


Activity Levels of Amine Transaminases Correlate with Active Site Hydrophobicity

Manideep Kollipara¹, Philipp Matzel¹, Uwe Bornscheuer², and Matthias Höhne^{1,*}

DOI: 10.1002/cite.202200062

 This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.



Supporting Information
available online

Amine transaminases (ATAs) are biocatalysts for the synthesis of chiral amines and can be identified in sequence databases by specific sequence motifs. This study shows that the activity level towards the model substrate 1-phenylethylamine can be predicted solely from the sequence. To demonstrate this, 15 putative ATAs with a different distribution of hydrophobic or hydrophilic amino acid side chains near the active site were characterized. Hydrophobic side chains were associated with a high activity level and were a better predictor of activity than global sequence identity to known ATAs with high or low activities. Enzyme stability investigations revealed that four out of the 15 ATAs showed a good operational stability.

Keywords: Biocatalysis, Chiral amines, Enzyme discovery, Function prediction, Transaminase

Received: May 16, 2022; *revised:* July 01, 2022; *accepted:* August 10, 2022

1 Introduction

Enzymes often feature high regio- and stereoselectivity and thus are attractive catalysts for the production of chiral intermediates [1–4]. For example, amine transaminases (ATAs) are employed for manufacturing primary amines [5–11] in an asymmetric synthesis from ketones (Fig. 1), which are desired intermediates in the manufacture of a large number of active pharmaceutical ingredients. ATAs use pyridoxal-5'-phosphate (PLP) as cofactor (Fig. 1a) to facilitate a complex series of catalytic steps in the reaction mechanism [12, 13]. In the first half reaction, the amino group of an amine donor substrate is transferred to PLP, resulting in a ketone or α -keto acid by-product, which dissociates from the enzyme. During the second half reaction, the cofactor-bound amino group is then transferred to the ketone substrate, yielding the desired corresponding amine product. A more detailed Scheme on the mechanism is presented in Fig. S1 in the Supporting Information (SI). Although transaminases are ubiquitous in all organisms, only a small fraction of them accepts amines and ketones besides their natural amino acid substrates [14, 15]. Importantly, enantiocomplementary (*S*)- and (*R*)-selective ATAs have been evolved in two different superfamilies, the PLP fold type I and IV, and they can be recognized by specific amino acid patterns [16–19]. The characteristic motif of (*S*)-selective ATAs includes an arginine residue that is located near to the PLP cofactor [15]. Its side chain can adopt multiple conformations: it can be oriented towards the PLP cofactor, or away from the active site [20]. The first case is observed

when it coordinates the carboxyl group of a bound amino acid substrate such as alanine (Fig. 1b, left) by forming a salt bridge with its positively charged guanidino group. On the contrary, rotation away from the cofactor creates a large binding pocket suitable to accommodate a hydrophobic ketone or amine substrate (Fig. 1b, right) [21]. The different sizes of the binding pockets explain the stereopreference of ATAs, in case that the substrate bears two differently sized substituents, since one orientation of the substrate fits better to the shape of the pockets and thus is energetically favorable. Based on this principle, molecular docking algorithms combined with molecular modeling can be used to predict the enantioselectivity and substrate specificity of transaminases with known structures [22].

The approach of structure-guided analysis of amino acids located around the active site is very helpful to improve the accuracy of function prediction of amino acid sequences in databases, and it was applied to identify various ATAs from database searches [17]. Within the characterized proteins, ATAs with a large range of specific activities were found

¹Manideep Kollipara, Dr. Philipp Matzel, Dr. Matthias Höhne
Matthias.Hoehne@uni-greifswald.de

University of Greifswald, Institute of Biochemistry, Protein Biochemistry, Felix-Hausdorff-Straße 4, 17489 Greifswald, Germany.

²Prof. Uwe Bornscheuer

University of Greifswald, Institute of Biochemistry, Dept. of Biotechnology & Enzyme Catalysis, Felix-Hausdorff-Straße 4, 17489 Greifswald, Germany.

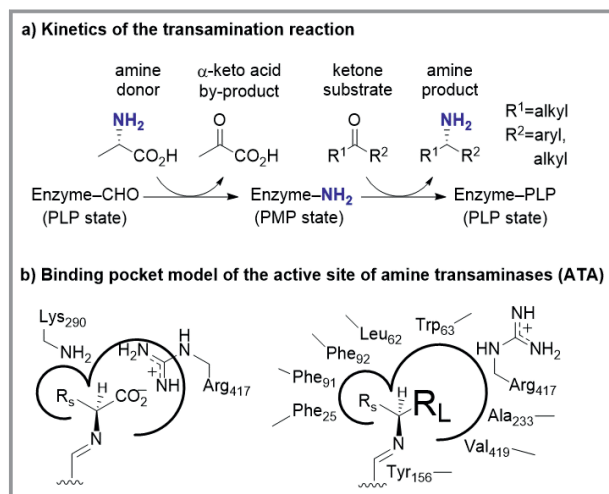


Figure 1. Schematic overview of the catalyzed reaction and substrate recognition of transaminases. a) The amino group shown in bold is first transferred to the enzyme and bound as pyridoxal-5'-amine phosphate (PMP) before it is passed on to the ketone to yield the amine product. b) The binding pockets of a (S)-selective ATA can either accept alanine (left) or an amine with two hydrophobic substituents of large and small size (R_L = alkyl, aryl; R_S = methyl or ethyl). Important residues that are part of the catalytic machinery (Lys290) or form the active site pockets are shown.

[15, 20]. An important remaining question is whether also the activity level can be predicted from the amino acid sequence as well, since this would aid the selection of potentially interesting new candidate enzymes during a database search. Although a quantitative and accurate prediction is currently not possible, the goal is to distinguish enzymes with medium to high-level activity from low active ones without applying complex computational or modeling techniques. In a previous study, two enzymes with very low specific activities were investigated in detail and compared to certain highly active ATAs [21]. From their sequence differences, the hypothesis was developed that the active site hydrophobicity is a key factor which affects the activity level. An active site mainly composed of hydrophobic side chains is associated with a high activity towards the standard substrate pair 1-phenylethylamine and pyruvate. This correlation was verified exemplarily by exchanging selected hydrophilic amino acids of the low active ATA from *Silicibacter* sp. TM1040 (TA-3FCR) by hydrophobic ones found in highly active enzymes, e.g., ATA from *Silicibacter pomeroyi* (TA-3HMU) [21]. As this hypothesis is derived from studying only one enzyme, the question remains whether this trend – the presence of hydrophobic side chains in the active site region facilitates a high specific activity and vice versa – can be generalized. Therefore, 15 putative transaminase sequences with a varying distribution of hydrophobic and hydrophilic amino acids surrounding the active site were selected from a protein database and characterized in this study. Besides the elucidation of their activities towards selected model substrates, the ATAs were also characterized

in terms of operational stability, since this property is also important when a suitable catalyst for a synthetic application has to be selected.

2 Materials and Methods

2.1 Chemicals and Biological Materials

All chemicals were purchased from Merck KGaA (Darmstadt, Germany), ABCR (Karlsruhe, Germany), Carl Roth (Karlsruhe, Germany), phage-resistant *Escherichia coli* BL21 (genotype: *fhuA2* [lon] *ompT gal* (λ DE3) [*dcm*] Δ *hsdS* [λ DE3 = λ *sBamHI*o Δ *EcoRI*-B *int*:: (*lacI*::*PlacUV5*::*T7 gene1*) *i21* Δ *nin5*) and the site-directed mutagenesis reagents were obtained from New England Biolabs (Ipswich, USA). The codon-optimized genes (see Supporting Information) encoding the putative transaminases TA-1H to TA-15F of this study were ordered from Genscript (New Jersey, USA) and obtained in the genscript IP free vector. The plasmids coding for TA-3FCR and TA-3HMU were available in our lab collection [20]. The oligonucleotides were ordered from Merck (Darmstadt, Germany).

2.2 Cultivation and Expression

The codon-optimized genes encoding the ATAs were cloned in the pET22b vector between the restriction sites *NdeI* and *BamHI* and verified by sequencing. *E. coli* BL21 (DE3) was transformed by the heat shock method with the plasmids containing the genes of interest. The protein expression was carried out in 1-L flasks using 250 mL Terrific Broth (TB) media with 100 μ g mL⁻¹ ampicillin at 37 °C and 160 rpm to an optical density OD₆₀₀ of 1.0 to 1.4, protein expression was then induced with IPTG (0.2 mM final concentration), and the cultures were shaken at 20 °C. After 18 h the cultures were centrifuged (4500 \times g, 20 min, 4 °C) and the supernatant was discarded.

2.3 Protein Purification

For purification, the cell pellets were washed with washing buffer (25 mL Buffer-A, 50 mM (HEPES), pH 8.0), and harvested (4500 \times g, 20 min, 4 °C). The obtained cell pellets were then resuspended in lysis buffer (25 mL Buffer-B, 50 mM HEPES, pH 8.0, 0.3 M NaCl, and 0.1 mM PLP). The cells were disrupted via ultrasonication for 5 min on ice followed by centrifugation to remove the cell debris (8500 \times g, 1 h, 4 °C). The supernatant was passed through a 0.45 μ m filter and the filtrate was loaded on a 5 mL Ni-NTA (nickel nitrilotriacetic acid) column (GE Health care), installed on an Äkta Purifier device. The column was preequilibrated with binding buffer (Buffer-C, 50 mM HEPES, pH 8.0, 0.3 M NaCl, 0.1 mM PLP, and 0.03 M imidazole). The histidine-

tagged proteins were eluted with elution buffer (Buffer-D, 50 mM HEPES, pH 8.0, 0.3 M NaCl, 0.3 M imidazole, and 0.1 mM PLP; 5 mL min⁻¹ flow rate) and the fractions containing the desired protein were collected. For desalting, size exclusion chromatography employing a 5 mL Sephadex desalting column (GE Healthcare) using Buffer-A was performed. SDS-PAGE was performed to check the quality of the purified enzymes (Fig. S2/Tab. S1 with protein yields). The purified enzyme solutions were stored at 4°C and the protein concentrations were determined by the Nano Quant method.

2.4 Point Mutations

Variants of TA-8H were constructed using the following methodology. Primers were designed with the desired mismatches to introduce the mutations. The reaction was set-up in a total reaction volume of 50 µL and contained 10 µL of 5X Q5 reaction buffer, 1 µL of 10 mM dNTPs, 2.5 µL of 10 µM forward primer, 2.5 µL of 10 µM reverse primer, 1 µL of template DNA (100 ng), 0.5 µL of Q5 Polymerase and 32.5 µL of molecular biology grade H₂O. PCR program: 1) initial denaturation at 98 °C, 1 min; 2) 30 cycles of a) denaturation at 98 °C, 10 s; b) annealing at 62–65 °C depending on the melting temperature of the primers, 30 s; c) polymerization at 72 °C, 45 s; 3) final extension at 72 °C, 5 min; 4) cooling/storage at 4 °C. A clean-up of the PCR products using the Macherey-Nagel PCR purification kit yielded the products in water. The ligation reaction was assembled in 10 µL total volume and contained 6 µL of purified PCR product, 1 µL of 10X T4 DNA ligase buffer, 1 µL each of T4 DNA ligase, T4 polynucleotide kinase (PNK), and *DpnI*. The mixture was incubated for 15 min at 25 °C and 45 min at 37 °C. Then, 5 µL of the ligation reaction was used to transform 50 µL of chemically competent *E. coli* Top10. The following primers were used. Variant Gly60Trp: FW_1 CGCGGTGCTGTGGGTTACCCAAG and RV_2 CACTTATGGACCTGCGCTC. Variant Gly60Tyr: FW_3 CGCGGTGCTGTATGTTACCCAAGTG and RV_4 ACTTATGGACCTGCGCTC. Variant Ala233Ser: FW_5 GATCATGGGTAGCGTTGGTGCGGC and RV_6 GGTTACCAATCATCGCC. The bold marked nucleotides code for the mutations.

2.5 Acetophenone Assay

The activity of the ATAs was investigated by the acetophenone assay and the alanine dehydrogenase assay (See Figs. S3 and S4 in the SI for a reaction scheme). The standard substrate **1** was combined with one of the α-keto acids **8–10** in the acetophenone assay [23], which was carried out in 96-well plates (Greiner bio-one UV-STAR microplates, F-Bottom). The reaction solution contained 2.5 mM amino donor (**S-1**), 1 mM of a α-keto acid, 50 mM HEPES pH 8,

and the kinetic measurement was started by adding 20 µL of the purified enzyme (0.5–1 mg mL⁻¹). The total volume was 200 µL per well and the reaction mixtures were shaken for 5 s (300 rpm, 2 mm) in the Tecan plate reader (Infinite 200). The formation of acetophenone was measured at 245 nm for 5 min at 30 °C. One unit (U) of ATA activity was defined as the formation of 1 µmol of acetophenone per minute. Three control reactions were included where enzyme, amine donor, or amino acceptor were replaced by buffer. Protein concentration was measured by the Roti nanoquant kit. The specific activity was calculated according to Eq. (1): the slope $\Delta Abs/\Delta t$ expressed in min⁻¹ was obtained from the kinetic absorbance measurement of the enzymatic reaction and was corrected by subtracting the slope of the control without amine substrate $\Delta Abs_c/\Delta t$; the constant 5.23 µmol⁻¹ represents the product of the absorption coefficient and pathlength and was obtained from a calibration experiment with varying acetophenone concentrations. c_p is the protein concentration in the reaction solution [mg mL⁻¹] and V the volume of the reaction solution, which was 200 µL. All the measurements were performed in triplicates.

$$A = \left(\frac{\Delta Abs}{\Delta t} - \frac{\Delta Abs_c}{\Delta t} \right) \frac{[\mu \text{ mol}]}{5.23 c_p V} \quad (1)$$

2.6 Alanine Dehydrogenase Assay

For amine substrates (**1–7**) the alanine dehydrogenase assay was performed as previously described [18] with the following reaction conditions: 2.5 mM amino donor (**1–7**), 1 mM pyruvate **8**, 2.5 vol% DMSO, 0.3 mg mL⁻¹ of recombinant AlaDH from *Thermus thermophilus* [18, 24], 5 µM methoxy-PMS, 1 mM NAD⁺, 0.3 mM XTT (2,3-Bis(2-methoxy-4-nitro-5-sulfophenyl)-2H-tetrazolium-5-carboxanilide) salt. The reaction was started by adding 40 µL of the purified transaminase solution at 30 °C. Two control reactions were included where AlaDH or the putative transaminase protein was substituted with buffer. Measurements were done in triplicate. The slope from the assay reaction was corrected by subtracting the absorbance of the control without transaminase. Corrected slopes were then normalized to that of the enzyme TA-3HMU to calculate relative activities.

2.7 Temperature Stability Assays

The thermal stability also referred to as ‘resting stability’ of the enzymes was investigated by incubating the purified enzyme (1 mg mL⁻¹) at temperatures between 30–70 °C for a defined time in a buffer containing 0.1 mM PLP and 50 mM HEPES at pH 8. Samples of the enzyme solution were collected at respective time intervals, cooled down to 30 °C and the activity was measured by the acetophenone assay as described above at 30 °C.

For assaying the stability under operating conditions similar to a biocatalysis reaction, also referred to as operating stability, the proteins were incubated with 100 mM L-alanine as an amino donor in 50 mM HEPES buffer, pH 8.0, containing 0.1 mM PLP and at temperatures between 30–60 °C. Before incubation at elevated temperatures, the initial activities at t_0 were determined at 30 °C. While incubating the enzyme at different temperatures, samples were collected at different time intervals and enzyme activities were measured by acetophenone assay as described above. The ratios of volumetric activities after and before heat treatment were then calculated to obtain the relative activities of the enzymes under operating conditions.

3 Results

3.1 Selection of Sequences and Substrates

The transaminases TA-3FCR and TA-3HMU are examples of enzymes with a very low versus a relatively high activity and are representatives of two subfamilies, which were dis-

covered in a previous bioinformatics analysis using the 3DM database software. In-line with the argumentation used for functional annotation, one would expect that sequences belonging to the TA-3HMU subfamily encode enzymes that are more likely to be highly active and vice versa. Members of the TA-3HMU subfamily typically have hydrophobic amino acids near to the catalytic lysine as shown in Fig. 1b. To investigate whether the hydrophobicity of active site residues rather than a global sequence similarity is a predictor of the activity level, non-typical sequences of each subfamily were chosen that have opposite properties of these active site region residues. Tab. 1 shows the most important active site residues of the selected sequences.

TA-1F to TA-8H are sequences that bear mainly hydrophilic side chains near to the active site similar to TA-3FCR, but five of the eight sequences show a slightly higher sequence similarity to TA-3HMU. For readability, the suffix “H” or “F” in the enzyme numbering denotes the subfamily, 3HMU versus 3FCR, to which the sequence belongs. On the contrary, six out of seven ATAs, whose active sites are composed of mainly hydrophobic residues, are more similar to TA-3FCR. See Figs. S5 and S6 in the SI for an amino acid

Table 1. Investigated proteins from TA-3FCR and TA-3HMU subfamilies with varying active-site residues.

Protein ^{a)}	Source organism	Uniprot-ID ^{b)}	Sequence identity ^{c)} [%] to		Sequence motif: amino acid at positions ^{d)}						
			TA-3FCR	TA-3HMU	25	63	91	92	233	417	419
TA-3FCR	<i>Silicibacter</i> sp. TM1040	Q1GD43	100	40	S	Y	Y	V	T	R	M
TA-1F	<i>Pseudomonas</i> sp.	J2S6R5	66	44	S	Y	Y	M	T	R	M
TA-2H	<i>Streptomyces sviveus</i>	B51413	35	40	M	Y	W	T	R	R	Y
TA-3H	<i>Streptomyces himastatinicus</i>	D9WWM5	34	40	M	Y	W	S	R	R	Y
TA-4H	<i>Rhizobium</i> sp.	J6DQH8	37	42	Q	W	Y	Y	S	R	S
TA-5H	<i>Pseudomonas aeruginosa</i>	A3L376	38	43	Y	W	F	C	S	R	I
TA-6H	<i>Rhodococcus vannielii</i>	E3I209	40	44	Q	W	Y	A	G	R	V
TA-7H	<i>Burkholderia cenocepacia</i>	A0KE01	44	44	V	W	Y	F	S	R	F
TA-8H	<i>Streptomyces clavuligerus</i>	E2PWJ4	39	41	T	G	W	G	A	R	N
TA-3HMU	<i>Silicibacter pomeroyi</i>	Q5LMU1	40	100	F	W	F	F	A	R	V
TA-9H	<i>Pseudomonas putida</i>	Q88KV9	43	64	F	W	F	F	A	R	V
TA-10F	<i>Burkholderia</i> sp.	I2IXB4	53	45	F	W	F	A	A	R	L
TA-11F	<i>Thalassospira xiamenensis</i>	K2L0P6	48	40	A	W	F	A	A	R	L
TA-12F	<i>Teredinibacter turnerae</i>	C5BK98	49	42	A	W	F	S	A	R	L
TA-13F	<i>Sphaerobacter thermophilus</i>	D1C7Z3	48	43	V	W	F	F	A	R	N
TA-14F	<i>Sphingomonas</i> sp.	F3WZ58	54	40	F	W	F	N	A	R	L
TA-15F	<i>Azospirillum amazonense</i>	G1XWD6	51	41	F	W	F	N	A	R	L

a) The letter F or H at the end indicates whether the protein belongs to the subfamily of TA-3FCR or TA-3HMU according to the phylogenetic tree in Fig. S5; b) Entrance code to retrieve the amino acid sequence online at <https://www.uniprot.org/>; c) Calculated from the multiple alignment given in Fig. S6. d) The position correspondent to TA-3FCR as assigned in the pdb file from the crystal structure. Corresponding positions for TA-3HMU are: 19, 59, 87, 88, 231, 420, 422. Letters are the standard one-letter abbreviations of amino acids. Amino acids with polar side chains are given in bold italic font.

alignment and a phylogenetic tree of the chosen sequences, and Tab. S2 for the amino acid identities of all proteins used in this study.

A range of amines **1–7** and α -keto acids **8–10** (Fig. 2) were chosen for the activity measurements. As few ATAs do not accept the standard substrate (*S*)-1-phenylethylamine **1** very well, the substrate selection included the aliphatic amines **4** and **6**. Furthermore, one additional interesting question is whether an ATA, where the active site is composed of preferentially hydrophilic amino acid side chains, would prefer a more hydrophilic substrate. This was investigated with substrates in which one methyl group is replaced by a hydroxyl group such as **3**, **5**, and **7** and α -keto acid **10**.

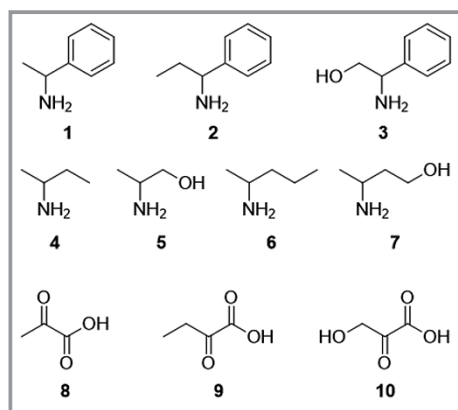


Figure 2. Structures of the investigated substrates.

3.2 Activities with α -Keto Acids

The gene expression of the 15 selected sequences yielded sufficient amounts of soluble protein for characterization and all enzymes were purified to near homogeneity (Fig. S2/Tab. S1). The first step in the characterization was the activity measurement with the acetophenone assay using the standard substrates **1** and **8**. For comparison, TA-3HMU was used as a reference point and its specific activity of 1.9 U mg^{-1} was set to 100% (Tab. 2). The distribution of the determined activities were in accordance with the proposed hydrophobicity hypothesis: Proteins having mainly hydrophilic active site residues showed one to two orders of magnitude lower activities than TA-3HMU (Tab. 2, Entries 1–8). On the contrary, the presence of hydrophobic active site residues is associated with high activities of 60–200%. Besides pyruvate, α -keto acids **9** and **10**, which bear an additional methyl or hydroxyl group, were assayed. A general trend can be seen with the sufficiently active enzymes TA-8H to TA-15F that the order of acceptance is $8 > 9 > 10$. The lower activities with the larger α -keto acids were observed in previous studies and this is in accordance with the binding pocket model of ATAs [25, 26] (Fig. 1a). Since an acidic substrate places its carboxyl group in the large binding pocket to form a salt bridge with the guanidino group of the ‘flipping’ arginine [21], the small

binding pocket has to accommodate the substrate’s side chain. With increasing size, steric hindrances will prohibit the substrate to approach the active site in a favorable position for reacting with the cofactor. This, however, does not explain why the hydroxy-substituted analogue **10** showed 3–10 times lower activities compared to the similarly sized **9**. One reason might be that substrates bearing a polar or charged functional group are usually solvated and need a fine-tuned network of interactions with the active site rather than merely the presence of hydrophilic side chains, which seems not sufficient for their efficient conversion.

Table 2. Relative enzyme activities towards α -keto acids.

Entry	Enzyme TA-	Relative specific activity ^{a)} [%] with substrate 1 and		
		8	9	10
1	3FCR	5	n.d.	n.d.
2	1F	2±0.2	3±0.1	1.4±0.1
3	2H	n.d.	5.5±0.4	2.5±0.1
4	3H	8.2±0.5	3.5±0.2	1.2±0.1
5	4H	n.d.	1±0.1	0.7±0.1
6	5H	2.3±0.2	2.8±0.2	1.2±0.1
7	6H	2.4±0.3	3.2±0.2	1.9±0.1
8	7H	8.3±0.9	2.3±0.3	1.4±0.07
9	8H	65±1.5	16±1.2	5.5±0.4
10	3HMU	100±5	20.5±1.5	7.3±0.5
11	9H	200±3.5	30±2	11±0.6
12	10F	59±3	15±0.9	5.7±0.3
13	11F	199±5	43±3	13±0.8
14	12F	61±2	10±0.6	6.2±0.4
15	13F	61±1.5	18±1.5	10.7±0.6
16	14F	59±2.5	15±1	7.1±0.4
17	15F	102±6	24±1.8	11±0.6

a) Specific activities were determined using the acetophenone assay. The specific activity of TA-3HMU with the substrate pair **1** and **8** was 1.9 U mg^{-1} and set as 100%. All other values were normalized to this activity. The errors represent the single standard deviation from a technical triplicate. n.d. – not detectable.

3.3 Activities with Amines

The alanine dehydrogenase assay allowed the characterization of the ATAs using different amines, resulting in a clear trend (Tab. 3). First, most ATAs showed the highest activity with the model substrate **1**. Only TA-10F showed a clear preference for the aliphatic amines **4** and **6** over **1**. Similar to the series of α -keto acids, amine **2** has a larger

Table 3. Relative enzyme activities towards amine substrates.

Protein TA-	Relative specific activity ^{a)} [%] with substrates 8 and						
	1	2	3	4	5	6	7
3FCR	6±0.7	3.2±0.2	1.9±0.1	16±1	6.3±0.7	9.4±1	6.3±0.4
1F	2.2±0.5	n.d.	n.d.	6.3±0.6	n.d.	n.d.	n.d.
2H	n.d.	n.d.	n.d.	4.4±0.6	3.2±0.4	6.3±0.3	2.5±0.1
3H	10±0.9	1.3±0.1	n.d.	1.9±0.2	0.7±0.1	n.d.	n.d.
4H	n.d.	n.d.	n.d.	3.8±0.5	0.9±0.1	8.8±0.4	n.d.
5H	6.6±1.5	1.6±0.1	n.d.	1.3±0.1	0.9±0.1	11±0.6	n.d.
6H	5.3±0.1	4.1±0.3	n.d.	3.2±0.3	1.5±0.1	5±0.3	n.d.
7H	9.3±0.7	4.5±0.3	n.d.	2.5±0.2	n.d.	2.5±0.2	n.d.
8H	46±2	13±0.8	8.8±0.6	25±1.6	9.4±0.8	7.5±0.6	3±0.2
3HMU	100±5	19±1.2	7.5±0.5	75±6	34±2.2	47±5	40±3
9H	198±8	9.4±0.6	6.3±0.3	63±4	16±1.1	56±3	25±1
10F	45±1	28±2.3	12±0.7	106±8	47±5	56±6	38±2.5
11F	214±27	22±1.5	16±1.2	32±2.5	13±0.9	19±1	16±1.4
12F	73±3	12.5±0.7	3.2±0.2	19±1	9.4±0.6	16±1	11.5±1
13F	68±4	37.5±1.5	22±2.3	19±2.7	13±1	28±2	25±1.6
14F	50±4	20±2.2	6.3±1	47±3	22±1.5	22±1.5	18±2
15F	90±3	16±1	8.2±0.5	66±5	38±3	38±3	35±4

a) Specific activities were determined using the alanine dehydrogenase assay. The specific activity of TA-3HMU with the substrate pair **1** and **8** was set as 100 %. All other values were normalized to this activity. The errors represent the single standard deviation from a technical triplicate. n.d. – not detectable.

substituent than a methyl group, which has to be accommodated in the small binding pocket. Compared to **1**, this resulted in a 2–20-fold drop in activity. The presence of the hydroxyl substituent in **3** also lead to 1.4–4-fold reduced activities compared to **2**, but it has a less dramatic influence as compared with the α -keto acid pair **9** and **10**. Substrate pairs **4** & **5** and **6** & **7** are analogues that differ only by an exchange of a methyl group, which is replaced by a hydroxyl group. On the contrary to substrates **10** and **3**, the hydroxyl group of **5** and **7** will be placed in the large binding pocket, as these substrates also bear a methyl group, which is very likely accommodated in the small pocket. Despite the alternative orientation and the differing distances of the hydroxyl groups to the amino group, an activity drop of up to 3-fold was observed with the hydroxyl-analogues. The activity decrease was less pronounced if the hydroxyl group is placed one carbon more distant to the amino group (**7** versus **5**). Contrary to expectations, the ATAs with more hydrophilic active site residues did not show a better acceptance for these substrates either.

3.4 Influence of Sequence Variations

For most of the studied sequences the hydrophobicity hypothesis was in accordance with the overall activity level. One protein, TA-8H, was predicted to have a lower activity because it does not show the typical hydrophobic active site pattern, but with 60 % relative activity to TA-3HMU it had an unexpected high activity. In its sequence, it has a glycine at position 63, which is very unusual because a conserved tryptophan or tyrosine is found in the TA-3HMU/3FCR subfamily. In addition, at position 233 there is an alanine instead of a hydrophilic residue expected in the less-active ATAs. To decipher which position might be responsible for the intermediate activity, variants were constructed via site-directed mutagenesis (Tab. 4) and activities measured with the standard substrate pair **1** and **8**. Since the Gly63Trp variant has approximately the same activity, it can be ruled out that the high activity of TA-8H is due to the presence of Gly63. In comparison, a tyrosine at position 63 reduced the activity to 60 %. A larger effect is seen at position 233: the introduced serine diminished the activity to 15 %. In the double variant, these effects are additive for the Tyr63-Ser233 double variant, but the Trp63-Ser233 variant still

showed 75 % activity of the wild type. This demonstrates that especially the combination of the hydrophilic residues Tyr63 and Ser233 leads to low activities, and the presence of Ala233 was most responsible for the relatively high activity of TA-8H.

Table 4. Enzyme activities of the variants of TA-8H.

Enzyme variant	Relative activity [%] ^{a)}
TA-8H	65±5
TA-8H-Gly60Tyr	36±3
TA-8H-Gly60Trp	69±7
TA-8H-Ala233Ser	10±6
TA-8H-Gly60Tyr/ Ala233Ser	5±0.6
TA-8H-Gly60Try/ Ala233Ser	47±4

a) Specific activities were measured by the acetophenone assay with the substrate pair **1** and **8** and normalized to the activity of TA-3HMU (100 % corresponds to a specific activity of 1.9 U mg⁻¹).

3.5 Stabilities

Finally, the stability under storage conditions or in the presence of the amino donor substrate alanine was investigated. Four out of the 15 proteins showed an acceptable stability, as they had 75–80 % residual activity after storage at 40 °C for 24 h, which is a typical reaction time in biocatalysis (Tab. 5; see also Tabs. S3 and S4 for the full data set of stability characterizations at different temperatures). In the presence of the amino donor alanine, which mimics operating conditions, their activities were reduced to 60 % at 40 °C and 30–50 % at 60 °C, representing a relatively good performance. Most wild-type proteins, such as the well-character-

Table 5. Stabilities of selected transaminases.

Protein TA-	Residual relative activity ^{a)} [%] after 24 h incubation under			
	Resting conditions ^{b)}		Operating conditions ^{c)}	
	40 °C	60 °C	40 °C	60 °C
7H	75	66	60	48
10F	80	40	60	50
13F	80	45	60	40
15F	77	60	57	31

a) Activities were measured by the acetophenone assay with the substrate pair **1** and **8**. Relative residual activity is the ratio of activity before and after the incubation step; b) The purified enzyme solution was incubated at the respective temperatures in the storage buffer (50 mM HEPES pH 8, 0.1 mM PLP) to determine the resting stability; c) To measure the operating stability, the storage buffer is supplemented with 100 mM alanine as amino donor.

ized ATA from *Vibrio fluvialis*, show a significant or complete activity loss after a few hours of catalysis [27]. This was explained by the following mechanism: due to the large excess of amino donor usually applied for equilibrium shift, the PMP state of the cofactor dominates and has an increased probability to dissociate from the protein compared to PLP. This accelerates inactivation by aggregation of the ATA [28]. The operating stabilities of the ATAs from this study are better than those reported for *Vibrio fluvialis* ATA [27]. They are in the same range as a wild-type ATA identified in a metagenomic library probably originating from *Pseudomonas* sp., TA-5LH9 [28], but less stable than the engineered variant of TA-5LH9 created by Börner and co-workers [29] and our recently characterized wild-type ATA from *Bulkholderia multivorans* [30], which both are tetrameric enzymes.

4 Conclusion

From previous work it was known that amine transaminases can be identified by protein sequence searches in databases. In this study we have shown that the relative activity can be predicted and it correlates with the hydrophobicity of selected residues surrounding the active site with hydrophobic residues being beneficial for activity towards amines. Hydrophilic substrate analogues bearing a hydroxyl group resulted in a reduced acceptance by all enzymes, which means that hydrophilic side chains present in half of the ATAs investigated did not promote acceptance of the more polar amino alcohol substrates. Future studies might elucidate structural features beneficial for converting amino alcohols or other types of hydrophilic substrates.

Supporting Information

Supporting Information for this article can be found under DOI: 10.1002/cite.202200062. This section includes additional references to primary literature relevant for this research [31,32].

We thank the German Research Foundation (DFG) for financial support (U.T.B., grant number: Bo1862/16-1; M.H., grant number: Ho4754/4-1). Open access funding enabled and organized by Projekt DEAL.

Symbols used

<i>A</i>	[U mg ⁻¹]	specific activity
<i>Abs</i>	[-]	absorbance
<i>c_p</i>	[mg mL ⁻¹]	protein concentration in the assay solution
<i>g</i>	[m s ⁻²]	gravity constant

OD_{600}	[-]	optical density at 600 nm
T_a	[°C]	annealing temperature
T_m	[°C]	melting temperature
U	[$\mu\text{mol min}^{-1}$]	unit of enzyme activity
V	[mL]	volume

Greek letters

λ_{max}	[nm]	absorbance maximum
------------------------	------	--------------------

Abbreviations

A	alanine
Ala	alanine
ATA	amine transaminase
ADH	alanine dehydrogenase
Arg	arginine
C	centigrade
C	cysteine
DNA	desoxyribonucleic acid
dNTPs	deoxynucleoside triphosphates (dATP , dCTP , dGTP , and dTTP)
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
dTTP	deoxythymidine triphosphate
DpnI	restriction enzyme from <i>Diplococcus pneumoniae</i> G41
F	phenylalanine
FWp	forward primer
G	glycine
Gly	glycine
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
I	isoleucine
IPTG	isopropyl- β -D-1-thiogalactopyranoside
L	leucine
Lys	Lysine
M	methionine
mU	milliunits
N	asparagine
Ni-NTA	nickel nitrilotriacetic acid
PCR	polymerase chain reaction
Phe	phenylalanine
PLP	pyridoxal-5'-phosphate
PMP	pyridoxal-5'-aminophosphate
Q	glutamine
R	arginine
rpm	revolutions per minute
RVp	reverse primer
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
S	serine
Ser	serine

T	threonine
Trp	tryptophan
Tyr	tyrosine
UV	ultraviolet
V	valine
Val	valine
W	tryptophan
XTT	2,3-Bis(2-methoxy-4-nitro-5-sulfophenyl)-2H-tetrazolium-5-carboxanilide
Y	tyrosine

References

- [1] R. N. Patel, *Coord. Chem. Rev.* **2008**, *252*, 659–701.
- [2] P. N. Devine, R. M. Howard, R. Kumar, M. P. Thompson, M. D. Truppo, N. J. Turner, *Nat Rev Chem* **2018**, *2*, 409–421.
- [3] D. Yi, T. Bayer, C. P. S. Badenhorst, S. Wu, M. Doerr, M. Höhne, U. T. Bornscheuer, *Chem. Soc. Rev.* **2021**, *50*, 8003–8049.
- [4] M. Höhne, P. Sondermann, N. J. Turner, E. M. Carreira, *Angew. Chem. Int. Ed.* **2017**, *56*, 8942–8973.
- [5] M. D. Patil, G. Grogan, A. Bommarius, H. Yun, *Catalysts* **2018**, *8*, 254.
- [6] E. E. Ferrandi, D. Monti, *World J. Microbiol. Biotechnol.* **2018**, *34*, 13. DOI: <https://doi.org/10.1007/s11274-017-2395-2>
- [7] I. V. Pavlidis, M. S. Weiss, M. Genz, P. Spurr, S. P. Hanlon, B. Wirz, H. Iding, U. T. Bornscheuer, *Nat. Chem.* **2016**, *8*, 1076–1082.
- [8] C. K. Savile, J. M. Janey, E. C. Mundorff, J. C. Moore, S. Tam, W. R. Jarvis, J. C. Colbeck, A. Krebber, F. J. Fleitz, J. Brands, P. N. Devine, G. W. Huisman, G. J. Hughes, *Science* **2010**, *329*, 305–309.
- [9] A. Gomm, E. O'Reilly, *Curr. Opin. Chem. Biol.* **2018**, *43*, 106–112.
- [10] M. Fuchs, J. E. Farnberger, W. Kroutil, *Eur. J. Org. Chem.* **2015**, *2015*, 6965–6982.
- [11] S. A. Kelly, S. Pohle, S. Wharry, S. Mix, C. C. R. Allen, T. S. Moody, B. F. Gilmore, *Chem. Rev.* **2018**, *118*, 349–367.
- [12] K. E. Cassimjee, B. Manta, F. Himo, *Org. Biomol. Chem.* **2015**, *13*, 8453–8464.
- [13] B. Manta, K. E. Cassimjee, F. Himo, *ACS Omega* **2017**, *2*, 890–898.
- [14] M. Höhne, S. Schätzle, H. Jochens, K. Robins, U. T. Bornscheuer, *Nat. Chem. Biol.* **2010**, *6*, 807–813.
- [15] F. Steffen-Munsberg, C. Vickers, H. Kohls, H. Land, H. Mallin, A. Nobili, L. Skalden, T. van den Bergh, H.-J. Joosten, P. Berglund, M. Höhne, U. T. Bornscheuer, *Biotechnol. Adv.* **2015**, *33*, 566–604.
- [16] A. Telzerow, J. Paris, M. Håkansson, J. González-Sabín, N. Ríos-Lombardía, H. Gröger, F. Moris, M. Schürmann, H. Schwab, K. Steiner, *ChemBioChem* **2021**, *22*, 1232–1242.
- [17] S. A. Kelly, S. Mix, T. S. Moody, B. F. Gilmore, *Appl. Microbiol. Biotechnol.* **2020**, *104*, 4781–4794.
- [18] F. Steffen-Munsberg, P. Matzel, M. A. Sowa, P. Berglund, U. T. Bornscheuer, M. Höhne, *Appl. Microbiol. Biotechnol.* **2016**, *100*, 4511–4521.
- [19] J. Jiang, X. Chen, D. Zhang, Q. Wu, D. Zhu, *Appl. Microbiol. Biotechnol.* **2015**, *99*, 2613–2621.
- [20] F. Steffen-Munsberg, C. Vickers, A. Thontowi, S. Schätzle, T. Tumlrirsch, M. Svedendahl Humble, H. Land, P. Berglund, U. T. Bornscheuer, M. Höhne, *ChemCatChem* **2013**, *5*, 150–153.

- [21] F. Steffen-Munsberg, C. Vickers, A. Thontowi, S. Schätzle, T. Meinhardt, M. Svedendahl Humble, H. Land, P. Berglund, U. T. Bornscheuer, M. Höhne, *ChemCatChem* **2013**, *5*, 154–157.
- [22] C. Ramírez-Palacios, H. J. Wijma, S. Thallmair, S. J. Marrink, D. B. Janssen, *J. Chem. Inf. Model.* **2021**, *61*, 5569–5580.
- [23] S. Schätzle, M. Höhne, E. Redestad, K. Robins, U. T. Bornscheuer, *Anal. Chem.* **2009**, *81*, 8244–8248.
- [24] Z. Váli, F. Kilar, S. Lakatos, S. A. Venyaminov, P. Zavadzsky, *Biochim. Biophys. Acta BBA – Enzymol.* **1980**, *615*, 34–47.
- [25] J.-S. Shin, B.-G. Kim, *J. Org. Chem.* **2002**, *67*, 2848–2853.
- [26] K. S. Midelfort, R. Kumar, S. Han, M. J. Karmilowicz, K. McConnell, D. K. Gehlhaar, A. Mistry, J. S. Chang, M. Anderson, A. Villalobos, J. Minshull, S. Govindarajan, J. W. Wong, *Protein Eng. Des. Sel.* **2013**, *26*, 25–33.
- [27] S. Chen, J. C. Campillo-Brocal, P. Berglund, M. S. Humble, *J. Biotechnol.* **2018**, *282*, 10–17.
- [28] T. Börner, S. Rämisch, E. R. Reddem, S. Bartsch, A. Vogel, A.-M. W. H. Thunnissen, P. Adlercreutz, C. Grey, *ACS Catal.* **2017**, *7*, 1259–1269.
- [29] T. Börner, S. Rämisch, S. Bartsch, A. Vogel, P. Adlercreutz, C. Grey, *ChemBioChem* **2017**, *18*, 1482–1486.
- [30] M. Kollipara, P. Matzel, M. Sowa, S. Brott, U. Bornscheuer, M. Höhne, *Appl. Microbiol. Biotechnol.* **2022**, *106*, 5563–5574.
- [31] M. V. Han, C. M. Zmasek, *BMC Bioinf.* **2009**, *10*, 356.
- [32] C. Gille, M. Fähling, B. Weyand, T. Wieland, A. Gille, *Nucl. Acids Res.* **2014**, *42*, 3–6.