Computational redesign of enzymes for regio- and enantioselective hydroamination

Li, Ruifeng; Wijma, Hein J; Song, Lu; Cui, Yinglu; Otzen, Marleen; Tian, Yu'e; Du, Jiawei; Li, Tao; Niu, Dingding; Chen, Yanchun

Published in:
Nature Chemical Biology

DOI:
10.1038/s41589-018-0053-0

Publication date:
2018

ARTICLES
https://doi.org/10.1038/s41589-018-0053-0
nature chemical biology

Computational redesign of enzymes for regio- and enantioselective hydroamination

Ruifeng Li[1,2,3,5], Hein J. Wijma[4,5], Lu Song[1,5], Yinglu Cui[1,2], Marleen Otzen[4], Yu'e Tian[1], Jiawei Du[1,3], Tao Li[1,3], Dingding Niu[1], Yanchun Chen[1,3], Jing Feng[1,3], Jian Han[1], Hao Chen[1], Yong Tao[1], Dick B. Janssen[4*] and Bian Wu[1*]

Introduction of innovative biocatalytic processes offers great promise for applications in green chemistry. However, owing to limited catalytic performance, the enzymes harvested from nature's biodiversity often need to be improved for their desired functions by time-consuming iterative rounds of laboratory evolution. Here we describe the use of structure-based computational enzyme design to convert *Bacillus* sp. YM55-1 aspartase, an enzyme with a very narrow substrate scope, to a set of complementary hydroamination biocatalysts. The redesigned enzymes catalyze asymmetric addition of ammonia to substituted acrylates, affording enantiopure aliphatic, polar and aromatic β-amino acids that are valuable building blocks for the synthesis of pharmaceuticals and bioactive compounds. Without a requirement for further optimization by laboratory evolution, the redesigned enzymes exhibit substrate tolerance up to a concentration of 300 g/L, conversion up to 99%, β-regioselectivity >99% and product enantiomeric excess >99%. The results highlight the use of computational design to rapidly adapt an enzyme to industrially viable reactions.

Innovations in applied biocatalysis and synthetic biology often require the availability of enzymes that surpass the performance of nature's protein catalysts in terms of stability, activity and selectivity[1]. Directed evolution using iterative rounds of mutagenesis and screening has emerged as an attractive and powerful strategy for adapting enzyme properties to specific needs[2,3]. Major successes include converting aminotransferases to practically useful biocatalysts[4,5], removing the carboxylate requirement in the case of amine dehydrogenases[6,7], tuning P450 enzymes for stereoselective hydroxylation of steroids[8], and evolving cytochrome *c* to catalyze carbon–silicon bond formation[9]. However, laboratory evolution usually requires high-throughput expression and experimental evaluation of large mutant libraries[10]. This can make the process time-consuming or even impractical in cases of complex expression and/or assay protocols. Consequently, there is a strong interest in improving library quality by focusing on positions and substitutions that are promising. The design of such 'smart libraries' can be based on the abundance of residues at corresponding positions among a group of homologous sequences or on inspection of 3D structures[11–18]. The final result of a directed evolution campaign usually consists of a few variants with a restricted set of mutations, which often can be explained by careful rational or computational comparison of wild-type and mutant structures[19]. This raises the question of whether physics-based computational design methods[20], which have been developed for de novo protein design and are capable of discovering new structures that stabilize predefined reactive enzyme–substrate complexes while exploring a huge sequence and conformational space, could be used to focus search space in directed evolution and quickly produce enzyme variants with a desired activity by allowing large jumps in function by multiple mutations.

We addressed this possibility by exploring enzymes for the asymmetric hydroamination of carbon–carbon double bonds, one of the top aspirational reactions in synthetic chemistry[21]. Direct asymmetric addition of ammonia to α,β-unsaturated carboxylic acids would be an attractive and atom-economic route for producing chiral β-amino acids, which exhibit unique bioactivities and are frequently used as building blocks of bioactive products, such as antifungal agents, antitumor agents, antibiotics, antiviral agents and insecticides (Supplementary Fig. 1)[22]. Consequently, various chemical processes for asymmetric C–N bond formation via conjugate addition have been developed during the past decade, using either metal catalysis or organocatalysis[23]. These methods rely on modified carbonyl compounds as acceptors, utilize amines instead of ammonia as the nitrogen nucleophiles, and require protective groups. Therefore, problems such as the use of expensive catalysts or chiral auxiliaries, the need for protection and deprotection steps, and the use of harsh reaction conditions have made these processes expensive and increasingly unsustainable[23]. Biocatalytic routes would offer environmentally friendly options[24–26], especially for hydroamination of acrylates by lyases, but ammonia lyases lack the required activity and/or regioselectivity[27]. Attempts to shift the α-regiochemical preference of arylalanine ammonia lyases to β-selectivity have met with some success, but the enzymes act only on cinnamate derivatives and even the best mutants have modest regioselectivity and low activity[28]. Thus, the development of efficient, sustainable and scalable processes toward enantiopure β-amino acids remains a formidable challenge[23].

In search of an enzyme that could serve as the starting point for engineering an effective catalyst for β-hydroamination of α,β-unsaturated carboxylic acids, we examined aspartase from *Bacillus* sp. YM55-1 (AspB, UniProt Q9LCC6), which catalyzes the reversible

[1]CAS Key Laboratory of Microbial Physiological and Metabolic Engineering, State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China. [2]State Key Laboratory of Transducer Technology, Chinese Academy of Sciences, Beijing, China. [3]University of Chinese Academy of Sciences, Beijing, China. [4]Department of Biochemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands. [5]These authors contributed equally: Ruifeng Li, Hein J., Lu Song. *e-mail: d.b.janssen@rug.nl; wub@im.ac.cn

deamination of aspartate to yield fumarate and ammonia[29]. Unlike methylaspartate ammonia lyase, which requires cations, and arylalanine ammonia lyases, which are dependent on the 4-methylideneimidazol-5-one moiety[30], AspB is a robust and cofactor-independent enzyme[31]. Its natural substrate, α-aspartate, carries a secondary carboxylate functionality, therefore we hypothesized that AspB could be transformed into more generic ammonia lyases active with diverse substituted acrylates by reshaping the enzyme's α-carboxylate binding pocket to accommodate other substituent groups while maintaining the functional interactions in the β-carboxylate binding pocket.

As an enzyme from primary metabolism, aspartase is one of the most specific enzymes known[31]. The enzyme lacks the promiscuous activity toward other substrates that is typically regarded as highly beneficial for directed evolution approaches[32]. Since the evolvability of specialist enzymes is more troublesome than that of generalist proteins with relaxed selectivity[33], it is not surprising that discovery of mutants by directed evolution of AspB required large libraries, with screening of thousands of variants, and still delivered only two mutants. A variant was found to exhibit minor activity toward β-asparagine (<0.06 U/mg) after screening 3,000 mutants[34], and another mutant showing activity toward crotonic acid ($k_{cat} = 0.09\ s^{-1}$) was identified by screening 300,000 clones[35]. Our initial attempts to obtain AspB variants by directed evolution with screening of single-site and pair-wise saturation libraries also did not yield any mutants with detectable activity. Selecting more residues for randomization exponentially increases the risk of incorporating lethal mutations and thus would further increase the number of variants that must be screened to discover desired improvements.

The recent development of algorithms for computational enzyme design[20] and our success in computational redesign of a solvent-compatible enzyme with relaxed selectivity (FRESCO approach)[36] and enantioselective enzymes[37] (CASCO approach) prompted us to tackle the problem of endowing this highly specialized enzyme with non-natural activities by a computation-aided strategy. By dissecting the enzyme–substrate complex into interactions relevant for substrate binding and for catalysis and through computationally optimizing the active site environment for novel substrates using mechanism-based geometric criteria and Rosetta Enzyme Design for energy calculations, we succeeded in transforming AspB into a set of tailored β-lyases that catalyze direct β-hydroamination of aliphatic, polar and aromatic α,β-unsaturated carboxylic acids to form the corresponding β-amino acids with full regiospecificity and enantiospecificity.

Results

Computational redesign of AspB. We first examined the reaction mechanism of AspB, which has been investigated by X-ray crystallography[38], site-directed mutagenesis[29] and quantum chemical calculations[39]. In the deamination direction, the general base Ser318 abstracts the *pro-R* proton from the C β atom of the aspartate (Fig. 1a)[38]. The negative charge on the enolate intermediate is stabilized by a network of hydrogen bonds between Thr101, Ser140, Thr141, Ser319 and the β-carboxylate group[39]. The collapse of the intermediate produces fumarate and ammonia (Fig. 1a,b). Substrate binding also involves interactions between the α-carboxylate and the protein environment (Asn142, Thr187, His188, Met321, Lys324 and Asn326)[39]. We hypothesized that hydroamination of α,β-unsaturated carboxylic acids other than fumarate could be achieved by reshaping the α-carboxylate binding pocket while maintaining interactions stabilizing negative charge on the β-carboxylate group. Accordingly, we used Rosetta Enzyme Design[37] to create active sites that accommodate alternative groups. Design constraints were defined to hold the substrate in a reactive or near-attack conformation and to preserve the interaction network of the β-carboxylate (Fig. 1c) while the α-carboxylate binding site was redesigned for new substituent groups by using the Rosetta energy function together with conformational sampling of side chain rotamers with a flexible backbone. This procedure predicted variants carrying sets of compatible mutations that optimize functional interactions between the new substrate and protein. A small number of designs were selected for experimental characterization based on the penalty score of the constraints and inspection of the designed structures (Fig. 1d).

Design of B19 for synthesis of (*R*)-β-aminobutanoic acid. The first target was an enzyme for production of β-aminobutanoic acid (**2**), which is a potent priming agent that provides broad-spectrum disease protection in at least 40 different plant species[40]. We reasoned that a hydrophobic environment could accommodate the methyl group. Because Asn142 and His188 function as the general acid for the leaving group ammonia[38], these amino acids were excluded from mutagenesis, and diversity was generated in silico by simultaneous substitutions of the other residues lining the α-carboxylate binding pocket (Thr187, Met321, Lys324 and Asn326) to any of the ten hydrophobic amino acids (alanine, cysteine, phenylalanine, glycine, isoleucine, leucine, methionine, proline, valine or tryptophan). To evaluate the success rate of this computational redesign strategy, we selected a relatively large set of 34 designs (Supplementary Table 2) for experimental validation. Measuring the production of crotonic acid (**1**) from β-aminobutanoic acid revealed that 14 of the 34 enzymes showed activity (Fig. 2a and Supplementary Table 3). Of these, we expressed the four most active designs in larger quantities and examined them for crotonic acid hydroamination in the presence of 0.3 M ammonia (Supplementary Fig. 2 and Supplementary Table 4)[35]. The best design (B19) contains four mutations (T187C, M321I, K324L and N326A), which modeling indicated had disrupted the native hydrogen-bonding network and promoted the hydrophobic interactions between the substrate's β-methyl group and the surrounding residues (Fig. 2b,c).

After enhancing the expression of design B19 in *Escherichia coli* (>5 g enzyme per liter culture) using high-density fermentation with inexpensive minimal medium, we examined its performance in practical synthesis. The whole-cell catalyst was added (final density $OD_{600} = 60$) to a 3.8-L reaction mixture containing 300 g/L crotonic acid, adjusted to pH 9 with ammonia. The reaction proceeded to >99% conversion in 8 h (Fig. 2d). From the mixture, 1.26 kg of pure (*R*)-β-aminobutanoic acid was obtained with an enantiomeric excess (e.e.) of 99% in 92% isolated yield (Supplementary Figs. 3 and 4). Although the economic feasibility of an industrial biocatalytic process needs to be examined on a case-by–case basis, there are some general threshold values that should be attained, including substrate concentration (>50 g/L), product yield (>90%), product profile (e.e. and regioselectivity >90%)[41,42]. The low catalyst production cost, the notably high substrate concentration, the absolute regioselectivity and enantioselectivity, the excellent catalyst productivity and the high product yield of the hydroamination route described herein are all suggestive of practical industrial applicability.

Tailored enzymes for synthesis of diverse β-amino acids. Encouraged by the successful design of an enzyme for the production of (*R*)-β-aminobutanoic acid, we examined the use of this computational approach for creating additional catalysts for other practically relevant conversions. The next target product we investigated was β-aminopentanoic acid (**4**), representing a common structural motif in various bioactive natural products, such as the cytotoxic cyclic depsipeptide obyanamide[43]. Because of the high success rate in the case of β-aminobutanoic acid, we constructed only five designs, three of which exhibited deamination activity toward β-aminopentanoic acid (Fig. 3a and Supplementary Tables 5 and 6). Intriguingly, B19 (like P4) was predicted by computational algorithm to possess, and indeed showed, promiscuous activity for β-aminopentanoic acid, but a more active design (T187C, M321I,

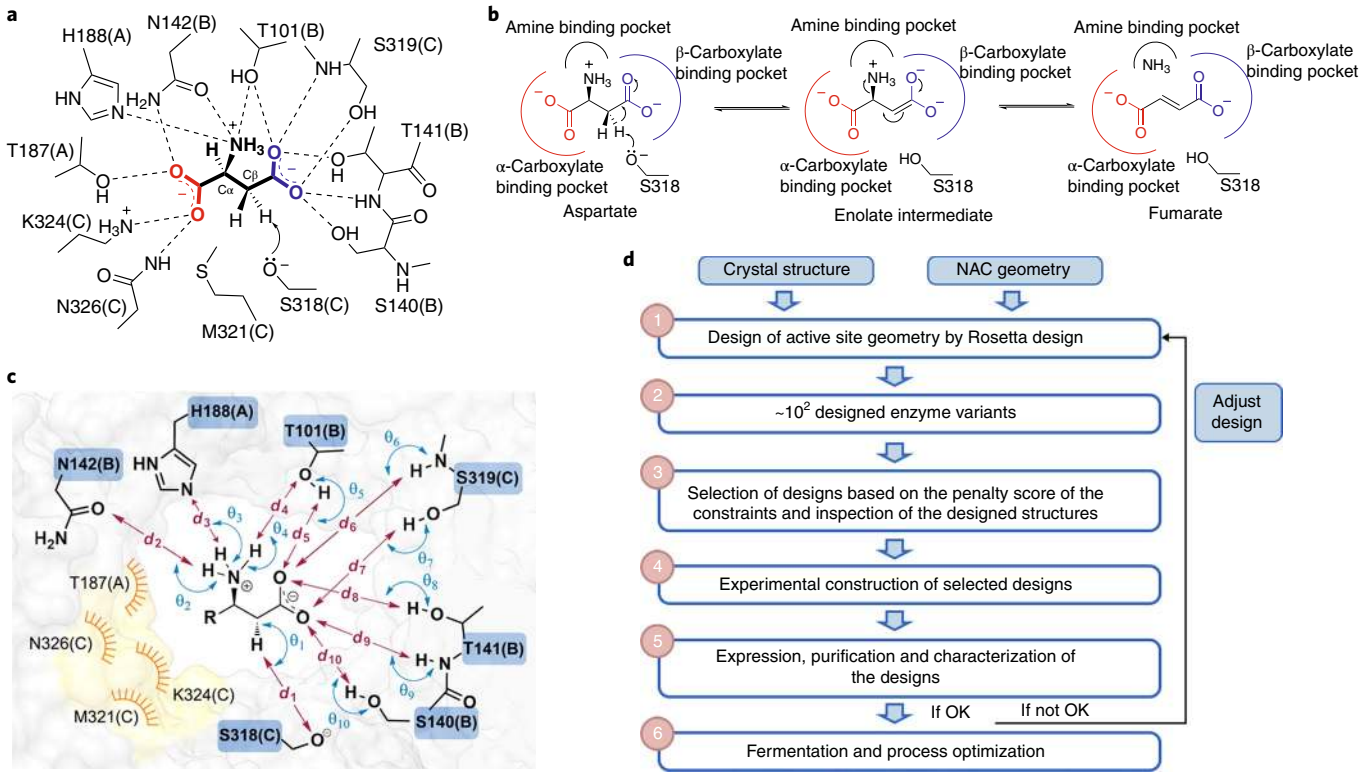

Fig. 1 | Computational active site redesign of AspB. a, Schematic representation of the interactions between aspartate and AspB, based on PDB 3R6V. Letters in parentheses are the chain names in PDB 3R6V. Hydrogen bonds are shown as dashed lines. **b**, Proposed reaction mechanism of AspB, in which the participation of the β-carboxylate group is sufficient for catalysis. **c**, Geometric constraints of enzyme–substrate interactions for catalytic activity with β-amino acids. Details are listed in Supplementary Table 1. **d**, A flowchart for computational redesign of AspB for alternative substrates.

K324L and N326C) for β-aminopentanoic acid was also identified (P1). The mutations in the Rosetta-predicted model conferred enhanced hydrophobic interactions around the ethyl group (Figs. 2b and 3b). P1 was also a practical biocatalyst (Supplementary Fig. 5 and Supplementary Table 7). Within 38 h, whole cells expressing the enzyme (OD_{600} = 60) converted (*E*)-2-pentenoic acid (**3**; 80 g/L) to (*R*)-β-aminopentanoic acid with e.e. >99% and 99% conversion as determined by HPLC, and the desired product was isolated with 84% yield (Fig. 3c and Supplementary Figs. 6 and 7).

Next we explored the use of computational design for the more polar compound β-asparagine (**6**). Asparagine to (*S*)-β-asparagine conversion is encountered in naturally occurring peptides[38], causing changes in conformation and altered biological activity. Accordingly, β-asparagine has been used for preparing peptide mimetics that were evaluated for their therapeutic potential[44]. Although (*S*)-β-asparagine resembles the native substrate aspartate, AspB has no detectable activity toward this compound[29]. To construct a new hydrogen-bonding network around the side chain amide, the residue Lys324, which confers hydrogen bonding interactions with the amide group of (*S*)-β-asparagine, was allowed to mutate in silico to any polar residue long enough to contact the substrate from this position (aspartate, glutamate, histidine, lysine, asparagine, glutamine, arginine or tyrosine), whereas the other α-carboxylate binding pocket residues (Thr187, Met321 and Asn326) were fully randomized to all 20 possible amino acids. Four of the six predicted designs (Supplementary Table 8) that we selected for experimental validation demonstrated β-asparagine deamination activity (Fig. 4a and Supplementary Table 9). The most active design, N5 (Supplementary Fig. 8 and Supplementary Table 10), contains three substitutions (M321I, K324N and N326C), for which the Rosetta-predicted model inferred a hydrogen bond between Thr187 and the amide oxygen atom and the formation of non-native hydrogen bonds between Asn324 and the amide group for substrate binding. Furthermore, Ile321 and Cys326 favor hydrophobic interactions with the aliphatic chain of β-asparagine (Figs. 2b and 4b). Cells expressing variant N5 (OD_{600} = 40) were used in gram-scale hydroamination of fumaric acid monoamide (**5**; 130 g/L), affording (*S*)-β-asparagine (e.e. >99%, 88% isolated yield), after complete conversion of the substrate in 24 h (Fig. 4c and Supplementary Figs. 9 and 10).

Finally, we sought to design an enzyme for the hydroamination of (*E*)-cinnamic acid (**7**) to produce (*S*)-β-phenylalanine (**8**), which is the chiral precursor of the potent antibiotic andrimid[45], the antiretroviral drug maraviroc (Selzentry, Pfizer)[46] and dapoxetine (Priligy, Johnson & Johnson) for the treatment of premature ejaculation[47]. Initial predictions introduced hydrophobic residues at the first shell of the binding pocket (Thr187, Met321, Lys324 and Asn326). We chose six designs (F1–F6, Supplementary Table 11) for experimental analysis, and two showed slight deamination activity for β-phenylalanine (Fig. 5a and Supplementary Table 12), but the cinnamate hydroamination activities were almost negligible. Considering that the binding pocket is formed by a flexible loop and that more drastic conformational alterations might be required, we included the second-shell residues (Ala192, Ala231 and Val232) in the computational mutagenesis, but this failed to deliver an improvement in catalysis (F7–F27, Supplementary Table 12). Assuming that expansion of the conformational search space was still needed beyond what is sampled when starting with the wild-type AspB structure, we adopted a different strategy. The Rosetta model of the slightly active design F4 (T187C, K324I and N326C) was subjected to a molecular dynamics simulation, with the expectation that this would remove structural bias and generate a more accurate backbone structure that would serve as a template for further design calculations (Supplementary Fig. 11). The molecular dynamics–refined model was used as the

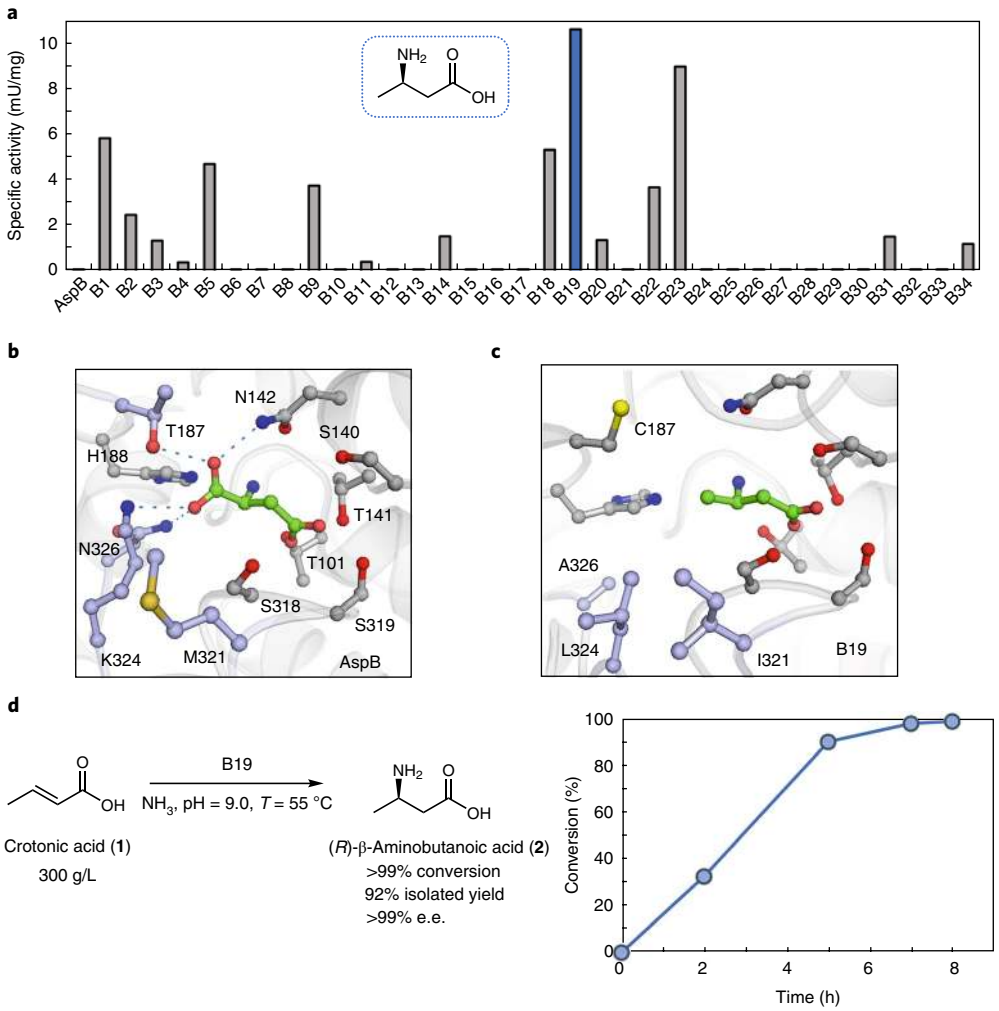

Fig. 2 | Computational redesign of AspB for (*R*)-β-aminobutanoic acid synthesis. a, The deamination activity of the 34 designed enzymes for the target (*R*)-β-aminobutanoic acid. Data are from one independent experiment. **b**, The active site of AspB with bound aspartate (PDB: 3R6V) showing the native hydrogen-bonding network. **c**, The Rosetta-predicted active site of B19 with (*R*)-β-aminobutanoic acid bound, highlighting disruption of the native hydrogen-bonding network and the predicted formation of a new hydrophobic binding pocket. **d**, The kilogram scale preparation of (*R*)-β-aminobutanoic acid using design B19. Data are from one independent experiment.

template for a second round of Rosetta design. To avoid steric hindrance with the substrate's phenyl group, mutations at positions 321 and 324 were rationally restricted to less bulky amino acids (methionine, isoleucine, valine, alanine, leucine, glycine or cysteine). Three more designs (F28–F30) were expressed and tested, resulting in the best variant, F29 (T187C, M321V, K324I and N326C). This design exhibited a higher activity in the deamination of β-phenylalanine (Fig. 5a and Supplementary Table 12). These mutations in the Rosetta-predicted model indicated a relaxed local conformational change of the loop and an enlarged binding pocket to accommodate the phenyl group (Fig. 5b,c).

We examined the use of design F29 for synthesis of (*S*)-β-phenylalanine from cinnamic acid (Supplementary Fig. 12 and Supplementary Table 13). Whereas conversion was lower than with the aliphatic unsaturated carboxylic acids owing to the thermodynamic equilibrium, the designed enzyme was still useful for β-amino acid synthesis. In a gram-scale preparative experiment, whole cells expressing F29 transformed (*E*)-cinnamic acid (15 g/L) into (*S*)-β-phenylalanine (e.e. >99%) with 50% conversion and 43% isolated yield (Fig. 5d and Supplementary Figs. 13 and 14). Unlike with natural and engineered phenylalanine aminomutases[28,48], no detectable amount of α-phenylalanine was formed. To our knowledge, this is the first reported enzyme with both absolute stereoselectivity and β-regioselectivity in the asymmetric hydroamination of cinnamic acid.

Discussion

Aspartase is a highly selective enzyme with no promiscuous activity, making it a notoriously difficult candidate as a starting point for delivery of a biocatalysis toolbox by traditional rational design or directed evolution. Nevertheless, the results reported here show that AspB can be computationally redesigned for different aliphatic, polar and aromatic acrylates. Using Rosetta design for in silico sampling of the sequence and conformational space that is accessible by partial active site randomization, we designed libraries, each encompassing a restricted number of variants, that delivered practically useful biocatalysts. The practical utility of the redesigned enzymes was demonstrated by the synthesis of four important chiral β-amino acids that are either bioactive compounds or building blocks for natural products and pharmaceuticals.

This biocatalytic system directly conjugates unactivated acrylates and the simplest nitrogen donor, ammonia, under mild conditions. Since there is no need for activation or deprotection and no formation of byproduct, these hydroamination reactions provide

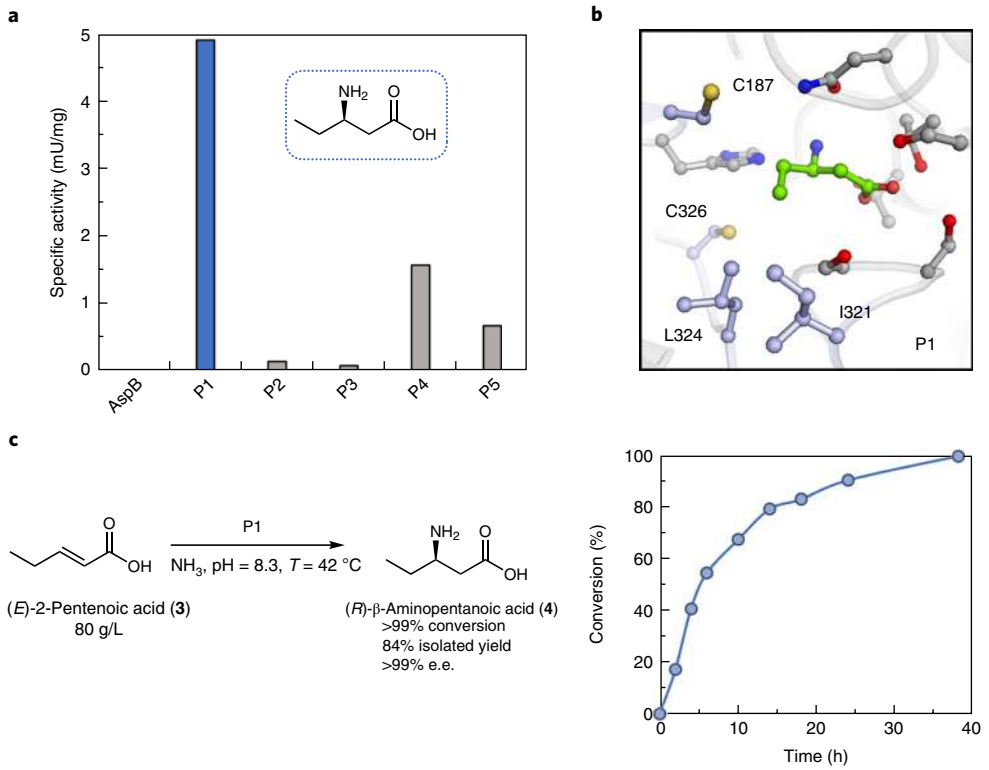

Fig. 3 | Computational redesign of AspB for (*R*)-β-aminopentanoic acid synthesis. a, The deamination activity of the designed enzymes for the target (*R*)-β-aminopentanoic acid. Data are from one independent experiment. **b**, The Rosetta-predicted active site of P1 with (*R*)-β-aminopentanoic acid bound, highlighting the predicted formation of a new hydrophobic binding pocket. **c**, The 10-gram-scale preparation of (*R*)-β-aminopentanoic acid using design P1. Data are from one independent experiment.

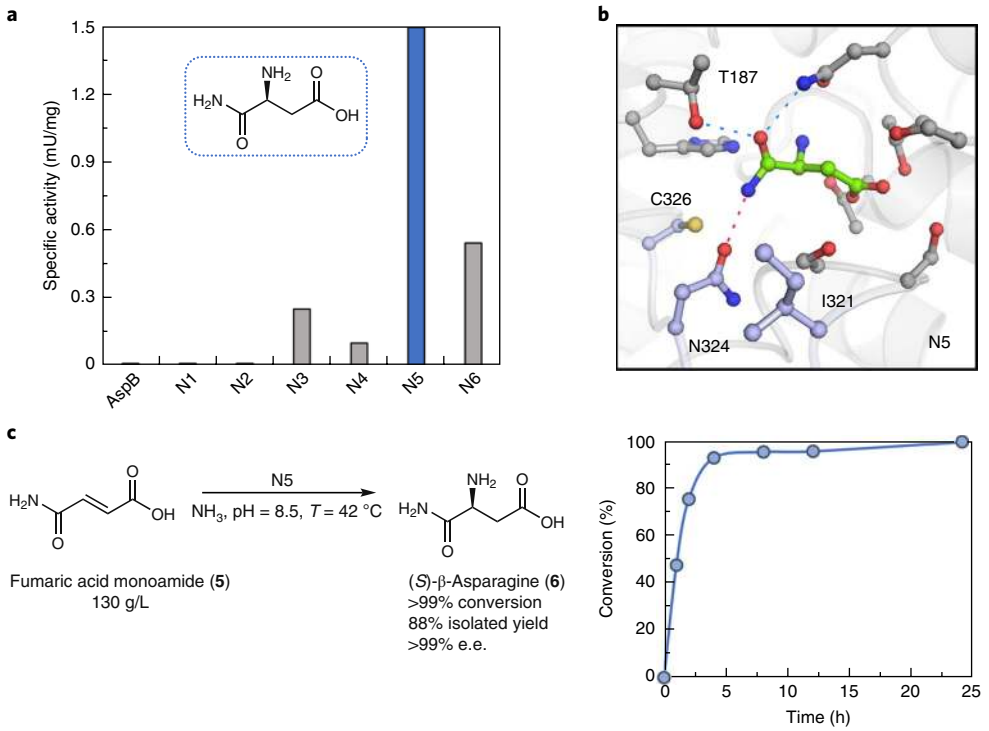

Fig. 4 | Computational redesign of AspB for (*S*)-β-asparagine synthesis. a, The deamination activity of the designed enzymes for the target β-asparagine. Data are from one independent experiment. **b**, The Rosetta-predicted active site of N5 bound with (*S*)-β-asparagine, highlighting a putative alternative hydrogen-bonding network. **c**, The gram-scale preparation of (*S*)-β-asparagine using design N5. Data are from one independent experiment.

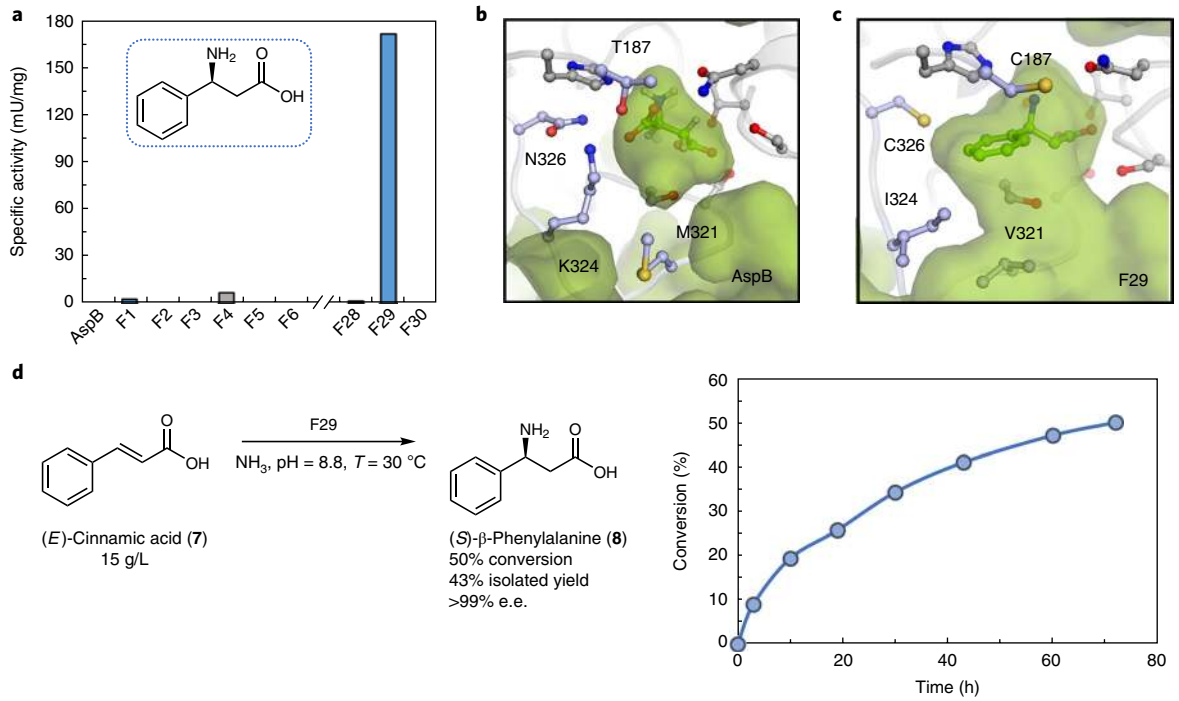

Fig. 5 | Computational redesign of AspB for (*S*)-β-phenylalanine synthesis. a, The deamination activity of the designed enzymes for the target β-phenylalanine. The left panel represents the variants from the first round of design and the right panel represents the designs based on a molecular dynamics–refined model. Data are from one independent experiment. **b**, The active site of AspB bound to aspartate (PDB: 3R6V), showing the tight fit of this substrate. **c**, The Rosetta-predicted active site of F29 bound to (*S*)-β-phenylalanine. The putative enlarged binding pocket of the F29 variant accommodates the substrate's phenyl group. **d**, The gram-scale preparation of (*S*)-β-phenylalanine using design F29. Data are from one independent experiment.

optimal atom efficiency for asymmetric β-amino acid synthesis, surpassing the processes of traditional metal catalysis and organocatalysis. Alongside the intrinsic advantage of biocatalysis in sustainability, the downstream processes also have low environmental impact. Product can be isolated via direct crystallization or simple ion-exchange operations, avoiding the need for column chromatography with organic solvents. Furthermore, the industrial potential and scalability of the system has been demonstrated in preparative experiments under commercially relevant protocols (up to kilogram scale). In principle, the computational redesign strategy may be applicable to the generation of ammonia lyase biocatalysts for the production of other β-amino acids of interest. Future challenges include the design of enantiocomplementary enzymes, which requires amination at the *si* face of the C β atom of acrylic acid and derivatives.

In all redesigned enzymes described above, the original β-carboxylate binding pocket of AspB was maintained because it is required for stabilizing the transition state of the reaction by interactions with delocalized negative charge, whereas new binding pockets for substituents were introduced. The loss of electrostatic interactions between the α-substituent and the protein environment explains the low affinity (high K_m values) of the designs for the new substrate, which at first sight may seem a disadvantage of the new ammonia lyases. However, for economic reasons, industrial biocatalysis must be performed at high substrate loadings anyway, and the elevated K_m values imply that reaction rates continue to increase with higher substrate concentration, in this case up to rates that clearly are industrially feasible[49]. Thus, in a campaign of computational enzyme design for practical biocatalysis, the catalytic effectiveness is much more important than the k_{cat}/K_m ratio, the traditional measure of enzyme performance[50].

In summary, the results described here demonstrate that a computational algorithm for de novo protein design can predict small sets of mutants for discovery of industrially viable hydroamination biocatalysts acting on structurally distinct acrylates. Large steps in function could be made rapidly by computational exploration of relevant sequence space, bypassing the need for multiple iterations of mutagenesis and testing and avoiding the need to employ high-throughput screening protocols with model substrates. As the performance of computational algorithms continues to improve, we foresee in silico design and screening becoming integrated into directed evolution protocols and reducing time- and cost-consuming laboratory work for creating industrially pertinent biocatalysts.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41589-018-0053-0.

Received: 14 October 2017; Accepted: 9 March 2018;
Published online: 21 May 2018

References

1. Bornscheuer, U. T. et al. Engineering the third wave of biocatalysis. *Nature* **485**, 185–194 (2012).
2. Reetz, M. T. Biocatalysis in organic chemistry and biotechnology: past, present, and future. *J. Am. Chem. Soc.* **135**, 12480–12496 (2013).
3. Nestl, B. M., Hammer, S. C., Nebel, B. A. & Hauer, B. New generation of biocatalysts for organic synthesis. *Angew. Chem. Int. Ed. Engl.* **53**, 3070–3095 (2014).
4. Savile, C. K. et al. Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science* **329**, 305–309 (2010).
5. Pavlidis, I. V. et al. Identification of (*S*)-selective transaminases for the asymmetric synthesis of bulky chiral amines. *Nat. Chem.* **8**, 1076–1082 (2016).

6. Abrahamson, M. J., Vázquez-Figueroa, E., Woodall, N. B., Moore, J. C. & Bommarius, A. S. Development of an amine dehydrogenase for synthesis of chiral amines. *Angew. Chem. Int. Ed. Engl.* **51**, 3969–3972 (2012).
7. 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).
8. Kille, S., Zilly, F. E., Acevedo, J. P. & Reetz, M. T. Regio- and stereoselectivity of P450-catalysed hydroxylation of steroids controlled by laboratory evolution. *Nat. Chem.* **3**, 738–743 (2011).
9. Kan, S. B., Lewis, R. D., Chen, K. & Arnold, F. H. Directed evolution of cytochrome *c* for carbon-silicon bond formation: bringing silicon to life. *Science* **354**, 1048–1051 (2016).
10. Blomberg, R. et al. Precision is essential for efficient catalysis in an evolved Kemp eliminase. *Nature* **503**, 418–421 (2013).
11. Jochens, H. & Bornscheuer, U. T. Natural diversity to guide focused directed evolution. *ChemBioChem* **11**, 1861–1866 (2010).
12. Bendl, J. et al. HotSpot Wizard 2.0: automated design of site-specific mutations and smart libraries in protein engineering. *Nucleic Acids Res.* **44**(W1), W479–W487 (2016).
13. Nobili, A. et al. Simultaneous use of in silico design and a correlated mutation network as a tool to efficiently guide enzyme engineering. *ChemBioChem* **16**, 805–810 (2015).
14. Lutz, S. Beyond directed evolution—semi-rational protein engineering and design. *Curr. Opin. Biotechnol.* **21**, 734–743 (2010).
15. Sebestova, E., Bendl, J., Brezovsky, J. & Damborsky, J. Computational tools for design smart libraries. in *Directed Evolution Library Creation* (Springer, New York, 2014).
16. Ebert, M. C. & Pelletier, J. N. Computational tools for enzyme improvement: why everyone can — and should — use them. *Curr. Opin. Chem. Biol.* **37**, 89–96 (2017).
17. Santiago, G. et al. Computer-aided laccase engineering: towards biological oxidation of arylamines. *ACS Catal.* **6**, 5415–5423 (2016).
18. Moroz, Y. S. et al. New tricks for old proteins: single mutations in a nonenzymatic protein give rise to various enzymatic activities. *J. Am. Chem. Soc.* **137**, 14905–14911 (2015).
19. Romero-Rivera, A., Garcia-Borràs, M. & Osuna, S. Computational tools for the evaluation of laboratory-engineered biocatalysts. *Chem. Commun. (Camb.)* **53**, 284–297 (2016).
20. Huang, P. S., Boyken, S. E. & Baker, D. The coming of age of de novo protein design. *Nature* **537**, 320–327 (2016).
21. Constable, D. J. C. et al. Key green chemistry research areas – a perspective from pharmaceutical manufacturers. *Green Chem.* **9**, 411–420 (2007).
22. Kudo, F., Miyanaga, A. & Eguchi, T. Biosynthesis of natural products containing β-amino acids. *Nat. Prod. Rep.* **31**, 1056–1073 (2014).
23. Ashfaq, M. et al. Enantioselective synthesis of β-amino acids: a review. *Med. Chem.* **5**, 295–309 (2015).
24. Liljeblad, A. & Kanerva, L. T. Biocatalysis as a profound tool in the preparation of highly enantiopure β-amino acids. *Tetrahedron* **62**, 5831–5854 (2006).
25. Rehdorf, J., Mihovilovic, M. D. & Bornscheuer, U. T. Exploiting the regioselectivity of Baeyer-Villiger monooxygenases for the formation of β-amino acids and β-amino alcohols. *Angew. Chem. Int. Ed. Engl.* **49**, 4506–4508 (2010).
26. Zhang, D. et al. Development of β-amino acid dehydrogenase for the synthesis of β-amino acids via reductive amination of β-keto acids. *ACS Catal.* **5**, 2220–2224 (2015).
27. Turner, N. J. Ammonia lyases and aminomutases as biocatalysts for the synthesis of α-amino and β-amino acids. *Curr. Opin. Chem. Biol.* **15**, 234–240 (2011).
28. Weise, N. J., Parmeggiani, F., Ahmed, S. T. & Turner, N. J. The bacterial ammonia lyase EncP: a tunable biocatalyst for the synthesis of unnatural amino acids. *J. Am. Chem. Soc.* **137**, 12977–12983 (2015).
29. Kawata, Y. et al. Cloning and over-expression of thermostable *Bacillus* sp. YM55-1 aspartase and site-directed mutagenesis for probing a catalytic residue. *Eur. J. Biochem.* **267**, 1847–1857 (2000).
30. 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).
31. Viola, R. E. L-Aspartase: new tricks from an old enzyme. *Adv. Enzymol.* **74**, 295–341 (2000).
32. Renata, H., Wang, Z. J. & Arnold, F. H. Expanding the enzyme universe: accessing non-natural reactions by mechanism-guided directed evolution. *Angew. Chem. Int. Ed. Engl.* **54**, 3351–3367 (2015).
33. Levin, K. B. et al. Following evolutionary paths to protein-protein interactions with high affinity and selectivity. *Nat. Struct. Mol. Biol.* **16**, 1049–1055 (2009).
34. Asano, Y., Kira, I. & Yokozeki, K. Alteration of substrate specificity of aspartase by directed evolution. *Biomol. Eng.* **22**, 95–101 (2005).
35. Vogel, A., Schmiedel, R., Hofmann, U., Gruber, K. & Zangger, K. Converting aspartase into a β-amino acid lyase by cluster screening. *ChemCatChem* **6**, 965–968 (2014).
36. Wu, B. et al. Versatile peptide C-terminal functionalization via a computationally engineered peptide amidase. *ACS Catal.* **6**, 5405–5414 (2016).
37. Wijma, H. J. et al. Enantioselective enzymes by computational design and in silico screening. *Angew. Chem. Int. Ed. Engl.* **54**, 3726–3730 (2015).
38. Fibriansah, G., Veetil, V. P., Poelarends, G. J. & Thunnissen, A. M. Structural basis for the catalytic mechanism of aspartate ammonia lyase. *Biochemistry* **50**, 6053–6062 (2011).
39. Zhang, J. & Liu, Y. A QM/MM study of the catalytic mechanism of aspartate ammonia lyase. *J. Mol. Graph. Model.* **51**, 113–119 (2014).
40. Cohen, Y., Vaknin, M. & Mauch-Mani, B. BABA-induced resistance: milestones along a 55-year journey. *Phytoparasitica* **44**, 513–538 (2016).
41. Lima-Ramos, J., Neto, W. & Woodley, J. M. Engineering of biocatalysts and biocatalytic processes. *Top. Catal.* **57**, 301–320 (2014).
42. Huisman, G. W. & Collier, S. J. On the development of new biocatalytic processes for practical pharmaceutical synthesis. *Curr. Opin. Chem. Biol.* **17**, 284–292 (2013).
43. Zhang, W. et al. Total synthesis and reassignment of stereochemistry of obyanamide. *Tetrahedron* **62**, 9966–9972 (2006).
44. Zhao, R. et al. Inhibition of the Bcl-xL deamination pathway in myeloproliferative disorders. *N. Engl. J. Med.* **359**, 2778–2789 (2008).
45. Fortin, P. D., Walsh, C. T. & Magarvey, N. A. A transglutaminase homologue as a condensation catalyst in antibiotic assembly lines. *Nature* **448**, 824–827 (2007).
46. Grayson, J. I., Roos, J. & Osswald, S. Development of a commercial process for (*S*)-β-phenylalanine. *Org. Process Res. Dev.* **15**, 1201–1206 (2011).
47. Owen, R. T. Dapoxetine: a novel treatment for premature ejaculation. *Drugs Today (Barc.)* **45**, 669–678 (2009).
48. Wu, B. et al. Mechanism-inspired engineering of phenylalanine aminomutase for enhanced β-regioselective asymmetric amination of cinnamates. *Angew. Chem. Int. Ed. Engl.* **51**, 482–486 (2012).
49. Eisenthal, R., Danson, M. J. & Hough, D. W. Catalytic efficiency and k_{cat}/K_M: a useful comparator? *Trends Biotechnol.* **25**, 247–249 (2007).
50. Fox, R. J. & Clay, M. D. Catalytic effectiveness, a measure of enzyme proficiency for industrial applications. *Trends Biotechnol.* **27**, 137–140 (2009).

Acknowledgements

We thank W. Szymanski for discussions. We thank for the 100 Talent Program grant (B.W.) and Biological Resources Service Network Initiative (ZSYS-012; B.W.) and a grant (SKT1604; C.Y.L.) from the Chinese Academy of Sciences, Natural Science Foundation of China grants (31601412 (B.W.), 21603013 (C.Y.L.)), and a BE-Basic grant (H.J.W. and D.B.J.) from the Dutch Ministry of Economic Affairs for the financial support.

Author contributions

D.B.J. and B.W. initiated the project. B.W., H.J.W. and Y. Cui performed the computational work. L.S., R.L., M.O., Y. Tian, J.D., T.L., D.N., Y. Chen and J.F. performed biocatalytic experiments. J.H., H.C. and Y. Tao developed high-density fermentation methods. R.L. performed preparative-scale synthesis of the amino acids. D.B.J. and B.W. provided supervision and input on experimental design and wrote the manuscript, which was revised and approved by all authors. R.L., H.J.W. and L.S. contributed equally to this work.

Competing interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41589-018-0053-0.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to D.B.J. or B.W.

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

Methods

General information. Commercial reagents were used as received: crotonic acid (Acros Organics), (*R*)-β-aminobutanoic acid (Alfa Aesar), (*RS*)-β-aminobutanoic acid (Shanghai Yuanye Bio-Technology Co.), (*E*)-2-pentenoic acid (Aldrich), (*RS*)-β-aminopentanoic acid (Ark Pharm), fumaric acid monoamide (Shanghai Kuanchao Technology Co.), (*S*)-β-asparagine (Hanhong chemical), (*E*)-cinnamic acid (Aladdin), (*S*)-β-phenylalanine (Damas-beta), (*RS*)-β-phenylalanine (Shangon Biotech Co.), 2,4-dinitrofluorobenzene (Sigma-Aldrich) and *N*-(2,4-dinitro-5-fluorophenyl)-L-valine amide (TCI). The purity of reagents was > 98%. HPLC analysis was performed on an Agilent 1200 instrument. NMR spectra were recorded on a Bruker AV500 spectrometer (500 MHz for ^{1}H NMR, 126 MHz for ^{13}C NMR).

Computational protein redesign. Computational redesign was carried out with the Rosetta Enzyme Design application[51] and the following command line options were used: -enzdes -cst_predock -cst_design -detect_design_interface -cut1 0.0 -cut2 0.0 -cut3 8.0 -cut4 10.0 -cst_min -chi_min -bb_min -packing::use_input_sc -packing::soft_rep_design -extrachi_cutoff 1 -design_min_cycles 3 -ex1:level 4 -ex2:level 4 -ex1aro:level 4 -ex2aro:level 4. An ensemble of different conformations of the substrate was generated by enumerating these under Yasara. Substrate rotamers were sampled around the canonical minimum dihedral angles (60°, –60° and 180°) with 5° intervals over 35° around the minima (for example, 42.5° to 77.5°).

The Rosetta Enzyme Design application positions substrates optimally for catalysis by applying forces between the bound substrate and catalytically important groups in the enzyme. The substrate geometry corresponded to near attack conformation criteria and was based on published QM/MM calculations[39] and the crystal structure with bound substrate (PDB 3R6V) (Supplementary Table 1). Rosetta enzyme design oriented the substrate optimally for deamination by applying these forces in silico.

Rosetta Enzyme Design uses a Monte Carlo algorithm in which it selects mutations and structural changes that decrease overall energy to generate 3D structures of designs. The designs were selected for experimental characterization on the basis of the following guidelines: (1) the sum of the penalty energies for the above constrains should not exceed 20 REU, (2) the active site must be organized, and maintain the original β-carboxylate hydrogen-bonding network, (3) the introduced mutations are not allowed to result in large cavities inside the protein model (visual inspection), and (4) there should not be more than two unsatisfied hydrogen bond donors or acceptors present in the design.

For molecular dynamics refinement, the 3D structure of the mutant F4 predicted by Rosetta was positioned in a rectangular simulation cell with at least 7.5 Å between protein and the periodic boundary of the simulation cell. All original water molecules present in the wild-type AspB structure (PDB: 3R6V) were put back unless they clashed sterically with the designed protein structure. The salt ions were positioned at electrostatically favorable positions by an algorithm implemented under Yasara. An energy minimization was carried out before the molecular dynamics simulation. Molecular dynamics simulation was carried out under Yasara with a leapfrog algorithm, with a time-step of 1.33 fs and a Berendsen thermostat to preserve constant pressure and temperature. The LINCS and SETTLE algorithms were used to constrain hydrogen atoms. The simulations were carried out using the Yamber3 force field. The temperature was increased from 5 to 298 K over 3 ps, followed by equilibration (2 ps) and production (5 ps).

Construction of AspB variants. Plasmid pET21a, containing the AspB sequence, was used as a template for QuikChange mutagenesis with Q5 PCR MasterMix (NEB). PCR was performed in 0.5 mL microcentrifuge tubes and the DpnI-treated PCR products were transformed into chemically competent *E. coli* TOP10. Incorporation of the mutations was confirmed by DNA sequencing.

Expression and purification of AspB variants for kinetic measurement. The enzymes were expressed from a pET21a vector in autoinduction medium[52], which comprised 1% tryptone, 0.5% yeast extract, 0.33% $(NH_4)_2SO_4$, 0.68% KH_2PO_4, 0.71% Na_2HPO_4, 0.024% $MgSO_4$, 0.2% glycerol (v/v), 0.05% glucose, 0.2% lactose and 50 μg/mL ampicillin at 30 °C using *E. coli* BL21 (DE3) for 24 h. The final OD_{600} typically reached ~4. After centrifugation (8,000 *g*, 15 min), the cells were lysed by sonication in a buffer containing 50 mM Tris·HCl (pH 7.5), 2 mM $MgCl_2$. Enzymes were purified by heat treatment at 60 °C for 30 min and centrifuged (12,000 *g*, 90 min) to remove precipitates. The purity of the protein obtained by this method was generally >90% (Supplementary Fig. 15). The enzyme solution was stored in aliquots at 4 °C until further use. Protein concentrations were determined with the Bradford assay and bovine serum albumin as the standard (Takara). The yield of purified protein was typically ~300 mg/L culture. The purified enzymes were used to measure deamination and hydroamination activities.

Determination of apparent melting temperatures. A fluorescence-based thermal stability assay was used to determine apparent melting temperatures[53]. A sample (20 μL) of a protein in buffer (50 mM Tris·HCl, 2 mM $MgCl_2$, pH 7.5) was mixed with 5 μL 100-fold diluted SYPRO Orange dye (Molecular Probes, Life Technologies, USA) in a thin-walled 96-well PCR plate. The plate was sealed with optical-quality sealing tape and heated in a CFX 96 real-time PCR detection system (Bio-Rad, Hercules, CA, USA) from 20° to 99 °C at a heating rate of 1.75 °C/min. The wavelengths for excitation and emission were 490 and 575 nm, respectively.

HPLC analysis of hydroamination reactions. Concentrations of β-aminobutanoic acid, β-aminopentanoic acid and β-asparagine were determined by HPLC analysis after DNFB (2,4-dinitrofluorobenzene) derivatization. Concentration of β-phenylalanine was determined by direct HPLC.

Derivatization. A mixture of 25 μL amino acid standard solutions or reaction solutions (concentration (amino acid + ammonia) < 50 mM) were mixed with 10 μL 1 M $NaHCO_3$, 40 μL DNFB (36.7 mM in acetone) and incubated at 60 °C for 30 min. The reaction was stopped by adding 20 μL of 1 M HCl and the precipitates were subsequently removed by centrifugation. To analyze the outcome of the reaction with respect to stereochemistry, 40 μL FDVA (*N*-(2,4-di-nitro-5-fluorophenyl)-L-valine amide) (36.7 mM in acetone) was added instead of DNFB.

HPLC analysis. General HPLC condition: column: Nucleosil C18 5 μ (250 × 4.6 mm); temperature: 25 °C; eluent A: 0.1% formic acid in water, eluent B: acetonitrile; flow rate: 1 ml·min^{-1}.

DNFB-β-aminobutanoic acid: detection: UV at 360 nm; gradient: 15% B → 50% B in 22 min, 50% B → 15% B from 22 min to 22.1 min, 15% B from 22.1 min to 26 min; retention time: 19.8 min.

DNFB-β-aminopentanoic acid: detection: UV at 360 nm; gradient: 15% B → 50% B in 22 min, 50% B → 15% B from 22 min to 22.1 min, 15% B from 22.1 min to 26 min; retention time: 22.3 min.

DNFB-β-asparagine: detection: UV at 360 nm; gradient: 15% B → 32%B in 20 min, 32% B → 15% B from 20 min to 20.1 min, 15% B from 20.1 min to 24 min; retention time: 14.0 min.

β-Phenylalanine: detection: UV at 210 nm; gradient: 5% in 0–11 min, 50% B in 11–18 min; retention time: 6.9 min.

FDVA-chiral β-aminobutanoic acid: detection: UV at 340 nm; isocratic flow: 35% B for 30 min; retention time (*S*): 12.4 min, retention time (*R*): 18.0 min. (Supplementary Fig. 4).

FDVA-chiral β-aminopentanoic acid: detection: UV at 340 nm; isocratic flow: 35% B for 30 min; retention time (*S*): 16.0 min, retention time (*R*): 28.9 min (Supplementary Fig. 7).

FDVA-chiral β-asparagine: detection: UV at 340 nm; gradient: 25% B → 39% B in 30 min, 39% B → 25% B from 30 min to 30.1 min, 25% B from 30.1 min to 35 min; retention time (*S*): 12.1 min, retention time (*R*): 13.9 min (Supplementary Fig. 10).

FDVA-chiral β-phenylalanine: detection: UV at 340 nm; gradient: 15% B → 50% B in 22 min, 50% B → 15% B from 22 min to 30 min, 15% B from 30 min to 35 min, retention time (*R*): 23.6 min, retention time (*S*): 27.3 min (Supplementary Fig. 14).

Deamination activities of computational designs for β-aminobutanoic acid. A solution of 300 mM (*RS*)-β-aminobutanoic acid and 100 mM Na_2HPO_4 was prepared, and the pH was adjusted to 8.0 by adding 5 M NaOH. To this solution (180 μL), purified AspB or mutant B1–B34 (20 μL, concentrations listed in Supplementary Table 3) was added. The reactions were performed at 55 °C for 1 h. The specific activity was determined by measuring the concentration of the crotonic acid formed after 1 h reaction time. One unit of enzyme is defined as the amount that converts 1 μmol of substrate per minute.

Deamination activities of the computational designs for β-aminopentanoic acid. A solution of 300 mM (*RS*)-β-aminopentanoic acid and 100 mM Na_2HPO_4 was prepared and the pH was adjusted to 8.0 by adding 5 M NaOH. To this solution (180 μL), purified AspB or mutant P1–P5 enzyme (20 μL, concentrations listed in Supplementary Table 6) was added and the reactions were performed at 55 °C for 2 h. The specific activity was determined by measuring the concentration of the formed (*E*)-2-pentenoic acid after 2 h reaction time. One unit of enzyme is defined as the amount that converts 1 μmol of substrate per minute.

Deamination activities of the computational designs for β-asparagine. A solution of 300 mM (*S*)-β-asparagine and 100 mM Na_2HPO_4 was prepared, and the pH was adjusted to 8.0 by adding 5 M NaOH. To this solution (90 μL), purified AspB or N1–N6 (10 μL, concentrations listed in Supplementary Table 9) was added, and the reactions were left at 55 °C for 15 min. The specific activity was determined by measuring the concentration of the formed fumaric acid monoamide after 15 min reaction time. One unit of enzyme is defined as the amount that converts 1 μmol of substrate per minute.

Deamination activities of the computational designs for β-phenylalanine. A solution of 50 mM (*S*)-β-phenylalanine and 100 mM Na_2HPO_4 was prepared, and the pH was adjusted to 8.0 by adding 5 M NaOH. To this solution (90 μL), purified AspB or F1–F30 (10 μL, concentrations are listed in Supplementary Table 12) was added, and the reactions were left at 55 °C for 2 h. The specific activity was determined by measuring the concentration of the formed (*E*)-cinnamic acid after

2 h reaction time. One unit of enzyme is defined as the amount that converts 1 μmol of substrate per minute.

Fed-batch fermentation of mutants. *E. coli* BL21 (DE3) containing the appropriate expression plasmid from a fresh LB-agar plate was used to prepare a seed culture by inoculation of 1 L LB broth in a 3-L flask, followed by cultivation to $OD_{600} = 2$ at 37 °C. Subsequent bioreactor cultivation was performed in a 20-L fermenter. A seed culture was inoculated 10% (v/v) into minimal medium, which comprised 0.34% KH_2PO_4, 0.4% Na_2HPO_4, 0.4% $(NH_4)_2SO_4$, 0.18% citric acid, 0.03% $MgSO_4$, 0.9% glucose, 50 μg/mL ampicillin and 1% (v/v) trace elements solution. This last contained 840 mg/L EDTA, 136 mg/L $CoCl_2$, 954 mg/L $MnCl_2$, 118 mg/L $CuCl_2$, 300 mg/L H_3BO_3, 213 mg/L Na_2MoO_4, 209 mg/L $Zn(CH_3COO)_2$, and 11.4 g/L ammonium ferric citrate. The whole fed-batch cultivation was divided into three phases. The first phase was started with an initial glucose concentration of 10 g/L at 37 °C. After the initial glucose was completely consumed, as indicated by a sudden increase of dissolved oxygen (DO), a feed with sterilized 50% (w/v) glucose solution was initiated (phase 2, Supplementary Fig. 16). This stepwise method avoids the accumulation of acetic acid. When the OD_{600} reached approximately 20, 30 mg/L IPTG was added to induce the expression of the mutant (phase 3). During the whole process, the pH was kept at 7.0 by the addition of 25% (v/v) ammonia solution when required. The DO level was kept at 30% of air saturation by controlling the cascading impeller speed. The aeration rate was kept at 1.0 $L \cdot L^{-1} \cdot min^{-1}$. The OD_{600} typically reached approximately 60 and the yield of AspB mutant enzyme was typically ~5 g/L culture as estimated from SDS–PAGE analysis and Bradford protein assays (Supplementary Fig. 17 and Supplementary Table 14). Whole cells were used for preparative-scale synthesis of the β-amino acids.

Synthesis and characterization. Synthetic procedures using B19, P1, N5 and F29 and compound characterization data are included in Supplementary Notes 1–4.

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

Data and code availability. The data supporting the findings of this study are available within the paper and the Supplementary Information files.

References

51. Rajagopalan, S. et al. Design of activated serine-containing catalytic triads with atomic-level accuracy. *Nat. Chem. Biol.* **10**, 386–391 (2014).
52. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
53. Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N. & Nordlund, P. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal. Biochem.* **357**, 289–298 (2006).

nature research

Corresponding author(s): D.B. Janssen and B. Wu

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

   Describe how sample size was determined. — This is not relevant to the study, because it is a protein design project.

2. Data exclusions

   Describe any data exclusions. — No data were excluded

3. Replication

   Describe whether the experimental findings were reliably reproduced. — The experimental findings were reliably reproduced

4. Randomization

   Describe how samples/organisms/participants were allocated into experimental groups. — This is not relevant to the study, because it is a protein design project.

5. Blinding

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis. — This is not relevant to the study, because it is a protein design project.

   Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

   For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☒	☐ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☒	☐ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☒	☐ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	Rosetta 3.7 and Yasara 15.3.8

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	No unique materials were used

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	No antibodies were used

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	No eukaryotic cell lines were used
b. Describe the method of cell line authentication used.	No eukaryotic cell lines were used
c. Report whether the cell lines were tested for mycoplasma contamination.	No eukaryotic cell lines were used
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	No eukaryotic cell lines were used

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	No animals were used

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	The study did not involve human research participants