The FEBS Journal

FEBS PRESS science publishing by scientists

The crystal structure of the tetrameric DABA-aminotransferase EctB, a rate-limiting enzyme in the ectoine biosynthesis pathway

Heidi Therese Hillier, Bjørn Altermark and Ingar Leiros

The Norwegian Structural Biology Centre (NorStruct), Department of Chemistry, Faculty of Science and Technology, UiT the Arctic University of Norway, Tromsø, Norway

Keywords
Chromohalobacter salexigens; crystal structure; DABA aminotransferase; EctB; osmoadaptation

Correspondence
I. Leiros, The Norwegian Structural Biology Centre (NorStruct), Department of Chemistry, Faculty of Science and Technology, UiT the Arctic University of Norway, N-9037, Tromsø, Norway
Tel: +47 95966895
E-mail: ingar.leiros@uit.no

(Received 14 November 2019, revised 30 January 2020, accepted 26 February 2020)

doi:10.1111/febs.15265

L-2,4-diaminobutyric acid (DABA) aminotransferases can catalyze the formation of amines at the distal ω-position of substrates, and is the intial and rate-limiting enzyme in the biosynthesis pathway of the cytoprotecting molecule (*S*)-2-methyl-1,4,5,6-tetrahydro-4-pyrimidine carboxylic acid (ectoine). Although there is an industrial interest in the biosynthesis of ectoine, the DABA aminotransferases remain poorly characterized. Herein, we present the crystal structure of EctB (2.45 Å), a DABA aminotransferase from *Chromohalobacter salexigens* DSM 3043, a well-studied organism with respect to osmoadaptation by ectoine biosynthesis. We investigate the enzyme's oligomeric state to show that EctB from *C. salexigens* is a tetramer of two functional dimers, and suggest conserved recognition sites for dimerization that also includes the characteristic gating loop that helps shape the active site of the neighboring monomer. Although ω-transaminases are known to have two binding pockets to accommodate for their dual substrate specificity, we herein provide the first description of two binding pockets in the active site that may account for the catalytic character of DABA aminotransferases. Furthermore, our biochemical data reveal that the EctB enzyme from *C. salexigens* is a thermostable, halotolerant enzyme with a broad pH tolerance which may be linked to its tetrameric state. Put together, this study creates a solid foundation for a deeper structural understanding of DABA aminotransferases and opening up for future downstream studies of EctB's catalytic character and its redesign as a better catalyst for ectoine biosynthesis. In summary, we believe that the EctB enzyme from *C. salexigens* can serve as a benchmark enzyme for characterization of DABA aminotransferases.

Database
Structural data are available in PDB database under the accession number 6RL5.

Introduction

The EctB enzyme is an L-2,4-diaminobutyric acid (DABA) aminotransferase (EC 2.6.1.76), found at the cross-roads between the transaminase (TA) world of synthesizing amine bonds [1] and the biosynthesis pathway of the interesting cytoprotectant (*S*)-2-methyl-1,4,5,6-tetrahydro-4-pyrimidinecarboxylic acid (ectoine)

Abbreviations
2-OG, 2-oxoglutaric acid; ADABA, *N*-γ-acetyldiaminobutyric acid; ASA, L-aspartate-β-semialdehyde; DABA, L-2,4-diaminobutyric acid; ectoine, (*S*)-2-methyl-1,4,5,6-tetrahydro-4-pyrimidine carboxylic acid; MSA, multiple sequence alignment; PLP, pyridoxal-5-phosphate; SEC, size exclusion chromatography; TA, transaminase; TCEP, tris(2-carboxyethyl)phosphine; β-ME, β-mercaptoethanol.

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.

[2]. EctB catalyzes the transfer of an amine group from a donor to an acceptor molecule [3] at the distal ω-position of an acceptor substrate (Fig. 1A), using different acceptor–donor substrate pairs in the forward and reverse direction.

In the context of osmoadaptation, EctB represents the first enzyme in the biosynthesis pathway of the small organic molecule ectoine [2,3]. This molecule is a member of a larger group of molecules known by several names, such as osmolytes, compatible solutes, cytoprotectants, and chemical chaperons [4,5], as cells actively accumulate these compounds to protect themselves against changes in salinity, freezing, thawing, and even radiation. Although ectoine can be accumulated to molar concentrations in the cytoplasm, it does not interfere with normal cellular processes [2,6-12]. Interestingly, ectoine has also been demonstrated to have a stabilizing effect on proteins [13,14] and has therefore gained interest as a highly valuable biotechnological additive. Ectoine and its derivative 5-hydroxyectoine are predominantly produced by *Bacteria* (but also found in some Archaea and Eukarya) [10-12,15,16] as a stress response to increases in environmental salinity, although production can also be triggered by other factors such as temperature and various bacterial growth phases [17,18].

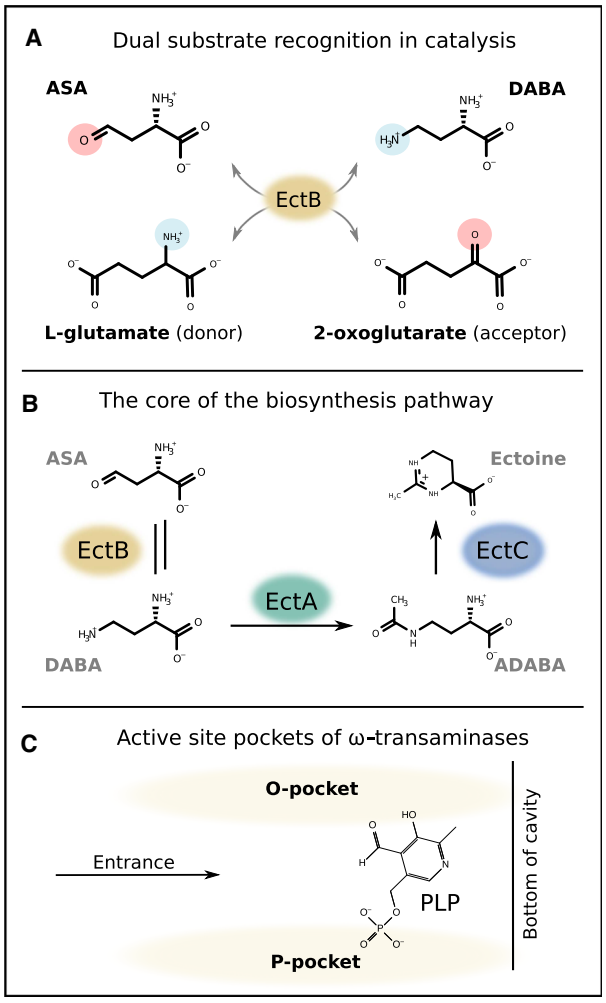

Fig. 1. Overview of EctB's known catalytic character. (A) Two different substrate pairs are recognized in the forward and reverse reactions. (B) The core of the ectoine biosynthesis pathway consists of the aminotransferase EctB, the acetyltransferase EctA, and the ectoine synthase EctC. (C) The active site of ω-TAs contains an O-pocket near the pyridoxal ring and a P-pocket near the phosphate moiety of the coenzyme PLP, to accommodate the different amine donor–acceptor substrate pairs. The illustration was created in INKSCAPE (www.inkscape.org).

The core of the ectoine biosynthesis pathway is a three-step process (Fig. 1B) [19,20]. Firstly, EctB catalyzes the reversible conversion of L-aspartate-β-semialdehyde (ASA) to DABA using the coenzyme pyridoxal-5-phosphate (PLP) and the cosubstrates L-glutamate or 2-oxoglutarate (2-OG) for the forward or reverse reaction, respectively [3]. Secondly, EctA (EC 2.3.1.178) converts DABA to *N*-γ-acetyldiaminobutyric acid (ADABA) using Acetyl-CoA as a cofactor. Lastly, EctC (EC 4.2.1.108) catalyzes the cyclization of ADABA to ectoine [21]. Although these three enzymes are responsible for the core of the pathway, many different bacteria exist that encode additional or alternative proteins that interact with either ectoine or other intermediates of this metabolic pathway. For example, some bacteria also encode an ectoine hydroxylase (EctD), which converts ectoine to the derivative 5-hydroxyectoine [10,22], and some may even encode the ectoine degradation enzymes DoeABCD [23]. We refer the reader to review articles for more details on the pathway and its phylogenetic distribution [12,20,24].

The interest in the biosynthesis of ectoine is increasing as demonstrated by the number of studies focused on whole-cell ectoine biosynthesis [25-29], and the recent studies on the biochemical details of EctC [21] and EctD [30]. However, literature on EctB is scarce, both in the context of ectoine production and as a general DABA aminotransferase, of which there are currently no representative structure models. Although EctB is the first of the three core enzymes of the ectoine biosynthesis pathway, it is reported to be the rate-limiting component of ectoine biosynthesis [31,32], and targeting of EctB by random mutagenesis has been shown to increase whole-cell production of ectoine [31]. It has also been shown that the ratio of EctB to EctA can affect the whole-cell production of ectoine [32].

To further optimize production of osmolytes such as ectoine, there is a need to fully characterize benchmark enzymes from a well-studied organism such as the

moderately halophilic *Chromohalobacter salexigens* DSM 3043. This bacterium displays broad salinity tolerance and is a natural producer of ectoine [25]. *Chromohalobacter salexigens* has also been suggested as an alternative organism for industrial production of ectoine [33], as it can grow at both lower salinity and lower temperatures than the currently used *Halomonas elongata*, which would reduce corrosion and energy use in industrial fermenters. Furthermore, the same study found *C. salexigens* to have a higher ectoine yield at lower culturing temperatures compared to a selection of other ectoine producers, including *H. elongata* [14]. A metabolic model for osmoadaptation in *C. salexigens* has also been published [34]. Put together, *C. salexigens* is emerging as one of the key organisms to study osmoadaptation, which increases our need to understand the basic biochemical and structural aspects of its ectoine-synthesizing enzymes.

The DABA aminotransferase EctB is a member of the class III PLP-dependent aminotransferase, which are expected to be functional dimers in solution [35]. In general, PLP-dependent enzymes are suitable biocatalysts for industry as the PLP coenzyme is continuously recycled in situ, which eliminates the costly addition of coenzymes during catalysis. The general reaction mechanism of substrate activation with PLP as a coenzyme is well described [36-38], where PLP is found in the active site of PLP-dependent enzymes covalently bound through a Schiff-base intermediate to a lysine residue. Class III ω-TAs, such as the herein presented EctB, must have dual substrate recognition mechanisms to differentiate between two substrate pairs [39,40] and be able to specifically catalyze the transfer of the distal amine group. Several studies show that the active site of many aminotransferases has two binding pockets, denoted according to their relative position to the PLP coenzyme (Figs 1C and 2), and that the orientation of substrates in the active site is often coordinated by one or two arginine residues that form strong directional bonds to the substrate's carboxylate group [39,41-46] by a so-called end-on geometry [47]. When the arginine–carboxylate bond is not required, some aminotransferases neutralize the arginine through a strong ionic bond with a glutamate residue (i.e., glutamate switch), while others have a flexible arginine that moves in and out of the active site (i.e., arginine switch; Fig. 2).

The particular arrangement of catalytic residues in the active site that give the DABA aminotransferase family their catalytic specificity has yet not been described. However, a bioinformatic review by Steffen-Munsberg *et al.* [40] has assigned 13 fingerprint residues to PLP-dependent enzymes with different reaction specificities, to aid further bioinformatic annotation of the PLP-dependent protein superfamily. The DABA aminotransferases represent one of these reaction specificities, and 13 residue positions are accordingly described to contain characteristic amino acids for this enzyme family. These residues should be generic for the DABA aminotransferases and most likely include residues that shape the two binding pockets (i.e., the O-pocket and P-pocket; Fig. 2) of the DABA aminotransferases. However, as a structural model for this enzyme family has up until now been lacking, there is uncertainty associated with the assignment of these functional residues. This study therefore gives needed structural insight to DABA aminotransferases by providing the first description of the binding pockets, opening up the possibility for future downstream mutational studies, optimization of whole-cell ectoine production, and will also improve bioinformatic annotation of this enzyme family.

Results

Purification and biochemical properties

Initial EctB purification trials were conducted with Tris/HCl buffer (pH 7.5); however, the purified EctB (10 $mg \cdot mL^{-1}$) did not absorb light at 420 nm, which results when PLP is covalently bound to the enzyme by a Schiff-base linkage. Furthermore, we could not observe the characteristic yellow hue, which is commonly described for PLP-bound enzymes. To ensure full saturation of PLP in EctB, we evaluated whether PLP could bind Tris/HCl on its own as it may react with primary amines [48]. Our findings show that the absorption maximum of PLP in water remains at 390 nm, indicating that PLP is found in its free aldehyde form, which is also observed when PLP is dissolved in 50 mM HEPES (pH 7.5). However, the absorption maximum shifts to 420 nm when PLP is dissolved in 50 mM Tris/HCl, which verifies that Tris/HCl may act as a PLP scavenger when used as a buffer with PLP-dependent enzymes (Fig. 3A). Furthermore, we chose to store the final enzyme with tris(2-carboxyethyl)phosphine (TCEP) as a reducing agent, since DTT and β-mercaptoethanol (β-ME) have been reported to interact with PLP [48]. Finally, when purifying EctB from *C. salexigens* we additionally supplement with 300 μM PLP [prior to lysis and prior to size exclusion chromatography (SEC)] which routinely gives a purified EctB with a yellow hue and an absorption maximum of 420 nm in 50 mM HEPES at pH 7.5.

The oligomeric state was verified by SEC and Blue-Native-PAGE (Fig. 3D). The calibrated analytical SEC

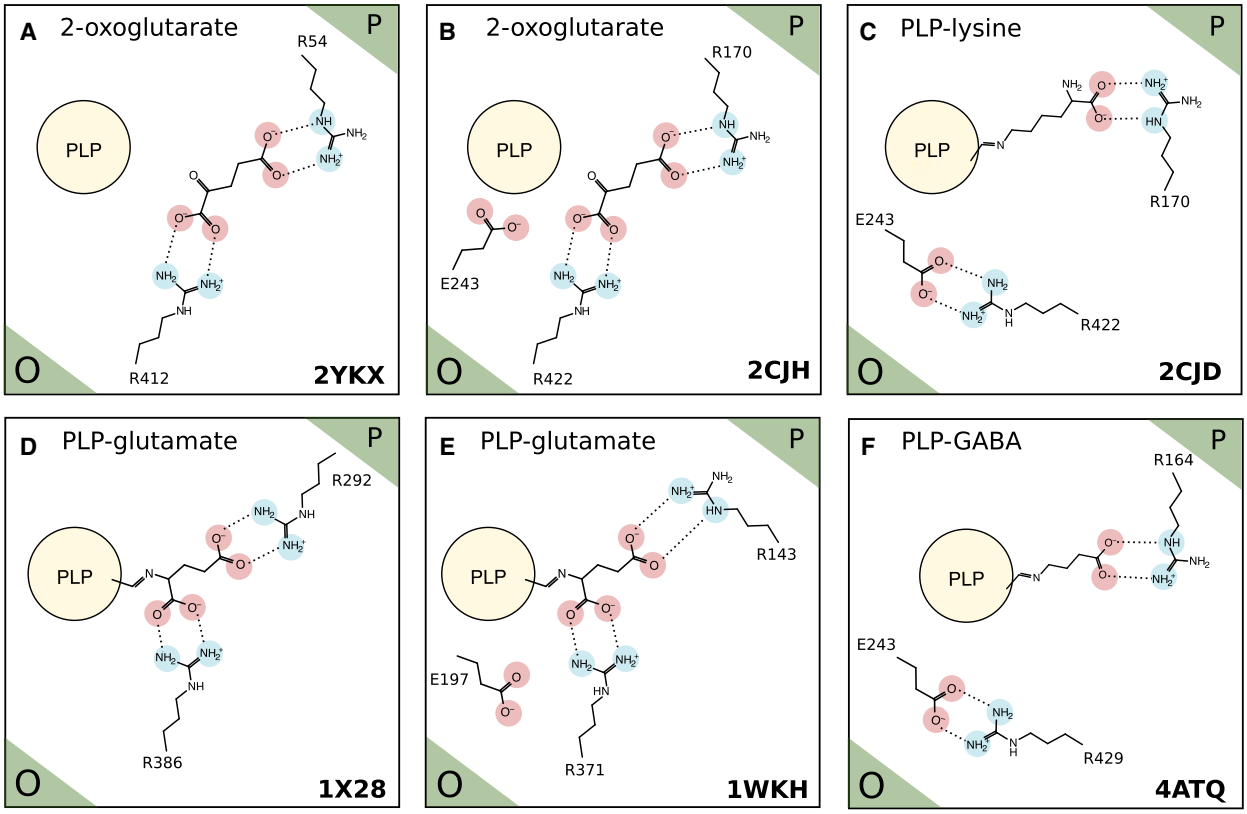

Fig. 2. The O-pocket (bottom left corner) and the P-pocket (top right corner) of various TAs and their catalytically important Arg (R) and Glu (E) residues that coordinate substrates in the active site. (A) Beta-selective aminotransferase from *Mesorhizobium* sp. LUK coordinating 2-OG (PDB ID: 2YKX). (B) Lysine aminotransaminase from *Mycobacterium tuberculosis* strain ATCC 25618 coordinating 2-OG (PDB ID: 2CJH). (C) Lysine aminotransferase from panel B now coordinating a lysine bound to PLP in the active site (PDB ID: 2CJD). (D) Aspartate aminotransferase from *Escherichia coli* strain K12 coordinating glutamate bound to PLP in the active site (PDB ID: 1X28) (E) Acetylornithine aminotransferase from *Thermus thermophilus* DSM 579 coordinating glutamate bound to PLP in the active site (PDB ID: 1WKH). (F) GABA aminotransferase from *Paenarthrobacter aurescens* strain TC1 coordinating γ-aminobutyrate (GABA) bound to PLP in the active site (PDB ID: 4ATQ). The sketches were made in INKSCAPE (www.inkscape.org) based on the respective deposited crustal structures.

Superdex 200 10/300 GL column gave an elution volume that corresponds to the apparent molecular weight of 172.3 kDa (Fig. 3E,F), demonstrating that EctB with the monomeric weight of 47 kDa should be a tetramer in solution, as the theoretical weight of the tetramer is 188 kDa. The BlueNative-PAGE also gave a strong single protein band between the 146- and 242-kDa protein marker.

To verify that EctB was catalytically active and to determine its optimal reaction conditions, we detected the reverse transamination of DABA to ASA with 2-OG (Fig. 4), as ASA is not commercially available. EctB was found to be most active at pH 8.0, although it was generally active in the broad range of pH 7.0 to pH 10.0. We found that EctB has an increased activity with increasing salt concentrations that plateaus at 0.3 M NaCl. Finally, we evaluated the optimal temperature for activity, which was found to increase up to 60 °C. No remaining activity was detected above 70 °C, which fits well with the melting temperature (T_m) found to be 71 °C by differential scanning calorimetry (DSC).

Assessment of the EctB structure

The template for molecular replacement was found using blastp [49] against the PDB database [50], where PDB ID: 1SF2 was chosen as the best template based on a preferential tradeoff between sequence identity (31%) and query coverage (97%; Table 1). The final *C. salexigens* DSM 3043 EctB model, refined to a resolution of 2.45 Å, contained a total of eight monomers and eight PLP coenzymes. The final statistics are summarized in Table 2. The final model had 93.7% of the residues within the most favorable region of the Ramachandran plot and only 0.6% of the residues in the least favorable region. Electron density for PLP

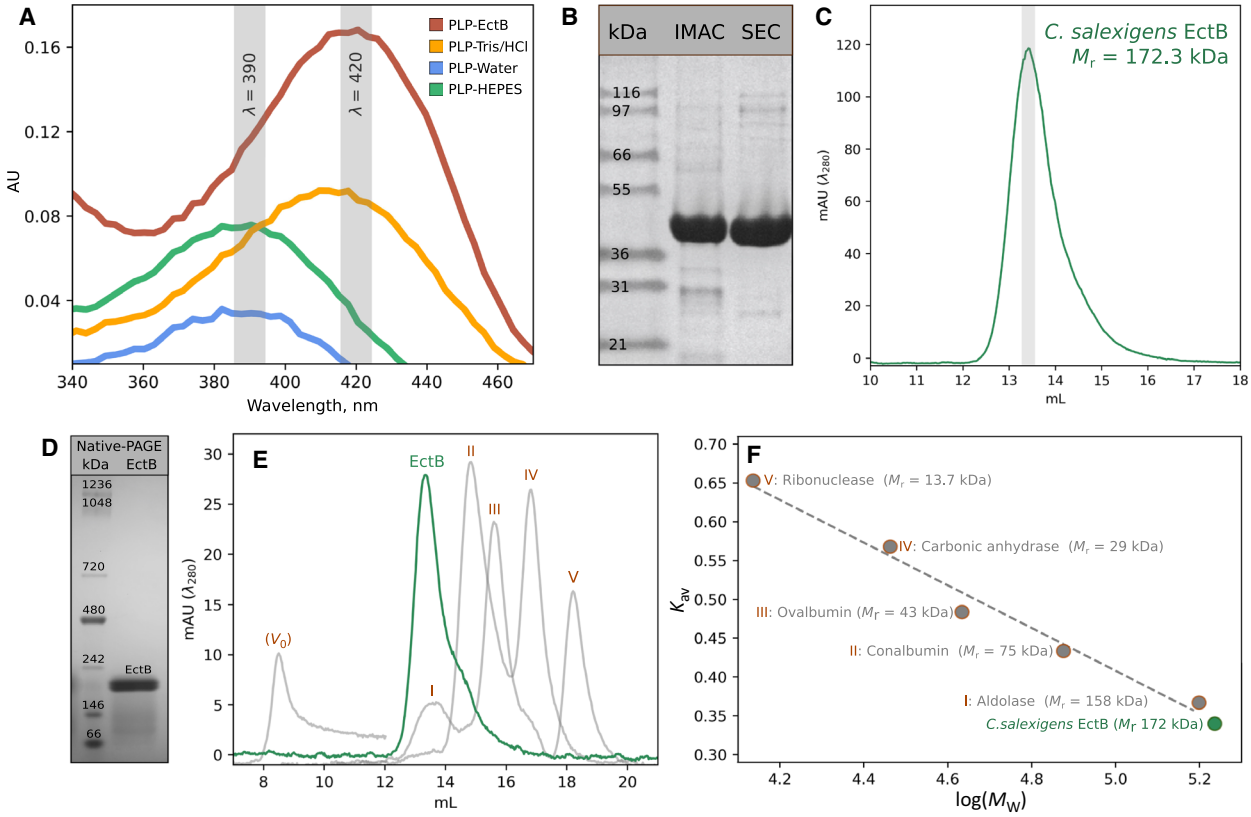

Fig. 3. Purification, optimization, and verification of the oligomeric state of EctB from *Chromohalobacter salexigens* DSM 3043. (A) UV-VIS spectroscopy of PLP under different conditions. PLP at final concentration of 200 μM dissolved in either HEPES buffer (green), water (blue), or Tris/HCl buffer (yellow). A purified sample of EctB in HEPES buffer is shown for comparison (red). (B) SDS/PAGE of EctB after purification by immobilized metal affinity chromatography (IMAC) and SEC, respectively. (C) Elution peak for EctB with the analytical SEC Superdex 200 10/300 GL column (D) BlueNative-PAGE of EctB with one single band between 146- and 242-kDa protein marker. (E) Calibration peaks (I–V) and void volume (V_0) for the SEC Superdex 200 10/300 GL (gray) column and elution peak for EctB (green). (F) The void volume (V_0) corrected calibration curves with K_{av} plotted against log(M_W) for the calibration proteins (I–V) shown in panel E and estimated log(M_W) of EctB. Figure panels A, C, E, and F were generated with matplotlib (www.matplotlib.org) and annotated in INKSCAPE (www.inkscape.org).

was observed and modeled for all eight chains. A covalent link between PLP and K267 is observed in six of the eight chains, but it is less apparent in the remaining two chains E and H. All chains are equally arranged around PLP, although some chains appear to be packed tighter than others, which causes some variation in the observed water coordination of PLP's phosphate moiety. The phosphate oxygens of chains B, D, E, and G are coordinated by two waters, while chains A, F, and H are coordinated by only one water molecule, and chain C has no observed density that fits water around the phosphate moiety. The electron density maps for all eight chains are intact except for a few N- and C-terminal residues that were excluded from the model. From the crystallized EctB construct of 429 amino acids, the final model includes the following chains with respective residues: Chain-A I5-G421, Chain-B T3-G421, Chain-C T3-G421, Chain-D T3-F420, Chain-E Q4-G421, Chain-F I5-G421, Chain-G Q2-A423, and Chain-H Q4-F420.

The EctB tetramer

The asymmetric unit of the final model contains eight polypeptide chains that form two tetramers, packed at a 27° angle. Accordingly, we found EctB to be a functional tetramer in solution which has also been reported for other class III TAs [53]. The EctB tetramer is a dimer of dimers, that is, two dimers come together to form the tetramer. To describe the interfaces within the quaternary structure, we will herein refer to the dimeric interface as the surface between two monomers and the tetrameric interface as the surface between two dimers (Fig. 5).

Analysis with PISA [54] reveals an extensive contact surface in the dimeric interface (e.g., 4713.7 Å^2, ~ 25%

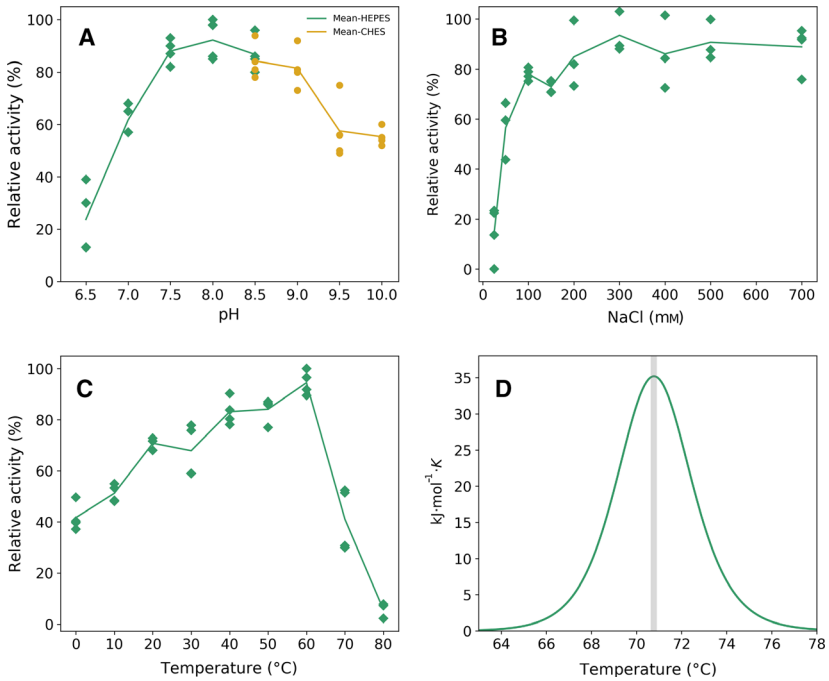

Fig. 4. Biochemical and biophysical characterization of EctB from *Chromohalobacter salexigens* DSM 3043. The effect of pH (A), NaCl (B), and temperature (C) on the relative activity of EctB was studied (with triplicate measurements; $n = 3$). In (A), HEPES was used for the pH range of 6.5–8.5 (green) and CHES was used for the pH range of 8.5–10.0 (yellow). (D) Thermal unfolding of EctB monitored by DSC. The figures were generated with matplotlib (www.matplotlib.org).

Table 1. *Chromohalobacter salexigens* DSM 3043 EctB protein sequence identity and query coverage from blastp against the PDB database, used to select PDB ID: 1SF2 as a template for molecular replacement.

Structural homologs[a]	PDB ID	Sequence ID (%)	Query cover (%)	RMSD (Å)[b]	Reference
7,8-diaminopelargonic acid aminotransferase	5UC7	32	59	2.06	Not published
GABA aminotransferase	1SF2	31	97	1.57	[51]
GABA aminotransferase	4ATP	29	98	1.64	[52]
Aspartate aminotransferase	2EO5	28	99	2.24	Not published

[a]EctB from *C. salexigens* DSM 3043 blastp results against the Protein Data Bank.; [b]Between the presented published structure and the herein presented EctB from *C. salexigens* DSM 3043.

of chain A). Using ConSurf [55] to align around 400 homologous protein sequences, we found the dimeric interface and areas surrounding the PLP-binding site to be highly conserved (Fig. 6A). We also observed a number of conserved residues facing the tetrameric interface (e.g., N136, N151, R155, Y184, and D191) that interact with the neighboring dimer to provide structural integrity (Fig. 6B). Interestingly, strong intermolecular bonds between charged side chains are abundant in EctB and characteristic for the overall quaternary structure. Furthermore, the abundance of strong intermolecular bonds in addition to a high aliphatic index [56] of 81.87 can rationalize EctB's high thermostability.

Fold and topology

The topology of the modeled chains is consistent with the TA fold type-I [35], commonly found as functional dimers. All eight chains in the two EctB tetramers are found to be highly similar (rmsd between 0.27 and 0.52 Å), with a PLP bound in each of the active sites. Each chain can be described by a small and a large folding domain (Fig. 7). The large domain (res. 80–314) forms the central region of each subunit and folds into a seven-stranded antiparallel β-sheet surrounded by nine α-helices. The domain has several important roles including structural integrity by forming bonds to the neighboring dimer in the tetrameric interface, it contains PLP-binding residues that keep the coenzyme retained during catalysis, and it provides the characteristic gating loop [52,57] that completes the active site of the neighboring chain in the dimer. The small domain (res. 1–79 and 315–423) wraps around the large domain and consists of two three-stranded β-sheets located at the N and C termini surrounded by two and three α-helices, respectively. The overall fold of the N-terminal end is less conserved among fold type-I TAs [35]. In EctB, the N terminus interacts

Table 2. Data collection and refinement statistics of EctB (PDB ID: 6RL5). Numbers in parentheses represent the highest resolution bin.

Data collection	EctB—data collection	Refinement	EctB—refinement
Beamline	ID23-1	R_{free} reflections	2703 (248)
Wavelength (Å)	0.99187	No. of non-H atoms	26 163
Temperature (K)	100	Protein	25 733
Space group	$P2_1$	Ligands	120
Unit cell parameters		Solvent	310
a, *b*, *c* (Å)	107.06, 124.78, 145.42	R_{work}/R_{free}	0.1850/0.2441
β (°)	98.95	Mean *B*-factor	63.3
Monomers in the asymmetric unit (#)	8	Protein	63.4
Oscillation range (°)	0.1	Ligands	69.6
Resolution range (Å)	49.07–2.45 (2.49–2.45)	Solvent	57.4
Unique reflections	136 967 (6796)	RMSD from ideal geometry	
Completeness (%)	98.9 (99.6)	Bond lengths (Å)	0.010
Multiplicity	3.4 (3.4)	Bond angles (°)	1.29
Wilson *B*-factor	59.8	Ramachandran statistics (%)	
$CC_{1/2}$	0.999 (0.401)	Favored	93.7
R_{pim} (%)	0.047 (1.070)	Allowed	5.8
$<I/\sigma(I)>$	10.6 (0.7)	Outliers	0.6
		PDB ID	6RL5

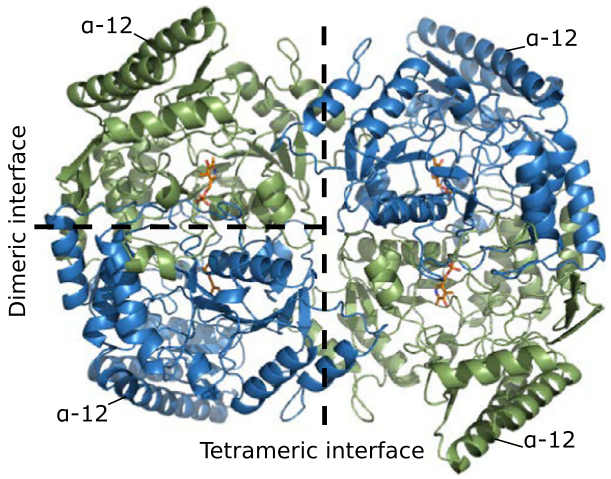

Fig. 5. The EctB tetramer from *Chromohalobacter salexigens* DSM 3043. One EctB tetramer can be described by the tetrameric interface between the two dimers, and the dimeric interface between the two chains in a dimer. For each chain, α-helix-12 is indicated for orientation purposes. PLP was captured in the crystal structure and accordingly positioned (orange) into all chains. The figure was generated in PYMOL (www.pymol.org) and annotated in INKSCAPE (www.inkscape.org).

tightly with the neighboring chain, providing structural integrity to the dimeric interface. It also plays a role in substrate specificity, by shaping the active site in its own chain in addition to contributing with an active site loop (preceding α-helix-3) to the neighboring chain (Fig. 7). The C-terminal regions are spatially located away from both interfaces and contain a high number of surface charges primarily involving residues in α-helix-12 and neighboring α-helix-14.

Recognition sites at the dimeric interface

There are two notable recognition sites at the dimer interface that contribute to EctB's structural integrity (Fig. 8 and Appendix S1). The sites consist of aromatic and hydrophobic residues that are generally found to be conserved among different TAs (Fig. S1), but also some residues that are specifically conserved among the DABA aminotransferases (Fig. S2). Herein, we will use chain A and chain B (*) to describe the two recognition sites. The first recognition site consists of H105*, W288*, P290* and G291* and V13, where G291* is generally conserved among TAs, while W288* and P290* are found to be specifically conserved among DABA aminotransferases. The second recognition site consists of the following residues in both chains A and B*: P109, P275, F297, and F300, where P275* forms a TA-conserved CH-π-bond with F297 (and *vice versa*). The surrounding residues P109 and F300 expand the hydrophobic contact surface between the chains of the dimer, where P109 is found to be conserved among DABA aminotransferases, although F300 appears unique for the EctB enzyme from *C. salexigens*. Interestingly, the second recognition site is found near the bound PLP coenzyme. We therefore suggest that the second recognition site might be structurally significant in terms of EctB's substrate specificity. Within the second site, we find the

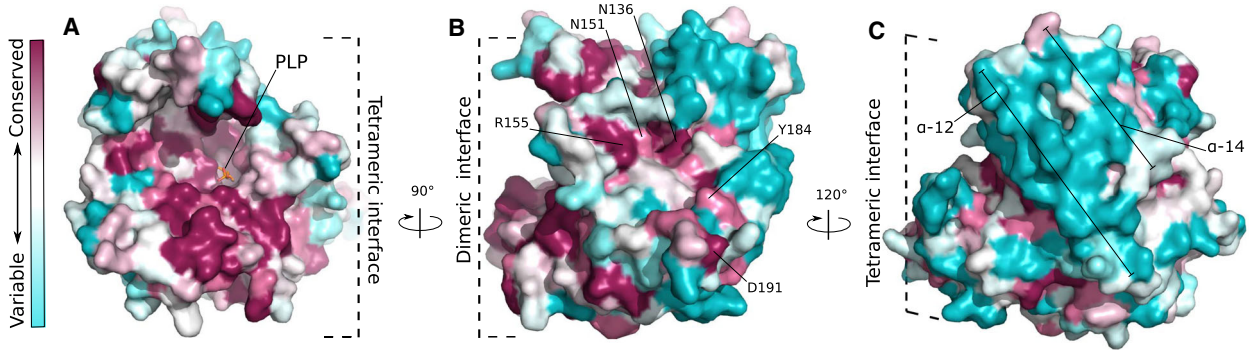

Fig. 6. Sequence conservation through ConSurf analysis mapped to the surface of one EctB monomer from *Chromohalobacter salexigens* DSM 3043. The color scheme describes the degree of conservation as indicated by the left legend, where teal = variable or nonconserved residues and purple = conserved residues. (A) Surface representation of the dimeric interface, looking directly at one chain. (B) The surface of the tetrameric interface, looking at one chain. The conserved residues (red) at the tetrameric interface form both hydrogen bonds and ionic bonds to the neighboring dimer (involving N136, N151, R155, Y184, and D191). (C) The surface with the least conserved sequences is found at α-helix-12 and α-helix-14, which face out toward the solvent and form multiple salt bridges. The figures were generated in PYMOL (www.pymol.org) from the output file generated by ConSurf and further annotated in INKSCAPE (www.inkscape.org).

(apparent) catalytically important R298 which is also located sequentially between F297 and F300 that seem critical for dimerization. It should be noted that W288 is a part of α-helix-10 preceding the gating loop and that the following P290, G291, F297, R298, and F300 are all located in the gating loop.

PLP-binding motif

The active site entrances are located on opposite sides of the dimeric interface (Fig. 9A). The active sites are completed by the gating loop provided by the neighboring chain in the dimer, which appears to be important for specificity in the substrate-binding pockets. PLP is located at the bottom of the active site cavities, where the pyridoxal ring is sandwiched between I240 and F138 (Fig. 9B). The nitrogen in the pyridoxal ring forms H-bonds to the carboxylate group of D238, a common motif for PLP-dependent enzymes, where D238 helps stabilize the carbanion intermediate state of PLP [38]. The position of D238 is further coordinated by H139. The Schiff-base formed between K267 and carbon-4 of the pyridoxal ring is located in close proximity to Q241, which likely plays an important role in PLP activation. The phosphate moiety of PLP is coordinated directly by H-bonds to the backbone amides of G111, T112, G295, and T296, and indirectly through water with S266 and V141 (Fig. 9C). The gating loop from the neighboring chain contributes to residues N294 and T296. The dipole moment of α-helix-4 creates a positive partial charge at the N terminus of the α-helix, which allows G111 to further stabilize PLP's phosphate moiety [58]. Put together, these coordinating interactions keep PLP anchored in the enzymes during catalysis [37].

An O-pocket and P-pocket confer dual substrate specificity

Aminotransferases, such as EctB, are described in the literature to have both an O-pocket and a P-pocket to accommodate substrates in the active site [43,45,57,59,60]. In EctB, the native substrates are ASA (acceptor) and L-glutamate (donor), or in the reverse reaction DABA (donor) and 2-OG (acceptor). Accordingly, in the *C. salexigens* EctB crystal structure, we identified two areas that appear to be substrate-binding pockets on either side of PLP (Figs 10 and 11). The P-pocket is in close proximity to the phosphate moiety of PLP, while the O-pocket is in closer proximity to the Schiff-base linkage between K267 and carbon-4 of the pyridoxal ring. Based on a multiple sequence alignment (MSA) of DABA aminotransferases and an MSA of TAs with different reaction specificities (Figs S1 and S2), we believe that some of the following observed residues found in the active site contribute to the dual substrate specificity of the herein presented EctB. The aromatic F154 occupies the space between the two binding pockets and is also located in close proximity to F138 that coordinates the pyridoxal ring. Both E205 and Q241 are found in other TAs in similar positions near the Schiff-base linkage and are thus likely important for the generic activation of PLP. The O-pocket is shaped by Y16, E210, E383, S385, and K393 that all point toward the center of it, directly above G46, A47, and G48, where

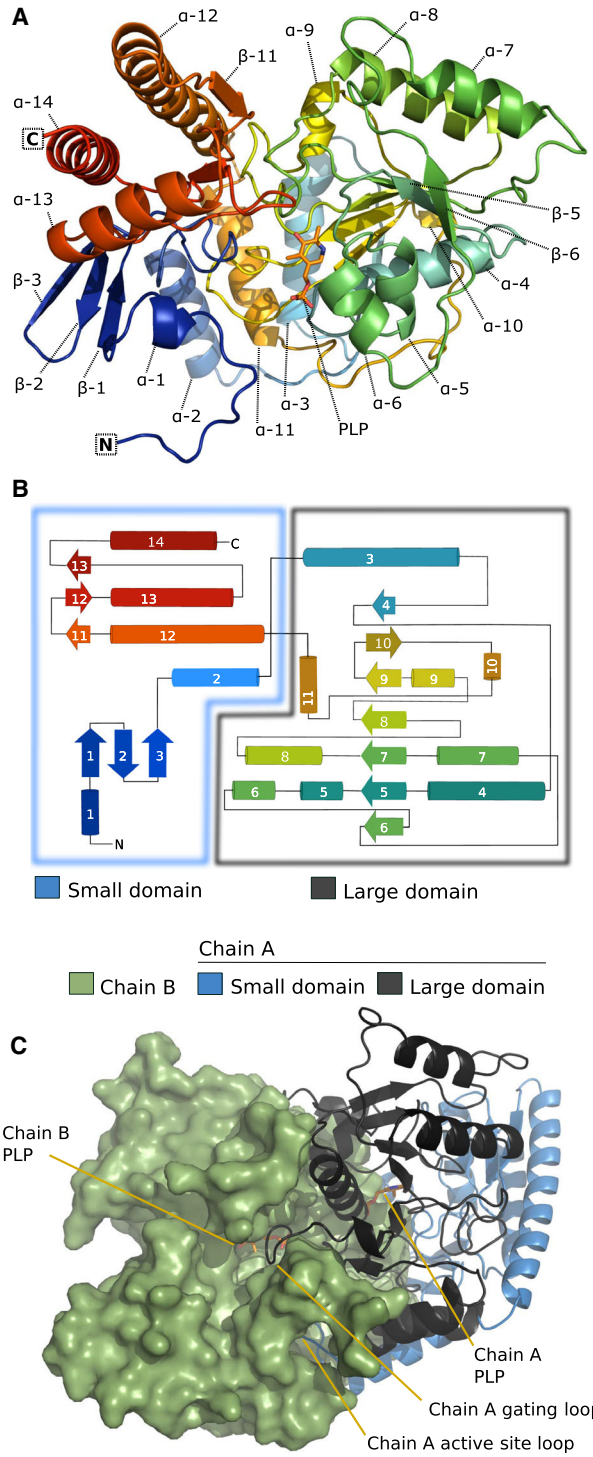

Fig. 7. Fold and topology of EctB from *Chromohalobacter salexigens* DSM 3043. (A) Cartoon representation of a monomer. (B) Topology diagram of one monomer. Secondary structure elements are shown in rainbow colors ranging from blue to red. The monomer consists of a large and a small domain, where the N and C termini are situated within the small domain. PLP is located on the interface between the domains. One chain contains 14 α-helices and 13 β-strands. (C) Illustration of the dimeric contact area. Both the small and large folding domains form contact with the neighboring chain at the dimeric interface. The large domain (black) binds PLP and forms the gating loop that completes the active site of the neighboring chain (green). Panels A and C were generated with PYMOL (www.pymol.org) and annotated in INKSCAPE, while panel B was created with INKSCAPE (www.inkscape.org).

all residues are part of the same chain that binds the corresponding catalytic PLP. The P-pocket is shaped by the gating loop provided by the neighboring chain, including the residues L76, D77, N294, T296, and R298. Of these, we suspect that the DABA aminotransferase conserved R298, located in the gating loop, plays a key role.

Review of the fingerprint residues for the DABA aminotransferases

Although thirteen fingerprint residues for the DABA aminotransferases have been suggested [40], there were no available structures at the time to verify their spatial orientation. Herein, we have therefore reviewed these fingerprint residues based on EctB's structural features at the dimer interface, the suggested substrate-binding pockets and the sequence conservation among DABA aminotransferases (Figs S1 and S2). Of the suggested fingerprint residues, we observe that P109 may be important for dimerization and that R298 is positioned centrally in the suggested O-pocket and thus likely interacts with the carboxylic group of substrates. We also find that the fingerprint residues Y16, G46, A47, G48, E210, E383, S385, and K393 shape the suggested O-pocket, that N294 is part of the gating loop that shapes the P-pocket, and that V141 interacts with PLP and is also located near the suggested P-pocket. Interestingly, V141 is commonly found as Arg in TAs with other reaction specificities near the phosphate moiety of PLP [61]. Fingerprint residues E210 and Q241 are likely important for the activation of PLP during catalysis, due to their structural position near the Schiff-base linkage between K267 and PLP. For the remaining two fingerprint residues, suggested to be K153 and Q292, we find that K153 is not conserved among DABA aminotransferases and that the EctB enzyme from *C. salexigens* contains a Gln at position 292. We therefore suggest that Q292 might not be vital for DABA aminotransferase substrate specificity. Furthermore, we notice that there is a positive charge conservation among the DABA aminotransferases at position 155 (R/K) that helps stabilize the tetrameric interface.

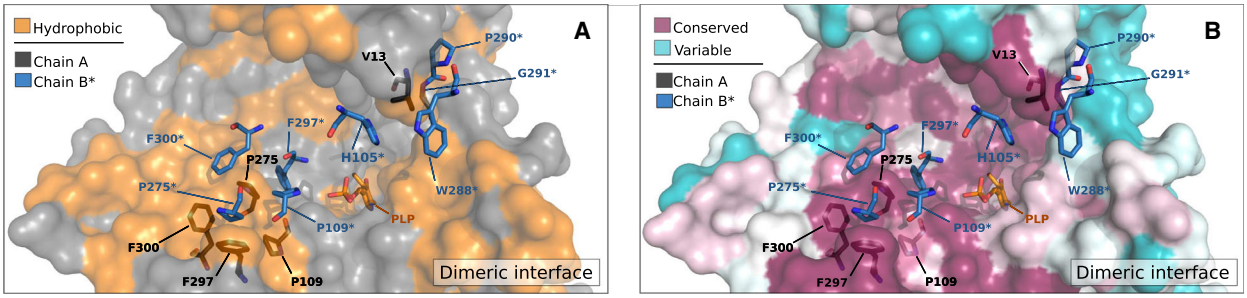

Fig. 8. Hydrophobic and aromatic sequence recognition in the dimer interface. The two hydrophobic recognition sites for dimer stabilization are located around the active site cavity (with a bound PLP). The first recognition site is composed of V13, H105*, W288*, P290*, and G291*, while the second site consists of P109, P275, F297, and F300 from both chains A and B*. (A) Hydrophobic residues (orange) and (B) sequence conservation through ConSurf analysis [ranging from variable (mint) to conserved (purple)] mapped onto the molecular surface of a monomer. The figure was generated in PYMOL (www.pymol.org) and annotated in INKSCAPE (www.inkscape.org).

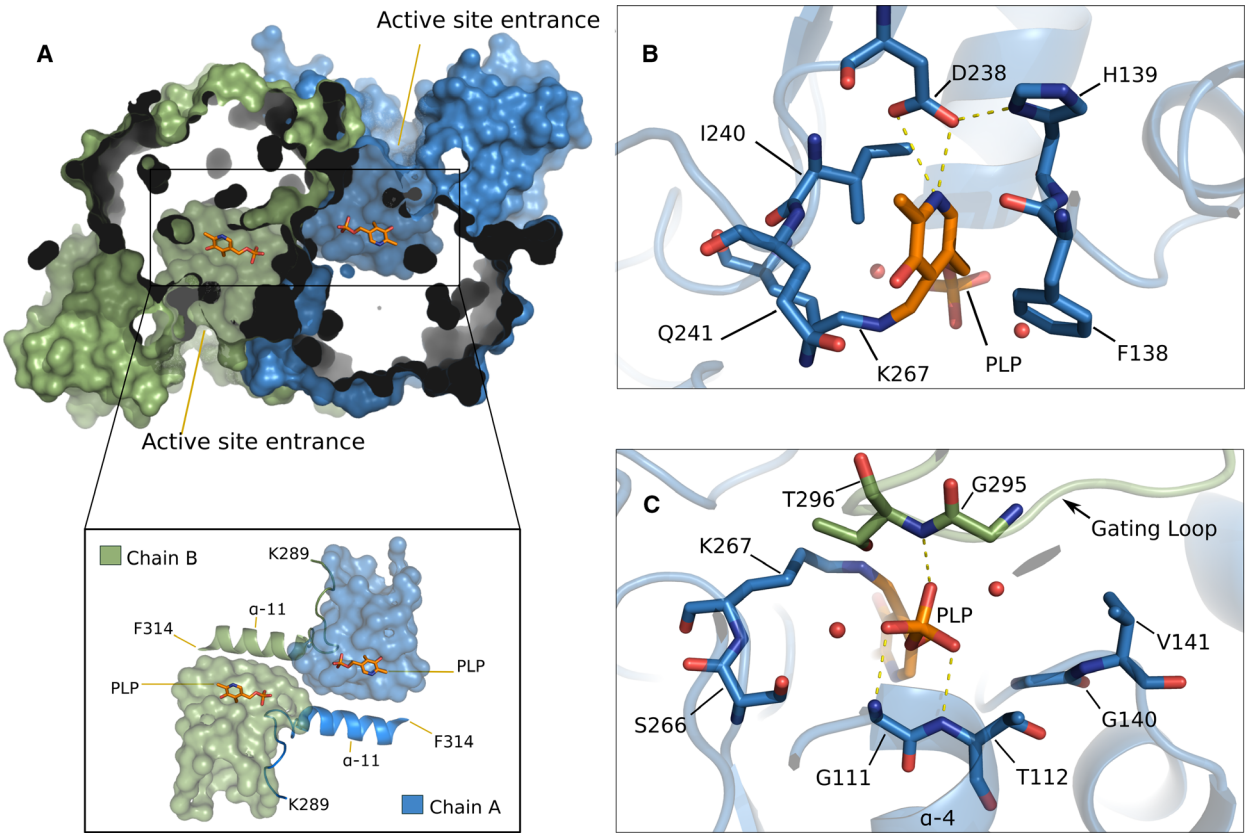

Fig. 9. The two active sites in one dimer, and coordination of PLP. (A) Slabbed views of the molecular surface of an EctB dimer with each chain colored individually (chain A in blue; chain B in green). The active site entrances with bound PLP are located on opposite sides of the dimeric interface. (B) Coordination of the pyridoxal ring moiety of PLP in EctB chain A. (C) Rotated views of PLP to show coordination of PLP's phosphate moiety in EctB chain A (blue). The figure was generated in PYMOL (www.pymol.org) and annotated in INKSCAPE (www.inkscape.org).

Discussion

In this study, we present the first crystal structure and biochemical characterization of the DABA aminotransferase EctB in the ectoine biosynthesis pathway, of the well-studied osmoadaptation organism *C. salexigens* DSM 3043. The EctB enzyme from *C. salexigens* serves as a benchmark enzyme for further annotation of the poorly characterized DABA aminotransferases, and future design of better biocatalysts for either

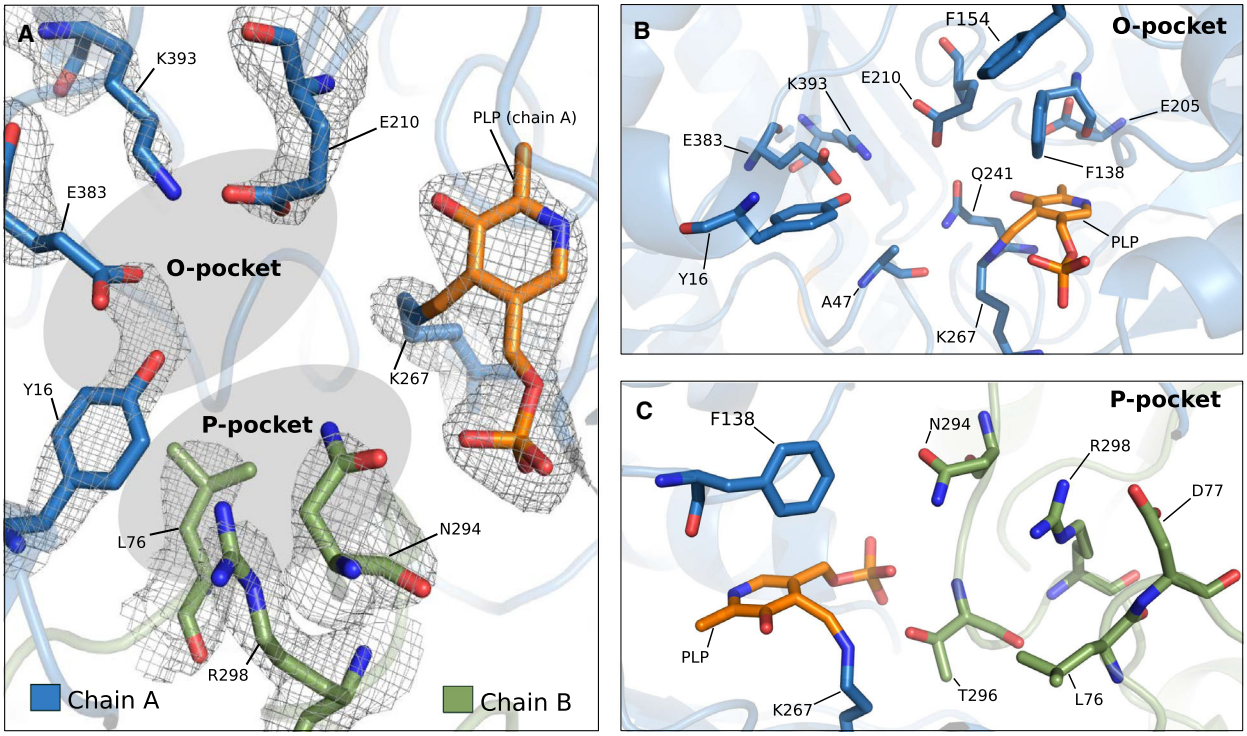

Fig. 10. Proposed substrate-binding pockets in the EctB DABA aminotransferase from *Chromohalobacter salexigens* DSM 3043. (A) The active site cavity contains two possible binding sites termed the O-pocket and the P-pocket that occupy opposite sides of the cavity between PLP and the solvent. Electron density shown around PLP and key residues. (B) The O-pocket is located near the Schiff-base linkage between PLP and K267. (C) The P-pocket is located closer to the phosphate moiety of PLP. The figure was generated in PYMOL (www.pymol.org) and annotated in INKSCAPE (www.inkscape.org).

increased production of ectoine, or potentially the synthesis of valuable amine bonds at the distal ω-position of substrates [1,62] that may be accepted by EctB from *C. salexigens*. Herein, we present a detailed investigation of EctB's oligomeric state and describe how dimerization is important for shaping the active site. We identify several residues that appear deterministic for the catalytic character of DABA aminotransferases and evaluate functional fingerprint residues that may aid further annotation of this protein family. Biochemical characterization reveals that the EctB enzyme from *C. salexigens* is a promising candidate for industrial application as it has (a) a broad pH tolerance with an activity optimum around pH 8.0, (b) high thermostability with a T_m of 71 °C, and (c) interestingly shown to be halotolerant with increased activity at higher salt concentrations.

The EctB enzyme from *C. salexigens* is a class III TA found to have a type-1 fold, thus expected to consist of functional dimers in solution [35]. In the literature, we find that EctB from both *Pseudomonas aeruginosa* PA01 and *Acinetobacter baumannii* has been reported as homotetramers by SEC [63,64], while EctB from the industrially relevant *H. elongata* was reported as a homohexamer [3]. These three EctBs are all purified or assayed in Tris/HCl buffers, which we found to act as a PLP scavenger. Our work shows that EctB from *C. salexigens* DSM 3043, purified in 50 mM HEPES (pH 7.5), is a homotetramer in solution. We propose several arguments that support this observation by analysis of the crystal structure and supporting biochemical data. From the structure, we find that EctB is a tetramer of two functional dimers, where each monomer in the dimeric unit is tightly associated with the other monomer at an extensive sequence-conserved contact surface. In contrast, the interface between the two dimers, that is, the tetrameric interface, is both less sequence-conserved and more loosely packed, although we do find some strong ionic bonds between the two dimers in EctB. This suggests that the tetrameric state is needed for EctB's overall structural integrity, while the dimeric unit is important for catalysis. The elution peak from the SEC shows that the relative molecular weight of EctB corresponds to 172.3 kDa, which is near the theoretical tetrameric size of EctB from *C. salexigens* of

Fig. 11. Specificity and binding residues in EctB from *Chromohalobacter salexigens* DSM 3043. The small domain (blue) is formed by the N and C terminus, while the large domain (gray) is formed from the central part of the polypeptide chain (see also Fig. 7). Conserved residues that stabilize the tetrameric interface (red) and residues that interact with PLP (yellow) are found in the large domain, while residues that seem important for specificity (green) and dimer association (pink) are located in both domains. The illustration was created in INKSCAPE (www.inkscape.org) based on the respective deposited crustal structure (6RL5) and the aligned sequence generated by SNAPGENE (www.snapgene.com).

188 kDa. Furthermore, the protein band in the Blue-Native-PAGE also shows that EctB in its native state gives a band between the 146- and 242-kDa protein marker, which again corresponds well with the theoretical tetrameric weight.

Positioning of the gating loop in the neighboring unit of the dimer appears to be critical for the correct reaction specificity and thus the overall catalytic character of EctB. We base this observation on two notions. First, we found that there are two notable recognition sites between the two units of the dimer, where many of the involved residues are part of the gating loop (i.e., W288, P290, G291, F297, and F300). Both sites are hydrophobic, and many of the specific recognition residues are conserved among DABA aminotransferases. Secondly, some of the gating loop residues (i.e., N294, T296, and R298) are centrally positioned in what appears to be the P-pocket (Fig. 10). This leads us to believe that modification of the gating loop might drastically effect EctB's catalytic character and could be an interesting target in downstream mutagenesis studies. In general, DABA aminotransferases coordinate substrates with both carboxylate and amino functional groups. Therefore, it is not surprising that the suggested P-pocket contains the centrally placed Arg298, positioned to coordinate a carboxylate group, as observed for other TAs (Fig. 2), while the suggested O-pocket contains the centrally placed Lys393 in close proximity to Glu210 which may function as a glutamate switch for dual substrate specificity (Fig. 12).

The incentive toward green chemistry has placed a higher demand on the development of resilient enzymes that can function in nonphysiological conditions, that is, extreme temperatures, high salt concentrations, and tolerance toward organic solvents [65]. The thermostable, pH tolerant, and halophilic properties of EctB, in combination with its ability to catalyze formation of amine bonds at the distal ω-position, make it a good candidate for industrial processes. Further inspection of EctB's structure reveals several determinants that may explain these properties. It is reported in the literature that halotolerant enzymes may have a high number of acidic residues at the surface [66], which can bind hydrated cations, and strengthen the hydration shell of the protein.

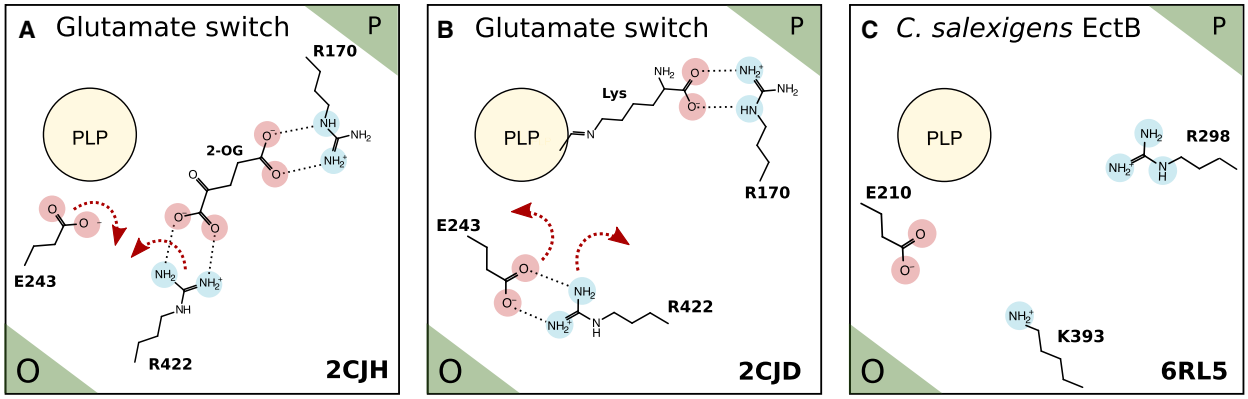

Fig. 12. Glutamate switch for dual substrate recognition in an (S)-selective TA from Mycobacterium tuberculosis compared to *Chromohalobacter salexigens* EctB DABA aminotransferase (A) The carboxylate groups of 2-OG coordinated through R422 from the O-pocket and R170 from the P-pocket (PDB ID: 2CJH). (B) The carboxylate group of lysine coordinated through R170 from the O-pocket, while R422 is switched 'off' by E243 (PDB ID: 2JCD). (C) The active site of the DABA aminotransferase EctB from *C. salexigens* (PDB ID: 6RL5). The illustration was created in INKSCAPE (www.inkscape.org) based on the respective deposited crystal structures.

Corroboratively, the EctB enzyme from *C. salexigens* has a slightly acidic pI of 6.17. More distinctly, we observe a large number of ionic bonds within each monomer and in the dimeric interface which are also, to some extent, found in the tetrameric interface. A study of an EctB from *H. elongata* [31] found increased whole-cell ectoine titer with the following point mutations: E36V, D180V, F320Y, and Q325R. When structurally compared to the EctB enzyme from *C. salexigens* DSM 3043, all point mutations locate to the protein surface away from the active site. These mutations likely increase structural flexibility through disruption of electrostatic interactions (E36V and D180V) and disruption of a hydrophobic patch (F320Y) or surface charge alterations (Q325R), respectively. Although these point mutations increase EctB's activity, there is likely a tradeoff between activity and stability which can be further explored with the herein presented crystal structure of EctB.

In summary, this study presents a detailed structural characterization of the DABA aminotransferase EctB, the first enzyme in the ectoine biosynthesis pathway from the well-studied organism *C. salexigens* DSM 3043. The EctB enzyme from *C. salexigens* serves as a benchmark enzyme for annotation of DABA aminotransferases and has been found to be the rate-limiting enzyme [31,32] for whole-cell ectoine production. The presented structure provides the groundwork for further optimization of the ectoine biosynthesis pathway by both rational and semirational design of the EctB enzyme. By targeting specific active site residues, or altering stabilizing interactions, the overall activity of EctB may be improved for increased ectoine production.

Material and methods

Cloning

The DABA TA (EctB) amplicon from *C. salexigens* DSM 3043 was sequence-optimized for protein expression in *Escherichia coli* and inserted into a pENTR221 vector by Thermo Fisher's GeneArt service. The final construct was flanked by an attL cloning site on the 5′ end and a hexahistidine tag followed by an attL cloning site on the 3′end. The construct was further cloned into a pDEST14 expression vector with a T7 promoter in an inducible *lac* operon (Invitrogen, Carlsbad, CA, USA) using the Gateway cloning system.

Expression

The EctB expression vector was transformed into chemically competent *E. coli* BL21 Star (DE3) cells (Invitrogen). Precultures were grown overnight in LB media (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl) with 100 $\mu g \cdot mL^{-1}$ ampicillin at 37 °C. The precultures were subsequently diluted (1 : 100) into TB media (2% (w/v) tryptone, 2.4% (w/v) yeast extract, 4% (v/v) glycerol, 17 mM KH_2PO_4, and 72 mM K_2HPO_4), with 100 $\mu g \cdot mL^{-1}$ ampicillin. The TB cultures were grown at 37 °C for 3 h, and protein expression was then induced by a final concentration of 1 mM IPTG and left to grow for an additional 18 h at 25 °C. The cells were pelleted by centrifugation at 9000 RCF and resuspended in lysis buffer (50 mM HEPES, 500 mM NaCl, and 15 mM imidazole adjusted to pH 7.5), prior to storage at −80 °C.

Purification

All steps of lysis and purification were carried out at 4 °C in buffers adjusted to pH 7.5. Frozen cells were thawed

with additional lysis buffer (supplemented with 200 μM PLP, lysozyme, DNase1, and a protease inhibitor tablet), and then lysed using a high-pressure cell disrupter (Constant Systems Ltd, Northants, UK) at 1.24 kPa and subsequently centrifuged at 30 000 RCF for 40 min. Lysate was loaded onto a 5 mL HisTrap column (GE Healthcare, Wood Dale, IL, USA) pre-equilibrated with lysis buffer, and then washed with 50 mL buffer A (50 mM HEPES, 500 mM NaCl, and 40 mM imidazole) followed by a gradient elution from 0% to 100% buffer B (50 mM HEPES, 500 mM NaCl, and 500 mM imidazole). The content of the eluted fractions was evaluated by SDS/PAGE. The eluted EctB fractions were supplemented with 100 μM PLP, prior to loading onto a Superdex 200 (26/200) gel filtration column (GE Healthcare) pre-equilibrated with buffer C (50 mM HEPES, 500 mM NaCl, 0.4 mM TCEP). SDS/PAGE was used to determine the final purity of EctB in buffer C, which was then either stored at 4 °C for crystallization trials or mixed with glycerol (10% (v/v) final concentration) and stored at −20 °C for subsequent activity assays.

Determination of the oligomeric state of EctB

The analytical size exclusion column Superdex 200 10/300 GL was calibrated with the following five proteins: ribonuclease (13.7 kDa), carbonic anhydrase (29 kDa), ovalbumin (43 kDa), conalbumin (75 kDa), and aldolase (158 kDa) in 50 mM HEPES pH 7.0 with 100 mM NaCl and 0.4 mM TCEP. Then, the EctB from *C. salexigens* was run on the column under the same conduction as the other five proteins. The same enzyme batch of purified EctB from *C. salexigens* was in parallel run in a BlueNative-PAGE to confirm the oligomeric state, as following. A total of six micrograms of EctB was mixed 1 : 1 with a loading buffer (100 mM BisTris pH 7.0, 1 M 6-aminocaproic acid, 15% v/v glycerol, and 1% w/v Coomassie Blue G-250) and loaded onto a NuPAGE 3–12% BisTris Protein Gel, and run at 80V for 7.5 h at 4 °C (power source limited to 12 mA and 25 W). The cathode buffer (15 mM BisTris pH 7.0, 50 mM tricine, and 0.02% w/v Coomassie Blue G-250) and the anode buffer (50 mM BisTris pH 7.0) were also precooled to 4 °C.

Coenzyme binding

Binding of the coenzyme PLP to EctB was determined by the intrinsic shift in the coenzyme's absorption maximum, which shifts from 390 to 420 nm when PLP is covalently bound through a Schiff-base intermediate in the enzyme's active site. The absorption maximum of EctB (319 μM) was compared to the absorption maximum of PLP (200 μM) using spectrophotometry, prepared in either 50 mM Tris/HCl (pH 7.5), 50 mM HEPES (pH 7.5), or ddH_2O.

Thermal stability

The melting temperature (T_m) of EctB, in 50 mM HEPES (pH 7.5) and 100 mM NaCl, was determined with the Nano-Differential Scanning Calorimeter III (Calorimetry Sciences Corporation). Both the reference buffer C and the EctB sample (1.4 $mg{\cdot}mL^{-1}$) were filtered and degassed for 15 min at 4 °C, prior to loading into the corresponding reference and sample chambers. The thermostability was measured at 3 atm from 4 to 90 °C with an increment of 1 $°C{\cdot}min^{-1}$.

Activity assay

Based on a previously published method [67], an end-point activity assay was established that detects the disappearance of the amino-donor 2-OG during EctB's conversion of DABA to ASA. The relative activity optimum was determined based on three adjustments of buffer C: (a) pH range of 6.5–10.0 at 20 °C, (b) NaCl concentrations from 0 to 700 mM at 20 °C, and (c) temperature range from 0 to 80 °C with a pH of 8.0. The assay was carried out in three steps. Firstly, the reaction was incubated for 3 min (15 μM EctB, 400 μM 2-OG, and 2 mM DABA) in buffer C. Secondly, the reaction was stopped by a 1 : 20 dilution into the 1,4-diamino-4,5-methylenedioxybenzene (DMB)-labeling mixture (700 mM HCl, 500 mM β-ME, 1 mM DMB, and 36 mM $NaSO_4$) and incubated at 50 °C for 4 h. Thirdly, the incubated mixture was cooled on ice for 5 min prior to lowering the pH by diluting (2 : 1) with 2.5 M BisTris (pH 7.2), and subsequently recording the fluorescence (λ_{ex} = 367 nm, λ_{em} = 444 nm). Each measurement was carried out with three-four parallels of the enzymatic reaction and the negative control (enzyme replaced by buffer C with 10% (v/v) glycerol).

Crystallization and data collection

EctB was crystallized by hanging drop vapor diffusion at 20 °C, over a 500 μL reservoir (0.2 M sodium malonate, 0.1 M BisTris propane pH 6.5, and 30% (w/v) PEG 3350). The drop consisted of 1.5 μL EctB (5.3 $mg{\cdot}mL^{-1}$) and 1.5 μL reservoir. The crystal was cryoprotected in the reservoir solution with 20% (v/v) glycerol added, and diffracted to 2.45 Å resolution at the ESRF beamline ID23-1.

Structure determination

The crystal structure was determined by molecular replacement applying a search model based on the dimer of *E. coli* γ-aminobutyrate aminotransferase (PDB ID: 1SF2; sequence identity of ~ 30%). The EctB model was improved using Swiss-Model [68], which threaded the majority of the sequence onto the search model. Phaser [69]

in the CCP4i interface [70] was run, where prior analyses indicated the presence of eight molecules in the asymmetric unit. A weak overall hit was improved through consecutive runs of rigid body and positional refinement in REFMAC5 [71]. Subsequently, BUCCANEER [72] autobuilt 3187 out of the possible 3384 amino acid residues, and brought the R_{work}/R_{free} down to ~ 0.32/~ 0.32. Several cycles of refinement in PHENIX.REFINE [73] interspersed with manual model building in COOT [74] gave the final model with R_{work}/R_{free} of 0.1850/0.2441, and excellent overall statistics considering the resolution (Table 2).

Bioinformatic analysis

Assembly analysis was performed with the PDBEPISA server [54] to estimate the buried surface area between monomers. Sequence conservation of the EctB enzyme from *C. salexigens* was visualized with the ConSurf Server [55], with the integrated psi-blast that included 400 sequences with 35–95% sequence identity. A MSA of DABA aminotransferases and an MSA of 14 TAs was generated with MUSCLE [75]. The DABA aminotransferases (Fig. S2) were collected from the curated Swiss-Prot database in UniProt [76], while the sequences of the 15 general TAs (Fig. S1) were gathered from the PDB database [50]. Structural illustrations were generated using PYMOL (www.pymol.org).

Acknowledgements

Provision of beam time at the BESSY II electron storage ring (Helmholtz-Zentrum Berlin) BL14.1 for testing crystals and at the European Synchrotron Radiation Facility (ESRF) ID23-1 for data collection is highly acknowledged. Travel funding for data collection was provided by Research Council of Norway [247732, 2016]. UiT The Arctic University of Norway is gratefully acknowledged for financial support.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

The study was conceived by BA and IL. The construct was designed by BA. HTH cloned, expressed, purified, and crystallized the protein. IL determined the structure, which was jointly refined by HTH and IL. HTH performed the biochemical characterization. HTH and IL drafted the manuscript, and all authors edited and reviewed it before submission.

References

1 Kelly SA, Pohle S, Wharry S, Mix S, Allen CCR, Moody TS & Gilmore BF (2018) Application of ω-transaminases in the pharmaceutical industry. *Chem Rev* **118**, 349–367.
2 Galinski EA, Pfeiffer H-P & Trüper HG (1985) 1,4,5,6-Tetrahydro-2-methyl-4-pyrimidinecarboxy acid. A novel cyclic acid from halophilic phototophic bacterua of the genus *Ectothiorhodospira*. *Eur J Biochem* **149**, 135–139.
3 Ono H, Sawada K, Khunajakr N, Tao T, Yamamoto M, Hiramoto M, Shinmyo A, Takano M & Murooka Y (1999) Characterization of biosynthetic enzymes for ectoine as a compatible solute in a moderately Halophilic Eubacterium, *Halomonas elongata*. *J Bacteriol* **181**, 91–99.
4 Yancey PH (2004) Compatible and counteracting solutes: protecting cells from the Dead Sea to the deep sea. *Sci Prog* **87**, 1–24.
5 Yancey PH (2005) Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *J Exp Biol* **208**, 2819–2830.
6 Brown AD (1976) Microbial water stress. *Bacteriol Rev* **40**, 803–846.
7 Lippert K & Gallinski EA (1992) Enzyme stabilization by ectoine-type compatible solutes: protection against heating, freezing and drying. *Appl Mircobiol Biotechnol* **37**, 61–65.
8 Botta C, Di Giorgio C, Sabatier AS & De Meo M (2008) Genotoxicity of visible light (400–800 nm) and photoprotection assessment of ectoin, L-ergothioneine and mannitol and four sunscreens. *J Photochem Photobiol B* **91**, 24–34.
9 Graf R, Anzali S, Buenger J, Pfluecker F & Driller H (2008) The multifunctional role of ectoine as a natural cell protectant. *Clin Dermatol* **26**, 326–333.
10 Widderich N, Hoppner A, Pittelkow M, Heider J, Smits SH & Bremer E (2014) Biochemical properties of Ectoine hydroxylases from extremophiles and their wider taxonomic distribution among microorganisms. *PLoS One* **9**, e93809.
11 Widderich N, Czech L, Elling FJ, Konneke M, Stoveken N, Pittelkow M, Riclea R, Dickschat JS, Heider J & Bremer E (2016) Strangers in the archaeal world: osmostress-responsive biosynthesis of ectoine and hydroxyectoine by the marine thaumarchaeon *Nitrosopumilus maritimus*. *Environ Microbiol* **18**, 1227–1248.
12 Czech L, Hermann L, Stoveken N, Richter AA, Hoppner A, Smits SHJ, Heider J & Bremer E (2018) Role of the extremolytes ectoine and hydroxyectoine as stress protectants and nutrients: genetics, phylogenomics, biochemistry, and structural analysis. *Genes (Basel)* **9**, 177.

13 Van-Thuoc D, Hashim SO, Hatti-Kaul R & Mamo G (2013) Ectoine-mediated protection of enzyme from the effect of pH and temperature stress: a study using *Bacillus halodurans* xylanase as a model. *Appl Microbiol Biotechnol* **97**, 6271–6278.
14 Kunte HJ, Lentzen G & Galinski EA (2014) Industrial production of the cell protectant ectoine: protection mechanism, processes and production. *Curr BIotechnol* **3**, 10–25.
15 Harding T, Brown MW, Simpson AG & Roger AJ (2016) Osmoadaptative strategy and its molecular signature in obligately halophilic heterotrophic protists. *Genome Biol Evol* **8**, 2241–2258.
16 Weinisch L, Kuhner S, Roth R, Grimm M, Roth T, Netz DJA, Pierik AJ & Filker S (2018) Identification of osmoadaptive strategies in the halophile, heterotrophic ciliate *Schmidingerothrix salinarum*. *PLoS Biol* **16**, e2003892.
17 Kuhlmann AU, Bursy J, Gimpel S, Hoffmann T & Bremer E (2008) Synthesis of the compatible solute ectoine in *Virgibacillus pantothenticus* is triggered by high salinity and low growth temperature. *Appl Environ Microbiol* **74**, 4560–4563.
18 Seip B, Galinski EA & Kurz M (2011) Natural and engineered hydroxyectoine production based on the *Pseudomonas stutzer*i ectABCD-*ask* gene cluster. *Appl Environ Microbiol* **77**, 1368–1374.
19 Louis P & Galinski EA (1997) Characterization of genes for the biosynthesis of the compatible solute ectoine from *Marinococcus halophilus* and osmoregulated expression in *Escherichia coli*. *Microbiology* **143**, 1141–1149.
20 Reshetnikov AS, Khmelenina VN, Mustakhimov II & Trotsenko YA (2011) Genes and enzymes of ectoine biosynthesis in halotolerant methanotrophs. *Methods Enzymol* **495**, 15–30.
21 Czech L, Hoppner A, Kobus S, Seubert A, Riclea R, Dickschat JS, Heider J, Smits SHJ & Bremer E (2019) Illuminating the catalytic core of ectoine synthase through structural and biochemical analysis. *Sci Rep* **9**, 364.
22 Bursy J, Pierik AJ, Pica N & Bremer E (2007) Osmotically induced synthesis of the compatible solute hydroxyectoine is mediated by an evolutionarily conserved ectoine hydroxylase. *J Biol Chem* **282**, 31147–31155.
23 Schwibbert K, Marin-Sanguino A, Bagyan I, Heidrich G, Lentzen G, Seitz H, Rampp M, Schuster SC, Klenk HP, Pfeiffer F *et al.* (2011) A blueprint of ectoine metabolism from the genome of the industrial producer *Halomonas elongata* DSM 2581 T. *Environ Microbiol* **13**, 1973–1994.
24 Pastor JM, Salvador M, Argandona M, Bernal V, Reina-Bueno M, Csonka LN, Iborra JL, Vargas C, Nieto JJ & Canovas M (2010) Ectoines in cell stress protection: uses and biotechnological production. *Biotechnol Adv* **28**, 782–801.
25 Rodriguez-Moya J, Argandona M, Iglesias-Guerra F, Nieto JJ & Vargas C (2013) Temperature- and salinity-decoupled overproduction of hydroxyectoine by *Chromohalobacter salexigens*. *Appl Environ Microbiol* **79**, 1018–1023.
26 Tanimura K, Nakayama H, Tanaka T & Kondo A (2013) Ectoine production from lignocellulosic biomass-derived sugars by engineered *Halomonas elongata*. *Bioresour Technol* **142**, 523–529.
27 Ning Y, Wu X, Zhang C, Xu Q, Chen N & Xie X (2016) Pathway construction and metabolic engineering for fermentative production of ectoine in *Escherichia coli*. *Metab Eng* **36**, 10–18.
28 Perez-Garcia F, Ziert C, Risse JM & Wendisch VF (2017) Improved fermentative production of the compatible solute ectoine by *Corynebacterium glutamicum* from glucose and alternative carbon sources. *J Biotechnol* **258**, 59–68.
29 Czech L, Poehl S, Hub P, Stoveken N & Bremer E (2018) Tinkering with osmotically controlled transcription allows enhanced production and excretion of ectoine and hydroxyectoine from a microbial cell factory. *Appl Environ Microbiol* **84**, e01772-17.
30 Czech L, Stoveken N & Bremer E (2016) EctD-mediated biotransformation of the chemical chaperone ectoine into hydroxyectoine and its mechanosensitive channel-independent excretion. *Microb Cell Fact* **15**, 126.
31 Chen W, Zhang S, Jiang P, Yao J, He Y, Chen L, Gui X, Dong Z & Tang SY (2015) Design of an ectoine-responsive AraC mutant and its application in metabolic engineering of ectoine biosynthesis. *Metab Eng* **30**, 149–155.
32 Giesselmann G, Dietrich D, Jungmann L, Kohlstedt M, Jeon EJ, Yim SS, Sommer F, Zimmer D, Muhlhaus T, Schroda M *et al.* (2019) Metabolic engineering of *Corynebacterium glutamicum* for high-level ectoine production: design, combinatorial assembly, and implementation of a transcriptionally balanced heterologous ectoine pathway. *Biotechnol J* **14**, e1800417.
33 Fallet C, Rohe P & Franco-Lara E (2010) Process optimization of the integrated synthesis and secretion of ectoine and hydroxyectoine under hyper/hypo-osmotic stress. *Biotechnol Bioeng* **107**, 124–133.
34 Piubeli F, Salvador M, Argandona M, Nieto JJ, Bernal V, Pastor JM, Canovas M & Vargas C (2018) Insights into metabolic osmoadaptation of the ectoines-producer bacterium *Chromohalobacter salexigens* through a high-quality genome scale metabolic model. *Microb Cell Fact* **17**, 2.
35 Schneider G, Kack H & Lindqvist Y (2000) The manifold of vitamin B6 dependent enzymes. *Structure* **8**, R1–R6.

36 John RA (1995) Pyridoxal phosphate-dependent enzymes. *Biochim Biopys Acta* **1248**, 81–96.
37 Denesyuk AI, Denessiouk KA, Korpela T & Johnson MS (2002) Functional attributes of the phosphate group binding cup of pyridoxal phosphate-dependent enzymes. *J Mol Biol* **316**, 155–172.
38 Toney MD (2011) Controlling reaction specificity in pyridoxal phosphate enzymes. *Biochim Biophys Acta* **1814**, 1407–1418.
39 Hirotsu K, Goto M, Okamoto A & Miyahara I (2005) Dual substrate recognition of aminotransferases. *Chem Rec* **5**, 160–172.
40 Steffen-Munsberg F, Vickers C, Kohls H, Land H, Mallin H, Nobili A, Skalden L, van den Bergh T, Joosten HJ, Berglund P *et al*. (2015) Bioinformatic analysis of a PLP-dependent enzyme superfamily suitable for biocatalytic applications. *Biotechnol Adv* **33**, 566–604.
41 Goto M, Miyahara I, Hayashi H, Kagamiyama H & Hirotsu K (2003) Crystal structures of branched-chain amino acid aminotransferase complexed with glutamate and glutarate: true reaction intermediate and double substrate recognition of the enzyme. *Biochemistry* **42**, 3725–3733.
42 Tripathi SM & Ramachandran R (2006) Direct evidence for a glutamate switch necessary for substrate recognition: crystal structures of lysine epsilon-aminotransferase (Rv3290c) from Mycobacterium tuberculosis H37Rv. *J Mol Biol* **362**, 877–886.
43 Schiroli D & Peracchi A (2015) A subfamily of PLP-dependent enzymes specialized in handling terminal amines. *Biochim Biophys Acta* **1854**, 1200–1211.
44 van Oosterwijk N, Willies S, Hekelaar J, Terwisscha van Scheltinga AC, Turner NJ & Dijkstra BW (2016) Structural basis of the substrate range and enantioselectivity of two (S)-selective omega-transaminases. *Biochemistry* **55**, 4422–4431.
45 Manta B, Cassimjee KE & Himo F (2017) Quantum chemical study of dual-substrate recognition in omega-transaminase. *ACS Omega* **2**, 890–898.
46 Palacio CM, Rozeboom HJ, Lanfranchi E, Meng Q, Otzen M & Janssen DB (2019) Biochemical properties of a Pseudomonas aminotransferase involved in caprolactam metabolism. *FEBS J* **286**, 4086–4102.
47 Mitchell JBO, Thornton JM, Singh J & Price SL (1992) Towards an understanding of the arginine-aspartate Interaction. *J Mol Biol* **226**, 251–262.
48 Geders TW, Gustafson K & Finzel BC (2012) Use of differential scanning fluorimetry to optimize the purification and crystallization of PLP-dependent enzymes. *Acta Crystallogr Sec F Biol Cryst Commun* **68**, 596–600.
49 Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**, 403–410.
50 Gilliland G, Berman HM, Weissig H, Shindyalov IN, Westbrook J, Bourne PE, Bhat TN & Feng Z (2000) The protein data bank. *Nucleic Acids Res* **28**, 235–242.
51 Liu W, Peterson PE, Carter RJ, Zhou X, Langston JA, Fisher AJ & Toney MD (2004) Crystal structures of unbound and aminooxyacetate-bound *Escherichia coli* γ-aminobutyrate aminotransferase. *Biochemistry* **43**, 10896–10905.
52 Bruce H, Nguyen Tuan A, Mangas Sanchez J, Leese C, Hopwood J, Hyde R, Hart S, Turkenburg JP & Grogan G (2012) Structures of a gamma-aminobutyrate (GABA) transaminase from the s-triazine-degrading organism *Arthrobacter aurescens* TC1 in complex with PLP and with its external aldimine PLP-GABA adduct. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **68**, 1175–1180.
53 Sayer C, Isupov MN, Westlake A & Littlechild JA (2013) Structural studies of *Pseudomonas* and *Chromobacterium* ω-aminotransferases provide insights into their differing substrate specificity. *Acta Crystallogr D Biol Crystallogr* **69**, 564–576.
54 Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797.
55 Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T & Ben-Tal N (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* **44**, W344–W350.
56 Ikai A (1980) Thermostability and aliphatic index of globular proteins. *J Biochem* **88**, 1895–1898.
57 Lyskowski A, Gruber C, Steinkellner G, Schurmann M, Schwab H, Gruber K & Steiner K (2014) Crystal structure of an (R)-selective ω-transaminase from *Aspergillus terreus*. *PLoS One* **9**, e87350.
58 Parker MH & Hefford MA (1998) Introduction of potential helix-capping residues into an engineered helical protein. *Biotechnol Appl Biochem* **28**, 69–76.
59 Malik MS, Park ES & Shin JS (2012) Features and technical applications of ω-transaminases. *Appl Microbiol Biotechnol* **94**, 1163–1171.
60 Han SW, Park ES, Dong JY & Shin JS (2015) Active-site engineering of ω-transaminase for production of unnatural amino acids carrying a side chain bulkier than an ethyl substituent. *Appl Environ Microbiol* **81**, 6994–7002.
61 Shin J-S & Kim B-G (2002) Exploring the active site of amine: pyruvate aminotransferase on the basis of the substrate-reactivity relationship: how the enzyme controls substrate specificity and sterioselectivity. *J Org Chem* **67**, 2848–2853.
62 Cairns R, Gomm A, Ryan J, Clarke T, Kulcinskaja E, Butler K & O'Reilly E (2019) Conversion of aldoses to valuable ω-amino alcohols using amine transaminase biocatalysts. *ACS Catalysis* **9**, 1220–1223.

63 Ikai H & Yamamoto S (1997) Identification and analysis of a gene encoding L-2,4-diaminobutyrate:2-ketoglutarate 4-aminotransferase Involved in the 1,3-diaminopropane production pathway in *Acinetobacter baumannii*. *J Bacteriol* **179**, 5118–5125.
64 Vandenende CS, Vlasschaert M & Seah SY (2004) Functional characterization of an aminotransferase required for pyoverdine siderophore biosynthesis in *Pseudomonas aeruginosa* PAO1. *J Bacteriol* **186**, 5596–5602.
65 Li PY, Zhang Y, Xie BB, Zhang YQ, Hao J, Wang Y, Wang P, Li CY, Qin QL, Zhang XY *et al*. (2017) Structural and mechanistic insights into the improvement of the halotolerance of a marine microbial esterase by increasing intra- and interdomain hydrophobic interactions. *Appl Environ Microbiol* **83**, e01286-01217.
66 Mevarech M, Frolow F & Gloss LM (2000) Halophilic enzymes: proteins with a grain of salt. *Biophys Chem* **86**, 155–164.
67 Kato S, Ikuta T, Hemmi H, Takahashi S, Kera Y & Yoshimura T (2012) Enzymatic assay for D-aspartic acid using D-aspartate oxidase and oxaloacetate decarboxylase. *Biosci Biotechnol Biochem* **76**, 2150–2152.
68 Waterhouse A, Rempfer C, Heer FT, Studer G, Tauriello G, Bordoli L, Bertoni M, Gumienny R, Lepore R, Bienert S *et al*. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296–W303.
69 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.
70 Potterton E, Briggs P, Turkenburg M & Dodson E (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr D Biol Crystallogr* **59**, 1131–1137.
71 Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355–367.
72 Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr* **62**, 1002–1011.
73 Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH & Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**, 352–367.
74 Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501.
75 Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797.
76 The UniProt Consortium (2018) The UniProt consortium. *Nucleic Acids Res* **47**, D506–D515.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. DABA aminotransferase from *Chromohalobacter salexigens* (CsEctB), aligned against 14 other transaminases with different substrate specificities.

Fig. S2. Multiple sequence alignment of 14 DABA aminotransferases with conserved residues highlighted.

Appendix S1. Supplementary PyMol session files.