Illuminating the catalytic core of ectoine synthase through structural and biochemical analysis

Laura Czech, Astrid Höppner, Stefanie Kobus, Andreas Seubert, Ramona Riclea, Jeroen S. Dickschat, Johann Heider, Sander H. J. Smits & Erhard Bremer

Ectoine synthase (EctC) is the signature enzyme for the production of ectoine, a compatible solute and chemical chaperone widely synthesized by bacteria as a cellular defense against the detrimental effects of osmotic stress. EctC catalyzes the last step in ectoine synthesis through cyclo-condensation of the EctA-formed substrate N-gamma-acetyl-L-2,4-diaminobutyric acid via a water elimination reaction. We have biochemically and structurally characterized the EctC enzyme from the thermo-tolerant bacterium *Paenibacillus lautus* (*Pl*). EctC is a member of the cupin superfamily and forms dimers, both in solution and in crystals. We obtained high-resolution crystal structures of the (*Pl*)EctC protein in forms that contain (i) the catalytically important iron, (ii) iron and the substrate N-gamma-acetyl-L-2,4-diaminobutyric acid, and (iii) iron and the enzyme reaction product ectoine. These crystal structures lay the framework for a proposal for the EctC-mediated water-elimination reaction mechanism. Residues involved in coordinating the metal, the substrate, or the product within the active site of ectoine synthase are highly conserved among a large group of EctC-type proteins. Collectively, the biochemical, mutational, and structural data reported here yielded detailed insight into the structure-function relationship of the (*Pl*)EctC enzyme and are relevant for a deeper understanding of the ectoine synthase family as a whole.

Ectoine [(S)−2-methyl-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid] and its derivative 5-hydroxyectoine [(4S, 5S)−5-hydroxy-2-methyl-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid] are representatives of a specialized group of organic osmolytes, the compatible solutes. These types of compounds are widely used by members of each domain of life as effective cytoprotectants, in particular against the detrimental effects caused by high external salinity or osmolarity on cellular hydration, physiology, and growth. Compatible solutes are well suited for this demanding task because their physico-chemical attributes make them compliant with cellular biochemistry. The function-preserving properties of these solutes allow their high-level cellular accumulation, a process raising the osmotic potential of the cytoplasm, which then in turn counteracts the high osmolarity-instigated efflux of water from the cell. At the same time, the solvent properties of the crowded cytoplasm are optimized for vital biochemical reactions, so that growth can occur under osmotically unfavorable circumstances. Compatible solutes also serve as stabilizers of proteins, macromolecular assemblies, and even entire cells, both in vitro and in vivo, a property that led to their description in the literature as chemical chaperones.

With the notable exception of their recent discovery in a restricted group of obligate halophilic protists and a few Archaea, ectoines are primarily synthesized by members of the *Bacteria*, as judged from the wide occurrence of the ectoine (*ectABC*) and 5-hydroxyectoine (*ectD*) biosynthetic genes in microbial genome sequences. Biosynthesis of ectoine proceeds from L-aspartate-β-semialdehyde, a central hub in bacterial amino acid and cell wall biosynthesis and is catalyzed by the sequential enzyme reactions of L-2,4-diaminobutyrate
transaminase (EctB; EC 2.6.1.76), L-2,4-diaminobutyrate acetyltransferase (EctA; EC 2.3.1.178), and ectoine synthase (EctC; EC 4.2.1.108)\textsuperscript{36}. 5-Hydroxyectoine is synthesized from ectoine by the ectoine hydroxylase (EctD; EC 1.14.11.55)\textsuperscript{33,37,38} through a regio- and stereo-specific hydroxylation reaction\textsuperscript{33}. In comparison with ectoine, this chemical modification endows 5-hydroxyectoine with enhanced, or additional, cellular protection against selected environmental challenges (e.g., heat and desiccation stress)\textsuperscript{21,22,39,40}. High-level expression of the ectoine/5-hydroxyectoine biosynthetic genes is typically triggered by increases in the external osmolarity and sometimes also by extremes in either high or low growth temperatures\textsuperscript{4,41–43}. Disruption of the \textit{ectABC(D)} genes causes sensitivity to such environmental challenges\textsuperscript{39,44}, highlighting the prominent role played by ectoines in the stress management of microbial cells\textsuperscript{4,5,45}.

The macromolecule stabilizing, cytoprotective, and anti-inflammatory effects of ectoines have generated considerable interest in the commercial and medical use of these natural products and led to an industrial scale biotechnological production process that delivers ectoines on the scale of tons. This process relies on the salt-tolerant bacterium \textit{Halomonas elongata} as a natural production host\textsuperscript{4,45–47}. In addition, recombinant cell factories are increasingly developed to offer additional means to produce these valuable compounds biotechnologically\textsuperscript{5,48–56}. Hence, both from the perspective of practical applications and basic science\textsuperscript{4,5,45}, a deeper understanding of the enzymatic and structural properties of the ectoine and 5-hydroxyectoine biosynthetic enzymes is desirable. This has already been achieved to a considerable degree for the ectoine hydroxylase (EctD)\textsuperscript{31,37,38,37}, a member of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenases enzyme super-family\textsuperscript{58}.

In comparison with EctD, the ectoine biosynthetic enzymes (EctABC)\textsuperscript{36,39} are biochemically and structurally far less understood. While the EctA and EctB enzymes have amino acid sequence-related counterparts in various biochemical pathways in microorganisms\textsuperscript{5,32,36}, the ectoine synthase (EctC) is considered the diagnostic enzyme for ectoine production\textsuperscript{30,31,60}. It can be used in database searches of microbial genome sequences to identify potential ectoine producers\textsuperscript{5}. EctC catalyzes the final step in the ectoine biosynthetic route and cyclizes the linear, EctA-produced, \(\mathrm{N}^\gamma\)-acetyl-L-2,4-diaminobutyric acid (\(\mathrm{N}^\gamma\)-ADABA) to ectoine through a water elimination reaction\textsuperscript{36,60} (Fig. 1a). Basic biochemical characterizations of ectoine synthases from the extremophiles \textit{H. elongata}. 

---

**Figure 1.** Ectoine synthase-catalyzed enzyme reaction scheme, purification and quaternary assembly of the (\textit{Pl}) EctC protein. (a) Scheme of the ectoine synthase catalyzed cyclo-condensation of the substrate \(\mathrm{N}^\gamma\)-ADABA to form ectoine. (b) Coomassie brilliant blue stained 15% SDS-polyacrylamide gel with different amounts of the affinity chromatography purified (\textit{Pl})EctC-Strep-tag II protein. The PageRuler Prestained Protein Ladder (Thermo Scientific) was used as a sizing marker. A picture of the original SDS-gel can be found in Fig. S1. (c) SEC-MALS analysis to determine the quaternary assembly of the purified (\textit{Pl})EctC-Strep-tag II protein.
elongata, Methylmicrobium alcaliphilum, and Acidiphilium cryptum\textsuperscript{63,64}, the cold-adapted marine bacterium Sphingopyxis alaskensis\textsuperscript{65}, and the nitriﬁng archaean Nitrosopumilus maritimus\textsuperscript{66} have been carried out, but a deeper understanding of this enzyme is lacking.

Crystallographic analysis of the psychrophilic S. alaskensis EctC protein [(Sa)EctC] showed that ectoine synthase is a member of the cupin superfamily\textsuperscript{69}. This family comprises a large group of pro- and eukaryotic proteins built on a common structural scaffold; its members can perform a variety of both enzymatic and non-enzymatic functions\textsuperscript{43-46}. Most of these proteins contain catalytically important divalent metals (e.g., iron, copper, zinc, manganese, cobalt, or nickel) that allow different types of chemistry to occur within the conﬁnes of an evolutionarily conserved tertiary structure\textsuperscript{66,67}. Biochemical studies conducted with the (Sa)EctC protein revealed for the ﬁrst time that ectoine synthase is also a metal-dependent enzyme, with Fe\textsuperscript{2+} as the physiologically most likely relevant catalyst\textsuperscript{69}. This ﬁnding has obvious ramifications with respect to the details of the not yet fully understood reaction mechanism catalyzed by the EctC enzyme (Fig. 1a).

One of our major aims is to understand the structure-function relationship of the ectoine synthase through biochemical and crystallographic analysis. In this respect, the structure of the (Sa)EctC protein that we reported recently\textsuperscript{69} provided only restricted functional information because it contained neither the catalytically important metal, nor the substrate or the reaction product. To signiﬁcantly advance our understanding of the key enzyme for ectoine biosynthesis, we explored the EctC protein from the thermo-tolerant bacterium Paenibacillus lautus (Pl)\textsuperscript{60} for biochemical and structural studies. We now report here crystallographic views into the catalytic core of the ectoine synthase prior and subsequent to enzyme catalysis. These crystal structures lay the framework for a proposal for the EctC-mediated cyclo-condensation reaction mechanism. Since the residues participating in metal, substrate, and reaction product binding are highly conserved among a very large collection of EctC-type proteins, the data provided here for the (Pl)EctC enzyme are relevant for a structural and functional understanding of the extended ectoine synthase family as a whole.

Results

Overproduction, purification, and oligomeric state of the (Pl)EctC ectoine synthase. Ectoine biosynthetic genes are present in microorganisms able to colonize ecological niches with rather different physico-chemical attributes\textsuperscript{4,5,30,31}. One of these ectoine-producing microorganisms is Paenibacillus lautus (Pl) strain Y4.12MC10, a Gram-positive spore-forming bacterium that was originally isolated from the effluent of the Obsidian hot spring in the Yellowstone National Park (USA). It can grow under laboratory conditions up to 50 °C\textsuperscript{68}. We therefore explored the EctC protein from P. lautus for biochemical and structural studies in the hope that the properties of the ectoine synthase from this thermo-tolerant strain might be more suitable for crystallographic analysis than the psychrophilic EctC protein derived from the cold-adapted bacterium S. alaskensis (Sa)\textsuperscript{69} which yielded crystal structures only in its apo-form\textsuperscript{69}. The (Pl)EctC protein possesses 130 amino acids, has a calculated molecular mass of 14.7 kDa, and possesses a calculated isoelectric point of 4.7.

We constructed an expression vector carrying a codon-optimized version of the P. lautus ectC gene for its heterologous expression in Escherichia coli that would lead to a recombinant protein fused at its C-terminus to a Strep-tag II afﬁnity peptide [(Pl)EctC-Strep-tag II] allowing its puriﬁcation by afﬁnity chromatography (Figs 1b and S1). The quaternary assembly of the puriﬁed (Pl)EctC-Strep-tag II protein was assessed by size exclusion chromatography followed by multi-angle light scattering (SEC-MALS); these experiments yielded a value for the molecular mass of 32.2 kDa of the (Pl)EctC-Strep-tag II protein in solution (Fig. 1c). Since the calculated molecular mass for the recombinant (Pl)EctC-Strep-tag II protein is 15.87 kDa, the ectoine synthase from P. lautus is a dimer in solution.

Biochemical and kinetic properties of the recombinant (Pl)EctC enzyme. Taking into account that the (Sa)EctC enzyme is an iron-dependent enzyme\textsuperscript{69}, we included 0.1 mM (NH\textsubscript{4})\textsubscript{2}Fe(SO\textsubscript{4})\textsubscript{2} into the buffer solution when we assessed its enzymatic activity. We ﬁrst determined some basic biochemical properties of the recombinant (Pl)EctC enzyme with respect to its temperature and pH optima and its tolerance towards salt (Fig. 2a–c). In keeping with the thermo-tolerant physiology of the (Pl)EctC enzyme, we explored the EctC protein from the thermo-tolerant bacterium Paenibacillus lautus (Pl)\textsuperscript{60} for biochemical and structural studies. We now report here crystallographic views into the catalytic core of the ectoine synthase prior and subsequent to enzyme catalysis. These crystal structures lay the framework for a proposal for the EctC-mediated cyclo-condensation reaction mechanism. Since the residues participating in metal, substrate, and reaction product binding are highly conserved among a very large collection of EctC-type proteins, the data provided here for the (Pl)EctC enzyme are relevant for a structural and functional understanding of the extended ectoine synthase family as a whole.

Results

Overproduction, purification, and oligomeric state of the (Pl)EctC ectoine synthase. Ectoine biosynthetic genes are present in microorganisms able to colonize ecological niches with rather different physico-chemical attributes\textsuperscript{4,5,30,31}. One of these ectoine-producing microorganisms is Paenibacillus lautus (Pl) strain Y4.12MC10, a Gram-positive spore-forming bacterium that was originally isolated from the effluent of the Obsidian hot spring in the Yellowstone National Park (USA). It can grow under laboratory conditions up to 50 °C\textsuperscript{68}. We therefore explored the EctC protein from P. lautus for biochemical and structural studies in the hope that the properties of the ectoine synthase from this thermo-tolerant strain might be more suitable for crystallographic analysis than the psychrophilic EctC protein derived from the cold-adapted bacterium S. alaskensis (Sa)\textsuperscript{69} which yielded crystal structures only in its apo-form\textsuperscript{69}. The (Pl)EctC protein possesses 130 amino acids, has a calculated molecular mass of 14.7 kDa, and possesses a calculated isoelectric point of 4.7.

We constructed an expression vector carrying a codon-optimized version of the P. lautus ectC gene for its heterologous expression in Escherichia coli that would lead to a recombinant protein fused at its C-terminus to a Strep-tag II afﬁnity peptide [(Pl)EctC-Strep-tag II] allowing its puriﬁcation by afﬁnity chromatography (Figs 1b and S1). The quaternary assembly of the puriﬁed (Pl)EctC-Strep-tag II protein was assessed by size exclusion chromatography followed by multi-angle light scattering (SEC-MALS); these experiments yielded a value for the molecular mass of 32.2 kDa of the (Pl)EctC-Strep-tag II protein in solution (Fig. 1c). Since the calculated molecular mass for the recombinant (Pl)EctC-Strep-tag II protein is 15.87 kDa, the ectoine synthase from P. lautus is a dimer in solution.

Biochemical and kinetic properties of the recombinant (Pl)EctC enzyme. Taking into account that the (Sa)EctC enzyme is an iron-dependent enzyme\textsuperscript{69}, we included 0.1 mM (NH\textsubscript{4})\textsubscript{2}Fe(SO\textsubscript{4})\textsubscript{2} into the buffer solution when we assessed its enzymatic activity. We ﬁrst determined some basic biochemical properties of the recombinant (Pl)EctC enzyme with respect to its temperature and pH optima and its tolerance towards salt (Fig. 2a–c). In keeping with the thermo-tolerant physiology of the (Pl)EctC enzyme, we explored the EctC protein from the thermo-tolerant bacterium Paenibacillus lautus (Pl)\textsuperscript{60} for biochemical and structural studies. We now report here crystallographic views into the catalytic core of the ectoine synthase prior and subsequent to enzyme catalysis. These crystal structures lay the framework for a proposal for the EctC-mediated cyclo-condensation reaction mechanism. Since the residues participating in metal, substrate, and reaction product binding are highly conserved among a very large collection of EctC-type proteins, the data provided here for the (Pl)EctC enzyme are relevant for a structural and functional understanding of the extended ectoine synthase family as a whole.

Biochemical and kinetic properties of the recombinant (Pl)EctC enzyme. Taking into account that the (Sa)EctC enzyme is an iron-dependent enzyme\textsuperscript{69}, we included 0.1 mM (NH\textsubscript{4})\textsubscript{2}Fe(SO\textsubscript{4})\textsubscript{2} into the buffer solution when we assessed its enzymatic activity. We ﬁrst determined some basic biochemical properties of the recombinant (Pl)EctC enzyme with respect to its temperature and pH optima and its tolerance towards salt (Fig. 2a–c). In keeping with the thermo-tolerant physiology of the (Pl)EctC enzyme, we explored the EctC protein from the thermo-tolerant bacterium Paenibacillus lautus (Pl)\textsuperscript{60} for biochemical and structural studies. We now report here crystallographic views into the catalytic core of the ectoine synthase prior and subsequent to enzyme catalysis. These crystal structures lay the framework for a proposal for the EctC-mediated cyclo-condensation reaction mechanism. Since the residues participating in metal, substrate, and reaction product binding are highly conserved among a very large collection of EctC-type proteins, the data provided here for the (Pl)EctC enzyme are relevant for a structural and functional understanding of the extended ectoine synthase family as a whole.

Biochemical and kinetic properties of the recombinant (Pl)EctC enzyme. Taking into account that the (Sa)EctC enzyme is an iron-dependent enzyme\textsuperscript{69}, we included 0.1 mM (NH\textsubscript{4})\textsubscript{2}Fe(SO\textsubscript{4})\textsubscript{2} into the buffer solution when we assessed its enzymatic activity. We ﬁrst determined some basic biochemical properties of the recombinant (Pl)EctC enzyme with respect to its temperature and pH optima and its tolerance towards salt (Fig. 2a–c). In keeping with the thermo-tolerant physiology of the (Pl)EctC enzyme, we explored the EctC protein from the thermo-tolerant bacterium Paenibacillus lautus (Pl)\textsuperscript{60} for biochemical and structural studies. We now report here crystallographic views into the catalytic core of the ectoine synthase prior and subsequent to enzyme catalysis. These crystal structures lay the framework for a proposal for the EctC-mediated cyclo-condensation reaction mechanism. Since the residues participating in metal, substrate, and reaction product binding are highly conserved among a very large collection of EctC-type proteins, the data provided here for the (Pl)EctC enzyme are relevant for a structural and functional understanding of the extended ectoine synthase family as a whole.
display in some cases notable differences in their kinetic properties but it should be noted that the various EctC enzymes were assessed under rather different buffer, pH, temperature, and salt concentrations (and partially also at different levels of purity). This makes a direct comparison with the kinetic parameters determined in this study for the \((\text{Pl})\)EctC enzyme difficult. We also note in this context that the \((\text{Pl})\)EctC enzyme is the first ectoine synthase for which a substantial substrate (or product) inhibition was detected (Fig. 2d).

Crystallization of the ectoine synthase in complex with iron, the substrate \(N^\gamma-\text{ADABA}\), and the enzyme reaction product ectoine. Crystals of the \((\text{Pl})\)EctC protein were obtained using commercial screens and by slightly optimizing the composition of the crystallization solution. Crystals were grown in a solution consisting of 0.2 M ammonium sulfate, 0.1 M phosphate citrate (pH 4.2), 20% (v/v) PEG 300, and 10% (v/v) glycerol. Since the ectoine synthase is an iron dependent enzyme, we added \(\text{Fe(II)Cl}_2\) to the crystallization solution of the purified \((\text{Pl})\)EctC protein to a final concentration of 4 mM and pre-incubated this mixture for at least 30 minutes prior to crystallization trials. Without this pre-incubation in the presence of iron, no crystals with a quality suitable for X-ray analysis could be grown. This approach yielded crystals which diffracted to a maximum of 1.6 Å for the \((\text{Pl})\)EctC::Fe protein complex (Table 1). To obtain crystals of the \((\text{Pl})\)EctC enzyme with its substrate \(N^\gamma-\text{ADABA}\) or its reaction product ectoine, we added these compounds to a final concentration of 40 mM and of 20 mM, respectively, to \((\text{Pl})\)EctC preparations pre-incubated with iron. Crystals which diffracted to a maximum of 2.0 Å were obtained for the \((\text{Pl})\)EctC::Fe/\(N^\gamma-\text{ADABA}\) complex and 2.5 Å for the \((\text{Pl})\)EctC::Fe/ectoine complex (Table 1).

The high-resolution dataset obtained for the \((\text{Pl})\)EctC::Fe crystal complex was phased using molecular replacement with the crystallographic data of the previously solved structure of the \((\text{Sa})\)EctC ectoine synthase (PDB entry 5BXX) as the search model. After several rounds of model building using COOT and subsequent refinement, the structure of the full-length \((\text{Pl})\)EctC protein was solved (Fig. 3a). This was a substantial improvement over the \(S.\ alaskensis\ EctC\) crystal structure where the spatial orientation of 22 amino acids from the \((\text{Sa})\)EctC COOH-terminus could not be localized in the electron density map, probably due to enhanced structural flexibility of this psychrophilic protein. The data- and refinement statistics for the \((\text{Pl})\)EctC::Fe complex are listed in Table 1.
prior to crystallization trials. Within the cupin fold 63–67 of the (Pl)EctC::Fe/ectoine, respectively.

Table 1. Data collection, phasing and refinement statistics for (Pl)EctC::Fe, (Pl)EctC::Fe/N-γ-ADABA, and (Pl)EctC::Fe/ectoine, respectively.

| Data collection | (Pl)EctC::Fe (PDB entry: 5ONM) | (Pl)EctC::Fe/N-γ-ADABA (PDB entry: 5ONN) | (Pl)EctC::Fe/ectoine (PDB entry: 5ONO) |
|-----------------|-----------------------------|---------------------------------|----------------------------------|
| Wavelength (Å)  | 0.953                       | 0.953                           | 0.953                            |
| Space group     | P 31 2 1                    | P 31 2 1                        | P 31 2 1                         |
| Cell dimensions | a, b, c (Å)                  | 71.13; 71.13; 68.66              | 70.91; 70.91; 68.54               | 71.41; 71.41; 68.83               |
| α, β, γ (°)     | 90.0, 90.0, 120.0            | 90.0, 90.0, 120.0               | 90.0, 90.0, 120.0                |
| Processing      |                             |                                 |                                 |
| Resolution (Å)  | 35.56–1.6 (1.7–1.6)          | 31.49–2.0 (2.1–2.0)              | 46.01–2.5 (2.6–2.5)               |
| Rmerge          | 6.5 (45.2)                  | 4.5 (6.89)                      | 8.2 (17.6)                       |
| Mean I/σ(I)     | 19.14 (1.70)                | 61.45 (47.39)                  | 8.2 (3.2)                        |
| Completeness (%)| 99.35 (97.83)               | 98.05 (97.58)                  | 99.32 (98.33)                    |
| Total reflections| 586136 (34379)             | 293346 (29757)                 | 145322 (15326)                   |
| Unique reflections| 31125 (3015)              | 13580 (1328)                   | 7282 (706)                       |
| Redundancy      | 18.8 (11.4)                 | 21.6 (22.4)                    | 5.4 (3.0)                        |
| Refinement      |                             |                                 |                                 |
| Resolution (Å)  | 35.56–1.6                    | 31.49–2.0                      | 46.01–2.5                        |
| No. of reflections | 545106                     | 281608                         | 36963                           |
| Rwork/Rfree    | 0.17/0.20                    | 0.17/0.21                     | 0.18/0.24                       |
| RMSD Bond length (Å) | 0.011                      | 0.016                         | 0.014                          |
| Bond angles (°) | 1.41                        | 1.80                          | 1.48                           |
| No. atoms       |                             |                                 |                                 |
| Protein         | 1115                        | 1115                          | 1115                           |
| Ligand          | 6                           | 12                            | 12                             |
| Water           | 133                         | 152                           | 42                             |
| B-factors       |                             |                                 |                                 |
| Protein (Å)     | 22.80                       | 21.60                         | 20.70                          |
| Ligand (Å)      | 31.20                       | 35.20                         | 86.10                          |

Once the 1.6 Å structure of the (Pl)EctC::Fe protein was completely refined, it was used to phase the other two datasets of the (Pl)EctC protein containing iron and either the substrate [(Pl)EctC::Fe/N-γ-ADABA] or the reaction product [(Pl)EctC::Fe/ectoine] of the ectoine synthase. A summary of the data collection statistics, refinement details, and model content of these two (Pl)EctC crystal structures is given in Table 1 as well. The (Pl)EctC-Strep-tag II protein carries an affinity tag attached to its carboxy-terminus; this tag consists of an octapeptide with an additional two amino acid linker (SA-WSHPGFEK) sequence. In the (Pl)EctC crystal structures the linker region and five amino acids of Strep-tag II sequence was visible, while the last three amino acids were disordered and were therefore not visible in the electron density. The crystallographic data of the three (Pl)EctC structures reported in this communication have been deposited in the RCSB Protein Data Bank (https://www.rcsb.org/) under accession number 5ONM [(Pl)EctC::Fe], 5ONN [(Pl)EctC::Fe/N-γ-ADABA], and 5ONO [(Pl)EctC::Fe/ectoine], respectively.

The (Pl)EctC::Fe/N-γ-ADABA crystals diffracted to very high resolution, but they suffered from severe radiation damage, a process, which lowered the diffraction rapidly during data collection. This phenomenon was encountered for more than 20 examined (Pl)EctC::Fe/N-γ-ADABA crystals. We therefore designed an experiment with a very low radiation dose and exposure time of the crystals to the X-ray beam. This yielded finally a dataset with a resolution of 2.0 Å in which no evidence for radiation damage was observed when the crystallographic data were processed. After manual rebuilding and several cycles of refinement, the (Pl)EctC::Fe/N-γ-ADABA crystal structure was finalized with an R-factor of 17.1% and an Rfree of 21.7%. The Ramachandran plot revealed that 99.3% of the residues are in the allowed region. Already during the first round of refinement of the (Pl)EctC::Fe/N-γ-ADABA complex, extra density was visible next to the iron atom, large enough to fit the N-γ-ADABA molecule (Fig. S2a). This allowed us to pinpoint and trace the spatial position of the N-γ-ADABA substrate within the active site of the ectoine synthase.

Crystals of the (Pl)EctC protein in complex with ectoine were obtained after the addition of 20 mM ectoine prior to crystallization trials. Within the cupin fold of the (Pl)EctC protein (Fig. 3a), clear density was visible for the ectoine molecule (Fig. S2b), that was subsequently manually placed and included during refinement. After manual rebuilding and several cycles of refinement, the structure of the (Pl)EctC::Fe/ectoine complex was finalized with an R-factor of 18.2%, an Rfree of 24%, and a resolution of 2.5 Å. The Ramachandran plot revealed that 99.3% obeyed the rules and fitted into the corresponding plot.
A comparison of the three (Pl)EctC crystal structures revealed a root-mean-square deviation (r.m.s.d.) ranging between 0.4 and 0.8 Å over 136 Cα atoms. Keeping in mind that crystal structures are snapshots of the three-dimensional space a given protein can potentially adopt, these data indicate that both the binding of the substrate N-γ-ADABA to the EctC protein pre-complexed with iron and the binding of the enzyme reaction product ectoine to the protein do not trigger any gross overall structural changes in the thermo-tolerant (Pl)EctC enzyme. A high degree of structural identity was also found between the thermo-tolerant (Pl)EctC protein and the cold-tolerant (Sa)EctC protein; an overall comparison of the three (Pl)EctC crystal structures with that of the (Sa)EctC protein revealed r.m.s.d. values ranging between 1.2 and 1.4 Å over 105 Cα atoms. These data indicate that the (Sa)EctC and the (Pl)EctC ectoine synthases display a highly similar three-dimensional structure despite the fact that these proteins were derived from microorganisms living in ecophysiologically rather different habitats; the effluent of a hot spring (P. latus) and permanently cold ocean waters (S. alaskensis).

Overall fold of the (Pl)EctC protein and analysis of its dimer interface. Since the monomers of the three (Pl)EctC crystal structures are nearly identical in overall shape, we only describe in the following section the overall structure for the high-resolution (1.6 Å) (Pl)EctC::Fe complex in detail. The structure of the (Pl)EctC protein consists of 11 β-strands (βI-βXI) and two α-helices (α-a and α-b) (Fig. 3a). The β-strands form two anti-parallel β-sheet regions consisting of βII, βIII, βV, βI, βVI, βIX and βXI. These sets of anti-parallel β-sheets are packed against each other, forming a cup-shaped β-sandwich with a topology characteristic for the widely found cupin-fold of proteins.

The structure of the psychrophilic (Sa)EctC protein has previously been solved in two different conformations, which were coined the open and semi-closed states. In the latter state, only part of the carboxy-terminus of the (Sa)EctC protein is visible in the electron density map, and it folds into a small helix (α-b) that closes the active site of the enzyme. The formation of the helix α-b induces a reorientation and shift of a long unstructured loop connecting the beta-sheets βIV and βVI of the (Sa)EctC protein, resulting in the formation of the stable β-strand βV. In contrast to the (Sa)EctC crystal structure, the COOH-terminus of the thermo-tolerant P. latus EctC protein (Fig. 3a) was completely resolved in the electron density map. The remaining segment of the previously unresolved part of the carboxy-terminus of the ectoine synthase from S. alaskensis flanks the cupin fold of the (Pl)EctC protein and protrudes out of the protein (Fig. 3a).
The (Pl)EctC protein is a dimer in solution as revealed by our SEC-MALS analysis (Fig. 1c). Since the asymmetric unit of the (Pl)EctC crystal revealed only a monomer, we inspected the crystal packing and analyzed the respective monomer/monomer interactions to elucidate the functional dimer within the crystal structure. The data resulting from this analysis show that the (Pl)EctC dimer in the crystal (Fig. 3b) is composed of two monomers arranged in a head-to-tail orientation; it is stabilized via strong interactions mediated by the N-terminus (sequence 1MIVKHS) from monomer A and 3-strand βIII from monomer B (sequence 1GMYYALB) (Fig. 3b–d). The interactions between these two β-strands rely primarily on backbone contacts (Fig. 3b–d). In addition to these interactions, some weaker hydrophobic interactions between the two monomers are also observed in some loop regions connecting the β-strands; these will probably play more subtle roles in dimer formation. Since the (Pl)EctC is a head-to-tail dimer, the interaction interface between the monomers occurs twice in the dimer assembly (Fig. 3b,c). As determined by PISA (Proteins, Interfaces, Structures and Assemblies) analysis, the (Pl)EctC monomers interact in the dimer assembly through an extensive surface area of 1501 Å² involving 16 hydrogen bonds and 4 salt bridges. The predicted substantial binding energy of −28.2 kcal mol⁻¹ of the two (Pl)EctC monomers indicates that these regions represent the predominant interface within the ectoine synthase dimer (Fig. 3b–d).

To identify the structurally closest homologs of the (Pl)EctC protein, we performed a DALI search which recovered, as expected, a variety of cupin-type proteins, most of which are functionally uncharacterized. Not surprisingly, the apo-form of the (Sa)EctC protein (PDB accession number: 5BXX) was found as the structurally closest homolog of (Pl)EctC; it had a Z-score of 21.1. Among the proteins with the highest Z-scores recovered by the DALI-search that had been biochemically and functionally studied were the KdgF²⁶ and DddK²⁷ crystal structures with Z-scores of 13.4 and 13.0, respectively. Like the (Pl)EctC protein, the cupin-type KdgF and DddK proteins are dimers that possess an overall topology and a dimer interface very similar to that observed for the ectoine synthase.

The KdgF protein from Halomonas sp. is an enzyme that catalyzes a step in the microbial metabolism of uronate sugars from two abundant sources of biomass, pectin and alginate. KdgF mediates the conversion of pectin- and alginate-derived 4,5-unsaturated mono-uronates to form linear ketonized forms. Interestingly, KdgF performs an enzyme reaction (hydrolysis of a cyclic molecule) opposite to that performed by EctC (cyclo-condensation of a linear metabolite). We note in this context that the ectoine synthase from Hal. elongata displays a hydrolytic activity towards synthetic ectoine derivatives with either reduced or expanded ring sizes but the equilibrium for the EctC-catalyzed N-γ-ADABA to ectoine biotransformation lies almost completely on the side of the cyclic condensation product ectoine. KdgF exhibits an amino acid sequence identity to (Pl)EctC of only 17%, but its crystal structure possesses an r.m.s.d. of 2.1 Å (over 109 Co atoms) to the ectoine synthase. As determined by PISA, the surface area of the dimer interface of KdgF is 1501 Å², and the monomers interact via 16 hydrogen bonds, and 4 salt bridges, yielding an overall predicted binding energy of the two KdgF monomers of about −30 kcal mol⁻¹. The crystal structure of KdgF contains a nickel ion, a metal that might have been acquired during the affinity chromatography of the His-tagged recombinant KdgF enzyme purified from E. coli cell extracts. As observed with other cupins, the KdgF enzyme is promiscuous with respect to the catalytically required metal, with Co²⁺ being the most effective catalyst among the tested divalent metals.

The DddK protein (PDB accession number: 5TFZ) from the marine bacterium Pelagibacter ubique HTCC1062 is a dimethylsulfoxoniopropionate (DMSP) lyase. This enzyme is involved in the catabolism of the organosulfur compound DMSP, an environmentally abundant organic osmolyte produced by marine algae, yielding the reaction products acrylate and the climate-active gas dimethylsulfide (DMS). The P. ubique DddK protein exhibits an amino acid sequence identity to (Pl)EctC protein of 16.3% and its crystal structure possesses an r.m.s.d. of 1.2 Å (over 126 Co atoms) to the ectoine synthase from P. laetus. The surface area of the dimer interface of DddK is 1556.1 Å², and the monomers interact with an overall predicted binding energy of −28 kcal mol⁻¹. Crystal structures of the DddK protein with either Ni²⁺ or Fe²⁺/Zn²⁺ were recovered, attesting again to the promiscuity of cupins with respect to the metal cofactor used for catalytic activity.

Structural features of the ectoine synthase and functional relevance of the iron-binding site for enzyme activity. Keeping in mind that the (Sa)EctC ectoine synthase is a metal-dependent enzyme, with Fe²⁺ as the physiologically most relevant catalyst, we added Fe(II)Cl₂ to the (Pl)EctC protein solution prior to the crystallization at a final concentration of 4 mM. Clear electron density was visible in the 1.6 Å structure for an atom with a strong electron density, which cannot be accounted for by a water molecule. To identify the probable nature of this ion, we modeled Mg²⁺, Ca²⁺, Fe²⁺, Ni²⁺, Co²⁺, and Zn²⁺ into the electron density and refined the (Pl)EctC crystal structure again. Only when we refined the (Pl)EctC structure modeled with Fe²⁺, we observed neither negative nor positive differences in electron density, indicating that iron is indeed the most probable element present in the crystallized (Pl)EctC protein.

Within the (Pl)EctC: Fe crystal structure, the iron atom is tetrahedrally ligated via interactions with the side chains of Glu-57, Tyr-84, and His-92 (Fig. 4a). The distance between the iron atom and the side chains of these three residues are 2.9 Å, 2.8 Å, and 2.9 Å, respectively. A water molecule completes the tetrahedral arrangement of the (Pl)EctC iron-binding site in the substrate-free (Pl)EctC: Fe crystal structure (Fig. 4a); it is present at a distance of 2.9 Å relative to that of the iron atom. In the previously reported (Sa)EctC crystal structure, no metal atom was visible but a water molecule occupied the same position that we observed here for the iron atom in the (Pl)EctC crystal structure (Fig. 4b). An overlay of the three iron-coordinating amino acid residues in the (Pl) EctC and (Sa)EctC₀ structures revealed that they are perfectly superimposable (Fig. 4b), indicating that (i) the iron-binding site in the ectoine synthase is already preformed in the absence of the catalytically important iron co-factor and that (ii) the binding of the iron atom does not seem to trigger substantial structural rearrangements in the overall fold of the enzyme.
If one considers the architecture of the iron-binding site of (P)EctC (Fig. 4a) with respect to the previously established consensus sequence for the amino acid sequences involved in metal coordination in cupin-type proteins, both common and distinct features are found. The first consensus cupin motif \[G(X)E(Y)(X)\] is altered in the (P)EctC protein to \(G(X)W(Y)(X)\), and the second consensus motif \([G(X)P(X)G(X)H(X)N]\) is changed in (P)EctC to \(G(X)P(X)G(X)Y(X)G(X)H\) (note: the letters in bold represent those residues that coordinate the metal) (Fig. 5a,b). Thus, in the first consensus cupin motif, none of the two canonical His residues is present in (P)EctC; however, the canonical Glu residue (Glu-57) is conserved (Fig. 5a,b). In the second cupin motif of the (P)EctC protein, a Tyr residue (Tyr-84) replaces the canonical His residue, and the motif is elongated to include another His residue (His-92) involved in iron binding (Fig. 4a). Variations in the metal-binding motifs of cupins occur frequently, but to the best of our knowledge, the one identified here for ectoine synthase is novel (Figs 4a and 5).

In Fig. 4a we have highlighted the positions of the two cupin motifs within the overall (P)EctC crystal structure, and we point out their position in EctC protein sequences in Fig. 5. By inspecting a recently reported extended amino acid sequence alignment of 582 EctC-type proteins, we found that minor variations in the amino acid sequences of the overall cupin motifs exist, but none of them affects the residues contacting the iron atom directly (Figs 4a and 5). This is highlighted in an abbreviated alignment of amino acid sequences of 15 randomly picked ectoine synthases from the previously reported dataset of 582 EctC-type proteins (Fig. 5).

The strict evolutionary conservation of the three iron-contacting residues in ectoine synthases attests to their likely critical role for enzyme function. To assess the individual contributions of the (P)EctC iron-binding residues Glu-57, Tyr-84, and His-92 (Fig. 4a) for the functionality of this enzyme, we replaced each of them individually with an Ala residue via site-directed mutagenesis. The three (P)EctC variants could be overproduced and purified with the same efficiency as the wild-type (P)EctC protein. Each of the individual Ala substitution mutations rendered the mutant (P)EctC proteins in essence catalytically inactive with remaining levels of enzyme activity.

**Figure 4.** The iron-binding site in the (P)EctC protein and enzyme activities of selected (P)EctC variants. (a) The iron (represented by an orange sphere) is coordinated by the side-chains of Glu-57, Tyr-84, and His-92 of the (P)EctC protein with distances of 2.9 Å, 2.8 Å, and 2.9 Å, respectively. The iron-binding site in the substrate-free (P)EctC crystal structure also contains a localized water molecule (blue sphere); it has a distance of 2.9 Å to the iron atom. The two conserved cupin-motifs include those residues that coordinate the metal ion \[G(X)E(Y)(X)\]. The iron-binding site in the (P)EctC crystal structure. The first cupin motif \(G(X)W(Y)(X)\) is shown in blue, and the second cupin motif \([G(X)P(X)G(X)Y(X)G(X)H]\) is represented as sticks. A water molecule (blue sphere) in the (P)EctC crystal structure occupies the same location as the Fe(II) ion (orange sphere) in the (P)EctC structure. (b) Overlay of the iron-binding site in the (P)EctC (shown in yellow) and (Sa)EctC (shown in grey) crystal structures. The three residues involved in the binding of the Fe(II) ion are depicted as sticks. A water molecule (blue sphere) in the (Sa)EctC crystal structure occupies the same location as the Fe(II) ion (orange sphere) in the (P)EctC structure. (c) Single amino acid substitution variants of the (P)EctC protein were assayed for their enzyme activity. Enzyme activity measurements were conducted under buffer conditions [20 mM HEPES buffer (pH 8.5), 50 mM NaCl, 0.1 mM (NH4)2Fe(SO4)2] optimized for the wild-type enzyme using 10 µg of protein and 10 mM of the substrate N-γ-ADABA. The enzyme assays were conducted at 30 °C and run for 30 min and the formation of ectoine was monitored by HPLC analysis. The enzyme activity of the mutant (P)EctC proteins is represented relative to that of the wild-type enzyme (set at 100% activity).
activity of 5.4% (Glu-57), 15.1% (Tyr-84), and 12.7% (His-92) (Fig. 4c). Notably, the same conclusion has been reached through mutant analysis of the corresponding putative iron-binding residues present in the (Sa)EctC protein.

**Structural features of the binding site for the N-γ-ADABA substrate.** In the (Pl)EctC::Fe/N-γ-ADABA crystal structure, the substrate for the ectoine synthase, N-γ-ADABA,16,59, is positioned in close proximity to the catalytically important iron atom within the cupin barrel (Fig. 6a) and the iron atom is bound in a fashion similar to that observed in the (Pl)EctC::Fe complex (Fig. 4a). The substrate N-γ-ADABA was added in large excess (40 mM) to the crystallization solution; however, the obtained crystal structure displayed only a partially bound molecule, which resulted in an occupancy of 68% after refinement. The N-γ-ADABA molecule is coordinated within the active site of the (Pl)EctC enzyme through six direct interactions with residues Trp-21, Arg-25, Asn-38, Thr-40, Tyr-52, and Glu-57. Interactions of N-γ-ADABA with the iron atom further stabilize it within the catalytic core of the ectoine synthase (Fig. 6a). In Fig. S3, we provide a numbering scheme for the atoms in the substrate N-γ-ADABA and the reaction product ectoine to aid the understanding of the following descriptions.

N-γ-ADABA is coordinated in the active site via two sets of interactions. First, a direct interaction occurs between the O atom of the acetyl group of N-γ-ADABA (acetamide oxygen) and the iron co-factor with a distance of 2.6 Å. The N5 atom of the substrate interacts with Glu-57 and with Tyr-52, its α-NH2 moiety interacts with Thr-40, and one of the carboxylic O atoms interacts with Asn-38. Both of the latter amino acid residues are part of β-sheet J1V. The carboxylate oxygens of N-γ-ADABA are also coordinated via interactions with a water molecule, which in turn is held in place via an interaction with the side chain of Arg-25 (Fig. 6a). A second set of interactions is observed for the C3 and C4 atoms of N-γ-ADABA, which interact with the side chain of Trp-21 (Fig. 6a). Notably, Trp-21 adopts a dual conformation in the crystal structure of the (Pl)EctC::Fe/N-γ-ADABA complex, in line with the observed partial occupancy of the crystals with the bound substrate. Only in one of these two conformations of Trp-21, its side chain is oriented towards the N-γ-ADABA substrate (52% occupancy) (Fig. 6a). The comparison of the (Pl)EctC::Fe and the (Pl)EctC::Fe/N-γ-ADABA structures therefore suggests that the presence of the N-γ-ADABA molecule within the substrate-binding site of the ectoine synthase induces the reorientation of the side-chain of Trp-21 to provide additional stabilizing contacts to the N-γ-ADABA molecule. The notion that the side chain of Trp-21 is critically involved in the stable positioning of the substrate is supported by data from a site-directed mutagenesis experiment in which we replaced Trp-21 with Ala (Fig. 6a,b). Despite the conspicuous spatial arrangement of Cys-104 side-chain (Fig. 6a) and the evolutionary conservation of this residue, the Cys-104/Ala substitution had in essence no effect on the catalytic activity of the (Pl)EctC enzyme (Fig. 4c). This result also corroborates the functional relevance of the amino acids whose site-directed change...
Figure 6. Crystallographic views into the catalytic core of the *P. lautus* ectoine synthase. (a) The iron- and substrate-binding network within the catalytic core of the (Pl)EctC protein. The side-chain of Trp-21 that adopts two different conformations in the (Pl)EctC::Fe/N-γ-ADABA crystal structure is emphasized in green. The N-γ-ADABA molecule is shown as yellow sticks. The iron is represented as an orange sphere and a water molecule (blue sphere) mediating indirect contacts between the N-γ-ADABA molecule and the side-chain of Arg-25 is highlighted. In addition, the spatial position of the side-chain of a conserved Cys residue (Cys-104) close to the active site, but not directly involved in enzyme catalysis is shown. (b) Details of the iron- and ectoine-bound state of the (Pl)EctC catalytic core are shown. In this structure, the side chain of Trp-21 adopts only a single conformation and provides stabilizing contacts to the ectoine ligand via cation-π interactions. The ectoine molecule is depicted as a yellow stick. (c) Cartoon representation of a structural overlay of the (Pl)EctC::Fe/N-γ-ADABA (gold) and (Pl)EctC::Fe/ectoine (green) crystal structures. The position of the substrate and reaction product within the confines of the cupin barrel is highlighted and the iron co-factor is shown as an orange sphere. In all panels, parts of the secondary structure elements of the (Pl)EctC protein are omitted in order to provide an unobstructed view into the catalytic core of the ectoine synthase.
impaired the enzyme function of the (Pl)EctC protein (Fig. 4c). Unexpectedly, the replacement of Cys-104 of the (Pl)EctC protein by a Ser residue, an amino acid often observed in orphan EctC-type proteins,1,30 in this amino acid position (Fig. 5b) yielded a (Pl)EctC variant that retained only approximately 60% of enzyme activity (Fig. 4c). This drop in enzyme activity is currently not understood in functional terms given the frequent occurrence of a Ser residue at position 104 of ectoine synthases.1,30

In the previously reported structural analysis of the S. alaskensis EctC protein, it was not possible to obtain (Sa)EctC crystals in complex with the natural substrate of the ectoine synthase, N-γ-ADABA. Instead, a chemically not fully defined compound (in all likelihood trapped by the recombinant (Sa)EctC protein either during purification in E. coli, or from the crystallization solution) was present.30 This C-6 molecule was modeled as hexandiol and it was argued by Widderich et al.30 that its spatial position within the cupin barrel of the (Sa)EctC protein could be used as a proxy for the actual substrate of the ectoine synthase, N-γ-ADABA. We now superimposed the (Pl)EctC::Fe/N-γ-ADABA (PDB accession code: 5ONN) and (Sa)EctC/hexandiol (PDB accession code: 5BXX) crystal structures; this yielded a root-mean-square deviation (r.m.s.d.) of 1.4 Å over 105 Cα atoms of the two crystal structures. The overlay showed that the N-γ-ADABA and the presumed hexandiol ligand occupy, indeed, a very similar position within the active site of the two studied ectoine synthases (Fig. S4).

**Structural features of ectoine binding within the (Pl)EctC active site.** We were able to obtain a (Pl)EctC crystal structure that contained both the iron ion and an ectoine molecule (Fig. 6b). The ectoine molecule is coordinated within the active site of the (Pl)EctC enzyme through five direct interactions with the following residues: Ser-23, Asn-38, Tyr-52, Glu-57, and Phe-106 (Fig. 6b). Ectoine exhibits a significantly different conformation to that of the bound N-γ-ADABA molecule. In particular, Tyr-52 and Glu-57 are no longer involved in any H-bonding interactions with N1 of ectoine (derived from the amidic N5 of N-γ-ADABA), but now both residues form H-bonds to N3 of ectoine, which is derived from the α-amino group of N-γ-ADABA. These new H-bonds towards N3 of ectoine also appear to be stronger than those formed originally with N-γ-ADABA, as suggested from their shorter distances (3.2 Å to Tyr-52, and 3.7 Å to Glu-57). This H-bonding interaction is further stabilized by an interaction with Phe-106 (Fig. 6b). In addition, the side-chain of Asn-38 makes a direct contact to the carboxylate of ectoine, and this side chain is in turn held in place through stabilizing interactions with the side-chain of Ser-23 (Fig. 6b). There may also be a direct interaction between the iron and the methyl group of the ectoine molecule (Fig. 6b).

The side chain of Trp-21 adopts a single conformation in the (Pl)EctC::Fe-ectoine structure and thereby provides additional stabilizing contacts to the ectoine ligand (compare Fig. 6a,b).

**A structural comparison of the substrate- and product-bound catalytic core of ectoine synthase.** When the crystal structures of the (Pl)EctC::Fe/N-γ-ADABA and the (Pl)EctC::Fe/ectoine complexes are overlaid, it becomes apparent that the substrate (N-γ-ADABA) and the reaction product (ectoine) occupy almost the same position within the (Pl)EctC active site (Fig. 6c). The (Pl)EctC::Fe/N-γ-ADABA structure reveals an extended conformation of the substrate within the catalytic core (Fig. 6a,c). The distance between the carbonyl C-atom and the α-amino group of the bound N-γ-ADABA molecule (4.1 Å) is far too large to allow a direct enzymatic attack to initiate ring closure. As a result, the N-γ-ADABA molecule needs to bend significantly in order to correctly position the two substituents of the substrate involved in ring closure (Fig. 1a) closely enough to form the intramolecular Schiff-base bond required to generate the enzyme reaction product ectoine. Hence, the observed mode of ectoine binding indicates the requirement of an extensive conformational change of the linear N-γ-ADABA molecule during enzyme catalysis to yield the spatial position of the cyclic ectoine molecule (Fig. 6a–c).

Because the N-γ-ADABA molecule is bound in an extended, and not in a pre-bent (Fig. 1a), conformation, the main chain region of N-γ-ADABA needs to be rearranged to a more bent conformation during the cyclo-condensation reaction. This requires large movements of the carbon atoms (2.0 Å, 1.2 Å, and 1.4 Å of the C-2, C-3 and C-4 of the N-γ-ADABA molecule) to yield the positions of the corresponding C-4, C-5, and C-6 atoms of the resulting ectoine molecule. Moreover, the plane described by the acetamide group of N-γ-ADABA indicates a required rotation of about 80° for superimposition with the plane of the amidinium group of ectoine captured in the (Pl)EctC crystal structure. These changes are associated with an outward spatial extension of the bound N-γ-ADABA molecule, which can only occur with an associated flip-over movement of the side chain of Trp-21 from the (partial) position observed in the (Pl)EctC::Fe/N-γ-ADABA structure (Fig. 6a) to that found in the ectoine-bound structure (Fig. 6b). We note, that the distance between ectoine and the side chain of Trp-21 in the “wrong” conformation (Fig. 6a,b) is only 1.6 Å, a configuration that would likely lead to molecular clashing.

The indole ring of Trp-21 of the (Pl)EctC protein stands out as the only amino acid side chain exhibiting a major conformational change between the substrate- and product-bound complexes (Fig. 6a,b). As noted above, this residue is required for enzyme activity (Fig. 4b). Although this residue appears to be present in a mixture of two conformations in the N-γ-ADABA bound (Pl)EctC structure, we assume that the one deviating from the product-bound state is important for catalysis. In the N-γ-ADABA-bound structure, the indole ring of Trp-21 is orientated towards and above the C3 and C4 atoms of the substrate and seems to act as a piston-like element by pushing the N-γ-ADABA molecule against the bottom of the active site cavity. This assumption is consistent with the presence of a significant conformational strain in the bound N-γ-ADABA molecule, as deduced from large deviations of the bond angles of the C2 and C3 atoms of the diaminobutyrate moiety in the crystal structure from the expected tetrahedral angles (104.5° and 117.5°, respectively, instead of the expected 109.5°). This indicates that Trp-21 might play an important role in exerting conformational constraints on the substrate, which may allow the N-γ-ADABA molecule to assume the extended conformation observed in the crystal structure (Fig. 6a). These structural constraints might also move the acetyl group of the substrate close enough towards the iron co-factor to support tautomeric rearrangements of the acetyl substituent.
Apart from the conformational flip of the side-chain of Trp-21 and the reorientation of the bound substrate with respect to the ectoine product, the active site moieties remain remarkably constant in their locations (Fig. 6a,b). Therefore, we assume that the flip-over of Trp-21 triggers the re-orientation of bound substrate during ring formation (or vice versa), and that this coordinated re-organization of the active site is a major driving force for catalysis. Consistent with its predicted critical role in enzyme function (Fig. 4b), we observed that Trp-21 is completely conserved in an amino acid sequence alignment of 437 EctC-type proteins encoded by ect gene clusters (Fig. 5a). The corresponding Trp residue in 145 orphan EctC-type proteins encoded outside canonical ect gene clusters (see below) is conserved as well, but in some cases its position within the EctC protein chain is shifted by two amino acid residues (Fig. 5b).

Proposal for the ectoine synthase-catalyzed reaction mechanism. By comparing the (Pl)EctC::Fe to the substrate- and product-bound (Pl)EctC structures (Fig. 6a,b), a catalytic cycle for the conversion of the substrate \( N^-\gamma\)-ADABA into the reaction product ectoine can be suggested (Fig. 7). The (Pl)EctC::Fe structure exhibits a tetrahedral coordination of the iron cofactor by the side chains of Glu-57, Tyr-84, His-92 and a water molecule (Fig. 4a). As can be observed in the (Pl)EctC::Fe/N^-\gamma\)-ADABA structure, the water is replaced by the amide carbonyl group of \( N^-\gamma\)-ADABA upon substrate binding (Fig. 6a). This coordination to the Lewis acidic Fe\(^{2+}\) cofactor stabilizes the amide functional group of the substrate in its charge-separated mesomeric form with a negative charge at the carbonyl oxygen and a positive charge at the amide nitrogen that is in turn stabilized by cation-\(\pi\)-interactions with the side chain of Trp-21. The resulting increased reactivity of the amide towards nucleophiles triggers the ring closure by attack of the \( \alpha\)-amino group of \( N^-\gamma\)-ADABA that may proceed with a simultaneous proton transfer from the \( \alpha\)-amino group or from one of the Fe-coordinating amino acids to the amide carbonyl oxygen. The nucleophilicity of the \( \alpha\)-amino group, which is mostly protonated at physiological pH, is increased by the hydrogen bonds to Thr-40 and further to Asn-38 (Fig. 6a,b) thereby allowing a proton transfer during the cyclization of \( N^-\gamma\)-ADABA to form ectoine. In the next step, the hydroxy group is extruded from the substrate to release the product ectoine (Fig. 7). A simultaneous back-transfer of the proton from Asn-38 via Thr-40 to the expelled hydroxy group may restore the initial binding situation at the central iron with one water ligand (Fig. 7); the product ectoine is finally released from the active site to allow for the initiation of a new catalytic cycle. We note however in this context that the expected water ligand of the Fe\(^{2+}\) ion was only confirmed for the substrate-free EctC crystal structure, but not for its ectoine-bound state as this latter crystal structure has a resolution (2.5 Å) that does not allow the positioning of water molecules with confidence.

The residues forming the entire EctC catalytic core are evolutionarily highly conserved. The explosion in the number of available microbial genome sequences allows one to place the salient features of a given protein within an evolutionary context. We have recently conducted an extensive analysis of the phylogeny of the ectoine synthase protein family (Fig. 5) by relying on the Integrated Microbial Genomes & Microbiomes...
reaction of the carbonyl group with the α-ectC genes altogether or contain the architecture of the catalytic core of ectoine synthase (Fig. 6a–c). The high-resolution crystal structures of the full-length EctC enzyme in complex with its iron-cofactor, its substrate, Pl the predicted ectoine-producing microorganisms5. However, there is a sub-group of EctC-type proteins that are not part of ect biosynthetic gene clusters or that can occur in addition to bona fide ectC genes (25% in the dataset examined by Czech et al.). A microbial strain (Pseudomonas syringae pv. syringae) with such an exclusive orphan ectC gene has been physiologically studied, and seems to be able to produced ectoine when surface sterilized leaves of the host plant of this pathogen were provided to the culture55.

Building on the extensive bioinformatic dataset reported by Czech et al.5 and on the salient features of the crystal structures of the (P)EctC protein that we present here (Figs 4 and 6), we now can focus on those ten residues involved in binding iron, the substrate or the product (Fig. 5a,b) in an evolutionary context by inspecting alignments of EctC proteins encoded within complete ectoine biosynthetic gene clusters (437 representatives) and those encoded by orphan ectC-type genes (145 representatives). We used the amino acid sequence of the (P)EctC protein as query for this search. The degree of amino acid sequence identity of EctC proteins encoded by ect biosynthetic gene clusters ranged between 90% (for Pseudomonas glaucescens) and 49% (for Streptomyces glaucescens) when 437 amino acid sequences of bona fide EctC proteins were aligned and compared with the amino acid sequence of the crystallized (P)EctC protein. Hence, bona fide EctC-type proteins are evolutionarily rather well conserved (Fig. 5a). When the (P)EctC amino acid sequence was compared with those of 145 orphan EctC-type proteins (Fig. 5b), the degree of amino acid sequence identity decreased and ranged between 42% (Burkholderia cepacia) and 37% (Roseobacter litoralis).

From an alignment of the 582 EctC-type proteins retrieved through the BLAST search of the IMG/M database6 we observed that 20 residues were completely conserved. After we excluded the EctC orphan sequences and only compared the 437 EctC-type proteins encoded by ect gene clusters, the number of completely conserved amino acid residues increased to 26 (Fig. 5a,b). Based upon the (P)EctC crystal structures (Fig. 6a,b), these conserved residues can be correlated to the following functions: three residues are involved in metal binding (Glu-57, Tyr-84, His-92), six residues are involved in coordinating the substrate N-γ-ADABA within the active site (Trp-21, Arg-25, Asn-38, Thr-40, Tyr-52, Gla-57), and five residues coordinate the reaction product ectoine (Ser-23, Asn-38, Tyr-52, Gla-57, Phe-106). The remaining conserved residues (Fig. 5a) might play either structural, or yet not recognized mechanistic roles. Notably, in the carboxy-terminal segment of EctC proteins, there are nine strictly conserved residues but only one of them (Phe-106) is involved in binding a ligand (ectoine) of the ectoine synthase (Fig. 5a). When one views this carboxy-terminal region in the overall EctC structure, it becomes apparent that it forms a lid over the entry to the cupin barrel (Fig. 8a,b). The strong conservation of the participating residues suggests a functionally important role of this presumed lid region. When one removes the 27 amino acid segment starting from the (P)EctC::Fe/γ-ADABA crystal structure, the cavity of the ectoine synthase active site would be large enough to accommodate the seven-membered ring of the non-natural homoectoine molecule. The H. elongata enzyme is also somewhat promiscuous in its biosynthetic activity as it can form the synthetic compatible solute 5-amino-3,4-dihydro-2H-pyrrole-2-carboxylate (ADPC) through the cyclic condensation of glutamine in a side reaction70. Because the EctC-mediated biotransformation of N-γ-ADABA into ectoine is practically an irreversible reaction, one wonders how it was possible in this study to obtain (P)EctC::Fe crystals in complex with N-γ-ADABA. However, crystallization with this substrate was carried out at pH 4.2, conditions under which the (P)EctC protein is practically enzymatically inactive (Fig. 2b).

By using an ectoine synthase from the thermo-tolerant bacterium P. lautus58, we were able to obtain high-resolution crystal structures of the full-length EctC enzyme in complex with its iron-cofactor, its substrate, and its product (Figs 3 and 6). The information obtained from this structural analysis illuminates for the first time the architecture of the catalytic core of ectoine synthase (Fig. 6a–c). The (P)EctC crystal structures presented here
likely represent catalytic states of the enzyme prior and subsequent to catalysis (Figs 3a, b, and 6), thereby providing the foundation for a proposal for the EctC-catalyzed cyclo-condensation reaction (Figs 1a and 7). The ten residues involved in binding of the iron co-factor, the $N$-$\gamma$-ADABA substrate, and the reaction product ectoine are evolutionarily highly conserved among a large group of EctC-type proteins $^{5,30,60}$ (Fig. 5a, b). Our structural and functional studies thus render the (P)EctC protein as a point of reference for the extended ectoine synthase family as a whole $^3$ and thereby provide a blueprint for further biochemical and physiological studies.

A previous bioinformatic analysis has indicated that ectoine synthases form a distinct branch within the cupin protein super-family $^{63,67}$. The three-dimensional structure of the ectoine synthase $^{40}$ (and this study) follows the basic design principles of proteins belonging to this super-family $^{63-87}$, and matches closely in its overall fold those of the KdgF and the DddK degradative enzymes. These two cupin-type proteins carry out enzymatic reactions that are different from each other $^{76,77}$ and from that catalyzed by the ectoine synthase (Figs 1a and 7). The (P)EctC monomer (Fig. 3a) and is part of an intricate network of interactions that position the $N$-$\gamma$-ADABA substrate, and the reaction product ectoine into the lumen of the cupin barrel (Fig. 6a, b). The (P)EctC protein is a head-to-tail dimer where backbone interactions between two $\beta$-sheets mediate the contacts between the monomers in the dimer assembly (Fig. 3b–d). A dimeric assembly has previously also been found through biochemical approaches for ectoine synthases from the bacteria $H. elongata$ $^{38}$ and $S. alaskensis$ $^{60,67}$ and from the archaean $N. maritimus$ $^{29}$, suggesting that dimer-formation is probably a general feature of ectoine synthases.

The catalytically critical Fe$^{2+}$ cation (Fig. 4a) is placed somewhat off-center within the central cavity of the (P)EctC monomer (Fig. 3a) and is part of an intricate network of interactions that position the $N$-$\gamma$-ADABA substrate within the active site of the enzyme close to the metal co-factor (Figs 6a and S4). Variations in the metal-binding motifs of cupins occur frequently $^{63,64,66,87}$, but the one we detected in the ectoine synthase (Figs 4a and 5a, b) is, to the best of our knowledge, novel. Cupins are often promiscuous with respect to the metal used for catalytic activity $^{76,77,82-84}$. For instance, crystals of the DddK DMSP lyase were found in one form to harbor Ni$^{2+}$ while a second crystal form harbored either Fe$^{2+}$ or Zn$^{2+}$ with different levels of occupancy of the same metal binding site $^{77}$. Likewise, although Ni$^{2+}$ was found in the KdgF crystal structure, reconstitution of the metal-depleted recombinant enzyme with various divalent metals identified Co$^{2+}$ as the most effective catalyst $^{77}$. We obtained crystals of the (P)EctC protein only after adding substantial Fe$^{2+}$ concentrations to the protein solution prior to crystallization and the previously reported crystal structure of the (Sa)EctC enzyme lacks a metal altogether $^{40}$. Despite this difference in metal content, the spatial position of the side chains of the three metal-coordinating residues in the active site of the (P)EctC and (Sa)EctC enzymes are super-imposable in their crystal structures.

Figure 8. Structural features of the proposed (P)EctC lid region. (a) Ribbon and surface representation of the (P)EctC::Fe/N-$\gamma$-ADABA crystal structure highlighting (in red) the lid-region formed by the carboxy-terminal segment of the ectoine synthase. The spatial position of the $N$-$\gamma$-ADABA substrate (shown in yellow sticks) and that of the iron atom (shown as an orange sphere) are highlighted within the EctC active site. (b) Surface representation of the (P)EctC::Fe/N-$\gamma$-ADABA crystal structure in which the lid region is highlighted in red. Relative to the representation of the (P)EctC crystal structure shown in (a), the crystal structure shown in (b) is slightly rotated in order to provide an expanded view onto the lid region. (c) View into the catalytic core of the (P)EctC::Fe/N-$\gamma$-ADABA crystal structure after the in silico removal of the lid region (amino acids 103 to 130). (d) Cross-section through the catalytic core of the (P)EctC::Fe/N-$\gamma$-ADABA crystal structure revealing a plausible entry tunnel for the $N$-$\gamma$-ADABA substrate (in yellow sticks) and its spatial positions to the catalytically critical iron atom (orange sphere) and a water molecule (blue sphere) that interacts with the $N$-$\gamma$-ADABA molecule.
A metal reconstitution experiment with the (Sa)EctC protein has previously revealed that the ectoine synthase can function with various divalent metals (Fe^{2+}, Zn^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Mn^{2+}), albeit with different levels of catalytic activity. Fe^{2+} served as the best-performing co-factor for the (Sa)EctC enzyme. Notwithstanding the apparent constraints in interpreting the specific type(s) of metal present in the enzymatically active (Sa)EctC and (Pl)EctC proteins, there can be no doubt that the ectoine synthase is metal-dependent. This conclusion is supported by the results of site-directed mutagenesis experiments targeting these residues in both the (Sa)EctC and (Pl)EctC proteins (Fig. 4c), and the complete conservation of the three metal-binding residues among the inspected 582 members of the extended EctC protein family5 (Fig. 5).

The ligand-binding site for ectoine in the (Pl)EctC enzyme differs significantly from those present in ectoine/5-hydroxyectoine-specific extracellular substrate-binding proteins operating in conjunction with either high-affinity ABC- (Ehu)5, or TRAP- (Tea and Ueh) type transporters.66,67 A complex network of interactions with the ectoine molecule is observed in the corresponding EhuB, TeaA, and UehA substrate-binding proteins. It involves several aromatic side-chains, which contribute strongly to high-affinity binding of ectoine with values for the dissociation constant (K_d) of about 1.6 μM, 0.2 μM, and 1.4 μM, respectively.65,66,68,69 In contrast, in the (Pl)EctC::Fe-ectoine complex, only the aromatic side chain of Trp-21 contributes to the stabilization of the ectoine molecule in the active site and the carboxylate of ectoine additionally interacts with the side chain of Asn-38 (Fig. 6b).

A key contributor for the co-ordination of the ectoine molecule within the substrate-binding site of the EhuB, TeaA, and UehA ligand-binding proteins is the interaction of the carboxylate of ectoine with the side chain of an Arg residue. Replacement of this Arg residue by Ala abrogates high-affinity ectoine binding53,65. While this interaction plays a central role in capturing ectoine stably by the extracellular substrate-binding proteins, it is not observed in the active site of the (Pl)EctC cytoplasmic enzyme (Fig. 6b). For ecophysiological reasons, a high substrate affinity is needed to scavenge ectoine through transport processes from scarce environmental sources.88–90 For its use either as an osmoprotectant86,91 or as a nutrient87,92. On the other hand, a low affinity of the (Pl)EctC enzyme for ectoine will be required to release it from the catalytic core once it has been formed by the cyclization of N-γ-ADABA (Figs 1a and 7). Hence, the differences observed in the architecture of the ectoine binding sites present in the EhuB, TeaA, and UehA substrate-binding proteins on one hand55–57 and the EctC enzyme on the other hand (Fig. 6a,b), most likely reflect functionally imposed constraints on protein structure.

In the previously reported crystal structures of the ectoine synthase from the cold-adapted marine microorganism S. alaskensis69, the carboxy-terminal segment of the (Sa)EctC protein appeared to be highly flexible. However, we have no indications to this effect through the crystallographic snapshots of the (Pl)EctC proteins derived from the thermo-tolerant bacterium P. latus.69 This region (27 amino acids of the 130-amino acids comprising (Pl)EctC protein) contains a surprisingly high number of strictly conserved residues among a large group of bona fide ectoine synthases (437 representatives) and orphan (145) EctC-type proteins (Fig. 5a,b). The crystal structures of the (Sa)EctC and (Pl)EctC proteins suggest that these conserved carboxy-terminal segments form a lid over the entry to the cupin barrel (Fig. 8a,b). The catalytic core of the ectoine synthase is positioned at a bottom of a deep tunnel (Fig. 8d), and one can therefore readily envision that movement of the lid provides access of the N-γ-ADABA substrate to the active site (Figs 3a and 6) and a subsequent exit route for the reaction product ectoine. It therefore seems plausible that the lid region shields the active site of the ectoine synthase from the external solvent and thereby provides a privileged space for the elimination of a water molecule during the cyclization step (Fig. 8d). Moreover, opening and closing of the lid domain may also serve as primers to study the substantial group of orphan EctC-type proteins5, both biochemically and structurally in the future to reveal their physiological function.
**Methods**

**Chemicals.** Ectoine was a kind gift from the bitop AG (Witten, Germany). Anhydrotetracycline hydrochloride (AHT), desithibiotin, and Strep-Tactin Superflow chromatography material for the purification of Strep-tag II labeled proteins were purchased from IBA GmbH (Göttingen, Germany). The commercially unavailable substrate for the ectoine synthase, N-α-acetyl-L-2,4-diaminobutyric acid (N-α-ADABA), was synthesized through alkaline hydrolysis of ectoine. It was purified from the byproduct N-α-acetyl-L-2,4-diaminobutyric acid (N-α-ADABA) by repeated chromatography on a silica gel column (Merck silica gel 60) using a gradient of ethanol/25% ammonia/water 50:1.2–10:1.2 as eluent. The identity and purity of the isolated N-α-ADABA was monitored by thin-layer chromatography and nuclear magnetic resonance spectroscopy (1H-NMR and 13C-NMR) on a Bruker AVIII-400 or DRX-500 NMR spectrometer as described. All chemicals used to synthesize and purify N-α-ADABA were purchased from either Sigma Aldrich (Steinheim, Germany) or Acros (Geel, Belgium). Numbering of the N-α-ADABA and ectoine follows the IUPAC rules, applying replacement nomenclature for N-α-ADABA as ”2-amino-5-aza-6-oxoheptanoic acid” (see Fig. S3).

**Recombinant DNA procedures and construction of plasmids.** The DNA sequences of the ectC genes from the genome sequence of *P. lautus* Y4.12MC10 (accession number: NC_013406) were used as the template for the synthesis of a codon-optimized version of ectC (LifeTechnologies; Darmstadt, Germany) for its heterologous expression in *E. coli*. The DNA sequence of this synthetic ectC gene has been deposited in the NCBI database under accession number KR002038. To allow the overproduction of the recombinant (PI)ECTC protein in *E. coli* and its purification via affinity chromatography, an overexpression plasmid was constructed, which directs the synthesis of a (PI)ECTC-Strep-tag II recombinant protein. For this purpose, the ectC-containing DNA fragment was retrieved from the plasmid provided by the supplier of the synthetic construct (LifeTechnologies), and inserted into the expression vector PAG-IBA3 (IBA GmbH, Göttingen, Germany). The ectC coding sequence was cloned into plasmid pAG-IBA3 in such a way that the resulting EctC protein is fused at its carboxy-terminus to a Strep-tag II affinity peptide (SA-WSHHPQFEK), thereby allowing the purification of the (PI)ECTC-Strep-tag II protein by affinity chromatography on a streptavidin affinity matrix (IBA GmbH, Göttingen, Germany). The transcription of the ectC gene in the resulting plasmid pWN14 ([*P. lautus* ectC]; constructed by N. Widderich) is mediated by the tet promoter present on the backbone of the expression vector pAG-IBA3 and controlled through the TetR repressor whose DNA-binding activity can be abrogated by adding the synthetic inducer AHT to the growth medium.

Variants of the codon-optimized ectC gene from *P. lautus* present on plasmid pWN14 were prepared by site-directed mutagenesis using the Q5 Site-Directed Mutagenesis Kit (New England BioLabs GmbH, Frankfurt a. M., Germany) with custom synthesized DNA primers purchased from Microsynth AG (Lindau, Germany). The DNA sequence of the entire coding region of each mutant ectC gene was determined by Eurofins MWG Ebersberg, Germany) to ensure the presence of the desired mutation and the absence of unwanted alterations. The following variants of the *P. lautus* ectC gene were constructed: pLC55 (Glu-57/Ala; GAA/GCA), pLC56 (Glu-99/Val; GGC/GGT), pLC57 (His-92/Ala; CAG/GCA), pLC58 (Cys-104/Ala; TGC/GCA), pLC59 (Cys-104/Ser; TGC/TCT), and pLC67 (Trp-21/Ala; TGG/GCG).

**Bacterial strains, media, and growth conditions.** The *E. coli* strain TOP10 (Invitrogen, Carlsbad, CA, USA) was used for the propagation of plasmids carrying ectC genes. Cultures of the plasmid-carrying *E. coli* strain were grown at 37 °C in Luria-Bertani (LB) liquid medium containing ampicillin (100 μg ml⁻¹). Heterologous overproduction of the plasmid-encoded *P. lautus* EctC Strep-tag II protein ([PI]ECTC) was carried out in the *E. coli* B strain BL21 in modified minimal medium A (MMA) containing 0.5% (w/v) glucose as the carbon source and 0.5% (w/v) casamino acids, 1 mM MgSO₄, and 3 mM thiamine as supplements. Mutant derivatives of the (PI)ECTC protein were overproduced and purified as described below for the corresponding wild-type proteins.

**Overproduction, purification and analysis of the quaternary assembly of EctC proteins.** Cells of the *E. coli* B strain BL21 harboring plasmid pWN14 (ectC⁺) were inoculated into modified MMA (1 L medium in a 2-L Erlenmeyer flask) to an OD₅₇₈ of 0.05 from an overnight culture. The cells were grown on an aerial shaker (set to 180 rpm) at 37 °C until the cultures reached an OD₅₇₈ of 0.5. At this time point the synthetic inducer AHT for the Tet repressor was added to a final concentration of 0.2 mg ml⁻¹ to trigger enhanced transcriptional activity of the tet promoter and thereby boost the expression of the plasmid-encoded ectC gene. After 2 h of further growth of the culture at 37 °C, the *E. coli* B strain BL21 (pWN14) cells were harvested by centrifugation (4,000 × g) and disrupted by passing them through a French Pressure cell (at 1 000 psi); a cleared cell lysate was prepared from these disrupted cells by ultracentrifugation (100 000 × g) at 4 °C for 45 min as described. Cleared cell extracts of the (PI)ECTC-Strep-tag II overproducing cultures were used to purify the recombinant proteins by affinity chromatography on Strep-Tactin affinity resin as detailed previously. The concentration of the (PI)ECTC protein in the individual fractions eluted from the Strep-Tactin Superflow affinity column was measured with the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific, Schwerte, Germany). The purity and apparent molecular mass of the (PI)ECTC protein was assessed by SDS–PAGE (15% polyacrylamide), and the PageRuler Prestained Protein Ladder (Thermo Fisher Scientific) was used as a reference to assess the electrophoretic mobility of the (PI)ECTC-Strep-tag II protein. The recombinant (PI)ECTC protein was concentrated to approximately 10 mg ml⁻¹ with Vivaspin 6 columns (Sartorius Stedim Biotech, Göttingen, Germany) with a 10-kDa molecular-weight cutoff value prior to crystallization trials.

To analyze the quaternary assembly of the (PI)ECTC protein, we used high-performance liquid chromatography coupled to multi-angle light scattering detection (HPLC–MALS). For these experiments, an Agilent Technologies system connected to a triple-angle light scattering detector (miniDAWN TREOS, Wyatt Technology Europe
μ the denatured proteins and the supernatant was subsequently analyzed for the formation of ectoine by HPLC monitored using an Infinity 1260 Diode Array Detector (DAD) (Agilent, Waldbronn, Germany). The amounts of the EctC-catalyzed enzyme reaction product ectoine in individual samples was integrated into an Agilent 1260 Infinity LC system (Agilent). The ectoine content of the samples was quantified using the OpenLAB software suite (Agilent) using commercially available ectoine (bitop AG, Witten, Germany) as the standard. For the ectoine synthase enzyme activity assays, three independently isolated (Pl)EctC protein preparations were used and each data-point from the individual protein preparations was assayed twice.

When mutant (Pl)EctC proteins were assayed for their enzyme activity, the buffer conditions optimized for the wild-type enzyme were used, but 10 μg of purified (Pl)EctC protein were employed and the reaction time was extended to 30 min. During the initial screening for the pH optimum of the (Pl)EctC enzyme, a buffer mixture of MES (pH 5.5), PIPES (pH 6.5), TES (pH 7.5), CHES (pH 8), HEPES (pH 8.5), and CAPS (pH 10) (20 mM each) was used. The pH values of these buffer solutions were adjusted with 38% HCl or 5 M NaOH at a temperature of (30°C) that was also used for the ectoine synthase enzyme reaction.

Data processing and structure determination. Data sets were collected from a single crystal of either (Pl)EctC::Fe, (Pl)EctC::Fe/N-γ-ADABA, and EctC::Fe/ectoine on beamline P13 at DESY (EMBL, Hamburg, Germany) and/or ID29 at the ESRF, Grenoble, France at 100 K. These data sets were processed using the XDS package27 and scaled with xscale38. Initial phases were obtained by molecular replacement using the program Phaser74 with the crystal structure of the S. alaskensis EctC protein (PDB entry 5BXX) (without taking its side chains into account) as a template80. Model building and refinement were performed using COOT99 and REFMAC5100. Data refinement statistics and model content are summarized in Table 1. The atomic coordinates and structure factors have been deposited in the Worldwide Protein Data Bank (PDB) (https://www.wwpdb.org/) under the following accession codes: for the (Pl)EctC::Fe complex, 5ONM for the (Pl)EctC::Fe/N-γ-ADABA complex, and 5ONO for (Pl)EctC::Fe/ectoine complex.

Figure preparation of crystal structures. Figures of the crystal structures of the (Pl)EctC protein were prepared using the PyMol software suite (www.pymol.org)103.

Database searches and phylogenetic analysis of EctC-type proteins. The amino acid sequence of the P. lautus EctC protein (accession number: YP_003245677) was used as the template for BLAST searches102 (of all finished sequences of the microbial database of the US Department of Energy Joint Genome Institute (http://jgi.doe.gov/))91. EctC-type amino acid sequences8 were compared using the MAFFT multiple amino acid sequence alignment server (http://mafft.cbrc.jp/alignment/server)103. This data set was then used to construct a rooted phylogenetic tree of EctC-type sequences74 by employing the iTOL software suit (http://itol.embl.de/). The dimer interface of the (Pl)EctC protein was analyzed using PISA74 (http://www.ebi.ac.uk/pdbe/pisa/). Structural
homologs of (P)EccC were searched using the DALI-web server (http://ekhidna.biocenter.helsinki.fi/dali_server/ start) using the (P)EccC::Fe crystal structure as the search query.

**Data Availability**

All data generated or analyzed during this study are included in this published article (and in its accompanying Supplementary Information). The atomic coordinates and structural factors for the crystal structures of the (P)EccC protein determined in this study have been deposited in the Protein Data Bank with accession codes 5ONM for the (P)EccC::Fe complex, 5ONF for the (P)EccC::Fe/N-γ-ADABA complex, and 5ONO for the (P)EccC::Fe/ectoine complex.

**References**

1. Galinski, E. A., Pfeiffer, H. P. & Trüper, H. G. 1,4,5,6-Tetrahydro-2-methyl-4-pyrimidinecarboxylic acid. A novel cyclic amino acid from halophilic phototrophic bacteria of the genus *Ectothiorhodospira*. *Eur J Biochem* 149, 135–139 (1985).
2. Inbar, L. & Lapidot, A. The structure and biosynthesis of new tetrahydrodipyrimidine derivatives in actinomycins D producer *Streptomyces parvulus*. Use of 13C- and 15N-labeled L-glutamate and 13C and 15N NMR spectroscopy. *J Biol Chem* 263, 16014–16022 (1988).
3. Kempf, B. & Bremer, E. Uptake and synthesis of compatible solutes as microbial stress responses to high osmolality environments. *Arch Microbiol* 170, 319–330 (1998).
4. Pastor, J. M. et al. Ectoecins in cell stress protection: uses and biotechnological production. *Biotechnol Adv* 28, 782–801, https://doi.org/10.1016/j.biotechadv.2010.06.005 (2010).
5. Czech, L. et al. Role of the extremeolysite ectoine and hydroxyectoine as stress protectants and nutrients: genetics, phylogenomics, biochemistry, and structural analysis. *J Biol (Basel)* 9, 177, https://doi.org/10.3390/jbi9040177 (2018).
6. Yancey, P. H. Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *J Exp Biol* 208, 2819–2830, https://doi.org/10.1242/jeb.01730 (2005).
7. Roessler, M. & Muller, V. Osmoadaptation in bacteria and archaea: common principles and differences. *Env Microbiol Rep* 3, 743–754 (2001).
8. Wood, J. M. et al. Osmosensing and osmoregulatory compatible solute accumulation by bacteria. *Comp Biochem Physiol A Mol Integr Physiol* 130, 467–461 (2001).
9. Street, T. O., Bolen, D. W. & Rose, G. D. A molecular mechanism for osmolyte-induced protein stability. *Proc Natl Acad Sci USA* 103, 13997–14002, 0606236103 (2006).
10. Bolen, D. W. & Baskakov, I. V. The osmophobic effect: natural selection of a thermodynamic force in protein folding. *J Mol Biol* 310, 955–963, https://doi.org/10.1006/jmbi.2001.4819 (2001).
11. Zaccaci, G. et al. Neutrons describe ectoine effects on water H-bonding and hydration around a soluble protein and a cell membrane. *Sci Rep* 6, 31434, https://doi.org/10.1038/srep31434 (2016).
12. Bremer, E. & Krammer, R. In *Bacterial Stress Responses* (eds Storz, G. & Hengge-Aronis, R.) 79–97 (ASM Press, 2000).
13. Wood, J. M. *Bacterial osmoregulation*: a paradigm for the study of cellular homeostasis. *Annu Rev Microbiol* 65, 215–238, https://doi.org/10.1146/annurev-micro-090110-102815 (2011).
14. Cayley, S., Lewis, B. A. & Record, M. T. Jr. Origins of the osmoprotective properties of betaine and proline in *Escherichia coli* K-12. *J Biol Chem* 268, 1586–1595 (1992).
15. van den Berg, J., Boersma, A. J. & Poonman, L. Microorganisms maintain crowding homeostasis. *Nat Rev Microbiol* 15, 309–318, https://doi.org/10.1038/nrmicro.2017.17 (2017).
16. Boch, J., Kempf, B. & Bremer, E. Osmoregulation in *Bacillus subtilis*: synthesis of the osmoprotectant glycine betaine from exogenously provided choline. *J Bacteriol* 176, 5364–5371 (1994).
17. Street, T. O., Krokenberg, K. A., Rosgen, J., Bolen, D. W. & Agard, D. A. Osmolyte-induced conformational changes in the Hsp90 molecular chaperone. *Protein Sci* 19, 57–65, https://doi.org/10.1002/pro.2028 (2010).
18. Roychoudhury, A., Bicker, A., Haussinger, D. & Oesterhelt, F. Membrane protein stability depends on the concentration of compatible solutes - a single molecule force spectroscopic study. *Biol Chem* 394, 1465–1474, https://doi.org/10.1515/hzsz-2013-0173 (2013).
19. Bourouf, S. et al. Glycine betaine-assisted protein folding in a *lycA* mutant of *Escherichia coli*. *J Biol Chem* 275, 1050–1056 (2000).
20. Ignatova, Z. & Gierasch, L. M. Inhibition of protein aggregation in vitro and in vivo by a natural osmoprotectant. *Proc Natl Acad Sci USA* 103, 13357–13361, 0603772103 (2006).
21. Tanne, C., Golovina, E. A., Hoekstra, F. A., Meffert, A. & Galinski, E. A. Glass-forming property of hydroxyectoine is the cause of its superior function as a dessication protectant. *Front Microbiol* 5, 150 (2014).
22. Manzanera, M. et al. Hydroxyectoine is superior to trehalose for anhydrobiotic engineering of *Pseudomonas putida* KT2440. *Appl Environ Microbiol* 68, 4328–4333 (2002).
23. Lippert, K. & Galinski, E. A. Enzyme stabilization by ectoine-type compatible solutes: protection against heating, freezing and drying. *Appl Microbiol Biotechnol* 37, 61–65 (1992).
24. Stadtmüller, S. S., Gorensek-Benitez, A. H., Guseman, A. J. & Pielak, G. J. Osmotic shock induced protein destabilization in living cells and its reversal by glycine betaine. *J Mol Biol* 429, 1155–1161, https://doi.org/10.1016/j.jmb.2017.03.001 (2017).
25. Chattopadhyay, M. K. et al. The chemical chaperone proline relieves the thermosensitivity of a *dnaK* deletion mutant at 42 degrees C. *J Bacteriol* 186, 8149–8152, https://doi.org/10.1128/JB.186.23.8149-8152.004 (2004).
26. Diamant, S., Eliahu, N., Rosenthal, D. & Goloubinoff, P. Chemical chaperones regulate molecular chaperones in *vivo* and in cells under combined salt and heat stresses. *J Biol Chem* 276, 39586–39591, https://doi.org/10.1074/jbc.M103081200 (2001).
27. Harding, T., Brown, M. W., Simpson, A. G. & Roger, A. J. Osmoadaptive strategy and its molecular signature in obligately halophilic heterotrophic protists. *Genome Biol Evol* 8, 2241–2258, https://doi.org/10.1093/gbe/evw152 (2016).
28. Weinsch, L. et al. Identification of osmoadaptive strategies in the halophile, heterotrophic ciliate *Schmidneiderocthus salinarum*. *PLoS Biol* 16, e2003892, https://doi.org/10.1371/journal.pbio.2003892 (2018).
29. Czech, L. & Bremer, E. With a pinch of extra salt - Did predatory protists steal genes from their food? *PLoS Biol* 16, e2003163, https://doi.org/10.1371/journal.pbio.2003163 (2018).
30. Widrich, N. et al. Strangers in the archael world: omostress-responsive biosynthesis of ectoine and hydroxyectoine by the marine thaumarchaeote *Nitrosopumilus maritimus*. *Env Microbiol* 18, 1227–1248 (2016).
31. Widrich, N. et al. Biochemical properties of ectoine hydroxyoxalates from extremophiles and their wider taxonomic distribution among microorganisms. *PLoS One* 9, e93809, https://doi.org/10.1371/journal.pone.0093809 (2014).
32. Louis, P. & Galinski, E. A. Characterization of genes for the biosynthesis of the compatible solute ectoine from *Marinococcus halophilus* and osmoregulated expression in *Escherichia coli*. *Microbiology* 143, 81–99 (1997).
33. Bursy, J., Pierik, A. J., Pica, N. & Bremer, E. Osmotically induced synthesis of the compatible solute hydroxyectoine is mediated by an evolutionarily conserved ectoine hydrolase. *J Biol Chem* 282, 31147–31155, https://doi.org/10.1074/jbc.M704023200 (2007).
69. Ting, L.

67. Dunwell, J. M., Purvis, A. & Khuri, S. Cupins: the most functionally diverse protein superfamily?

66. Dunwell, J. M., Culham, A., Carter, C. E., Sosa-Aguirre, C. R. & Goodenough, P. W. Evolution of functional diversity in the cupin superfamily.

63. Agarwal, G., Rajavel, M., Gopal, B. & Srinivasan, N. Structure-based phylogeny as a diagnostic for functional characterization of the cupin superfamily.

61. Moritz, K. D., Amendt, B., Witt, E. M. H. J. & Galinski, E. A. The hydroxyectoine gene cluster of the non-halophilic acidophile Methylomicrobium alcaliphilum Ac106.

52. Czech, L., Stöveken, N. & Bremer, E. EctD-mediated biotransformation of the chemical chaperone ectoine into hydroxyectoine and 5-hydroxy-5-methyl-1,3-dioxolane.

55. Perez-Garcia, F., Ziert, C., Risse, J. M. & Wendisch, V. F. Improved fermentative production of the compatible solute ectoine by Pseudomonas putida F10.

56. Czech, L., Poehl, S., Hub, P., Stoeveken, N. & Bremer, E. Tinkering with osmotically controlled transcription allows enhanced production and excretion of the compatible solute adpc.

50. Becker, J. Ectoine metabolism and structure of the prokaryotic cell. Microb Cell Fact 12, 110, https://doi.org/10.1186/s12934-015-0051-0 (2016).

45. Kunte, H. J., Lentzen, G. & Galinski, E. Industrial production of the cell protectant ectoine: protection, mechanisms, processes, and products. Carbohydr Polym 94, 72–86, https://doi.org/10.1016/j.carbpol.2012.05.005 (2013).

40. Bursy, J. Systems metabolic engineering of Corynebacterium glutamicum for production of the chemical chaperone ectoine. Biotechnol J 12, e10417, https://doi.org/10.1002/biot.201300417 (2017).

36. Ono, H. Evolution of functional diversity in the cupin superfamily.

25. Czech, L., Stöveken, N. & Bremer, E. EctD-mediated biotransformation of the chemical chaperone ectoine into hydroxyectoine and 5-hydroxy-5-methyl-1,3-dioxolane.

26. Islam, M. S., Leissing, T. M., Chowdhury, R., Hopkinson, R. J. & Schofield, C. J. 2-Oxoglutarate-dependent oxygenases.

20. Dunwell, J. M., Purvis, A. & Khuri, S. Cupins: the most functionally diverse protein superfamily?

14. Garcia-Estepea, R. et al. The cupin7 gene, which is involved in the synthesis of the compatible solute xyloside, is essential for thermoprotection of the halophilic bacterium Chromohalobacter salexigens. J Bacteriol 188, 3774–3784, https://doi.org/10.1128/JB.00136-06 (2006).

10. Kol, S. et al. Molecular mechanisms of the salt stress response in Streptomyces coelicolor. Appl Environ Microbiol 76, 2574–2581, https://doi.org/10.1128/AEM.01992-10 (2010).

4. Kuhlmann, A. U., Bursy, J., Gimpel, S., Hoffmann, T. & Bremer, E. Synthesis of the compatible solute ectoine in Chromohalobacter salexigens.

3. Czech, L., Stöveken, N. & Bremer, E. EctD-mediated biotransformation of the chemical chaperone ectoine into hydroxyectoine and 5-hydroxy-5-methyl-1,3-dioxolane.

2. Czech, L., Stöveken, N. & Bremer, E. EctD-mediated biotransformation of the chemical chaperone ectoine into hydroxyectoine and 5-hydroxy-5-methyl-1,3-dioxolane.

1. Czech, L., Stöveken, N. & Bremer, E. EctD-mediated biotransformation of the chemical chaperone ectoine into hydroxyectoine and 5-hydroxy-5-methyl-1,3-dioxolane.

0. Czech, L., Stöveken, N. & Bremer, E. EctD-mediated biotransformation of the chemical chaperone ectoine into hydroxyectoine and 5-hydroxy-5-methyl-1,3-dioxolane.
71. Reshetnikov, A. S., Khmelenina, Y. N., Mustakhimov, I. I. & Trotsenko, Y. A. Genes and enzymes of ectoine biosynthesis in halotolerant marine methanotrophs. *Methods Enzymol.* **495**, 15–30, https://doi.org/10.1016/B978-0-12-386905-0.00002-4 (2011).
72. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132, https://doi.org/10.1107/S0907444904019185 (2004).
73. Kobus, S., Widdrich, N., Hoepner, A., Bremer, E. & Smits, S. H. J. Overproduction, crystallization and X-ray diffraction analysis of ectoine synthase from the cold-adapted marine bacterium *Sphingopyxis alakensis*. *Acta Cryst.* F**71**, 1027–1032 (2015).
74. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797, https://doi.org/10.1016/j.jmb.2007.05.022 (2007).
75. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545–549, https://doi.org/10.1093/nar/gkq366 (2010).
76. Hobbs, J. K. et al. EdgF: the missing link in the microbial metabolism of uronate sugars from pectin and alginic acid. *Proc Natl Acad Sci USA* **113**, 6184–6189, https://doi.org/10.1073/pnas.1524214113 (2016).
77. Schnickner, N. J., De Silva, S. M., Todd, J. D. & Dey, M. Structural and biochemical insights into dimethylsulfoxoniopropionate cleavage by cofactor-bound DddK from the prolific marine bacterium *Pelagibacter*. *Biochemistry* **56**, 2873–2885, https://doi.org/10.1021/acs.biochem.6b00997 (2017).
78. Brov, S. et al. Abiotic stress protection by ecologically abundant dimethylsulfoxoniopropionate and its natural and synthetic derivatives: insights from *Bacillus subtilis*. *Environ Microbiol* **7**, 2362–2378, https://doi.org/10.1111/1462-2920.12698 (2015).
79. Widdrich, N., Bremer, E. & Smits, S. H. J. The ectoine hydroylase: a non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase. *Encyclopedia of Inorganic and Bioinorganic Chemistry*, Online (2016).
80. Kurz, M., Burch, A. Y., Seip, B., Lindow, S. E. & Gross, H. Genome-driven investigation of compatible solute biosynthesis pathways of *Pseudomonas syringae* pv. syringae and their contribution to water stress tolerance. *Appl Environ Microbiol* **76**, 5452–5462, https://doi.org/10.1128/AEM.00686-10 (2010).
81. Nordberg, H. et al. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Res* **42**, D26–D31, https://doi.org/10.1093/nar/gkt1069 (2013).
82. Silvennoinen, L., Sandalova, T. & Schneider, G. The polypeptidyl cyclase RemC from *Streptomyces resistomyceticus* contains an unusual octahedral zinc binding site. *FEBS Lett* **583**, 2917–2921, https://doi.org/10.1016/j.febslet.2009.07.061 (2009).
83. Hantai, J. et al. Biochemical and structural characterization of a novel bacterial manganese-dependent hydroxynitrile lyase. *FEBS J* **280**, S851–S858, https://doi.org/10.1111/febs.12501 (2013).
84. Gopal, B., Madan, L. L., Betz, S. F. & Kossiakoff, A. A. The crystal structure of a quercetin 2,3-dioxygenase from *Enterococcus faecalis*. *J Mol Biol* **395**, 1–12, https://doi.org/10.1016/j.jmb.2010.04.028 (2010).
85. Hanekop, N. et al. Crystal structure of the ligand-binding protein EluB from *Sporosarcina villosa* reveals substrate recognition for the compatible solutes ectoine and hydroxyectoine. *J Mol Biol* **374**, 1237–1250, https://doi.org/10.1016/j.jmb.2007.09.071 (2007).
86. Kuhlmann, S. I., Terwisscha van Scheltinga, A. C., Bienert, R., Kunte, H. J. & Ziegler, C. A 1.55 A structure of the ectoine binding domain from *Sphingopyxis alaskensis*. *Acta Crystallogr D Biol Crystallogr* **62**, 374–377, https://doi.org/10.1107/S0907444911056414 (2016).
87. Schulz, A. et al. Feeding on compatible solutes: a substrate-induced pathway for uptake and catabolism of ectoines and its genetic control by EnuR. *Environ Microbiol* **19**, 926–946, https://doi.org/10.1111/1462-2920.13414 (2017).
88. Schulz, A. et al. Transcriptional regulation of ectoine catabolism in response to multiple metabolic and environmental cues. *Env Microbiol* **19**, 4599–4619 (2017).
89. Kunte, H. J., Galinski, E. A. & Trüper, G. H. A modified FMOC-method for the detection of amino acid-type osmolytes and tetrahydropyrimidines (ectoines). *J Microbiol Meth* **17**, 129–136 (1993).
90. Miller, J. H. Experiments in molecular genetics. (Cold Spring Harbor Laboratory, 1972).
91. Hoepner, A., Widdrich, N., Bremer, E. & Smits, S. H. J. Overexpression, crystallization and preliminary X-ray crystallographic analysis of the ectoine hydroylase from *Sphingopyxis alakensis*. *Acta Cryst.* F**70**, 493–496 (2014).
92. Kabsch, W. XDS. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132, https://doi.org/10.1107/S0909449909047337 (2010).
93. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* **66**, 133–144, https://doi.org/10.1107/S0909449909047337 (2010).
94. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–500, https://doi.org/10.1107/S0909444910007493 (2010).
95. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355–367, https://doi.org/10.1107/S0907444911001314 (2011).
96. Delano, W. L. *The PyMol molecular graphics system*. (Delano Scientific, 2002).
97. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J Mol Biol* **215**, 403–410 (1990).
98. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772–780, https://doi.org/10.1093/molbev/msx010 (2013).
99. Lutunic, I. & Bork, P. Interactive Tree Of Life2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res* **39**, W475–478, https://doi.org/10.1093/nar/gkr201 (2011).

**Acknowledgements**

We cordially acknowledge the staff of beamline P13 (DESY, EMBL, Hamburg, Germany) for their kind and helpful support during data collection and are equally thankful to the staff of the ID29 beamline of the European Synchrotron Radiation Facility (Grenoble, France) for expertly providing synchrotron radiation facilities. We greatly appreciate the generous gift of ectoine by the bitop AG (Witten, Germany) for our study and the helpful discussions with Wolfgang Buckel. We thank Jochen Sohn for his excellent technical assistance during EctC protein overproduction and purification, and acknowledge N. Widdrich for the construction of the ectC overexpression plasmid pWN14. We are grateful to Vickie Koogler for her help in the language editing of our manuscript and to our colleague Lutz Schmitt (Institute of Biochemistry, Heinrich–Heine Universität Düsseldorf) for his kind support of this project. This work was financially supported in part by grants from...
the Deutsche Forschungsgemeinschaft (DFG) through the SFB 987 (Philipps-University Marburg; to E.B.), by the LOEWE program of the state of Hessen (via the Center for Synthetic Microbiology, Marburg; to J.H. and E.B.), by the Institute of Biochemistry (Heinrich-Heine-University Düsseldorf; to S.H.J.S.) and the DFG-funded Transregio research consortium TRR51 (Rheinische Friedrich Wilhelms-Universität Bonn; to J.S.D.). L.C. gratefully acknowledges the receipt of a Ph.D fellowship from the International Max Planck Research School for Environmental, Cellular and Molecular Microbiology (IMPRS-Mic, Marburg).

Author Contributions
E.B. and S.H.J.S. conceived and coordinated the study and evaluated all data. L.C. performed overexpression, purification, functional characterization of proteins, constructed mutants and analyzed the resulting data. A.H., S.K. and S.H.J.S. designed and performed crystallization experiments, solved the crystal structures and analyzed the data. R.R. and J.S.D. synthesized reagents and analyzed their identity and purity. A.S. performed mass spectrometry. J.H. and J.S.D. made the proposal for the ectoine synthase reaction mechanism. E.B., L.C., S.H.J.S., J.H. and J.S.D. wrote the manuscript with input from other authors.

Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-36247-w.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019