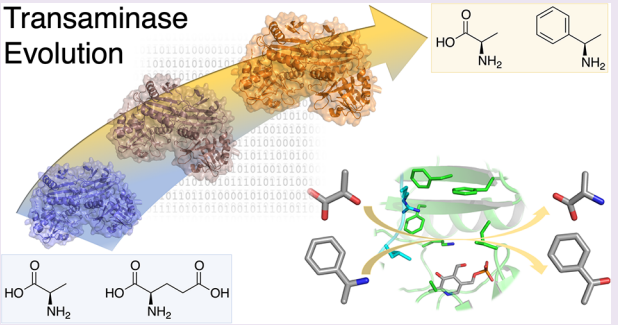
Creation of (R)-Amine Transaminase Activity within an α-Amino Acid Transaminase Scaffold

Moritz Voss, Chao Xiang, Jeremy Esque, Alberto Nobili, Marian Menke, Isabelle André, Matthias Höhne, Uwe Bornscheuer

Creation of (*R*)-Amine Transaminase Activity within an α-Amino Acid Transaminase Scaffold

Moritz Voss, Chao Xiang, Jérémy Esque, Alberto Nobili, Marian J. Menke, Isabelle André, Matthias Höhne, and Uwe T. Bornscheuer*

ABSTRACT: The enzymatic transamination of ketones into (*R*)-amines represents an important route for accessing a range of pharmaceuticals or building blocks. Although many publications have dealt with enzyme discovery, protein engineering, and the application of (*R*)-selective amine transaminases [(*R*)-ATA] in biocatalysis, little is known about the actual *in vivo* role and how these enzymes have evolved from the ubiquitous α-amino acid transaminases (α-AATs). Here, we show the successful introduction of an (*R*)-transaminase activity in an α-amino acid aminotransferase with one to six amino acid substitutions in the enzyme's active site. Bioinformatic analysis combined with computational redesign of the D-amino acid aminotransferase (DATA) led to the identification of a sextuple variant having a specific activity of 326 milliunits mg^{-1} in the conversion of (*R*)-phenylethylamine and pyruvate to acetophenone and D-alanine. This value is similar to those of natural (*R*)-ATAs, which typically are in the range of 250 milliunits mg^{-1}. These results demonstrate that (*R*)-ATAs can evolve from α-AAT as shown here for the DATA scaffold.



The family of pyridoxal 5′-phosphate (PLP)-dependent enzymes is a prime example for the natural diversity of enzyme activities that evolved within one structural scaffold. The superfamily of aspartate transaminases (fold type I) is the largest of seven known fold types. Besides enzymes catalyzing transamination, it contains activities that belong to EC classes 1−5, including racemization, decarboxylation, aldolase reactions, and a lipoamide-dependent oxidative decarboxylation and the C−C bond hydrolase kynureninase. Furthermore, different substrate specificities and stereopreferences are found among enzymes in these fold types.[1]

Transaminases are exclusively found in fold types I and IV and catalyze the reversible transfer of an amino group from an amino donor to an aldehyde, prochiral keto acid, or ketone, resulting in a (chiral) amine.[2] Transaminases are usually homodimers, with the active site located in the dimer interface. The transamination mechanism is a ping-pong bi-bi mechanism based on two half-reactions.[3] At the beginning of the catalytic cycle, the PLP is covalently bound as a Schiff base to the ε-group of the catalytic lysine in the enzyme's active site (called the internal aldimine). The first half of the reaction is initiated by the nucleophilic attack of the amino donor (e.g., alanine) where the amino donor replaces the lysine and forms the external aldimine. After abstraction of a proton from the Cα atom of the amino donor by the catalytic lysine, the planar quinonoid intermediate is formed. Via the transfer of the abstracted proton to C4′ of the cofactor, the chiral ketimine is created, which ultimately is hydrolyzed. This completes the first half-reaction with the release of the ketone while pyridoxamine 5′-phosphate (PMP) remains in the active site. In the second half-reaction, the reaction occurs in reverse order, starting with binding of the amino acceptor (e.g., a ketone) substrate.

Transaminases are highly enantioselective and substrate specific. α-Amino acid aminotransferases (AAT) are ubiquitous enzymes present in all organisms and exclusively convert α-amino acids or the respective α-keto acids, where the α-carboxylate function serves as an important recognition moiety for the substrate.[2] In contrast, the smaller subgroup of amine transaminases (ATAs) converts substrates that lack the α-carboxylic acid group, making ATAs suitable for the synthesis of industrially relevant chiral amines from prochiral ketones.[4,5] Interestingly, ATAs still accept alanine and pyruvate as co-substrates. The acceptance of both substrates with and without α-carboxylate in the same active site requires a dual-substrate recognition. In ATAs, this is facilitated by the side chain of an arginine residue, which stabilizes the negative charge of the α-carboxylate by a salt bridge, but also has the flexibility of "flipping out" of the active site in case an amine or ketone

Received: November 1, 2019
Accepted: January 28, 2020
Published: January 28, 2020

substrate lacking the α-carboxylate is binding as documented in the literature.[2,6,7]

Fold type IV of PLP-dependent enzymes is comprised of 4-amino-4-deoxychorismate lyases (ADCL) and three classes of transaminases: L-branched chain amino acid aminotransferases (BCAT), D-amino acid aminotransferases (DATA), and (*R*)-selective amine transaminases [(*R*)-ATA].[8] These different transaminase classes have a high degree of structural similarity in their overall fold but differ substantially in terms of their amino acid sequence, substrate scope, stereopreference, and activity.

The (*R*)-ATA is a group of enzymes discovered rather recently, since their activity was first reported in 2006 by Iwasaki et al., who described that the bacterial strain *Arthrobacter* sp. KNK168 can perform the whole-cell asymmetric synthesis of (*R*)-3,4-dimethoxyamphetamine.[9] In 2010, Savile et al. published the amino acid sequence of ATA117, a homologue of the transaminase that accounts for the (*R*)-ATA activity in *Arthrobacter* sp. KNK168, together with the extensively engineered variant ATA117-11Rd.[4] This variant harbors 27 mutations and was developed by Codexis and Merck & Co. for the asymmetric synthesis of the antidiabetic drug (*R*)-sitagliptin.[4] In parallel, Höhne et al. identified 17 novel (*R*)-ATAs by developing an algorithm for the sequence motif-based screening of protein data banks based on rational assignments.[8] At the time of that study, no protein or structural information for (*R*)-ATAs had been reported. Therefore, it was challenging to predict a sequence motif that correlates with (*R*)-ATA activity. The authors interrogated the protein sequences and structures of ADCLs, BCATs and DATAs, assuming that (*R*)-ATAs would also be a member of fold type IV. They based this assumption on the substrate coordination of DATA and BCAT, which share an overall similar active site architecture as assigned for the potential (*R*)-ATA.

Recently, α-AATs of fold type IV showing side activities toward (*R*)-amines were characterized.[10−14] Pavkov-Keller et al. described the DATA from *Curtobacterium pusillum* (*CPU*TA1),[13] and the thermostable BCAT from *Thermobaculum terrenum* (*TaTT*) was described by Bezsudnova et al.[11,14] Both of these transaminases exhibit the expected α-AAT activity, and significant activity toward (*R*)-amines was also observed, which distinguishes them from typical DATAs and BCATs. In addition, their sequence motifs differ from typical DATAs and BCATs, reflecting their unusual (*R*)-amine acceptance.[11]

Intrigued by the results of Höhne et al., we investigated whether it is possible to create (*R*)-ATAs from an α-transaminase by introducing key amino acid substitutions. Our major motivation was to develop a deeper understanding of how (*R*)-ATAs might have evolved by natural evolution within the α-AAT scaffold. Our overall engineering strategy was built upon the guidelines developed by Höhne et al. in combination with the now available crystal structures, which we and other groups have determined in the meantime [Protein Data Bank (PDB) entries 3WWH, 4CE5, 4CHI, 4CMD, and 6FTE].[15−19] These structures enable identification of key residues for (*R*)-amine coordination in the fold type IV class and eventually allow light to be shed on the evolutionary origin of (*R*)-ATAs. Key differences between (*R*)-ATAs and α-AATs were observed in the shape and polarity of the active site and the substrate coordination. On the basis of this structural information, we used a combination of sequence analysis and computational protein design studies to introduce (*R*)-ATA activity into the scaffold of an α-AAT from the fold type IV class of PLP-dependent enzymes. We chose the BCAT from *Escherichia coli* (*e*BCAT) and the DATA from *Bacillus subtilis* as α-AAT scaffolds, because these enzymes are well-studied, they do not show activity toward the typical ATA benchmark substrate 1-phenylethylamine (PEA), and their crystal structures had been determined (PDB entries 1IYE and 3DAA, respectively).[20,21]

RESULTS AND DISCUSSION

A Single Mutation Generates Initial (*R*)-ATA Activity on the *e*BCAT Scaffold. The *e*BCAT serves as a potential scaffold for the introduction of (*R*)-ATA activity, because the overall active site is composed like those in (*R*)-ATAs. As shown in Figure 1, the active site of both transaminases

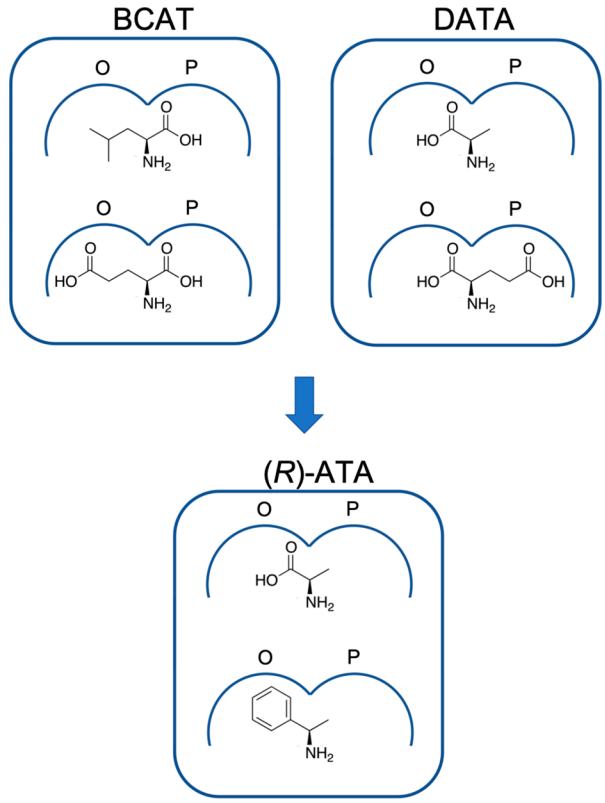


Figure 1. Pocket model of the active sites and the corresponding substrates of BCAT, DATA, and (*R*)-ATA.

consists of two binding pockets. The large binding pocket (O-site; active site region above PLP's hydroxyl group) of the *e*BCAT and (*R*)-ATAs can coordinate nonpolar alkyl and aryl groups, as well as carboxylic acids (the "dual-substrate recognition" mentioned above). Despite their similar architecture, the active site residues and properties differ according to their functions. The small binding pocket (P-site) of the *e*BCAT is restricted to accommodate the polar carboxylate function of L-amino acids, whereas the small binding pocket (P-site) of (*R*)-ATAs is limited to small nonpolar methyl or ethyl substituents.

To introduce the acceptance of (*R*)-PEA into the *e*BCAT scaffold, the mutation of the restricted P-site should be prioritized. Due to the reported activity toward L-phenylglycine, it can be assumed that the O-site can accept the phenyl

function of (*R*)-PEA.[22] Because the α-carboxylate in the P-site of *e*BCAT is coordinated without the direct coordination by basic amino acid side chains (further details are described in the Supporting Information), the acceptance of methyl substituents should be feasible by minor mutations in the hydrogen bond network. The substitution of the α-carboxylate function from an L-amino acid by a methyl group would already result in an (*R*)-amine as a switch in the designation of the absolute configuration takes place according to the CIP nomenclature.

Therefore, residues G38, R40, Y95, and R97 were chosen for mutagenesis in *e*BCAT, because these residues can account for the subtle α-carboxylate coordination in the P-site. To introduce the acceptance of (*R*)-PEA lacking the mandatory carboxylate function, the corresponding amino acids of the (*R*)-ATA from *Aspergillus fumigatus* were introduced, resulting in the G38V, R40S, Y95F, and R97E single mutants. Unfortunately, none of the single mutants showed detectable activity in the classical acetophenone assay with (*R*)-PEA as the amino donor and pyruvate as the acceptor.[23] Notably, PMP formation by the *e*BCAT R97E variant was observed, once the amino donor was added to the purified enzyme solution. This is evidence that, although no product formation was detected in the acetophenone assay, the first half-reaction was completed, and the amino donor is in principle accepted as it has reacted with the PLP cofactor to form the PMP. To enable catalysis, α-ketoisocaproate (the α-keto acid corresponding to the natural substrate leucine) was used as the amino acceptor instead of pyruvate. We were pleased to observe in this case an initial activity of 3.5 milliunits (mU) mg^{-1} in the spectrophotometric assay. Molecular modeling analysis revealed that the introduction of the R97E mutation created more space in the O-site, which could favor the binding of the phenyl ring in this pocket and the binding of the methyl group in the P-site (Figure S17). The R97E mutation caused the loss of activity in the native *e*BCAT reaction (with L-leucine and α-ketoglutarate in the glutamate dehydrogenase assay), because only 0.8% of WT activity (130 mU mg^{-1}) was detected. Unfortunately, further attempts to increase the activity of the R97E variant based on rational design or random mutagenesis did not lead to better variants (see the Supporting Information for details). Therefore, we considered using the D-amino acid transferases (DATA) as an alternative and complementary scaffold to the *e*BCAT.

Symmetrically with Respect to *e*BCAT, Y88 Is the Key Position for Allowing Initial (*R*)-ATA Activity in DATA. The D-amino acid transaminase (DATA, PDB entry 3DAA) from *B. subtilis* catalyzes the transfer between several D-amino acids and their corresponding α-keto acids. D-Alanine serves as the native amino donor for the synthesis of D-glutamate, both essential components for bacterial cell wall synthesis.[21] Therefore, the substrate specificity of only one of the two half-reactions needs to be altered. In comparison to our engineering efforts with BCAT, the challenge with DATA is that the coordination of the substrate's α-carboxylate must not be completely destroyed but needs to be modified in a way that both types of substrates, with or without an α-carboxylate, will be accepted.

The O-site of the DATA is exclusively coordinating the α-carboxylate function of the substrates by the polar side chains of residues Y31, R98, and H100. Therefore, Peisach et al. introduced the term "carboxylate trap" for this pocket.[21] The R98 side chain seems to have the strongest impact on D-alanine coordination, because the arginine residue forms a salt bridge with the α-carboxylate of alanine. Mutational studies by Kishimoto et al. showed that the R98M mutation yielded decreases in the specific activity (V_{max}) toward D-alanine and α-ketoglutarate of 4 orders of magnitude.[24] Besides R98, Y31 was also shown to drastically impact the D-alanine acceptance by Barber et al.[25] H100 also participates in the coordination of the α-carboxylate and completes the carboxylate trap of the DATA in the O-site.

The P-site harbors the methyl side chain of D-alanine and the γ-carboxylate function of D-glutamate. The P-site is formed by residues V33, S240, and T242.[26] These residues create a rather small binding pocket, but it is large enough to accommodate a wide variety of D-amino acid side chains.[21] Whether the glutamate's γ-carboxylate is coordinated by active site amino acids or remains in the solvent is currently not known.

To introduce (*R*)-PEA activity into the DATA scaffold, we chose the first coordination sphere of the active site as the target residues for mutation. In *e*BCAT, the R97E variant was the only beneficial one, leading to the introduction of (*R*)-PEA acceptance and hence activity. The corresponding position in the DATA scaffold is Y88, and we speculated that a substitution with glutamate could also have a beneficial effect here. Additionally, we sought to install a similar dual-substrate recognition mechanism as observed in ω-transaminases such as lysine-ε-aminotransferase (PDB entries 2CJD and 2CJH). In this enzyme, a positively charged guanidino group from R422 coordinates the α-carboxylate of α-ketoglutarate. To accept substrates lacking the α-carboxylate (e.g., the terminal amino group of lysine), a flexible glutamate residue E243 changes its orientation into a position where it neutralizes the positive charge of the arginine's guanidino group.[2,27] We speculated that a similar mechanism should also work in our DATA scaffold for the dual-substrate recognition (see also Figure S16). For this, the R98 residue could "switch" toward the introduced glutamate at position 88, coordinate with the γ-carboxylate for balancing the positive charge, and thus allow a nonpolar phenyl function to be coordinated in the O-site. Indeed, the glutamate substitution resulted in 12 mU mg^{-1} of activity in the acetophenone assay. Other single mutations or combinations did not further increase the (*R*)-PEA activity. Mutating residues F26 and Y31 also led to initial activities in the acetophenone assay (see Table S4).

On the basis of the gained knowledge that the DATA scaffold in principle can evolve toward (*R*)-PEA and its active site tolerates several mutations, we decided to follow a computer-aided enzyme design approach combined with amino acid co-evolution analysis. This should guide the selection of most beneficial amino acid mutations to broaden the substrate selectivity of the DATA. Here, the O-site was prioritized in our engineering strategy, because a more promiscuous binding pocket would be needed for the accommodation of the phenyl moiety of (*R*)-PEA.[28] This could be achieved by mutating the carboxylate trap forming residues Y31, R98, and H100 for (*R*)-PEA acceptance while keeping the D-alanine coordination intact.

The Combination of Computational Protein Design and Sequence Analysis Led to Wild Type-like (*R*)-ATA Activity in DATA. Starting from the crystal structure of D-amino acid aminotransferase determined in complex with the natural substrate D-alanine (PDB entry 3DAA),[21] we constructed a three-dimensional (3D) model of the enzyme

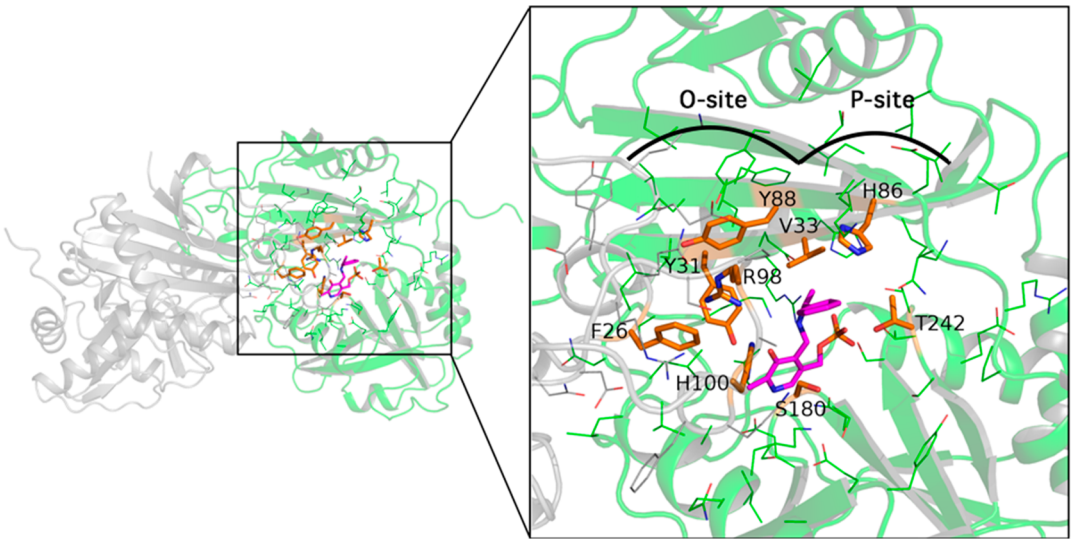


Figure 2. View of the overall three-dimensional organization of D-amino acid aminotransferase. On the left, the dimeric structural organization is shown with protein as a cartoon (green and gray). The right panel is a close-up of the catalytic site showing the mutable residues (orange sticks) and the flexible ones (green and gray lines) during RosettaDesign. In each panel, the PLP-(*R*)-phenylethylamine is represented as magenta sticks.

in complex with our substrate of interest, the PLP-(*R*)-phenylethylamine. Next, we undertook the computational redesign of the enzyme's active site using RosettaDesign[29] to adapt it to the recognition of PLP-(*R*)-phenylethylamine. Upon visual analysis of the amino acid residues surrounding PLP-(*R*)-phenylethylamine, we selected nine mutable (or designable) positions: Y31, V33, H86, Y88, S180, and T242 from the main subunit of the catalytic site (green cartoon in Figure 2) and F26, R98, and H100 from the adjacent subunit (gray cartoon in Figure 2). The remaining amino acid residues located within an 8 Å sphere centered on the ligand were defined as flexible through a rotamer library to allow side chain rearrangements during the design procedure. To explore combinatorial space, we then considered an enzyme design protocol based on the sampling of the nine positions by performing 20,000 independent runs of Rosetta *Enzyme_design*, which takes into account the backbone and side chain flexibility of "mutable" and "flexible" residues. The 3D models corresponding to the 20,000 designed sequences were subsequently rescored toward PLP-D-alanine derived from the crystallographic complex. For each complex, the total Rosetta score was plotted against the ligand energy contribution (ligand Rosetta score) for the PLP-(*R*)-phenylethylamine and the PLP-D-alanine ligands (Figure S15). The top-scoring sequences toward both ligands (corresponding to the bottom left quadrants) were further analyzed, leading to the identification of a small library of common sequences, of which the five top-scoring ones were selected for experimental testing. These sequences were found to contain seven or eight mutations, essentially composed of aromatic and aliphatic amino acid substitutions (Table S5).

Unfortunately, none of the proposed variants [M1−M5 (Table S5)] showed the desired (*R*)-PEA activity in the acetophenone assay with pyruvate as the amine acceptor. Interestingly, variant M2 exhibited some (*R*)-PEA acceptance when we assayed PMP formation after incubating the purified variant with (*R*)-PEA, which indicated a completed first half-reaction. The lack of activity in the acetophenone assay can be explained by the fact that the mutations might have had a detrimental effect on D-alanine coordination.

To identify mutations with a detrimental effect regarding the pyruvate acceptance from computational design results, we refined the variant M2 based on 3DM alignment statistics (Figures S3−S11). 3DM is a database that aligned sequences on the basis of their superfamily structure.[2,30] This allows the comparison of sequences with a low degree of sequence identity, because the alignment is based on the 3D position of the respective residue. The "D-aminotransferase PLP (2015)" database employed here includes 21,980 aligned sequences. Interestingly, only a few sequences containing W98 or W100 and no sequence harboring F88 were found. After these three mutations had been removed from the M2 variant, the resulting variant M2-4 (Y31F/H86F/S180A/T242I) resulted in an activity of 18 mU mg^{-1} (Table 1). Despite this still low activity, the variant did not show fast PMP formation upon incubation with (*R*)-PEA as observed with M2. However, when the Y88F substitution was included again [resulting in variant M2-5 (Y31F/H86F/Y88F/S180A/T242I)], rapid (*R*)-PEA-based PMP formation was restored and the activity increased to 185 mU mg^{-1} for M2-5. This clearly highlights the

Table 1. Activities of DATA Variants and Apparent Melting Temperatures

variant	(*R*)-amine activity[a] (mU mg^{-1})	native activity[b] (mU mg^{-1})	app. T_m (°C)
WT	0.2 ± 0.1	610 ± 30	70.2
Y31F	8 ± 1	208 ± 5	69.9
H86F	0.2 ± 0.1	nm	nm
Y88E	12 ± 1	nm	62.9
Y88F	0.6 ± 0.1	nm	nm
S180A	0 ± 1	nm	nm
M2-3	41 ± 4	210 ± 20	64.4
M2-4	18.3 ± 0.1	nm	69.9
M2-5	185 ± 7	17 ± 1	65.7
M2-6	326 ± 3	9 ± 4	69.1
M2	0.3 ± 0.1	6 ± 4	nm

[a]Initial activity measured in the acetophenone assay [(*R*)-PEA and pyruvate]. [b]Initial activity measured in the D-amino acid oxidase assay with D-glutamate and pyruvate. nm, not measured.

beneficial effect of Y88F on the (*R*)-PEA acceptance, similar to the results obtained for the *e*BCAT scaffold. Interestingly, the Y88F mutation had no activity as a single variant and is not present in any variant in the 3DM data set (Figure S7). A further increase in activity was achieved by combining the M2-5 variant with the H100L mutation, which is also present in 15% of the aligned sequences in the superfamily of the 3DM data set (Figure S9). The activity of this final variant is 326 mU mg^{-1} [M2-6 (Y31F/H86F/Y88F/H100L/S180A/T242I)] (Table 1). We were pleased to find that the incorporation of the six mutations of M2-6 had no detrimental influence on the apparent melting point of the DATA enzyme (Table 1).

Interestingly, the introduced mutations beneficial for the (*R*)-amine activity gradually decreased the activity toward the native substrate D-glutamate (Table 1). Whereas variant M2-3 (Y31F/H86F/Y88F) shows significant activity toward (*R*)-PEA and retains 34% of the native activity, the final variant M2-6 shows almost complete depletion of the native DATA activity but high activity toward (*R*)-PEA.

In all of these variants, residue R98 was kept unmutated, because this residue is probably the most important one in the "carboxylate trap" to enable D-alanine acceptance.[24] Additionally, the side chain of R98 lies in a position similar to that of R126 in *A. fumigatus*, which was shown to be an important residue for mediating the dual-substrate recognition and D-alanine acceptance.[7]

Overall, the mutations introduced into the M2-6 variant were of a nonpolar nature (aromatic and aliphatic amino acid side chains), leading to a notable increase in the active site volume with respect to that of the parental enzyme, favoring thus binding of the bulkier and more hydrophobic (*R*)-phenylethylamine compared to D-alanine.

To better understand the increased activity of variant M2-6 at molecular and energetic levels, binding free energy calculations were performed using the MM/PBSA method on the wild type enzyme and the M2-6 variant in complex with PLP-D-alanine or PLP-(*R*)-PEA (Figure 3 and Table S6). The contribution of the active site amino acid residues to the binding free energy is shown in Figure 3, whereas Table S6 lists all key residues (hot spots).

These calculations predicted that the PLP-D-alanine binding free energy is nearly of the same order of magnitude for both wild type and M2-6 enzymes (−50 ± 11 and −41 ± 9 kcal mol^{-1}, respectively). Consistent with the experimental results, the binding free energy value predicted for PLP-(*R*)-PEA in M2-6 was found to be −52 ± 10 kcal mol^{-1}, while it was predicted to be only −26 ± 7 kcal mol^{-1} in the wild type enzyme. Analysis of the amino acid contribution to the binding free energy revealed that most of designable residues (Y31, R98, H100, and T242) contribute in fact unfavorably to PLP-(*R*)-PEA binding in the wild type enzyme (Figure 3 and Table S6). Interestingly, analysis of MD simulation also suggested that the torsion of the ketimine (rectangle in Figure 3B) could be essential for favoring a catalytically productive orientation of the substrate. These results appear to be consistent with the reaction mechanism in which the ketimine is converted to the quinonoid (planar structure).

CONCLUSIONS

The ubiquitous α-AATs are believed to be the evolutionary origin of ATAs. Therefore, we studied two members of the α-AAT of fold type IV of PLP-dependent enzymes, *e*BCAT and DATA, and their respective variants for the acceptance of the benchmark amine substrate (*R*)-PEA. The *e*BCAT was initially thought to be a suitable scaffold for the acceptance of (*R*)-PEA. We hypothesized that with a change in the P-site in the active site, the acceptance of a methyl function instead of the α-carboxylate (of the L-amino acid substrate) would lead to an (*R*)-ATA. However, the single mutation R97E was enabling only a very low level of (*R*)-PEA acceptance while depleting the native *e*BCAT activity. A further increase in activity by additional mutations could not be achieved. In contrast, we were able to evolve the DATA toward a moderately active (*R*)-ATA by mutating residues in the active site. By employing an *in silico* approach combining computational protein design and sequence co-evolution analysis, we finally identified a DATA variant harboring six mutations, yielding a total activity of 326 mU mg^{-1} toward (*R*)-PEA, while maintaining overall stability. Although from our experiments it is not possible to trace back the exact way in which ATAs have occurred during natural evolution, the shown DATA engineering deepens our understanding of how substrate specificity in α-AAT is affected and can be changed toward accepting amines. In addition, our results are in agreement with the evolutionary IAD model (innovation−amplification−divergence).[31] According to this theory, the original enzyme-encoding gene acquires mutations that introduce side activities (innovation) while maintaining its original function. This leads to a promiscuous enzyme, because it is catalyzing the original reaction (maybe with decreased activity) and the newly acquired side activity. In our study, variants like Y31F and M2-3 showed promiscuous activity. The divergence toward M2-6 is accompanied by the depletion of the native DATA activity.

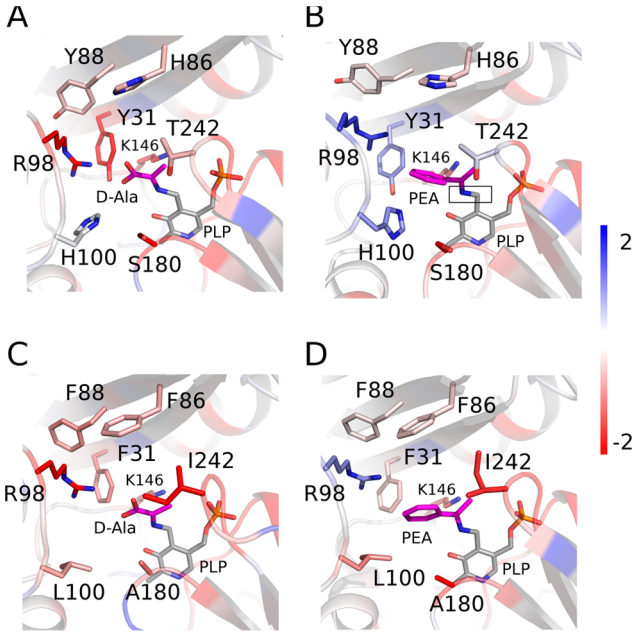


Figure 3. Binding free energy prediction using the MM/PBSA method. The contribution for each amino acid residue is projected on the 3D structures of the four complexes, i.e., wild type complexed with (A) PLP-D-alanine or (B) PLP-(*R*)-phenylethylamine and M2-6 complexed with (C) PLP-D-alanine or (D) PLP-(*R*)-phenylethylamine. The black rectangle in panel B highlights the torsion of the imine (ketimine form). Each snapshot was taken from the end of the MD simulation (5 ns), and the free energy values are given in kilocalories per mole.

METHODS

Plasmid Constructs. The *e*BCAT-encoding gene (GenBank entry YP_026247.1) was subcloned in the pET28b vector between the NdeI and XhoI restriction sites, resulting in a C-terminal His-tagged variant. The DATA gene (GenBank entry AAA22252.1) was cloned in the pGASTON vector between the NdeI and BamHI restriction sites with a C-terminal His tag. The D-amino acid oxidase from the *Trigonopsis variabilis* (*Tv*_DAAO; GenBank entry CAA90322.1) gene was ordered codon-optimized (for *E. coli* codon usage) from BioCat GmbH cloned into the pETDuet-1 vector between the EcoRI and HindIII restriction sites in the multiple cloning site 1 with an N-terminal His tag. The open reading frames and translated protein sequences of the three genes are given in the Supporting Information.

Site-Directed Mutagenesis. The creation of *e*BCAT and DATA mutants was performed using the QuikChange protocol (Agilent) using complementary primers with the desired codon and 3′ and 5′ homologue flanking regions. The primers were designed by using the standard settings of the web-based program PrimerX and are listed in Tables S1 and S2. For the polymerase chain reaction (PCR), 0.25 ng μL^{-1} template plasmid (carrying the DATA or eBCAT gene), 0.5 μM forward and reverse primers, 0.2 mM dNTPs, and 16 mU μL^{-1} *PfuPlus*! DNA polymerase (roboklon) in *Pfu* reaction buffer were used. The PCR was performed as follows: (a) 95 °C for 2 min, (b) 25 cycles at 95 °C for 30 s, 63 °C for 1 min, and 72 °C for 1 min kbp^{-1}, and (c) 72 °C for 10 min. The resulting PCR product was directly digested with the DpnI restriction enzyme (20 $\mu L\ mL^{-1}$; NEB) for 2 h at 37 °C followed by subsequent heat inactivation at 80 °C for 20 min. The digested PCR product was transformed in chemically competent *E. coli* TOP10 cells (Invitrogen), and single colonies were sequenced (Eurofins MWG GmbH) to verify the introduced mutation. The introduction of additional mutations was realized by iterating the described protocol.

Protein Expression and Purification of the Enzyme Variants. For the expression of the *e*BCAT and DATA variants, the plasmid constructs were transformed into *E. coli* BL21(DE3) cells, and these were incubated overnight at 37 °C in a 5 mL LB medium (Lysogeny Broth) preculture supplemented with the respective antibiotics (100 $\mu g\ mL^{-1}$ ampicillin or 50 $\mu g\ mL^{-1}$ kanamycin for the *e*BCAT and DATA, respectively). One milliliter of the preculture was used for the inoculation of 100 mL of TB medium (supplemented with the corresponding antibiotic) and incubated at 37 °C and 180 rpm. The expression of the *e*BCAT and DATA variants was induced at an optical density of approximately 0.6 at 600 nm with 0.2% L-rhamnose or 0.2 mM isopropyl β-D-thiogalactopyranoside (IPTG), respectively, and each was incubated at 20 °C overnight. The *Tv*_DAAO gene was expressed using the described protocol of Barber et al. at 17 °C overnight in TB medium.[25] The cells were harvested by centrifugation (20 min at 4,000 *g* and 4 °C).

For purification, the *e*BCAT and DATA variants were resuspended in 50 mM HEPES buffer (pH 7.5) containing 0.1 mM PLP, 0.3 M NaCl, and 0.01 M imidazole and lysed via ultrasonic treatment (50% pulse, 50% power, 2 × 5 min; Sonoplus HD2070, Bandelin Electronic GmbH). The lysate was clarified by centrifugation (1 h at 10,000 *g* and 4 °C) and purified by immobilized metal affinity chromatography with the following buffers: washing buffer [50 mM HEPES buffer (pH 7.5) containing 0.1 mM PLP, 0.3 M NaCl, and 0.02 M imidazole] and elution buffer [50 mM HEPES buffer (pH 7.5) containing 0.1 mM PLP, 0.3 M NaCl, and 0.3 M imidazole]. The protein was desalted in 50 mM HEPES buffer (pH 7.5) and 0.1 mM PLP using the PD-10 desalting column (GE Healthcare). The *Tv*_DAAO was lysed and purified in an analogous manner, but in 50 mM phosphate buffer (pH 8.0) with 20 μM FAD^+. The purified and desalted proteins were stored at −20 °C in 30% glycerol.

Acetophenone Assay. The activities of the purified *e*BCAT or DATA variants regarding the conversion of 1-phenylethylamine were determined photometrically by applying the acetophenone assay on an Infinite 200 PRO (TECAN) or FLUOstar Omega (BMG LABTECH GmbH) plate reader in UV-transparent microtiter plates (UV-Star, Greiner Bio-One GmbH).[23] The assay was performed with 2.5 mM (*S*)- or (*R*)-PEA as amine donors and 2.5 mM pyruvate or 2.5 mM α-ketoisocaproate as amine acceptors in 1.25−2.5% dimethyl sulfoxide (DMSO) and 50 mM CHES buffer (pH 9.0) at 30 °C. The formation of acetophenone was quantified by following the increase in absorption at 245 nm over time. One unit (U) was defined as the formation of 1 μmol of acetophenone per minute. All measurements were performed in triplicate.

Glutamate Dehydrogenase Assay. The activities of the purified *e*BCAT variants regarding the form of NADH were determined photometrically by applying the glutamate dehydrogenase assay on the FLUOstar Omega (BMG LABTECH GmbH) plate reader in microtiter plates (200 μL assay volume). The assay was performed with 2.5 mM amino donor (L-leucine), 2.5 mM α-ketoglutarate as the amine acceptor, 5 U mL^{-1} glutamate dehydrogenase (Sigma-Aldrich), 500 μM NAD^+, and 50 mM CHES buffer (pH 9.0). The formation of NADH was detected by following the increase in absorption at 340 nm over time.[32]

D-Amino Acid Oxidase Assay. The native activity of the purified DATA variants was measured using the D-amino acid oxidase assay described by Barber et al.[25] The reaction was performed in microtiter plates (200 μL assay volume) with 2.5 mM D-glutamate as the amine donor and 2.5 mM pyruvate as the amine acceptor, 0.7 U mL^{-1} D-amino acid oxidase from *T. variabilis* (*Tv*_DAAO), 22 U mL^{-1} HRP (Sigma-Aldrich), and 50 μM Ampliflu Red in 50 mM CHES buffer (pH 9.0) with 0.5% DMSO and 2.5% ethanol as co-solvents. The formation of Resorufin was quantified by following the increase in absorption at 560 nm over time on the FLUOstar Omega (BMG LABTECH GmbH) plate reader.

Analysis of the Half-Reaction by PMP Formation. The formation of PMP was determined photometrically on the FLUOstar Omega (BMG LABTECH GmbH) plate reader in microtiter plates (200 μL assay volume). The reaction was performed with 50 μL of purified transaminase, 2.5 mM (*R*)-PEA in 1.25% ethanol, and 50 mM CHES buffer (pH 9.0). The formation of PMP was detected by following the change in absorbance from 290 to 470 nm.

Melting Point Determination. The apparent melting points of most variants were measured with the nanoDSF instrument Prometheus NT.48 (NanoTemper) and represents the inflection point of the tryptophan fluorescence ratio (350 nm to 330 nm) with a 0.1 °C min^{-1} temperature ramp from 20 to 95 °C. For the analysis, the purified enzyme at a concentration of 1 mg mL^{-1} in 50 mM HEPES buffer (pH 7.5) with 0.1 mM PLP was used. The apparent melting points of the DATA variants are listed in Table 1 and Table S4.

Computational Enzyme Design. The crystal structure of the D-amino acid aminotransferase (DATA) from *Bacillus* sp. YM-1 (PDB entry 3DAA)[21] was used as a scaffold for the computational enzyme redesign. This crystal structure was determined in the dimeric form with a high resolution of 1.90 Å, containing 277 amino acids for each subunit. All missing residues, only the N- and C-termini, compared to the UniProt sequence (P19938), were modeled using Modeler 9.19.[33]

The crystal structure also formed a complex with PLP-D-alanine. The enzyme complexed with PLP-(*R*)-PEA was built in two steps. (i) PLP-(*R*)-PEA was built and geometrically optimized using the Avogadro software,[34] and (ii) PLP-(*R*)-PEA was superimposed on the enzyme−PLP-D-alanine complex. For both PLP-D-alanine and PLP-(*R*)-PEA, RESP charges were obtained after fitting partial charges (ESP) computed from Hartree−Fock theory (HF/6-31G*) by using Gaussian G09.[35] RESP charges were included in the ligand parameters for further calculations (Figure S14).

Only the PLP-(*R*)-PEA complex was used for design using Rosetta 3.9. A fast relaxation (energy minimization) was first carried out on the protein side chains and the ligand [PLP-(*R*)-phenylethylamine]. Then, the design was performed on nine selected mutable residues belonging to the first shell of ligand coordination: Y31, V33, H86, Y88, S180, and T242 from the main subunit (catalytic site) and F26, R98, and H100 from the adjacent subunit. All other residues surrounding the ligand in a sphere of 8 Å were considered flexible (through the use of a rotamer library), whereas the rest of the protein was fixed. Harmonic constraints were applied to maintain the nitrogen

of the catalytic lysine (K146) close (~3 Å) to the nitrogen bonding the phenylethylamine to PLP during the relaxation and the design. From these calculations, 20,000 amino acid sequences and their corresponding 3D structures were extracted as output. For each corresponding 3D structure (20,000), superimposition of PLP-D-alanine was performed to build complexes (enzyme−PLP-D-alanine). Fast relaxation was performed as previously described to rearrange the side chains around the new ligand, and the new complexes [enzyme−PLP-(*R*)-PEA] were scored with the same energy function as the enzyme−PLP-D-alanine complex.

MD Simulations and Molecular Mechanics Poisson−Boltzmann Surface Area (MM/PBSA) Calculations. The mutant models (M2-6) were built using Modeller9.19.[33] Models with the lowest DOPE score[36] were kept for further MD simulations. All complexes were first minimized in vacuum to release steric clashes. MD simulations were performed using the AMBER ff14SB[37,38] force field for enzymes and GAFF for the ligands using pmemd.CUDA[39] of AMBER16 software. The system was protonated using the propka web server to set the experimental pH to 9. The catalytic lysine was kept neutral as expected in the reaction mechanism with the ketimine. MD simulations were carried out at a constant temperature (300 K) and a constant pressure (1 bar) using the Berendsen algorithm.[40] The integration time step was 2 fs, and covalent bonds involving hydrogens were constrained using SHAKE.[41] The nonbonded pair list was updated heuristically. Long-range electrostatic interactions were treated using the particle mesh Ewald (PME)[42] approach. Nonbonded interactions were treated with a 9 Å direct space cutoff. All enzyme systems were neutralized with Na^+ ions (minimal salt condition),[43] in explicit TIP3P water molecules.[44] Periodic boundary conditions (PBCs) were applied on primary cubic boxes having minimal distances from the dimers of 10 Å. The water molecules and counterions were energy-minimized and equilibrated at 100 K around the constrained solute for 100 ps in the *NVT* ensemble; the entire system was then heated incrementally over 100 ps from 100 to 300 K in 5 K steps with harmonic positional restraints of 25.0 kcal mol^{-1} $Å^{-2}$ on the solute atoms. The MD simulations were continued in the *NPT* ensemble. The positional restraints were gradually removed over 250 ps and followed by the production phase. In all stages, the imine bond (N−C) was kept close to the catalytic lysine by setting a distance constraint of 3.5 Å using a force constant of 50 kcal mol^{-1} $Å^{-2}$. MD snapshots were saved every 10 ps. The MD simulations were carried out for a total of 5 ns for all complexes.

MM/PBSA calculations were performed using MMPBSA.py software and default parameters for Poisson−Boltzmann as described by Miller et al.[45] The calculations were performed on 20 snapshots equitably extracted between 1 and 5 ns.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acschembio.9b00888.

A description of additional *e*BCAT protein engineering studies, the nucleotide and protein sequences of the *e*BCAT, DATA, and *Tv*_DAAO enzymes, lists of mutants, and energies from the *in silico* experiments, and figures of 3DM alignments and molecular modeling (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Uwe T. Bornscheuer − *Department of Biotechnology & Enzyme Catalysis, Institute of Biochemistry, Greifswald University, 17487 Greifswald, Germany;* orcid.org/0000-0003-0685-2696; Phone: (+49) 3834 420 4367; Email: uwe.bornscheuer@uni-greifswald.de

Authors

Moritz Voss − *Department of Biotechnology & Enzyme Catalysis, Institute of Biochemistry, Greifswald University, 17487 Greifswald, Germany*

Chao Xiang − *Department of Biotechnology & Enzyme Catalysis, Institute of Biochemistry, Greifswald University, 17487 Greifswald, Germany*

Jérémy Esque − *Toulouse Biotechnology Institute (TBI), Université de Toulouse, CNRS, INRA, INSA, F-31077 Toulouse, France*

Alberto Nobili − *Department of Biotechnology & Enzyme Catalysis, Institute of Biochemistry, Greifswald University, 17487 Greifswald, Germany*

Marian J. Menke − *Department of Biotechnology & Enzyme Catalysis, Institute of Biochemistry, Greifswald University, 17487 Greifswald, Germany*

Isabelle André − *Toulouse Biotechnology Institute (TBI), Université de Toulouse, CNRS, INRA, INSA, F-31077 Toulouse, France;* orcid.org/0000-0001-6280-4109

Matthias Höhne − *Protein Biochemistry, Institute of Biochemistry, Greifswald University, 17487 Greifswald, Germany;* orcid.org/0000-0002-2542-725X

Complete contact information is available at: https://pubs.acs.org/10.1021/acschembio.9b00888

Author Contributions

M.V. and C.X. contributed equally to this work. U.T.B. and M.H. initiated and supervised the project. The computational design and molecular modeling of enzymes were performed by I.A. and J.E. Initial experiments were performed by A.N. M.V., C.X., A.N., and M.J.M. created and biochemically characterized all enzyme variants. U.T.B., M.H., I.A., M.V., C.X., and J.E. analyzed the results. U.T.B., M.H., and M.V. drafted the manuscript, and M.V., J.E., I.A., M.H., and U.T.B. wrote the manuscript. The manuscript was revised and approved by all authors.

Funding

The authors are grateful for funding by the DFG (Grants HO 4754/3-1 and BO 1862/16-1) and the China Scholarship council for a Ph.D. stipend to C.X. (Grant 201808330394). A.N. especially thanks the European Union (KBBE-2011-5, Grant 289350) for financial support within the European Union Seventh Framework Programme.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are grateful to Y. Tao and M. Stricker for their help in studying the *e*BCAT enzyme. This work was granted access to the HPC resources on the TGCC-Curie and Occigen supercomputers and the Computing mesocenter of Région Midi-Pyrénées (CALMIP, Toulouse, France).

■ ABBREVIATIONS

BCAT, L-branched chain amino acid aminotransferase; DATA, D-amino acid aminotransferase; PEA, 1-phenylethylamine

■ REFERENCES

(1) Eliot, A. C., and Kirsch, J. F. (2004) Pyridoxal Phosphate Enzymes: Mechanistic, Structural, and Evolutionary Considerations. *Annu. Rev. Biochem. 73*, 383−415.

(2) Steffen-Munsberg, F., Vickers, C., Kohls, H., Land, H., Mallin, H., Nobili, A., Skalden, L., van den Bergh, T., Joosten, H. J., Berglund, P., Höhne, M., and Bornscheuer, U. T. (2015) Bioinformatic Analysis of a PLP-Dependent Enzyme Superfamily Suitable for Biocatalytic Applications. *Biotechnol. Adv. 33*, 566−604.
(3) Henson, C. P., and Cleland, W. W. (1964) Kinetic Studies of Glutamic Oxaloacetic Transaminase Isozymes. *Biochemistry 3*, 338−345.
(4) Savile, C. K., Janey, J. M., Mundorff, E. C., Moore, J. C., Tam, S., Jarvis, W. R., Colbeck, J. C., Krebber, A., Fleitz, F. J., Brands, J., Devine, P. N., Huisman, G. W., and Hughes, G. J. (2010) Biocatalytic Asymmetric Synthesis of Chiral Amines from Ketones Applied to Sitagliptin Manufacture. *Science 329*, 305−309.
(5) Gomm, A., and O'Reilly, E. (2018) Transaminases for Chiral Amine Synthesis. *Curr. Opin. Chem. Biol. 43*, 106−112.
(6) Manta, B., Cassimjee, K. E., and Himo, F. (2017) Quantum Chemical Study of Dual-Substrate Recognition in ω-Transaminase. *ACS Omega 2*, 890−898.
(7) Skalden, L., Thomsen, M., Höhne, M., Bornscheuer, U. T., and Hinrichs, W. (2015) Structural and Biochemical Characterization of the Dual Substrate Recognition of the (*R*)-Selective Amine Transaminase from *Aspergillus Fumigatus*. *FEBS J. 282*, 407−415.
(8) Höhne, M., Schätzle, S., Jochens, H., Robins, K., and Bornscheuer, U. T. (2010) Rational Assignment of Key Motifs for Function Guides in Silico Enzyme Identification. *Nat. Chem. Biol. 6*, 807−813.
(9) Iwasaki, A., Yamada, Y., Kizaki, N., Ikenaka, Y., and Hasegawa, J. (2006) Microbial Synthesis of Chiral Amines by (*R*)-Specific Transamination with *Arthrobacter* sp. KNK168. *Appl. Microbiol. Biotechnol. 69*, 499−505.
(10) Boyko, K. M., Stekhanova, T. N., Nikolaeva, A. Y., Mardanov, A. V., Rakitin, A. L., Ravin, N. V., Bezsudnova, E. Y., and Popov, V. O. (2016) First Structure of Archaeal Branched-Chain Amino Acid Aminotransferase from *Thermoproteus Uzoniensis* Specific for L-Amino Acids and *R*-Amines. *Extremophiles 20*, 215−225.
(11) Bezsudnova, E. Y., Boyko, K. M., Nikolaeva, A. Y., Zeifman, Y. S., Rakitina, T. V., Suplatov, D. A., and Popov, V. O. (2019) Biochemical and Structural Insights into PLP Fold Type IV Transaminase from *Thermobaculum Terrenum*. *Biochimie 158*, 130−138.
(12) Zeifman, Y. S., Boyko, K. M., Nikolaeva, A. Y., Timofeev, V. I., Rakitina, T. V., Popov, V. O., and Bezsudnova, E. Y. (2019) Functional Characterization of PLP Fold Type IV Transaminase with a Mixed Type of Activity from *Haliangium Ochraceum*. *Biochim. Biophys. Acta, Proteins Proteomics 1867*, 575−585.
(13) Pavkov-Keller, T., Strohmeier, G. A., Diepold, M., Peeters, W., Smeets, N., Schürmann, M., Gruber, K., Schwab, H., and Steiner, K. (2016) Discovery and Structural Characterisation of New Fold Type IV-Transaminases Exemplify the Diversity of This Enzyme Fold. *Sci. Rep. 6*, 38183.
(14) Bezsudnova, E. Y., Dibrova, D. V., Nikolaeva, A. Y., Rakitina, T. V., and Popov, V. O. (2018) Identification of Branched-Chain Amino Acid Aminotransferases Active towards (*R*)-(+)-1-Phenylethylamine among PLP Fold Type IV Transaminases. *J. Biotechnol. 271*, 26−28.
(15) Guan, L. J., Ohtsuka, J., Okai, M., Miyakawa, T., Mase, T., Zhi, Y., Hou, F., Ito, N., Iwasaki, A., Yasohara, Y., and Tanokura, M. (2015) A New Target Region for Changing the Substrate Specificity of Amine Transaminases. *Sci. Rep. 5*, 10753.
(16) Sayer, C., Martinez-Torres, R. J., Richter, N., Isupov, M. N., Hailes, H. C., Littlechild, J. A., and Ward, J. M. (2014) The Substrate Specificity, Enantioselectivity and Structure of the (*R*)-Selective Amine: Pyruvate Transaminase from *Nectria Haematococca*. *FEBS J. 281*, 2240−2253.
(17) Thomsen, M., Skalden, L., Palm, G. J., Höhne, M., Bornscheuer, U. T., and Hinrichs, W. (2014) Crystallographic Characterization of the (*R*)-Selective Amine Transaminase from *Aspergillus Fumigatus*. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 70*, 1086−1093.
(18) Łyskowski, A., Gruber, C., Steinkellner, G., Schürmann, M., Schwab, H., Gruber, K., and Steiner, K. (2014) Crystal Structure of an (*R*)-Selective ω-Transaminase from *Aspergillus Terreus*. *PLoS One 9*, No. e87350.
(19) Telzerow, A., Paris, J., Håkansson, M., González-Sabín, J., Ríos-Lombardía, N., Schürmann, M., Gröger, H., Morís, F., Kourist, R., Schwab, H., and Steiner, K. (2019) Amine Transaminase from *Exophiala Xenobiotica* - Crystal Structure and Engineering of a Fold IV Transaminase That Naturally Converts Biaryl Ketones. *ACS Catal. 9*, 1140−1148.
(20) Goto, M., Miyahara, I., Hayashi, H., Kagamiyama, H., and Hirotsu, K. (2003) Crystal Structures of Branched-Chain Amino Acid Aminotransferase Complexed with Glutamate and Glutarate: True Reaction Intermediate and Double Substrate Recognition of the Enzyme. *Biochemistry 42*, 3725−3733.
(21) Peisach, D., Chipman, D. M., Van Ophem, P. W., Manning, J. M., and Ringe, D. (1998) Crystallographic Study of Steps along the Reaction Pathway of D-Amino Acid Aminotransferase. *Biochemistry 37*, 4958−4967.
(22) Bommer, M., and Ward, J. M. (2013) A 1-Step Microplate Method for Assessing the Substrate Range of l-α-Amino Acid Aminotransferase. *Enzyme Microb. Technol. 52*, 218−225.
(23) Schätzle, S., Höhne, M., Redestad, E., Robins, K., and Bornscheuer, U. T. (2009) Rapid and Sensitive Kinetic Assay for Characterization of ω-Transaminases. *Anal. Chem. 81*, 8244−8248.
(24) Kishimoto, K., Yoshimura, T., Soda, K., and Esaki, N. (1997) Mutation of Arginine 98, Which Serves as a Substrate-Recognition Site of D-Amino Acid Aminotransferase, Can Be Partly Compensated for by Mutation of Tyrosine 88 to an Arginyl Residue. *J. Biochem. 122*, 1182−1189.
(25) Barber, J. E. B., Damry, A. M., Calderini, G. F., Walton, C. J. W., and Chica, R. A. (2014) Continuous Colorimetric Screening Assay for Detection of D-Amino Acid Aminotransferase Mutants Displaying Altered Substrate Specificity. *Anal. Biochem. 463*, 23−30.
(26) Walton, C. J. W., Parmeggiani, F., Barber, J. E. B., McCann, J. L., Turner, N. J., and Chica, R. A. (2018) Engineered Aminotransferase for the Production of D-Phenylalanine Derivatives Using Biocatalytic Cascades. *ChemCatChem 10*, 470−474.
(27) Mani Tripathi, S., and Ramachandran, R. (2006) Direct Evidence for a Glutamate Switch Necessary for Substrate Recognition: Crystal Structures of Lysine ε-Aminotransferase (Rv3290c) from *Mycobacterium Tuberculosis* H37Rv. *J. Mol. Biol. 362*, 877−886.
(28) Li, R., Wijma, H. J., Song, L., Cui, Y., Otzen, M., Tian, Y., Du, J., Li, T., Niu, D., Chen, Y., Feng, J., Han, J., Chen, H., Tao, Y., Janssen, D. B., and Wu, B. (2018) Computational Redesign of Enzymes for Regio- and Enantioselective Hydroamination Article. *Nat. Chem. Biol. 14*, 664−670.
(29) Liu, Y., and Kuhlman, B. (2006) Rosetta Design Server for Protein Design. *Nucleic Acids Res. 34*, W235−W238.
(30) Nobili, A., Gall, M. G., Pavlidis, I. V., Thompson, M. L., Schmidt, M., and Bornscheuer, U. T. (2013) Use of "small but Smart" Libraries to Enhance the Enantioselectivity of an Esterase from Bacillus Stearothermophilus towards Tetrahydrofuran-3-Yl Acetate. *FEBS J. 280*, 3084−3093.
(31) Bergthorsson, U., Andersson, D. I., and Roth, J. R. (2007) Ohno's Dilemma: Evolution of New Genes under Continuous Selection. *Proc. Natl. Acad. Sci. U. S. A. 104*, 17004−17009.
(32) Steffen-Munsberg, F., Matzel, P., Sowa, M. A., Berglund, P., Bornscheuer, U. T., and Höhne, M. (2016) *Bacillus Anthracis* ω-Amino Acid:Pyruvate Transaminase Employs a Different Mechanism for Dual Substrate Recognition than Other Amine Transaminases. *Appl. Microbiol. Biotechnol. 100*, 4511−4521.
(33) Webb, B., and Sali, A. (2016) Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinf. 54*, 5.6.1−5.6.37.
(34) Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., and Hutchison, G. R. (2012) Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminf. 4*, 17.

(35) Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J. A., Jr., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, N. J., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, Ö., Foresman, J. B., Ortiz, J. V., Cioslowski, J., and Fox, D. J. (2009) *Gaussian 09*, Gaussian, Inc., Wallingford, CT.

(36) Shen, M., and Sali, A. (2006) Statistical Potential for Assessment and Prediction of Protein Structures. *Protein Sci. 15*, 2507−2524.

(37) Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. A. (1995) Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc. 117*, 5179−5197.

(38) Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput. 11*, 3696−3713.

(39) Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S., and Walker, R. C. (2013) Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput. 9*, 3878−3888.

(40) Berendsen, H. J. C. C., Postma, J. P. M. M., van Gunsteren, W. F., Dinola, A., and Haak, J. R. (1984) Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys. 81*, 3684−3690.

(41) van Gunsteren, W. F., and Berendsen, H. J. C. (1977) Algorithms for Macromolecular Dynamics and Constraint Dynamics. *Mol. Phys. 34*, 1311−1327.

(42) Darden, T., York, D., and Pedersen, L. (1993) Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys. 98*, 10089−10092.

(43) Åqvist, J. (1990) Ion-Water Interaction Potentials Derived from Free Energy Perturbation Simulations. *J. Phys. Chem. 94*, 8021−8024.

(44) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys. 79*, 926−935.

(45) Miller, B. R., McGee, T. D., Swails, J. M., Homeyer, N., Gohlke, H., and Roitberg, A. E. (2012) MMPBSA. Py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput. 8*, 3314−3321.