The compatible solute-binding protein OpuAC from *Bacillus subtilis*: ligand-binding, site-directed mutagenesis and crystallographic studies

Sander H.J. Smits¹,‡, Marina Höing²,‡, Justin Lecher¹, Mohamed Jebbar³,⁺, Lutz Schmitt¹,* and Erhard Bremer²

‡: Both authors contributed equally to this work

1) Institute of Biochemistry, Heinrich Heine University Duesseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany
2) Laboratory for Microbiology, Department of Biology, Philipps University Marburg, Karl-von-Frisch Str., 35032 Marburg, Germany
3) Departement Osmoregulation chez les Bacteries, Universite de Rennes 1, UMR-CNRS 6026, Rennes, Frances

⁺ Present address: Laboratory of Extreme Environments, Microbiology, University of West Brittany (Brest), European Institute of Marine studies, Technopôle Brest-Iroise, F-29280 Plouzané, France
Running title: Analyzing the ligand-binding site of OpuAC

Proposed section: Genetics and Molecular Biology

* Corresponding author:
Lutz Schmitt
Institute of Biochemistry
Heinrich Heine University Duesseldorf
Universitätsstr. 1
40225 Düsseldorf, Germany
Phone: +49(0)211-81-10773
Fax: +49(0)211-81-15310
Email: lutz.schmitt@uni-duesseldorf.de
Abstract

In the soil bacterium *Bacillus subtilis*, five transport systems work in concert to mediate the import of various compatible solutes that counteract the deleterious effects of increases in the osmolarity of the environment. Among these five systems, the ABC transporter OpuA, which catalyses the import of glycine betaine and proline betaine has been studied in detail in the past. Here, we demonstrate that OpuA is capable of importing the sulfobetaine dimethylsulfonylacetate (DMSA). Since OpuA is a classic ABC importer that relies on a substrate-binding protein priming the transporter with specificity and selectivity, we analyzed the OpuA-binding protein, OpuAC, by structural and mutational means with respect to DMSA binding. The determined crystal structures of OpuAC in complex with DMSA at 2.8 Å resolution and a detailed mutational analysis of these residues revealed a hierarchy within the amino acids participating in substrate binding. This finding is different to other binding proteins that recognize compatible solutes. Furthermore important principles that enable OpuAC to specifically bind various compatible solutes were uncovered.
The soil bacterium *Bacillus subtilis* is equipped with five transport systems (Opu: osmoprotectant uptake) that allow the import of a large number of compatible solutes (4, 5, 25). Compatible solutes are low-molecular weight organic osmolytes that balances the osmotic potential of the cytoplasm with that of the environment. Three of the five compatible solutes transport systems (OpuA, OpuC and OpuD) mediate the uptake of the widespread found osmoprotectant glycine betaine (22, 24). Glycine betaine can also be synthesized by *B. subtilis* from the precursor choline (3), which is acquired from the environment via the osmoregulated OpuB and OpuC transporters (23). The Opu transport systems also mediate the osmoregulated uptake of several other compatible solutes (4, 19, 25). For instance, proline betaine is taken up by *B. subtilis* via the OpuA and OpuC transporters (B. Kempf and E. Bremer; unpublished results).

OpuD is a secondary transporter that belongs to the BCCT-family (*Betaine-Choline-Carnitine-Transporter*) of uptake systems (22). In contrast, OpuA, OpuB and OpuC are members of the ABC-(ATP binding cassettes) transporters, which use the energy released by ATP hydrolysis to transport substrates against a concentration gradient (16, 42). In general, ABC-transporters are composed of four modules. The two nucleotide-binding domains and two transmembrane domains can be arranged in any possible combination. However, ABC-import
systems such as OpuA, OpuB and OpuC contain a fifth module, a substrate-binding protein. This substrate-binding protein captures the substrate with high affinity and delivers it to the cognate transport system for subsequent ATP-dependent import. In Gram-negative bacteria binding proteins diffuse freely in the periplasmic space, while they are lipid-anchored in the cytoplasmic membrane in Gram-positive bacteria such as *B. subtilis* (1, 24, 26). However, it was recently shown that binding proteins can even be fused to the transmembrane domain of the ABC-transporter (36, 43).

Despite this variation, all substrate-binding proteins from ABC-transporters analyzed by X-ray crystallography today display a bilobal architecture. The ligand-binding site is located in a deep cleft situated between these two lobes and residues located on both lobes usually contribute to substrate binding (45). Based on structural and kinetic investigations, a “Venus fly-trap” mechanism was proposed to explain the ligand-binding mechanism on a molecular level (32, 39). Here, substrate-binding proteins undergo constant opening-closing motions in the absence of the ligand and the amino acids connecting both domains act as pivot point in such a hinge-bending motion. Upon ligand binding, the equilibrium between the open and closed state of the binding protein is shifted towards the so-called “liganded-closed” state and the ligand is bound in a cleft located between both domains.

The ABC transporter OpuA from *B. subtilis* (19) has been analyzed
functionally and structurally by \textit{in vivo} or \textit{in vitro} studies of either the whole transporter or its isolated components (17, 18, 20, 24, 26). The OpuA system consists of the cytoplasmic membrane-associated ATPase OpuAA (18), the integral membrane transport component OpuAB (17) and the extracellular ligand binding protein OpuAC (24). This latter protein is tethered to the cytoplasmic membrane via a lipid modification at its amino-terminus (26). The crystal structure of OpuAC in complex with two ligands, glycine betaine or proline betaine, has been reported recently (20). The ligand-binding pocket of OpuAC is formed by three tryptophan residues arranged in a “prism-like” geometry suitable to coordinate the positive charge of the trimethylammonium group of glycine betaine or the dimethylammonium group of proline betaine by cation-π interactions. Additionally, hydrogen bonds with the carboxylate moiety of the ligand are formed. Structural differences between the OpuAC/glycine betaine and OpuAC/proline betaine complexes occur within the ligand-binding pocket that allow a structural explanation for the drastic affinity differences of OpuAC for these two ligands. The $K_D$ for the binding of glycine betaine by OpuAC is $17 \pm 1 \mu M$, whereas the $K_D$ for the binding of proline betaine is $295 \pm 27 \mu M$ (20).

**Dimethylsulfonioacetate (DMSA), the closest sulfonium analog of glycine betaine** (Fig. 1) is found as a secondary osmolyte in certain species of marine algae (7, 9). Previous studies have shown that DMSA (also referred to as sulfobetaine or dimethylthetin) (7) can function as an osmoprotectant for \textit{E. coli} where it is accumulated via the ProP and ProU compatible solute uptake systems.
(9). DMSA also serves as an osmoprotectant for *Pseudomonas aeruginosa* PAO1 (10) and the lactic acid bacterium *Tetragenococcus halophila* (2). Furthermore, DMSA is a substrate for the periplasmic binding protein from the glycine betaine and choline transporter OusB from *Erwinia crysanthemi* (8). Interestingly, uptake of DMSA in *Sinorhizobium meliloti* is toxic and it becomes only osmoprotective in mutants that are unable to dimethylate this sulfobetaine (38).

To further analyze the binding principles of compatible solutes to OpuAC, it is desirable to assess the importance of single tryptophans participating in the formation of the Trp-prism and other amino acids contributing to ligand binding (20). Therefore, we have performed a mutational study of the ligand binding-site. Furthermore, we present the crystal structure of OpuAC in complex with the compatible solute dimethylsulfonioacetate (DMSA), an efficient osmoprotectant for *B. subtilis* and a substrate of the OpuA transporter.
MATERIALS AND METHODS

**Bacterial strains, plasmids and culture conditions.** The *E. coli* strains used in this study were maintained on Luria-Bertani medium (33) and were propagated at 37°C. For the selection of *E. coli* strains carrying derivatives of the expression vector pASK-IBA6 (IBA, Göttingen, Germany), ampicillin (100 µg ml⁻¹) was added to the liquid cultures and agar plates. Overproduction of the *B. subtilis* OpuAC protein and its mutant derivatives was carried out in the *E. coli* strain BL21 (F gal met r' m' hsdS(λDE3) (Stratagene, La Jolla, CA, USA). For OpuAC overproduction, the plasmid-carrying BL21 strain was propagated in a defined minimal medium (MMA) (33) supplemented with 100 µg ampicillin ml⁻¹, 0.2% (w/v) Casamino Acids and 0.5% (w/v) of glucose as carbon source. Mutant derivatives of the *opuAC*-expression plasmid pMH24 were recovered after transformation into Epicurian coli® XL1-Blue (recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F’ proAB lacI°ZΔM15Tn10(Tet’)] (Stratagene, La Jolla, CA, USA). The *B. subtilis* strains RMKB24 [Δ(opuA::erm)4 Δ(opuBD::tet)23 opuC-20::Tn10(spc) Δ(opuD::neo)2] and RMKB34 [opuA + Δ(opuBD::tet)23 opuC-20::Tn10(spc) Δ(opuD::neo)2] are derivatives of the wild type strain JH642 (trpC pheA1) (J. Hoch, Scripps Research Institute, CA, USA). The genetic construction of these two *B. subtilis* mutants has been described by Kappes et al. (23).

*B. subtilis* strains were grown in Spizizen’s minimal medium (SMM) with 0.5% (w/v) glucose as the carbon source and supplemented with L-tryptophan.
(20 µg ml⁻¹), L-phenylalanine (18 µg ml⁻¹) and a solution of trace elements (15). When required, the osmotic strength of SMM was increased by the addition of NaCl from a 5 M stock solution. Experiments that continuously monitored the growth of the \textit{B. subtilis} cultures, 20 ml pre-warmed SMM-Medium containing 1.2 M NaCl in a 100 ml Erlenmeyer flask was inoculated with a late-exponential-phase pre-culture grown in SMM with 0.4 M NaCl to an \textit{OD}_{578} of 0.1. These cultures were grown in a shaking water bath (set at 200 rpm) at 37°C. Compatible solutes (glycine betaine, proline betaine and dimethylsulfonioacetate) were added to \textit{B. subtilis} cultures to a final concentration of 1 mM each, as required.

**Chemicals.** Glycine betaine was purchased from Sigma-Aldrich (Munich, Germany), proline betaine was obtained from Extrasynthèse (Genay Cedex, France) and dimethylsulfonioacetate was synthesized as described by Ferger and Vigneaud (13).

**Overexpression and purification of the recombinant OpuAC protein in \textit{E. coli}**. Plasmid pMH24 is a \textit{B. subtilis} opuAC⁺ derivative of the expression vector pASK-IBA6 (IBA, Göttingen, Germany). In this recombinant plasmid, the \textit{opuAC} coding region (without its own signal sequence and the codon specifying the amino-terminal cysteine residue of the mature OpuAC protein) (26) is positioned under the control of the anhydrotetracycline-inducible \textit{tet} promoter present in the vector pASK-IBA6. This allows induction of the transcription of the
opuAC gene to high levels in the host strain BL21. The opuAC coding region has been inserted in pASK-IBA6 in-frame with an upstream ompA signal sequence and the codons for a Strep-TagII affinity peptide. This allowed the secretion of the Strep-TagII-OpuAC 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, Göttingen, Germany). Strain BL21(pMH24) in 5 liters defined MMA to an OD$_{578}$ = 0.1 was inoculated from an overnight culture of BL21(pMH24) prepared in the same medium. opuAC transcription was induced at an OD$_{578}$ = 0.7 of the culture by the addition of 0.2 µg ml$^{-1}$ anhydrotetracycline. Growth of the culture of BL21(pMH24) was then continued for 1.5 h at 37°C with avid stirring. Subsequently, cells were harvested by centrifugation (10 min, 3,000 x g). To release periplasmic proteins from the BL21(pMH24) cells, the cell pellet was re-suspended in 50 ml of ice-cold buffer P (50 mM Tris (tris(hydroxymethyl)aminomethan)-HCl pH 8.0, 100 mM NaCl, 500 mM sucrose) and incubated for 30 min on ice. Soluble periplasmic proteins were isolated by two subsequent centrifugation steps. First, the supernatant was centrifuged for 15 min at 21,000 x g to remove cellular debris. Subsequently, the supernatant was re-centrifuged for 60 min at 120,000 x g to remove denatured proteins. The cleared, soluble periplasmic protein fraction was then loaded onto a 10 ml-Strep-Tactin column (IBA, Göttingen, Germany), pre-equilibrated with 10 bed volumes of buffer W (50 mM Tris-HCl, 100 mM NaCl, pH 8.0). After the column was washed with 10 bed volumes of buffer W, bound proteins were eluted from the affinity resin with buffer E (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 2.5 mM
Two forms of the recombinantly produced OpuAC were released from the periplasmic fraction: (i) the non-processed OmpA-Strep-TagII-OpuAC fusion and (ii) the processed Strep-TagII-OpuAC form. To remove the OmpA signal sequence and the Strep-TagII from unprocessed OmpA-Strep-TagII-OpuAC and the Strep-TagII from processed Strep-TagII-OpuAC, OpuAC containing fractions were pooled and incubated overnight at 23°C with 0.5 U Factor Xa (Novagen, Darmstadt, Germany) per 10 µg of OpuAC in buffer E in the presence of 4 mM CaCl₂. Complete removal of the OmpA signal sequence and the Strep-TagII from OpuAC was verified by SDS-PAGE. OpuAC was concentrated to a volume of approximately 500 µl using VIVASPIN 4 (Vivascience, Hannover, Germany) concentrator columns (exclusion size, 10 kDa). Subsequently, the protein was passed through a HiTrapQ anion exchange column (GE Healthcare, Munich, Germany) to remove Factor Xa from the protein preparation. The column was washed with a buffer containing 25 mM Tris-HCl and 25 mM NaCl (pH 8.3). OpuAC does not bind to the HiTrapQ material and passes through the column, whereas Factor Xa bound to the HiTrapQ material. Finally, isolated OpuAC was dialyzed against 2 x 5 liters of 10 mM Tris-HCl (pH 7.0) at 4°C overnight and stored at 4°C until further use. In general, approximately 1.5 mg of pure OpuAC protein was obtained per liter of E. coli culture. The functionality of the purified OpuAC protein was assessed by fluorescence spectroscopy using changes in the intrinsic tryptophan fluorescence of OpuAC upon substrate binding (e.g.
glycine betaine) as detailed by Horn et al. (20). Protein concentrations were estimated based on the theoretical molar extinction coefficient of OpuAC yielding the following correlation: $A_{280}=1.0$ corresponds to $0.5 \text{ mg ml}^{-1}$ OpuAC. OpuAC used for crystallization experiments was concentrated to approximately $10 \text{ mg ml}^{-1}$ using VIVASPIN 4 (Vivascience, Hannover, Germany) concentrator columns (exclusion size, 10 kDa).

**Site directed mutagenesis of the opuAC gene.** To determine the individual contribution of the amino acids forming the Trp prism (Trp$^{72}$, Trp$^{178}$ and Trp$^{225}$) and His$^{230}$ to the stability of the OpuAC/glycine betaine, OpuAC/proline betaine and OpuAC/DMSA complexes (20), the corresponding codons in the *opuAC* gene were changed via site-directed mutagenesis using the QuikChange site directed-mutagenesis kit (Stratagene, La Jolla, CA, USA) and custom synthesized mutagenic primers (Biomers, Ulm, Germany). These experiments were conducted with the *opuAC*+ plasmid pMH24. The entire coding region of the mutant *opuAC* genes was sequenced to ensure the presence of the desired mutation and the absence of undesired alterations in the *opuAC* coding region. Double and triple mutants were generated from the plasmids bearing the corresponding single or double mutations at the desired positions. The following mutant *opuAC* variants were generated on plasmid pMH24: pMH26 (Trp$^{72}\rightarrow\text{Ala}$ [TGG$\rightarrow$GCG]), pMH27 (Trp$^{72}\rightarrow\text{Phe}$ [TGG$\rightarrow$TTT]), pMH28 (Trp$^{72}\rightarrow\text{Tyr}$ [TGG$\rightarrow$TAT]), pMH29 (Trp$^{178}\rightarrow\text{Ala}$ [TGG$\rightarrow$GCG]), pMH30 (Trp$^{178}\rightarrow\text{Phe}$ [TGG$\rightarrow$TTT]), pMH31 (Trp$^{178}\rightarrow\text{Tyr}$ [TGG$\rightarrow$TAT]), pMH32 (Trp$^{225}\rightarrow\text{Ala}$ [TGG$\rightarrow$GCG]), pMH33 (Trp$^{225}\rightarrow\text{Phe}$ [TGG$\rightarrow$TTT]), pMH34 (Trp$^{225}\rightarrow\text{Tyr}$ [TGG$\rightarrow$TAT]), pMH35 (His$^{230}\rightarrow\text{Ala}$ [CAC$\rightarrow$GCA]), pMH36 (His$^{230}\rightarrow\text{Phe}$ [CAC$\rightarrow$TTC]), pMH37 (His$^{230}\rightarrow\text{Tyr}$ [CAC$\rightarrow$TAT]), pMH38 (Trp$^{72}\rightarrow\text{Ala}$ [TGG$\rightarrow$GCG], His$^{230}\rightarrow\text{Ala}$ [CAC$\rightarrow$GCA]), pMH39 (Trp$^{72}\rightarrow\text{Phe}$ [TGG$\rightarrow$TTT], His$^{230}\rightarrow\text{Ala}$ [CAC$\rightarrow$GCA]), pMH40 (Trp$^{72}\rightarrow\text{Tyr}$ [TGG$\rightarrow$TAT], His$^{230}\rightarrow\text{Ala}$ [CAC$\rightarrow$GCA]), pMH41 (Trp$^{72}\rightarrow\text{Ala}$ [TGG$\rightarrow$GCG], His$^{230}\rightarrow\text{Phe}$ [CAC$\rightarrow$TTC]), pMH42 (Trp$^{72}\rightarrow\text{Phe}$ [TGG$\rightarrow$TTT], His$^{230}\rightarrow\text{Phe}$ [CAC$\rightarrow$TTC]), pMH43 (Trp$^{72}\rightarrow\text{Tyr}$ [TGG$\rightarrow$TAT], His$^{230}\rightarrow\text{Phe}$ [CAC$\rightarrow$TTC]), pMH44 (Trp$^{72}\rightarrow\text{Ala}$ [TGG$\rightarrow$GCG], His$^{230}\rightarrow\text{Tyr}$ [CAC$\rightarrow$TAT]), pMH45 (Trp$^{72}\rightarrow\text{Phe}$ [TGG$\rightarrow$TTT], His$^{230}\rightarrow\text{Tyr}$ [CAC$\rightarrow$TAT]), pMH46 (Trp$^{72}\rightarrow\text{Tyr}$ [TGG$\rightarrow$TAT], His$^{230}\rightarrow$
Determination of the dissociation constants of the OpuAC/compatible solute complexes. The dissociation affinity of the OpuAC/glycine betaine, proline betaine or DMSA complexes was determined as described by Horn et al. (20). In brief, the intrinsic tryptophan fluorescence of OpuAC was monitored from 300 nm to 450 nm using a Cary Eclipse fluorescence spectrometer (Varian, Surrey, UK). The excitation wavelength was set to 295 nm, slit width of 5 nm and the temperature was maintained at room temperature ($22 \pm 1^\circ$ C) using a circulating water bath. Different amounts of the substrates were titrated to 1 ml OpuAC samples (250 nM in 10 mM Tris-HCl, pH7.0) and fluorescence was measured after equilibration (5 min). Changes in the maximum emission wavelength (glycine betaine or proline betaine), determined by an automated
peak search routine, or changes in the fluorescence intensity (DMSA) were
plotted against substrate concentration after background correction. Upon
binding of glycine betaine or proline betaine to OpuAC, a blue shift of $\lambda_{em,max}$ from
345 nm in the absence of ligand to 336 nm under substrate saturation conditions
was observed. The changes of the emission maxima or fluorescence intensity
due to the concentration of bound substrates could by analyzed using a 1:1
binding site model employing equation 1 (glycine betaine and proline betaine) or
equation 2 (DMSA):

$$\lambda_{em,max} = \lambda_{em,max0} + (\Delta \lambda_{em,max} \times [S_0]/([S_0] + K_D))$$  equation 1

Here, $\lambda_{em,max}$ is the emission wavelength maximum for a given substrate
concentration; $\lambda_{em,max0}$ is the emission wavelength maximum without substrate;
$\Delta \lambda_{em,max}$ is the maximal emission wavelength maximum shift; $S_0$ is the substrate
concentration; $K_D$ is the dissociation constant.

$$F = F_0 + (\Delta F \times [S_0]/([S_0] + K_D))$$  equation 2

Here, $F$ is the fluorescence intensity for a given substrate concentration; $F_0$ is the
fluorescence intensity without substrate; $\Delta F$ is the maximal change in
fluorescence intensity; $S_0$ is the substrate concentration; $K_D$ is the dissociation
constant. All $K_D$ measurements of OpuAC and its mutant derivatives that are
summarized in Table 1 represent the average of at least three independent
measurements, with a standard deviation given as errors.

**Crystallization of the OpuAC/DMSA complex, data collection and**
model refinement. Crystals of the OpuAC/DMSA complex were obtained under conditions similar to the ones described for the glycine betaine and proline betaine complexes (20). Prior to crystallization, OpuAC (at a concentration of 10 mg ml\(^{-1}\)) was incubated with 3 mM DMSA. 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\(_4\)OAc and 15% (w/v) PEG 4000. Crystal plates appeared at room temperature after several weeks, with final dimensions of 200 µm x 100 µm x 30 µm. Crystals were transferred into cryo-buffer (150 mM TrisHCl (pH 8.3), 20% (w/v) ethylene glycol, 200 mM NH\(_4\)OAc, 20% (w/v) PEG 4000) and flash-frozen in liquid nitrogen. Diffraction data were collected at 100 K at the EMBL beam line BW7A at DESY, Hamburg. Data were indexed and scaled with XDS and further analyzed using the CCP4 program package (6). The structure was solved by molecular replacement using AMORE (35) with the OpuAC/glycine betaine monomer (20) as search model. Four monomers were found in the asymmetric unit and the initial structure was further improved by manual rebuilding in 2F\(_o\)-F\(_c\) and 1F\(_o\)-F\(_c\) electron density maps using COOT (12) and subsequent rounds of refinement employing REFMAC5 (34). During the initial refinement cycles strict NCS restraints (27) were maintained, which were released in the last five cycles of refinement. The quality of the model was verified with the MolProbity server (www.molprobity.biochem.duke.edu/) and is summarized in Table 2. R\(_F\) and R\(_{free}\) values are 28.5 % and 36.4 %, respectively. Although at the higher end of the range expected at this resolution (28), they are still within the limits and the
quality of the electron density map allowed a detailed analysis of the structure.

**Protein Data Bank Accession Code.** Coordinates of the OpuAC/DMSA complex have been deposited in the RCSB Protein Data Bank under accession code 3CHG.
DMSA is a substrate for the OpuA transporter and confers osmoprotection in *B. subtilis*. Previous growth studies and transport assays have shown that DMSA is also an effective osmoprotectant for *B. subtilis* and is acquired by the cell via the OpuA, OpuC and OpuD osmolyte transport systems (G. Nau-Wagner, M. Jebbar and E. Bremer; unpublished results). Hence, glycine betaine (22-24) and its sulfur analog DMSA are taken up via the same three transport systems. We analyzed the uptake of DMSA via the OpuA transporter by growth experiments. Both glycine betaine and DMSA (provided at a concentration of 1 mM each) were very effective osmoprotectants in strain RMBKB34 that is OpuA+, but defective in the compatible solute uptake systems OpuB, OpuC and OpuD (Fig. 2A). When the OpuA system is inactivated by a gene disruption mutation in an otherwise OpuB+, OpuC− and OpuD− background (RMBKB24), osmoprotection by glycine betaine is completely blocked as has been reported previously by Kappes *et al.* (22) and osmoprotection by DMSA is greatly reduced (Fig. 2B). We conclude from these growth experiments, that DMSA is a substrate for the OpuA transporter, but that a fourth uptake route for DMSA seems to operate in *B. subtilis* that remains to be identified. Since DMSA can enter the cell via the OpuA system, this sulfobetaine should be recognized by the ligand-binding protein (OpuAC) of the OpuA transporter.
DMSA is bound by the purified *B. subtilis* OpuAC protein. We overexpressed the *B. subtilis* opuAC gene in *E. coli* and purified the recombinant OpuAC protein by affinity chromatography to homogeneity (data not shown). To determine the affinities of glycine betaine, proline betaine and DMSA to the purified OpuAC protein, an intrinsic Trp-fluorescence based binding assay was employed. A spectra and the corresponding binding curve for DMSA is shown in Figure 3. Binding of glycine betaine and proline betaine to OpuAC resulted in a blue shift of the emission spectra of 9 nm (glycine betaine, data not shown) and 6 nm (proline betaine, data not shown), respectively. This shift in emission maximum was subsequently used to determine the dissociation constants of the complexes according to equation 1 (see MATERIALS AND METHODS). A 1:1 binding isotherm described the experimental data adequately and $K_D$ values could be calculated to $22 \pm 4 \ \mu M$ and to $267 \pm 6 \ \mu M$ for glycine betaine and proline betaine, respectively (Table 1). These values are in very good agreement with those previously determined (glycine betaine: $K_D= 17 \pm 1 \ \mu M$; proline betaine: $K_D= 295 \pm 27 \ \mu M$) for these two OpuAC substrates by Horn *et al.* (20). In contrast to the binding of glycine betaine and proline betaine to OpuAC, binding of DMSA to OpuAC did only induce a marginal blue shift of the emission maximum ($2 \ \text{nm}; \ data \ not \ shown$). Therefore, the decrease in fluorescence intensity was used to determine the binding constant according to equation 2 (see MATERIALS AND METHODS) assuming again a 1:1 binding isotherm (Fig. 3B). Here, a calculated $K_D$ value of $102 \pm 11 \ \mu M$ was determined, an affinity that is between the values determined for the other two OpuAC ligands (Table 1).
From a chemical point of view, the structures of the individual ligands (Fig. 1) do not provide any ready explanation for these differences in affinity. All three ligands contain a carboxylate moiety and a delocalized positive charge. Since the $K_D$ value of DMSA is in between the one for glycine betaine and proline betaine, the nature of the delocalized positive charge does not seem to be relevant for the apparent affinity differences (Table 1).

**Crystal structure of OpuAC with its ligand DMSA.** To gain inside into the molecular determinants that govern binding of DMSA by OpuAC, we crystallized this ligand-binding protein in the presence of DMSA and determined the crystal structure of the OpuAC/DMSA complex at a resolution of 2.8 Å. The structure was solved by molecular replacement using the recently determined OpuAC/glycine betaine structure (20) as a search model and refined using REFMAC5 (34). A summary of the data collection statistics and refinement details as well as the model content are given in Table 2. As expected, the overall fold of OpuAC in complex with DMSA (RCSB Protein Data Bank accession code 3CHG) is similar to that of the recently published OpuAC complexes containing either glycine betaine (RCSB Protein Data Bank accession code 2B4L) or proline betaine (RCSB Protein Data Bank accession code 2B4M), respectively (20). The OpuAC/DMSA complex exhibits the characteristic bilobal protein fold observed for many binding proteins of prokaryotic ABC transport systems (32, 39, 45).
The quality of the initial electron density map of the OpuAC protein with the bound DMSA allowed an unambiguous placement of the ligand and thereby the localization of the sulfonium moiety of DMSA despite the medium resolution (2.8 Å) of the overall OpuAC/DMSA structure. In contrast to the structure of the OpuAC/glycine betaine complex (20), the asymmetric unit of the OpuAC/DMSA crystal structure contained four protomers. Since the root mean square deviation of the individual protomers in the asymmetric unit was smaller than 1 Å, the description of the structure will be restricted to a single protomer (monomer D). Two of these protomers are related via non-crystallographic symmetry, which was used during the initial steps of structure refinement but released in the last cycles of refinement. As shown in Fig. 4, the overall architecture of the DMSA binding site was identical to the OpuAC/glycine betaine complex (20) and is composed of three tryptophans (Trp<sup>72</sup>, Trp<sup>178</sup>, Trp<sup>225</sup>) and one histidine (His<sup>230</sup>). Additionally, the carboxylate of DMSA interacts with the backbone amids of Gly<sup>26</sup> and Ile<sup>27</sup> via hydrogen bonds (3.5 Å and 2.9 Å, respectively), as has been previously observed both in the OpuAC/glycine betaine and OpuAC/proline betaine complexes (20). These two hydrogen bonds together with the interaction of DMSA with His<sup>230</sup> (distance of 3.2 Å) fix the carboxylate moiety of DMSA within the ligand-binding site (Fig. 4). The dimethylsulfonylum group of DMSA interacts with the individual tryptophans of the “Trp-prism” (Trp<sup>72</sup> Trp<sup>178</sup> Trp<sup>225</sup>) (20) via cation-π interactions (31). All distances range between 3.5-4.0 Å, perfectly fitting the van der Waals interactions and fulfilling the requirements of cation-π interactions (30). However, a closer inspection reveals that only 19 cation-π interactions...
interactions and 6 van der Waals interactions are present in the DMSA complex, while 22 cation-\(\pi\) interactions were determined for the OpuAC/glycine betaine complex (20). In contrast, only 6 cation-\(\pi\) and 15 van der Waals interactions are found in the OpuAC/proline betaine complex (13). More important, however, is the fact that the interaction distance between His\(^{230}\) and the ligands glycine betaine, DMSA and proline betaine is 2.6 Å, 3.5 Å and 4.7 Å, respectively. The distance of His\(^{230}\) to proline betaine is even beyond the effective distance of a salt bridge. Horn et al. (20) used this distance argument, to explain the drastically lower affinity of OpuAC for proline betaine (\(K_D = 295 \mu M \pm 27 \mu M\)) compared to the affinity of OpuAC for glycine betaine (\(K_D = 17 \mu M \pm 1 \mu M\)). In light of the OpuAC/DMSA structure reported here and the previously reported analysis of the OpuAC/glycine betaine and OpuAC/proline betaine complexes (20), the combination of different numbers of cation-\(\pi\) and van der Waals interactions contribute significantly to ligand binding. Furthermore, important for substrate binding appears also the presence (in the case of glycine betaine and DMSA) or the absence (in the case of proline betaine) of an interaction between the ligand and His\(^{230}\).

To compare the positioning of the three OpuAC substrates within the ligand binding sites, we superimposed the OpuAC/glycine betaine, OpuAC/proline betaine and OpuAC/DMSA crystal structures. As shown in Figure 5, the ligand-binding site of OpuAC/DMSA complex matches almost perfectly the OpuAC/glycine betaine and the OpuAC/proline betaine structure. Next to the slightly different conformations of His\(^{230}\) in the three structures (Fig. 5), the
most important difference between the OpuAC structures is the conformation of the indole moiety of Trp$^{178}$. In the OpuAC/proline betaine complex, it is flipped nearly 180° with respect to the position in the OpuAC/DMSA and the OpuAC/glycine betaine complex. Thus, it is tempting to speculate that the orientation of this side chain might contribute to the overall affinity of OpuAC to either its high-(glycine betaine), medium-(DMSA) or low-affinity (proline betaine) ligands.

Site-directed mutagenesis of the OpuAC ligand-binding pocket. The analysis of the three OpuAC structures in complex with the various ligands provides a molecular framework to describe the interactions and affinities of different compatible solutes to OpuAC. To analyze the contribution of individual amino acid residues within the OpuAC binding pocket to ligand binding, we performed a site-directed mutagenesis study. We mutagenized the opuAC$^+$ overexpression plasmid pMH24 using the QuikChange site directed-mutagenesis kit (Stratagene) and a set of mutagenic DNA-primers. In total, we generated 16 opuAC mutants (see MATERIALS AND METHODS). Each of these mutant opuAC genes were overexpressed in strain BL21 and the variant OpuAC proteins were purified to homogeneity by affinity chromatography. The purified mutant OpuAC proteins were analyzed for binding and affinity to glycine betaine, proline betaine and DMSA by using fluorescence spectroscopy and these data are summarized in Table 1. For a structural summary see Figure 4.
The generated mutations can be principally sub-divided into four classes: (i) The three tryptophan residues forming the “Trp-prism” (Trp\textsuperscript{72}, Trp\textsuperscript{178}, and Trp\textsuperscript{225}) were mutated individually to alanine residues. (ii) The three Trp residues were separately mutated to either phenylalanine or tyrosine residues. (iii) We also simultaneously changed the three tryptophan residues forming the “Trp-prism” to either Phe or Tyr. (iv) His\textsuperscript{230} was substituted to alanine.

To determine the influence of Trp-residues within the OpuAC binding site on complex stability, we assessed individual Ala substitutions for substrate binding. As shown in Table 1, mutation of any of the three tryptophans to alanine resulted in a complete loss of ligand binding. This is different from the situation found in the glycine betaine/proline betaine binding protein ProX from \textit{E. coli}. Here, three Trp-residues, arranged in a “box-like” structure, constitute the binding surface for the trimethlyammonium headgroup of glycine betaine and the dimethylammonium-headgroup of proline betaine via cation-π interactions (40). Two of these Trp-residues (Trp\textsuperscript{65} and Trp\textsuperscript{140}) can be changed to Ala-residues with modest effects on substrate binding. However, the replacement of Trp\textsuperscript{188} with Ala results in a complete loss of binding of glycine betaine (40). Consequently, in the “box-like” arrangement of the Trp-residues found within the binding site of ProX, only a single Trp-residue is critical for substrate binding. The other two Trp-residues appear to stabilize the substrate within the ligand-binding pocket (33).

As elaborated by Dougherty and co-workers (11, 31), the strength of a
cation-π interaction between a ligand and a protein decreases from Trp to Tyr to Phe, thus following the decrease in the electronegative potential of the indole, phenole and benzole side chain of the amino acids. We therefore individually changed Trp^{72}, Trp^{178} and Trp^{225} of OpuAC to either Phe or Tyr-residues. Exchange of a single tryptophan to either Phe or Tyr resulted in a complex response with respect to ligand binding and this was dependent on the tryptophan mutated and the ligand analyzed (Table 1). These substitutions caused in general substantial decreases in affinity of OpuAC for its three ligands and in several cases no substrate binding could be detected at all (Table 1). This result is surprising, since the site-directed change of the Trp-residues to either Phe- or Tyr-residues within the *E. coli* ProX ligand binding site, has essentially no influence on ligand binding (40). Furthermore, mutational analysis of the aromatic residues within the binding site of the ectoine/hydroxyectoine binding protein EhuB from *Sinorhizobium meliloti* (14, 21) revealed that the strength of the cation-π interaction is of key importance for the efficiency of substrate binding. An aromatic box composed of Phe^{24}, Tyr^{60} and Phe^{80} forms a central part of the ligand binding site of the EhuB protein allowing substrate binding with *K_D* values in a low μM range (14). Substitutions of these aromatic residues by Trp, the amino acid with the strongest electronegative potential and hence best suited for cation-π interactions (11, 31), created super-binding variants of EhuB that bind both ectoine and hydroxyectoine with *K_D* values in a low nM range (14).

Simultaneous change of either two tryptophans (Trp^{72/178}, Trp^{72/225} or Trp^{178/225}) or all three tryptophans to either Phe- or Tyr-residues completely
abolished ligand binding (Table 1). This clearly demonstrates that a single Trp-
residue paired with two other aromatic amino acids is not sufficient for OpuAC to
bind any of the three substrates tested. We are thus tempted to speculate that
the “Trp-prism” found in OpuAC has been evolutionary optimized for ligand
binding in such a way that only minor variations are permitted. This argument is
strengthened by our data base searches. We aligned the amino acid sequence of
64 OpuAC related proteins (Figure 6) and found that Trp$^{72}$, Trp$^{178}$ and Trp$^{225}$ are
completely conserved, regardless whether the proteins align directly with OpuAC,
whether the alignment requires the inversion of N- and C-terminal domains (20)
or whether the ligand-binding portion is fused to the transmembrane domain of
the corresponding ABC transport systems (Figure 6; further details are provided
in the Figure legend).

A rather surprising result is obtained when Trp$^{178}$ is changed to Tyr. This
substitution causes a drastic decrease in glycine betaine binding, abolishes
DMSA binding, but increases substantially the binding of proline betaine.
Currently, we have no biochemical or structural explanation for these findings.

The analysis of the crystal structures of the OpuAC/glycine betaine and
OpuAC/proline betaine complexes suggested that an additional hydrogen bond
between the carboxylate of glycine betaine and His$^{230}$ was responsible for the
17fold higher affinity of OpuAC for glycine betaine than for proline betaine (20). In
a His$^{230}$ to Ala substitution, this critical hydrogen bond will be abolished, thus
predicting that the binding affinity of OpuAC for glycine betaine should be
strongly decreased and should approach that of proline betaine. The data
summarized in Table 1 demonstrate that this is indeed the case and thus support
the prediction made by Horn et al. (20) based on the interpretation of the OpuAC
structures. Since His$^{230}$ makes also contacts to the carboxylate of DMSA binding
of the sulfobetaine is also substantially reduced by the replacement of His$^{230}$ with
Ala (Table 1). These findings are consistent with the view that interactions
between the carboxylate of the ligands and His$^{230}$ make important contributions
to the overall affinity of the OpuAC protein to its ligands. Thus, His$^{230}$ has a prime
role in modulating affinity of the OpuAC/compatible solute complexes. Although
not as completely conserved as the three tryptophan residues forming the “Trp-
prism” of OpuAC, changes of His$^{230}$ in OpuAC-related proteins occur only by
amino acids that are capable of forming salt bridges or hydrogen bonds (20).

CONCLUSIONS

Site-directed mutagenesis of compatible solute binding proteins ((14, 40)
and this study) has demonstrated that individual amino acids within the aromatic
scaffold make different contributions to ligand binding. Furthermore, the strength
of the cation-π interaction is a key factor for the efficiency of ligand binding (14).
In addition to cation-π interactions, interactions between the carboxylate of the
substrates and the ligand binding protein permit the precise positioning of the
compatible solute within the binding site. As shown in this study, loss of the
interaction between His$^{230}$ of OpuAC and the carboxylate of glycine betaine or
DMSA results in a substantial drop of affinity (Table 1). Thus, the correct
positioning of the ligand within the binding cavity requires molecular interactions involving both the positively charged head group and the negatively charged carboxylic tail of these organic solutes. Therefore, a limited set of molecular interactions is used in various compatible solute-binding proteins to precisely position the ligand within the binding site and a modulation of the interplay between these interactions generates different hierarchies in substrate affinity.
We are indebted to Matthew Groves for his excellent support during data acquisition at beamline BW-7B, EMBL Outstation Hamburg. M. H. is a recipient of a PhD fellowship from the International Max-Planck Research School (MPI Marburg). This work was supported by the Fonds der chemischen Industrie (to E. B.), the Deutsche Forschungsgemeinschaft (SFB 395 to E. B.), the Max-Planck Institut für terrestrische Mikrobiologie (Marburg) (to E.B) and grants of the Heinrich Heine University Düsseldorf (to L. S.).
References


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.


40. Schiefner, A., J. Breed, L. Bosser, S. Kneip, J. Gade, G. Holtmann, K.


**Table legends**

TABLE 1. Using a fluorescence-based assay, the dissociation constants of glycine betaine, proline betaine and DMSA were determined for the wild type protein and each of the generated OpuAC variants. Reported K_D values are the average of at least three independent experiments with the standard deviation reported as error. >> 5 mM: no binding of ligand was detected up to the highest concentration employed in the assay (5 mM). In the case of low affinity binders such as W72F or W178Y the final substrate concentration was 2- to 3fold higher than the K_D value. Abbreviations: A: alanine, F: phenylalanine, Y: tyrosine.

TABLE 2. Crystal parameters and data collection statistics are derived from SCALEPACK (37). Refinement statistics were obtained from REFMAC5 (34) and Ramachandran analysis was performed using MolProbity. Data in parentheses correspond to the highest resolution shell (2.85- 2.80 Å).

\[ R_{sym} = \frac{\sum_{hkl} \sum_{i} I_i(hkl) - \langle I(hkl) \rangle}{\sum_{hkl} \sum_{i} I_i(hkl)} \]  
\[ R_f = \frac{\sum_{hkl} \|F_{obs} - |F_{calc}|\|}{\sum_{hkl} |F_{obs}|} \]  
\[ R_{free} \] is calculated as \( R_f \) but for 5% randomly chosen reflections that were omitted from all refinement steps. All amino acids located in the disallowed region of the Ramachandran plot (0.9% or 9 residues) are involved in crystal contacts.
TABLE 1. Dissociation constants of glycine betaine, proline betaine and DMSA to the wild type OpuAC protein and its mutant derivatives.

<table>
<thead>
<tr>
<th>AMINO ACID</th>
<th>MUTATION</th>
<th>GLYCINE BETAIN</th>
<th>PROLINE BETAIN</th>
<th>DMSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td></td>
<td>22 µM ± 4 µM</td>
<td>267 µM ± 6 µM</td>
<td>102 µM ± 11 µM</td>
</tr>
<tr>
<td>Trp&lt;sup&gt;72&lt;/sup&gt;</td>
<td>A</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>1.4 mM ± 0.4 mM</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>750 µM ± 8 µM</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
</tr>
<tr>
<td>Trp&lt;sup&gt;178&lt;/sup&gt;</td>
<td>A</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>14 µM ± 1.4 µM</td>
<td>243 µM ± 28 µM</td>
<td>56 µM ± 17 µM</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>548 µM ± 263 µM</td>
<td>58 µM ± 27 µM</td>
<td>&gt;&gt; 5 mM</td>
</tr>
<tr>
<td>Trp&lt;sup&gt;225&lt;/sup&gt;</td>
<td>A</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>308 µM ± 18 µM</td>
<td>1.93 mM ± 0.03 mM</td>
<td>&gt;&gt; 5 mM</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>67 µM ± 22 µM</td>
<td>1.25 mM ± 0.28 mM</td>
<td>425 µM ± 30 µM</td>
</tr>
<tr>
<td>His&lt;sup&gt;230&lt;/sup&gt;</td>
<td>A</td>
<td>392 µM ± 92 µM</td>
<td>491 µM ± 195 µM</td>
<td>259 µM ± 55 µM</td>
</tr>
<tr>
<td>Trp&lt;sup&gt;72/178&lt;/sup&gt;</td>
<td>F</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
</tr>
<tr>
<td>Trp&lt;sup&gt;72/225&lt;/sup&gt;</td>
<td>F</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
</tr>
<tr>
<td>Trp&lt;sup&gt;72/178/225&lt;/sup&gt;</td>
<td>F</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
<td>&gt;&gt; 5 mM</td>
</tr>
</tbody>
</table>
TABLE 2. Data collection and refinement statistics for the OpuAC/DMSA complex.

<table>
<thead>
<tr>
<th>Crystal parameters at 100 K</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
</tr>
<tr>
<td>Unit Cell parameters</td>
</tr>
<tr>
<td>a, b, c (Å)</td>
</tr>
<tr>
<td>α, β, γ (deg.)</td>
</tr>
</tbody>
</table>

B. Data collection and processing

| Wavelength (Å)                              | 0.98 |
| Resolution (Å)                              | 20-2.8 (2.85-2.8) |
| Mean redundancy                             | 2.4 |
| Unique reflections                          | 24,818 |
| Mosaicity (°)                               | 0.4 |
| Completeness (%)                            | 93.0 (96.5) |
| I/σ                                         | 6.8 (3.0) |
| R<sub>merge</sub> <sup>a</sup>              | 16.4 (28.8) |

C. Refinement

| R<sub>F</sub><sup>b</sup> (%)               | 28.5 |
| R<sub>free</sub><sup>c</sup> (%)           | 36.4 |
| Overall B-factor from Wilson scaling (Å²)  | 27.4 |
| rmsd from ideal                             |     |
| Bond lengths (Å)                            | 0.07 |
| Bond angles (deg.)                          | 1.12 |
| Average B-factors (Å²)                      | 27.85 |
| Ramachandran plot                           |     |
| Most favored (%)                            | 89.2 |
| Allowed (%)                                 | 9.9 |
| Generously allowed (%)                      |     |
| Disallowed (%)                              | 0.9 |

D. Model content

| Monomers/ASU                                | 4   |
| Protein residues                            | 20-272 |
| Ligand                                     | Four DMSA |
**Figure legends**

FIG. 1. Chemical structures of the OpuAC substrates used in this study.

FIG. 2. Osmoprotective effects of the compatible solutes glycine betaine and DMSA for *B. subtilis*. A. The OpuA⁺ (OpuB⁻ OpuC⁻ OpuD⁻) strain RMKB34 was grown in SMM with 1.2 M NaCl (– ■ –), 1.2 M NaCl with 1 mM glycine betaine (– ● –) and 1.2 M NaCl with 1 mM DMSA (– ▲ –). B. The OpuA⁻ (OpuB⁻ OpuC⁻ OpuD⁻) strain RMKB24 was grown in SMM with 1.2 M NaCl (– ■ –), 1.2 M NaCl with 1 mM glycine betaine (– ▲ –) and 1.2 M NaCl with 1 mM DMSA (– ● –). Cultures (20 ml) were inoculated to an OD₅₇₈ of 0.1 from overnight cultures pre-grown in SMM with 0.4 M NaCl and were propagated in 100-ml Erlenmeyer flasks in a shaking water bath (220 rpm) at 37°C. Cell growth was monitored over time by measuring the OD₅₇₈.

FIG. 3. Ligand binding of OpuAC with DMSA (A, B). (A) Emission spectra of the protein in the absence (red line) or presence (black, dashed line) of 1 mM substrate. (B) Equilibrium binding titration experiments with DMSA.

FIG. 4. View of the ligand-binding pocket of the OpuAC/DMSA complex. Interactions between the OpuAC protein and its ligand DMSA are highlighted by dashed lines. Highlighted are the three tryptophans (Trp⁷₂, Trp¹₇₈, Trp²₂₅) and the histidine residue (His²₃₀), which constitute the binding pocket. Amino acids given
in single letter code in brackets indicate the mutations performed in this study.

FIG. 5. View of the superpositioning of the ligand-binding sites of the OpuAC/glycine betaine, OpuAC/proline betaine and OpuAC/DMSA complexes. Residues involved in glycine betaine coordination are shown in green, residues involved in proline betaine binding are shown in orange and the residues involved in DMSA binding are shown in purple. For simplicity, the backbone contacts of the ligands with Gly$^{26}$ and Ile$^{27}$ have been omitted from the representation.

FIG. 6. Domain organization of glycine betaine binding proteins related to OpuAC from *B. subtilis*. Data base searches using the BLAST program showed that there are four classes of ligand binding proteins that are related to OpuAC from *B. subtilis*. The OpuAC-protein from *B. subtilis* is shown with the residues involved in binding of the trimethylammonium-headgroup of glycine betaine (W72, W178, W225) and the carboxylate of glycine betaine (G26, I27, H230). Group 1 contains those proteins that align directly with the OpuAC-protein. An example is the glycine betaine binding protein GbuC from *Listeria monocytogenes* (29). Group 2 is composed of proteins that align with the OpuAC-protein when the N- and C-terminal domains are inverted. An example is the glycine betaine binding protein OtaC from the archaeon *Methanosarcina mazei* (41). Binding protein domains that are fused to the transmembrane domain of the ABC-transport system and contain the domain inversion form group 3. An example is the glycine betaine binding/transmembrane protein OpuBC (also referred to as BusAB) from
Lactococcus lactis (36, 44). Finally, group 4 of the OpuAC related proteins contain those examples where the transmembrane domain is fused to a duplicated binding protein domain both of which contain the domain inversion. This type of fusion protein was first noticed by van der Heide and Poolman (43).

An example of this group of OpuAC related-proteins is present in Streptomyces coelicolor (NP_625895). But in contrast to the other mentioned glycine betaine binding proteins, the substrate specificity of this fused binding protein has not been experimentally assessed. For the various alignments, the N-terminal and C-terminal domains of OpuAC were split between the amino acids 168/169 as initially described by Horn et al. (20).
Glycine betaine

Proline betaine

DMSA