# JMB

# Crystal Structure of the Ligand-Binding Protein EhuB from *Sinorhizobium meliloti* Reveals Substrate Recognition of the Compatible Solutes Ectoine and Hydroxyectoine

# Nils Hanekop[1], Marina Höing[2], Linda Sohn-Bösser[2], Mohamed Jebbar[3], Lutz Schmitt[1]* and Erhard Bremer[2]

[1]*Institute of Biochemistry, Heinrich Heine University Duesseldorf, Universitaetsstrasse 1, 40225 Duesseldorf, Germany*

[2]*Laboratory for Microbiology, Department of Biology, Philipps University Marburg, Karl-von-Frisch Strasse, 35032 Marburg, Germany*

[3]*UMR-CNRS 6026, Départment Osmorégulation chez les Bactéries, Université de Rennes I, Ave. du Général Leclerc, 35042 Rennes, France*

In microorganisms, members of the binding-protein-dependent ATP-binding cassette transporter superfamily constitute an important class of transport systems. Some of them are involved in osmoprotection under hyperosmotic stress by facilitating the uptake of "compatible solutes". Currently, the molecular mechanisms used by these transport systems to recognize compatible solutes are limited to transporters specific for glycine betaine and proline betaine. Therefore, this study reports a detailed analysis of the molecular principles governing substrate recognition in the Ehu system from *Sinorhizobium meliloti*, which is responsible for the uptake of the compatible solutes ectoine and hydroxyectoine. To contribute to a broader understanding of the molecular interactions underlying substrate specificity, our study focused on the substrate-binding protein EhuB because this protein binds the ligand selectively, delivers it to the translocation machinery in the membrane and is thought to be responsible for substrate specificity. The crystal structures of EhuB, in complex with ectoine and hydroxyectoine, were determined at a resolution of 1.9 Å and 2.3 Å, respectively, and allowed us to assign the structural principles of substrate recognition and binding. Based on these results, site-directed mutagenesis of amino acids involved in ligand binding was employed to address their individual contribution to complex stability. A comparison with the crystal structures of other binding proteins specific for compatible solutes revealed common principles of substrate recognition, but also important differences that might be an adaptation to the nature of the ligand and to the demands on protein affinity imposed by the environment.

© 2007 Published by Elsevier Ltd.

*Keywords:* compatible solutes; osmoprotection; substrate-binding protein; ABC transporter; X-ray crystallography

*Edited by I. B. Holland*

*Corresponding author. E-mail address:
lutz.schmitt@uni-duesseldorf.de.
Present address: M. Jebbar, UMR-CNRS 6197, Laboratory of Extreme Environments Microbiology, European Institute for Marine Studies, University of West Brittany, Technopole Brest-Iroise, 29280 Plouzané, France.

Abbreviations used: ABC, ATP-binding cassette; MAD, multiple anomalous dispersion; PEG, polyethylene glycol; PDB, Protein Data Bank; SeMet, selenomethionine; SBP, substrate-binding protein.

## Introduction

In their habitats, microorganisms have to face diverse environmental conditions, which are frequently subject to sudden changes. Therefore, microorganisms have evolved complex strategies to sense and respond to changes in their environment. One of the most important parameters affecting the growth and survival of microorganisms is the osmolality of the external medium, because the level of cellular water is solely determined by osmotic processes due to the lack of systems for active water transport in most Bacteria

and Archaea.[1,2] It is therefore crucial for micro-organisms to control the net flux of water across the plasma membrane by actively adjusting the pool of osmotically active solutes in the cytoplasm to keep the level of cellular water, as well as turgor, within physiologically acceptable limits. To counteract the loss of cell water under hyperosmotic conditions, many microorganisms amass large quantities of a particular class of organic osmolytes, the so-called "compatible solutes"[1,3,4]—a process referred to as osmoprotection. These compatible solutes can be accumulated to high intracellular concentrations by *de novo* synthesis or uptake from the environment without disturbing vital cellular functions.[1,5,6] Although compatible solutes are used in all three kingdoms of life to cope with high-osmolality environments, they belong to only a few classes of compounds, reflecting the fundamental restraints on these kinds of solutes that are highly compatible with macromolecular and cellular functions.[1,4,7]

Besides their role in regulating the content of cell water and turgor, compatible solutes also stabilize the native conformation of proteins upon thermal or high-ionic-strength stress both *in vitro*[8,9] and *in vivo*.[10] The exact mechanism that compatible solutes employ to achieve this effect is not entirely under-stood but is generally explained in terms of the "preferential exclusion model."[11] This hypothesis explains the stabilization of protein structures and conformation by compatible solutes with their exclusion from the immediate hydration shell of proteins due to unfavorable interactions with the protein surface.[12,13] This nonhomogeneous distribu-tion of compatible solutes results in a thermody-namic force that minimizes the surface of the protein and the amount of excluded water, stabilizing the native conformation of the protein.

In case of hyperosmotic stress, many micro-organisms rely on the uptake of compatible solutes from their environment using dedicated transport systems.[1,14] One example of such a high-affinity transporter involved in osmoprotection is the Ehu system from the Gram-negative soil bacterium *Sinorhizobium meliloti*, which belongs to the bind-ing-protein-dependent ATP-binding cassette (ABC) transporter superfamily.[15] The Ehu system accom-plishes the uptake of the tetrahydropyrimidines ectoine and hydroxyectoine (see insets in Fig. 1), which are synthesized by various halophilic and halotolerant bacteria in response to high salinity.[16–18] It is composed of the cytoplasmic-membrane-asso-ciated ATPase EhuA, the membrane-spanning per-meases EhuC and EhuD, and the substrate-binding protein (SBP) EhuB. In *S. meliloti*, the Ehu system accounts for 95% of all ectoine uptake activity and is absolutely required for osmoprotection in the pre-sence of ectoine, although ectoine is catabolized even under hyperosmotic conditions.[19]

In binding-protein-dependent ABC transporters, which are only found in Archaea and Bacteria, SBPs play a central role in substrate translocation, as they bind their ligands selectively and deliver them to the translocation machinery in the membrane.[15,20]
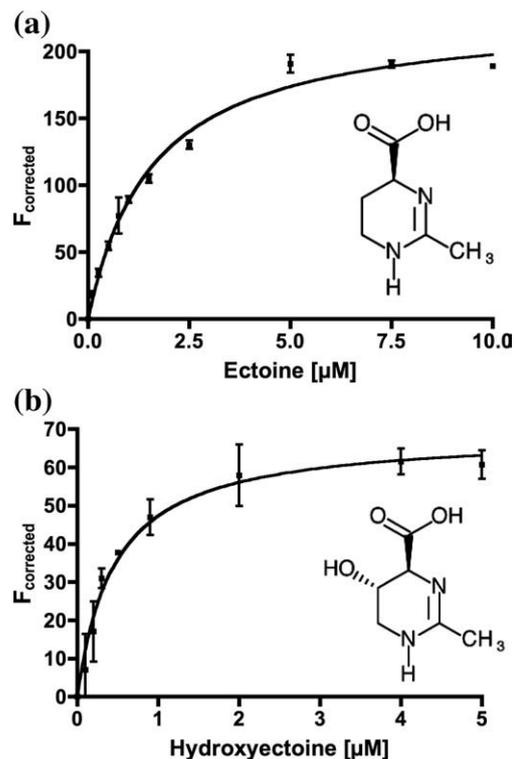


**Fig. 1.** Equilibrium binding titration experiments with ectoine (a) or hydroxyectoine (b). The chemical structures of the two compatible solutes are shown in the inset. Assays were performed and data were analyzed as outlined in Materials and Methods.

Therefore, these proteins are thought to ensure substrate specificity and directionality of transport in these systems.[21] In Gram-negative bacteria, SBPs reside in the periplasmic space in the form of soluble proteins. In contrast, SBPs of Gram-positive Bacteria and Archaea are normally attached to the plasma membrane via an N-terminal lipid anchor[21–23] or are occasionally fused to the membrane-spanning permease.[24]

Structural studies on various SBPs revealed that they can be subdivided into two globular domains or lobes of similar topology, which are connected by one, two or three segments of the polypeptide chain, forming a hinge.[25] A detailed analysis of SBPs in their substrate-bound state showed that in the so-called closed conformation, the two globular do-mains are in close contact, forming a rigid structure with the bound substrate buried in the cleft between the two lobes.[25] In contrast, crystal-structure deter-mination of SBPs in their ligand-free open con-formation revealed structures with the two globular domains separated from each other.[26,27]

In case of SBPs specific for compatible solutes, three different systems have so far been studied on a structural level, all of which are components of glycine-betaine and proline-betaine transport sys-tems.[20,28,29] Interestingly, all these proteins (irre-spective of their archeal or bacterial origin) reveal similar solutions for binding compatible solutes that

are normally excluded from protein surfaces. In all the systems studied, cation–π interactions between a special set of aromatic residues and the delocalized positive charge of the substrate make a major contribution to ligand binding. Substrate binding is further stabilized via salt bridges and/or hydrogen bonds of the proteins with the negatively charged carboxylate group of the ligands. Despite employing similar mechanisms for ligand binding, each protein analyzed has its own architecture of the substrate-binding site. In the SBPs ProX from *Escherichia coli*[28] and OpuAC from *Bacillus subtilis*,[20] three tryptophan residues play a central role in ligand binding. Together, these residues form a binding site with a negative surface potential that is tailored to fit the delocalized positive charge of the substrates. In ProX from *E. coli*, the three tryptophan residues are arranged in a "Trp box," whereas they form a "Trp prism" in OpuAC from *B. subtilis*. Ligand binding is further facilitated by the formation of salt bridges and/or hydrogen bonds with the negatively charged carboxylate group of the substrates. Similar to the two other systems studied, substrate binding in ProX from *Archaeoglobus fulgidus*[28] mainly relies on four tyrosine residues and a main-chain carbonyl group that together form an aromatic box that perfectly fits the quaternary amine of the substrates. Moreover, salt bridges and hydrogen bonds with the negatively charged carboxylate group of the substrates assist ligand binding in this system.

To extend our knowledge on the molecular mechanisms of substrate specificity in binding-protein-dependent ABC transporters participating in the uptake of compatible solutes beyond systems translocating glycine betaine and proline betaine,[20,28,29] we analyzed the ectoine/hydroxyectoine-binding protein EhuB from *S. meliloti* in terms of X-ray crystallography combined with site-directed mutagenesis. The rational behind this study was to analyze the molecular basis for substrate specificity in SBPs binding compatible solutes other than glycine betaine and proline betaine, because so far the molecular mechanisms used to achieve substrate specificity in these systems have remained unclear. Interestingly, a BLAST search for EhuB homologues in the Protein Data Bank (PDB) revealed no detectable homology to the three compatible-solute-specific SBPs of known structure. In the search, the highest matches were obtained for the histidine-binding protein (a 23% identity with an *E*-value of $2 \times 10^{-7}$; Swiss-Prot entry P0AEU0); the lysine-, arginine-, ornithine-binding protein (a 25% identity with an *E*-value of $1 \times 10^{-6}$; Swiss-Prot entry P02911); and the glutamine-binding protein (a 24% identity with an *E*-value of $3 \times 10^{-4}$; Swiss-Prot entry P0AEQ3). Unfortunately, the observed similarity was too low to draw any conclusions regarding the molecular basis of ligand binding in EhuB.

Therefore, this study reports the crystal structure of the SBP EhuB from *S. meliloti* in complex with its substrates ectoine and hydroxyectoine at a resolution of 1.9 Å and 2.3 Å, respectively. This enabled us to determine residues involved in ligand binding

and to analyze their individual contribution to substrate binding by site-directed mutagenesis. Taken together, these results reveal the molecular mechanisms of ligand binding in this SBP, which, in combination with the results of other studies,[20,28,29] significantly contributes to the understanding of substrate specificity in binding-protein-dependent ABC transporters involved in osmoprotection.

## Results and Discussion

EhuB was heterologously overexpressed and purified as described in Materials and Methods. The functionality of EhuB, which eluted as a single peak corresponding to monomeric species in a size-exclusion chromatography experiment (data not shown), was verified by intrinsic tyrosine fluorescence.[30] EhuB does not contain tryptophans, but six tyrosine residues, which are evenly distributed throughout the sequence. Structure determination of EhuB (see below), in complex with its substrates, revealed that two out of its six tyrosine residues are located within the substrate-binding site (Tyr60) or at the interface of the two domains (Tyr200) and are thus likely to change their fluorescence properties upon ligand binding. Therefore, we rationalized that intrinsic tyrosine fluorescence might be a suitable approach to determine the steady-state affinities of ectoine and hydroxyectoine. As shown in Fig. 1, increasing concentrations of the ligands resulted in a change in intrinsic tyrosine fluorescence, which could be analyzed by a standard 1:1 Langmuir-binding isotherm after background correction. Dissociation constants ($K_d$) were calculated to be $1.6 \pm 0.3$ μM (ectoine) and $0.5 \pm 0.1$ μM (hydroxyectoine), respectively. The $K_d$ value for ectoine determined by intrinsic tyrosine fluorescence is in excellent agreement with the value derived from steady-state equilibrium dialysis measurements ($0.5 \pm 0.2$ μM),[19] while no such data exist for hydroxyectoine. Thus, the steady-state binding experiments demonstrate that the heterologously expressed protein is functional and that EhuB binds its two ligands with high affinity.

### Overall structure of EhuB

The structure of EhuB, in complex with ectoine, was solved at 2.1 Å resolution by multiple anomalous dispersion (MAD) phasing of a selenomethionine (SeMet) derivative crystal. The structures of the native crystals of EhuB, in complex with ectoine (which diffracted to higher resolution) and hydroxyectoine, were subsequently solved by rigid body refinement and molecular replacement at a resolution of 1.9 Å and 2.3 Å, respectively. Data statistics, refinement details and model content are summarized in Table 1. The final model contains all 256 residues plus the last C-terminal residue of the Factor Xa cleavage site, 1 ligand, 4 putative $Cd^{2+}$ and 123 (ectoine) or 57 (hydroxyectoine) water molecules, respectively. The rmsd of the structures of the EhuB/ectoine and EhuB/hydroxyectoine com-

**Table 1.** Data collection and refinement statistics

| | Ectoine | Hydroxyectoine | Ectoine SeMet | | |
| --- | --- | --- | --- | --- | --- |
| | | | Inflection | Maximum | Remote |
| A. *Crystal parameters at 100 K* | | | | | |
| Space group | $P4_32_12$ | $P4_32_12$ | $P4_32_12$ | | |
| $a = b$, $a = c$ [Å] | 57.37, 161.88 | 57.14, 161.58 | 57.41, 162.12 | | |
| $\alpha = \beta = \gamma$ [°] | 90 | 90 | 90 | | |
| | | | | | |
| B. *Data collection and processing* | | | | | |
| Wavelength [Å] | 1.05 | 1.05 | 0.9795 | 0.9791 | 0.95 |
| Resolution [Å] | 20–1.9 (1.95–1.9) | 20–2.3 (2.36–2.3) | 20–2.1 | 20–2.1 | 20–2.1 |
| Mean redundancy | 3.7 | 2.5 | 6.9 | 7.3 | 6.9 |
| Unique reflections | 21,659 | 12,062 | 16,572 | 16,544 | 16,632 |
| Completeness [%] | 97.3 (98.6) | 95.5 (96.3) | 99.9 | 99.9 | 99.9 |
| $I/\sigma$ | 11.4 (3.0) | 13.6 (4.6) | 24.3 | 25.4 | 23.8 |
| $R_{sym}$[a] | 11.8 (39.1) | 5.5 (20.2) | 8.2 | 9.4 | 8.5 |
| | | | | | |
| C. *Refinement* | | | | | |
| $R_f$[b] [%] | 22.8 (26.4) | 23.5 (24.1) | | | |
| $R_{free}$[b] [%] | 24.9 (31.4) | 26.1 (26.7) | | | |
| rmsd bond [Å] | 0.009 | 0.007 | | | |
| rmsd angle [°] | 1.23 | 1.04 | | | |
| Average $B$-factor [Å$^2$] | 28.4 | 34.3 | | | |
| Ramachandran plot [%] | | | | | |
| Most favored | 94.0 | 92.6 | | | |
| Allowed | 6.0 | 7.4 | | | |
| | | | | | |
| D. *Model content* | | | | | |
| Monomers/asymmetric unit | 1 | 1 | | | |
| Protein residues | 0–256 | 0–256 | | | |
| Ligand | One ectoine | One hydroxyectoine | | | |
| Ions | 4 Cd$^{2+}$ | 4 Cd$^{2+}$ | | | |
| Water molecules | 123 | 57 | | | |

Crystal parameters and data collection statistics are derived from SCALEPACK.[46] Refinement statistics were obtained from REFMAC5,[49] and Ramachandran analysis was performed using PROCHECK.[54] Data in parentheses correspond to the highest-resolution shell (1.95–1.90 Å for EhuB/ectoine and 2.36–2.30 Å for EhuB/hydroxyectoine, respectively).
[a] $R_{sym}$ is defined as $R_{sym} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$.
[b] $R_f = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$. $R_{free}$ is calculated as $R_f$, but for 5% randomly chosen reflections that were omitted from all refinement steps.

plex is 0.27 Å over 257 $C_\alpha$ atoms. Thus, the general description of the structure will be restricted to the EhuB/ectoine complex, which is composed of two $\alpha/\beta$ domains connected by a hinge region that is formed by two β-strands. Domain 1 (green in Fig. 2), encompassing residues 1–95 and 203–256 and harboring the N- and C-termini, contains eight α-helices and two additional β-strands packed against a 5-stranded β-sheet. The strand order of the central β-sheet is $\beta_B$, $\beta_A$, $\beta_C$, $\beta_L$ and $\beta_D$, with strand $\beta_L$ running antiparallel (see inset in Fig. 2). Domain 2 (residues 100–197; orange in Fig. 2) is also composed of a 5-stranded β-sheet, against which four α-helices are packed. The strand order is similar to Domain 1, with the order $\beta_J$, $\beta_F$, $\beta_G$, $\beta_I$ and $\beta_H$, with strand $\beta_I$ running antiparallel (see inset in Fig. 2). This clearly documents that EhuB is a Class-II-binding protein.[31] Both domains are connected by two antiparallel β-strands (strands $\beta_E$ and $\beta_K$; gray in Fig. 2), which are connected by a disulfide bond between Cys97 and Cys201. A second disulfide bond between Cys86 and Cys253 is located in Domain 1 and is highlighted in ball-and-stick representation in Fig. 2. Furthermore, two *cis* peptide bonds are present; the first, between the two prolines at positions 22 and 23, is located in Domain 1, while the second, located between A192 and P193, is found in Domain 2. The

importance of the *cis* peptide bonds will be further discussed (see below). The ligands ectoine and hydroxyectoine are bound in a deep cleft located between the two domains. The delocalized positive charge of the ligands interacts via cation–π interactions with three aromatic amino acids that are highlighted in ball-and-stick representation in Fig. 2. This type of interaction has been observed in many proteins such as compatible-solutes-binding proteins (ProX from *E. coli*[28] or *A. fulgidus*,[29] or OpuAC from *B. subtilis*[20]), a phosphocholine transfer protein[32] or an acetylcholine-binding protein.[33]

A DALI search revealed a number of proteins with high structural similarity to EhuB. Among the highest matches was the lysine-, arginine-, ornithine-binding protein [$Z = 25.9$, with 227 of the 239 $C_\alpha$ atoms aligning with an rmsd of 2.2 Å; PDB entry 1LST],[34] while the glutamate receptor 2 fragment (PDB entry 1GR2) displayed a $Z$ score of 19.0 (rmsd = 3.4 Å for 212 out of 249 $C_\alpha$ atoms aligned).[35] Many other hits, including other SBPs or LysR-type regulator proteins, were found. Interestingly, ProX from *E. coli*,[28] which binds the compatible solutes glycine betaine and proline betaine (PDB entry 1R9L), displayed a $Z$ score of only 8.5 (rmsd = 4.4 Å for 182 out of 309 $C_\alpha$ atoms), while the one from *A. fulgidus* displayed a higher similarity ($Z$ score = 10.8,
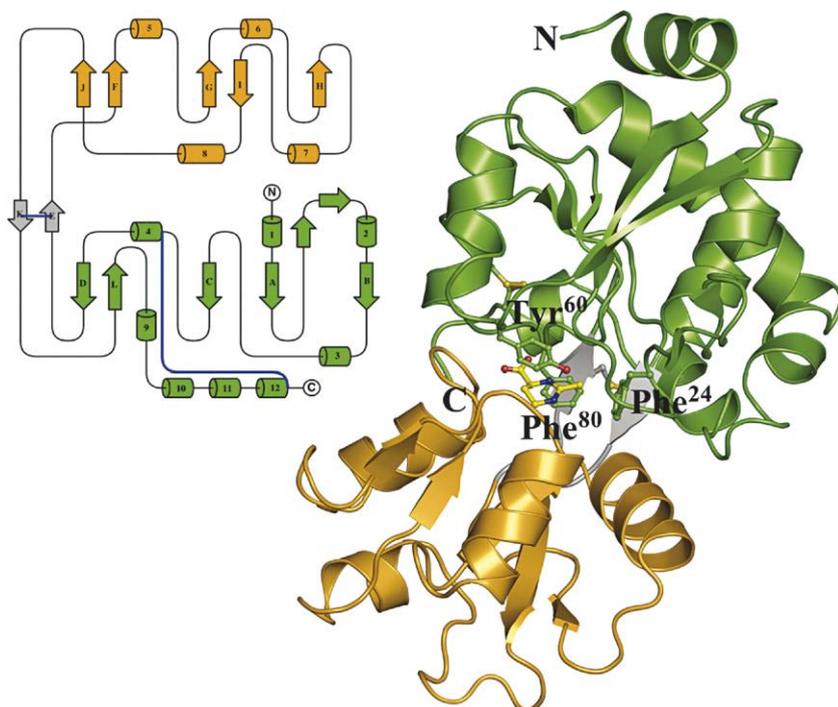
**Fig. 2.** Overall structure of the EhuB/ectoine complex. The two domains of EhuB are colored in green (residues 1–95 and 203–256) and orange (residues 100–197), respectively, and ectoine is shown in yellow. The two β-sheets forming the hinge (residues 96–99 and 198–202) are depicted in gray. N- and C-termini are labeled. Ectoine and the residues (Phe24, Tyr60 and Phe80) forming the aromatic-ligand-binding box are shown in ball-and-stick representation. The two disulfide bridges between residues Cys97/Cys201 and Cys86/Cys253 are highlighted in ball-and-stick representation. The inset shows the topology of the protein. Color coding is derived from the overall structure. The two disulfide bonds are indicated by blue lines.

rmsd = 3.6 Å for 185 out of 265 $C_\alpha$ atoms; PDB entry 1SW1). Despite the fact that these Z scores are still significant, the reduced score and the higher rmsd show that ProX and EhuB, although capable of recognizing compatible solutes that serve osmoprotective purposes, are only distantly related with respect to their structure.

### Metal-ion-binding sites

During the optimization of crystals of EhuB, we noticed that $Cd^{2+}$ was required. In the absence of divalent ions, no crystal formation occurred, and substitution of $Cd^{2+}$ for other divalent heavy metal ions dramatically reduced crystal quality. In most cases, even no crystals were formed in the presence of other metal ions. In the final structure, four heavy atoms that were modeled as $Cd^{2+}$ were observed in the asymmetric unit, all remote from the ligand-

binding site and involved in interactions of protein molecules from different asymmetric units (Fig. 3). Due to the missing functional relevance of the bound ions (see below), their nature was not verified by an anomalous difference electron density map, but the strict requirement for $Cd^{2+}$ in the crystallization trials supports our assumption that the four heavy atoms are indeed $Cd^{2+}$. Since we have not observed any influence of $Cd^{2+}$ on ligand binding or *in vivo* transport, one can exclude any functional importance of these ions. Rather, the EhuB structure provides another example of the importance of cadmium ions during protein crystallization.[36]

### Architecture of the ligand-binding site

An analysis of the ligand-binding site of EhuB, in complex with ectoine and hydroxyectoine (stereoviews of the ligand-binding sites are shown in
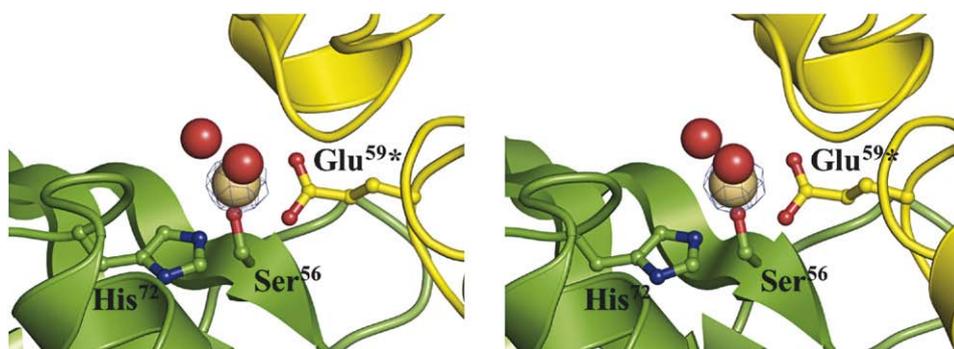


**Fig. 3.** Stereoview of one of the four $Cd^{2+}$-binding sites observed in the crystal structures of EhuB/ectoine and EhuB/hydroxyectoine, respectively. Ligands coordinating the metal ion are shown in ball-and-stick representation. A star highlights the residue, which originates from the symmetry-related mate. $Cd^{2+}$ is shown as a light tan sphere, and the two water molecules are shown as red spheres. The radius of the spheres relates to half of their corresponding van der Waals radius. A $2F_o - F_c$ composite-annealed omit map contoured at 5σ is displayed for the putative cadmium ion.

Fig. 4), reveals that the conformations of the residues forming the substrate-binding site are almost identical. In both structures, the ligands are bound by EhuB employing various modes of interactions. The delocalized positive charge of the ligands is coordinated via cation–π interactions with Phe24, Tyr60 and Phe80. This arrangement is reminiscent of the Trp box and the Trp prism observed in ProX from *E. coli*[28] and OpuAC from *B. subtilis*,[20] respectively. Additionally, a salt bridge is formed between the carboxyl moiety of the ligand and Arg85.

The carboxyl moiety is further stabilized by hydrogen bonds with the backbone amide of Phe80 and Thr133. This suggests that a ligand like in the small molecule structure[37] is bound to the protein in its zwitterionic form. Salt bridges between the imido moiety and Glu21, as well as hydrogen bonds of the former with the carbonyl group of the peptide backbone of Gly78, complete the binding site. In the case of hydroxyectoine (Fig. 4d), additional hydrogen bonds are formed between the hydroxyl group of the ligand and the side chain of Glu134, as well as
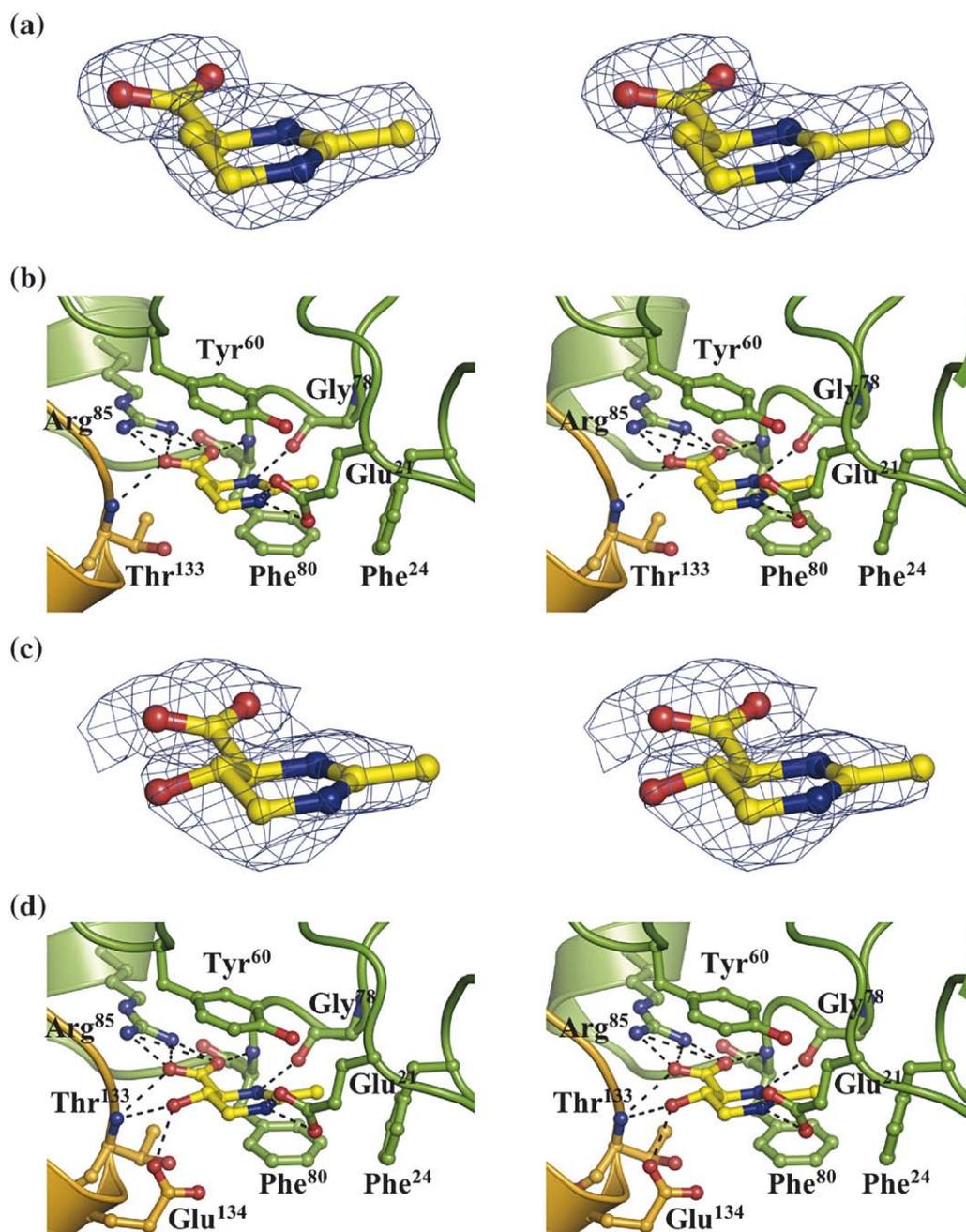


**Fig. 4.** Stereoview of the ligands ectoine (a) and hydroxyectoine (c) and the architecture of the ligand-binding sites of the EhuB/ectoine (b) and EhuB/hydroxyectoine (d) complexes. $2F_o−F_c$ composite-annealed omit maps contoured at 1σ are shown as a blue mesh in (a) and (c). Residues originating from Domain 1 of EhuB are shown in a green ball-and-stick representation, and residues from Domain 2 are shown in an orange ball-and-stick representation. Black dashed lines highlight salt bridges and hydrogen bonds between the ligands and protein residues.

with the backbone amide of Thr133. These additional interactions explain the slightly higher affinity of EhuB to hydroxyectoine (0.5±0.1 μM) compared to ectoine (1.6±0.3 μM). Interestingly, the observed protein–ligand interactions are highly asymmetrical. In the case of the ectoine complex, six (green in Fig. 4b) of the seven amino acids interacting with the ligand originate from Domain 1 of EhuB, while six (green in Fig. 4d) of the eight interacting amino acids stem from Domain 1 in the hydroxyectoine complex. In light of the "Venus-flytrap" mechanism,[38] the asymmetrical distribution of the residues forming the ligand-binding site in EhuB shows that one (ectoine) or three (hydroxyectoine) interactions of the bound ligand with Domain 2 are enough to shift the equilibrium from the open conformation to the closed conformation of the SBP. A brief comparison with other binding proteins reveals that a highly asymmetrical distribution of residues involved in substrate binding can also be found in other SBPs (e.g., six of eight hydrogen bonds of the ligand and the greater proportion of van der Waals contracts in the leucine-, isoleucine-, valine-binding protein originate from Domain 1[36]), but often a more evenly distributed pattern of interactions is observed. In the compatible-solute-binding protein ProX from *E. coli*, for example, the interaction ratio is 2:4 (Domain 1: Domain 2), while it is 3:5 in ProX from *A. fulgidus* and 3:3 in OpuAC from *B. subtilis*.[20,28,29] Additionally, in HisJ and the ribose-binding protein, similar even distributions are present (7:5 in HisJ and 3:4 in the ribose-binding protein).

Similar to the other known crystal structures of binding proteins specific for compatible solutes, no water molecules are observed in the vicinity of the bound ligands. This is in clear contrast to other SBPs such as HisJ,[39] the glutamine-binding protein[40] or the lysine-, arginine-, ornithine-binding protein,[34] which employ water molecules to fine-tune specificity and allow a certain degree of promiscuity in substrate binding. Therefore, the crystal structure of EhuB, the fourth of an SBP operational in osmoprotection, supports the notion that these SBPs contain ligand-binding-site architectures that are optimized for selectivity without relying on additional "coligands" such as water molecules.

## Mutational analysis of the ligand-binding site of EhuB

The crystal structures of the EhuB/ectoine and EhuB/hydroxyectoine complexes provide a molecular framework to determine the amino acids responsible for ligand binding and to understand the principles of recognition and stability. However, it is not possible to deduce their individual contributions. Therefore, we have analyzed these amino acids by site-directed mutagenesis to gain insights into their role in complex stability. A summary of the mutants examined and their influence on the steady-state affinities of the ectoine and hydroxyectoine complexes are summarized in Table 2. An inherent problem of the above mentioned strategy—espe-

**Table 2.** Summary of the mutational studies performed on the residues forming the ligand-binding site of EhuB

| Amino acid | Mutation | Ectoine | | Hydroxyectoine | |
|---|---|---|---|---|---|
| | | $K_d$ (μM) | Binding | $K_d$ (μM) | Binding |
| | Wild type | 1.6±0.3 | = | 0.5±0.1 | = |
| Phe24 | A | 1.3±0.6 | = | 0.4±0.1 | = |
| | Y | <0.02 | +++ | <0.02 | +++ |
| | W | <0.02 | +++ | <0.02 | +++ |
| | D | 6.7±2.3 | = | 2.5±1.5 | = |
| | E | 1.4±0.3 | = | 5.1±0.8 | = |
| Tyr60 | A | 6±1.9 | = | 1.3±0.4 | = |
| | F | 2.1±1.1 | = | 1.1±0.3 | = |
| | W | 0.11±0.02 | ++ | 0.16±0.02 | ++ |
| | D | | --- | | --- |
| | E | | --- | | --- |
| Phe80 | A | | --- | | --- |
| | Y | <0.02 | +++ | <0.02 | +++ |
| | W | <0.02 | +++ | <0.02 | +++ |
| | D | 43±3 | − | 94±23 | − |
| | E | 6.5±0.6 | = | 59±8 | − |
| Glu21 | A | 59±15 | − | 56±15 | − |
| Arg85 | A | | --- | | --- |
| Thr133 | A | | --- | | --- |
| Glu134 | A | 1.2±0.3 | = | | --- |
| Phe24/Tyr60 | A/A | | --- | | --- |

(=) Binding as wild type; (++) increased affinity compared to the wild-type protein; (+++) superbinder; (−) decreased binding affinity compared to the wild-type protein; (---) no binding detectable.
$K_d$ values represent the average value of at least three independent experiments.

cially in tightly constructed binding sites such as the one encountered in EhuB—is that substituting a residue might cause structural changes that in turn can affect the affinity of the mutant protein for its ligands. Despite this drawback, the steady-state affinities of EhuB mutants yield valuable insights into the individual role of the residues forming the ligand-binding site for complex stability (see below), which, in many cases, can be convincingly explained with the help of the structures of the wild-type complexes, although the conclusions have to be supported by further structural studies on the individual mutants if a detailed quantitative analysis is required.

First, we analyzed the aromatic box composed of Phe24, Tyr60 and Phe80 and replaced these amino acids one by one with alanine, tyrosine/phenylalanine, tryptophan, aspartate or glutamate. While alanine in positions 24 and 60 was well tolerated, replacement in position 80 completely abrogated the binding of ectoine and hydroxyectoine, which points to a special role of this aromatic residue in ligand binding. Furthermore, the nondetectable affinity of EhuB for its substrates upon substitution of Phe24 and Tyr60 for alanine suggests that, next to Phe80, at least one additional aromatic residue has to be present in the aromatic box. Taken together, this points to some kind of hierarchy within the aromatic box and to the high importance of Phe80 in stabilizing the bound ligand. Mutation of Phe24 into tyrosine or tryptophan, of Tyr60 into phenylalanine or tryptophan, and of Phe80 into tyrosine or

tryptophan demonstrated a behavior expected for aromatic amino acids involved in cation–π interactions. As pointed out by Ma and Dougherty,[41] the strength of cation–π interactions increases from Phe to Tyr to Trp.[42] Thus, increasing the strength of the cation–π interaction by placing a tyrosine or tryptophan at position 24 or 80, or a tryptophan at position 60, creates a "superbinder" with a 10-fold- (Y60W) to 50-fold-higher affinity (F24Y or F24W). Such a behavior is in clear contrast to ProX from *E. coli*.[28] Here, any aromatic amino acid is tolerated within the Trp box without significantly altering the affinity of the ligand. Mutations of the aromatic residues of EhuB lining the ligand-binding pocket into negatively charged amino acids (Asp and Glu) do not drastically change the thermodynamics of ligand binding at positions 24 and 80. Likely, the salt bridge formed between the carboxylate and the delocalized positive charge of ectoine or hydroxyectoine might be able to compensate for the lost cation–π interaction. However, in position 60, a different outcome is observed with respect to negatively charged amino acids because mutation of Tyr60 into Glu or Asp completely deleted the ability of EhuB to bind ectoine or hydroxyectoine. Based on our structures, this behavior was not unexpected, because the hydroxyl group of Tyr60 forms a hydrogen bond with the carboxyl moiety of Glu21 (a distance of 2.7 Å). Therefore, it seems plausible that the mutation of Tyr60 into Asp or Glu results in a corrupted substrate-binding site incapable of binding the ligand due to the electrostatic repulsion introduced. This interpretation of our data is supported by the substitution of Tyr60 with phenylalanine and alanine, which shows that removing the hydrogen bond between Tyr60 and Glu21 (phenylalanine) and, additionally, the cation–π interaction with the imido moiety of the ligand (alanine) does not drastically change the affinity of the protein for its ligands. Next to cation–π interactions, the importance of electrostatic interactions for substrate binding in EhuB is manifested in the importance of Arg85, which is located at the interface of the two domains and forms a salt bridge with the carboxyl moiety of the bound ligand (a distance of 2.8 Å). The drastic mutation of Arg85 into alanine abrogates ligand binding completely. Moreover, the nondetectable affinity of EhuB for its substrates in the *T133A* mutant and the selective loss of hydroxyectoine binding upon alanine substitution of Glu134, which forms a hydrogen bond with the hydroxyl group of hydroxyectoine (a distance of 2.8 Å), points to an important role of this part of the ligand-binding site for complex formation. This assumption is supported by the hydrogen bond network formed by the residues Thr133, Glu134 and Glu100, which links a peripheral α-helix (Thr133 and Glu134) to the central β-sheet (Glu100) of Domain 2. Nevertheless the tight hydrogen bond network of these residues makes it difficult to predict the outcome of removing one of these residues from the network. In summary and compared with the results of the mutational

studies performed on ProX from *E. coli*,[28] the situation in EhuB seems to be much more complex and composed of an interplay between different amino acids, patterns of interactions and secondary effects, which could not be predicted from the crystal structures alone. Only our detailed mutational analysis has started to shed light on the underlying molecular principles that enable EhuB to bind ectoine and hydroxyectoine with high specificity and affinity.

## A search for structural homologs of EhuB

A BLAST search for homologous bacterial proteins revealed a number of putative ectoine/hydroxyectoine-binding proteins. Selected hits are summarized in Fig. 5. For simplicity, only the important parts of the individual sequences containing those amino acids that participate in ligand binding are given. Evidently, those residues that coordinate the carboxyl moieties of ectoine and hydroxyectoine, the imido moiety of the ligands and the hydroxyl moiety of hydroxyectoine (Glu21, Arg85 and Glu134) are well conserved among all protein sequences analyzed, even in the candidate protein from *Agrobacterium tumefaciens* or *Desulfovibrio vulgaris*, which share only a 34% and a 28% sequence identity, respectively. Those amino acids forming the aromatic box in EhuB (Phe24, Tyr60 and Phe80) are also well conserved with respect to their aromatic nature. Phe24 is either Phe or Tyr in all cases analyzed. Based on our mutational studies, exchange of Phe with Tyr in position 24 would, however, drastically increase the affinity of the ligand (factor of 100 in EhuB; see Table 2). Pairwise comparison of the aromatic amino acids composing the ligand-binding box reveals an interesting pattern. Only in two cases (*Rhodobacter sphaeroides* and *Burkholderia fungorum*) a wild-type-like arrangement is observed, while in all other cases, an increase in ligand affinity is predicted based on the mutational studies summarized in Table 2. For example, the exchange of Phe/Tyr in position 24 in *Desulfitobacterium dehalogenans* would increase the affinity by roughly 100-fold, while the Tyr/Phe exchange in position 60 has hardly an effect on complex stability. Therefore, our crystal structures not only serve as a template for modeling related proteins but also might be useful for precisely determining the amino acids involved in ligand binding.

Two disulfide bridges (Cys97/Cys201 and Cys86/Cys253) and two *cis* peptide bonds (positions 22/23 and 192/193) are present in the EhuB structure. The two cysteines at positions 97 and 201 are conserved in 4 out of 12 cases, while two cysteines are present at positions 86 and 253 in 9 cases of the proteins shown in Fig. 5. This strongly suggests that, in these cases, the disulfide bridges are formed as well. The disulfide bridge between Cys97 and Cys201, conserved in 30% of the sequences shown in Fig. 5, locks the two strands of the hinge region. Thus, this disulfide bridge very likely enhances the stability of the intrastrand

|  |  | Glu²¹ Phe²⁴ | Tyr⁶⁰ |
|---|---|---|---|
| *Sinorhizobium meliloti* EhuB[1] | | N E P P **F** T A | **E Y** G A |
| *Agrobacterium tumefaciens*[2] | (34%) | N E V P Y S Y | P F G T |
| *Bordetella parapertussis*[3] | (39%) | N E A P Y A V | E F G S |
| *Desulfitobacterium dehalogenans*[4] | (43%) | N E K P Y A Y | E F A S |
| *Pseudomonas putida*[5] | (38%) | N E T P **F** A Y | E W G S |
| *Streptomyces coelicolor*[6] | (40%) | G E V P **F** G Y | E F G S |
| *Arthrobacter sp.*[7] | (45%) | N E P P Y T Q | P Y E S |
| *Burkholderia fungorum*[8] | (73%) | N E P P **F** T A | **E Y** G A |
| *Mesorhizobium loti*[9] | (83%) | N E P P Y T A | **E Y** G A |
| *Thermobifida fusca*[10] | (36%) | N E A P **F** G Y | A W D G |
| *Rhodobacter sphaeroides*[11] | (83%) | N E P P **F** T A | **E Y** G A |
| *Desulfovibrio vulgaris*[12] | (28%) | I E P P Y A F | D F N R |
| *Paracoccus denitrificans*[13] | (76%) | N E P P Y T A | **E Y** G A |
| **Consensus** | | N E x P **F** x / **Y** | |

|  |  | Phe⁸⁰ Arg⁸⁵ | Glu¹³⁴ |
|---|---|---|---|
| *Sinorhizobium meliloti* EhuB | | L **F** M K P E **R** C | T E E |
| *Agrobacterium tumefaciens* | (34%) | Q N I S P E **R** C | N N V |
| *Bordetella parapertussis* | (39%) | M Y V T P E **R** C | I E A |
| *Desulfitobacterium dehalogenans* | (43%) | M **F** I T P E **R** A | I E Y |
| *Pseudomonas putida* | (38%) | M Y I T P A **R** C | V E L |
| *Streptomyces coelicolor* | (40%) | M Y I N P D **R** C | A E I |
| *Arthrobacter sp.* | (45%) | L **F** M K Q S **R** C | F E E |
| *Burkholderia fungorum* | (73%) | L **F** M T P Q **R** C | T E Q |
| *Mesorhizobium loti* | (83%) | L **F** M T P Q **R** C | T E Q |
| *Thermobifida fusca* | (36%) | M Y I T P E **R** C | V E Q |
| *Rhodobacter sphaeroides* | (83%) | L **F** M K P E **R** C | T E E |
| *Desulfovibrio vulgaris* | (28%) | M **F** I T P K **R** I | V E Q |
| *Paracoccus denitrificans* | (76%) | L **F** M K P E **R** C | T E E |
| **Consensus** | | **F** x x P x **R** / **Y** | |

**Fig. 5.** Sequence alignment of EhuB with selected binding proteins identified by a BLAST search. For simplicity, only the sequences of the regions involved in ligand binding are shown. Numbers in parentheses indicate the sequence identity with EhuB. Highly conserved residues are highlighted in bold, and the consensus sequence is given below the individual alignment. Swiss-Prot entry codes are as follows: 1 (NP_436954), 2 (NP_535233), 3 (AAK11558), 4 (NP_746541), 5 (NP_627057), 6 (ZP_00414893), 7 (ZP_00281678), 8 (NP_107510), 9 (YP_288364), 10 (YP_011827), 11 (YP011827) and 12 (ZP_00629206).

hydrogen bonds and likely forces this β-sheet to act as a single unit during domain closure upon ligand binding. The high conservation of the disulfide bond between Cys86 (helix 4; see inset of Fig. 2) and Cys253 (C-terminal helix 12; see inset of Fig. 2) implies that the disulfide bond restricts this region of the protein conformationally and might even have a functional role. Although this region is rather remote from the ligand-binding site at first glance, Arg85, which is located N-terminally to this disulfide bond, is of prime importance for ligand binding. Clearly, this poses an extreme importance on these cysteines, and sequence analysis further supports this fact.

Prolines at positions 23 and 193, which favor the formation of a *cis* peptide bond, are conserved in 12 cases (position 23) and 5 cases (position 253) of the proteins analyzed. The complete conservation of Pro23 suggests an important function of this *cis* peptide bond, which seems to be necessary to position Phe24 in a proper conformation to interact with ectoine or hydroxyectoine via cation–π interactions. Thus, replacement of proline with any other amino acids would disfavor the formation of a *cis* peptide bond and consequently remove one important residue from the cation–π interaction network.

## Structural comparison with other SBPs specific for compatible solutes

Currently, three crystal structures of binding proteins participating in the uptake of compatible solutes are available in the PDB: ProX from the Gram-negative bacterium *E. coli*,[28] ProX from the hyperthermophilic archeon *A. fulgidus*[29] and OpuAC from the Gram-positive soil bacterium *B. subtilis*.[20] Like EhuB, all these proteins employ cation–π interactions for ligand binding via a set of aromatic residues—a Trp box in ProX from *E. coli* composed of three Trp, an aromatic box in ProX from *A. fulgidus* composed of four Tyr and a Trp prism in OpuAC from *B. subtilis* composed of three Trp—but in contrast to EhuB, the binding proteins mentioned above bind glycine betaine with high affinity. Therefore, we analyzed these structures to further extract some of the principles of ligand recognition in binding proteins specific for compatible solutes. Structural superimposition using LSQMAN[43] resulted in an rmsd of 2.13 Å for 134 $C_\alpha$ atoms (EhuB and ProX(1R9L) from *E. coli*), 2.15 Å for 137 $C_\alpha$ atoms (EhuB and ProX(1SW2) from *A. fulgidus*) and 2.10 Å for 128 $C_\alpha$ atoms (EhuB and OpuAC(2B4L) from *B. subtilis*). As expected for a structural comparison of SBPs,[44] the central

structural elements such as the 5-stranded β-sheets agree well among these proteins (data not shown). Differences are mainly observed at the periphery and at the interface of both domains. Despite the existing structural differences among these compatible-solute-specific binding proteins, similar principles of protein–ligand interactions are used. In all proteins, the negatively charged carboxyl moiety of the ligands is bound via salt bridges and/or hydrogen bonds, while the delocalized positive charge is stabilized within the binding pocket via cation–π interactions. The architecture of the aromatic residues participating in cation–π interactions in the ligand-binding sites are summarized in Fig. 6. For simplicity, ProX from *A. fulgidus* has been omitted.

In contrast to the other binding proteins studied, EhuB employs additional salt bridges and hydrogen bonds to further stabilize the delocalized positive charge of the ligand in the substrate-binding site of the protein. Mutational studies on EhuB (this study) and ProX from *E. coli*[28] have demonstrated that the different architecture results in a different hierarchy and promiscuity of the binding site. Moreover, although similar sets of molecular interactions are used, clear differences with respect to the number, the nature and the spatial arrangement of the aromatic amino acids are evident. In summary, this comparison suggests that each of the proteins analyzed structurally follows similar principles for efficient binding of compatible solutes and that the differences in the detailed arrangement of the ligand-binding sites are a consequence of the nature of the ligands and the demands on protein affinity imposed by the environment. Consequently, the ligand-binding site has adopted and presents a perfectly designed pocket to accommodate the individual compatible solutes.

## Conclusions

In this study, we have presented the crystal structures of EhuB, a SBP from the soil bacterium *S. meliloti*, in complex with the compatible solutes ectoine and hydroxyectoine at a resolution of 1.9 Å and 2.3 Å, respectively. This is the first report of an

SBP that is operational in osmoprotection, which is not specific for glycine betaine and proline betaine. Thus, our structures allow a broader insight into the mechanisms these proteins use to bind substrates that are normally excluded from protein surfaces. Our structures—as expected based on the previously determined structures[20,28,29]—revealed a similar ligand-binding site architecture, with the negatively charged carboxyl moiety of the ligands bound via salt bridges and hydrogen bonds and with the delocalized positive charge of the ligands stabilized within the substrate-binding site via cation–π interactions. Unlike the other proteins studied so far, EhuB showed an additional stabilization of the delocalized positive charge of the ligands via salt bridges and hydrogen bonds. This special feature might account for the higher tolerance of EhuB for the exchange of key residues (as suggested by mutational studies) compared to ProX from *E. coli*. In the latter, substitution of the residues forming the Trp box with nonaromatic residues in 7 out of 12 cases resulted in a mutant protein with no detectable affinity for its ligands. In contrast to that, mutagenesis of the aromatic box of EhuB suggests a much higher tolerance, with only three out of nine nonaromatic mutants unable to bind the ligands. In summary, our structures of the EhuB/ectoine and EhuB/hydroxyectoine complexes have shed additional light on the molecular mechanisms of ligand binding in SBPs specific for compatible solutes. It revealed common principles of substrate recognition, but also important differences that point to a certain degree of flexibility of the substrate-binding site, which is not surprising having the different nature of the ligands and the different demands on protein affinity imposed by the environment in mind.

## Materials and Methods

### Culture conditions for bacterial strains

*E. coli* strains were grown aerobically in Luria–Bertani medium at 37 °C.[45] For the selection of *E. coli* strains carrying derivatives of the expression vector pASK-IBA6 (IBA, Göttingen, Germany), ampicillin (100 μg/ml) was
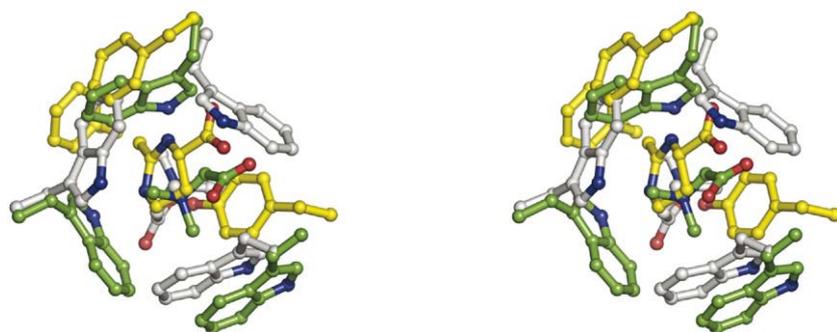


**Fig. 6.** Stereoview of a structural alignment of the EhuB/ectoine complex with SBPs specific for compatible solutes, ProX from *E. coli* (PDB entry 1R9L)[28] and OpuAC from *B. subtilis* (PDB entry 2B4L).[20] EhuB and ectoine are shown in yellow; ProX and its ligand glycine betaine are shown in gray; and OpuAC and its ligand glycine betaine are shown in green. For simplicity, only the compatible solutes and amino acid residues forming the aromatic boxes, which coordinate the ligands via cation–π interactions, have been included in the representation.

added to the cultures. Overproduction of the *S. meliloti* EhuB protein and its mutant derivatives was carried out in a defined minimal medium (MMA).[45]

## Overproduction and purification of the recombinant EhuB protein in *E. coli*

Plasmid pLB22 is an *S. meliloti ehuB*⁺ derivative of the expression vector pASK-IBA6. In this recombinant plasmid, the *ehuB* coding region (without its own signal sequence) is positioned under the control of the anhydrotetracycline-inducible *tet* promoter present in the vector pASK-IBA6, thereby allowing induction of the transcription of the *ehuB* gene to high levels in the heterologous host *E. coli*.[19] The *ehuB* coding region has been inserted in-frame with an upstream *ompA* signal sequence and a Strep-TagII affinity peptide. This allowed the secretion of the Strep-TagII–EhuB fusion protein into the periplasm of *E. coli*, where it could be released from by cold osmotic shock and recovered by affinity chromatography on Strep-Tactin Sepharose (IBA).[19] EhuB was overproduced in the *E. coli* strain BL21 (F⁻ gal met r⁻ m⁻ hsdS (λDE3); Stratagene, La Jolla, CA, USA) carrying plasmid pLB22. Cells were grown in defined minimal medium (MMA) supplemented with 100 μg/ml ampicillin, 0.5% (wt/vol) glucose and 0.2% (wt/vol) casamino acids, as detailed by Jebbar *et al.*[19] Purification of the recombinant EhuB protein followed the procedure published by Jebbar *et al.*[19] In general, approximately 3 mg of pure EhuB protein was obtained per liter of culture. The purified protein was concentrated to approximately 10 mg/ml using VIVASPIN 4 (VIVASCIENCE, Hanover, Germany) concentrator columns (exclusion size, 10 kDa) and dialyzed extensively at 4 °C against 5 l of 10 mM Tris (tris(hydroxymethyl)aminomethane)–HCl (pH 7) to remove desthiobiotin and salt from the protein. The concentrated EhuB was analyzed by SDS-PAGE for purity and stored at 4 °C until further use. The functionality of the purified protein was assessed by fluorescence spectroscopy using the intrinsic tyrosine fluorescence of EhuB.

## Site-directed mutagenesis of the *ehuB* gene

Site-directed mutagenesis of the *ehuB* gene was carried out with the QuikChange site-directed mutagenesis kit (Stratagene) using plasmid pLB22 as the DNA template and custom-synthesized mutagenic DNA primers. The crystal structure of the EhuB protein liganded with either ectoine or hydroxyectoine showed that the residues Phe24, Tyr60, Phe80, Glu21, Arg85, Thr133 and Glu134 were involved in ligand binding. To assess the importance of individual residues for the binding of either ectoine or hydroxyectoine, we replaced each of these amino acids with codons encoding the amino acid Ala, Phe/Tyr, Trp, Glu or Asp. This generated the following mutant *ehuB* plasmids: pMH1 (Phe24→Ala [TTT→GCG]); pMH2 (Phe24→Tyr [TTT→TAT]); pMH3 (Phe24→Trp [TTT→TGG]); pMH4 (Phe24→Asp [TTT→GAT]); pMH5 (Phe24→Glu [TTT→GAA]); pMH6 (Tyr60→Ala [TAC→GCG]); pMH7 (Tyr60→Phe [TAC→TTT]); pMH8 (Tyr60→Trp [TAC→TGG]); pMH9 (Tyr60→Asp [TAC→GAT]); pMH10 (Tyr60→Glu [TAC→GAA]); pMH11 (Phe80→Ala [TTT→GCG]); pMH12 (Phe80→Tyr [TTT→TAT]); pMH13 (Phe80→Trp [TTT→TGG]); pMH14 (Phe80→Asp [TTT→GAT]); pMH15 (Phe80→Glu [TTT→GAA]); pMH16 (Glu21→Ala [GAG→GCG]); pMH17 (Arg85→Ala [CGA→GCG]); pMH18 (Thr133→ Ala [ACC→GCG]); pMH19 (Glu134→Ala [GAG→GCG]).

A double mutant with an exchange of the amino acids Phe24 and Tyr60 with Ala was prepared by remutating the plasmid pMH1 (Phe24 into Ala) to yield plasmid pMH21 (Phe24 into Ala; Tyr60 into Ala). The entire coding region of each mutant *ehuB* gene was sequenced to ensure the presence of the desired mutation and the absence of unwanted additional alterations in the *ehuB* coding region.

## Determination of the binding constants of EhuB proteins for ectoine and hydroxyectoine by fluorescence spectroscopy

The native EhuB protein contains six tyrosine residues but no tryptophans. We therefore employed the intrinsic tyrosine fluorescence of EhuB to monitor the binding of the ligands ectoine and hydroxyectoine. Ectoine [(*S*)-2-methyl-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid] and hydroxyectoine [(*S,S*)-2-methyl-5-hydroxy-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid] were purchased from bitop (Witten, Germany). Fluorescence spectroscopy measurements were performed using a Cary Eclipse Fluorescence Photometer (VARIAN, Palo Alto, CA, USA). The excitation wavelength was set to 280 nm (with a slid width of 5 nm), and temperature was maintained at 25 °C using a circulating water bath. EhuB protein solution (900 μl of 200 nM; in 10 mM Tris–HCl, pH 7) was mixed with 100 μl of a substrate solution (in 10 mM Tris–HCl, pH 7) with the appropriate concentration of the ligand ectoine or hydroxyectoine. The mixture was then incubated for 1.5 min to allow equilibration before the actual measurement. The EhuB concentration employed in the assay was 18 nM for the *F24Y, F24W, Y60W, F80Y* and *F80W* mutants, and 180 nM for all other mutants. Emission spectra were monitored from 280 nm to 410 nm using the program Cary Eclipse Scan (VARIAN). Upon substrate binding, a decrease in the fluorescence intensity of EhuB was detected, and this change in emission spectrum was used to determine the $K_d$ values for the EhuB proteins for either ectoine or hydroxyectoine by plotting the peak area from 295 nm to 400 nm against the substrate concentration. For convenience, data were converted into fluorescence increase and analyzed by assuming a standard one-site-binding model according to:

$$F_{cor} = F_0 + (\Delta F[S_0]/([S_0] + K_d))$$

where $F_{cor}$ = fluorescence intensity for a given substrate concentration; $F_0$ = fluorescence intensity without substrate; $\Delta F$ = maximal change in fluorescence intensity; $S_0$ = substrate concentration; and $K_d$ = binding constant.

All $K_d$ measurements of EhuB and its mutant derivatives represent the average of at least three independent measurements, with the standard deviation given as errors.

## Crystallization

Crystals of wild-type and SeMet-substituted EhuB, in complex with one of its substrates ectoine or hydroxyectoine, were obtained using a vapor diffusion technique at 4 °C. Prior to crystallization, the protein solution in 10 mM Tris, pH 7.0 (10–40 mg/ml for the wild-type protein and 7 mg/ml for the SeMet-substituted protein), was incubated for 30 min on ice with a final concentration of 1 mM ectoine or hydroxyectoine, respectively. Following incubation, equal amounts of protein and reservoir solution were mixed and incubated in a hanging drop setup at 4 °C

to allow crystallization. The reservoir solution contained 100 mM sodium acetate (pH 4.0), 100 mM cadmium chloride and 14% (wt/vol) polyethylene glycol (PEG) 5000 MME for the wild-type EhuB and 22% (wt/vol) PEG 5000 MME for the SeMet-substituted EhuB. Crystal needles typically grew in 4–6 weeks and reached a final size of 500 μm × 80 μm × 40 μm for the wild-type protein. For data collection, crystal needles were separated using micro-tools, and single crystals were transferred into a suitable cryobuffer and flash frozen in liquid nitrogen. In case of wild-type EhuB with bound ectoine, the optimized cryobuffer contained 100 mM sodium acetate (pH 4.0), 100 mM cadmium chloride, 21% (wt/vol) PEG 5000 MME and 20% (vol/vol) ethylene glycol. For wild-type EhuB complexed with hydroxyectoine and for SeMet-substituted protein complexed with ectoine, the best results were obtained with the above cryobuffer containing 30% (wt/vol) PEG 5000 MME.

### Data collection, structure determination and refinement

A three-wavelength MAD data set from a single crystal of SeMet-substituted EhuB complexed with ectoine, and native data sets from single crystals of EhuB in complex with ectoine and hydroxyectoine, respectively, were collected under cryogenic conditions at the beamline BW-6 (DESY, Hamburg, Germany) using a MarCCD detector. All data sets were processed using DENZO and SCALEPACK,[46] as summarized in Table 1. Initial phases were determined by three-wavelength MAD. In brief, four out of five expected selenium positions were located and used to calculate the initial phases employing SOLVE,[47] followed by statistical density modification and auto-mated model building using RESOLVE.[47] The resulting model for the structure of SeMet-substituted EhuB complexed with ectoine covered approximately 80% of the protein and was completed by manual building into $1F_{obs}-F_{cal}$ and $2F_{obs}-F_{cal}$ electron density maps using O.[48] After several rounds of manual rebuilding and refinement of the model using O and REFMAC5,[49] respectively, the substrate ectoine, as well as four heavy atoms that were modeled as cadmium ions, were included. Further rounds of refinement with REFMAC5, including TLS refinement,[50] resulted in a model for the structure of SeMet-substituted EhuB in complex with ectoine at a resolution of 2.1 Å. To extend the structural model of EhuB complexed with ectoine to a resolution of 1.9 Å, the model for the SeMet-substituted protein was refined against the native data set. In short, rigid body refinement using the previously determined model of SeMet-substituted EhuB without the substrate and cad-mium ions, followed by several rounds of manual rebuilding and refinement, was carried out using O and REFMAC5. Subsequently, the substrate and cadmium ions were included. Due to imperfections in the native data set, which might be due to beginning radiation damage, the occupancy of cadmium ions was refined manually by decreasing the occupancy in steps of 10%, followed by refinement with REFMAC5 until a minimum in the $1F_{obs}-F_{cal}$ difference density had been reached. After further rounds of rebuilding and refinement, automated water picking using ARP/wARP[51] was performed with a conservative cutoff of 3.2σ, followed by visual inspection of all water molecules, which were only kept in case of clear density. The remaining positive-difference density was not modeled although it was indicative of, for example, PEG molecules, since these areas were all located at the surface of the protein and not in functionally important regions of the protein. Additional difference density observed at the disulfide bridges and a few exposed acidic residues are most likely the first signs of beginning radiation damage.[52] The final model of the structure of EhuB, in complex with ectoine, was obtained after further rounds of rebuilding and refinement, which included TLS refinement in the last round of refinement.

The structure of EhuB complexed with hydroxyectoine was solved by molecular replacement with AMoRe using the previously determined model of EhuB with bound ectoine as template. After several rounds of manual rebuilding and refinement, the hydroxyectoine and the four cadmium ions were included in the model. Following further rounds of rebuilding and refinement, automated water picking and final refinement of the structure were performed as described for EhuB complexed with ectoine.

For cross-validation, 5% of all data sets were excluded from the refinement to calculate the $R_{free}$ value.[53] The quality of the obtained models was validated with the program PROCHECK.[54]

### Data bank searches, sequence alignments and structure alignments

Sequence homologues of EhuB were searched via the National Center for Biotechnology Information†. Struc-tural homologues of EhuB in the PDB were identified via the DALI server at the European Bioinformatics Institute‡. Sequence alignments were performed using CLUSTALW§. Structure alignments were performed using LSQMAN,[43] employing standard settings.

### Figure preparation

Structure figures were prepared using PYMOL‖.

### PDB accession code

Coordinates have been deposited in the PDB under accession codes 2Q88 (EhuB/ectoine) and 2Q89 (EhuB/hydroxyectoine).

† www.ncbi.nlm.nih.gov/BLAST/
‡ www.ebi.ac.uk/dali/
§ www.ebi.ac.uk/clustalw/
‖ http://pymol.sourceforge.net/

## References

1. Kempf, B. & Bremer, E. (1998). Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Arch. Microbiol.* **170**, 319–330.

2. Horn, C., Jenewein, S., Sohn-Bosser, L., Bremer, E. & Schmitt, L. (2005). Biochemical and structural analysis of the *Bacillus subtilis* ABC transporter OpuA and its isolated subunits. *J. Mol. Microbiol. Biotechnol.* **10**, 76–91.

3. Csonka, L. N. & Epstein, W. (1996). Osmoregulation. In *Escherichia coli and Salmonella. Cellular and Molecular Biology* (Neidhard, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M. & Umbarger, H. E., eds), pp. 1210–1223, American Society for Microbiology Press, Washington, DC.

4. da Costa, M. S., Santos, H. & Galinski, E. A. (1998). An overview of the role and diversity of compatible solutes in Bacteria and Archaea. *Adv. Biochem. Eng. Biotechnol.* **61**, 117–153.

5. Record, M. T., Jr., Courtenay, E. S., Cayley, D. S. & Guttman, H. J. (1998). Responses of *E. coli* to osmotic stress: large changes in amounts of cytoplasmic solutes and water. *Trends Biochem. Sci.* **23**, 143–148.

6. Record, M. T., Jr., Courtenay, E. S., Cayley, S. & Guttman, H. J. (1998). Biophysical compensation mechanisms buffering *E. coli* protein–nucleic acid interactions against changing environments. *Trends Biochem. Sci.* **23**, 190–194.

7. Le Rudulier, D., Strom, A. R., Dandekar, A. M., Smith, L. T. & Valentine, R. C. (1984). Molecular biology of osmoregulation. *Science*, **224**, 1064–1068.

8. Canovas, D., Borges, N., Vargas, C., Ventosa, A., Nieto, J. J. & Santos, H. (1999). Role of Ngamma-acetyldiaminobutyrate as an enzyme stabilizer and an intermediate in the biosynthesis of hydroxyectoine. *Appl. Environ. Microbiol.* **65**, 3774–3779.

9. Lippert, K. & Galinski, E. A. (1992). Enzyme stabilization by ectoine-type compatible solutes: protection against heating, freezing and drying. *Appl. Microbiol. Biotechnol.* **37**, 61–65.

10. Bourot, S., Sire, O., Trautwetter, A., Touze, T., Wu, L. F., Blanco, C. & Bernard, T. (2000). Glycine betaine-assisted protein folding in a *lysA* mutant of *Escherichia coli. J. Biol. Chem.* **275**, 1050–1056.

11. Arakawa, T. & Timasheff, S. N. (1985). The stabilization of proteins by osmolytes. *Biophys. J.* **47**, 411–414.

12. Bolen, D. W. & Baskakov, I. V. (2001). The osmophobic effect: natural selection of a thermodynamic force in protein folding. *J. Mol. Biol.* **310**, 955–963.

13. Qu, Y., Bolen, C. L. & Bolen, D. W. (1998). Osmolyte-driven contraction of a random coil protein. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 9268–9273.

14. Sleator, R. D. & Hill, C. (2002). Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence. *FEMS Microbiol. Rev.* **26**, 49–71.

15. Davidson, A. L. & Chen, J. (2004). ATP-binding cassette transporters in bacteria. *Annu. Rev. Biochem.* **73**, 241–268.

16. Calderon, M. I., Vargas, C., Rojo, F., Iglesias-Guerra, F., Csonka, L. N., Ventosa, A. & Nieto, J. J. (2004). Complex regulation of the synthesis of the compatible solute ectoine in the halophilic bacterium *Chromohalobacter salexigens* DSM 3043T. *Microbiology*, **150**, 3051–3063.

17. Kuhlmann, A. U. & Bremer, E. (2002). Osmotically regulated synthesis of the compatible solute ectoine in *Bacillus pasteurii* and related *Bacillus* spp. *Appl. Environ. Microbiol.* **68**, 772–783.

18. Reshetnikov, A. S., Khmelenina, V. N. & Trotsenko, Y. A. (2006). Characterization of the ectoine biosynthesis genes of haloalkalotolerant obligate methanotroph "*Methylomicrobium alcaliphilum 20Z*". *Arch. Microbiol.* **184**, 286–297.

19. Jebbar, M., Sohn-Bosser, L., Bremer, E., Bernard, T. & Blanco, C. (2005). Ectoine-induced proteins in *Sinorhizobium meliloti* include an Ectoine ABC-type transporter involved in osmoprotection and ectoine catabolism. *J. Bacteriol.* **187**, 1293–1304.

20. Horn, C., Sohn-Bosser, L., Breed, J., Welte, W., Schmitt, L. & Bremer, E. (2006). Molecular determinants for substrate specificity of the ligand-binding protein OpuAC from *Bacillus subtilis* for the compatible solutes glycine betaine and proline betaine. *J. Mol. Biol.* **357**, 592–606.

21. Boos, W. & Lucht, J. M. (1996). Periplasmic binding protein-dependent ABC transporters. In *Escherichia coli and Salmonella. Cellular and Molecular Biology* (Neidhard, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M. & Umbarger, H. E., eds), pp. 1175–1209, American Society for Microbiology Press, Washington, DC.

22. Gilson, E., Alloing, G., Schmidt, T., Claverys, J. P., Dudler, R. & Hofnung, M. (1988). Evidence for high affinity binding-protein dependent transport systems in Gram-positive bacteria and in *Mycoplasma*. *EMBO J.* **7**, 3971–3974.

23. Kempf, B., Gade, J. & Bremer, E. (1997). Lipoprotein from the osmoregulated ABC transport system OpuA of *Bacillus subtilis*: purification of the glycine betaine binding protein and characterization of a functional lipidless mutant. *J. Bacteriol.* **179**, 6213–6220.

24. van der Heide, T. & Poolman, B. (2000). Osmoregulated ABC-transport system of *Lactococcus lactis* senses water stress via changes in the physical state of the membrane. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 7102–7106.

25. Quiocho, F. A. & Ledvina, P. S. (1996). Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. *Mol. Microbiol.* **20**, 17–25.

26. Bjorkman, A. J. & Mowbray, S. L. (1998). Multiple open forms of ribose-binding protein trace the path of its conformational change. *J. Mol. Biol.* **279**, 651–664.

27. Magnusson, U., Chaudhuri, B. N., Ko, J., Park, C., Jones, T. A. & Mowbray, S. L. (2002). Hinge-bending motion of D-allose-binding protein from *Escherichia coli*: three open conformations. *J. Biol. Chem.* **277**, 14077–14084.

28. Schiefner, A., Breed, J., Bosser, L., Kneip, S., Gade, J., Holtmann, G. *et al.* (2004). Cation–pi interactions as determinants for binding of the compatible solutes glycine betaine and proline betaine by the periplasmic ligand-binding protein ProX from *Escherichia coli*. *J. Biol. Chem.* **279**, 5588–5596.

29. Schiefner, A., Holtmann, G., Diederichs, K., Welte, W. & Bremer, E. (2004). Structural basis for the binding of compatible solutes by ProX from the hyperthermophilic archaeon *Archaeoglobus fulgidus*. *J. Biol. Chem.* **275**, 48270–48281.

30. Lakowicz, J. R. (1999). *Principles of Fluorescence Spec-*

*troscopy*, 2nd edit. Kluwer Academic Publishers, New York.

31. Fukami-Kobayashi, K., Tateno, Y. & Nishikawa, K. (1999). Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history. *J. Mol. Biol.* **286**, 279–290.

32. Roderick, S. L., Chan, W. W., Agate, D. S., Olsen, L. R., Vetting, M. W., Rajashankar, K. R. & Cohen, D. E. (2002). Structure of human phosphatidylcholine transfer protein in complex with its ligand. *Nat. Struct. Biol.* **9**, 507–511.

33. Brejc, K., van Dijk, W. J., Klaassen, R. V., Schuurmans, M., van Der Oost, J., Smit, A. B. & Sixma, T. K. (2001). Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature*, **411**, 269–276.

34. Oh, B. H., Pandit, J., Kang, C. H., Nikaido, K., Gokcen, S., Ames, G. F. & Kim, S. H. (1993). Three-dimensional structures of the periplasmic lysine/arginine/ornithine-binding protein with and without a ligand. *J. Biol. Chem.* **268**, 11348–11355.

35. Armstrong, N., Sun, Y., Chen, G. Q. & Gouaux, E. (1998). Structure of a glutamate-receptor ligand-binding core in complex with kainate. *Nature*, **395**, 913–917.

36. Trakhanov, S., Kreimer, D. I., Parkin, S., Ames, G. F. & Rupp, B. (1998). Cadmium-induced crystallization of proteins: II. Crystallization of the *Salmonella typhimurium* histidine-binding protein in complex with L-histidine, L-arginine, or L-lysine. *Protein Sci.* **7**, 600–604.

37. Inbar, L., Frolow, F. & Lapidot, A. (1993). The conformation of new tetrahydropyrimidine derivatives in solution and in the crystal. *Eur. J. Biochem.* **214**, 897–906.

38. Mao, B., Pear, M. R., McCammon, J. A. & Quiocho, F. A. (1982). Hinge-bending in L-arabinose-binding protein. The "Venus's-flytrap" model. *J. Biol. Chem.* **257**, 1131–1133.

39. Yao, N., Trakhanov, S. & Quiocho, F. A. (1994). Refined 1.89-A structure of the histidine-binding protein complexed with histidine and its relationship with many other active transport/chemosensory proteins. *Biochemistry*, **33**, 4769–4779.

40. Sun, Y. J., Rose, J., Wang, B. C. & Hsiao, C. D. (1998). The structure of glutamine-binding protein complexed with glutamine at 1.94 A resolution: comparisons with other amino acid binding proteins. *J. Mol. Biol.* **278**, 219–229.

41. Ma, J. C. & Dougherty, D. A. (1997). The cation–pi interaction. *Chem. Rev.* **97**, 1303–1324.

42. Mecozzi, S., West, A. P., Jr & Dougherty, D. A. (1996). Cation–pi interactions in aromatics of biological and medicinal interest: electrostatic potential surfaces as a useful qualitative guide. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 10566–10571.

43. Kleywegt, G. J. (1999). Experimental assessment of differences between related protein crystal structures. *Acta Crystallogr. Sect. D*, **55**, 1878–1884.

44. Wilkinson, A. J. & Verschueren, K. H. G. (2003). Crystal structures of periplasmic solute binding proteins in ABC transport complexes illuminate their function. In *ABC Proteins from Bacteria to Man* (Holland, I. B., Cole, S. P. & Kuchler, K., eds), Academic Press, Amsterdam.

45. Miller, J. H. (1992). *A Short Course in Bacterial Genetics. A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

46. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology* (Carter, C. W. & Sweet, R. M., eds), vol. 276, Academic Press, London.

47. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallogr. Sect. D*, **55**, 849–861.

48. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. Sect. A*, **47**, 110–119.

49. Murshudov, G., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D*, **53**, 240–255.

50. Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. Sect. D*, **57**, 122–133.

51. Lamzin, V. S. & Wilson, K. S. (1993). Automated refinement of protein models. *Acta Crystallogr. Sect. D*, **49**, 129–147.

52. Burmeister, W. P. (2000). Structural changes in a cryo-cooled protein crystal owing to radiation damage. *Acta Crystallogr. Sect. D*, **56**, 328–341.

53. Brünger, A. T. (1992). Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–475.

54. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.