

# Exploring the Stability and Substrate Profile of Transaminase from *Silicibacter pomeroyi* with Ancestral Sequence Reconstruction

Luyao Zhao, Bhu-Bhud Thongrakon, Trishnamoni Gautom, Viktor Sahlberg, and Per Berglund\*

Dedicated to the memory of Prof Iván Lavandera, University of Oviedo, Spain.

Amine transaminases (ATAs), belonging to the class III transaminases within the superfamily of pyridoxal-5'-phosphate-dependent enzymes, catalyze transamination reactions between amino donors and amino acceptors. These enzymes are particularly appealing for their role in stereospecific synthesis of chiral amines. However, the stability of most ATAs is not satisfying, limiting their suitability for industrial applications. Among them, the amine transaminase from *Silicibacter pomeroyi* (*Sp*-ATA) has drawn attention due to its high activity and broad substrate scope under mild conditions and high pH. Nevertheless, maintaining the activity at higher temperatures is a challenge. Previous studies to enhance

enzyme function through directed evolution have shown promising results, yet predicting the cooperative effects of individual stabilizing mutations remains challenging. An alternative strategy is ancestral sequence reconstruction (ASR), which is based on gene sequences to create a more or less artificial phylogenetic tree. This study aims to leverage ASR techniques to explore the thermostability, solvent tolerance, and substrate profile of *Sp*-ATA, to find more stable transaminases. By using *Sp*-ATA as a template and incorporating insights from ancestral sequences, this strategy offers a promising approach for developing robust biocatalysts suitable for industrial applications.

## 1. Introduction

Biocatalysis has emerged as an important technology in recent years, especially in asymmetric synthesis, due to the stereoselectivity, regioselectivity, and chemoselectivity of the biocatalysts, which make the process more environmentally friendly and simplified.<sup>[1–6]</sup> Chiral amines and their derivatives are widely found in many active pharmaceutical ingredients (API) and fine chemicals.<sup>[7–9]</sup> They are high-value synthetic targets, as these compounds often contain one or more stereocenters, and controlling the overall stereochemistry is quite challenging. Biocatalysis provides a more sustainable route for the synthesis of high-purity chiral amines and their derivatives.<sup>[3,10,11]</sup> Currently, biocatalysts used for the production of chiral amines mainly include hydrolases,<sup>[12,13]</sup> oxidoreductases,<sup>[14,15]</sup> and transferases.<sup>[16]</sup> Amine dehydrogenases (AmDHs) and imine reductases (IREDs) have emerged with significant research focus over the

past decade due to their catalytic role in the amination of ketones.<sup>[17–20]</sup> Despite their potential as alternatives to traditional methods for chiral amine synthesis, transaminases still hold significant potential for further investigation.<sup>[21–23]</sup> Transaminases (TAs, EC 2.6.1.X) are enzymes that use the coenzyme pyridoxal 5'-phosphate (PLP, a derivative of vitamin B6) to catalyze the transfer of amino groups to carbonyl groups, playing a crucial role in the synthesis of chiral amines and amino acids.<sup>[17,24]</sup> PLP forms a Schiff base with a lysine residue in the active site in the first half-cycle of the reaction as part of a “ping-pong bi-bi” mechanism. The amino donor undergoes deamination to produce the donor product and pyridoxamine-5'-phosphate (PMP). As the amino acceptor is aminated, PMP releases the amino group, forming PLP and generating a new amine or amino acid.<sup>[25,26]</sup> TAs exhibit a broad substrate range, efficient cofactor regeneration, and excellent stereoselectivity, giving these enzymes advantages not seen in other enzyme classes.<sup>[10,20,23]</sup> TAs play important roles in industry and can reduce the need for expensive homogeneous catalysts while lowering the high energy consumption associated with downstream purification processes.<sup>[22]</sup>

Amine transaminases (ATAs) are distinguished by their ability to aminate a wide range of ketones, aldehydes, and other carbonyl compounds, enabling the asymmetric synthesis of chiral amines from prochiral ketones without requiring a carboxyl group in the substrate. ATAs are one of the most versatile options for chiral amine synthesis.<sup>[27]</sup> Among all ATAs, one particularly interesting enzyme is a transaminase from *Silicibacter pomeroyi* (*Sp*-ATA, 3HMU) where a highly conserved arginine residue is

L. Zhao, B.-B. Thongrakon, T. Gautom, V. Sahlberg, P. Berglund  
Department of Industrial Biotechnology  
School of Engineering Sciences in Chemistry, Biotechnology and Health  
KTH Royal Institute of Technology  
SE-10691 Stockholm, Sweden  
E-mail: perbe@kth.se

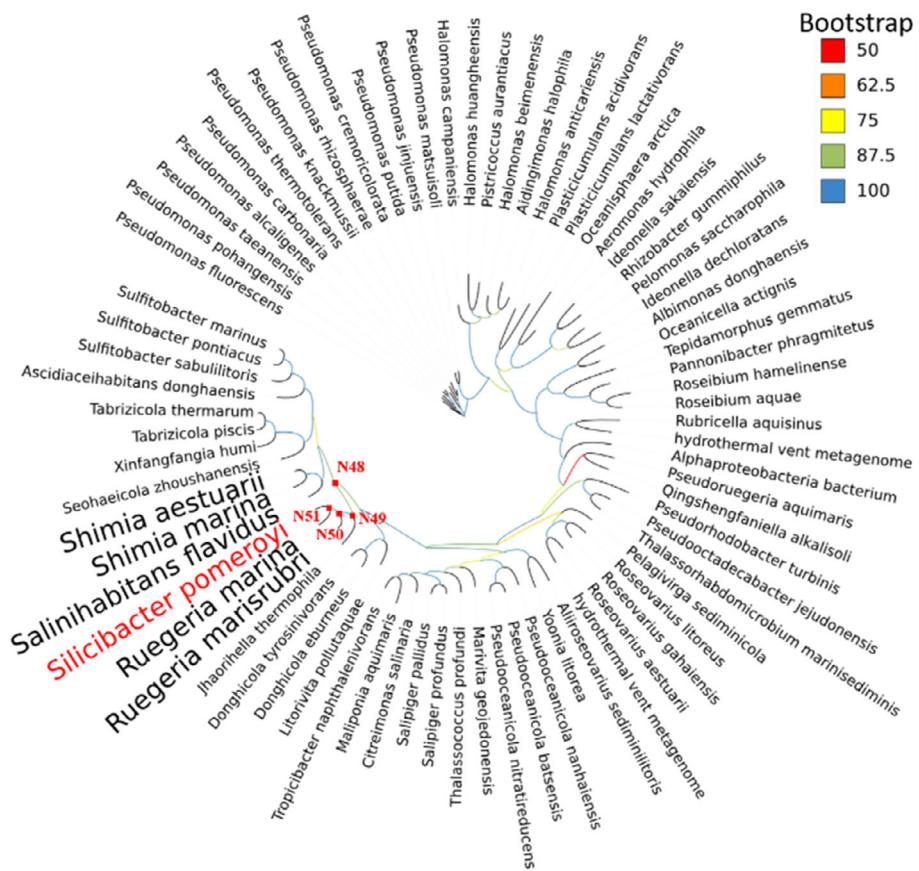
 Supporting information for this article is available on the WWW under <https://doi.org/10.1002/cbic.202500155>

 © 2025 The Author(s). ChemBioChem published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

typically responsible for binding the carboxyl group at the active site of transaminases. This flexible “flipping” Arg417 creates a cavity in the active site that can accommodate the amino/ketone group.<sup>[27-30]</sup> This flexibility gives *Sp*-ATA a broad substrate scope. However, despite its excellent catalytic properties, *Sp*-ATA has limitations in stability.<sup>[31]</sup> Studies have shown that *Sp*-ATA exhibits poor stability under extreme conditions such as high temperatures and organic solvents, and instability of the cofactor PLP.<sup>[23,32]</sup> Industrial processes demand enzymes that maintain stability and catalytic efficiency under nonphysiological conditions over time. Enzymes with poor stability tend to deactivate, requiring frequent replenishment and complex enzyme immobilization steps, which increase production costs.<sup>[33,34]</sup> Additionally, industrial biocatalysis frequently involves harsh conditions, such as high substrate concentrations, non-aqueous systems, and extreme pH values,<sup>[35,36]</sup> making enzyme stability crucial for maintaining activity and preventing downtime.<sup>[37]</sup> Thermostability is widely regarded as a key factor in enhancing overall enzyme stability, as it strengthens protein structure and improves resistance to extreme environments.<sup>[38]</sup> Studies show that thermostable enzymes possess enhanced hydrogen bonding networks, hydrophobic cores, and disulfide bonds, making them more resilient under industrial conditions.<sup>[39]</sup> Enhancing enzyme stability ensures that it maintains catalytic activity under these conditions,

preventing downtime caused by enzyme deactivation, thus improving production efficiency. However, *Sp*-ATA's catalytic activity diminishes significantly in these environments, restricting its industrial applicability.<sup>[40]</sup>

To overcome these limitations, we aim to improve the stability of *Sp-ATA*. A primary objective in engineering transaminases is improving thermostability through various strategies, including charge-charge interactions, proline substitutions, and the stabilization of protein quaternary structures by preventing dimer dissociation.<sup>[41–45]</sup> For instance, using B-factor analysis in combination with the proline rule, Land et al. improved the thermostability of the amine transaminase from *Chromobacterium violaceum*.<sup>[41]</sup> Cao et al. increased the thermostability and activity of a transaminase from *Aspergillus terreus* by identifying four surface-exposed residues hypothesized to enhance thermal stability.<sup>[44]</sup> Enzyme functions have been enhanced by directed evolution in the past, but it is challenging to predict the effects of combining individual stabilizing mutations.<sup>[46–49]</sup> Computational methods have also been applied, but these are often time-consuming and require a detailed structure to generate training data sets.<sup>[25,50]</sup> Therefore, an alternative method using ancestral sequence reconstruction (ASR) has emerged as a potential solution,<sup>[51]</sup> which leverages evolutionary history to identify residues critical for maintaining protein function. Although less frequently used, ASR has shown potential as a tool



**Figure 1.** Phylogenetic tree of the 75 transaminase sequences selected from UniProtKB using BLAST with *Sp-ATA* (PDB: 3HMU) as a query. Sequences were filtered by length (400–600 residues) and identity thresholds (90% for selection, 97% for redundancy reduction). The tree was constructed using IQ-TREE with the JTT model and 1000 bootstrap replicates. Ancestral sequences were reconstructed with FastML, and the nodes of the same are marked with red dots. Branch colors represent bootstrap values.

for engineering thermostable proteins.<sup>[52–54]</sup> Previous studies have demonstrated that ASR can generate transaminases with broader substrate scope and significantly improved catalytic activity, with up to 20-fold increases in efficiency compared to modern enzymes.<sup>[55]</sup> Notably, ASR-derived transaminases have exhibited enhanced conversion of industrial feedstocks such as 12-amino-dodecanoic acid, a key precursor for nylon-12.<sup>[52]</sup> Structural analyses of the reconstructed enzymes revealed that substrate preference was influenced by broader structural features beyond the active site, highlighting the potential of ASR as a powerful alternative to traditional enzyme engineering.<sup>[55]</sup> Building on these findings, we propose to screen and reconstruct more stable ancestral enzymes from a gene library based on the amino acid sequences of *Sp*-ATA. By expressing and characterizing some of these ancestral enzymes, we aim to validate whether ASR can be applied to enhance the stability of transaminases.

## 2. Results and Discussion

### 2.1. ASR of ATAs

The enzymes in our panel exhibit 60%–95% sequence identity with the benchmark enzyme, *Sp*-ATA. All amino acid sequences of these enzymes are provided in the supplementary materials. To identify ancestral enzymes with potentially higher thermostability, we selected sequences from bacteria that thrive at temperatures above 40 °C for ASR. The resulting phylogenetic tree (Figure 1) was analyzed, and marginal reconstruction algorithms with optimized likelihood scores were applied in FastML to generate ancestral proteins. Ancestral sequences were chosen from the four closest nodes to *Sp*-ATA (from closest to farthest: N51, N50, N49, and N48) for further study. Among them, N48 and *Sp*-ATA have the lowest sequence identity of only 87.1%. Moving toward the youngest ancestor, N51, the sequence identity increases: from N49 with 88.7% to N50 with 92.2% and N51 showing the highest sequence identity of 94.3% (Table 1). The decision to reconstruct these specific ancestors is based on their evolutionary proximity to *Sp*-ATA. By selecting these closer nodes, we aim to retain a similar substrate range to that of *Sp*-ATA, as they are expected to exhibit more similar structural and functional characteristics compared to ancestors from more distant nodes within the phylogenetic tree. This assumption relies on the principle that genetic divergence accumulates over time, making more distant ancestors likely to show greater differences from the contemporary protein.<sup>[56]</sup>

Enzymes	Sequence identity [%]	RMSD [Å]
N48	87.1	0.90
N49	88.7	0.81
N50	92.2	0.57
N51	94.3	0.67

### 2.2. Kinetics and Substrate Range of the Ancestral Enzyme Variants

After successfully expressing and purifying soluble ancestral proteins (Figure S2, Supporting Information), further assays were

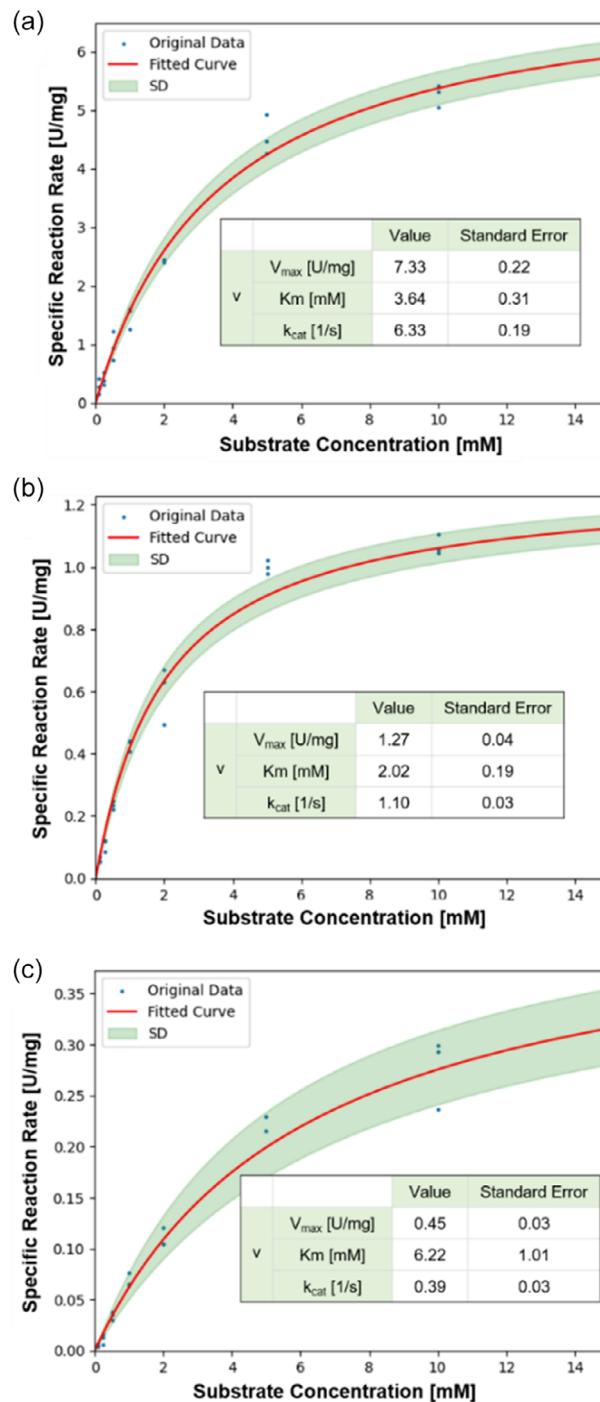
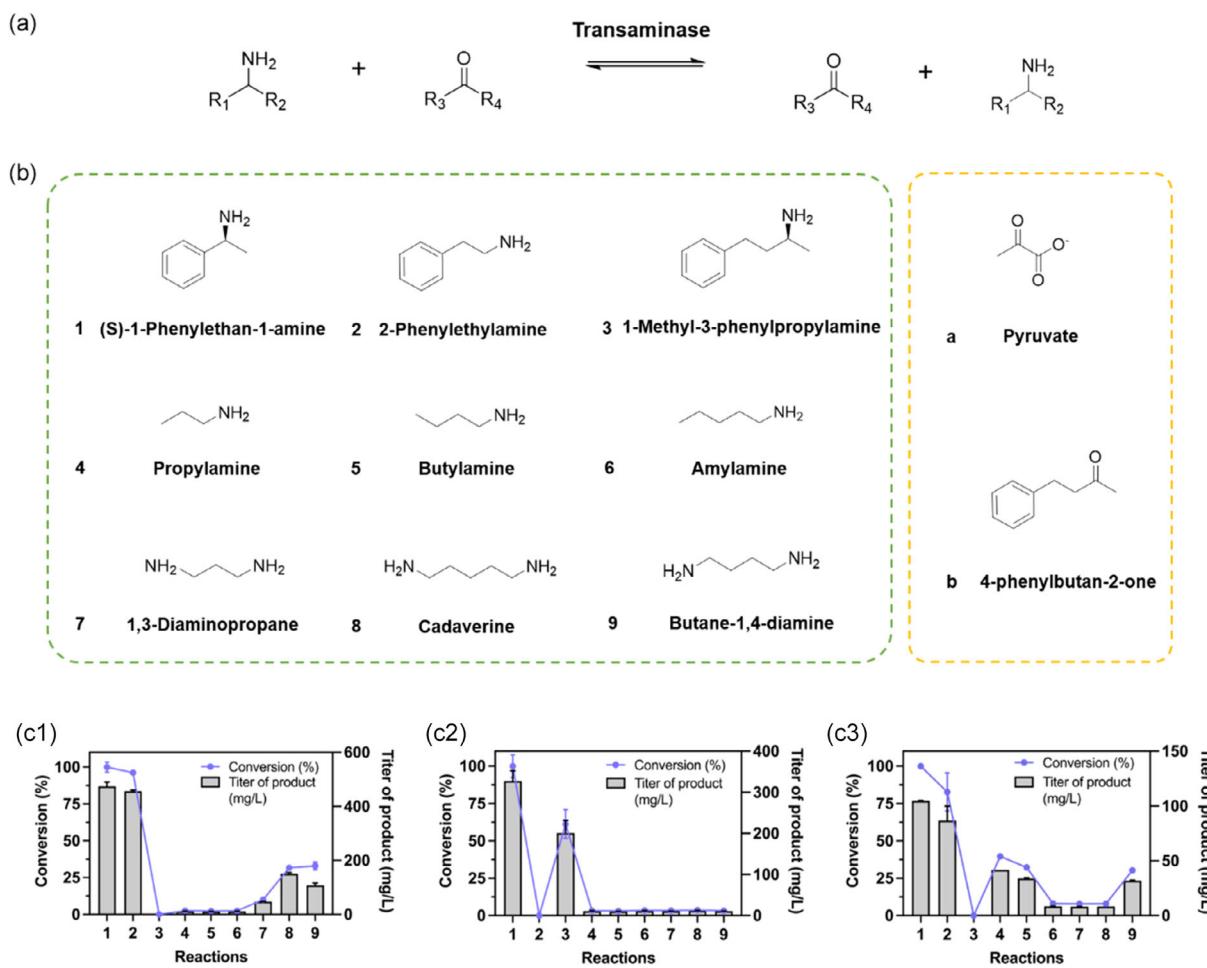


Figure 2. The Michaelis-Menten kinetic constants of a) *Sp*-ATA, b) N49, and c) N50. Kinetic parameters were determined using the acetophenone assay with PEA as the amine donor and Na-pyruvate as the amino acceptor. Reactions were performed in triplicate with six PEA concentrations (0.1–15 mM) and initiated by adding 10 µg enzyme with 5 mM PLP. Acetophenone formation was monitored at 245 nm, and standard deviations were typically <20%.

conducted to determine the optimal pH for the ancestral enzyme variants. The activities of the ancestral enzymes N48, N49, N50, and N51 were tested in 50 mM Tris-HCl buffer at pH 7.5, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer at pH 8.2, and *N*-cyclohexyl-2-aminoethanesulfonic acid (CHES) buffer at pH 9.0. The acetophenone assay was used to measure activity and confirm the selectivity of the selected ancestral enzymes. To ensure the functional relevance of the selected sequences, we experimentally tested the activity of the reconstructed ancestral enzymes. N51, N50, and N49 exhibited confirmed *S*-selective transaminase activity (Figure S3, Supporting Information). This experimental verification mitigates concerns about misclassification due to automatic annotation.<sup>[32]</sup> Additionally, structural analysis based on homology modeling described in the following section further supports the expected activity and selectivity of these enzymes. The results showed that all ancestral enzymes, except N48 (which, despite being soluble, exhibited no activity, likely due to improper folding or active site structural constraints), along with the wild-type *Sp*-ATA, performed best in the pH 9.0 CHES buffer (Figure S3, Supporting Information), consistent with previous results.<sup>[30]</sup> As for N51, although its phylogenetic position suggests activity similar to the wild type, its

significantly lower activity may be attributed to factors beyond sequence similarity, such as the need for post-translational modifications,<sup>[57]</sup> folding issues, or molecular interactions. N51 was excluded from further analyses due to its significantly lower activity compared to *Sp*-ATA wild type. To compare the catalytic efficiency of ancestral enzymes with *Sp*-ATA, we measured the kinetic parameters ( $V_{max}$ ,  $K_m$  and  $k_{cat}$ ) of *Sp*-ATA, N49, and N50 with the substrate (*S*)-PEA. As shown in Figure 2, the ancestral enzymes displayed decreased catalytic efficiency relative to *Sp*-ATA, which may indicate either a reduction in overall catalytic ability or that (*S*)-PEA is not their optimal substrate. A plausible explanation could be that these thermostable enzymes tend to have a more rigid structure at room temperature, requiring higher temperatures for optimal activity.<sup>[56,58]</sup> Additionally, evolutionary changes in the ancestral proteins may have led to the loss of activity, particularly in N48, which is the most ancestral of the four proteins. While the phylogenetic tree provides insights into evolutionary relationships, it does not always predict functional outcomes, especially when unexpected structural or dynamic changes occur.<sup>[59]</sup> Further modifications in the ancestral proteins might contribute to reduced activity, potentially due to new interactions between different regions (e.g., the active site), which could enhance stability but



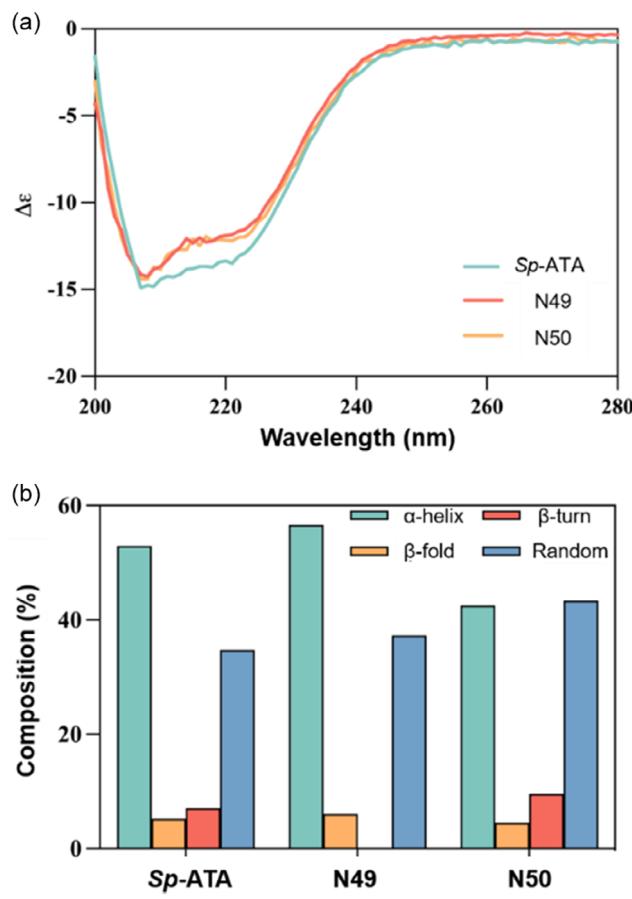
**Figure 3.** Determination of substrates scope. a) Schematic diagram of the transaminase reaction. b) Amino donors (1–9) and amino acceptors (a, b) used in the catalytic reaction. c) Catalytic performance of *Sp*-ATA (c1), N49 (c2), and N50 (c3) on various substrates. Titer of product ( $\text{mg L}^{-1}$ ) represents the concentration of the target compound produced. Reactions were conducted at 30 °C in 50 mM CHES buffer (pH 9.0) with 10 mM amine acceptors (a, b), 5 mM amine donors (1–9), 1 mM PLP, and 0.5 mg purified enzyme. Product titers ( $\text{mg L}^{-1}$ ) were quantified via HPLC after 20 h. All reactions were conducted in triplicates.

compromise activity. These findings are consistent with the broader goal of engineering enzymes for enhanced thermal stability, where improved stability may come at the cost of reduced catalytic activity due to structural rigidity or altered active site interactions.

The activities of the ancestral variants were then screened against nine amine donors, which included three aromatics (Figure 3b, 1–3), three aliphatic monoamines (Figure 3b, 4–6), and three aliphatic diamines (Figure 3b, 7–9), all at 5 mM concentration. For the aromatic amine donors, sodium pyruvate served as the amine acceptor. For the aliphatic donors, the aromatic acceptor 4-phenyl-2-butanone (PEB) was used to allow detection with a high-performance liquid chromatography (HPLC) UV detector (Figure 3b). Results displayed in Figure 3c indicate the product conversion rates with substrates 2–9 (the reactions numbered as the applied substrates) relative to (S)-PEA. Compared to *Sp*-ATA, N49 showed an interesting loss of conversion for 2 but a significant increase for 3. This shift in substrate preference may be attributed to structural changes in the active site, particularly the repositioning of key amino acids influencing substrate binding. Specifically, the outward-facing F92 (F94) side chain in N49 may increase the space in the binding pocket, making it easier for larger substrates like substrate 3 to enter. In contrast, substrate 2, which is smaller, may require a tighter binding pocket, and the altered pocket structure in N49 may reduce or eliminate its binding affinity. Additionally, N49 exhibited nearly no reactivity with reactions 7–9 compared to *Sp*-ATA. N50 demonstrated alternative substrate selectivity, with enhanced activity for reactions 4 and 5 and markedly reduced activity for 7 and 8. This shift suggests that changes in the active site architecture may have optimized N50 for binding smaller aliphatic amines. Notably, in N50, the inward-facing R417 (R426) may cause the active pocket to be more suited for smaller, straight-chain amine substrates, enhancing the binding affinity for substrates like propylamine and butylamine. This phenomenon has been confirmed in previous studies, where thermostable enzymes exhibited superior activity at higher temperatures compared to their less thermostable counterparts. Additionally, sequence and structural changes in the ancestral variants may also contribute to a decrease in activity.<sup>[58,60]</sup> We propose that structural changes in the active site have led to alterations in the substrate range of the ancestral enzymes. To better understand how conformational change affects the substrate domain and activity of N49 and N50, we analyzed the secondary structure of *Sp*-ATA, N49, and N50 using circular dichroism spectroscopy (Figure 4a). Compared to *Sp*-ATA, both N49 and N50 show distinct differences in the composition of their secondary structure elements ( $\alpha$ -helix,  $\beta$ -sheet,  $\beta$ -turn, and random coils) (Figure 4b). These structural differences may significantly impact the enzyme's substrate domain and activity, contributing to the observed shifts in substrate preference.

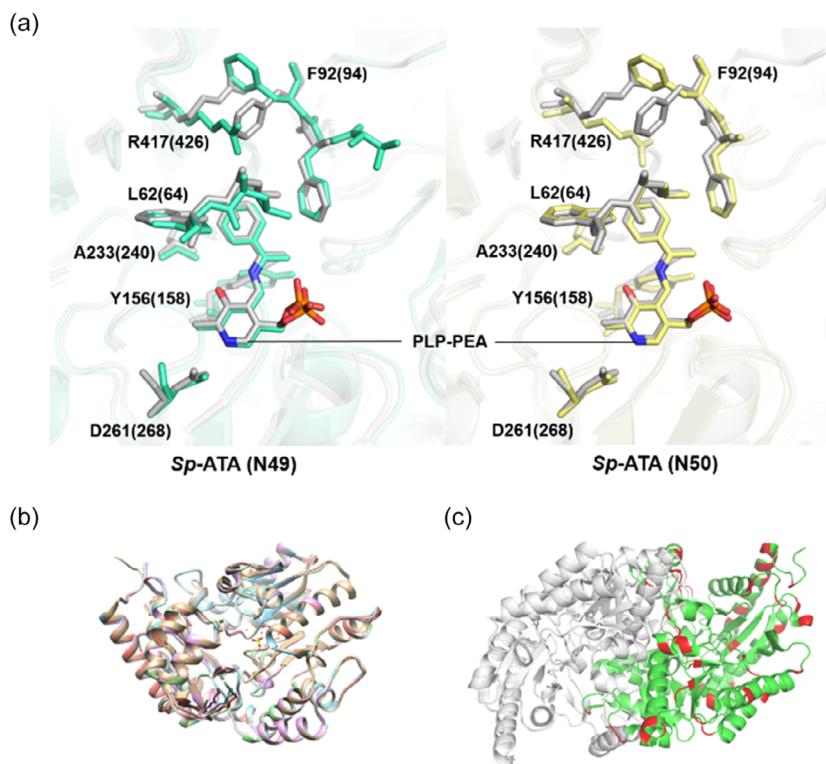
### 2.3. Computational Analysis of Structural Differences and Active Site Conservation in Ancestral Proteins

To visualize the impact of structural differences in the active site, homology modeling was employed. Alignment of the homology-modeled structures for the ancestral proteins with the wild-type



**Figure 4.** Conformation changes of ancestral enzymes compared to *Sp*-ATA. a) Circular dichroism spectra of *Sp*-ATA, N49, and N50 at 200–280 nm using a Chirascan CD spectrometer. Protein samples (0.1 mM) were measured in a 1 mm path-length cuvette. Spectra were averaged over five scans, baseline-corrected, and plotted using Prism GraphPad. b) Corresponding secondary structure composition.

*Sp*-ATA showed a high degree of similarity (Figure 5b). For a more precise analysis, docking simulations with PLP-PEA were conducted to assess the active site, yielding an interesting observation. As reported by Steffen-Munsberg et al.<sup>[28]</sup> the residues forming the active site—including Lys258, responsible for forming an internal aldimine with PLP, Asp259, which stabilizes and orients PLP during catalysis, and Arg417, which plays a role in substrate binding and specificity—are conserved across all four ancestral proteins as well as the wild-type *Sp*-ATA (Figure 5a). From a structural perspective, this suggests that the observed changes in activity and stability in the ancestral enzymes may arise from alterations in regions outside the active site. However, compared to *Sp*-ATA, the residue numbering in the ancestral proteins has changed, although their positions are nearly similar. The numbers in parentheses correspond to the ancestral proteins, while those outside correspond to the *Sp*-ATA wild type. Additionally, further analysis of the ancestral sequences through multiple sequence alignment (MSA) (Figure S2, Supporting Information) revealed amino acid insertions in certain regions, which are illustrated in one subunit of N49 (Figure S3, Supporting Information) and are normal occurrences in ancestral enzymes. Furthermore, in the same study on *Sp*-ATA, Steffen-Munsberg et al.<sup>[28]</sup> highlighted



**Figure 5.** Structural comparison of *Sp*-ATA and the ancestral enzymes. a) Docking simulations of the ancestral enzyme and the PLP-PEA ligand complex, the structures of *Sp*-ATA, N49, and N50 are shown in gray, cyan, and blue, respectively, with the corresponding PLP-PEA ligand in matching colors. Residue numbers in parentheses represent the ancestral enzymes, while those without parentheses represent *Sp*-ATA. b) Alignment of homology-modeled ancestral structures with *Sp*-ATA. c) Dimeric structure of *Sp*-ATA, with red regions highlighting sequence differences among the four ancestral enzymes, N48, N49, N50, and N51. Homology models were built in YASARA using hm\_build.mcr with a dimeric state. Ancestral sequences and templates were identified via PSI-BLAST, and 25 models per ancestor were generated and refined. Hybrid structures were analyzed, and molecular docking used these models alongside the *Sp*-ATA crystal structure (PDB: 3HMU) with PLP (PDB: 3FCR). Structures were aligned in YASARA, with SO<sub>4</sub> and obstructive H<sub>2</sub>O removed. PLP protonation was adjusted, a 5 Å simulation cell was set, and energy minimization was performed using the AMBER03 force field. Fifty docking simulations were run with AutodockLGA to evaluate binding poses. All images were generated using PyMOL.

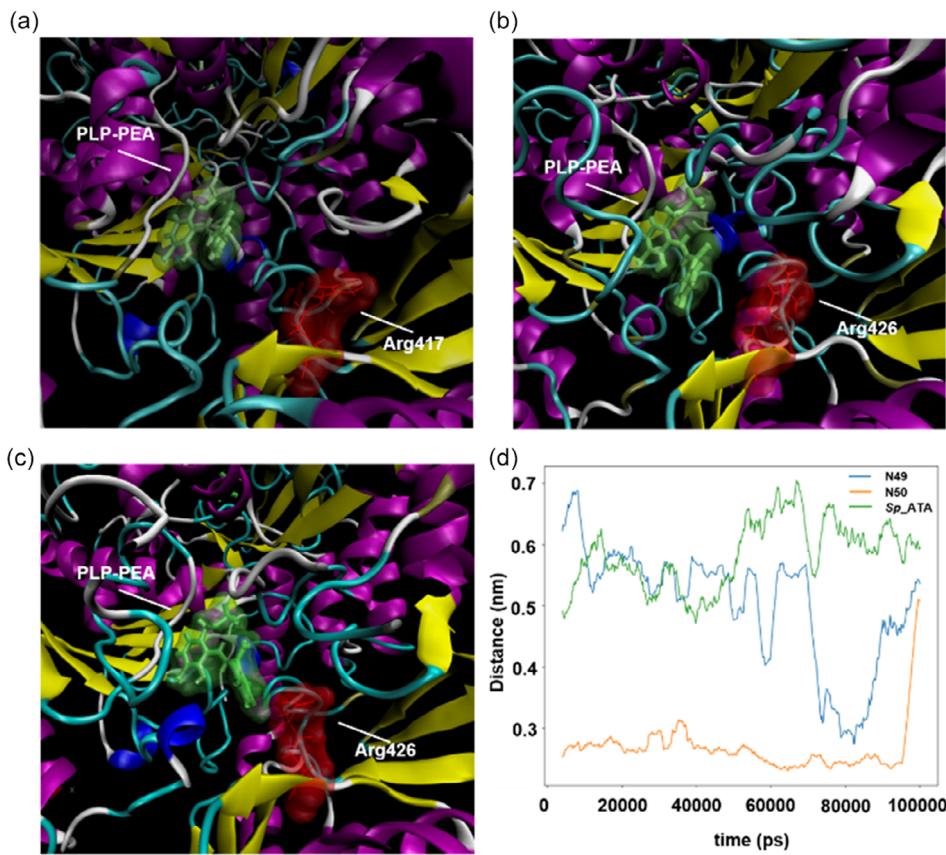
the key role of the “flipping” arginine residue (R417), which is crucial for dual substrate recognition. When R417 points towards the PLP, it can accept α-carboxylic substrates. Conversely, when the arginine flips upwards in the active site, it creates a larger binding pocket, enabling the binding of larger substrates, such as (S)-PEA. When each ancestral protein is aligned with *Sp*-ATA, this arginine adopts a more extended conformation towards PLP-PEA across all ancestral proteins compared to its conformation in *Sp*-ATA (Figure 5a). This may suggest that (S)-PEA is less favored as an amino donor for these enzymes.

To further validate the docking results and our hypothesis, molecular dynamics (MD) simulations were performed on the docked structures of *Sp*-ATA, N49, and N50 with PLP-PEA. The simulations revealed distinct conformational differences in the “flipping” arginine residue, further supporting our observations. As shown in Figure 6, in *Sp*-ATA, the “flipping” arginine extends outward, positioned farther from the PLP-PEA ligand. In N49, the arginine exhibits partial inward flipping, while in N50, it adopts the most pronounced inward conformation. Additionally, analysis of the shortest pairwise distance between the flipping arginine and PLP-PEA (Figure 6d) corroborates these structural differences: *Sp*-ATA exhibits the largest distance, followed by N49, while N50 has the shortest distance. These findings reinforce

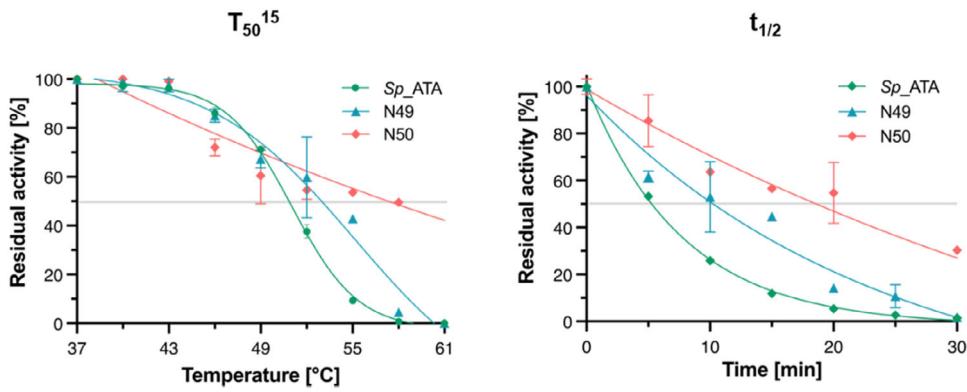
our hypothesis that the altered flipping behavior of this arginine residue in ancestral proteins may influence their substrate specificity and catalytic properties.

#### 2.4. The Ancestral Variants Exhibit a Significant Increase in Thermostability

Research has frequently highlighted the greater stability of ancestral enzymes compared to their modern counterparts. To investigate this stability, we measured half-inactivation temperature ( $T_{50}^{15}$ ) and the half-life ( $t_{1/2}$ ) of the ancestral enzymes (Figure 7). For instance, N49 showed a  $T_{50}^{15}$  of 54.9 °C, an increase of about 4 °C over the  $T_{50}^{15}$  of the wild-type *Sp*-ATA, which is 51 °C (Figure 7a,c). Additionally, the half-life of N49 extended twofold, from 5 to 10.3 min (Figure 7d,f). N50, in comparison, exhibited even more significant stability enhancements, with a  $T_{50}^{15}$  of 6 °C increase and a half-life extended by 3.7 times to 18.6 min (Figure 7b,e). These results suggest the potential of ancestral reconstruction as an effective strategy for enhancing enzyme stability. To address concerns regarding potential renaturation effects after heat treatment, we ensured strict consistency in temperature control and conducted activity assays immediately after cooling. While renaturation could theoretically influence measured



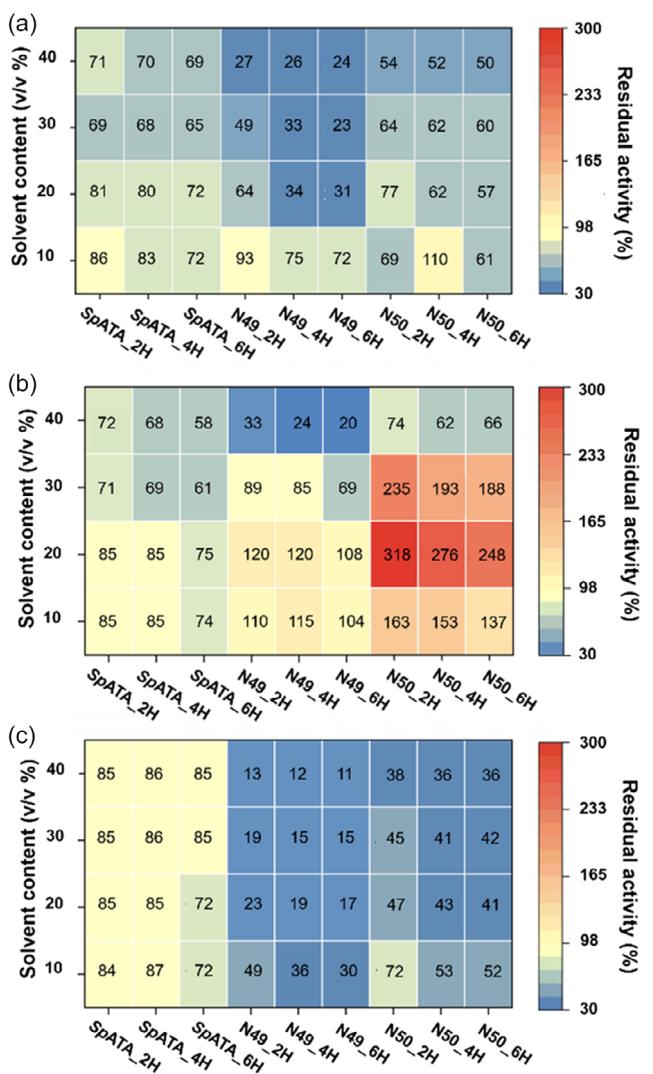
**Figure 6.** MD simulation of the docked PLP-PEA complex with *Sp*-ATA, N49, and N50. a–c) Structural visualization of the PLP-PEA complex (green surface) and the “flipping” arginine (red surface) in a) *Sp*-ATA, b) N49, and c) N50. In *Sp*-ATA, the arginine extends outward, positioned farthest from the PLP-PEA. In N49, it partially flips inward, while in N50, it adopts the most inward conformation. d) The shortest pairwise distance between the “flipping” arginine and PLP-PEA shows a trend consistent with the structures: *Sp*-ATA has the largest distance, followed by N49, with N50 having the shortest. MD simulation details: Simulations were performed using GROMACS 2024.2 with the AMBER99SB-ILDN force field for proteins and TIP3P for water. Ligands were parameterized using GRAFF via ACPYPE. Systems were solvated in a triclinic box ( $\approx 1500 \text{ nm}^3$ ), neutralized with  $\text{Cl}^-$  ions, minimized, equilibrated (NVT/NPT, 200 ps at 300 K), and subjected to a 100 ns production run. Visualization was performed with VMD, and quantitative analysis used GROMACS gmx pairdist, plotting results in Python.



**Figure 7.** The thermostability of *Sp*-ATA and its ancestral variants (N49, N50, and N51) was evaluated by determining the half-inactivation temperature ( $T_{50}^{15}$ ) and the half-life ( $t_{1/2}$ ).  $T_{50}^{15}$  denotes the temperature at which enzyme activity is reduced to 50% after 15 min of heat exposure, whereas the  $t_{1/2}$  represents the time required for enzyme activity to decrease to 50% at 55 °C. The enzymes were incubated at 55 °C for up to 60 min, followed by cooling, and their activity was measured at 22 °C. An exponential decay model in GraphPad Prism 10 was used to analyze the data and determine the point of 50% activity loss. Furthermore, the enzymes were subjected to 15-min incubations at temperatures ranging from 37 to 61 °C and then cooled before assessing activity. All experiments were conducted in triplicate.

activity, the observed significant increase in thermostability among ancestral enzymes strongly suggests an inherent stabilization effect rather than a transient recovery. The marked increases in

$T_{50}^{15}$  and half-life highlight the robustness of the enzyme variants under thermal stress, indicating that evolutionary adaptations retained in ancestral enzymes can confer substantial stability



**Figure 8.** Enzyme tolerance to different concentrations of a) MeOH, b) DMSO, and c) acetonitrile after incubation for 2, 4, and 6 h at 25 °C. Purified enzymes were incubated in a mixture of organic solvent (0%, 10%, 20%, 30%, and 40% v/v) and CHES buffer. Residual enzyme activity was then determined using an acetophenone assay and measured by HPLC. Reactions without enzymes as blanks, while those without organic solvents were controls. All reactions were performed in triplicates.

advantages. This stability may stem from structural features that enhance robustness at elevated temperatures, a hypothesis that is worth further investigation to identify specific structural contributors to this enhanced stability. The superior stability of N49 and N50 makes them particularly promising for industrial processes, thus broadening the practical applications of transaminases. While ancestral reconstruction offers stability benefits, a well-known limitation of the ASR method is the challenge of linking structural modifications to functional outcomes.<sup>[61]</sup> Numerous amino acid substitutions are present in the ancestral proteins, scattered across different protein regions—particularly in loops and  $\alpha$ -helices (Figure 5c and S1, Supporting Information). Although pinpointing the exact contribution of each substitution remains challenging, the distribution of these substitutions may guide further rational design, focusing on regions that could improve the thermal stability of other transaminases.

## 2.5. Organic Solvents Stability

To determine the enzyme tolerance to the reaction media, the wild type, N49, and N50 were incubated in 10, 20, 30, and 40% concentrations of MeOH, dimethyl sulfoxide (DMSO), and acetonitrile for 2, 4, and 6 h, followed by a 20 h reaction to measure the acetophenone yield (Figure 8). Results showed that the presence of DMSO significantly enhanced the activity of ancestral enzymes N49 and N50, a phenomenon not observed in the wild-type Sp-ATA (Figure 8b). Specifically, N49 displayed an increase in activity in 10%–30% DMSO, with activity more than triple after a 2-h incubation in 20% DMSO. N50 also showed slight activity enhancements in 10% and 20% DMSO. However, both ancestral enzymes experienced activity loss in 40% of DMSO. In contrast, the wild-type Sp-ATA performed poorly in all DMSO concentrations, retaining only about half of its activity in 40% DMSO. In the presence of MeOH, the wild-type Sp-ATA exhibited only a minor decrease in catalytic activity with increasing concentration and incubation time (Figure 8a). For instance, after a 6-h incubation in 40% MeOH, Sp-ATA retained around 70% of its activity. However, N49 showed significant declines in activity with increasing MeOH concentration and incubation time. After 6 h in 10% MeOH, N49 retained 72% of its activity, but this dropped to 31% in 20% MeOH and only 24% in 40% MeOH. In acetonitrile, the ancestral enzymes showed the poorest tolerance, while the wild-type Sp-ATA displayed some degree of resistance (Figure 8c). This variation in enzyme stability and activity across different solvents can be attributed to how the solvents interact with the protein. Acetonitrile is a polar aprotic solvent, which can destabilize the enzyme's native conformation, causing protein unfolding and reduced activity.<sup>[62]</sup> In contrast, DMSO, also a polar aprotic solvent with strong hydrogen bond-accepting ability,<sup>[63]</sup> possibly stabilizes the enzyme's secondary and tertiary structures by preserving essential surface hydration through its S=O structure, thereby enhancing enzyme stability and activity.<sup>[64]</sup> Overall, this experiment reveals differences in reaction medium tolerance between the ancestral enzymes and the wild-type Sp-ATA, which provides understanding and insights for developing industrial enzymes.

## 3. Conclusion

This study utilized ASR with the amino acid sequence of Sp-ATA as a template. Four ancestral enzymes (N48, N49, N50, and N51) were selected from branches closest to Sp-ATA for further investigation. The study showed that all four ancestral enzymes were successfully expressed in *Escherichia coli* (*E. coli*). Among these, N49 and N50 exhibited promising catalytic activity and substrate specificity distinct from Sp-ATA, whereas N48 showed no activity with the amino donor (S)-PEA, and N51 had negligible activity. The ancestral enzymes N49 and N50 also demonstrated improved thermal stability with an increase in  $T_{50}^{15}$  values by 3.9 and 6 °C, respectively, relative to Sp-ATA, and their half-lives ( $t_{1/2}$ ) at 55 °C were extended 2.06 to 3.72 times before reaching inactivation. Furthermore, we compared solvent tolerance across the ancestral enzymes and found that both N49 and N50 displayed a preference for DMSO. In lower concentrations of

DMSO, they exhibited significantly enhanced catalytic activity. However, their tolerance to methanol and acetonitrile was lower than that of the wild-type enzyme. These findings provide valuable insights into enzyme adaptability to specific solvents and strategies for enhancing enzyme thermal stability, thereby supporting the use of biocatalysts in industrial applications. Nonetheless, a deeper understanding of enzyme structure will be crucial to precisely control modifications, preserving high activity and substrate specificity while boosting stability.

## 4. Experimental Section

### Materials

All chemicals of analytical or higher grade were purchased from Sigma-Aldrich Sweden. Yeast extract, tryptone, agar and ampicillin were purchased from Merck Sweden. Agarose and NaCl were obtained from VWR Sweden. The Ni/NTA His Bind resin for His-trap column was purchased from Merck Sweden. The PD10 column was purchased from Cytiva Sweden.

### Medium, Plasmids, and Strains

The lysogeny broth medium (LB, 10 g L<sup>-1</sup> NaCl, 10 g L<sup>-1</sup> tryptone, and 5 g L<sup>-1</sup> yeast extract, pH 7.0) was utilized for growing *E. coli* cultures. Protein expression was conducted in terrific broth (TB, 16 g L<sup>-1</sup> tryptone, 10 g L<sup>-1</sup> yeast extract, NaCl, pH 7.0). Agar was added to a final concentration of 1.5% to prepare solid medium. 5 g L<sup>-1</sup> Ampicillin was added to the medium at a concentration of 100 µg mL<sup>-1</sup>. The strains *E. coli* XL1-Blue and BL21(DE3) were used for plasmid amplification and expression respectively.

### Sequence Collection

The amino acid sequence of *Sp*-ATA (PDB: 3HMU) was used as a query to retrieve homologous transaminase sequences. The search was conducted against the UniProtKB database (<https://www.uniprot.org/>) using the basic local alignment search tool (BLAST) with default parameters. Unique sequences were identified and cleaned using the CD-HIT<sup>[65]</sup> program with a 90% sequence identity threshold to ensure nonredundant data. The resulting sequences were manually reviewed to remove false positives, including those shorter than 400 or longer than 600 residues, as well as sequences with clear duplicates or missing fragments. Further data cleaning was performed by removing sequences with more than 97% identity to *Sp*-ATA. Furthermore, an automated script was created using Python, incorporating Matplotlib 3.7.1,<sup>[66]</sup> NumPy 1.26.2,<sup>[67]</sup> and Pandas 1.5.3 module (Zenodo DOI: 10.5281/zenodo.13819579). The script was used to filter out sequences from microorganisms with optimal growth temperatures below 40 °C. Furthermore, we validated the presence of the PLP-dependent transferase domain using InterPro annotations,<sup>[68]</sup> confirming that all sequences belong to the same structural and phylogenetic family as *Sp*-ATA. These filtering criteria minimized the risk of including unrelated or nonfunctional sequences and ensured a robust foundation for MSA and phylogenetic tree construction. The final dataset contained 75 sequences.

### MSA and Phylogenetic Analysis

A multiple sequence alignment (MSA) of the dataset was generated using MAFFT<sup>[69]</sup> with the FFT-NS-I × 1000 algorithm (fast fourier transform-based progressive alignment with iterative refinement), which is known for its high alignment accuracy and efficiency in

handling large-scale datasets. This advanced alignment strategy incorporates FFT and iterative optimization to achieve high computational efficiency while maintaining alignment precision. Given the scale of the dataset and the need for both speed and accuracy, the FFT-NS-I × 1000 algorithm was particularly suitable for the analysis of this work. Multiple runs of the algorithm were performed to obtain consistent results, which confirmed the stability of the outcome. Additionally, the alignment results were compared with those from ClustalW, yielding similar outcomes and further validating the accuracy of the FFT-NS-I × 1000 algorithm. Gaps within the aligned sequences were removed, and a maximum likelihood (ML) phylogenetic tree was constructed in IQ-TREE using the Jones-Taylor-Thornton substitution (JTT) model with 1000 bootstrap replicates. ASR was performed with FastML,<sup>[70]</sup> and the ancestral tree was visualized in iTOL.<sup>[71]</sup> Homology modeling of the ancestral sequences was carried out in YASARA (version 23.12.24)<sup>[72]</sup> and visualized using PyMOL (version 4.6.0).<sup>[73]</sup>

### Expression and Purification

The genes encoding ancestral enzymes were synthesized and cloned into pET22b (+) with ampicillin resistance by Twist Bioscience (San Francisco, USA). The plasmids were transformed into *E. coli* XL1-Blue. For *Sp*-ATA, the gene was previously cloned into pET22b (+) with ampicillin resistance. For overexpression, plasmids pET22b-*Sp*-ATA, pET22b-N48, pET22b-N49, pET22b-N50, and pET22b-N51 were transformed into *E. coli* BL21(DE3). Colonies were grown in LB with 50 µg mL<sup>-1</sup> ampicillin at 37 °C, 220 RPM. Protein expression was induced with 0.2 mM IPTG at OD<sub>600</sub> = 0.6, followed by incubation at 20 °C, 200 RPM for 20 h. Cells were harvested (4,000 × g, 20 min, 4 °C), resuspended (1:1.5, w/v<sup>-1</sup>) in lysis buffer, lysed via sonication, and centrifuged (15 000 × g, 20 min, 4 °C) to collect the supernatant. For purification, 1 mL of Ni-NTA agarose was loaded into a column, and the cell lysate was allowed to pass the column. The column was then washed with washing buffers 1 and 2 (50 mM NaCl, and 20 mM or 50 mM imidazole) to remove contaminants. Target protein was eluted with 0.5 mL of buffer with 50 mM NaCl, and 300 mM imidazole. Desalting was performed using a PD-10 column. Purified proteins were analyzed on 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE, Bio-Rad TGX gel), run at 220 V for 30 min in 1× SDS-Tris buffer. SeeBlue Plus2 Pre-stained Protein Standard was used as standard and Sigma Laemmli 2× Concentrate were used as sample buffers (20 µL per sample). For dimer formation, enzymes were incubated with 5 mM PLP at 4 °C overnight. Then, excess PLP was removed using PD10 column with either 50 mM HEPES buffer (pH 8.2) or 50 mM CHES buffer (pH 9.0).

### Enzyme Activity Assay

The conversion of (S)-1-phenylethylamine (PEA) to acetophenone<sup>[74]</sup> was monitored at 245 nm using an extinction coefficient of 12 mM<sup>-1</sup> cm<sup>-1</sup>. Acetophenone formation was measured every 5 s over a 3-min period with a Cary50 ultraviolet (UV)-visible spectrophotometer, using 1 mL UV cuvettes. The assay was conducted at 22 °C, with a typical reaction mixture (1 mL) consisting of 5–50 µg of an enzyme, 2.5 mM PEA, and 2.5 mM sodium pyruvate in different buffers. To determine the optimal pH of the ancestral enzyme, the reactions were carried out in 50 mM Tris-HCl buffer at pH 6.0 and pH 7.5, in HEPES buffer at pH 8.2, and in CHES buffer at pH 9.0. All measurements were performed in triplicate, and protein samples were incubated at 22 °C for 1 h before the assay.

### Kinetic Analysis

Kinetic assays were performed in triplicate using a microplate reader. The acetophenone assay was performed to assess the kinetic analysis.

The PEA was an amine donor, and Na-pyruvate was used as an amino acceptor. To obtain the Michaelis–Menten constant ( $K_{M}^{app}$ ), six different concentrations of (S)-PEA, ranging from 0.1 to 15 mM, were used. Reactions were started by adding 10 µg enzyme with 5 mM PLP. The generation of acetophenone followed over time by measuring the 245 nm absorbance. The standard deviations for the determined kinetic parameters were typically <20%.

### Secondary Structure Assessment

Circular dichroism (CD) spectra were recorded using a Chirascan CD spectrometer. Protein samples were prepared at a concentration of 0.1 in 50 mM CHES buffer at pH 9.0, with a sample volume of 400 µL. The measurements were carried out in a 1 mm path-length cuvette. The UV range spanned from 200 to 280 nm, using a quartz cell with a 0.2 mm path length. The scan speed was set to 50 nm min<sup>-1</sup>, with a response time of 1 s, and a total of five scans were accumulated. Baseline correction was applied to all spectra, and the data was subsequently plotted using Prism GraphPad (version 10.2).

### Thermostability Determination

Both the half-inactivation temperature ( $T_{50}^{15}$ ) and half-life ( $t_{1/2}$ ) values characterize the thermostability of the enzyme.  $T_{50}^{15}$  refers to the temperature at which the enzyme activity was reduced to half of the original activity after heat treatment at a continuous temperature for 15 min. The enzyme solution was incubated for 15 min at temperatures of 37, 40, 43, 46, 49, 52, 55, 58, and 61 °C and then cooled on ice for 10 min. Using the Levenberg–Marquardt iterative algorithm, the data were fitted to a four-parameter Boltzmann sigmoidal function. The formula is presented in Equation (1).

$$R = A + (B - A)/(1 + e^{(T_m - T)/C}) \quad (1)$$

where  $A$  and  $B$  represent pretransition and post-transitional residual activity, respectively, at temperature  $T$ .  $C$  is the slope factor, controlling the steepness of the curve or the temperature sensitivity.  $T$  is the current temperature.  $T_0$  is the temperature in the absence of treatment. In the absence of treatment,  $T_m = T_0$ , when there is no treatment.

Similarly,  $t_{1/2}$  is defined as the time when the residual activity of *Sp*-ATA and its ancestors was reduced to 50% of its original activity. For the study, we measured the  $t_{1/2}$  at 55 °C. The purified *Sp*-ATA, N49, N50, and N51 variants were incubated for 0–60 min at 55 °C and then cooled on ice for 10 min. The enzyme activity test was performed at 22 °C for 3 min with a spectrophotometer under 245 nm. The data was fitted with an exponential function model:  $[Y = (Y_0 - \text{Plateau}) * \exp(-K*X) + \text{Plateau}]$  ( $Y_0$ , Y-value at  $X = 0$ , same units as Y; Plateau, Y-value at infinite time, same units as Y; and K, rate constant, reciprocal of X-axis units) by nonlinear regression using GraphPad Prism 10 (GraphPad Software Inc., La Jolla, CA, USA) and then the 50% of relative enzyme activity were determined.

### Substrate Scope Screening

To evaluate the substrate scope of ancestral enzymes compared to *Sp*-ATA, nine amine donors were selected. The substrate spectrum of the ancestral enzymes was determined at 30 °C in a 0.5 mL reaction mixture containing 50 mM CHES buffer (pH 9.0), 10 mM acceptor (sodium pyruvate and PEB for R4-9), the nine amine donors (Figure 3b), 1 mM PLP, and 0.5 mg of purified enzyme. The conversion of (S)-PEA (following acetylation) to acetophenone was measured after 20 h via HPLC, using a flow rate of 1 mL min<sup>-1</sup>, a mobile phase of 20/80% v v<sup>-1</sup> acetonitrile and water with 1% trifluoroacetic acid, and detection at 245 nm.

### Solvent Tolerance

The impact of different concentrations of water-miscible organic solvents (0%, 10%, 20%, 30%, and 40% v v<sup>-1</sup>), viz., DMSO, methanol, and acetonitrile, on enzyme activity was assessed by incubating purified enzymes in a mixture of organic solvent and CHES buffer for 2, 4, and 6 h at 25 °C. Residual enzyme activity was then determined using an acetophenone assay and measured by HPLC. A reaction without enzyme served as the blank, while a reaction without organic solvent was used as the control. All reactions were performed in triplicates.

### Homology Modeling and Molecular Docking Simulations

The homology modeling in YASARA used the macro hm\_build.mcr, modifying the oligo state from 4 to 2 for dimeric structures. It identified ancestral sequences and templates through PSI-BLAST. After secondary structure predictions, 25 models per ancestor were built and refined in two steps. The best models were combined into hybrid structures for final analysis. Molecular docking simulations utilized the homology-modeled structures and crystal structure 3HMU for *Sp*-ATA, incorporating the essential cofactor PLP from PDB: 3FCR. Structures were aligned in YASARA, with SO<sub>4</sub><sup>2-</sup> and obstructive H<sub>2</sub>O molecules removed. The PLP protonation state was adjusted, and a simulation cell extending 5 Å was defined and filled with water. The AMBER03 force field was used for energy minimization. Pro-(S)-quinonoid complexes of PEA (PLP\_PEA) were constructed, maintaining specific dihedral angles. Fifty independent docking simulations were executed with dock\_run.mcr and AutodockLGA, evaluating binding poses and modes for PLP\_PEA across structures.

### MD Simulation

MD simulations were performed on the docked structures of *Sp*-ATA, N49, and N50, each complexed with PLP\_PEA and PLP at the dimer active site. All simulations were conducted using GROMACS 2024.2 (Zenodo DOI: 10.5281/zenodo.11148655). The AMBER99SB-ILDN force field was used to parametrize the protein, while the TIP3P was employed for water. Ligands were parametrized using the GRAFF force field<sup>[75,76]</sup> via ACPYPE.<sup>[77]</sup> Each system was placed in a triclinic periodic box with a volume of ≈1500 nm<sup>3</sup>, filled with TIP3P water molecules, and neutralized with Cl<sup>-</sup> ions. Energy minimization was performed using the steepest descent algorithm for 50 000 steps or until convergence at −1000 kJ nm mol<sup>-1</sup>. The systems were equilibrated under NVT (constant volume and temperature equilibration run), followed by NPT (constant pressure and temperature equilibration run), for 200 ps each at 300 K. A 100 ns production run was then carried out, and the resulting trajectories were analyzed. The VMD program<sup>[78]</sup> was used to visualize the docked substrate and the “flipping” arginine. Quantitative analysis was conducted using the GROMACS gmx pairdist function, and the results were plotted using Python 3.9.16 with Matplotlib 3.7.1.

### Acknowledgements

The authors are grateful for financial support from KTH Royal Institute of Technology.

### Conflict of Interest

The authors declare no conflict of interest.

## Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

**Keywords:** ancestral sequence reconstruction • *Silicibacter pomeroyi* • stability •  $\omega$ -transaminase

- [1] P. N. Devine, R. M. Howard, R. Kumar, M. P. Thompson, M. D. Truppo, N. J. Turner, *Nat. Rev. Chem.* **2018**, *2*, 409.
- [2] E. L. Bell, W. Finnigan, S. P. France, A. P. Green, M. A. Hayes, L. J. Hepworth, S. L. Lovelock, H. Niikura, S. Osuna, E. Romero, *Nat. Rev. Methods Primers* **2021**, *1*, 1.
- [3] R. A. Sheldon, J. M. Woodley, *Chem. Rev.* **2018**, *118*, 801.
- [4] J. B. Pyser, S. Chakrabarty, E. O. Romero, A. R. Narayan, *ACS Central Sci.* **2021**, *7*, 1105.
- [5] R. Wohlgemuth, *Curr. Opin. Biotechnol.* **2010**, *21*, 713.
- [6] K. Faber, K. Faber, *Biotransformations In Organic Chemistry: A Textbook*, Vol. 6, Springer, Heidelberg **2011**.
- [7] V. Kumar, V. Bansal, A. Madhavan, M. Kumar, R. Sindhu, M. K. Awasthi, P. Binod, S. Saran, *Bioengineered* **2022**, *13*, 4309.
- [8] S. Simić, E. Zukić, L. Schmermund, K. Faber, C. K. Winkler, W. Kroutil, *Chem. Rev.* **2021**, *122*, 1052.
- [9] G. Rossino, M. S. Robescu, E. Licastro, C. Tedesco, I. Martello, L. Maffei, G. Vincenti, T. Bavaro, S. Collina, *Chirality* **2022**, *34*, 1403.
- [10] S. Mathew, D. Renn, M. Rueping, *ACS Catal.* **2023**, *13*, 5584.
- [11] P. D. de Maria, *Curr. Opin. Green Sustainable Chem.* **2021**, *31*, 100514.
- [12] E. Bustó, V. Gotor-Fernández, V. Gotor, *Chem. Rev.* **2011**, *111*, 3998.
- [13] F. van Rantwijk, R. A. Sheldon, *Tetrahedron* **2004**, *60*, 501.
- [14] L. S. Vidal, C. L. Kelly, P. M. Mordaka, J. T. Heap, *Biochim. et Biophys. Acta BBA-Proteins Proteomics* **2018**, *1866*, 327.
- [15] S. W. May, *Curr. Opin. Biotechnol.* **1999**, *10*, 370.
- [16] K. Faber, W. Kroutil, *Curr. Opin. chem. Biol.* **2005**, *9*, 181.
- [17] M. S. Malik, E.-S. Park, J.-S. Shin, *Appl. Microbiol. Biotechnol.* **2012**, *94*, 1163.
- [18] B. Yuan, D. Yang, G. Qu, N. J. Turner, Z. Sun, *Chem. Soc. Rev.* **2024**, *53*, 227.
- [19] J. Mangas-Sánchez, *Biocatalysis in Asymmetric Synthesis*, Elsevier, **2024**, pp. 209–236.
- [20] J. Liu, W. Kong, J. Bai, Y. Li, L. Dong, L. Zhou, Y. Liu, J. Gao, R. T. B. Allen, N. J. Turner, *Chem. Catal.* **2022**, *2*, 1288.
- [21] M. D. Patil, G. Grogan, A. Bommarius, H. Yun, *Catalysts* **2018**, *8*, 254.
- [22] S. A. Kelly, S. Pohle, S. Wharry, S. Mix, C. C. Allen, T. S. Moody, B. F. Gilmore, *Chem. Rev.* **2018**, *118*, 349.
- [23] S. Mathew, H. Yun, *ACS Catal.* **2012**, *2*, 993.
- [24] F. Ruggieri, J. C. Campillo-Brocal, S. Chen, M. S. Humble, B. Walse, D. T. Logan, P. Berglund, *Sci. Rep.* **2019**, *9*, 16946.
- [25] F. Guo, P. Berglund, *Green Chem.* **2017**, *19*, 333.
- [26] K. E. Cassimjee, B. Manta, F. Himo, *Org. Biomol. Chem.* **2015**, *13*, 8453.
- [27] F. Steffen-Munsberg, C. Vickers, H. Kohls, H. Land, H. Mallin, A. Nobili, L. Skalden, T. van den Bergh, H.-J. Joosten, P. Berglund, *Biotechnol. Adv.* **2015**, *33*, 566.
- [28] 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.
- [29] L. M. Merz, L. M. van Langen, P. Berglund, *ChemCatChem* **2023**, *15*, e202201174.
- [30] F. Steffen-Munsberg, C. Vickers, A. Thontowi, S. Schätzle, T. Tumlirsch, M. Svedendahl Humble, H. Land, P. Berglund, U. T. Bornscheuer, M. Höhne, *ChemCatChem* **2013**, *5*, 150.
- [31] M. Kollipara, P. Matzel, M. Sowa, S. Brott, U. Bornscheuer, M. Höhne, *Appl. Microbiol. Biotechnol.* **2022**, *106*, 5563.
- [32] M. Höhne, S. Schätzle, H. Jochens, K. Robins, U. T. Bornscheuer, *Nat. Chem. Biol.* **2010**, *6*, 807.
- [33] C. García-Galan, Á. Berenguer-Murcia, R. Fernandez-Lafuente, R. C. Rodrigues, *Adv. Synth. Catal.* **2011**, *353*, 2885.
- [34] A. S. Bommarius, M. F. Paye, *Chem. Soc. Rev.* **2013**, *42*, 6534.
- [35] P. V. Iyer, L. Ananthanarayanan, *Proc. Biochem.* **2008**, *43*, 1019.
- [36] M. N. Gupta, *Eur. J. Biochem.* **1992**, *203*, 25.
- [37] C. Xu, S. Tong, L. Sun, X. Gu, *Carbohydr. Polym.* **2023**, *321*, 121319.
- [38] C. Vieille, G. J. Zeikus, *Microbiol. Mol. Biol. Rev.* **2001**, *65*, 1.
- [39] F. Rigoldi, S. Donini, A. Redaelli, E. Parisini, A. Gautieri, *APL Bioeng.* **2018**, *2*, 011501.
- [40] S. Prasad, I. Roy, *Recent pat. Biotechnol.* **2018**, *12*, 33.
- [41] H. Land, J. C. Campillo-Brocal, M. Svedendahl Humble, P. Berglund, *ChemBioChem* **2019**, *20*, 1297.
- [42] Z. Xie, L. Zhai, D. Meng, Q. Tian, Z. Guan, Y. Cai, X. Liao, *3 Biotech* **2020**, *10*, 1.
- [43] A. V. Gribenko, M. M. Patel, J. Liu, S. A. McCallum, C. Wang, G. I. Makhatadze, *Proc. Natl. Acad. Sci.* **2009**, *106*, 2601.
- [44] J.-R. Cao, F.-F. Fan, C.-J. Lv, H.-P. Wang, Y. Li, S. Hu, W.-R. Zhao, H.-B. Chen, J. Huang, L.-H. Mei, *Front. Chem.* **2021**, *9*, 664156.
- [45] A. Q. Stiving, Z. L. VanAernum, F. Busch, S. R. Harvey, S. H. Sarni, V. H. Wysocki, *Anal. Chem.* **2018**, *91*, 190.
- [46] Y. Wang, P. Xue, M. Cao, T. Yu, S. T. Lane, H. Zhao, *Chem. Rev.* **2021**, *121*, 12384.
- [47] S. Atsumi, J. C. Liao, *Appl. Environ. Microbiol.* **2008**, *74*, 7802.
- [48] C.-Y. Liu, L. C. Severin, C.-J. Lyu, W.-L. Zhu, H.-P. Wang, C.-J. Jiang, L.-H. Mei, H.-G. Liu, J. Huang, *Biochem. Eng. J.* **2021**, *167*, 107926.
- [49] R. R. Chirumamilla, R. Muralidhar, R. Marchant, P. Nigam, *Mol. Cell. Biochem.* **2001**, *224*, 159.
- [50] D. Yi, T. Bayer, C. P. Badenhorst, S. Wu, M. Doerr, M. Höhne, U. T. Bornscheuer, *Chem. Soc. Rev.* **2021**, *50*, 8003.
- [51] R. E. S. Thomson, S. E. Carrera-Pacheco, E. M. J. Gillam, *J. Biol. Chem.* **2022**, *298*, 102435.
- [52] K. Prakinee, S. Phaisan, S. Kongjaroon, P. Chaiyen, *JACS Au* **2024**, *4*, 4571.
- [53] M. A. Spence, J. A. Kaczmarski, J. W. Saunders, C. J. Jackson, *Curr. Opin. Struct. Biol.* **2021**, *69*, 131.
- [54] Y. Gumulya, J.-M. Baek, S.-J. Wun, R. E. Thomson, K. L. Harris, D. J. Hunter, J. B. Behrendorff, J. Kulig, S. Zheng, X. Wu, *Nat. Catal.* **2018**, *1*, 878.
- [55] M. Wilding, T. S. Peat, S. Kalyaanamoorthy, J. Newman, C. Scott, L. S. Jermiin, *Green Chem.* **2017**, *19*, 5375.
- [56] M. Di Giulio, *J. Theor. Biol.* **2003**, *221*, 425.
- [57] F. Zappacosta, G. Sannia, L. A. SAVOY, G. Marino, P. Pucci, *Eur. J. Biochem.* **1994**, *222*, 761.
- [58] J. G. Zeikus, C. Vieille, A. Savchenko, *Extremophiles* **1998**, *2*, 179.
- [59] K. Sjölander, *Bioinformatics* **2004**, *20*, 170.
- [60] M. Rahban, S. Zolghadri, N. Salehi, F. Ahmad, T. Haertlé, N. Rezaei-Ghaleh, L. Sawyer, A. A. Saboury, *Int. J. Biol. Macromol.* **2022**, *214*, 642.
- [61] J. B. Joy, R. H. Liang, R. M. McCloskey, T. Nguyen, A. F. Poon, *PLoS Comput. Biol.* **2016**, *12*, e1004763.
- [62] N. Doukyu, H. Ogino, *Biochem. Eng. J.* **2010**, *48*, 270.
- [63] A. Panuszko, P. Brzudziak, M. Śmiechowski, M. Stasiulewicz, J. Stefaniak, J. Stangret, *J. Mol. Liq.* **2019**, *294*, 111661.
- [64] A. M. Klabinov, *Nature* **2001**, *409*, 241.
- [65] L. Fu, B. Niu, Z. Zhu, S. Wu, W. Li, *Bioinformatics* **2012**, *28*, 3150.
- [66] J. D. Hunter, *Comp. Sci. Eng.* **2007**, *9*, 90.
- [67] C. R. Harris, K. J. Millman, S. J. Van Der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, *Nature* **2020**, *585*, 357.
- [68] M. Blum, A. Andreeva, L. C. Florentino, S. R. Chuguransky, T. Grego, E. Hobbs, B. L. Pinto, A. Orr, T. Paysan-Lafosse, I. Ponamareva, *Nucleic Acids Res.* **2025**, *53*, D444.
- [69] K. Katoh, J. Rozewicki, K. D. Yamada, *Briefings Bioinf.* **2019**, *20*, 1160.
- [70] H. Ashkenazy, O. Penn, A. Doron-Faigenboim, O. Cohen, G. Cannarozzi, O. Zomer, T. Pupko, *Nucleic Acids Res.* **2012**, *40*, W580.
- [71] I. Letunic, P. Bork, *Nucleic Acids Res.* **2021**, *49*, W293.
- [72] H. Land, M. S. Humble, *Protein Eng.: Methods Prot.* **2018**, *43*.
- [73] W. L. DeLano, *CCP4 Newslet. Protein Crystall.* **2002**, *40*, 82.
- [74] S. Schätzle, M. Höhne, E. Redestad, K. Robins, U. T. Bornscheuer, *Anal. Chem.* **2009**, *81*, 8244.
- [75] J. Wang, W. Wang, P. A. Kollman, D. A. Case, *J. Mol. Graphics modell.* **2006**, *25*, 247.
- [76] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157.
- [77] A. W. Sousa da Silva, W. F. Vranken, *BMC Res. Notes* **2012**, *5*, 1.
- [78] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33.

Manuscript received: February 21, 2025

Revised manuscript received: April 18, 2025

Version of record online: May 30, 2025