

Structures of the substrate-binding protein provide insights into the multiple compatible solute binding specificities of the *Bacillus subtilis* ABC transporter OpuC

Yang DU¹, Wei-Wei SHI¹, Yong-Xing HE, Yi-Hu YANG, Cong-Zhao ZHOU² and Yuxing CHEN²

Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230027, People's Republic of China

The compatible solute ABC (ATP-binding cassette) transporters are indispensable for acquiring a variety of compatible solutes under osmotic stress in *Bacillus subtilis*. The substrate-binding protein OpuCC (Opu is osmoprotectant uptake) of the ABC transporter OpuC can recognize a broad spectrum of compatible solutes, compared with its 70% sequence-identical paralogue OpuBC that can solely bind choline. To explore the structural basis of this difference of substrate specificity, we determined crystal structures of OpuCC in the apo-form and in complex with carnitine, glycine betaine, choline and ectoine respectively. OpuCC is composed of two $\alpha/\beta/\alpha$ globular sandwich domains linked by two hinge regions, with a substrate-binding pocket

located at the interdomain cleft. Upon substrate binding, the two domains shift towards each other to trap the substrate. Comparative structural analysis revealed a plastic pocket that fits various compatible solutes, which attributes the multiple-substrate binding property to OpuCC. This plasticity is a gain-of-function via a single-residue mutation of Thr⁹⁴ in OpuCC compared with Asp⁹⁶ in OpuBC.

Key words: *Bacillus subtilis*, compatible solute, crystal structure, fluorescence spectrum, substrate-binding protein.

INTRODUCTION

Water utilization plays a pivotal role in cytoplasmic hydration, growth and survival for cells [1,2]. Exposure of micro-organisms to an environment of high salinity will result in excessive water efflux and destruction of the positive turgor that is generally regarded as the driving force for cell growth and division [3,4]. To counteract the unfavourable water effects, micro-organisms have to actively cope with frequent alteration in the availability of water [5]. The most common strategy against high osmolality is the intracellular amassing of osmotically active compounds, the so-called compatible solutes, through uptake and/or biosynthesis pathways [2]. These compatible solutes function as osmoprotectants that do not interfere with cell physiology, even at a high concentration in the cytoplasm [4,6]. Through fine adjustment of the pool of compatible solutes, micro-organisms maintain a proper intracellular water content to ensure the restoration of turgor and sound growth of cells under osmotic stress [1]. The rigorous restraints of metabolic inertness for such osmoprotectants define a very limited spectrum of compounds, such as GB (glycine betaine), ectoine and trehalose. In addition to osmoprotectants, compatible solutes were identified to possess multifaceted functions, including thermoprotection, and stabilization of proteins and cell components against the denaturing effects of high ionic strength [7–9].

The Gram-positive bacterium *Bacillus subtilis* is widespread in the upper layers of the soil, where frequent variations of water content due to flooding or drought often affects the physiology of cells [5,10]. To adapt to different osmotic stresses, the bacterium

has evolved an orchestrated Opu (osmoprotectant uptake) system to allow acquisition of diverse compatible solutes [2]. The Opu system has five members, including ABC (ATP-binding cassette) transporters (OpuA, OpuB and OpuC) or secondary transporters (OpuD and OpuE) [2,7]. The three ABC transporters OpuA, OpuB and OpuC possess divergent kinetic properties to scavenge the compatible solutes from the environment [11,12]. OpuA exhibits high affinity towards GB ($K_d = 17 \mu\text{M}$) and moderate affinity towards PB (proline betaine; $K_d = 295 \mu\text{M}$), mainly contributing to GB uptake in *B. subtilis* [3]. OpuB and OpuC, which share no sequence similarity with OpuA, are two paralogues due to a gene duplication event. However, the two transporters differ significantly in substrate specificity, affinity and transport capacity [12]. OpuB is a dedicated transporter for choline with high affinity ($K_d = 1 \mu\text{M}$) [12]. In contrast, OpuC exhibits transport activity for multiple compatible solutes, comprising both trimethylammonium derivatives and sulfonium compounds, listed as GB, PB, carnitine, choline, γ -butyrobetaine, choline-O-sulfate and ectoine etc. [12–15]. Remarkably, several compatible solutes, including carnitine and ectoine, are exclusively transported by OpuC, suggesting its indispensable role of osmoregulation in *B. subtilis* [13–15]. To our knowledge, OpuC appears to possess the broadest spectrum of substrate specificity among the characterized compatible solute ABC transporters.

The OpuC system consists of five proteins: two cytoplasmic membrane-associated ATPases, OpuCA, that power the translocation, two membrane-spanning components, OpuCB/OpuCD, that provide a substrate-specific pathway and an extracellular SBP (substrate-binding protein), OpuCC, that determines the substrate

Abbreviations used: ABC, ATP-binding cassette; GB, glycine betaine; Opu, osmoprotectant uptake; PB, proline betaine; RMSD, root mean square deviation; SBP, substrate-binding protein.

¹ These authors contributed equally to this work.

² Correspondence may be addressed to either of these authors (email zcz@ustc.edu.cn or cyxing@ustc.edu.cn).

Co-ordinates and structural factors of the apo-form, and the carnitine-, glycine betaine-, choline- and ectoine-bound forms have been deposited in the PDB under the accession codes 3PPN, 3PPO, 3PPP, 3PPQ and 3PPR respectively.

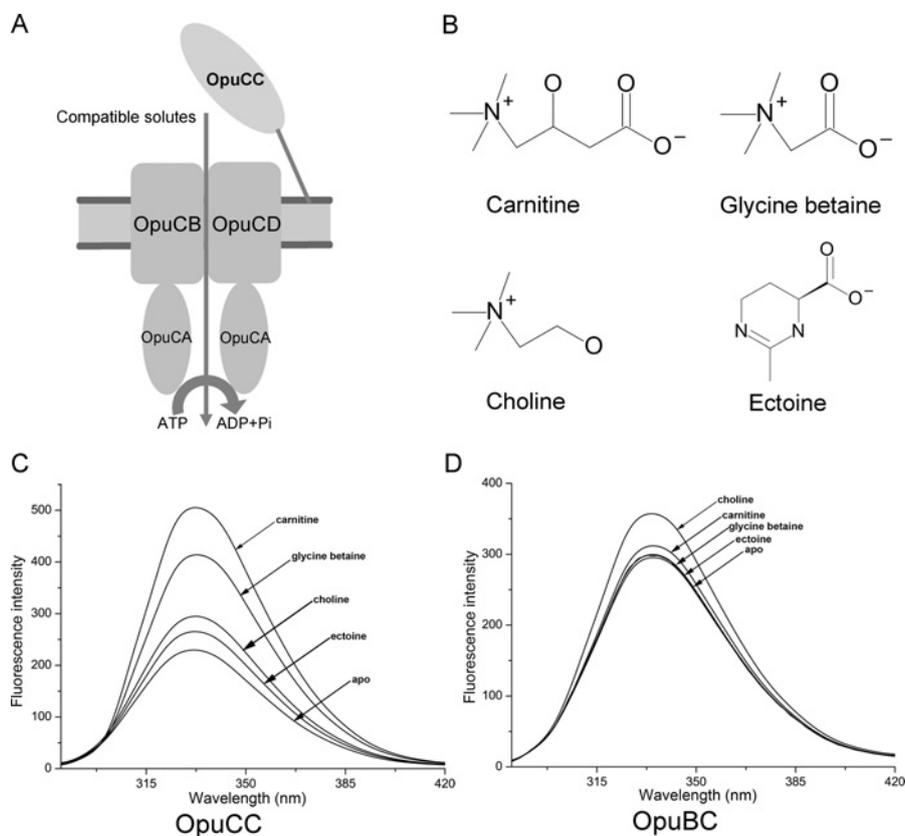


Figure 1 The OpuC transporter and substrate-binding assays

(A) Schematic diagram of the *B. subtilis* OpuC system. The OpuC comprises five components, which are the SBP OpuCC, two membrane-spanning proteins OpuCB/OpuCD and two copies of the cytoplasmic ATPase OpuCA. (B) Molecular structures of the four representative substrates utilized by OpuC. Fluorescence spectra of (C) OpuCC and (D) OpuBC upon substrate binding. The spectra are distinguished by labelling the lines using arrows.

specificity (Figure 1A) [10,12]. OpuCC is a 34 kDa lipoprotein that is tethered to the exterior of the cytoplasmic membrane via a modified N-terminal residue, Cys²¹ [12]. Compared with the SBPs of known structure, OpuCC shares a highest sequence identity of 32% with the GB/PB-binding protein ProX from *Archaeoglobus fulgidus* [16], but is significantly different from others such as *B. subtilis* OpuAC [3] and the choline/acetylcholine-binding protein ChoX of *Sinorhizobium meliloti* [17]. In contrast with the stringent substrate specificity of other SBPs, OpuCC serves as a multiple compatible-solute-binding protein. To understand the structural basis of its multiple-substrate-binding property and the difference from its paralogue OpuBC, we solved four structures of OpuCC in complex with carnitine [2.7 Å (1 Å = 0.1 nm)], GB (2.4 Å), choline (1.9 Å) and ectoine (2.1 Å) respectively. In addition, the 2.3 Å structure of OpuCC in the apo-form enabled us to trace structural rearrangement upon substrate binding. Comparative structural analysis in combination with site-directed mutagenesis revealed that the multiple-substrate-binding property is a gain-of-function due to a single residue mutation.

EXPERIMENTAL

Fluorescence measurements

Protein fluorescence spectra were monitored at wavelengths from 285 to 420 nm using an RF-5301PC spectrofluorophotometer (Shimadzu). Considering the higher content of tyrosine (5%) compared with tryptophan (1%) residues in OpuCC, we used

an excitation wavelength at 275 nm (at a slit-width of 3 nm) to measure the intrinsic fluorescence of tyrosine side chains. A quartz-cuvette containing 300 µl of protein sample (1 µM) was incubated for 10 min at 25 °C before measurement. Carnitine (L-carnitine), GB, choline and ectoine at a final concentration of 500 µM were added to the protein sample respectively. The apo-form OpuCC was used as the control.

Cloning, expression and purification of OpuCC

The coding sequence of OpuCC from the *B. subtilis* genome was cloned into a pET28a-derived vector. A His₆ (hexa-histidine) tag was added to the N-terminus of each recombinant protein before overexpression in the *Escherichia coli* BL21-RIL (DE3) strain (Novagen) in 2×YT culture medium (16 g of tryptone, 10 g of yeast extract and 5 g of NaCl per litre). Cells were grown at 37 °C to an *A*₆₀₀ of 0.6. Expression of recombinant proteins was induced with 0.2 mM IPTG (isopropyl β-D-thiogalactopyranoside) and cell growth was continued for 20 h at 16 °C before harvesting. Cells were collected by centrifugation at 4000 g for 20 min, and resuspended in lysis buffer [20 mM Tris/HCl (pH 8.0) and 150 mM NaCl]. After 5 min of sonication (six cycles, 30 s per cycle, 1:1 pulse/pause at 50% amplitude of 750 W) and centrifugation at 12 000 g for 25 min, the supernatant containing the target protein was collected and loaded on to a Ni-NTA (Ni²⁺-nitrilotriacetate) column (GE Healthcare) equilibrated with binding buffer [20 mM Tris/HCl (pH 8.0) and 150 mM NaCl]. Target protein was eluted with 200 mM imidazole, and further loaded on to a Superdex 75 column (GE Healthcare) equilibrated

Table 1 Crystal parameters, data collection and refinement statistics

Numbers in parentheses are for the outer shell.

Measurements	Apo-form	Substrate-bound forms			
		Carnitine	GB	Choline	Ectoine
Data collection					
Resolution range (Å)	50.00–2.30 (2.34–2.30)	38.41–2.70 (2.85–2.70)	50.00–2.40 (2.44–2.40)	38.22–1.91 (2.02–1.91)	50.00–2.10 (2.18–2.10)
Space group	$P2_12_12_1$	$P22_12_1$	$P22_12_1$	$P22_12_1$	$P22_12_1$
Unit-cell parameters (Å)	a = 89.28 b = 89.67 c = 90.39 $\alpha = \beta = \gamma = 90^\circ$	a = 61.89 b = 91.07 c = 115.23 $\alpha = \beta = \gamma = 90^\circ$	a = 61.82 b = 91.78 c = 115.33 $\alpha = \beta = \gamma = 90^\circ$	a = 61.73 b = 90.15 c = 115.66 $\alpha = \beta = \gamma = 90^\circ$	a = 61.96 b = 92.58 c = 115.81 $\alpha = \beta = \gamma = 90^\circ$
Wavelength (Å)	0.9194	1.5418	1.5418	0.9701	0.9998
Total reflections	241156	51741	122284	338317	262849
Unique reflections	33035	17555	24956	49630	38654
Redundancy	7.3 (7.4)	2.9 (2.7)	4.9 (3.8)	6.8 (6.4)	6.8 (5.7)
Completeness (%) (overall/outer shell)	99.9/100.0	95.2/85.5	94.3/83.0	98.5/96.2	98.2/97.3
$\langle I/\sigma(I) \rangle$	19.0 (6.6)	9.8 (2.8)	11.1 (3.4)	12.2 (5.0)	19.4 (4.6)
R_{merge}^* (%)	9.2 (30.5)	8.5 (44.3)	9.6 (34.4)	9.2 (24.1)	8.6 (35.6)
Refinement statistics					
$R_{\text{factor}}^\dagger$ (%)	21.6	21.2	21.6	22.5	19.8
R_{free}^\ddagger (%)	23.9	23.7	25.3	24.2	23.1
Number of atoms	4399	4496	4493	4468	4490
RMSD from targets§					
Bond lengths (Å)	0.016	0.014	0.019	0.014	0.011
Bond angles (°)	1.406	1.383	1.638	1.343	1.137
Average B factors	25.8	10.4	19.1	12.6	18.0
Ramachandran plot (%)	98.51/1.49/0.00	99.26/0.74/0.00	99.26/0.74/0.00	99.26/0.74/0.00	99.07/0.93/0.00

* $R_{\text{merge}} = \frac{\sum_h \sum_i |I_{hi} - \langle I_h \rangle|}{\sum_h \sum_i I_{hi}}$, where I_{hi} is the i th observation of the reflection h , whereas I_h is the mean intensity of reflection h .

† $R_{\text{factor}} = \sum \frac{|F_{\text{obs}} - F_{\text{calc}}|}{F_{\text{obs}}}$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes respectively.

‡ R_{free} was calculated with 5% randomly selected reflections. No refinement was done on the 5% of randomly selected reflections at any stage.

§ RMSD of bond lengths or bond angles from ideal geometry.

|| Percentage of residues in most favoured/generously allowed/disallowed regions of Ramachandran plot, according to MOLPROBITY.

with 20 mM Tris/HCl (pH 8.0) and 150 mM NaCl. Fractions containing the target protein were collected and concentrated to 20 mg/ml. Protein purity was evaluated by electrophoresis.

Crystallization, data collection and processing

Crystals of the apo- and substrate-bound forms were grown at 289 K with hanging-drop vapour-diffusion techniques by mixing 1 μ l of the 100 mg/ml protein sample with an equal volume of the reservoir solutions. All of the four substrate-bound forms (5 mM of each substrate added) were crystallized in the conditions of 28% (v/v) PEG [poly(ethylene glycol)] 400, 0.2 M CaCl₂ and 0.1 M sodium Hepes (pH 7.5). Apo-form crystals were obtained in the drop containing 32% (v/v) Jeffamine M-600 and 0.1 M Hepes (pH 7.0). The diffraction images of the apo- and substrate-bound forms were recorded at 100 K in a liquid-nitrogen stream using the beamline at the Shanghai Synchrotron Radiation Facility (for apo-, choline- and ectoine-bound forms) or in-house Rigaku MM007 X-ray collection systems at the University of Science and Technology of China (for carnitine- and GB-bound forms). The datasets were integrated and scaled with the program HKL2000 (for the apo-, GB- and ectoine-bound forms) or MOSFLM 7.0.4 (for the carnitine- and choline-bound form) [18], and scaled with SCALA [19].

Structure solution and refinement

The structures in substrate-bound forms were determined by molecular replacement with MOLREP [20] using the co-ordinates

of GB-bound ProX from *A. fulgidus* which shares 32% sequence identity over 270 residues (PDB code 1SW2). The structure of the apo OpuCC has substantial conformational change compared with that of the substrate-bound form. Thus domain II (residues 136–240) of the GB-bound OpuCC co-ordinate was used as the search model to determine the phase of the apo-form. The initial models were further refined by using the maximum likelihood method implemented in REFMAC5 [21] as part of CCP4 program suite [22], and rebuilt interactively by using the σ_A -weighted electron-density maps with coefficients $2mF_o - DF_c$ and $mF_o - DF_c$ in the program COOT [23]. TLS groups were analysed at the TLSMD server (<http://skuld.bmsc.washington.edu/~tlsmd/>) and the optimum number of groups was applied during the refinement [24]. The final models were evaluated with the programs MOLPROBITY [25] and PROCHECK [26]. The data collection and structure refinement statistics are listed in Table 1. All structure figures were prepared with PyMOL (<http://www.pymol.org>).

RESULTS AND DISCUSSION

Multiple-substrate-binding specificities of OpuCC

OpuCC from *B. subtilis* can transport a variety of compatible solutes [1,2]. In the present study, four osmoprotectants were applied to the binding assays and structural studies. These substrates (Figure 1B) are critical to the osmoregulation of *B. subtilis*. Carnitine (γ -trimethyl- β -hydroxy-butyrobetaine) is a widespread osmoprotectant and a carrier of acyl moieties [28]. GB, as a fully N-methyl-substituted derivative of

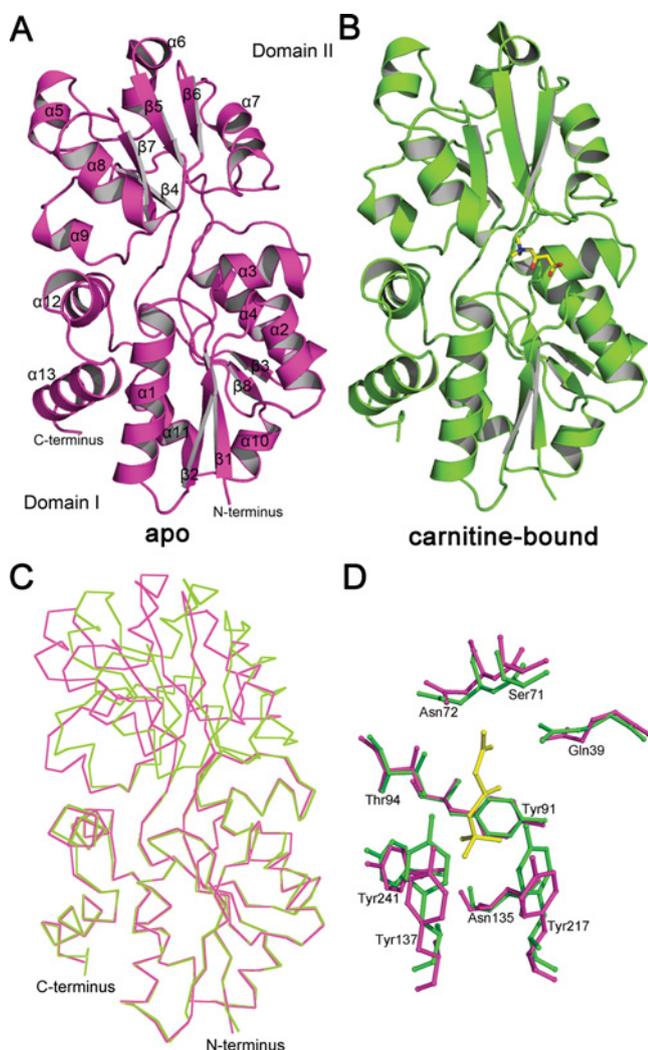


Figure 2 Overall structures and structural comparison of the apo- and substrate-bound forms

Overall structures of OpuCC in (A) apo- and (B) carnitine-bound forms. Secondary structural elements are labelled sequentially in the apo-form. Superimposition of the apo-form (magenta) against the carnitine-bound form (green) with (C) the overall structure and (D) the carnitine-binding site. The crystal of the carnitine-bound form was obtained by co-crystallizing OpuCC with 5 mM carnitine.

glycine, is the most popular osmoprotectant and an effective protectant against abiotic stress [29]. As the precursor of GB, choline is indispensable for *B. subtilis* since the bacterium is not capable of synthesizing GB *de novo* [12]. Ectoine (1,4,5,6-tetrahydro-2-methyl-4-pyrimidinecarboxylic acid) is a natural compatible solute differing from the quaternary ammonium compatible solute family, and has been found in several species of bacteria, especially halophilic micro-organisms [30]. As reported previously, the equilibrium dissociation constants (K_d) of OpuCC towards the four substrates are 5 μ M (carnitine) [13], 6 μ M (GB) [10], 28 μ M (choline) [12] and 1.5 mM (ectoine) [15] respectively. In the present study we used fluorescence enhancement assays to compare the substrate-binding specificities of OpuCC and OpuBC. The addition of each substrate to the apo-form OpuCC led to an increase in fluorescence intensity at a fixed maximal emission wavelength (332 nm) (Figure 1C). In contrast, only addition of choline, but

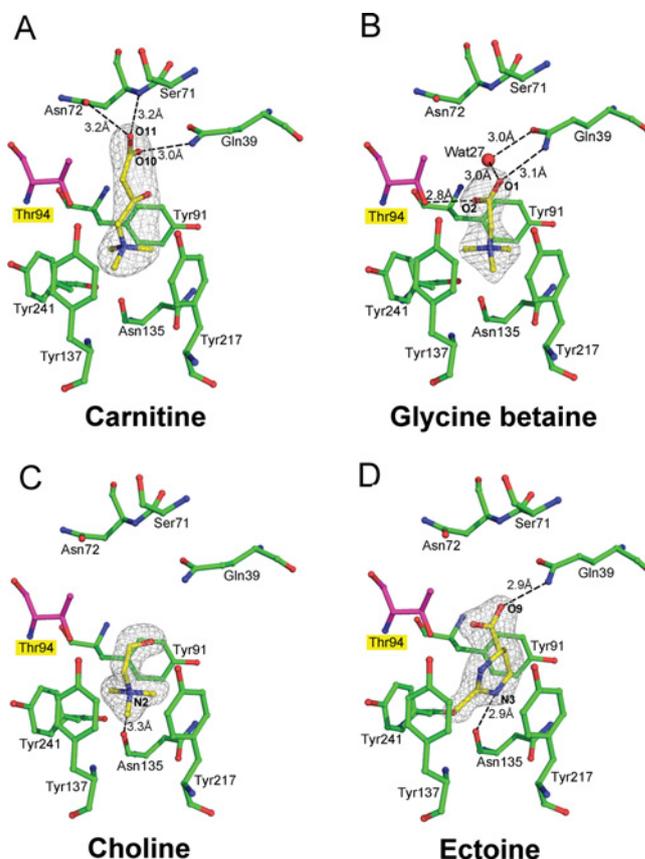


Figure 3 Binding patterns of various substrates

Binding sites of (A) carnitine, (B) GB, (C) choline and (D) ectoine. All of the four substrates are coloured in yellow with $2mF_o - DF_c$ density map contoured at 1.0 σ . Thr⁹⁴ is coloured in magenta. Atoms of the four substrates contributing to hydrogen bonds are labelled according to PDB files.

not carnitine, GB or ectoine, to the apo-form OpuBC triggered a significant increase of fluorescence intensity (Figure 1D).

Overall structure of OpuCC and the induced fit upon substrate binding

To elucidate the structural basis for the multiple-substrate-binding property and the different substrate specificity from OpuBC, we solved the structures of OpuCC in the apo-form (2.3 Å) and substrate-bound forms, in complex with carnitine (2.7 Å), GB (2.4 Å), choline (1.9 Å) and ectoine (2.1 Å) respectively. They belong to $P2_12_12_1$ (apo) or $P22_12_1$ (substrate-bound) space group, with two molecules in an asymmetric unit. The overall structure exhibits a typical topology in an F-III cluster for bacterial SBPs [31]. OpuCC is an elliptically bilobal protein with approximate dimensions of $60 \times 40 \times 30$ Å³. The protein is built of two $\alpha/\beta/\alpha$ globular sandwich domains. Domain I comprises residues 1–135 and 241–303, which is formed by a central four-stranded β -sheet ($\beta_1, 2, 3, 8$) flanked by eight α -helices (α_1 –4, α_{10} –13). Domain II (residues 136–240) is made up of a central four-stranded β -sheet (β_4 –7) surrounded by five α -helices (α_5 –9). The two domains are linked by two hinge regions (residues 127–137 and 226–246) (Figure 2A). Comparison of the overall structures of the apo-form and the carnitine-bound form yielded an RMSD (root mean square deviation) of 1.6 Å over 267 residues. The flexible hinge regions between the two domains enable the apo-form to adopt a relatively

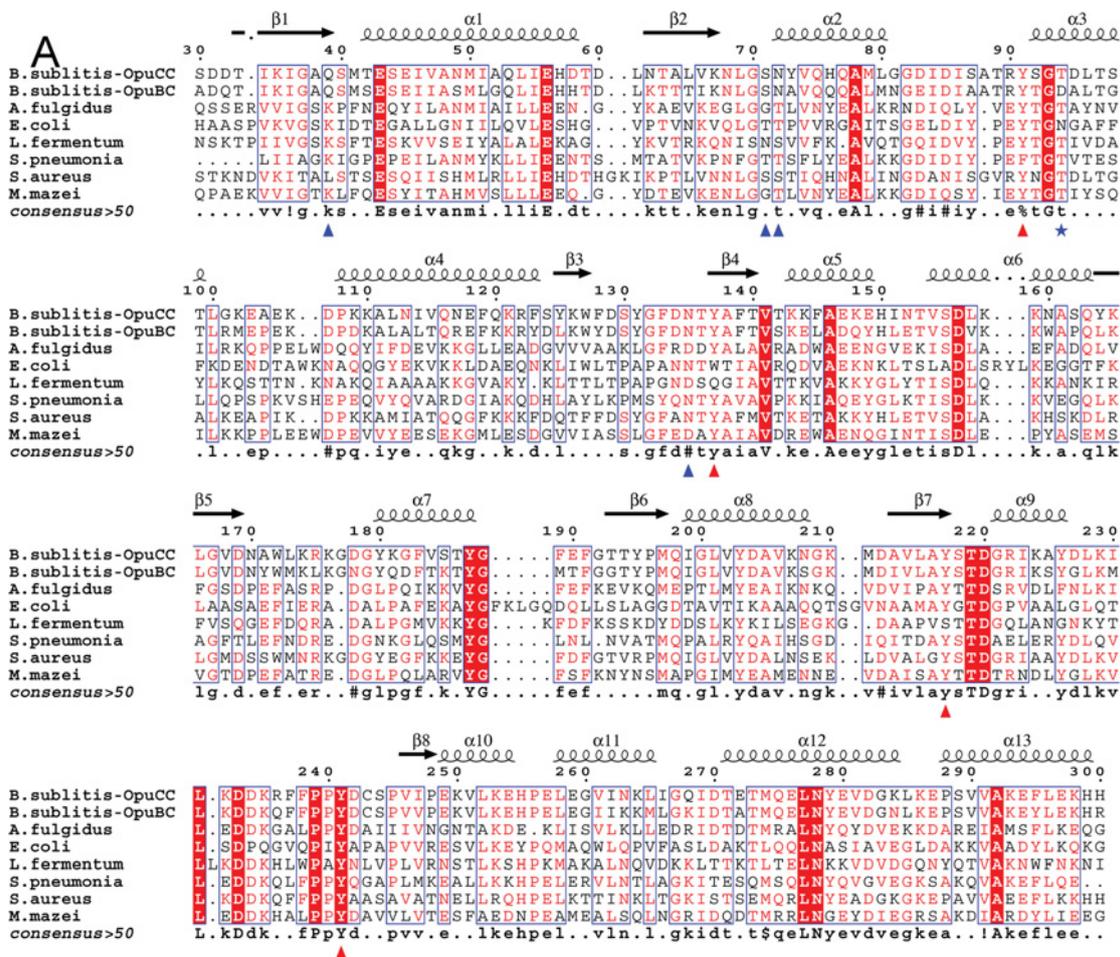


Figure 4 Multiple sequence alignment

(A) Sequence alignment of OpuCC (GenBank[®] accession number AF009352) against the homologues from the archaea *A. fulgidus* (GenBank[®] accession number AAB90259, 32% sequence identity with OpuCC), *Methanosarcina mazei* (GenBank[®] accession number AAM29994, 32% sequence identity with OpuCC), the Gram-negative bacterium *Escherichia coli* (GenBank[®] accession number CAR17002, 27% sequence identity with OpuCC), probiotic *Lactobacillus fermentum* (GenBank[®] accession number FJ025800, 29% sequence identity with OpuCC), the pathogens *Streptococcus*

loose conformation. Upon substrate binding, obvious rigid body movements make the two domains shift towards each other and trap the substrate at the interdomain cleft (Figures 2B and 2C).

Structural plasticity of the substrate-binding pocket of OpuCC

Similar to the structures of other SBPs, the substrate-binding pocket of OpuCC is located at the cleft between domains I and II (Figure 2B). The pocket has a side wall resembling an 'aromatic girdle' which is composed of four tyrosine residues, Tyr⁹¹, Tyr¹³⁷, Tyr²¹⁷ and Tyr²⁴¹, with their phenyl rings almost perpendicular to each other. The residue Asn¹³⁵ at the bottom and residues Gln³⁹, Ser⁷¹, Asn⁷² and Thr⁹⁴ at the top, together with the 'aromatic girdle', make up a cage to accommodate the substrate (Figure 2D). This pocket is quite different from that of OpuAC, which is composed of three tryptophan residues in a prism-like geometry to fit the positively charged trimethylammonium moiety through cation- π interactions [3].

The residues described above contribute differently to bind various substrates. In the structure of carnitine-, GB- and choline-bound forms, the three substrates share a common head group of trimethylammonium to interplay with the four tyrosine residue side wall of the pocket mediated by cation- π and hydrophobic interactions. However, different combinations of residues participate in binding to substrates of different polar moieties and lengths of the main chain. The pocket is almost fully occupied by the long carbon chain of carnitine, with the atoms O-10 and O-11 of the carboxy group extended to Gln³⁹-N ϵ 2, Ser⁷¹-N α and Asn⁷²-O δ 1 to form three hydrogen bonds (Figure 3A). In the GB-bound form, the carbon chain of GB is too short to make hydrogen bonds with Ser⁷¹ and Asn⁷² (Figure 3B). Alternatively, O-1 and O-2 in the polar moiety of GB form three hydrogen bonds with Thr⁹⁴-O γ 1 and Gln³⁹-N ϵ 2, and a water molecule Wat²⁷ respectively. Both of the polar moieties of GB and carnitine are fixed by three hydrogen bonds, which determine their comparable binding affinity to OpuCC. In the choline-bound form, the short carbon chain of choline makes the polar moiety incapable of making direct contact with surrounding residues, resulting in its much lower binding affinity compared with that of carnitine and GB (Figure 3C). The absence of hydrogen bonds to the polar moiety of choline makes its trimethyl group shift approximately 0.4 Å towards Asn¹³⁵, and form a hydrogen bond between N-2 and Asn¹³⁵-O α . As a special compatible solute in structure from the above, ectoine possesses a pyrimidine ring. In addition to a hydrogen bond between Gln³⁹-N ϵ 2 and O-9 which stabilizes the polar moiety, a hydrogen bond between Asn¹³⁵-O and N-3, and the π - π interaction to Tyr²¹⁷ fix the pyrimidine ring (Figure 3D). The much-reduced interactions between ectoine and the pocket lead to the lower affinity of OpuCC towards ectoine.

Sequence analysis indicates that the residues interacting with the trimethyl group are highly conserved, whereas the residues fixing the polar group are variable (Figure 4A). Although superimposition of the pocket of the apo-form to that of the substrate-bound form only yields an RMSD of approximately 0.7 Å over nine C α atoms of the substrate-binding residues, obvious conformational changes of the pocket upon substrate binding are observed, mainly for the conserved aromatic residues

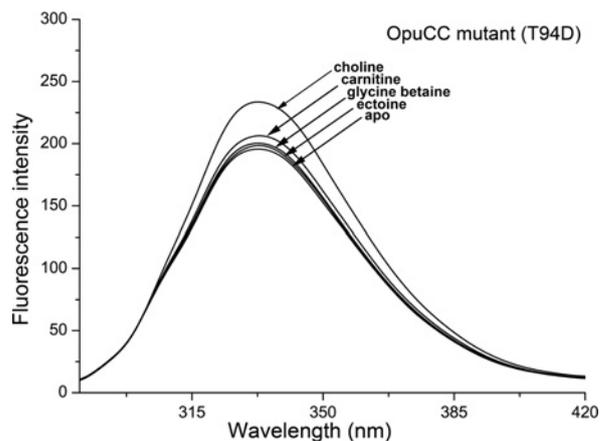


Figure 5 Fluorescence spectra of the mutant OpuCC^{T94D}

The spectra are shown in a 'bottom to the top' order for the apo-, ectoine-, GB-, carnitine- and choline-bound forms.

(Figure 2D). As a result of induced fit, the side chains of residues Gln³⁹, Ser⁷¹ and Asn⁷² are shifted towards the substrate. Meanwhile, the side chains of Tyr¹³⁷ and Tyr²¹⁷ are flapped towards the substrate to form a tighter 'tyrosine girdle' for stabilizing the trimethyl moiety via cation- π and hydrophobic interactions. Once Tyr⁹¹ was mutated to phenylalanine, tryptophan or alanine, the increase of fluorescence intensity upon addition of carnitine was more or less reduced (Figure 4B). In contrast with the slight decrease in the binding affinity for the Y91F and Y91W mutants, the Y91A mutant completely lost the binding activity. These results suggested the indispensability of integrity of the 'tyrosine girdle'.

The multiple-substrate-binding property of OpuCC is a gain-of-function resulting from a single-residue mutation

Since OpuCC and OpuBC in *B. subtilis* were duplicated from a primordial protein, how did OpuCC evolve the binding property towards multiple compatible solutes? Sequence alignment revealed that they shared almost all of the same substrate-binding residues, except for Thr⁹⁴ in OpuCC compared with Asp⁹⁶ in OpuBC (Figure 4A). The hydroxy side chain of Thr⁹⁴ of OpuCC is close to the polar moiety of various substrates, and even forms a hydrogen bond with GB (Figure 3). It is compatible with either the carboxy or hydroxy group of the polar moieties. Replacement of Thr⁹⁴ with Asp⁹⁶ will introduce a negatively charged carboxy group at this site, which will cause electrostatic repulsion against the adjacent negatively charged carboxy group of carnitine, GB or ectoine (Figures 3A, 3B and 3D). This probably accounts for the incapability of OpuBC to bind these three substrates. In contrast, the hydroxy group of choline is not excluded by the carboxy group of Asp⁹⁶ (Figure 3C). To validate the role of Thr⁹⁴, we constructed the mutant OpuCC^{T94D} and applied it to fluorescence spectra assays. Upon addition of various substrates, OpuCC^{T94D} shares a quite similar pattern of fluorescence spectrum to that of OpuBC (Figure 5 compared with 1D). Only choline can

pneumoniae (GenBank® accession number AAL00480, 38% sequence identity with OpuCC) and *Staphylococcus aureus* (GenBank® accession number EES94881, 57% sequence identity with OpuCC), and the paralogue OpuBC (GenBank® accession number AAC14358, 70% sequence identity with OpuCC) from *Bacillus subtilis*. The alignment was performed using Multalin [32] and ESPript [33] (<http://multalin.toulouse.inra.fr/multalin/>). The residues interacting with trimethyl and polar groups are marked by red and blue triangles respectively. Thr⁹⁴ in OpuCC is indicated by a blue star. (B) The fluorescence spectra of the wild-type OpuCC and mutants Y91F, Y91W and Y91A. The black and red lines represent the spectra of the apo-form and the carnitine-bound form respectively.

trigger obvious changes of fluorescence intensity of OpuCC^{T94D}, whereas carnitine, GB and ectoine cannot. Therefore the multiple-substrate-binding property of OpuCC can be attributed to this single residue mutation, which might be a successful evolutionary selection of *B. subtilis* to uptake more types of compatible solute from diverse milieu.

Conclusions

In the present study, the issues concerning the structure and function of the compatible-solute-binding protein OpuCC were addressed. Crystal structures of OpuCC in complex with different substrates revealed a plastic pocket that fits multiple compatible solutes, which provide structural insights into its multiple-substrate-binding property. This plasticity is a gain-of-function via a single residue mutation of Thr⁹⁴ in OpuCC compared with Asp⁹⁶ in OpuBC, which is responsible for the distinct substrate specificity between OpuCC and its paralogue OpuBC. Taken together, these investigations on the multiple-compatible-solute specificities of OpuCC illustrate a principle of how a plastic binding pocket for multiple ligands is achieved at the protein structural level.

AUTHOR CONTRIBUTION

Yang Du and Wei-Wei Shi carried out the molecular cloning, site-directed mutagenesis, protein purification, crystallization and biochemical analysis. Yang Du and Yong-Xing He performed data collection and structure determination. Yi-Hu Yang contributed to biochemical analysis. Yang Du, Yuxing Chen and Cong-Zhao Zhou wrote the paper. Yuxing Chen and Cong-Zhao Zhou supervised the project.

ACKNOWLEDGEMENT

We are grateful to the staff at the Shanghai Synchrotron Radiation Facility for the help with data collection.

FUNDING

This work was supported by the Ministry of Science and Technology of China [grant number 2009CB918804]; and the National Natural Science Foundation of China [grant number 30870488].

REFERENCES

- Wood, J. M., Bremer, E., Csonka, L. N., Kraemer, R., Poolman, B., Van Der Heide, T. and Smith, L. T. (2001) Osmosensing and osmoregulatory compatible solute accumulation by bacteria. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **130**, 437–460
- Kempf, B. and Bremer, E. (1998) Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Arch. Microbiol.* **170**, 319–330
- Horn, C., Sohn-Bosser, L., Breed, J., Welte, W., Schmitt, L. and Bremer, E. (2006) Molecular determinants for substrate specificity of the ligand-binding protein OpuAC from *Bacillus subtilis* for the compatible solutes glycine betaine and proline betaine. *J. Mol. Biol.* **357**, 592–606
- Wood, J. M. (1999) Osmosensing by bacteria: signals and membrane-based sensors. *Microbiol. Mol. Biol. Rev.* **63**, 230–262
- Hoffmann, T., Boiangiu, C., Moses, S. and Bremer, E. (2008) Responses of *Bacillus subtilis* to hypotonic challenges: physiological contributions of mechanosensitive channels to cellular survival. *Appl. Environ. Microbiol.* **74**, 2454–2460
- Roesser, M. and Muller, V. (2001) Osmoadaptation in bacteria and archaea: common principles and differences. *Environ. Microbiol.* **3**, 743–754
- Holtmann, G. and 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
- Arakawa, T. and Timasheff, S. N. (1985) The stabilization of proteins by osmolytes. *Biophys. J.* **47**, 411–414
- Caldas, T., Demont-Caulet, N., Ghazi, A. and Richarme, G. (1999) Thermoprotection by glycine betaine and choline. *Microbiology* **145**, 2543–2548
- Kappes, R. M., Kempf, B. and Bremer, E. (1996) Three transport systems for the osmoprotectant glycine betaine operate in *Bacillus subtilis*: characterization of OpuD. *J. Bacteriol.* **178**, 5071–5079
- Smits, S. H., Hoing, M., Lecher, J., Jebbar, M., Schmitt, L. and Bremer, E. (2008) The compatible-solute-binding protein OpuAC from *Bacillus subtilis*: ligand binding, site-directed mutagenesis, and crystallographic studies. *J. Bacteriol.* **190**, 5663–5671
- Kappes, R. M., Kempf, B., Kneip, S., Boch, J., Gade, J., Meier-Wagner, J. and Bremer, E. (1999) Two evolutionarily closely related ABC transporters mediate the uptake of choline for synthesis of the osmoprotectant glycine betaine in *Bacillus subtilis*. *Mol. Microbiol.* **32**, 203–216
- Bremer, R. and Kappes, M. E. (1998) Response of *Bacillus subtilis* to high osmolarity: uptake of carnitine, crotonobetaine and -butyrobetaine via the ABC transport system OpuC. *Microbiology* **144**, 83–90
- Nau-Wagner, G., Boch, J., Le Good, J. A. and Bremer, E. (1999) High-affinity transport of choline-O-sulfate and its use as a compatible solute in *Bacillus subtilis*. *Appl. Environ. Microbiol.* **65**, 560–568
- Jebbar, M., von Blohn, C. and Bremer, E. (1997) Ectoine functions as an osmoprotectant in *Bacillus subtilis* and is accumulated via the ABC-transport system OpuC. *FEMS Microbiol. Lett.* **154**, 325–330
- Schiefner, A., Holtmann, G., Diederichs, K., Welte, W. and Bremer, E. (2004) Structural basis for the binding of compatible solutes by ProX from the hyperthermophilic archaeon *Archaeoglobus fulgidus*. *J. Biol. Chem.* **279**, 48270–48281
- Oswald, C., Smits, S. H., Hoing, M., Sohn-Bosser, L., Dupont, L., Le Rudulier, D., Schmitt, L. and Bremer, E. (2008) Crystal structures of the choline/acetylcholine substrate-binding protein ChoX from *Sinorhizobium meliloti* in the liganded and unliganded-closed states. *J. Biol. Chem.* **283**, 32848–32859
- Powell, H. R. (1999) The Rossmann Fourier autoindexing algorithm in MOSFLM. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55**, 1690–1695
- Evans, P. R. (1993) Data reduction. In *Proceedings of CCP4 Study Weekend: Data Collection and Processing*, pp. 114–122, Daresbury Laboratory, Warrington
- Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.* **30**, 1022–1025
- Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53**, 240–255
- Collaborative Computational Project No. 4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760–763
- Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132
- Painter, J. and Merritt, E. A. (2006) TLSMD web server for the generation of multi-group TLS models. *J. Appl. Crystallogr.* **39**, 109–111
- Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, III, W. B., Snoeyink, J., Richardson, J. S. and Richardson, D. C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291
- Reference deleted
- Peluso, G., Barbarisi, A., Savica, V., Reda, E., Nicolai, R., Benatti, P. and Calvani, M. (2000) Carnitine: an osmolyte that plays a metabolic role. *J. Cell. Biochem.* **80**, 1–10
- Chen, T. H. and Murata, N. (2008) Glycinebetaine: an effective protectant against abiotic stress in plants. *Trends Plant Sci.* **13**, 499–505
- Graf, R., Anzali, S., Buenger, J., Pfluecker, F. and Driller, H. (2008) The multifunctional role of ectoine as a natural cell protectant. *Clin. Dermatol.* **26**, 326–333
- Berntsson, R. P., Smits, S. H., Schmitt, L., Slotboom, D. J. and Poolman, B. (2010) A structural classification of substrate-binding proteins. *FEBS Lett.* **584**, 2606–2617
- Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**, 10881–10890
- Gouet, P., Robert, X. and Courcelle, E. (2003) ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* **31**, 3320–3323