem.* **28**, 139–142 (1980).
- Tömböly, C. et al. Structure-activity study on the Phe side chain arrangement of endomorphins using conformationally constrained analogues. *J. Med. Chem.* **47**, 735–743 (2004).
- Hruby, V. J. Peptide science: exploring the use of chemical principles and interdisciplinary collaboration for understanding life processes. *J. Med. Chem.* **46**, 4215–4231 (2003).
- Blaskovich, M. A. Unusual amino acids in medicinal chemistry. *J. Med. Chem.* **59**, 10807–10836 (2016).

- Almhjell, P. J., Boville, C. E. & Arnold, F. H. Engineering enzymes for noncanonical amino acid synthesis. *Chem. Soc. Rev.* **47**, 8980–8997 (2018).
- Burk, M. J., Gross, M. F. & Martinez, J. P. Asymmetric catalytic synthesis of  $\beta$ -branched amino acids via highly enantioselective hydrogenation of  $\alpha$ -enamides. *J. Am. Chem. Soc.* **117**, 9375–9376 (1995).
- Zhang, S.-Y., Li, Q., He, G., Nack, W. A. & Chen, G. Stereoselective synthesis of  $\beta$ -alkylated  $\alpha$ -amino acids via palladium-catalyzed alkylation of unactivated methylene C ( $sp^3$ )-H bonds with primary alkyl halides. *J. Am. Chem. Soc.* **135**, 12135–12141 (2013).
- Chen, G. et al. Ligand-enabled  $\beta$ -C-H arylation of  $\alpha$ -amino acids using a simple and practical auxiliary. *J. Am. Chem. Soc.* **137**, 3338–3351 (2015).
- Bornscheuer, U. T. et al. Engineering the third wave of biocatalysis. *Nature* **485**, 185–194 (2012).
- Wu, S., Snajdrova, R., Moore, J. C., Baldenius, K. & Bornscheuer, U. T. Biocatalysis: enzymatic synthesis for industrial applications. *Angew. Chem. Int. Ed.* **60**, 88–119 (2021).
- Xue, Y. P., Cao, C. H. & Zheng, Y. G. Enzymatic asymmetric synthesis of chiral amino acids. *Chem. Soc. Rev.* **47**, 1516–1561 (2018).
- Buller, R. et al. From nature to industry: harnessing enzymes for biocatalysis. *Science* **382**, eadh8615 (2023).
- Boville, C. E. et al. Engineered biosynthesis of  $\beta$ -alkyl tryptophan analogues. *Angew. Chem. Int. Ed.* **130**, 14980–14984 (2018).
- Raj, H. et al. Engineering methylaspartate ammonia lyase for the asymmetric synthesis of unnatural amino acids. *Nat. Chem.* **4**, 478–484 (2012).
- Roff, G. J., Lloyd, R. C. & Turner, N. J. A versatile chemo-enzymatic route to enantiomerically pure  $\beta$ -branched  $\alpha$ -amino acids. *J. Am. Chem. Soc.* **126**, 4098–4099 (2004).
- Xian, M., Alaux, S., Sagot, E. & Gefflaut, T. Chemoenzymatic synthesis of glutamic acid analogues: substrate specificity and synthetic applications of branched chain aminotransferase from *Escherichia coli*. *J. Org. Chem.* **72**, 7560–7566 (2007).
- Gu, X. et al. Synthesis of the constrained glutamate analogues (2S,1'R,2'R) and (2S,1'S,2'S)-2-(2'-carboxycyclobutyl)glycines L-CBG-II and L-CBG-I by enzymatic transamination. *Tetrah. Lett.* **47**, 193–196 (2006).
- Liao, C. & Seebeck, F. P. Asymmetric  $\beta$ -methylation of L- and D- $\alpha$ -amino acids by a self-contained enzyme cascade. *Angew. Chem. Int. Ed.* **59**, 7184–7187 (2020).
- Li, F., Yang, L. C., Zhang, J., Chen, J. S. & Renata, H. Stereoselective synthesis of  $\beta$ -branched aromatic  $\alpha$ -amino acids by biocatalytic dynamic kinetic resolution. *Angew. Chem. Int. Ed.* **60**, 17680–17685 (2021).
- Dunham, N. P. et al. Transaminase-catalyzed synthesis of  $\beta$ -branched noncanonical amino acids driven by a lysine amine donor. *J. Am. Chem. Soc.* **146**, 16306–16313 (2024).
- Wang, T.-C. et al. Stereoselective amino acid synthesis by photo-biocatalytic oxidative coupling. *Nature* **629**, 98–104 (2024).
- 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 (2018).
- Turner, N. J. Ammonia lyases and aminomutases as biocatalysts for the synthesis of  $\alpha$ -amino and  $\beta$ -amino acids. *Curr. Opin. Chem. Biol.* **15**, 234–240 (2011).
- Barros, J. & Dixon, R. A. Plant phenylalanine/tyrosine ammonia-lyases. *Trends Plant Sci.* **25**, 66–79 (2020).
- MacDonald, M. J. & D' Cunha, G. B. A modern view of phenylalanine ammonia lyase. *Biochem. Cell Biol.* **85**, 273–282 (2007).
- Nelson, R. P. Immobilized Microbial Cells. U.S. Patent US 3,957,580 (1976).
- de Lange, B. et al. Asymmetric synthesis of (S)-2-indolinecarboxylic acid by combining biocatalysis and homogeneous catalysis. *ChemCatChem* **3**, 289–292 (2011).
- Hardegger, L. A. et al. Toward a scalable synthesis and process for EMA401, Part III: using an engineered phenylalanine ammonia lyase enzyme to synthesize a non-natural phenylalanine derivative. *Org. Process Res. Dev.* **24**, 1763–1771 (2020).
- Weise, N. J. et al. Intensified biocatalytic production of enantiomerically pure halophenylalanines from acrylic acids using ammonium carbamate as the ammonia source. *Catal. Sci. Technol.* **6**, 4086–4089 (2016).
- Zhu, L. et al. Cloning, expression and characterization of phenylalanine ammonia-lyase from *Rhodotorula glutinis*. *Biotechnol. Lett.* **35**, 751–756 (2013).
- Renard, G., Guilleux, J.-C., Bore, C., Malta-Valette, V. & Lerner, D. Synthesis of L-phenylalanine analogs by *Rhodotorula glutinis*. Bio-conversion of cinnamic acids derivatives. *J. Biotechnol.* **14**, 673–678 (1992).
- Hodgins, D. S. Yeast phenylalanine ammonia-lyase: purification, properties, and the identification of catalytically essential dehydroalanine. *J. Biol. Chem.* **246**, 2977–2985 (1971).
- Kawasaki Watanabe, S., Hernandez-Velazco, G., Iturbe-Chinas, F. & Lopez-Munguia, A. Phenylalanine ammonia lyase from *Sporidiobolus parvoseus* and *Rhodospiridium toruloideis*: application for phenylalanine and tyrosine deamination. *World. J. Microb. Biot.* **8**, 406–410 (1992).
- Lois, R., Dietrich, A., Hahlbrock, K. & Schulz, W. A phenylalanine ammonia-lyase gene from parsley: structure, regulation and identification of elicitor and light responsive cis-acting elements. *EMBO J.* **8**, 1641–1648 (1989).
- Moffitt, M. C. et al. Discovery of two cyanobacterial phenylalanine ammonia lyases: kinetic and structural characterization. *Biochemistry* **46**, 1004–1012 (2007).
- Lovelock, S. L. & Turner, N. J. Bacterial *Anabaena variabilis* phenylalanine ammonia lyase: a biocatalyst with broad substrate specificity. *Bioorg. Med. Chem.* **22**, 5555–5557 (2014).
- Benche, L. C. et al. Expanding the substrate scope of phenylalanine ammonia-lyase from *Petroselinum crispum* towards styrylalanines. *Org. Biomol. Chem.* **15**, 3717–3727 (2017).
- Filip, A. et al. Tailored mutants of phenylalanine ammonia-lyase from *Petroselinum crispum* for the synthesis of bulky L- and D-arylanines. *ChemCatChem* **10**, 2627–2633 (2018).
- Tork, S. D. et al. The production of L- and D-phenylalanines using engineered phenylalanine ammonia lyases from *Petroselinum crispum*. *Sci. Rep.* **9**, 20123 (2019).
- Nagy, E. Z. et al. Mapping the hydrophobic substrate binding site of phenylalanine ammonia-lyase from *Petroselinum crispum*. *ACS Catal* **9**, 8825–8834 (2019).
- Rowles, I. et al. Engineering of phenylalanine ammonia lyase from *Rhodotorula graminis* for the enhanced synthesis of unnatural L-amino acids. *Tetrahedron* **72**, 7343–7347 (2016).
- Ahmed, S. T., Parmeggiani, F., Weise, N. J., Flitsch, S. L. & Turner, N. J. Engineered ammonia lyases for the production of challenging electron-rich L-phenylalanines. *ACS Catal* **8**, 3129–3132 (2018).
- Bartsch, S. & Bornscheuer, U. T. Mutational analysis of phenylalanine ammonia lyase to improve reactions rates for various substrates. *Protein Eng. Des. Sel.* **23**, 929–933 (2010).
- Easton, C. J. & Hutton, C. A. Synthesis of each stereoisomer of [3-2 H 1] phenylalanine and evaluation of the stereochemical course of the reaction of (R)-phenylalanine with (S)-phenylalanine ammonia-lyase. *J. Chem. Soc. Perkin Trans 1* **1994**, 3545–3548 (1994).
- Lovelock, S. L., Lloyd, R. C. & Turner, N. J. Phenylalanine ammonia lyase catalyzed synthesis of amino acids by an MIO-cofactor independent pathway. *Angew. Chem. Int. Ed.* **126**, 4740–4744 (2014).
- Szymanski, W. et al. Phenylalanine aminomutase-catalyzed addition of ammonia to substituted cinnamic acids: a route to enantiopure  $\alpha$ - and  $\beta$ -amino acids. *J. Org. Chem.* **74**, 9152–9157 (2009).

52. Bata, Z. et al. Substrate tunnel engineering aided by X-ray crystallography and functional dynamics swaps the function of MIO-enzymes. *ACS Catal* **11**, 4538–4549 (2021).
53. Jiang, Y. et al. Substrate positioning dynamics involves a non-electrostatic component to mediate catalysis. *J. Phys. Chem. Lett.* **14**, 11480–11489 (2023).
54. Mehmood, R., Qi, H. W., Steeves, A. H. & Kulik, H. J. The protein's role in substrate positioning and reactivity for biosynthetic enzyme complexes: the case of SyrB2/SyrB1. *ACS Catal.* **9**, 4930–4943 (2019).
55. Dunitz, J., Lehn, J. & Wipff, G. Stereochemistry of reaction paths at carbonyl centres. *Tetrahedron* **30**, 1563–1572 (1974).
56. Tork, S. D., Nagy, E. Z. A., Tomoiagă, R. B. & Bencze, L. C. Engineered, scalable production of optically pure L-phenylalanines using phenylalanine ammonia-lyase from *Arabidopsis thaliana*. *J. Org. Chem.* **88**, 852–862 (2022).
57. Tork, S. D. et al. Towards a general approach for tailoring the hydrophobic binding site of phenylalanine ammonia-lyases. *Sci. Rep.* **12**, 10606 (2022).
58. Wu, S. et al. A growth selection system for the directed evolution of amine-forming or converting enzymes. *Nat. Commun.* **13**, 7458 (2022).
59. Xiao, G. et al. Highly enantioselective Ni-catalyzed asymmetric hydrogenation of  $\beta,\beta$ -disubstituted acrylic acids. *Org. Chem. Front.* **9**, 4472–4477 (2022).
60. Neese, F. Software update: the ORCA program system—version 5.0. *Wires. Comput. Mol. Sci.* **12**, e1606 (2022).
61. Grimme, S., Hansen, A. & Ehlert, S. Mewes, J.-M. r2SCAN-3c: a “Swiss army knife” composite electronic-structure method. *J. Chem. Phys.* **154**, 064103 (2021).
62. Lu, T. & Chen, F. Multiwfn: a multifunctional wavefunction analyzer. *J. Comput. Chem.* **33**, 580–592 (2012).
63. Anandakrishnan, R., Aguilar, B. & Onufriev, A. V. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **40**, W537–W541 (2012).
64. Case, D. A. et al. AmberTools. *J. Chem. Inf. Model.* **63**, 6183–6191 (2023).
65. Grossfield, A. WHAM: the weighted histogram analysis method, version 2.0. 9 (accessed 15 November 2013). Available at: <http://membrane.urmc.rochester.edu/content/wham/> (2013).
66. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED 2: new feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613 (2014).
67. Maglic, J. B. & Lavendomme, R. MoloVol: an easy-to-use program for analyzing cavities, volumes and surface areas of chemical structures. *J. Appl. Crystallogr.* **55**, 1033–1044 (2022).
68. Miehlich, B., Savin, A., Stoll, H. & Preuss, H. Results obtained with the correlation energy density functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **157**, 200–206 (1989).
69. Weigend, F. & Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **7**, 3297–3305 (2005).
70. Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **8**, 1057–1065 (2006).
71. Grimme, S., Ehrlich, S. & Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **32**, 1456–1465 (2011).
72. Eastman, P. et al. OpenMM 8: molecular dynamics simulation with machine learning potentials. *J. Phys. Chem. B* **128**, 109–116 (2023).

## Acknowledgements

We thank Prof. Dr. Andreas Link and Maria Hühner for access to the 400 MHz NMR spectrometer and Ina Menyes for her support with the HPLC analytics. This work was supported by the European Union's Horizon 2020 Research and Innovation Program under Grant Agreement No. 814650 for the project SynBio4Flav, and supported by the start funding of Huazhong University of Science and Technology to Guifa Zhai (3034514105). We thank Siwen Yuan and Feng Li for specific rotation data collection.

## Author contributions

C. Sun, G. Zhai, and U.T. Bornscheuer designed the project. C. Sun, G. Zhai, G. Lu, B. Chen, G. Li and Y. Brack performed the experiment; C. Sun, G. Zhai, G. Lu, Y. Wu, G. Li, B. Chen, Y. Brack, D. Yi, Y. Ao, S. Wu, R. Wei, Y. Sun and U.T. Bornscheuer analyzed and discussed the results. C. Sun, G. Zhai, G. Lu and U.T. Bornscheuer wrote the manuscript with inputs from all authors.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-52613-x>.

**Correspondence** and requests for materials should be addressed to Chenghai Sun, Guifa Zhai or Uwe T. Bornscheuer.

**Peer review information** *Nature Communications* thanks Laszlo Csaba Bencze, Yang Yang, Zhongyue Yang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024