doi:10.1016/j.jmb.2005.12.085 J. Mol. Biol. (2006) 357, 592–606

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Molecular Determinants for Substrate Specificity of the Ligand-binding Protein OpuAC from *Bacillus subtilis* for the Compatible Solutes Glycine Betaine and Proline Betaine

Carsten Horn[1], Linda Sohn-Bösser[2], Jason Breed[3], Wolfram Welte[3] Lutz Schmitt[1]* and Erhard Bremer[2]

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

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

[3]*Department of Biology University of Konstanz Universitätsstr. 10, 78457 Konstanz, Germany*

**Corresponding author*

Compatible solutes play a decisive role in the defense of microorganisms against changes in temperature and increases in osmolarity in their natural habitats. In *Bacillus subtilis*, the substrate-binding protein (SBP)-dependent ABC-transporter OpuA serves for the uptake of the compatible solutes glycine betaine (GB) and proline betaine (PB). Here, we report the determinants of compatible solute binding by OpuAC, the SBP of the OpuA transporter, by equilibrium binding studies and X-ray crystallography. The affinity of OpuAC/GB and OpuAC/PB complexes were analyzed by intrinsic tryptophan fluorescence and the K_D values were determined to be 17($\pm$1) μM for GB and 295($\pm$27) μM for PB, respectively. The structures of OpuAC in complex with GB or PB were solved at 2.0 Å and 2.8 Å, respectively, and show an SBP-typical class II fold. The ligand-binding pocket is formed by three tryptophan residues arranged in a prism-like geometry suitable to coordinate the positive charge of the trimethyl ammonium group of GB and the dimethyl ammonium group of PB by cation-π interactions and by hydrogen bonds with the carboxylate moiety of the ligand. Structural differences between the OpuAC/GB and OpuAC/PB complexes occur within the ligand-binding pocket as well as across the domain–domain interface. These differences provide a structural framework to explain the drastic differences in affinity of the OpuAC/GB and OpuAC/PB complexes. A sequence comparison with putative SBP specific for compatible solutes reveals the presence of three distinct families for which the crystal structure of OpuAC might serve as a suitable template to predict the structures of these putative compatible solute-binding proteins.

Keywords: substrate-binding protein; osmoprotection; cation-π; Trp-prism; ligand specificity

Introduction

The intracellular water content of microorganisms exceeds that of the environment resulting in the build-up of a hydrostatic pressure, the turgor.[1] An increase in the salinity of the environment will trigger osmotically instigated water fluxes across the permeability barrier of the cell, the cytoplasmic membrane. To avoid excessive water efflux, plasmolysis, molecular crowding of the cytoplasm and cessation of growth in high-salinity environments,[2] many Bacteria and Archaea amass large amounts of

Present addresses: C. Horn, CellGenix Technology Transfer GmbH, Am Flughafen 16, 79108 Freiburg, Germany; J. Breed, Astra Zeneca, Mereside, Macclesfield SK10 4TG, UK.

Abbreviations used: ABC, ATP-binding cassette; GB, glycine betaine; MAD, multiple anomalous dispersion; Opu, osmoprotectant uptake; PB, proline betaine; SEC, size-exclusion chromatography; SBP, substrate-binding protein; TMD, transmembrane domain.

E-mail address of the corresponding author: lutz.schmitt@uni-duesseldorf.de

a particular class of low molecular mass organic osmolytes, the so-called compatible solutes.[3–6] This can be accomplished either through biosynthesis or uptake from the environment through dedicated transport systems.[4,5,7] Compatible solutes do not interfere with cellular functions and they can be amassed up to molar concentrations in the cytoplasm. Accumulation of compatible solutes triggers water entry into the cell, thereby allowing a re-hydration of the cytoplasm, a restoration of turgor and subsequently the resumption of cell growth under unfavorable osmotic conditions.

In addition to their well-established role as osmoprotectants, compatible solutes function as protein stabilizers both *in vitro*[8] and *in vivo*.[9] This property of compatible solutes is generally explained in terms of the preferential exclusion model.[10,11]

One of the most widely used compatible solutes in nature is the trimethyl-ammonium compound glycine betaine (GB).[12] Microorganisms that can acquire preformed compatible solutes from the environment have a growth advantage in high-osmolality and high-salinity habitats.[3–5,7] In the Gram-positive soil bacterium *Bacillus subtilis*, five compatible solute transport systems are operational that, together, allow the acquisition of 12 preformed compatible solutes and of choline, the precursor for the synthesis of GB.[13] One of these transport systems is OpuA (Opu, osmoprotectant uptake),[14] a member of the binding protein-dependent ATP-binding cassette (ABC) superfamily of transporters.[15–17]

The OpuA system[14] consists of a cytoplasmic membrane-associated ATPase (OpuAA),[18] the membrane-spanning substrate translocator protein OpuAB,[19] and the extracellular ligand-binding protein OpuAC.[20] This latter protein is tethered to the outside of the cytoplasmic membrane *via* a lipid modification at a Cys-residue at the mature N terminus of OpuAC.[14,20] OpuAC binds GB with high affinity,[20] and delivers it to the OpuAB/OpuAA complex for the release of substrate from the ligand-binding protein and further ATP-dependent substrate translocation into the cytoplasm.[18,19] The OpuA system contributes significantly to GB import when this compatible solute is used by *B. subtilis* for osmoprotective purposes,[13,21] and serves for its uptake when it is used either as an effective heat[22] or cold stress protectant.[23] In addition to GB, the compatible solute proline betaine (PB) functions as an osmoprotectant for *B. subtilis* and is acquired by the cell *via* the OpuA transporter.[13] Preliminary binding assays have revealed that PB is a substrate for the OpuAC ligand-binding protein (B. Kempf and E.B., unpublished results).

The ligand-binding proteins of binding protein-dependent ABC-transport systems either diffuse freely in the periplasm of Gram-negative bacteria or are tethered with a lipid tail to the outside of the cytoplasmic membrane of Gram-positive bacteria.[15,20,24]

Crystallographic analysis of a considerable number of substrate-binding proteins has shown that they are generally composed of two rigid globular lobes that are connected by a variable number of hinge regions.[25–27] In the ligand-free conformation, the so-called open form of the binding protein, the two globular domains are linked flexibly by the hinge regions.[28,29] Upon substrate binding, a conformational change is induced in the hinge that generally causes a large rigid body motion that moves both globular domains towards each other.[26,30] This hinge bending motion engulfs the ligand in a predefined cleft between the two globular domains of the substrate-binding protein and largely occludes it from the bulk solvent. The ligand-bound form of the substrate-binding protein is referred to as the closed conformation.[26]

The substrate-binding protein confers substrate specificity to a given bacterial binding protein-dependent ABC-transporter.[15] These types of binding proteins generally recognize and bind their ligands with high affinity, usually with K_D values in the low micromolar or even nanomolar range.[15,31] This is true also for the purified OpuAC ligand-binding protein of the *B. subtilis* OpuA transporter, which binds radiolabelled GB *in vitro* within an apparent K_D of 6 μM, as assessed by an ammonium sulfate-precipitation assay.[20]

Ligand-binding proteins from bacterial ABC transport systems that are involved in compatible solute acquisition have to bind compatible solutes with specificity and high affinity.[32,33] Therefore, these types of binding proteins have to avoid the exclusion of their ligands within the binding pocket that otherwise is generally observed for proteins.[10,11] The recent crystallographic analysis of the GB/PB ligand-binding protein ProX from *Escherichia coli*[32] and the Archaeon *Archaeoglobus fulgidus*[33] have provided an initial insight into how such a task can be accomplished. In the *E. coli* ProX protein,[32] a rectangular aromatic box formed by three Trp-residues provides an evenly structured negative surface to accommodate the bulky and positively charged trimethylammonium head-group of GB and the dimethylammonium group of PB *via* cation-π interactions.[34] The carboxylic groups of the GB and PB ligands are stabilized within the binding pocket through hydrogen bond interactions.[32] Although the amino acid sequences of the ProX proteins from *E. coli* and *A. fulgidus* are only distantly related, a similar solution for GB/PB binding has been realized in the *A. fulgidus* ProX protein.[33] Here, a girdle of four Tyr residues is key to binding of the quaternary amines of the ligands *via* cation-π interactions together with a contact to a main-chain carbonyl oxygen atom. Stabilization of the carboxylic groups of GB and PB is accomplished *via* either hydrogen bonds or salt-bridges.[33]

The amino acid sequence of the OpuAC protein from *B. subtilis* is only very distantly related to those of the ligand-binding proteins ProX from *E. coli* or *A. fulgidus*. To further extend our understanding of

the specific binding of compatible solutes by ligand-binding proteins of ABC transport systems, we have determined the crystal structure of the OpuAC protein with its natural ligands GB or PB at a resolution of 2.0 Å and 2.8 Å, respectively. This allowed us to deduce the molecular determinants that govern ligand binding by the OpuAC protein. In database searches, we found many bacterial and archaeal amino acid sequences related to the *B. subtilis* OpuAC protein. However, in most cases an efficient alignment of OpuAC with these protein sequences could be accomplished only when the N and C-terminal segments of the OpuAC protein were inverted. Remarkably, the three Trp residues that are key for substrate binding *via* cation-π interactions are conserved completely within the many proteins aligned with the *B. subtilis* OpuAC amino acid sequence. The OpuAC crystal structure represented here can thus be considered as a lead structure for many putative GB/PB ligand-binding proteins.

Results and Discussion

The OpuAC protein is tethered to the outside of the *B. subtilis* cytoplasmic membrane *via* a lipid modification at its N terminus.[20] Such a lipid modification would obviously interfere with crystallization attempts of OpuAC. We therefore made use of a previously constructed MalE-OpuAC hybrid protein from which the OpuAC part (without lipid modification) can be released by proteolytic cleavage.[20] OpuAC was then purified to apparent homogeneity as described in Materials and Methods. The oligomeric state of the purified OpuAC protein was verified by size-exclusion chromatography (SEC) and showed a symmetric peak with an apparent mass of 28 kDa corresponding to monomeric OpuAC (calculated mass 30.3 kDa; data not shown).

Equilibrium binding experiments of OpuAC with the natural substrates glycine betaine and proline betaine

The binding properties of OpuAC towards its physiological ligands GB and PB were analyzed by equilibrium binding experiments following intrinsic tryptophan fluorescence[35] in 10 mM Tris–HCl (pH 7.3). OpuAC contains eight tryptophan residues spread throughout the sequence. In the absence of a ligand, an emission maximum at 346 nm was detected (Figure 1(a), red curve) which is characteristic for solvent-exposed Trp.[36] The addition of a ligand, either GB or PB, resulted in an increase of fluorescence that was accompanied by a blue shift of the emission maximum. Under ligand saturation conditions, this shift amounted to 10 nm (Figure 1(a), black curve). Such an observation is indicative of a decrease in local polarity of at least part of the Trp residues, and can be rationalized by a ligand-induced shielding.

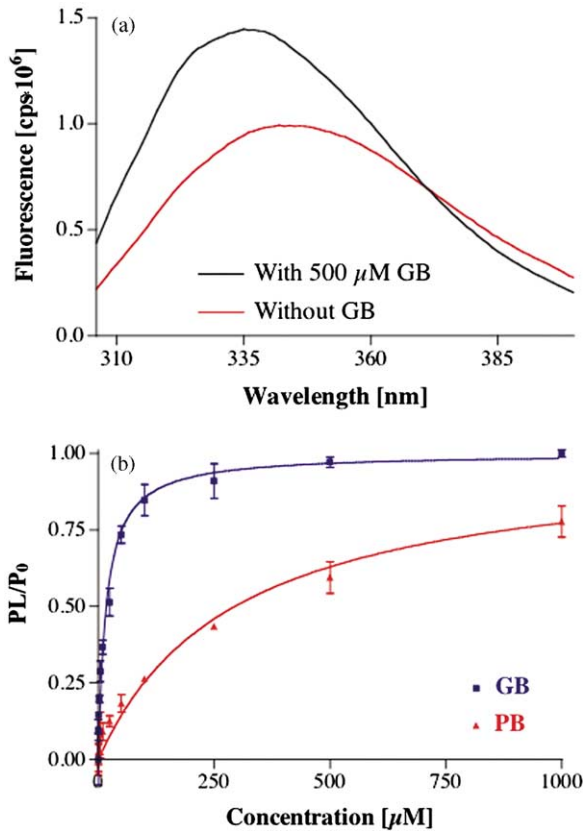

Figure 1. Binding properties of OpuAC/ligand complexes. (a) Emission spectra of OpuAC in the absence (red line) or in the presence of 0.5 mM GB (black line). (b) Equilibrium binding titration experiments with GB (■) or PB (▲). The shift of the emission maximum in the presence of different amounts of ligand was analyzed according to equation (1).

The shift of the emission maximum was normalized and used to determine the affinity of the OpuAC/ligand complexes according to equation (1) (Figure 1(b)). The dissociation constants of OpuAC/GB and OpuAC/PB complexes were calculated to be 17(±1) μM and 295(±27) μM, respectively. The determined value of the OpuAC/GB complex is in agreement with the value (6 μM) determined by Kempf *et al.* using a radioactive filter assay,[20] while no reference value for the dissociation constant of the OpuAC/PB complex exists. The results of the equilibrium binding experiments demonstrate that the dissociation constants of the OpuAC/GB and OpuAC/PB complexes differ by more than one order of magnitude, despite the fact that both ligands are natural substrates of OpuAC and share similar molecular characteristics, i.e. a carboxylate group and a fully methylated nitrogen atom. The strong difference in the affinity of the *B. subtilis* OpuAC protein for its natural ligands GB and PB is different from the situation observed for the binding of these compatible solutes by the *E. coli* and *A. fulgidus* ProX proteins, which both bind GB and PB with similar low affinities.[32,33,37–39]

Crystal structure determination of the OpuAC/GB and OpuAC/PB complexes

The structures of OpuAC in complex with the compatible solutes GB and PB were determined in the closed liganded conformation. Initial phases of a SeMet-substituted OpuAC/GB crystal were obtained from a three-wavelength multiple anomalous dispersion (MAD) experiment (see Materials and Methods and Table 1). Since SeMet-substituted and native crystals of OpuAC/GB grew in different space groups, the SeMet crystal structure was refined to an R_{free} value below 30% and subsequently used as molecular replacement template for the native data set. After refinement, the native OpuAC/GB structure was used to determine and refine the crystal structure of the OpuAC/PB complex (Table 1). A summary of the data statistics and refinement details as well as the model content are given in Table 1.

Overall structure of the OpuAC/GB and OpuAC/PB complexes

The overall structure of the OpuAC/GB complex is shown in Figure 2. The first four N-terminal residues were disordered in the crystal structure and have not been included in the model. Since OpuAC is tethered to the plasma membrane *via* a lipid modification at the N-terminal Cys,[20] one would expect a high intrinsic flexibility in this part of the protein, explaining the observed disorder.

The OpuAC protein can be divided into two globular domains (red and light magenta in Figure 2), which are formed by two central, five-stranded β-sheets flanked by several α-helices.[26] This is in contrast to the closely related ProX protein from *E. coli*[32] and *A. fulgidus*,[33] which contained only a four-stranded β-sheet. Domain 1 (residues 4–94 and 253–272) harbors the N terminus and the C terminus, and consists of a β-pleated sheet (strands S_A, S_B, S_C, S_D and S_J) that is packed against five helices (H1–H4 and H13). In domain 2 (residues 100–247) the β-pleated sheet (strands S_E, S_F, S_G, S_H and S_I) is flanked by eight helices (H5–H12). Residues 95–99 and 248–252 form the hinge region connecting both domains. The structure of OpuAC in complex with PB revealed an identical overall conformation (not shown). This fold is generally conserved for substrate-binding proteins (SBPs). Based on the topology of the five β-strands (order BACJD, see Figure 2, inset), OpuAC belongs to class

Table 1. Crystal parameters and data collection statistics are derived from SCALEPACK

	GB native	PB native	GB SeMet		
A. *Crystal parameters at 100 K*					
Space group	$P2_12_12_1$	$P2_1$		$P2_1$	
Unit cell parameters					
a, *b*, *c* (Å)	29.82, 88.49, 95.71	88.56, 28.31, 102.86			
α,β,γ (deg.)	90, 90, 90	90, 93.87, 90		90, 94.94, 90	
B. *Data collection and processing*					
			Infl.	*Max*	*Remote*
Wavelength (Å)	1.05	1.05	0.9787	0.97	0.95
Resolution (Å)	65–2.0 (2.05–2.0)	100–2.8 (2.87–2.8)	20–2.5	20–2.5	20–2.5
Mean redundancy	11.2	5.6			
Unique reflections	16,355	10,300	39,539	34,505	34,419
Completeness (%)	96.5 (80.9)	82.4 (70.7)	96.4	96.1	95.8
I/σ	5.0 (2.3)	10.5 (2.2)	10.1	9.6	9.6
R_{sym}[a]	10.4 (22.2)	10.0 (31.2)	9.8	9.9	10.4
C. *Refinement*					
R_F[b] (%)	20.1 (21.8)	23.1 (36.9)			
R_{free}[c] (%)	25.5 (27.8)	28.3 (41.0)			
rmsd from ideal					
Bond lengths(Å)	0.009	0.009			
Bond angles (deg.)	1.129	1.133			
Average *B*-factor (Å^2)	31.5	25.6			
Ramachandran plot[a]					
Most favored (%)	91.6	89.6			
Allowed (%)	7.9	10.0			
Generously allowed (%)	0.0	0.2			
Disallowed (%)	0.4	0.2			
D. *Model content*					
Monomers/ASU	1	2			
Protein residues	4–272	9–272			
Ligand	One GB	Two PB			
Ethylene glycol	8				
Water molecules	175				

Refinement statistics were obtained from REFMAC5 and Ramachandran analysis was performed using PROCHECK. Values in parentheses correspond to the highest resolution shell (2.05–2.0 Å for the GB complex and 2.87–2.8 Å for the PB complex).

[a] $R_{sym} = <\Sigma_{h'}|I_{h'} - I_h|> /(\Sigma I_h)$.

[b] $R_F = \Sigma||F_o| - |F_c||/(\Sigma|F_o|)$, where F_c is the calculated structure factor.

[c] R_{free} is as R_F but calculated for 5% randomly chosen reflections that were omitted from the refinement procedure.

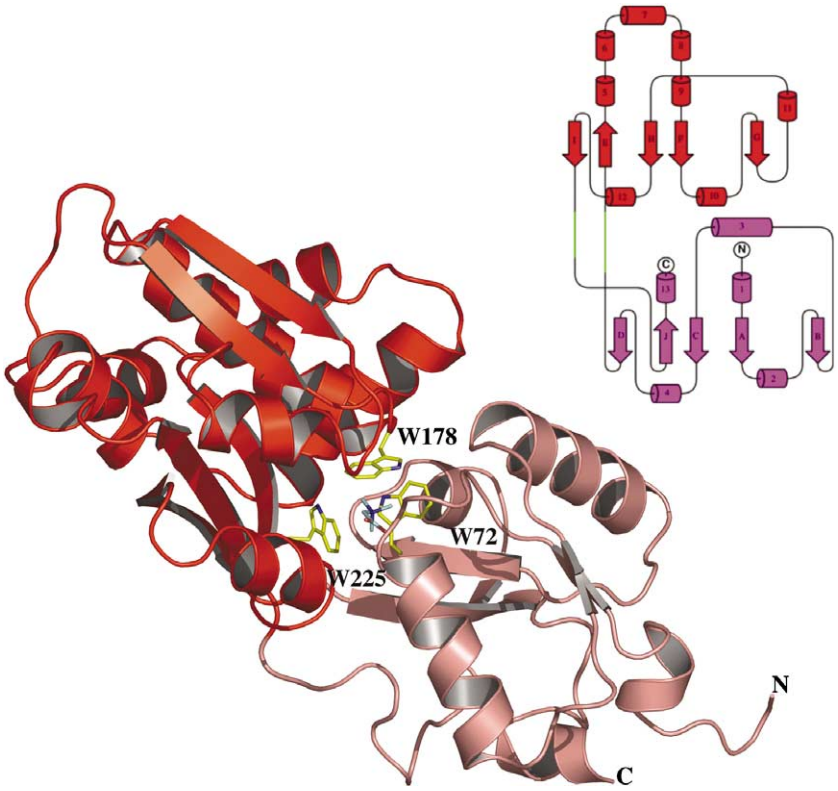

Figure 2. Overall structure of the OpuAC/GB complex. Domain 1 (residues 1–96 and 251–272) and domain 2 (residues 97–250) are colored light magenta and red. N and C termini are labeled at the bottom. The ligand, which is coordinated by three Trp residues colored yellow, is shown in ball-and-stick representation. The inset shows the topology of OpuAC.

II of the SBP superfamily.[25] The ligand GB, which is shown as sticks in Figure 2, is localized in a cleft between the two domains and flanked by strands S_E and S_I. Three tryptophan residues (shown as yellow sticks in Figure 2), Trp72, Trp178, and Trp225 stabilize the ligand inside the cleft.

Architecture of the ligand-binding site

The ligands, GB and PB, are bound in a deep groove between the two domains, which is composed of three Trp residues: Trp72 at the C terminus of S_C, Trp178 N-terminal of H8 and Trp225 C-terminal of S_H. The indole ring of Trp225 lies parallel with the β-pleated strands of domain 2 and is oriented perpendicular to the indole ring of Trp72 (Figure 2). The indole ring of Trp178 connects this geometry to an imaginary prism, the so-called Trp-prism, whereby it is tilted to the base of the prism (Figure 3 and Supplementary Data Figure S1). The geometry of the Trp-prism is perfectly suited to coordinate the ligand by two distinct contributions: (i) cation-π interaction[34] between the aromatic ring system of the indole moieties and the positive charge of the quaternary ammonium group that is delocalized across the four bound carbon atoms of GB and PB; and (ii) van der Waals interactions. Recently, the structure of ProX, the SBP of the compatible solute ABC-transporter ProU from *E. coli*, in complex with GB or PB was reported and exhibited similar features in ligand coordination.[32]

In addition to the cation-π interaction between the Trp-prism and the trimethyl ammonium group of GB and PB, the carboxylate group of the ligands forms hydrogen bonds with the backbone amide groups of Gly26 and Ile27 (Figure 3 and Supplementary Data Figure S1). In the OpuAC/GB complex, a third interaction, a hydrogen bond with the side-chain of His230 (Figure 3(a)), was observed. In the case of the OpuAC/PB complex, the distance between the side-chain of His230 and the carboxylate moiety of PB has increased from 2.6 Å to 4.7 Å. While the distances of all other interactions between the ligand, GB or PB, and OpuAC remain constant within experimental error, the loss of the hydrogen bond between His230 and PB might explain the 17-fold higher dissociation constant of PB. Using the relation $\Delta G = -RT\ln K$, the difference in dissociation constants of the two compatible solutes, 17(±1) μM for the GB/OpuAC and 295(±27) μM for the PB/OpuAC complex, accounts for a $\Delta\Delta G$ value of 6.7 kJ/mol. Since the energy of a hydrogen bond lies between 4 kJ/mol and 20 kJ/mol, the loss of the single hydrogen bond between the carboxylate moiety of PB and His230 might explain the 17-fold increase in affinity of the two complexes. Furthermore, no water molecule is located within the ligand-binding cavity. In contrast, the oligopeptide transporter OppA[40] uses water molecules to fine-tune its substrate specificity, demonstrating that specificity for the compatible solutes GB and PB by OpuAC is achieved solely by proper placement of the amino acid backbone and side-chains.

A superposition of GB and PB complexes is shown in Figure 4. As evident, the carboxyl moieties of GB (cyan) and PB (green) are displaced, while

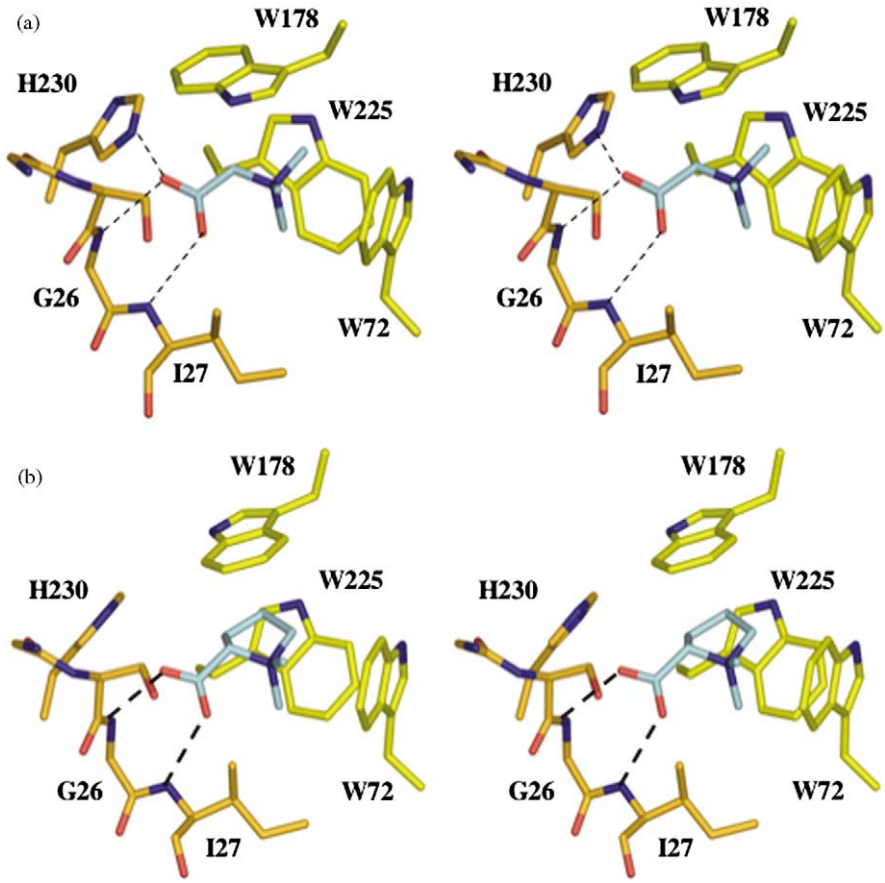

Figure 3. Stereo view of the ligand-binding pocket of OpuAC. (a) Coordination of GB and (b) coordination of PB. Interactions between the ligands, GB and PB respectively, are highlighted by broken green lines. For a Figure including a $1F_o - F_c$ omit map, please see Supplementary Data Figure S1.

the trimethyl ammonium groups align fairly well. Also, Trp72, Trp225, and His230 coincide, but Trp178 and the backbone region between Gly26 and Ile27 are displaced. Most striking is the 180° flip of the side-chain of Trp178. As a consequence of the different orientation of the side-chain of Trp178, no cation-π interaction but eight van der Waals contacts are formed between PB and Trp178 (data not shown). Taken together, the strongly reduced affinity of OpuAC for PB is a result of the missing

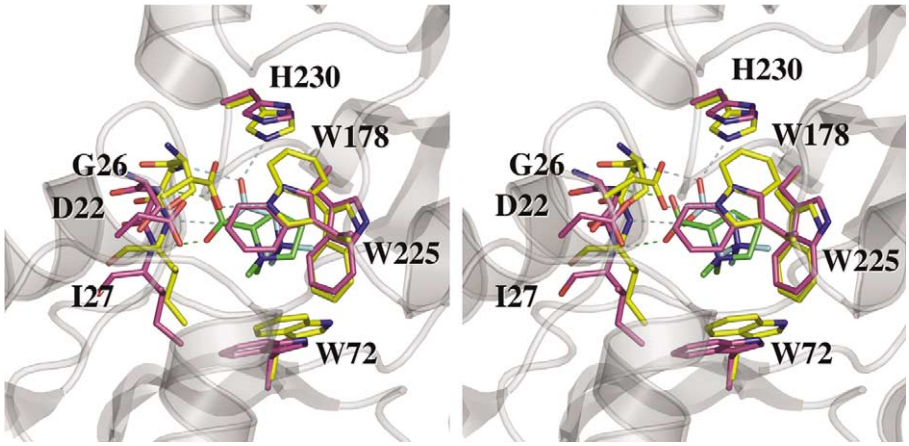

Figure 4. Stereo view of the superposition of the ligand-binding sites of the OpuAC/GB and OpuAC/PB complexes. Residues involved in GB coordination are shown in light blue and residues involved in PB binding are shown in magenta. Distances (in Å) are summarized in Table 1.

hydrogen bond between His230 and the ligand, which creates a subtle difference in the architecture of the ligand-binding pocket, most evident in the 180° flip of the indole side-chain of Trp178.

Domain–domain interactions stabilizing the closed conformation

The residues of OpuAC forming the ligand-binding pocket are localized in both domains in a rather even distribution. Gly26, Ile27, and Trp72 belong to domain 1, while Trp178, Trp225, and His230 are part of domain 2. As noted above, a major structural difference between the OpuAC/GB and OpuAC/PB complexes is the additional hydrogen bond, which involves His230 located in domain 2. Since this additional hydrogen bond is present also in the OpuAC/GB crystal structure derived from a selenomethionine-substituted protein, which contained two monomers in the asymmetric unit in an arrangement very similar to the OpuAC/PB complex, an influence of crystal packing on the ligand–protein interaction can be excluded (data not shown). This additional interaction influences the network of interactions between both domains (Supplementary Data Figure S2), which stabilize the closed conformation of OpuAC. In the OpuAC/GB complex four interactions between the two domains were evident (Supplementary Data Figure S2A), whereas only three were observed in the OpuAC/PB complex (Supplementary Data Figure S2B). As a result, different levels of interactions are observed across the domain–domain interface in the OpuAC/GB and OpuAC/PB complexes. Taken together, these changes form the structural basis for the 17-fold decreased affinity for the two ligands, despite their similar chemical nature.

Comparison of the OpuAC structure with that of the *E. coli* ProX protein

To further evaluate structural principles governing the binding of the compatible solutes GB and PB, we compared the crystal structures of OpuAC from *B. subtilis* and ProX from *E. coli* (Figure 5).[32] As shown in Figure 5(a), both proteins align well with an rmsd of 1.83 Å using 251 C^{α} atoms. Also evident is a domain dislocation:[25] domain 1 of OpuAC, which harbors the N terminus and the C terminus of the protein, superimposes with domain 2 of ProX and *vice versa*. Superposition of both proteins without employing the domain dislocation resulted in an rmsd of 2.4 Å using 157 C^{α} atoms. On the other hand, both domains of OpuAC align with each other with an rmsd of 3.1 Å using 51 C^{α} atoms. This emphasizes the presence of a domain dislocation in OpuAC and ProX. However, even after considering domain dislocation, one region of the structural alignment showed some degree of diversity. This stretch is located N-terminal to one of the two hinges of OpuAC and ProX (residues 95–99 and 231–235, respectively) and corresponds to a 17 residue insertion in ProX. On the basis of this superposition, a structure-based sequence alignment (Figure 7(b)) was constructed. Here, domain 1 of OpuAC is shown in red and domain 2 in dark blue, while domain 1 of ProX is shown in light blue and domain 2 in orange. To obtain this structure-based alignment, the sequences of OpuAC and ProX were split according to their domain boundaries. Despite the structural similarity of both proteins, which according to the DALI server† belong to the leucine-arginine-ornithine (LAO)[41] subfamily of SBPs,[27] there is only 19% identity on the amino acid level (indicated by asterisks in Figure 5(b)). According to the sequence homology between the transmembrane part of the ABC-transporter, OpuAB and ProW, OpuAC and ProX belong to the OTCN subfamily of SBP-dependent ABC-transporters.[42] Taking this ProX-specific insertion into account, one might speculate that a domain dislocation has occurred in the two proteins. The domain dislocation is further supported by the amino acid residues forming the ligand-binding pocket (Figure 5(c)). Two of the three Trp residues forming the Trp-prism of the *B. subtilis* protein (Trp72 and Trp225) and the residues coordinating the carboxylate group of GB (His230 and the backbone amide groups of Gly26 and Ile27) have counterparts located in the opposite domain of ProX (Trp188, Trp65, His69, Gly141, and Cys142). In contrast, Trp178 of OpuAC is located in domain 2, while the corresponding Trp of ProX, Trp140, is also located in domain 2.

It is also interesting to note that OpuAC contains only a single *cis*-peptide bond at position 226. This coincides with the *cis*-peptide bond of ProX[32] at position 66. As in ProX, this *cis*-peptide bond allows a sharp bend of the backbone, generating the proper environment for Trp225, which corresponds to Trp65 in ProX. However, the second *cis*-peptide bond in ProX at position 189 has no counterpart in OpuAC, which would be position 73. Thus, no restraint is imposed on the conformational flexibility of the adjacent Trp72 as it was proposed for ProX.[32] Nevertheless, both, Trp72 of OpuAC and Trp188 of ProX align fairly well (Figure 5(c)) and the restriction of Trp188 by a *cis*-peptide bond is not necessary for the formation of the Trp-prism. Another striking difference from ProX is the absence of disulfide bonds in the structure of OpuAC. The conserved motif G**C**NPGWG**C** of ProX,[32] where the two Cys residues forming a disulfide bond are highlighted in bold, possesses the homologue sequence G**I**DPGSG**I** in OpuAC and its five closest homologues, with the Cys replaced by Ile (highlighted in bold). Asp22 located in this sequence stretch forms a hydrogen bond with Trp178 across the domain–domain interface. This hydrogen bond does not exist in the OpuAC/PB structure and Trp178 has undergone a 180° flip (Figure 4 and Supplementary Data Figure S2). Thus,

† www.ebi.ac.uk/dali/

despite the absence of disulfide bonds in OpuAC, a conformational restraint is imposed on Trp178 by a hydrogen bond across the domain–domain interface. Since both types of restraints, an intra-domain disulfide bond or an inter-domain hydrogen bond, originate from the same stretch of amino acid residues, one might conclude that the conservation throughout related SBPs serves to create a switch to stabilize the closed conformation of the SBP and ensure stable ligand binding either *via* disulfide bond formation (ProX) or domain–domain interactions (OpuAC).

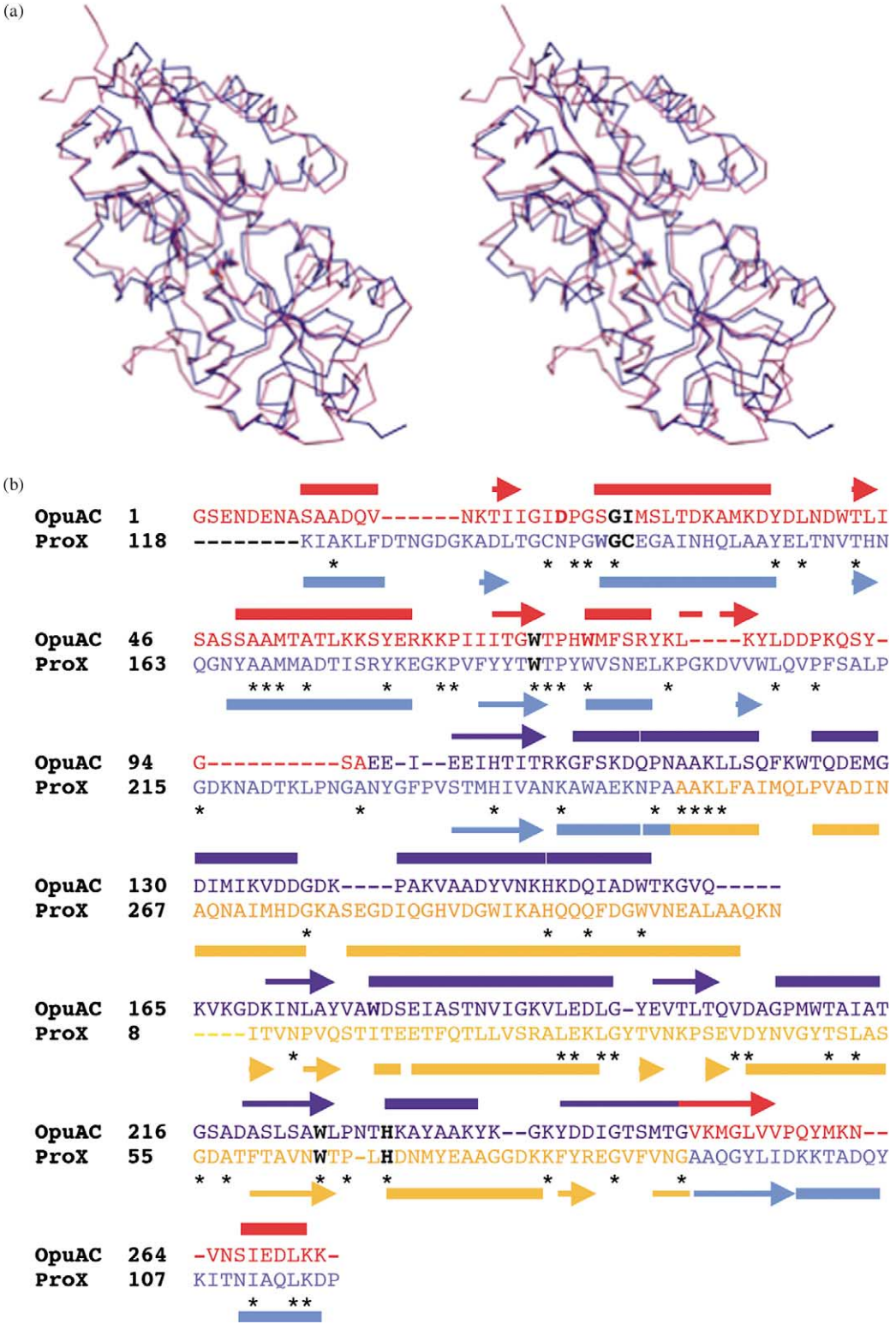

Figure 5 (*legend next page*)

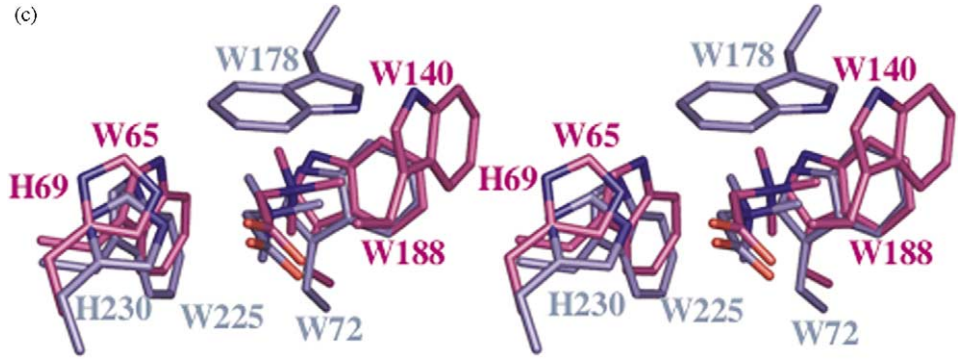

Figure 5. (a) Stereo view of the superposition of the OpuAC/GB (blue) and ProX/GB (magenta) complexes. The individual ligands are shown in stick representation. (b) Sequence alignment of OpuAC and ProX based on the structural alignment shown in (a). Residues involved in ligand binding are highlighted in black. Secondary structure elements (α-helix as squares and β-sheets as arrows) of OpuAC and ProX are shown above the corresponding sequences. Conserved residues are marked by an asterisk. (c) Stereo representation of the superposition of the ligand binding site of OpuAC (blue) and ProX (magenta). For further details, see the text.

Comparison with related compatible solute-binding proteins

OpuAC-related proteins were identified *via* a BLAST search in the NCBI database. A representative set of these proteins is summarized in Figure 6. These proteins align directly with the sequence of OpuAC (Figure 6(a)) or after splitting OpuAC into two sequence stretches, which are different from the domain boundaries (Figure 6(b) and discussed below). The examples given in Figure 6 include Gram-positive and Gram-negative bacteria, cyanobacteria and archaea. In the case of Gram-positive bacteria, the alignment contains binding proteins, which are either tethered to the plasma membrane *via* a lipid modification at the processed N terminus,[24,43] or fused to the C terminus of the transmembrane domain (TMD) of the ABC-transporters.[44,45] An example for these fused systems is OpuA from *Lactococcus lactis*.[42,44] In archaea, SBPs can be anchored to the plasma membrane *via* an N-terminal transmembrane segment,[31] or through a lipid modification at a Cys residue at the mature N terminus.[33] For clarity, the transmembrane segment of archaeal SBP and the TMD of the fused SBP of Gram-positive bacteria have been omitted from the alignment.

Immediately evident from the alignment is the absolute conservation of the three Trp, residues (Trp72, Trp178, and Trp225) forming the Trp-prism (Figure 6). No other aromatic amino acid is tolerated in these positions, implying an absolute need for Trp. This is in contrast to ProX, in which the Trp residues forming the "Trp-box" can be conservatively exchanged to Tyr, phe or even to Ala in two cases.[32] Furthermore, Asp22 (Supplementary Data Figure S2) is conserved (47% of all sequences). In the other sequences, the Asp is conservatively exchanged to Glu (53%). This further supports our notion that Asp22 is important. Another residue providing greater stability in the case of the OpuAC/GB complex is His230. Although not conserved as

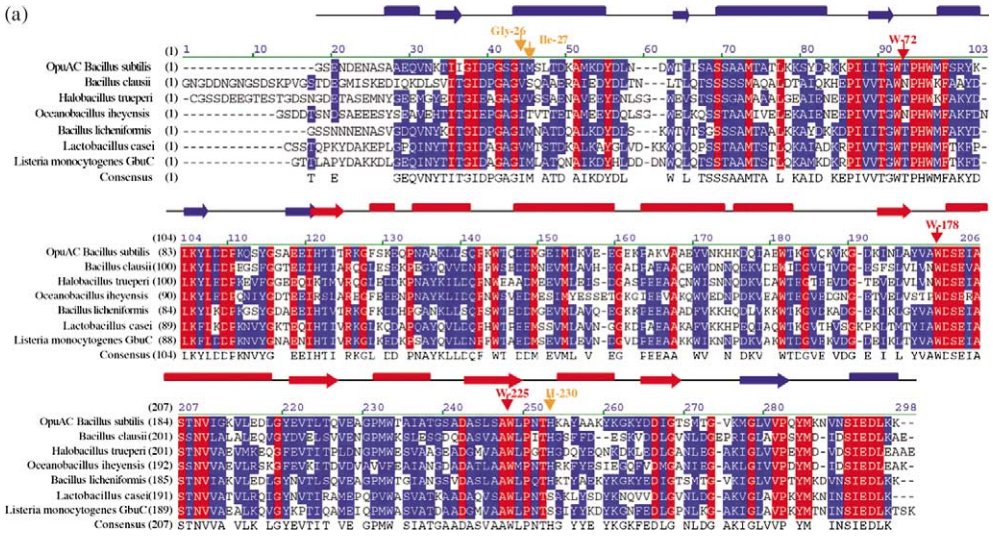

Figure 6 (*legend next page*)

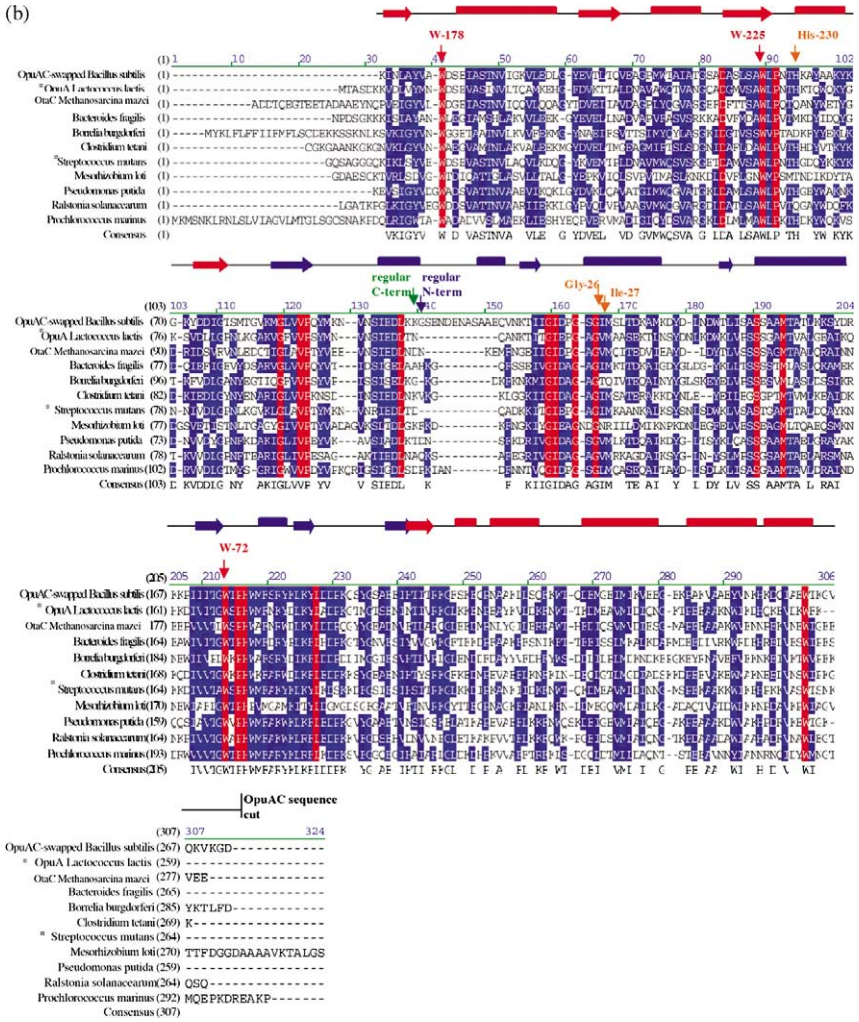

Figure 6. Sequence alignment of OpuAC from *B. subtilis* with putative glycine betaine binding proteins from various organisms. For simplicity, only the name of the organism is given. In the case of archaea, the transmembrane-spanning segments anchoring the proteins in the cytoplasmic membrane have been removed. In the case of binding proteins from bacteria (highlighted by an asterisk), which are fused to the transmembrane region of the transporter, the TMD has been deleted as well. Secondary structure elements of OpuAC are given above the sequence. To visualize the individual domains of OpuAC, secondary structure elements of domain 1 are given in blue and secondary structure elements of domain 2 are given in red. Residues involved in ligand binding are indicated. Conserved residues have been highlighted in red. (a) Direct sequence alignments with other putative glycine betaine-binding proteins. (b) Sequence alignments with glycine-binding proteins after the sequence of OpuAC was split between amino acid residues 168/169. Further, putative glycine betaine-binding proteins are given in Supplementray Data Figure S3.

strictly as Asp22 and the three Trp residues, exchange of His230 occurs only by amino acids that are capable of forming salt-bridges or hydrogen bonds. Another interesting residue is Pro227, which enables *cis*-peptide bond formation at residue 226. Like the three Trp residues, Pro227 is absolutely conserved among all SBP analyzed. Obviously, one has to expect a *c*is-peptide bond at the corresponding position (position 226 in OpuAC).

The sequence alignments presented can be subdivided in three distinct families. The first family (Figure 6(a)) can be aligned directly with OpuAC and comprises proteins such as the SBPs from *Bacillus licheniformis* or *Listeria monocytogenes*. The *L. monocytogenes* GbuC protein is the ligand-binding protein of a functionally characterized GB transporter system.[46] The second family includes SBPs such as ProX from *E. coli* (Figure 5(b)) or *A. fulgidus* (not shown), which align with OpuAC after taking the possibility of SBPs to undergo domain dislocations, which is described in detail by Fukami-Kobayashi *et al.*, into account.[25] This suggests that this sort of alignment might be useful to identify putative GB-specific SBPs. However, and to our big surprise, the largest family, family 3 (Figure 6(b)), was identified only after the sequence of OpuAC was split into two parts (part 1, residues 4–168, and part 2, residues 169–273). This split did not coincide with the structural domain definitions (structural domain 1, residues 4–94 and 253–272; structural domain 2, residues 100–247). The relation of family 3 therefore relies on a, so far unidentified, split of the sequence of SBPs, which

we term sequence domains (for further discussion see below). SBPs belonging to this class are, for example, the GB-binding protein OpuBC from *L. lactis*,[44,45] or OtaC from *Methanosarcina mazei*.[47] OpuBC from *L. lactis* is an example of ABC-transporters containing a fusion of SBP and TMD.[42] The putative GB-binding proteins with such a fusion of the TMD with the SBP that have been identified in our sequence search are indicated by an asterisk in Figure 6(b) and comprise a substantial number of the putative GB-specific proteins.

As shown in Figure 7(a), the two sequence domains (red and blue), which are the foundation of family 3, are located within the conserved lobes of OpuAC. A structural superposition using the central β-sheet as anchor points revealed an rmsd of 2.3 Å between both sequence domains (Figure 7(b)), implying that some sort of duplication might have arisen from such an ancestor protein. The five helices located in sequence domain 1 have no counterpart in sequence domain 2. This region was identified by Fukami-Kobayashi *et al*. as a point of divergency in SBPs due to different insertions occurring at this location.[25] Further support of the relevance of the identified sequence domains arises from a comparison of the Trp residues forming the Trp-prism in OpuAC. As shown in Figure 7(c), Trp72 and Trp225 align fairly well based on the superposition shown in Figure 7(b). On the other hand, Trp172 has no structural counterpart. This is reminiscent of the situation in ProX from *E. coli*. As pointed out above, Trp172 is not conserved between

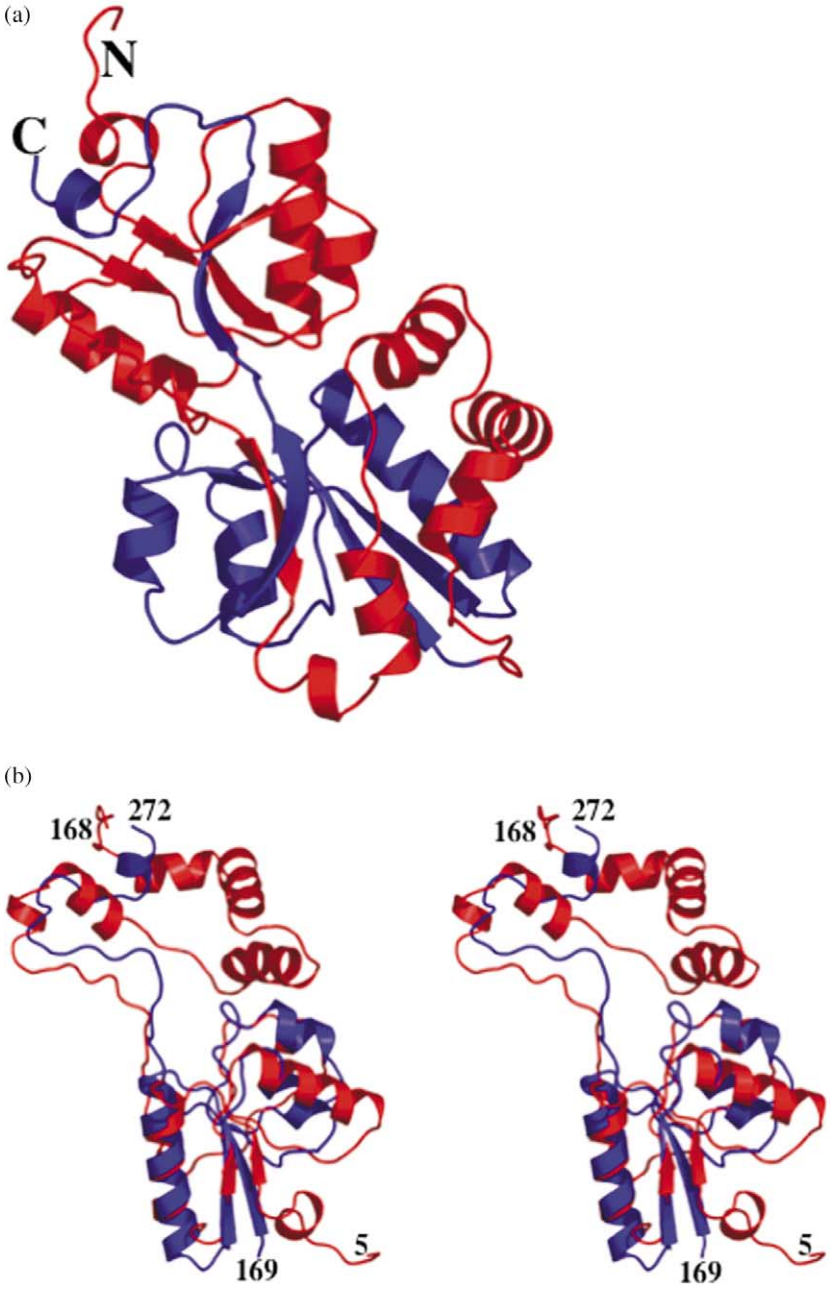

Figure 7 (*legend next page*)

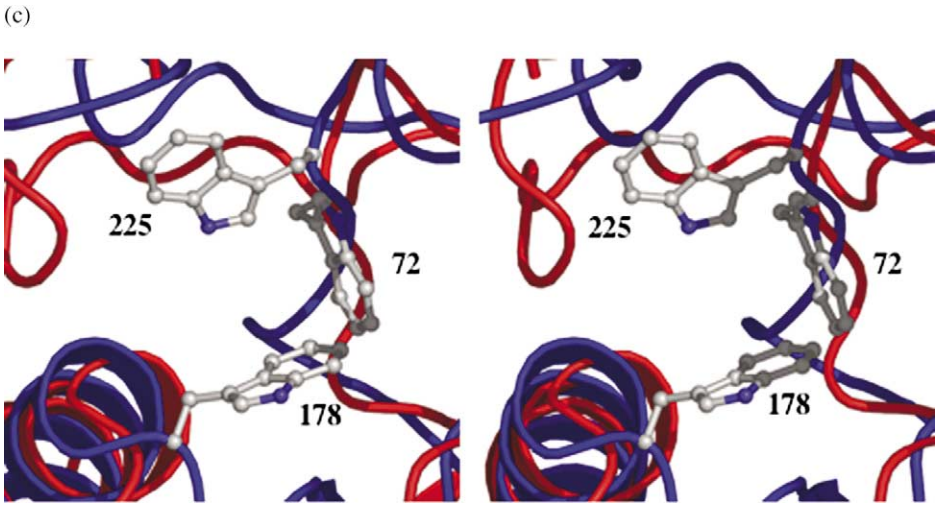

Figure 7. (a) Crystal structure of OpuAC. The color-coding highlights the parts of the protein that have been used for the sequence alignment shown in Figure 6(c). The N and C termini of OpuAC are indicated. (b) Stereo view of a structural superposition of two sequence-based regions of OpuAC. Positions of the amino acid residues where the protein was split for the alignment are indicated. For the superposition, amino acid residues 17–77 and 171–231 were used (rmsd of 2.3 Å for 60 C^{α} atoms). (c) Stereo view of the three Trp residues forming the "Trp-prism" in OpuAC. Orientation is identical to (b).

OpuAC and ProX, while Trp72 and Trp225 coincide with Trp188 and Trp65. Obviously, the evolutionary aspects of the identified sequence domains need further inspection, but the sequence and structural analysis have established clearly that the crystal structure of OpuAC might serve as a template for three different families of SBPs able to bind compatible solutes.

Conclusions

Here, we have described the structural characterization of the SBP OpuAC from *B. subtilis* and its functional interaction with its two main substrates, the compatible solutes GB and PB. Although similar in chemical structure, the equilibrium binding constants differed by a factor of ~20. Based on the crystal structures, which were determined at 2.0 Å (GB complex) and 2.8 Å (PB complex), it became evident that a Trp-prism coordinates the delocalized positive charge of the ligand. However, different interactions between the amino acid residues that form the ligand-binding cavity of OpuAC and the carboxy moiety of GB and PB suggest a structural explanation for the different binding constants. Furthermore, the observed subtle difference in ligand coordination resulted in an alternate pattern of interactions in the domain–domain interface, which correlates difference in K_D with structural changes. Thus, the difference in affinity of the two substrates can now be explained structurally. Using the OpuAC amino acid sequence as a template for database searches, it became evident that the Trp residues forming the Trp-prism in OpuAC as well as other amino acid residues participating either in ligand binding or domain–domain interactions are highly conserved in all proteins inspected. The *B. subtilis* OpuAC protein is thus a suitable model structure for a wide variety of binding proteins with a putative specificity for GB and PB, including the examples that are based on the newly identified sequence domains.

Materials and Methods

Protein expression and purification

A soluble form of OpuAC (residues 22–274),[20] lacking its signal peptide and Cys21, was produced as a fusion protein with the maltose-binding protein (MBP) and a factor Xa cleavage site using the expression plasmid pBKB76. *E. coli* BL21 (DE3) cells harboring pBKB76 were cultivated in minimal medium containing 2% (w/v) glucose, 31 mM Na_2HPO_4, 19 mM KH_2PO_4, 25 mM NH_4Cl, 5 mM sodium citrate, 10 mM $MgSO_4$, 20 mg/l of $FeCl_3 \cdot 6H_2O$, 4 mg/l of $Zn(OAc)_2 \cdot 2H_2O$, 10 mg/l of thiamine/HCl and 100 μg/ml of ampicillin at 30 °C and 220 rpm. Protein expression was induced at $A_{550} = 0.8$ with 1 mM IPTG for 3 h. Subsequently, cells were harvested by centrifugation (4000*g*, 15 min, 4 °C) and lysed by ultrasonication. The cytosolic protein fraction was purified as described.[20] In brief, the fusion protein was isolated by amylose-affinity chromatography. OpuAC was liberated by factor Xa proteolysis and separated from MBP and factor Xa by anion-exchange chromatography. To obtain homogenous protein, an additional SEC step was introduced using a Superdex 200 column (16/60 prep grade, Amersham Pharmacia, Freiburg) equilibrated with TBS buffer (10 mM Tris–HCl, pH 7.3). Fractions corresponding to monomeric OpuAC were pooled if purity was greater than 99% (as judged from SDS-PAGE), concentrated to 10 mg/ml by ultrafiltration (Amicon Ultra, 5 kDa MWCO, Millipore, Eschborn, Germany) and stored at 4 °C until further use. Protein concentration was determined spectroscopically at 280 nm using a theoretical molar extinction coefficient of 60,900 $M^{-1}\,cm^{-1}$.

For seleno-methionine (SeMet) substituted OpuAC, cells were grown in glucose minimal medium (see above), supplemented prior to induction with 100 mg each of Lys, Phe and Thr (10 mg/ml) and 50 mg each of Ile, Leu and Val (5 mg/ml) to inhibit methionine biosynthesis. Incorporation of SeMet was achieved by subsequent addition of 60 mg of L-SeMet (6 mg/ml). SeMet-substituted OpuAC was purified and stored as described above.

Ligand binding experiments

To assess the binding affinity of OpuAC for GB (purchased from Sigma) or PB (purchased from Extrasynthese, France), intrinsic tryptophan fluorescence of OpuAC was monitored from 305 nm to 450 nm using a Flurolog (Horiba, Edison, NJ, USA). The excitation wavelength was set to 295 nm, slit-width of 5 nm, and temperature was maintained at 22($\pm$1) °C using a circulating waterbath. Different amounts of GB or PB in 10 mM Tris–HCl (pH 7.3) (from a 0.5, 5.0 or 50 mM stock solution) were added to OpuAC samples (250 nM in 10 mM Tris–HCl, pH 7.3) and fluorescence was measured 5 min after ligand addition to allow for equilibration. Spectra in the presence and in the absence of protein were subtracted and the maximum emission wavelength ($\lambda_{em,max}$) was determined by an automated peak search routine using the program DataMax (Jobin Yvon-Spex Instruments, Edison, NJ, USA). Upon binding of ligand to OpuAC, a blue shift of $\lambda_{em,max}$ from 345 nm in the absence of ligand to 336 nm under conditions of ligand saturation was observed. This shift was normalized according to equation (1) and used to determine the dissociation constants of OpuAC/GB or OpuAC/PB complexes, respectively. All data represent the average of three independent measurements with the standard deviations given as errors.

$$\frac{\lambda-\lambda_0}{\lambda_\infty-\lambda_0}=\frac{L_0}{K_D-L_0} \quad (1)$$

Here, λ is the maximum emission wavelength. Indices 0 and ∞ indicate the absence of ligand and conditions of ligand saturation, respectively. L_0 denotes the ligand concentration and K_D is the dissociation constant of OpuAC/GB or OpuAC/PB complexes, respectively.

Crystallization

Crystals of OpuAC were obtained by the hanging-drop technique. Prior to crystallization, OpuAC at a concentration of 10 mg/ml was incubated with 1 mM GB or 3 mM PB for 30 min on ice. Subsequently, 1 µl of protein solution was mixed with 1 µl of reservoir solution and 0.5 µl of 100 mM L-cysteine. The reservoir solution contained 100 mM Tris–HCl (pH 8.25), 150 mM NH_4OAc and 15% (w/v) PEG 4000. Crystal needles appeared at 274 K between 48 h and 72 h, and grew to their final dimensions (300 µm$\times$100 µm$\times$50 µm) within two weeks. Appropriate crystals were transferred into cryo-buffer (150 mM Tris–HCl (pH 8.25), 20% (w/v) ethylene glycol, 200 mM NH_4OAc, 20% (w/v) PEG 4000) and flash-frozen in liquid nitrogen. Under identical conditions, SeMet-substituted OpuAC/GB produced only small needles that showed a weak and anisotropic diffraction pattern. Crystal quality and dimensions were improved by crystallizing at 295 K and by refining the concentrations of NH_4OAc to 50 mM and PEG 4000 to 12% (w/v). Crystals of the SeMet-substituted OpuAC/GB were transferred directly into cryo-buffer and flash-frozen in liquid nitrogen.

Data collection, structure determination and refinement

A three-wavelength MAD data set from a single SeMet-substituted OpuAC/GB crystal and native data sets from single OpuAC/GB and OpuAC/PB crystals were collected at beamline BW-6 (DESY, Hamburg) with an MAR CCD and were processed using DENZO and SCALEPACK.[48] After localization of 16 of the 20 expected SeMet sites of the MAD data set using SHELXD,[49] an initial electron density of OpuAC/GB was determined using SOLVE and RESOLVE.[50] After the initial model was built in O,[51] REFMAC 5 was used in the subsequent refinement steps.[52] The model for the SeMet-substituted OpuAC/GB structure was refined to 2.7 Å by manual building into the $F_{obs}-F_{cal}$ and $2F_{obs}-F_{cal}$ electron densities. Since the SeMet-substituted and native crystals of OpuAC/GB crystallized in different space groups (Table 1), it was necessary to refine the model of the SeMet-substituted crystal to an R_{free} value below 30%. The structure of the native OpuAC/GB data set was subsequently solved by molecular replacement using AMoRe with the refined structure of the SeMet-substituted OpuAC/GB crystal as template.[53] Model building and refinement were again performed iteratively in O and REFMAC5.

Automated water picking was performed using ARP/wARp[54] with a cut-off of 3σ for individual water molecules that were checked manually for appropriate density. 5% of the data was excluded from refinement to calculate the R_{free} value for cross-validation. The quality of the model was verified with PROCHECK,[55] and the resulting parameters are summarized in Table 1. The OpuAC/PB model was calculated by molecular replacement using AMoRE and the native OpuAC/GB structure as template. Model building and refinement were again performed iteratively in O and REFMAC5. In the early stages of refinement, strict non-crystallographic symmetry between the two monomers was applied, which was released in the last four cycles of refinement.

Data Bank search and sequences alignment

Protein homologues of OpuAC were searched *via* the National Center for Biotechnology Information†. Sequence alignments were performed using CLUSTAL W, taking possible domain dislocations (derived from the structural boundaries of both domains of OpuAC) into account using the Vector NTI software package. The sequence domains were detected upon manual inspection of sequences, which had been aligned dierectly. Here, it became evident that, for example, the Trp residues forming the ligand-binding pocket could not be aligned properly without applying a split of OpuAC at the indicated positions.

Structure alignments

Structure alignments were performed using LSQMAN employing standard settings.

Figure preparation

Structure Figures were prepared using PYMOL‡.

† NCBI, www.ncbi.nlm.nih.gov/BLAST/
‡ http://pymol.sourceforge.net/

Protein Data Bank accession codes

Coordinates have been deposited in the RCSB Protein Data Bank under accession codes 2B4L (OpuAC/GB) and 2B4M (OpuAC/PB).

Acknowledgements

We thank Alexander Wiedemann, Clemens Schneeweis, Nils Hanekop, Robert Ernst, Jelena Zaitseva, and Stefan Jenewein for support and stimulating discussions. We are indebted to Uli Ermler and Harmut Michel, MPI of Biophysics, for in-house facilities and Gleb Bourenkov, BW-6, DESY/Hamburg, for his constant support and advice during data collection and refinement. We are indebted to Vickie Koogle for her help with editing the manuscript. L.S. thanks Robert Tampé and the members of the Institute of Biochemistry for support and encouragement. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 395 to E.B. (project B1)), the Max-Planck Institute for terrestrial microbiology (Marburg) to E.B., the Fond der chemischen Industrie to E.B., and SPP 1070, (grant BR576/5-1 to E.B. and grant Schm1279/4-1 to L.S.).

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2005.12.085

References

1. Wood, J. M. (1999). Osmosensing by bacteria: signals and membrane-based sensors. *Microbiol. Mol. Biol. Rev.* **63**, 230–262.
2. 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. *TIBS*, **23**, 143–148.
3. Galinski, E. A. & Trüper, H. G. (1994). Microbial behaviour in salt-stressed ecosystems. *FEMS Microbiol. Rev.* **15**, 95–108.
4. Kempf, B. & Bremer, E. (1998). Uptake and synthesis of compatible solutes as microbial stress responses to high osmolality environments. *Arch. Microbiol.* **170**, 319–330.
5. Morbach, S. & Kramer, R. (2002). Body shaping under water stress: osmosensing and osmoregulation of solute transport in bacteria. *ChemBioChem*, **3**, 384–397.
6. Roessler, M. & Müller, V. (2001). Osmodaptation in bacteria and archaea: common principles and differences. *Environ. Microbiol.* **3**, 742–754.
7. Bremer, E. & Krämer, R. (2000). Coping with osmotic challenges: osmoregulation through accumulation and release of compatible solutes in bacteria. In *Bacterial Stress Responses* (Storz, G. & Hengge-Aronis, R., eds), pp. 79–97, ASM Press, Washington, DC.
8. Lippert, K. & Galinski, E. A. (1992). Enzyme stabilization by ectoine-type compatible solutes: protection against heating, freezing and drying. *Appl. Microbiol. Biotechn.* **37**, 61–65.
9. 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.
10. Arakawa, T. & Timasheff, S. N. (1985). The stabilization of proteins by osmolytes. *Biochem. J.* **47**, 411–414.
11. 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.
12. Le Rudulier, D., Strom, A. R., Dandekar, A. M., Smith, L. T. & Valentine, R. C. (1984). Molecular biology of osmoregulation. *Science*, **224**, 1064–1068.
13. Bremer, E. (2002). Adaptation to changing osmolarity. In Bacillus subtilis *and its Closest Relatives* (Sonenshein, A. L., Hoch, J. A. & Losick, R., eds), pp. 385–391, ASM Press, Washington, DC.
14. Kempf, B. & Bremer, E. (1995). OpuA, an osmotically regulated binding protein-dependent transport system for the osmoprotectant glycine betaine in *Bacillus subtilis*. *J. Biol. Chem.* **270**, 16701–16713.
15. 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. *et al*., eds), vol. 1, pp. 1175–1209, ASM Press, Washington, DC.
16. Davidson, A. L. & Chen, J. (2004). ATP-binding cassette transporters in bacteria. *Annu. Rev. Biochem.* **73**, 241–268.
17. Schmitt, L. & Tampe, R. (2002). Structure and mechanism of ABC transporters. *Curr. Opin. Struct. Biol.* **12**, 754–760.
18. Horn, C., Bremer, E. & Schmitt, L. (2003). Nucleotide dependent monomer/dimer equilibrium of OpuAA, the nucleotide-binding protein of the osmotically regulated ABC transporter OpuA from *Bacillus subtilis*. *J. Mol. Biol.* **334**, 403–419.
19. Horn, C., Bremer, E. & Schmitt, L. (2005). Functional overexpression and in vitro re-association of OpuA, an osmotically regulated ABC-transport complex from *Bacillus subtilis*. *FEBS Letters*, **579**, 5765–5768.
20. 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.
21. Kappes, R. M., Kempf, B. & Bremer, E. (1996). Three transport systems for the osmoprotectant glycine betaine operate in *Bacillus subtilis*: characterization of OpuD. *J. Bacteriol.* **178**, 5071–5079.
22. Holtmann, G. & Bremer, E. (2004). Thermoprotection of *Bacillus subtilis* by exogenously provided glycine betaine and structurally related compatible solutes: involvement of Opu transporters. *J. Bacteriol.* **186**, 1683–1693.
23. Brigulla, M., Hoffmann, T., Krisp, A., Völker, A., Bremer, E. & Völker, U. (2003). Chill induction of the SigB-dependent general stress response in *Bacillus subtilis* and its contribution to low-temperature adaptation. *J. Bacteriol.* **185**, 4305–4314.
24. 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.

25. 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.
26. 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.
27. Tam, R. & Saier, M. H., Jr (1993). Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol. Rev.* **57**, 320–346.
28. 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.
29. 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.
30. Sharff, A. J., Rodseth, L. E., Spurlino, J. C. & Quiocho, F. A. (1992). Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. *Biochemistry*, **31**, 10657–10663.
31. Albers, S. V., Elferink, M. G., Charlebois, R. L., Sensen, C. W., Driessen, A. J. & Konings, W. N. (1999). Glucose transport in the extremely thermoacidophilic *Sulfolobus solfataricus* involves a high-affinity membrane-integrated binding protein. *J. Bacteriol.* **181**, 4285–4291.
32. Schiefner, A., Breed, J., Bosser, L., Kneip, S., Gade, J., Holtmann, G. *et al*. (2004). Cation-π 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.
33. 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.
34. Ma, J. C. & Dougherty, D. A. (1997). The cation-π interaction. *Chem. Rev.* **97**, 1303–1324.
35. Lakowicz, J. R. (1999). *Principles of Fluoresccence Spectroscopy* (2nd edit.), Kluwer Academic, New York.
36. Vivian, J. T. & Callis, P. R. (2001). Mechanisms of tryptophan fluorescence shifts in proteins. *Biophys. J.* **80**, 2093–2109.
37. May, G., Faatz, E., Villarejo, M. & Bremer, E. (1986). Binding protein dependent transport of glycine betaine and its osmotic regulation in *Escherichia coli* K12. *Mol. Gen. Genet.* **205**, 225–233.
38. Haardt, M., Kempf, B., Faatz, E. & Bremer, E. (1995). The osmoprotectant proline betaine is a major substrate for the binding-protein-dependent transport system ProU of *Escherichia coli* K-12. *Mol. Gen. Genet.* **246**, 783–786.
39. Barron, A., Jung, J. U. & Villarejo, M. (1987). Purification and characterization of a glycine betaine binding protein from *Escherichia coli*. *J. Biol. Chem.* **262**, 11841–11846.
40. Tame, J. R. H., Sleigh, S. H., Wilkinson, A. J. & Ladbury, J. E. (1996). The role of water in sequence-independent ligand binding by an oligopeptide transporter protein. *Nature Struct. Biol.* **3**, 998–1001.
41. 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.
42. van der Heide, T. & Poolman, B. (2002). ABC transporters: one, two or four extracytoplasmic substrate-binding sites? *EMBO Rep.* **3**, 938–943.
43. Sutcliffe, I. C. & Russell, R. R. (1995). Lipoproteins of gram-positive bacteria. *J. Bacteriol.* **177**, 1123–1128.
44. Obis, D., Guillot, A., Gripon, J. C., Renault, P., Bolotin, A. & Mistou, M. Y. (1999). Genetic and biochemical characterization of a high-affinity betaine uptake system (BusA) in *Lactococcus lactis* reveals a new functional organization within bacterial ABC transporters. *J. Bacteriol.* **181**, 6238–6246.
45. 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. USA*, **97**, 7102–7106.
46. Ko, R. & Smith, L. T. (1999). Identification of an ATP-driven, osmoregulated glycine betaine transport system in *Listeria monocytogenes*. *Appl. Environ. Microbiol.* **65**, 4040–4048.
47. Roessler, M., Pfluger, K., Flach, H., Lienard, T., Gottschalk, G. & Muller, V. (2002). Identification of a salt-induced primary transporter for glycine betaine in the methanogen *Methanosarcina mazei* Go1. *Appl. Environ. Microbiol.* **68**, 2133–2139.
48. 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, pp. 307–326, Academic Press, London.
49. Schneider, T. R. & Sheldrick, G. M. (2002). Substructure solution with SHELXD. *Acta Crystallog. sect. D*, **58**, 1772–1779.
50. Terwilliger, T. C., Berendzen, J. & Automated, M. A. D. (1999). and MIR structure solution. *Acta Crystallog. sect. D*, **55**, 849–861.
51. 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 Crystallog. sect. A*, **47**, 110–119.
52. Murshudov, G., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–255.
53. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallog. sect. D*, **50**, 157–163.
54. Lamzin, V. S. & Wilson, K. S. (1993). Automated refinement of protein models. *Acta Crystallog. sect. D*, **49**, 129–147.
55. 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. Crystallog.* **26**, 283–291.

Edited by I. B. Holland

(Received 27 September 2005; received in revised form 22 December 2005; accepted 29 December 2005)
Available online 13 January 2006