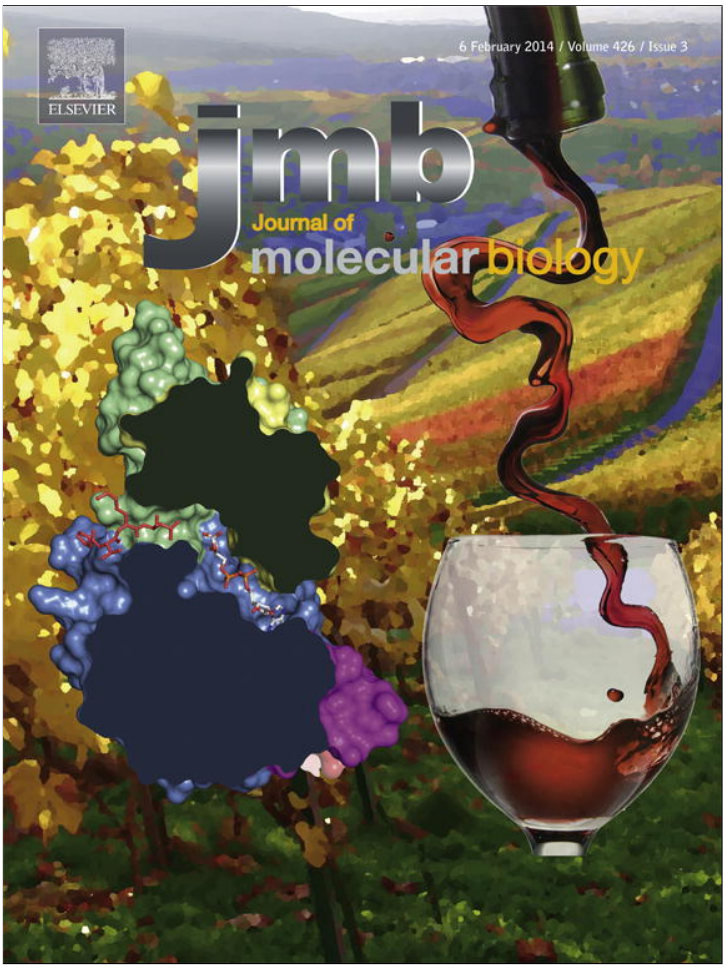

Article

Molecular Dynamics Simulations and Structure-Guided Mutagenesis Provide Insight into the Architecture of the Catalytic Core of the Ectoine Hydroxylase

Nils Widderich[1,2,3,†], Marco Pittelkow[1,†], Astrid Höppner[4], Daniel Mulnaes[5], Wolfgang Buckel[1,2,6], Holger Gohlke[5], Sander H.J. Smits[7] and Erhard Bremer[1,2]

*1 - **Department of Biology,** Laboratory for Microbiology, Philipps-University Marburg, 35037 Marburg, Germany*
*2 - **LOEWE Center for Synthetic Microbiology,** Philipps-University Marburg, 35037 Marburg, Germany*
*3 - **Max Planck Institute for Terrestrial Microbiology Emeritus Group R. K. Thauer,** 35043 Marburg, Germany*
*4 - **X-ray Facility and Crystal Farm,** Heinrich-Heine-University Düsseldorf, 40225 Düsseldorf, Germany*
*5 - **Institute for Pharmaceutical and Medicinal Chemistry,** Heinrich-Heine-University Düsseldorf, 40225 Düsseldorf, Germany*
*6 - **Max Planck Institute for Terrestrial Microbiology,** 35043 Marburg, Germany*
*7 - **Institute of Biochemistry,** Heinrich-Heine-University Düsseldorf, 40225 Düsseldorf, Germany*

***Correspondence to Erhard Bremer:** bremer@staff.uni-marburg.de*
http://dx.doi.org/10.1016/j.jmb.2013.10.028
Edited by D. Case

Abstract

Many bacteria amass compatible solutes to fend-off the detrimental effects of high osmolarity on cellular physiology and water content. These solutes also function as stabilizers of macromolecules, a property for which they are referred to as chemical chaperones. The tetrahydropyrimidine ectoine is such a compatible solute and is widely synthesized by members of the *Bacteria*. Many ectoine producers also synthesize the stress protectant 5-hydroxyectoine from the precursor ectoine, a process that is catalyzed by the ectoine hydroxylase (EctD). The EctD enzyme is a member of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily. A crystal structure of the EctD protein from the moderate halophile *Virgibacillus salexigens* has previously been reported and revealed the coordination of the iron catalyst, but it lacked the substrate ectoine and the co-substrate 2-oxoglutarate. Here we used this crystal structure as a template to assess the likely positioning of the ectoine and 2-oxoglutarate ligands within the active site by structural comparison, molecular dynamics simulations, and site-directed mutagenesis. Collectively, these approaches suggest the positioning of the iron, ectoine, and 2-oxoglutarate ligands in close proximity to each other and with a spatial orientation that will allow the region-selective and stereo-specific hydroxylation of (4*S*)-ectoine to (4*S*,5*S*)-5-hydroxyectoine. Our study thus provides a view into the catalytic core of the ectoine hydroxylase and suggests an intricate network of interactions between the three ligands and evolutionarily highly conserved residues in members of the EctD protein family.

Introduction

The exposure of cell-walled microorganisms to high osmolarity surroundings poses a considerable challenge to the physiology of the bacterial cell since it invariably elicits water efflux. The resulting loss of turgor impairs growth and threatens viability [1,2]. Many microorganisms offset these detrimental effects by amassing, either through synthesis or through uptake, compatible solutes, a class of organic osmolytes with special physiochemical properties [3–5]. Compatible solutes do not only serve as water-attracting osmolytes [1,2] but they are also highly compliant with cellular physiology and can promote and preserve the functionality of individual proteins [6–8], of protein complexes [9], and even of entire cells [10], a property for which the term "*chemical chaperone*" has been coined in the literature [11,12].

Ectoine [(4*S*)-2-methyl-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid] is an important representative of

the compatible solutes [3,13]. Its properties as a superb stabilizer of macromolecules [14] and excellent cytoprotectant [15–17] led to efforts to produce it biotechnologically on an industrial scale using highly salt tolerant microorganisms [15,18] and fostered commercial applications (e.g., in skin-care products, as PCR enhancers, or as protein and cell stabilizers); the use of ectoine for medical purposes is also currently being pursued [15–17]. Ectoine is synthesized almost exclusively by members of the *Bacteria*, and they produce it, or import it, as a protectant against stress caused by high salinity/osmolarity [16,18–25] or to relieve cellular challenges imposed by extremes in high or low growth temperatures [19,26–28]. Ectoine is synthesized from the precursor L-aspartate-β-semialdehyde, a central intermediate of microbial amino acid metabolism [29,30], by three successive steps that are catalyzed by the EctB, EctA, and EctC enzymes [23,31].

Ectoine can be classified either as a heterocyclic amino acid or as a partially hydrogenated pyrimidine derivative [22]. A substantial subgroup of the ectoine producers also synthesize a hydroxylated derivative of ectoine, 5-hydroxyectoine [(4*S*,5*S*)-5-hydroxy-2-methyl-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid] [32], through a region-selective and stereo-specific enzymatic reaction from ectoine [20,33]. Despite their close relatedness in chemical structure, 5-hydroxyectoine often possesses stabilizing and stress-protecting properties superior to those of ectoine [19,34–36]. Efforts in several laboratories led to the identification of the structural gene (*ectD*) for the ectoine hydroxylase (EctD) in various microorganisms [20,27,37], and database searches for EctD-related proteins showed that this enzyme is evolutionarily well conserved [20,38].

The biochemical purification and enzymatic characterization of the EctD protein from the moderate halophile *Virgibacillus salexigens* [20] and from the soil bacterium *Streptomyces coelicolor* [19] identified this enzyme as a member of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily (EC 1.14.11). Members of this superfamily are versatile biocatalysts whose enzyme reactions include hydroxylations, dimethylations, desaturations, epimerizations, cyclizations, halogenations, and ring expansions [39–41]. Despite these diverse biochemical activities, the so far determined crystal structures of members of this family are all closely related, and common principles of their enzyme reaction mechanism have emerged [40,42,43]. In this vein, EctD follows the enzyme reaction scheme (e.g., for the taurine dioxygenase TauD [44,45]) of this class of enzymes in catalyzing the O_2-dependent hydroxylation of ectoine. This reaction relies on a mononuclear iron catalyst and uses 2-oxoglutarate as a co-substrate; EctD thereby forms CO_2, succinate, and 5-hydroxyectoine (Fig. 1) [20,38].

The crystal structure of the biochemically characterized EctD from *V. salexigens* (formally taxonomically classified as *Salibacillus salexigens*) [20] (PDB accession code: 3EMR) is the only available structure of an ectoine hydroxylase [38]. It revealed a protein fold for the monomeric EctD [20,38] that is commonly observed among non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenases [39–41,43]. The EctD crystal structure [38] consists of a double-stranded β-helix (DSBH) core, also referred to as the jelly roll or cupin fold [40], which is decorated with and stabilized by a number of α-helices. The DSBH in EctD is formed by four-stranded antiparallel β-sheets and is arranged in form of a β-sandwich (Fig. 2a).

The EctD structure also revealed the unambiguous positioning of the iron ligand within the active site, but it lacked the substrate ectoine and the co-substrate 2-oxoglutarate [38]. Within the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily, the catalytically important iron ligand is typically coordinated by side chains from a conserved HxD/E…H motif, the so-called 2-His-1-carboxylate facial triad [39,40,42,43]. The HxD/E…H motif is also present in the EctD protein [20], and the iron co-factor is coordinated by the functional groups of His146, Asp148, and His248 (Fig. 2a) that protrude into the ligand-binding cavity (Fig. 2b). The iron-coordinating residues His146 and Asp148 are part of a strictly conserved signature sequence motif for EctD-type hydroxylases of 17 amino acids in length (extending from Phe143 to Pro159) [20,38]. This consensus sequence spans an extended α-helix and a linked short β-strand within the EctD crystal structure (Fig. 2a), thereby lining one site of the presumed active site of the ectoine hydroxylase [38].

Since a deeper understanding of EctD is desirable, both with respect to basic science and industrial applications, we used the crystal structure of the *V. salexigens* EctD protein as a template for molecular docking experiments, molecular dynamics (MD) simulations, and structure-guided site-directed mutagenesis to provide deeper insight into the architecture of the catalytic core of the EctD enzyme. Emerging from these studies is a model that allows the positioning of the substrate ectoine and the co-substrate 2-oxoglutarate within the active site of the ectoine hydroxylase in such a way that it will permit the region-selective and stereo-specific hydroxylation of (4*S*)-ectoine to (4*S*,5*S*)-5-hydroxyectoine.

Results and Discussion

A view of the protein space of EctD-type enzymes

We used the *V. salexigens* EctD protein (accession number: AY935522) [20] as the query sequence

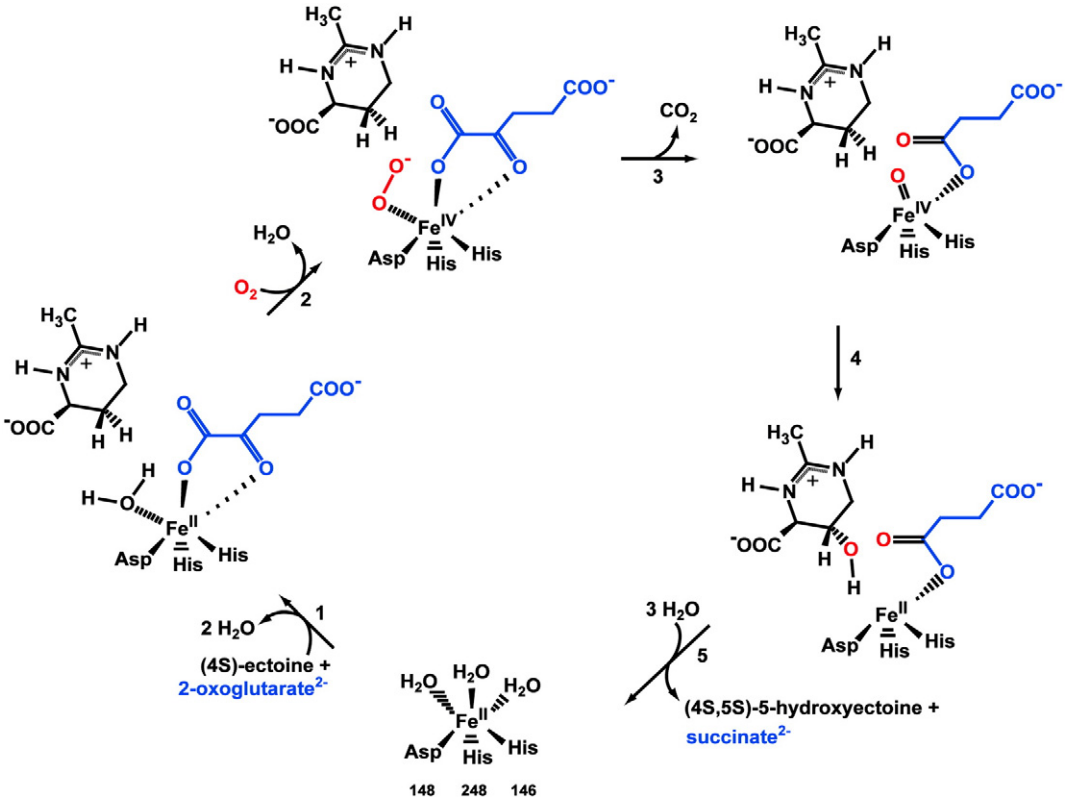

Fig. 1. EctD-catalyzed hydroxylation of ectoine. In members of non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily, a mononuclear ferrous iron center catalyzes the oxidative decarboxylation of the co-substrate 2-oxoglutarate, and this reaction is coupled with a two-electron oxidation of the substrate [40,43]. The co-substrate 2-oxoglutarate and the substrate appear to bind in a sequential fashion leading to the formation of a ternary enzyme-2-oxoglutare-substrate complex to which molecular oxygen can then bind. The proposal for the EctD-catalyzed hydroxylation of ectoine shown here is adapted from the well-established enzyme reaction mechanisms of the taurine dioxygenase TauD [44,45]: (1) The substrate-free EctD enzyme contains Fe^{II}, 6-fold-coordinated by the 2-His-1-carboxylate facial triad formed by the side chains of His146, Asp148, and His248, as well as three water molecules [38]. Two of the water molecules are replaced from the iron center upon addition of 2-oxoglutarate. Subsequently, the substrate ectoine is envisioned to bind to the EctD enzyme. (2) Upon substrate binding, dioxygen replaces the third water molecule from the iron center and oxidizes Fe^{II} to Fe^{IV}. (3) The formed Fe^{IV}-peroxo species attacks with its anion the carbonyl-C of 2-oxoglutarate to yield succinate, CO_2, and the ferryl species (Fe^{IV} = O) [46]. The ferryl species then hydroxylates ectoine to (4*S*,5*S*)-5-hydroxyectoine. (4) This may occur either by abstraction of the *pro-S* hydrogen at C5 by Fe^{IV} = O and re-addition of the hydroxyl radical from Fe^{III}-OH to the ectoine radical intermediate with retention of configuration (re-bound mechanism) or by insertion of the ferryl oxygen atom into the C–H bond (two-state reactivity mechanism) [47]. (5) Finally, the product (4*S*,5*S*)-5-hydroxyectoine and the co-product succinate are released. Three water molecules will occupy then again the empty coordination site of Fe^{II}.

in a BLAST search of the Joint Genome Institute database of microorganisms with a finished genome sequence. This search retrieved 184 proteins related to the *V. salexigens* EctD protein. An alignment of these amino acid sequences (Supplementary Fig. S1) with the ClustalW algorithm [48] showed that they possess a degree of sequence identity that ranges from 71% for the EctD protein from *Bacillus pseudofirmus* OF4 to 49% for the EctD protein from *Planctomyces maris* (DSM8797).

Since 5-hydroxyectoine is synthesized directly from ectoine [20], each of the microorganisms that possess an *ectD* gene also possess *ectC*, the structural gene for the ectoine synthase, the diagnostic enzyme for the ectoine biosynthetic route [18,21,23,24,31,49,50]. This strict correlation between the presence of the EctD and EctC proteins was ascertained by using the EctC amino acid sequence of *V. salexigens* [20] as the search query (data not shown). In each of the retrieved EctD-related proteins, the previously identified signature sequence (F*X*WHSDFETWH*X*EDGM/LP) for ectoine hydroxylases [20,38] is fully conserved. The *V. salexigens* EctD protein consists of 300 amino acid residues, and this is approximately the typical size of EctD-type proteins (Supplementary Fig. S1). Among the 185 aligned amino acid sequences, 29 amino acid

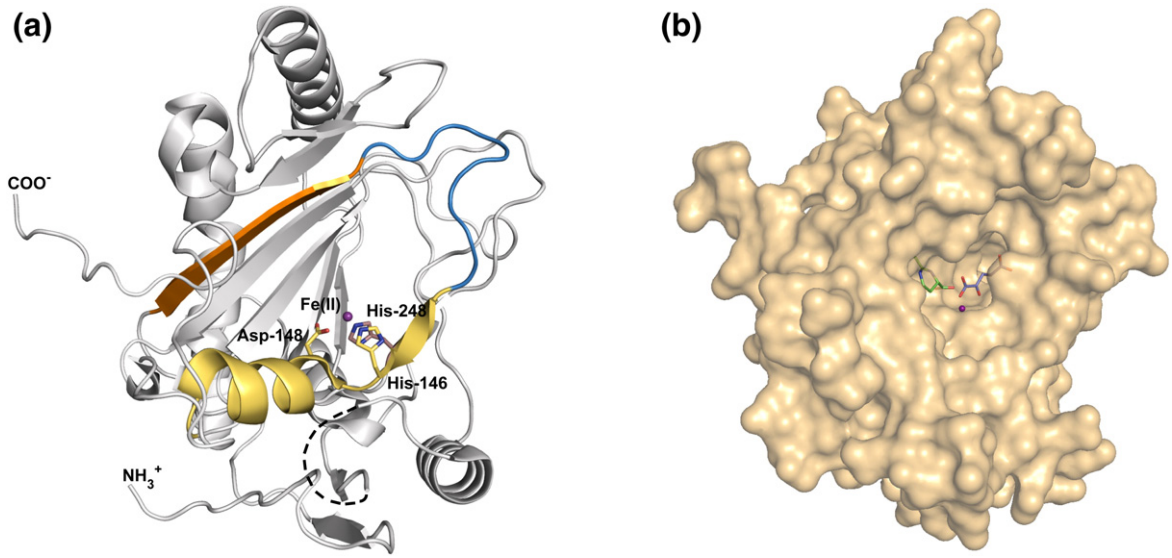

Fig. 2. Overall structure of the ectoine hydroxylase EctD from the moderate halophile *V. salexigens*. (a) The crystal structure of the EctD protein (PDB accession code: 3EMR) [38] is shown in a ribbon presentation format. The region of EctD serving as the signature sequence of ectoine hydroxylases [20,38] (highlighted in yellow) and its connection (marked in blue) to a β-sheet (highlighted in orange) containing residues important for ectoine and 2-oxoglutarate binding are emphasized. The iron(II) species is represented as a purple sphere. The region indicated by the broken line represents a flexible loop (Gly195–Leu211) that is not resolved in the EctD crystal structure [38]. (b) Surface representation of the EctD protein containing the iron(II) ligand [38] and the modeled 2-oxoglutarate and ectoine ligands.

residues are strictly conserved, including 14 residues that are part of the ectoine hydroxylase signature sequence [20,38]. Taken together, the inspection of the 185 aligned amino acid sequences revealed a strong evolutionary conservation of EctD-type proteins and a strict correlation of their occurrence with the presence of the ectoine biosynthetic enzymes in the inspected microbial genomes. Thus, we suggest that the 184 proteins identified in our database analysis are all *bona fide* ectoine hydroxylases. Each of these proteins bears the characteristic residues of enzymes that belong to the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily [40–43].

The iron-binding site

To probe the functional role of His146, Asp148, and His248 in the *V. salexigens* EctD protein for iron binding and enzyme activity [20], we separately replaced each of these amino acids by an Ala residue. None of these EctD variants was catalytically active (Table 1) nor did any of the mutant proteins contain iron (data not shown). In contrast, recombinant preparations of the wild-type *V. salexigens* EctD enzyme isolated from *Escherichia coli* cells by affinity chromatography on a Strep-Tactin resin typically had an iron content up to 0.9 mol per mole of protein [20].

In some members of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily, a Glu residue can functionally substitute for the Asp residue forming part of the iron-binding site [40,42,43]. Thus, we replaced Asp148 by a Glu residue; however, this EctD variant did not show enzyme activity either (Table 1). Collectively, we conclude from these data that no variations in the side chains forming the iron(II)-binding center of the EctD enzyme are tolerated and that all three residues forming the 2-His-1-carboxylate facial triad in the EctD protein [38] are required for iron binding and enzyme activity.

The 2-oxoglutarate-binding site

Non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenases crystallized in complex with either 2-oxoglutarate or analogs possess a similarly structured binding site for this co-substrate [40,43]. Within this binding site, the 2-oxo carboxylate interacts with the iron atom in a bidentate manner; the C-5 carboxylate group points toward the inside of the active-site pocket and is bound via electrostatic and hydrogen bonding interactions. Based on the information gained from available crystal structures of members of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenases [40,43], the spatial positioning of 2-oxoglutarate relative to the iron ligand in the EctD active site [38] should be predictable with a high degree of confidence.

Table 1. Kinetic properties of investigated EctD mutant derivatives.

Mutant	K_m (mM ectoine)	V_{max} (U mg^{-1})	Proposed role in the binding of the following ligands
Wild type	5.9 ± 0.3	6.2 ± 0.2	—
H146A	—	—	Fe(II)
D148A	—	—	Fe(II)
D148E	—	—	Fe(II)
H248A	—	—	Fe(II)
F95A	—	—	2-Oxoglutarate
R131A	—	—	2-Oxoglutarate
N133A	—	—	2-Oxoglutarate
F143A	—	—	2-Oxoglutarate
F143Y	—	—	2-Oxoglutarate
F143W	—	—	2-Oxoglutarate
S250A	—	—	2-Oxoglutarate
R259A	—	—	2-Oxoglutarate
R259K	—	—	2-Oxoglutarate
R259H	—	—	2-Oxoglutarate
R259Q	—	—	2-Oxoglutarate
Q129A	—	—	Ectoine
W152A	—	—	Ectoine
W152F	—	—	Ectoine
W152Y	—	—	Ectoine
S165A	17.3 ± 1.0	2.3 ± 0.6	Ectoine
F242A	—	—	Ectoine
F242Y	19.6 ± 0.5	1.7 ± 0.3	Ectoine
F242W	—	—	Ectoine
F263A	—	—	Ectoine
F263Y	8.0 ± 0.2	3.7 ± 0.4	Ectoine
F263W	—	—	Ectoine
P198A	7.1 ± 0.7	4.8 ± 0.4	Loop region
S205A	6.8 ± 0.2	5.4 ± 0.2	Loop region
A163C	6.0 ± 0.1	5.9 ± 0.4	Not involved in ligand binding
S167A	6.9 ± 0.2	6.0 ± 0.1	Not involved in ligand binding
S244C	6.2 ± 0.4	6.0 ± 0.5	Not involved in ligand binding
N261A	5.5 ± 0.4	6.4 ± 0.3	Not involved in ligand binding
V265A	16.2 ± 0.5	3.4 ± 0.2	Not involved in ligand binding
V265T	9.7 ± 0.3	5.6 ± 0.4	Not involved in ligand binding
V265L	24.6 ± 0.9	1.6 ± 0.2	Not involved in ligand binding
A163C, S244C	4.5 ± 0.6	8.1 ± 0.8	Not involved in ligand binding

The enzymatic activity of the wild-type EctD enzyme and its mutant derivatives was assayed as described in Materials and Methods. The kinetic parameters were determined by systematically varying the substrate (ectoine) concentration between 0 and 40 mM. EctD mutants that showed no activity under these conditions were considered to be catalytically inactive (–).

A search via the Dali server [51] for crystal structures related to EctD revealed the human peroxisomal phytanoyl-CoA 2-hydroxylase PhyH [52] (also known as PAXH) [53] and its relative PhyHD1A [54] as the most closely related structures to EctD with *Z*-scores of 22.6 and 22.0, respectively, and RMSD values of 2.5 Å (over 190 C^{α} atoms). While the structural relatedness of the human PhyH and PhyHD1A to EctD is substantial, the degree of their amino acid sequence identity is moderate (19% and 24%, respectively). Both the PhyH and PhyHD1A proteins have been crystallized in complex with 2-oxoglutarate [52,54], and we used the PhyHD1A crystal structure (PDB accession code: 2OPW) [54] as a point of reference to gain information about the potential binding site of 2-oxoglutarate within the EctD enzyme.

The side chains of Arg257, Ser248, and Trp174 of PhyHD1A position 2-oxoglutarate in the vicinity of the catalytically important iron ligand (Fig. 3a). The corresponding residues in EctD are Arg259, Ser250, and Phe143 (Fig. 3b). We separately replaced these residues in EctD by an Ala residue and found that each of these three EctD mutant proteins was catalytically inactive (Table 1). The 2-oxoglutarate ligand present in the PhyHD1A crystal structure contacts directly, via a charge-assisted hydrogen bond, the side chain of Trp174 (Fig. 3a). Such an interaction is not found in the predicted 2-oxoglutarate-binding site of EctD; instead, Phe143 located on the opposite side of the 2-oxoglutarate molecule makes a stacking interaction with this ligand via its aromatic side chain (Fig. 3b). We replaced Phe143 with the aromatic residues Tyr or Trp, but these EctD variants were enzymatically inactive (Table 1). Hence, Phe143 seems to play a specific role for the activity of the ectoine hydroxylase. This suggestion is supported by the complete conservation of this residue among 185 inspected EctD-type proteins (Supplementary Fig. S1). Taking the structural comparisons and the site-directed mutagenesis data together, we suggest that Arg259, Ser250, and Phe143 constitute the 2-oxoglutarate-binding site of the ectoine hydroxylase (Fig. 3b).

Besides the interactions outlined above, the 2-oxoglutate molecule also interacts with Arg131 of the EctD protein. The exact positioning of the Arg131 side chain within the ligand-binding cavity is stabilized through a cation–π interaction with the side chain of Phe95, as well as an interaction with the side chain of Asn133 (Fig. 3b). This latter interaction ensures the exact positioning of the Arg131 side chain. The lack of the Asn133 side chain results in the loss of activity (Table 1) likely due to the higher flexibility of Arg131. As to the former interaction, whereas Arg131 is located within the DSBH core of the EctD protein, Phe95 is positioned in a β-sheet flanking this fold. Hence, the Arg131/Phe95 interaction probably stabilizes the conformation of the EctD protein as a whole. In turn, loss of this interaction will likely lay the active site of the EctD protein more open and will thereby probably render it unable to bind any of its substrates. Consistent with this suggestion, Ala substitution mutations of Arg131 or Phe95 led to enzymatically inactive EctD variants (Table 1). We are thus tempted to speculate that Arg131 senses the presence of the 2-oxoglutarate

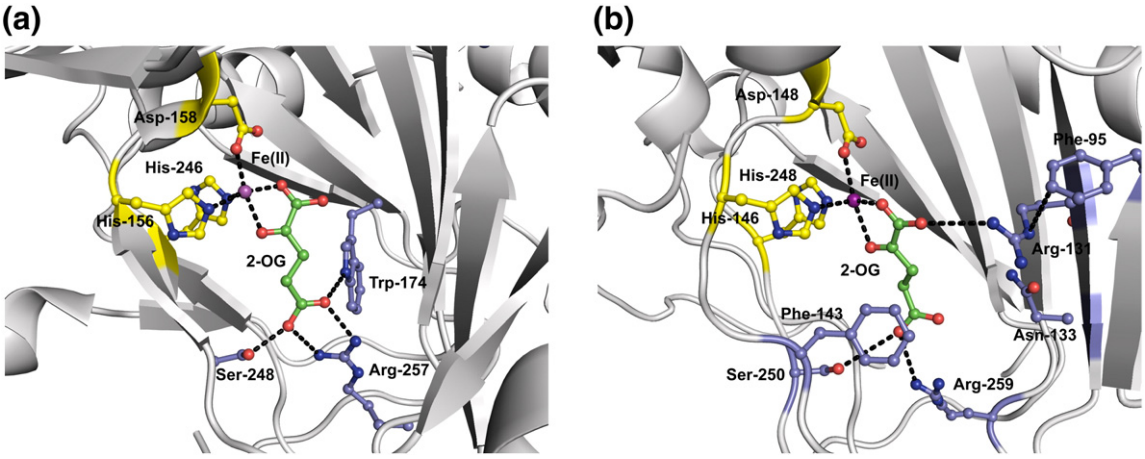

Fig. 3. Comparison of the iron(II) and 2-oxoglutarate-binding site of EctD with that present in PhyHD1A. (a) Crystal structure of a human phytanoyl-CoA dioxygenase (PhyHD1A) (PDB accession code: 3OBZ) with highlighted residues involved in the binding of iron(II) (purple sphere) and the co-substrate 2-oxoglutarate (2-OG) [54]. (b) Model of the EctD ectoine hydroxylase highlighting those residues predicted to be involved in 2-oxoglutarate (2-OG) and iron (purple sphere) binding.

co-substrate and ensures, via its interaction with Phe95, a more compact (or stable) structure of the EctD protein that is then primed for the hydroxylation of ectoine.

Inspection of the amino acid sequences of the aligned 185 EctD-type proteins (Supplementary Fig. S1) revealed that, besides the amino acid residues involved in iron binding (His146, Asp148, and His248), the amino acids predicted to be involved in 2-oxoglutarate binding (Arg259, Ser250, Phe143, Arg131, Asn133, and Phe95) (Fig. 3b) are also completely conserved (with one exception at position Ser250).

MD simulations and structure-guided mutagenesis to locate the ectoine-binding site

One of the hallmarks of compatible solutes is their preferential exclusion from the immediate hydration shell of proteins due to unfavorable interactions between these compounds and the protein backbone [6,9]. Despite the preferential exclusion of compatible solutes from protein surfaces, high-affinity interactions between these compounds and defined ligand-binding sites do occur as demonstrated by bacteria acquiring these compounds through efficient transport processes [1,3,28,55,56]. Four crystal structures for ectoine and 5-hydroxyectoine ligand-binding proteins that are associated with either bacterial ABC transporters (EhuB and OpuCC) [57,58] or TRAP transporters (UehA and TeaA) [59,60] have been reported. Although the architecture of the ectoine/5-hydroxyectoine-binding sites in these proteins varies somewhat [59], common principles for ligand binding have emerged. The carboxylate moiety of ectoine or 5-hydroxyectoine is typically coordinated within the ligand-binding site through a salt bridge with a positively charged residue. Additional salt bridges between the amidino moieties of the ectoine or 5-hydroxyectoine ligands and the substrate-binding protein provide an additional level of stabilization. A key factor contributing to ligand binding are cation–π interactions [61] that are formed between the delocalized positive charges of the ectoine and 5-hydroxyectoine molecules and aromatic side-chain residues of the ligand-binding protein. Furthermore, hydrophobic interactions and hydrogen bonds support ligand binding [57–60].

In order to generate suggestions for the location of the ectoine-binding site in EctD, we employed MD simulations. This technique has recently been successfully applied to identify ligand-binding sites within a Src kinase [62] and an alanopine dehydrogenase [63].

To generate an EctD:ectoine starting structure for the MD simulations, we first modeled the loop region between Gly195 and Leu211 of the EctD protein as these residues are missing in the available crystal structure of the *V. salexigens* ectoine hydroxylase (see Materials and Methods for details) [38]. Disordered regions corresponding to this missing loop in EctD also occur in the ligand-free crystal structures of the human phytanoyl-CoA 2-hydroxylase PhyH [52], the asparagine hydroxylase AsnO from *S. coelicolor* [64], and the AsnO-related VioC L-arginine oxygenase from *Streptomyces vinaceus* [65]. These flexible regions became ordered and structurally resolved once the PhyH, AsnO, and VioC proteins were crystallized with their corresponding ligands. The loop regions in these enzymes might thus serve as a lid to enclose the active site and shield the reactants from the surrounding

solvent during enzyme catalysis [43,52,64,65]. After careful relaxation of this EctD model, we then placed the ectoine molecule manually at a site (Supplementary Fig. S2) that we deemed suitable for ligand binding based upon principles revealed through the structural analysis of the EhuB, OpuCC, UehA, and TeaA ectoine/hydroxyectoine-binding proteins [57–60].

The *in silico* constructed EctD:ectoine complex structure (Supplementary Fig. S2) was then subjected to three independent MD simulations of 300 ns length each. These simulations revealed minor structural changes in the protein part when one does not consider the modeled loop region and the C-terminus of EctD, with RMSD values of the coordinates of C^{α} atoms (RMSD) of almost always <2.0 Å (Supplementary Fig. S3a and Fig. S4a and b). The C-terminus of the EctD protein showed a higher mobility with an RMSD up to 6 Å (data not shown); even larger movements are observed in the modeled loop region (RMSD up to 18 Å; Supplementary Fig. S3c). Although the loop moves toward the binding site in the course of the MD simulations, in none of the three trajectories did we observe that the loop actually closed the binding site as a lid, nor were interactions between loop residues and the ectoine ligand observed.

In the MD simulations, the ectoine ligand initially relocates from its manually chosen starting position (Supplementary Fig. S2) to a nearby sub-pocket formed by the side chains of residues Gln129, Trp152, Ser165, Phe242, and Phe263 characterized by RMSD values between 5 and 6 Å (Fig. 4a). It occasionally moves back toward the starting position (RMSD ≈ 3 Å) after which it returns again to the above-mentioned sub-pocket (Fig. 4a). The movement of ectoine to the sub-pocket was observed in all three independent trajectories (Fig. 4a and Supplementary Fig. S4a and b) and thereby indicated that the ectoine molecule was highly mobile, particularly at the end of simulation 1 (Fig. 4a) and around 75 and 225 ns in simulation 3 (Supplementary Fig. S4b). This high mobility of the ectoine ligand suggests a low binding affinity of the EctD protein for its substrate; this is consistent with a high K_m value, which is between 2.6 and 6 mM, depending on the particular assay conditions used and the enzyme studied [19,20]. We note that, despite the high mobility of the ectoine molecule in the MD simulations, the ligand repeatedly returns to the above-mentioned sub-pocket (as indicated by RMSD values between 5 and 6 Å; Fig. 4a and Supplementary Fig. S4b), suggesting that this sub-pocket is the preferred location for ectoine binding by the EctD protein (Fig. 4d).

Within this presumed ligand-binding site (Fig. 4d), ectoine preferentially orients such that its methyl group points to the backbone region of Ser130 at the bottom of the sub-pocket [with a distance of <5 Å between the ligands methyl group and C^{α} of Ser130 observed in ~50% of all conformations (Fig. 4b)]; its carboxylate group forms hydrogen bonds with the side chain of Ser165 in ~15% of all conformations [considering a cutoff distance of 3.2 Å between the donor and acceptor atoms (Fig. 4c)]. This binding mode leads to infrequent hydrogen bonds between the amidino moiety of ectoine and the side chain of Gln129 (~2% occurrence; data not shown). Interactions of the carboxylate group with other potential hydrogen bond donors in the binding site region, in particular, Arg131 and Ser167, are negligible (below 0.1% occurrence in all three trajectories). In this orientation, the *pro-S* hydrogen at C5 of the tetrahydropyrimidine ring of ectoine is oriented toward the catalytic iron atom (Fig. 4d), fully consistent with the fact that this hydrogen of the ectoine molecule gets replaced by a hydroxyl group during the EctD-catalyzed reaction cycle leading to the formation of (4*S*,5*S*)-5-hydroxyectoine [20,33].

To substantiate the ectoine-binding site model derived from the MD simulations (Fig. 4d), we carried out a series of site-directed mutagenesis experiments and tested the resulting EctD variants in enzyme activity assays (Table 1). Direct interactions are predicted to occur between the ectoine substrate and the side chains of Gln129 and Ser165; in addition, Trp152, Phe242, and Phe263 form the wall of the above-mentioned sub-pocket (Fig. 4d). The individual substitution of Gln129, Trp152, Phe242, and Phe263 with an Ala residue resulted in a complete loss of enzyme activity, and the replacement of Ser165 with an Ala residue substantially impaired the functioning of the EctD protein (Table 1). The conservative substitution of Trp152 with either Phe or Tyr residues resulted in enzymatically inactive EctD mutant proteins (Table 1). This can be readily understood in view of our model because the second ring of the Trp152 side chain is not present in the Phe or Tyr residues but it is this second ring that is in contact with the ectoine molecule (Fig. 4d).

Consistent with the suggestion of a functional role of Phe242 and Phe263 for ectoine binding (Fig. 4d), their individual replacement with Ala residues abolished EctD enzyme activity (see above; Table 1). We also exchanged both Phe242 and Phe263 with either Tyr or Trp residues. The Phe242/Tyr mutant is about 3-fold less active than the wild-type enzyme, whereas the Phe263/Tyr variant possesses wild-type enzymatic activity (Table 1). Conversely, the substitution of either Phe242 or Phe263 by Trp residues completely abolished the enzyme activity of the EctD mutant proteins (Table 1). Our data thus indicate that the architecture of the EctD ectoine-binding site does not tolerate larger structural modifications such as the insertion of the bulky side chain of a Trp residue. The ectoine ligand also contacts Ser130; since this contact is

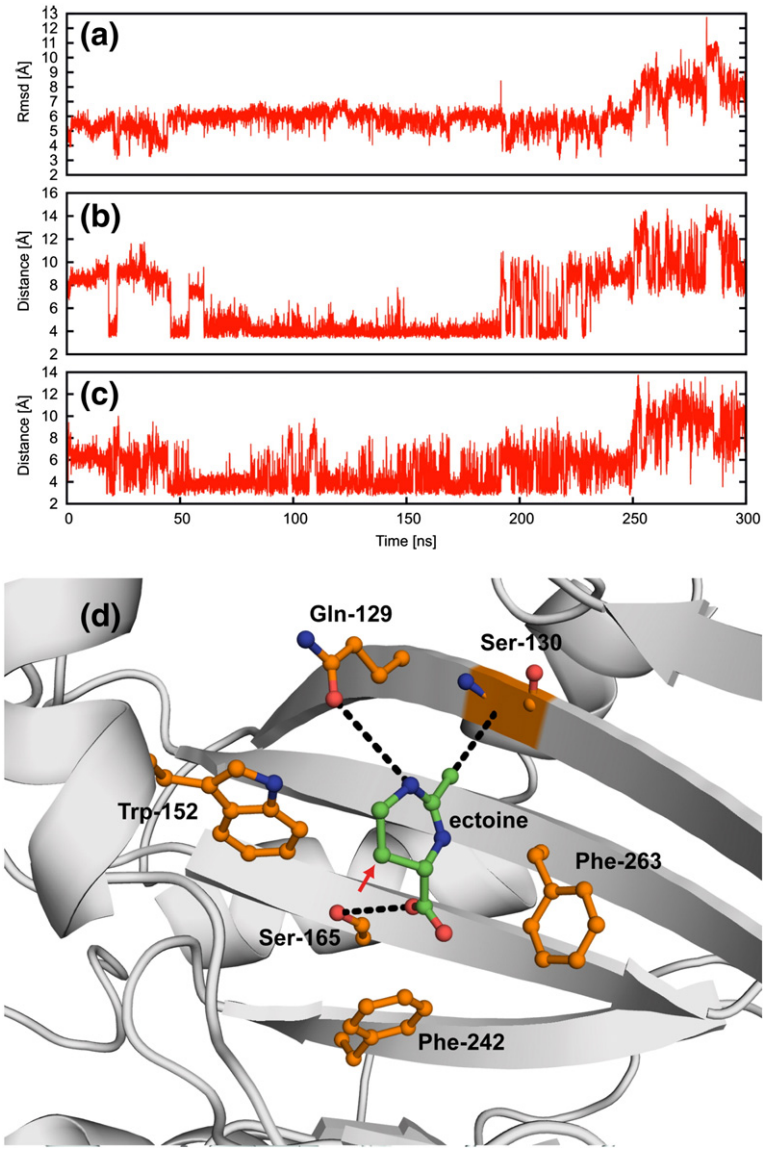

Fig. 4. MD simulation of ectoine binding by the EctD protein.(a–c) Structural parameters during the MD simulation 1 of the EctD:ectoine complex. (a) All-atom RMSD of ectoine after superimposing the core region (see Materials and Methods) of EctD on itself. (b) Distance between the methyl carbon of ectoine and the C^{α} atom of Ser130. (c) Respective minimal distance between the oxygen atoms of the carboxylate group of ectoine and the hydroxyl oxygen of Ser165. (d) Representative structure for the binding of ectoine within the EctD active site showing the dominant interactions to Ser130 and Ser165 and the occasional hydrogen bond to Gln129; in addition, Trp152, Phe242, and Phe263 are shown, which contribute to the formation of the sub-pocket containing the ectoine ligand. The read arrowhead indicates the position of the C5 atom within the ectoine molecule that becomes hydroxylated in the course of the enzyme reaction.

mediated via the protein main chain of EctD (Fig. 4d), no site-directed mutagenesis of Ser130 was carried out.

Taken together, site-directed mutagenesis of practically all residues implicated by the MD simulation in ectoine binding yield EctD variants that are either completely enzymatically inactive or strongly impaired in their function (Table 1). Furthermore, all those residues that we implicate in ectoine binding are very strongly conserved in the compiled 185 EctD-type proteins (Supplementary Fig. S1).

In our site-directed mutagenesis experiments, we specifically targeted residues that were deemed to be important for the binding of the iron, 2-oxoglutarate, and ectoine ligands and the replacement of these functionally important residues yielded almost invariably catalytically inactive EctD variants

(Table 1). While this is not surprising, it may raise the concern that essentially any amino acid substitution in the EctD protein would render the ectoine hydroxylase non-functional. We therefore carried out a series of additional mutagenesis studies of residues whose side chains either are located further away from the catalytic core of the EctD enzyme or protrude into the cupin barrel but do not make direct contacts to any of the ligands. All 10 additional mutants constructed, including residues Pro198 and Ser205 that are present within the flexible loop region of the EctD protein [38], were enzymatically active (Table 1).

Description of the architecture of the active site of the ectoine hydroxylase

The crystal structure of the EctD enzyme [38] adheres to the common protein fold of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily where a series of antiparallel β-sheets form one side of a barrel-like structure [40,43] (Fig. 1a). The side chains of those residues that we implicate in binding of the iron, 2-oxoglutarate, and ectoine ligands all protrude into the inside of the cupin barrel (Fig. 5) forming the deep ligand-binding cavity of the EctD enzyme (Fig. 2b). Key contributions to the precise positioning of these ligands are mediated by residues present on an extended α-helix and a flanking short β-strand, which contain most of the residues making up the signature sequence of EctD-type hydroxylases [20,38], and β-sheet number I, which forms the other side of the cupin barrel (Fig. 1a). The extended α-helix and its flanking short β-strand comprise residues important for iron (His146 and Asp148), 2-oxoglutarate (Phe143), and ectoine binding (Trp152) (Fig. 5). β-Sheet number I contains three strictly conserved residues (Gln129, Arg131, and Asn133) of EctD-type proteins (Supplementary Fig. S1). Gln129 is suggested to make a direct contact to the ectoine ligand (Fig. 4d), Arg131 is implicated in contacting the 2-oxoglutarate

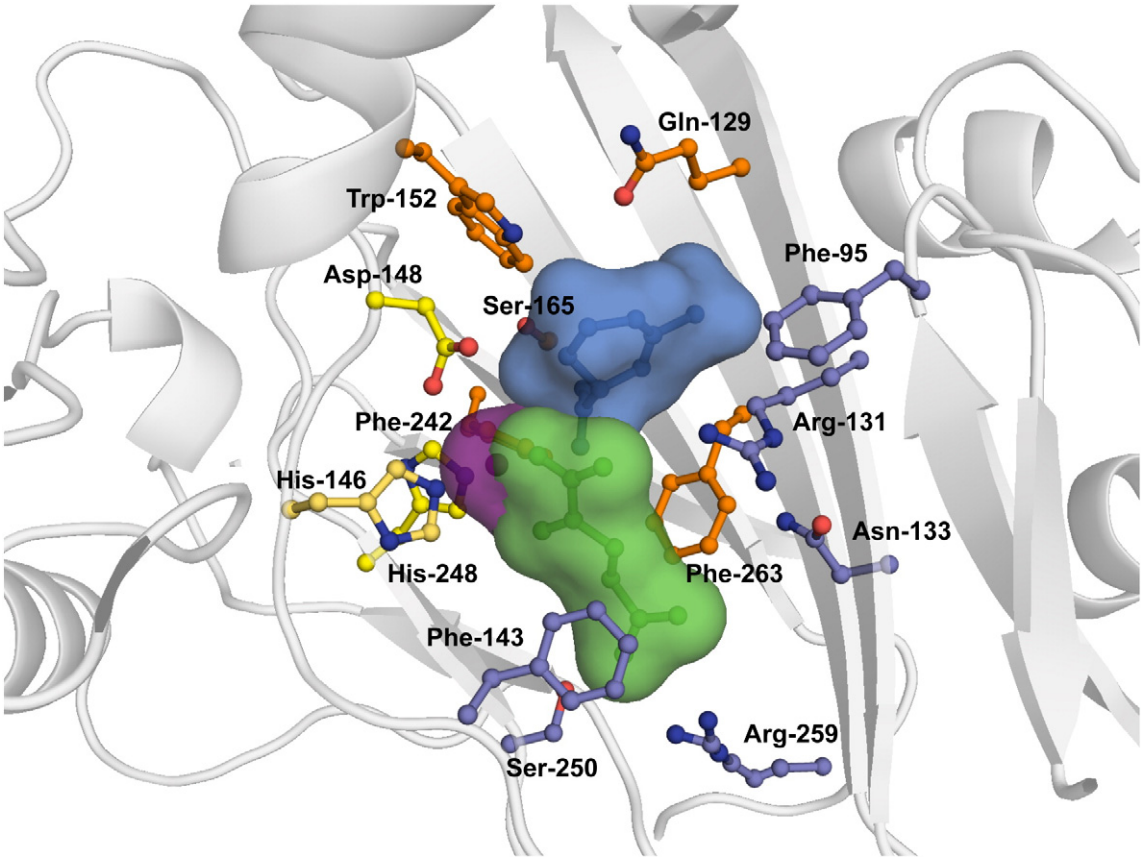

Fig. 5. Model for the catalytic core of the EctD hydroxylase. This model summarizes the predictions made by MD simulations for the positioning of the substrate ectoine (shown as blue surface), the docking of the co-substrate 2-oxoglutarate (shown as green surface), and the experimentally determined binding site for the iron(II) ligand (shown as magenta surface) [38]. The amino acids involved in binding of ectoine (orange), 2-oxoglutarate (blue), and iron (yellow) are shown in a ball-and-stick representation.

co-substrate, and Asn133 is thought to stabilize the proper conformation of the side chain of Arg131 (Fig. 3b).

Concluding remarks

We performed MD simulations starting from an *in silico* model of ectoine bound to the *V. salexigens* EctD enzyme that was inspired by the architecture of the ectoine-binding site in crystal structures of solute receptors operating in conjunction with transport systems for this compatible solute [57–60]. However, the conducted MD simulations suggest a coordination of the ectoine ligand within the active site of the ectoine hydroxylase (Fig. 4d) that is different from that observed in the aforementioned solute receptor proteins. In particular, cation–π interactions between the delocalized positive charge of the ectoine ring and aromatic residues present in the ligand-binding proteins [57–60] are not observed in the predicted ectoine-binding site of EctD (Fig. 4d). It is highly likely that these different modes of ectoine binding will contribute to the strong differences in affinity for the ectoine ligand, which are in the micromolar range for the EhuB, UehA, and TeaA solute receptors [57,59,60] and in the millimolar range for the two biochemically characterized EctD enzymes [19,20].

When one views our model for the EctD active site as a whole (Fig. 5), an intricate network of interactions between the iron, 2-oxoglutarate, and ectoine ligands and their binding partners within the catalytic core of the ectoine hydroxylase becomes apparent. This network of interactions mediates the positioning of the ectoine substrate in such a way that the C5-carbon atom in the pyrimidine ring of ectoine can be hydroxylated to (4*S*,5*S*)-5-hydroxyectoine. This hydroxylation reaction (Fig. 1) occurs with high region-selectivity and stereo-specificity both *in vivo* (e.g., in *Streptomyces parvulus*) [33] and *in vitro* by the purified *V. salexigens* EctD enzyme [20] that we have studied here. Nevertheless, the architecture of the EctD active site (Fig. 5) is probably flexible enough to allow the accommodation and hydroxylation of synthetic ectoine derivatives with reduced or expanded ring sizes [66,67] that then might possess stress-protective and protein-stabilizing properties different, and perhaps even superior, from those exhibited by their natural counterparts ectoine and 5-hydroxyectoine. The residues that we link to the binding of the iron catalyst, the 2-oxoglutarate co-substrate, and the substrate ectoine (Fig. 5) are all highly conserved among members of the EctD family of dioxygenases (Supplementary Fig. S1), suggesting that the data presented here for the EctD protein from the moderate halophile *V. salexigens* will have general functional implications for ligand binding and catalysis by ectoine hydroxylases.

Materials and Methods

Chemicals

Ectoine and 5-hydroxyectoine were kindly provided from Dr. Thomas Lentzen and Dr. Irina Bagyan (Bitop AG, Witten, Germany). 2-Oxoglutarate (disodium salt) was purchased from Sigma-Aldrich (St. Louis, MO, USA). Anhydrotetracycline hydrochloride (AHT), desthiobiotin, and Strep-Tactin Superflow chromatography material were obtained from IBA GmbH (Göttingen, Germany).

Culture conditions

E. coli strains were maintained on LB agar plates and were grown in LB liquid medium at 37 °C under aeration [68]. To select for *E. coli* DH5α (Invitrogen, Karlsruhe, Germany) cells harboring pBJ10 (*ectD*$^{+}$)-derived plasmids containing the *V. salexigens ectD* gene [20] and its mutant derivatives, we added ampicillin (100 μg ml^{-1}) to liquid and solid media. Minimal medium A [68] containing 0.5% (w/v) glucose as carbon source, 0.5% (w/v) casamino acids, 1 mM $MgSO_4$, and 3 mM thiamine was used for overproduction of the *V. salexigens* EctD protein and its mutant variants.

Site-directed mutagenesis of the *ectD* gene

To identify amino acids that might be involved in binding of the different substrates of the ectoine hydroxylase, we performed site-directed mutagenesis of the cloned *ectD* gene from *V. salexigens* [20] using the QuikChange Lightning site-directed mutagenesis kit (Stratagene, La Jolla, CA) and custom-synthesized mutagenic primers (typical length: 25 nucleotides) purchased from Biomers (Ulm, Germany). Plasmid pBJ10 [20] was used as the DNA template for the mutagenesis procedure. To ascertain the presence of the desired mutation and the absence of unwanted alterations in the coding region of the recombinant *ectD* gene, we determined the DNA sequence of the entire coding region of each mutant (Eurofins MWG GmbH, Ebersberg, Germany). The following *ectD* variants were generated in the parent plasmid pBJ10: pMP1 (Phe^{143} → Ala [TTT → GCG]), pMP2 (Arg^{259} → Ala [CGT → GCG]), pMP3 (Phe^{143} → Tyr [TTT → TAC]), pMP4 (Arg^{259} → Lys [CGT → AAG]), pMP5 (Arg^{259} → His [CGT → CAC]), pMP6 (Arg^{259} → Gln [CGT → CAA]), pMP7 (Phe^{263} → Ala [TTC → GCC]), pMP8 (Asn^{261} → Ala [AAT → GCT]), pMP9 (Asn^{133} → Ala [AAT → GCT]), pMP10 (Arg^{131} → Ala [CGA → GCA]), pMP11 (Trp^{152} → Ala [TGG → GCG]), pMP12 (Gln^{129} → Ala [CAA → GCA]), pMP13 (Ser^{167} → Ala [TCT → GCT]), pMP14 (Phe^{95} → Ala [TTT → GCG]), pMP15 (Phe^{143} → Trp [TTT → TGG]), pMP17 (His^{146} → Ala [CAT → GCT]), pMP18 (Asp^{148} → Glu [GAT → GAG]), pMP19 (Asp^{148} → Ala [GAT → GCT]), pMP20 (His^{248} → Ala [CAT → GCT]), pMP21 (Phe^{263} → Tyr [TTC → TAC]), pMP22 (Phe^{263} → Trp [TTC → TGG]), pMP23 (Trp^{152} → Phe [TGG → TTC]), pMP24 (Trp^{152} → Tyr [TGG → TAC]), pMP26 (Ser^{205} → Ala [TCA → GCA]), pMP29 (Pro^{189} → Ala [CCT → GCT]), pMP30 (Ser^{250} → Ala [TCA → GCA]), pWN1 (Phe^{242} → Ala [TTT → GCT]), pWN2 (Phe^{242} → Tyr [TTT → TAT]) pWN3 (Phe^{242} → Trp [TTT → TGG]). pWN4 (Ala^{163} → Cys

[GCC → TGC]), pWN5 (Ser^{244} → Cys [AGT → TGT]), pWN6 (Ala^{163} → Cys [GCC → TGC]; Ser^{244} → Cys [AGT → TGT]), pWN16 (Val^{265} → Ala [GTG → GCG]), pWN17 (Val^{265} → Thr [GTG → ACG]), pWN18 (Val^{265} → Leu [GTG → CTG]), and pWN19 (Ser^{165} → Ala [AGT → GCT]).

Overproduction and purification of recombinant EctD proteins

Plasmid pBJ10 is a derivative of the pASK-IBA3 expression vector (IBA) and harbors the wild-type *ectD* gene from the moderate halophile *V. salexigens* (DSM 11483^T) [20]. In this plasmid, transcription of the otherwise osmotically inducible *ectD* gene [20] is mediated by the anhydrotetracycline (AHT)-inducible *tet*-promoter present on the pASK-IBA3 expression vector; hence, the production of the EctD protein can be triggered by adding the inducer (AHT) of the TetR repressor protein to the culture medium of *E. coli* strains. To permit the purification of the EctD protein by affinity chromatography, we fused the *ectD* coding region present on plasmid pBJ10 at its 3′ end to a short DNA sequence encoding a *Strep*-tag-II peptide. Overexpression of the recombinant *ectD* gene and its mutant derivatives was performed in the host *E. coli* strain DH5α essentially as described by Reuter *et al.* [38]. Typically, 1 l of minimal medium A was inoculated to an OD_{578} of 0.1 from an overnight culture of DH5α (pBJ10 or its mutant derivatives) propagated in the same medium. When the culture (grown at 37 °C) had reached mid-log phase (OD_{578} of about 0.7), the high-level expression of *ectD* was triggered by the addition of AHT (final concentration of 0.2 mg ml^{-1}), the growth temperature was reduced to 35 °C, and the cells were then grown for additional 2 h. The cells were harvested by centrifugation (10 min, 5000 rpm, 4 °C), and the cell pellets were stored at −20 °C until further use. Purification of EctD was attained by affinity chromatography on a Strep-Tactin Superflow column as detailed previously [20]. Similar yields of the wild-type and mutant EctD proteins were obtained by this purification scheme, indicating that the single amino mutations introduced into EctD did not grossly affect protein stability. The purified EctD proteins were shock-frozen in liquid nitrogen and stored at −80 °C until their enzyme activity was determined via an HPLC-based assay [20,26].

EctD enzyme assays

The enzymatic activity of EctD and its mutant derivatives was assayed in 30 µl reaction volumes that typically contained 10 mM TES buffer (pH 7.5), 10 mM 2-oxoglutarate, 6 mM ectoine, 1 mM $FeSO_4$, and various amounts (between 2 and 15 µg) of the recombinantly produced EctD protein. In contrast to previously used assay conditions [20], the reactive oxygen species scavenger catalase was left out. To determine the kinetic parameters of the *V. salexigens* EctD wild-type enzyme and its mutant derivatives, we varied the concentration of ectoine in the assay between 0 and 40 mM. The enzyme reaction mixtures were incubated at 32 °C for 20 min in a thermomixer (Eppendorf, Hamburg, Germany) with vigorous shaking to provide enough oxygen to the O_2-dependent EctD enzyme. The enzyme reactions were stopped by adding 30 µl of 100% acetonitrile to the reaction mixtures, and the samples were immediately centrifuged to remove the denatured EctD protein (10 min, 4 °C, 32,000*g*). Formation of 5-hydroxyectoine from the substrate ectoine was determined by loading 20 µl of the supernatant onto a GROM-SIL 100 Amino-1PR column (125 mm × 4 mm, 3-µm particle size; GROM, Rottenburg-Hailfingen, Germany) attached to a UV-visible detector system (LINEAR UVIS 205; SYKAM, Fürstenfeldbruck, Germany). Ectoine and 5-hydroxyectoine were monitored by their absorbance at 210 nm [20,21] and quantitated with the ChromStar 7.0 software package (SYKAM). The iron content of EctD protein preparations produced in *E. coli* by recombinant techniques was assayed as previously described [69].

Modeling of the flexible loop region (Gly195–Leu211) of the EctD protein

In order to generate a starting structure for MD simulations of the ectoine ligand bound to the EctD enzyme of *V. salexigens*, we took the coordinates of the EctD protein atoms from the PDB entry 3EMR [38]. In the reported EctD structure, residues Gly195 to Leu211 have not been resolved, which indicates the presence of a mobile loop region [38]. In order to generate coordinates of this region by homology modeling, we searched suitable template structures in the PDB by means of a BLAST search. This resulted in PDB entries 2WBO (9% sequence identity according to the Dali Structural Alignment Server), 2WBP (10%), and 2WBQ (10%), that is, crystal structures of the non-heme iron(II) oxygenase VioC from *S. vinaceus* involved in the synthesis of the antibiotic viomycin [65]. As may be expected from the low sequence identity of the EctD and VioC proteins, we did not succeed in generating a multiple sequence alignment of the template and target sequences that resulted in meaningful models, neither by way of automatic alignments with ClustalW or Modeller nor by way of a structural alignment using the Dali Server, nor by combining automatic and manual alignment. Thus, we resorted to the modeling of the missing loop residues of the EctD protein with the automatic loop refinement of Modeller [70]. Initially, Modeller's automodel class was used with a 3-fold slow refinement for 300 iterations with a molpdf ("energy" computed by Modeller's objective function) limit of 106; subsequently, the loops in each model were refined using the dope_loopmodel [71] class with the slow md_level, giving 25 models in total.

Each model was processed using Anolea [72], Dope [73], and PROCHECK [74] to determine structure quality. Anolea provides a residue-wise score; Dope, a residue-wise score and a global score. For PROCHECK a residue-wise score was calculated as the maximum deviation of dihedral angles, bond angles, and bond lengths from optimal values for that residue. The residue-wise scores are smoothed over a window of 9 residues, and "Ensemble *Z*-scores" are computed by comparing these scores for each residue in a model to corresponding residues in all models; averaging over all residues yield the global "Ensemble Scores". Finally, a "Composite Ensemble Score" is computed as the average of the Anolea, Dope, and PROCHECK scores. The "best" model was defined as the one with the lowest "Composite Ensemble Score". This model is characterized by an average Anolea energy per residue of −10.6, a global Dope score of −1.8903,

and a Ramachandran plot with 90.8% of the residues in the core region, 8.8% in the allowed region, 0.0% in the generously allowed region, and 0.4% in the disallowed region. This model was then used for further generation of the *in silico* EctD/ectoine complex structure. In this model, the loop is not occluding the binding site region.

MD simulations

MD simulations were performed with the AMBER 11 suite of programs [75] using the GPU accelerated code of *pmemd* [76]. For the EctD protein, the force field by Cornell *et al.* [77] was used with modifications suggested by Simmerling *et al.* [78]. For ectoine and 2-oxoglutarate, the general amber force field was used [79]; partial charges were generated by the RESP procedure [80]. The catalytic metal ion was modeled with parameters for Mg^{2+}. The two His residues coordinating the metal ion were protonated at $N^{\delta 1}$ (His248) and $N^{\varepsilon 2}$ (His146); all other His residues were protonated at $N^{\varepsilon 2}$. Visual inspection confirmed that the metal ion remained in the binding site throughout the simulations. The structures were placed into an octahedral periodic box of TIP3P water molecules [81] such that the distance between the edges of the water box and the closest atom of the EctD protein was at least 11 Å. This resulted in a system size of ~40,000 atoms. The particle mesh Ewald method [82] was used to treat long-range electrostatic interactions. Bond lengths involving bonds to hydrogen atoms were constrained using SHAKE [83]. The timestep for all MD simulations was 2 fs, with a direct-space, non-bonded cutoff of 8 Å.

In order to relax the modeled loop, we energy minimized the *apo* structure of EctD by 500 steps of conjugate gradient minimization, applying harmonic restraints with force constants of at least 5 kcal mol^{-1} $Å^{-2}$ to all solute atoms. Applying harmonic restraints with force constants of 5 kcal mol^{-1} $Å^{-2}$ to all solute atoms, we carried out NVT-MD for 50 ps, during which the system was heated from 100 K to 300 K. Subsequent NPT-MD was used for 150 ps to adjust the solvent density. Finally, the force constants of the harmonic restraints on solute atom positions were gradually reduced to 1 kcal mol^{-1} $Å^{-2}$ during 50 ps of NVT-MD, followed by 10 ns of NVT-MD with all atoms, but the modeled loop restrained in that way, again followed by 200 ns of NVT-MD without any restraints.

The structure obtained at the end of this procedure was aligned to the PDB entry 3EMR. The coordinates of the catalytic metal ion were copied, and ectoine was then placed in the presumed active-site region of the EctD protein in a way that its carboxylate group forms a salt bridge with residue Arg131 and that its methyl group points toward residue Phe263. This orientation was chosen because the ectoine ligand assumes a similar configuration with the binding sites of the high-affinity ligand-binding proteins EhuB (PDB accession code: 2Q88) [57], UehA (PDB accession code: 3FXB) [59], and TeaA (PDB accession code: 2VPN) [60]. Finally, 2-oxoglutarate was placed in the structure using the program COOT [84]; this co-substrate remained largely in its predicted binding site throughout the simulations (RMSD mostly around 3 Å; Supplementary Fig. S3b). This structure was again thermalized following the procedure described above for relaxing the loop region except that now the force constants of the harmonic restraints on solute atom positions were gradually reduced to zero during 100 ps of NVT-MD.

At this point, three NVT-MD simulations for production were spawned, at 300.0, 300.1, and 300.2 K, respectively. Trajectories of 300 ns length were separately analyzed with the program *ptraj*, with conformations extracted every 20 ps. The core region of the protein was defined to be all residues but those of the loop and the mobile C-terminus.

Database searches and alignment of EctD amino acid sequences

Proteins that are homologous to the EctD protein from *V. salexigens* were searched for via the Web server of the DOE Joint Genome Institute[‡]. The amino acid sequences of the retrieved EctD-type proteins were aligned using ClustalW [48].

Preparation of figures

Figures of the experimentally determined structure of the *V. salexigens* EctD protein (PDB accession code: 3EMR) [38] or of conformations extracted from the MD trajectories were prepared using the PyMOL software package[§]. Graphs were prepared with gnuplot[‖].

Acknowledgements

We thank Jochen Sohn for excellent technical assistance and Georg Lentzen and Irina Bagyan (Bitop AG) for their kind gifts of ectoine and 5-hydroxyectoine. We greatly appreciate the expert help of Vickie Koogle in the language editing of our manuscript and thank Lutz Schmitt for helpful discussions. N.W. and E.B. are very grateful to Rolf Thauer for his continued support. E.B. greatly valued the hospitality and kind support of Tom Silhavy during a sabbatical at the Department of Molecular Biology of Princeton University (Princeton, NJ, USA). We thank the anonymous reviewers of our manuscript for their thoughtful advice. Financial support for this study was generously provided by grants from the Deutsche Forschungsgemeinschaft through the SFB 987 (to E.B.), by the LOEWE program of the state of Hessen (via the Center for Synthetic Microbiology, Marburg) (to E.B. and W.B.), the Max Planck Institute for Terrestrial Microbiology (to W.B. and N.W.), by the Fonds der Chemischen Industrie (to E.B.), by the initiative “Fit for Excellence” of the Heinrich-Heine-University of Düsseldorf (to H.G., A.H., and S.H.J.S.), and by the Ministry of Innovation, Science, and Research of North Rhine-Westphalia and Heinrich-Heine-University, Düsseldorf, for a scholarship to D.M. within the CLIB-Graduate Cluster Industrial Biotechnology. Computational support for the work of H.G. was provided by the “Zentrum für Informations und

Medientechnologie" at the Heinrich-Heine-University. N.W. is a member of the International Max Planck Research School for Environmental, Cellular and Molecular Microbiology (Marburg) and gratefully acknowledges its support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jmb.2013.10.028.

Received 30 April 2013;
Received in revised form 29 September 2013;
Accepted 23 October 2013
Available online 30 October 2013

Keywords:
dioxygenase;
compatible solute;
osmotic stress;
molecular modeling

† N.W. and M.P. contributed equally to this work.

‡ http://www.jgi.doe.gov/
§ http://www.pymol.org
‖ http://www.gnuplot.info

Abbreviations used:
DSBH, double-stranded β-helix; MD, molecular dynamics.

References

[1] Bremer E, Krämer R. Coping with osmotic challenges: osmoregulation through accumulation and release of compatible solutes. In: Storz G, Hengge-Aronis R, editors. Bacterial stress responses. Washington: ASM Press; 2000. p. 79–97.
[2] Wood JM. Bacterial osmoregulation: a paradigm for the study of cellular homeostasis. Annu Rev Microbiol 2011;65:215–38.
[3] Kempf B, Bremer E. Uptake and synthesis of compatible solutes as microbial stress responses to high osmolality environments. Arch Microbiol 1998;170:319–30.
[4] Wood JM, Bremer E, Csonka LN, Krämer R, Poolman B, van der Heide T, et al. Osmosensing and osmoregulatory compatible solute accumulation by bacteria. Comp Biochem Physiol Part A Mol Integr Physiol 2001;130:437–60.
[5] Yancey PH. Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. J Exp Biol 2005;208:2819–30.
[6] Street TO, Bolen DW, Rose GD. A molecular mechanism for osmolyte-induced protein stability. Proc Natl Acad Sci U S A 2006;103:13997–4002.
[7] Ignatova Z, Gierasch LM. Inhibition of protein aggregation *in vitro* and *in vivo* by a natural osmoprotectant. Proc Natl Acad Sci U S A 2006;103:13357–61.
[8] Bourot S, Sire O, Trautwetter A, Touze T, Wu LF, Blanco C, et al. Glycine betaine-assisted protein folding in a *lysA* mutant of *Escherichia coli*. J Biol Chem 2000;275:1050–6.
[9] Street TO, Krukenberg KA, Rosgen J, Bolen DW, Agard DA. Osmolyte-induced conformational changes in the Hsp90 molecular chaperone. Protein Sci 2010;19:57–65.
[10] Manzanera M, Garcia de Castro A, Tondervik A, Rayner-Brandes M, Strom AR, Tunnacliffe A. Hydroxyectoine is superior to trehalose for anhydrobiotic engineering of *Pseudomonas putida* KT2440. Appl Environ Microbiol 2002;68:4328–33.
[11] Diamant S, Eliahu N, Rosenthal D, Goloubinoff P. Chemical chaperones regulate molecular chaperones *in vitro* and in cells under combined salt and heat stresses. J Biol Chem 2001;276:39586–91.
[12] Chattopadhyay MK, Kern R, Mistou MY, Dandekar AM, Uratsu SL, Richarme G. The chemical chaperone proline relieves the thermosensitivity of a *dnaK* deletion mutant at 42 degrees C. J Bacteriol 2004;186:8149–52.
[13] da Costa MS, Santos H, Galinski EA. An overview of the role and diversity of compatible solutes in Bacteria and Archaea. Adv Biochem Eng Biotechnol 1998;61:117–53.
[14] Lippert K, Galinski EA. Enzyme stabilization by ectoine-type compatible solutes: protection against heating, freezing and drying. Appl Microbiol Biotechnol 1992;37:61–5.
[15] Lentzen G, Schwarz T. Extremolytes: natural compounds from extremophiles for versatile applications. Appl Microbiol Biotechnol 2006;72:623–34.
[16] Pastor JM, Salvador M, Argandona M, Bernal V, Reina-Bueno M, Csonka LN, et al. Ectoines in cell stress protection: uses and biotechnological production. Biotechnol Adv 2010;28:782–801.
[17] Graf R, Anzali S, Buenger J, Pfluecker F, Driller H. The multifunctional role of ectoine as a natural cell protectant. Clin Dermatol 2008;26:326–33.
[18] Schwibbert K, Marin-Sanguino A, Bagyan I, Heidrich G, Lentzen G, Seitz H, et al. A blueprint of ectoine metabolism from the genome of the industrial producer *Halomonas elongata* DSM 2581^T. Environ Microbiol 2011;13:1973–94.
[19] Bursy J, Kuhlmann AU, Pittelkow M, Hartmann H, Jebbar M, Pierik AJ, et al. Synthesis and uptake of the compatible solutes ectoine and 5-hydroxyectoine by *Streptomyces coelicolor* A3(2) in response to salt and heat stresses. Appl Environ Microbiol 2008;74:7286–96.
[20] Bursy J, Pierik AJ, Pica N, Bremer E. Osmotically induced synthesis of the compatible solute hydroxyectoine is mediated by an evolutionarily conserved ectoine hydroxylase. J Biol Chem 2007;282:31147–55.
[21] Kuhlmann AU, Bremer E. Osmotically regulated synthesis of the compatible solute ectoine in *Bacillus pasteurii* and related *Bacillus* spp. Appl Environ Microbiol 2002;68:772–83.
[22] Galinski EA, Pfeiffer HP, Trüper HG. 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 1985;149:135–9.
[23] Louis P, Galinski EA. Characterization of genes for the biosynthesis of the compatible solute ectoine from *Marinococcus halophilus* and osmoregulated expression in *Escherichia coli*. Microbiology 1997;143:1141–9.
[24] Reshetnikov AS, Khmelenina VN, Mustakhimov II, Kalyuzhnaya M, Lidstrom M, Trotsenko YA. Diversity and phylogeny of the ectoine biosynthesis genes in aerobic, moderately halophilic methylotrophic bacteria. Extremophiles 2011;15:653–63.

[25] Jebbar M, von Blohn C, Bremer E. Ectoine functions as an osmoprotectant in *Bacillus subtilis* and is accumulated via the ABC-transport system OpuC. FEMS Microbiol Lett 1997;154:325–30.
[26] Kuhlmann AU, Bursy J, Gimpel S, Hoffmann T, Bremer E. Synthesis of the compatible solute ectoine in *Virgibacillus pantothenticus* is triggered by high salinity and low growth temperature. Appl Environ Microbiol 2008;74:4560–3.
[27] Garcia-Estepa R, Argandona M, Reina-Bueno M, Capote N, Iglesias-Guerra F, Nieto JJ, et al. The *ectD* gene, which is involved in the synthesis of the compatible solute hydroxyectoine, is essential for thermoprotection of the halophilic bacterium *Chromohalobacter salexigens*. J Bacteriol 2006;188:3774–84.
[28] Kuhlmann AU, Hoffmann T, Bursy J, Jebbar M, Bremer E. Ectoine and hydroxyectoine as protectants against osmotic and cold stress: uptake through the SigB-controlled betaine-choline-carnitine transporter-type carrier EctT from *Virgibacillus pantothenticus*. J Bacteriol 2011;193:4699–708.
[29] Stöveken N, Pittelkow M, Sinner T, Jensen RA, Heider J, Bremer E. A specialized aspartokinase enhances the biosynthesis of the osmoprotectants ectoine and hydroxyectoine in *Pseudomonas stutzeri* A1501. J Bacteriol 2011;193:4456–68.
[30] Lo CC, Bonner CA, Xie G, D'Souza M, Jensen RA. Cohesion group approach for evolutionary analysis of aspartokinase, an enzyme that feeds a branched network of many biochemical pathways. Microbiol Mol Biol Rev 2009;73:594–651.
[31] Ono H, Sawada K, Khunajakr N, Tao T, Yamamoto M, Hiramoto M, et al. Characterization of biosynthetic enzymes for ectoine as a compatible solute in a moderately halophilic eubacterium, *Halomonas elongata*. J Bacteriol 1999;181:91–9.
[32] Inbar L, Lapidot A. The structure and biosynthesis of new tetrahydropyrimidine derivatives in actinomycin D producer *Streptomyces parvulus*. Use of ^{13}C- and ^{15}N-labeled L-glutamate and ^{13}C and ^{15}N NMR spectroscopy. J Biol Chem 1988;263:16014–22.
[33] Inbar L, Frolow F, Lapidot A. The conformation of new tetrahydropyrimidine derivatives in solution and in the crystal. Eur J Biochem 1993;214:897–906.
[34] Borges N, Ramos A, Raven ND, Sharp RJ, Santos H. Comparative study of the thermostabilizing properties of mannosylglycerate and other compatible solutes on model enzymes. Extremophiles 2002;6:209–16.
[35] Ablinger E, Hellweger M, Leitgeb S, Zimmer A. Evaluating the effects of buffer conditions and extremolytes on thermostability of granulocyte colony-stimulating factor using high-throughput screening combined with design of experiments. Int J Pharm 2012;436:744–52.
[36] Van-Thuoc D, Hashim SO, Hatti-Kaul R, Mamo G. Ectoine-mediated protection of enzyme from the effect of pH and temperature stress: a study using *Bacillus halodurans* xylanase as a model. Appl Microbiol Biotechnol 2013;97:6271–8.
[37] Prabhu J, Schauwecker F, Grammel N, Keller U, Bernhard M. Functional expression of the ectoine hydroxylase gene (*thpD*) from *Streptomyces chrysomallus* in *Halomonas elongata*. Appl Environ Microbiol 2004;70:3130–2.
[38] Reuter K, Pittelkow M, Bursy J, Heine A, Craan T, Bremer E. Synthesis of 5-hydroxyectoine from ectoine: crystal structure of the non-heme iron(II) and 2-oxoglutarate-dependent dioxygenase EctD. PLoS One 2010;5:e10647.
[39] Hausinger RP. FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes. Crit Rev Biochem Mol Biol 2004;39:21–68.
[40] Aik W, McDonough MA, Thalhammer A, Chowdhury R, Schofield CJ. Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. Curr Opin Struct Biol 2012;22:691–700.
[41] Prescott AG, Lloyd MD. The iron(II) and 2-oxoacid-dependent dioxygenases and their role in metabolism. Nat Prod Rep 2000;17:367–83.
[42] Straganz GD, Nidetzky B. Variations of the 2-His-1-carboxylate theme in mononuclear non-heme FeII oxygenases. ChemBioChem 2006;10:1536–48.
[43] Hangasky JA, Taabazuing CY, Valliere MA, Knapp MJ. Imposing function down a (cupin)-barrel: secondary structure and metal stereochemistry in the alphaKG-dependent oxygenases. Metallomics 2013;5:287–301.
[44] Knauer SH, Hartl-Spiegelhauer O, Schwarzinger S, Hanzelmann P, Dobbek H. The Fe(II)/alpha-ketoglutarate-dependent taurine dioxygenases from *Pseudomonas putida* and *Escherichia coli* are tetramers. FEBS J 2012;279:816–31.
[45] Grzyska PK, Appelman EH, Hausinger RP, Proshlyakov DA. Insight into the mechanism of an iron dioxygenase by resolution of steps following the FeIV = HO species. Proc Natl Acad Sci U S A 2010;107:3982–7.
[46] Riggs-Gelasco PJ, Price JC, Guyer RB, Brehm JH, Barr EW, Bollinger JM, et al. EXAFS spectroscopic evidence for an Fe = O unit in the Fe(IV) intermediate observed during oxygen activation by taurine:alpha-ketoglutarate dioxygenase. J Am Chem Soc 2004;126:8108–9.
[47] Schwarz H. Chemistry with methane: concepts rather than recipes. Angew Chem Int Ed Engl 2011;50:10096–115.
[48] Thompson JD, Plewniak F, Thierry JC, Poch O. DbClustal: rapid and reliable global multiple alignments of protein sequences detected by database searches. Nucleic Acids Res 2000;28:2919–26.
[49] Saum SH, Müller V. Growth phase-dependent switch in osmolyte strategy in a moderate halophile: ectoine is a minor osmolyte but major stationary phase solute in *Halobacillus halophilus*. Environ Microbiol 2008;10:716–26.
[50] Vargas C, Argandona M, Reina-Bueno M, Rodriguez-Moya J, Fernandez-Aunion C, Nieto JJ. Unravelling the adaptation responses to osmotic and temperature stress in *Chromohalobacter salexigens*, a bacterium with broad salinity tolerance. Saline Systems 2008;4:14.
[51] Holm L, Rosenström P. Dali server: conservation mapping in 3D. Nucleic Acids Res 2010;38:W545–9.
[52] McDonough MA, Kavanagh KL, Butler D, Searls T, Oppermann U, Schofield CJ. Structure of human phytanoyl-CoA 2-hydroxylase identifies molecular mechanisms of Refsum disease. J Biol Chem 2005;280:41101–10.
[53] Schofield CJ, McDonough MA. Structural and mechanistic studies on the peroxisomal oxygenase phytanoyl-CoA 2-hydroxylase (PhyH). Biochem Soc Trans 2007;35:870–5.
[54] Zhang Z, Kochan GT, Ng SS, Kavanagh KL, Oppermann U, Schofield CJ, et al. Crystal structure of PHYHD1A, a 2OG oxygenase related to phytanoyl-CoA hydroxylase. Biochem Biophys Res Commun 2011;408:553–8.
[55] Ziegler C, Bremer E, Krämer R. The BCCT family of carriers: from physiology to crystal structure. Mol Microbiol 2010;78:13–34.
[56] Bremer E. A look into the aromatic cage. Environ Microbiol 2011;3:1–5.

[57] Hanekop N, Höing M, Sohn-Bösser L, Jebbar M, Schmitt L, Bremer E. Crystal structure of the ligand-binding protein EhuB from *Sinorhizobium meliloti* reveals substrate recognition of the compatible solutes ectoine and hydroxyectoine. J Mol Biol 2007;374:1237–50.
[58] Du Y, Shi WW, He YX, Yang YH, Zhou CZ, Chen Y. Structures of the substrate-binding protein provide insights into the multiple compatible solute binding specificities of the *Bacillus subtilis* ABC transporter OpuC. Biochem J 2011;436:283–9.
[59] Lecher J, Pittelkow M, Zobel S, Bursy J, Bonig T, Smits SH, et al. The crystal structure of UehA in complex with ectoine-A comparison with other TRAP-T binding proteins. J Mol Biol 2009;389:58–73.
[60] Kuhlmann SI, Terwisscha van Scheltinga AC, Bienert R, Kunte HJ, Ziegler C. 1.55 Å structure of the ectoine binding protein TeaA of the osmoregulated TRAP-transporter TeaABC from *Halomonas elongata*. Biochemistry 2008;47:9475–85.
[61] Dougherty DA. Cation-pi interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. Science 1996;271:163–8.
[62] Shan Y, Kim ET, Eastwood MP, Dror RO, Seeliger MA, Shaw DE. How does a drug molecule find its target binding site? J Am Chem Soc 2011;133:9181–3.
[63] Gohlke H, Hergert U, Meyer T, Mulnaes D, Grieshaber MK, Smits SH, et al. Binding region of alanopine dehydrogenase predicted by unbiased molecular dynamics simulations of ligand diffusion. J Chem Inf Model 2013;53: 2493–8.
[64] Strieker M, Kopp F, Mahlert C, Essen LO, Marahiel MA. Mechanistic and structural basis of stereospecific Cbeta-hydroxylation in calcium-dependent antibiotic, a daptomycin-type lipopeptide. ACS Chem Biol 2007;2:187–96.
[65] Helmetag V, Samel SA, Thomas MG, Marahiel MA, Essen LO. Structural basis for the erythro-stereospecificity of the L-arginine oxygenase VioC in viomycin biosynthesis. FEBS J 2009;276:3669–82.
[66] Schnoor M, Voss P, Cullen P, Boking T, Galla HJ, Galinski EA, et al. Characterization of the synthetic compatible solute homoectoine as a potent PCR enhancer. Biochem Biophys Res Commun 2004;322:867–72.
[67] Witt EM, Davies NW, Galinski EA. Unexpected property of ectoine synthase and its application for synthesis of the engineered compatible solute ADPC. Appl Microbiol Biotechnol 2011;91:113–22.
[68] Miller JH. Experiments in molecular genetics. New York: Cold Spring Harbor Laboratory; 1972.
[69] Lovenberg W, Buchanan BB, Rabinowitz JC. Studies on the chemical nature of *Clostridial ferredoxin*. J Biol Chem 1963;238:3899–913.
[70] Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 1993;234:779–815.
[71] Fiser A, Do RK, Sali A. Modeling of loops in protein structures. Protein Sci 2000;9:1753–73.
[72] Melo F, Feytmans E. Assessing protein structures with a non-local atomic interaction energy. J Mol Biol 1998;277:1141–52.
[73] Shen MY, Sali A. Statistical potential for assessment and prediction of protein structures. Protein Sci 2006;15:2507–24.
[74] Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–91.
[75] Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, et al. The Amber biomolecular simulation programs. J Comput Chem 2005;26:1668–88.
[76] Gotz AW, Williamson MJ, Xu D, Poole D, Le Grand S, Walker RC. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized Born. J Chem Theory Comput 2012;8:1542–55.
[77] Cornell WD, Cieplak CI, Bayly IR, Gould IR, Merz KM, Ferguson DM, et al. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc 1995;117:5179–97.
[78] Simmerling C, Strockbine B, Roitberg AE. All-atom structure prediction and folding simulations of a stable protein. J Am Chem Soc 2002;124:11258–9.
[79] Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J Comput Chem 2004;25:1157–74.
[80] Bayly CI, Cieplak P, Cornell WD, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for determining atom-centered charges: the RESP model. J Phys Chem 1993;97:10269–80.
[81] Jorgensen WL, Chandrasekhar J, Madura J, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–35.
[82] Darden T, York D, Pedersen L. Particle mesh Ewald: an $N \cdot \log(N)$ method for Ewald sums in large systems. J Chem Phys 1993;98:10089–92.
[83] Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical-integration of Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. J Comput Phys 1977;23:327–41.
[84] Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr Sect D Biol Crystallogr 2004;60:2126–32.