

## Rational engineering of *Luminiphilus sylvensis* (R)-selective amine transaminase for the acceptance of bulky substrates

Eleni Konia,<sup>a</sup> Konstantinos Chatzicharalampous,<sup>a</sup> Athina Drakonaki,<sup>a</sup> Cornelia Muenke,<sup>b</sup>  
Ulrich Ermler,<sup>b</sup> Georgios Tsiotis<sup>a</sup> and Ioannis V. Pavlidis<sup>a,\*</sup>

<sup>a</sup> Department of Chemistry, University of Crete, Voutes University Campus, 70013, Heraklion, Greece.

<sup>b</sup> Max Planck Institute of Biophysics, Max-von-Laue-Straße 3, 60438 Frankfurt am Main, Germany.

\*Corresponding author: [ipavlidis@uoc.gr](mailto:ipavlidis@uoc.gr) (I.V. Pavlidis)

### 1 Materials and Methods

#### 1.1 Protein expression and purification for structure elucidation experiments

The gene for the branched-chain amino acid aminotransferase of *Luminiphilus sylvensis* NOR5-1B (NCBI Sequence ID: EED34093.1) was expressed in *Escherichia coli* BL21 (DE3) cells containing the expression vector GamPro\_pGASTON, which encodes the sequence of the amine transaminase, including an additional C-terminal His6 tag (SGSHHHHHH).<sup>1</sup> The recombinant protein consists of 302 amino-acid residues with a molecular weight of 33.79 kDa.

An overnight culture was used for the inoculation of the LB medium. The cells were grown at 30°C in 2 L LB medium containing 0.1 mg mL<sup>-1</sup> ampicillin, until an optical density at 600 nm (OD<sub>600</sub>) of 0.4 was reached. The temperature was then reduced to 20°C and the cells were further incubated until they reached an OD<sub>600</sub> of 0.5-0.6. Expression of the protein was induced by the addition of 0.2% (w/w) L-(+)-rhamnose. Cell cultures were further incubated at 20°C for 20 h following induction to express the protein. Cells were harvested by centrifugation at 9,000 x g and 4°C for 10 min (JLA8.1000 rotor, Avanti J-26 XP, Beckman Coulter). Cell pellets were flash frozen in liquid nitrogen and stored at -80°C until use.

Cells were resuspended in HEPES buffer (50 mM), 300 mM NaCl, pH 7.5 containing 0.1 mM pyridoxal-5'-phosphate (PLP) and 15 mM imidazole in a ratio of 5 mL buffer per 1 g cells. After homogenization of the cells, 20 µg mL<sup>-1</sup> DNase I and 1 mM phenylmethanesulfonyl fluoride (PMSF) were added. Cells were lysed by passing through a microfluidizer (Microfluidics Corporation) for four times at 8,000 psi. The cell debris was removed by centrifugation at 10,000 x g and 4°C for 20 min. The pellet was discarded and the supernatant was ultracentrifuged at 200,000 x g, 4°C for 60 min.

The supernatant was collected and filtered through a 0.45 µm syringe filter (RephiLe) and loaded onto a nickel-NTA column (GE Healthcare) with a flow rate of 2 mL min<sup>-1</sup> which has been equilibrated with

50 mM HEPES, 300 mM NaCl, pH 7.5 containing 0.1 mM PLP and 15 mM imidazole. After washing with three column volumes of the same buffer but increasing the imidazole concentration to 30 mM at a flow rate of 2 mL min<sup>-1</sup>, the protein was eluted with the same buffer including 300 mM imidazole. The eluted protein was then desalted by ultrafiltration using 30 kDa cut-off Amicon centrifugal filter. The desalted protein was in HEPES buffer (50 mM), pH 7.5 containing 0.1 mM PLP. The final step of purification was achieved using a Superdex 200 size exclusion chromatography column (GE Healthcare) coupled to an ÄKTA Purifier system (GE Healthcare). The column was equilibrated with HEPES (50 mM), pH 7.5 containing 0.1 mM PLP.

The purified recombinant proteins (20 µg) were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) using 4-12% Novex Tris-glycine gels (Invitrogen) and MOPS running buffer. Proteins were visualized by Coomassie blue staining. Blue native (BN)-PAGE was performed using the Novex 4-16% Bis-Tris gels (Invitrogen) according to manufacturer's instructions.

## 1.2 Crystallization

Initial crystallization hits were obtained with a variety of PEG based conditions (JBScreen Classic HTSI; PACT HTS Jena Bioscience). Suitable crystals of recombinant protein appeared after 4 days in solutions containing concentrated protein (19.7 mg mL<sup>-1</sup>) and 20% PEG 3350, 0.2 M NaNO<sub>3</sub> at 4°C. The length the crystals was >150 µm and had a very bright yellow color, suggesting the existence of bound PLP in the crystallized proteins.

## 1.3 Data collection and X-ray crystallographic analysis

For cryoprotection, a solution consisting of 10% glycerol was used. X-ray diffraction data were collected at 100 K and a wavelength of 1 Å on beamline PXII of the Swiss-Light-Source, Villigen, Switzerland. Their processing was performed with XDS/XSCALE (**Table S3**).<sup>2</sup> Phases were determined with the molecular replacement method using PHENIX.<sup>3</sup> As search model served the structure of the branched-chain aminotransferase from the thermophilic archaea *Geoglobus acetivorans* (pdb-code: 5cm0) which is the closest known structure (sequence identity: 41%, E-value 3.912e-61).<sup>4</sup> After successful molecular replacement (LLG 1236, TFZ = 35.3) and after a few refinement rounds a new model with the correct sequence was created with the Swiss-Model server.<sup>5</sup> Refinement was continued with PHENIX thereby applying the NCS and TLS options. The obtained model was manually corrected as well as PLP and solvent molecules incorporated within COOT.<sup>6</sup> The quality of the model was evaluated by COOT and MOLPROBITY.<sup>7</sup> The atomic coordinates and structure factors of the (R)-selective amine transaminase from *L. syltensis* have been deposited in the Protein Data Bank, www.pdb.org with ID code 7p3t.

#### 1.4 Bioinformatic analysis

All bioinformatic analysis was performed in YASARA Structure suite v. 20.10.4 (<http://www.yasara.org>) and all Figures with protein structures were prepared with PyMol v.0.99 (<http://www.pymol.org>). The crystal structure was initially refined in a water cell with a 5 Å radius around the atoms of the protein, at pH 7.5 and 25°C, with the YAMBER3 force field. The refinement simulated 500 ps, taking snapshots every 25 ps. The structure with the lowest energy was selected for further studies, after inspection of the catalytic center and that the PLP and the catalytic lysine are in proper orientation. The quality of the refined structure was evaluated with the MolProbity platform.<sup>7</sup> The quinonoid intermediates were constructed manually by building the molecule on the PLP, and the resulting structure was gradually energy minimized; initially, the atoms of the substrate were energy minimized while all rest were fixed and then, the quinonoid molecule was remaining free for energy minimization, apart from the nitrogen of pyridine ring, which was acting as anchor. In the third round, energy minimization was performed with all the atoms of both the protein and the quinonoid being free.

#### 1.5 Mutagenesis polymerase chain reaction

The primers of the mutagenesis polymerase chain reaction (PCR) were designed to introduce the desired mutations on the sequence of the template gene (**Table S4**), and they were subsequently interrogated via OligoCalc for self-complementarity.<sup>8</sup> PCR reactions were performed using 0.25 µL of a high-fidelity DNA polymerase kindly provided by Minotech biotechnology (IMBB-FORTH), with its 5x HF polymerase buffer (containing 7.5 mM MgCl<sub>2</sub>), 0.2 mM dNTPs, 0.5 µM of each primer, 0.2 ng/µL parental plasmid in 25 µL final volume. The amplification was performed as follows: a) 95°C, 30 s; b) 25 cycles: 98°C, 10 s; 65°C, 30 s; 72°C, 1 min 30 s; 3) 72°C, 3 min 30 s. The PCR product was detected by 0.8% (w/w) agarose gel electrophoresis and then the parental plasmid was digested by DpnI for 1 h at 37°C. The restriction enzyme was deactivated at 80°C for 10 min before transforming the produced plasmid into chemo-competent *E. coli* Top10 cells. After obtaining colonies on an LB-agar plate supplemented with 100 mg mL<sup>-1</sup> ampicillin, the plasmid was isolated using Macherey-Nagel™ NucleoSpin Plasmid Kit. The correct constructs were confirmed by DNA sequencing (Genewiz, DE) and then the plasmids were transformed into chemo-competent *E. coli* BL21 (DE3) cells for expression.

#### 1.6 Protein expression and purification for biochemical characterization

The cultivations for the protein expression and purification for biochemical characterization were based on the protocol of Paragraph 2.1, with some minor alterations. After expression, cells from

500 mL cultivation were harvested by centrifugation (10,000 rpm, 10 min, 4°C) and resuspended in 10 mL phosphate buffer (50 mM, pH 7.5) containing 0.1 mM PLP, 0.3 M NaCl, 15 mM imidazole (lysis buffer). The cell suspension was sonicated on ice (4 x 30 s, 30% amplitude, 0.5 cycle) and cell debris was removed by centrifugation (10,000 rpm, 10 min, 4°C).

The recombinant protein was purified by metal affinity chromatography. Using the ÄKTA start FPLC system, the Ni-NTA 5 mL Protino® column was equilibrated with the lysis buffer. The crude lysate was filtered via 0.2 µm filters (Millipore) and it was subsequently loaded onto the column with 1.5 mL min<sup>-1</sup> flow. Non-specifically bound proteins were removed by application of four column volumes of the same buffer with 30 mM imidazole. The enzyme was eluted from the column by the application of three column volumes of the same buffer with 300 mM imidazole. The purified enzyme fractions were pooled and concentrated using Amicon® centrifugal filters (30 kDa molecular weight cut-off), and it was subsequently desalted using a HiTrap™ 15 mL column equilibrated with phosphate buffer (50 mM, pH 7.5) containing 0.1 mM PLP. The purified enzyme solutions were stored at -20°C in 25% glycerol until use.

### 1.7 Photometric determination of activity

Specific activities were determined using a direct, continuous photometric assay based on the conversion of 1-phenylethylamine (PEA) to acetophenone, which can be directly detected spectrophotometrically at 245 nm in a 96-well-plate.<sup>9</sup> Activities were measured in 200 µL total volume in HEPES buffer (50 mM, pH 7.5), 2 mM sodium pyruvate, 1 mM amine and a final concentration of 0.05 mM pyridoxal-5'-phosphate (PLP). Dimethyl sulfoxide (DMSO), 5 % (v/v) final concentration was used to aid the solubilization of the substrates. The spectroscopic change at 245 nm was recorded for 5 min at 37°C. The early data points that define a linear rate of product formation were used to estimate the initial rate of the reaction (mM ketone/min). The concentration of the enzyme used was calculated via the Bradford assay. The extinction coefficients used for the calculations were estimated by standard curves and were found 14.7 M<sup>-1</sup> cm<sup>-1</sup> for acetophenone and 10.18 M<sup>-1</sup> cm<sup>-1</sup> for propiophenone and butyrophenone. For the different pH measurements, the measurements in pH 9.0 and 10.0 were performed in CHES 50 mM buffer with 0.05 mM PLP. All measurements were performed in a Thermo Scientific™ Multiskan Sky Microplate Spectrophotometer in triplicates.

### 1.8 Pyridoxamine phosphate formation analysis

The inactive variants were investigated for their activity in the first half-reaction, the formation of the pyridoxamine phosphate (PMP), based on the assay of Voss et al.<sup>10</sup> Shortly, in a UV-Star 96-well plate, in 200 µL of reaction volume, 50 µL of purified LS\_ATA variant were added in 147.5 µL CHES

buffer (50 mM, pH 9.0) and 2.5  $\mu$ L of ethanol, with and without (*R*)-PEA (final concentration: 2.5 mM). The spectrum was monitored from 290 nm to 470 nm in a Thermo Scientific™ Multiskan Sky Microplate Spectrophotometer. The spectra were normalized per mg of protein per mL or reaction.

### 1.9 Kinetic resolution experiments and gas chromatography analysis

In 2 mL final volume, 8 mM of (*R,S*)-PEA or (*R,S*)-PPA or 4 mM (*R*)-PBA were added in HEPES buffer (50 mM, pH 8.0, 0.1 mM PLP) with 5% (v/v) DMSO as co-solvent, while pyruvate (16 mM for the racemates, 8 mM for the (*R*)-PBA) was used as amino acceptor. The reaction was initiated by the addition of 60  $\mu$ g wild-type LS\_ATA or LS\_ATA/V37A and was run at 35°C, 750 rpm for 24 h. In specific time intervals, an aliquot of 200  $\mu$ L was removed and 15  $\mu$ L of 5 M NaOH was added with a subsequent vortexing step, to terminate the reaction and deprotonate the amines. Then the reactants were extracted with ethyl acetate (2x 200  $\mu$ L) with vortexing, the organic phase was dried anhydrous Na<sub>2</sub>SO<sub>4</sub> and then analyzed by gas chromatography in a Shimadzu Nexis GC-2030 system equipped with the chiral column Cyclodex-B (30 m x 0.25 mm x 0.25  $\mu$ m) from Agilent and He as carrier gas with 1.4 mL min<sup>-1</sup> column flow. Injector and detector temperatures were 250°C. The following program was used: 80°C initial temperature, heated to 120°C with 2°C min<sup>-1</sup> and then to 220°C with 20°C min<sup>-1</sup>, where it is held for 3 min. Retention times: 13.8 min for (*R*)-PEA, 14.0 min for (*S*)-PEA, 15.4 min for acetophenone, 18.7 min for (*R*)-PPA, 18.9 min for (*S*)-PPA, 20.8 min for propiophenone, 23.5 min for (*R*)-PBA and 24.6 min for butyrophenone.

## 2 Tables and Figures

**Table S1:** Specific activity (U/mg) and apparent enantioselectivity ( $E_{app}$ ) of LS\_ATA against PEA (1 mM) and pyruvate (2 mM) at different pH. Experimental conditions described in paragraph 1.7.

pH	(R)-PEA	(S)-PEA	$E_{app}$
8	$0.27 \pm 0.05$	n.d.	n.d.
9	$0.47 \pm 0.03$	$0.007 \pm 0.004$	67
10	$0.56 \pm 0.06$	$0.008 \pm 0.001$	70

*n.d.* Not detectable

**Table S2:** Activity of LS\_ATA against (R)-amine (1 mM) and pyruvate (2 mM) at HEPES (50 mM, pH 8.0, 0.05 mM PLP). No activity was detected against (R)-PBA with any of the variants with the photometric assay. Experimental conditions described in paragraph 1.7.

LS_ATA Variant	(R)-PEA [mU mg <sup>-1</sup> ]	(R)-PPA [mU mg <sup>-1</sup> ]
Wild type	$274 \pm 54$	n.d.
V37A	$93 \pm 16$	$45 \pm 26$
S248A	n.d.	n.d.
T249A	$138 \pm 36$	n.d.
V37A/S248A	n.d.	n.d.
V37A/T249A	n.d.	$1.8 \pm 0.7$

*n.d.* Not detected

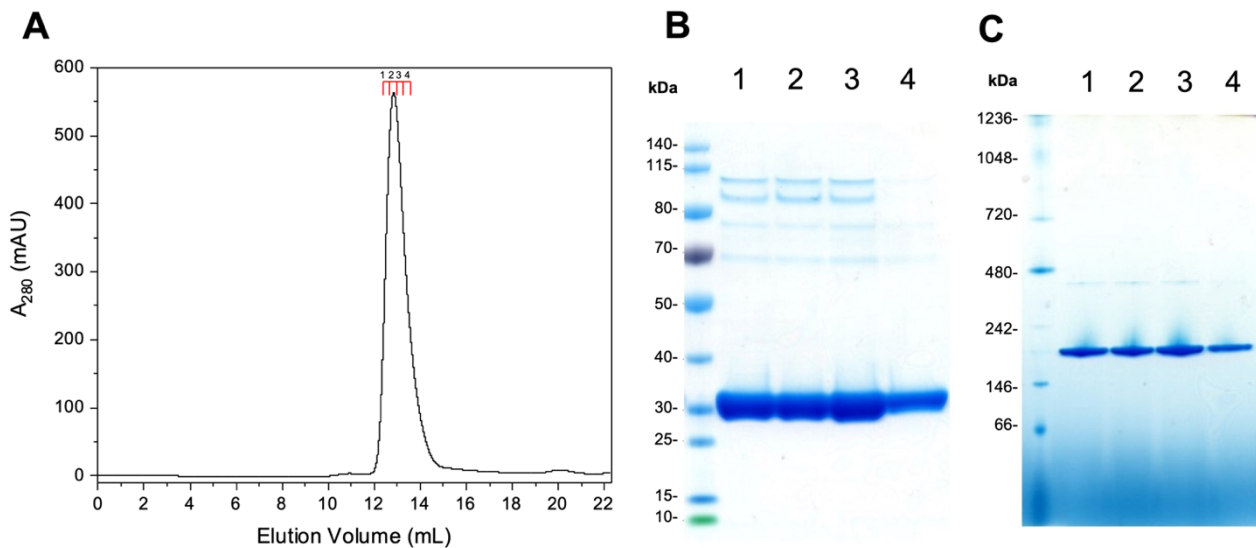
**Table S3.** Data-collection and processing statistics. Values in parentheses are for the outermost resolution shell.

<b>Crystal</b>	<b>LS-ATA</b>
<b>Crystallization</b>	
Protein solution	50 mM HEPES pH 7.5, 0.1 mM PLP (19.7 mg/mL protein)
Crystallization conditions	20 % (w/v) PEG 3350, 0.2 M NaNO <sub>3</sub>
Cryo conditions	+ 10% (v/v) glycerol
<b>Data collection</b>	
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Wavelength [Å]	1.0
Resolution [Å]	50.0 - 1.6 (1.7 - 1.6)
Unit cell a, b, c [Å]	86.8, 144.6, 151.8
Redundancy	9.1 (5.5)
Completeness [%]	99.7 (98.1)
R <sub>sym</sub> [%]	3.6 (87.9)
I/σ(I)	32.7 (2.0)
CC <sub>1/2</sub>	100 (71.9)
<b>Refinement statistics</b>	
<b>Unique reflections</b>	249902
Subunits per asym. unit	6
No. atoms: polypeptide, cofactors, solvent	14360, 126, 1274
R <sub>working</sub> , R <sub>free</sub> (%)	16.9, 18.5
B <sub>average</sub> (Å <sup>2</sup> ): polypeptide, ligands, solvent	32.7, 31.0, 35.8
<b>Clashscore</b>	5.1
R.m.s. deviation	
bond lengths (Å)	0.009
bond angles (°)	1.12
Ramachandran Plot: favored, outliers (%)	98.2, 0.06

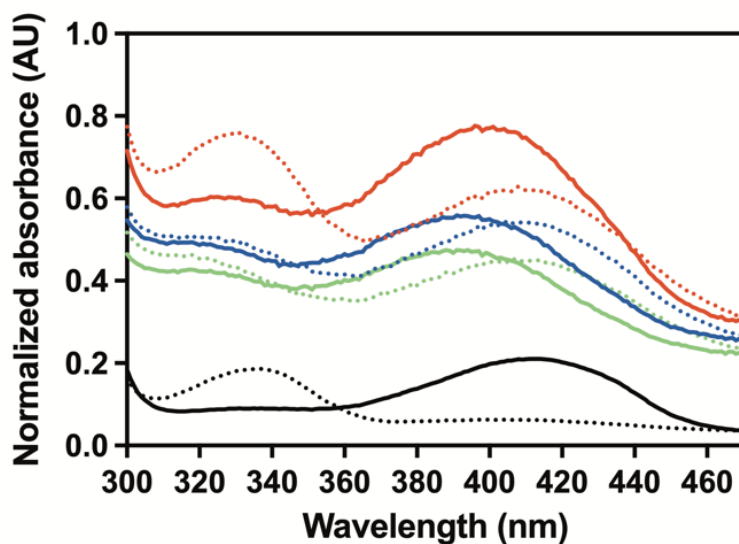
**Table S4.** Primers used for the mutagenesis PCR.

<b>V37A_fw</b>	GTT TTT GAT GCT GTT AGC GCC TGG
<b>V37A_rv</b>	GGC GCT AAC AGC ATC AAA AAC ACC
<b>S248A/T249A_fw</b>	GTT TAC CTG TGC CGC CGC AGG C
<b>S248A/T249A_rv</b>	GCG GCG GCA CAG GTA AAC ACT TC
<b>S248A_fw</b>	GTT TAC CTG TGC CAC CGC AGG C
<b>S248A_rv</b>	CGG TGG CAC AGG TAA ACA CTT CAT C
<b>T249A_fw</b>	GTT CTG CCG CAG GCG GTG
<b>T249A_rv</b>	GCG GCA GAA CAG GTA AAC ACT TC





**Fig. S1:** Size exclusion chromatography revealed a main peak at about 13 min (**A**), which is the Ni-NTA-purified LS\_ATA, as shown from the main band around 33 kDa in SDS-PAGE of fractions 1-4 (**B**). The blue native polyacrylamide gel electrophoresis showed that soluble LS\_ATA also exists in hexameric form (**C**, expected MW: 202 kDa).



**Fig. S2:** Absorption spectra of the LS\_ATA variants in resting state (internal aldimine, line) and after addition of (R)-PEA (dotted line). Wild type, black; S248A, red; V37A/S248A, blue, V37A/T249A green. Experimental details in ESI.

### **3 Further Notes**

- The crystal structure was deposited in PDB under the code: 7p3t.
- All primary and secondary data are available at: <https://fairdomhub.org/publications/629>.

## 4 References

- 1 M. Höhne, S. Schätzle, H. Jochens, K. Robins and U.T. Bornscheuer, *Nat. Chem. Biol.*, 2010, **6**, 807–813.
- 2 W. Kabsch, *Acta Crystallogr. D Biol. Cryst.*, 2010, **66**, 125-132.
- 3 P.V. Afonine, R.W. Grosse-Kunstleve, V.B. Chen, J.J. Headd, N.W. Moriarty, J.S. Richardson, D.C. Richardson, A. Urzhumtsev, P.H. Zwart and P. D. Adams, *J. Appl. Crystallogr.*, 2010, **43**, 669-676.
- 4 M.N. Isupov, K.M. Boyko, J.M. Sutter, P. James, C. Sayer, M. Schmidt, P. Schonheit, A.Y. Nikolaeva, T.N. Stekhanova, A.V. Mardanov, N.V. Ravin, E.Y. Bezsudnova, V.O. Popov, and J.A. Littlechild, *Front. Bioeng. Biotechnol.*, 2019, **7**, Art. No. 7.
- 5 T. Schwede, J. Kopp, N. Guex and M.C. Peitsch, *Nucleic Acids Res.*, 2003, **31** (13), 3381-3385.
- 6 P. Emsley and K. Cowtan, *Acta Crystallogr. D Biol. Cryst.*, 2004, **60**, 2126-2132.
- 7 I.W. Davis, A. Leaver-Fay, V.B. Chen, J.N. Block, G.J. Kapral, X. Wang, L.W. Murray, W.B. Arendall 3rd, J. Snoeyink, J.S. Richardson, and D.C. Richardson, *Nucleic Acids Res.*, 2007, **35**, W375-W383.
- 8 W.A. Kibbe, *Nucleic Acids Res.*, 2007, **35**, W43-W46.
- 9 S. Schätzle, M. Höhne, E. Redestad, K. Robins and U.T. Bornscheuer, *Anal. Chem.*, 2009, **81**(19), 8244-8248.
- 10 M.Voss, C. Xiang, J. Esque, A. Nobili, M. J. Menke, I. André, M. Höhne, U.T. Bornscheuer, *ACS Chem. Biol.*, 2020, **15**, 416-424.