Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallographic analysis of the C-terminal domain of the *Escherichia coli* lipoprotein BamC

Kelly H. Kim, Suraaj Aulakh, Wendy Tan and Mark Paetzel*

Molecular Biology and Biochemistry, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada

Correspondence e-mail: mpaetzel@sfu.ca

Received 15 July 2011
Accepted 18 August 2011

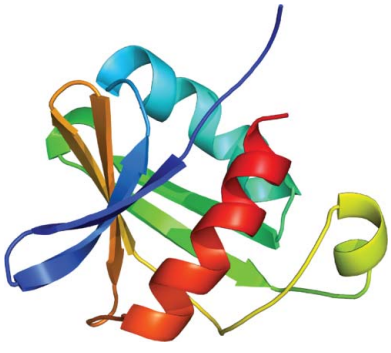
PDB Reference: C-terminal domain of BamC, 3sns.

In Gram-negative bacteria, the BAM complex catalyzes the essential process of assembling outer membrane proteins. The BAM complex in *Escherichia coli* consists of five proteins: one β-barrel membrane protein, BamA, and four lipoproteins, BamB, BamC, BamD and BamE. Here, the crystal structure of the C-terminal domain of *E. coli* BamC ($BamC_C$: Ala224–Ser343) refined to 1.5 Å resolution in space group *H*3 is reported. $BamC_C$ consists of a six-stranded antiparallel β-sheet, three α-helices and one 3_{10}-helix. Sequence and surface analysis reveals that most of the conserved residues within $BamC_C$ are localized to form a continuous negatively charged groove that is involved in a major crystalline lattice contact in which a helix from a neighbouring $BamC_C$ binds against this surface. This interaction is topologically and architecturally similar to those seen in the substrate-binding grooves of other proteins with BamC-like folds. Taken together, these results suggest that an identified surface on the C-terminal domain of BamC may serve as an important protein-binding surface for interaction with other BAM-complex components or substrates.

1. Introduction

The outer membrane (OM) of Gram-negative bacteria contains β-barrel outer membrane proteins (OMPs) that fulfil a number of roles including nutrient uptake, antibiotic resistance, cell adhesion and maintenance of membrane selective permeability (Silhavy *et al.*, 2010; Delcour, 2009; Bos *et al.*, 2007). Proper folding and insertion of these OMPs into the OM is essential for the membrane's structure and function, and hence for cell viability (Silhavy *et al.*, 2010; Paschen *et al.*, 2005; Schleiff & Soll, 2005; von Heijne, 2011). The OMPs are synthesized in the cytosol with an N-terminal signal sequence that directs them to the inner membrane (IM; von Heijne, 2011; Volokhina *et al.*, 2009). The protein is then transported across the IM *via* the Sec translocation system and the signal sequence is subsequently removed by signal peptidase (Bos *et al.*, 2007; Gatsos *et al.*, 2008). The processed OMPs are released into the periplasm, guided to the OM by chaperones such as SurA, Skp and DegP, and finally folded and inserted into the OM (Bos *et al.*, 2007; Gatsos *et al.*, 2008; Sklar, Wu, Kahne *et al.*, 2007). In *Escherichia coli*, this final folding and insertion process is catalyzed by the BAM (β-barrel assembly machinery) complex (Fitzpatrick & McInerney, 2005; Gentle *et al.*, 2004; Hagan *et al.*, 2011; Voulhoux *et al.*, 2003; Voulhoux & Tommassen, 2004). Systems homologous to the BAM complex can also be found in the mitochondria and chloroplasts of eukaryotes (Paschen *et al.*, 2005; Gentle *et al.*, 2004).

The *E. coli* BAM complex consists of five proteins: BamA, BamB, BamC, BamD and BamE (Hagan *et al.*, 2011). BamA is a β-barrel integral membrane protein (*i.e.*, it is an OMP) with a long periplasmic tail consisting of five polypeptide transport-associated (POTRA) motifs (Gentle *et al.*, 2005; Misra, 2007). BamB–E are lipoproteins that are anchored to the periplasmic face of the OM *via* an amino-terminally attached lipid (Silhavy *et al.*, 2010; Bos *et al.*, 2007). While BamA and BamD are essential for OMP biogenesis and cell viability, the absence of BamB, BamC or BamE results in decreased efficiency of OMP folding and assembly (Gentle *et al.*, 2004; Charlson *et al.*, 2006; Malinverni *et al.*, 2006; Sklar, Wu, Gronenberg *et al.*, 2007). Mutagenesis and pull-down studies have shown that BamB and



BamD independently and directly interact with the POTRA motifs of BamA (Malinverni *et al.*, 2006; Kim *et al.*, 2007). While BamE has been shown to stabilize the interaction between BamA and BamD (Sklar, Wu, Gronenberg *et al.*, 2007), BamC cannot be co-immunoprecipitated with BamA unless BamD is present (Malinverni *et al.*, 2006). It has also been shown that the C-terminal region of BamD is necessary for BamCD interaction (Malinverni *et al.*, 2006). However, the oligomeric structure and stoichiometry of the BAM complex is not yet clearly understood.

The detailed molecular mechanism of how the BamA–E proteins work together as a complex is being actively investigated from both structural and functional perspectives. Structural information on the individual BAM proteins has started to emerge in recent years (Kim *et al.*, 2007; Albrecht & Zeth, 2010; Clantin *et al.*, 2007; Gatzeva-Topalova *et al.*, 2008, 2010; Heuck *et al.*, 2011; Kim *et al.*, 2011; Kim & Paetzel, 2011; Knowles *et al.*, 2008, 2011; Noinaj *et al.*, 2011; Sandoval *et al.*, 2011; Vanini *et al.*, 2008; Warner *et al.*, 2011). Here, we present the structural analysis of the C-terminal domain of *E. coli* BamC. It should be noted that structures of both the N-terminal and the C-terminal domains of BamC ($BamC_N$ and $BamC_C$, respectively) were reported while we were in the process of determining our structure (Warner *et al.*, 2011; Albrecht & Zeth, 2011). Analysis of $BamC_C$ and its crystal contacts in space group *H*3, together with comparisons with previously reported BamC structures (space group $P2_1$ and the NMR solution structure) and structural homologues, provides insights into potential $BamC_C$ interactions.

2. Materials and methods

2.1. Cloning and mutagenesis

For a full-length construct, a 957-base-pair DNA fragment coding for residues Ser26–Lys344 of *E. coli* BamC was amplified from *E. coli* K-12 genomic DNA using the forward primer 5′-TAT ACC ATG GCC AGT TCT GAC TCA CGC-3′ and the reverse primer 5′-TAT AAA GCT TTT ATT ACT TGC TAA ACG CAG-3′, which contain *Nde*I and *Xho*I restriction sites, respectively. The PCR product was then ligated into vector pET28a (Novagen). The resulting construct contains a cleavable N-terminal hexahistidine affinity tag (with an SSGLVPRGSHM linker between the tag and Ser26). Subsequent DNA sequencing confirmed that the insert matched the sequence reported in the Swiss-Prot database (accession No. P0A903). A surface-entropy reduction mutant (BamC_SERM_E182A_Q183A_K186A) was created to promote crystallization. The site-directed mutagenesis reaction was performed using the Phusion kit (Finnzymes) using the primers 5′-GCG TAA CCT GGC AGC GGC GGG CGC ACC GGT TGC AGA CGC GG-3′ and 5′-CCG CGT CTG CAA CCG GTG CGC CCG CCG CTG CCA GGT TCA GC-3′.

2.2. Expression and purification of BamC

The expression plasmid coding for the BamC_SERM protein was transformed into *E. coli* BL21 (λDE3) and used to inoculate (1:100 back dilution) 2 l Luria Bertani medium containing kanamycin (50 µg ml^{-1}). Cultures were grown at 310 K until the $OD_{600\,nm}$ reached 0.6. The culture was then induced with 1 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) for 3 h. Cells were harvested by centrifugation (6000*g* for 10 min) and subsequently lysed using an Avestin Emulsiflex-3C cell homogenizer in buffer *A* (20 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl). The resulting lysate was clarified by centrifugation (45 000*g*) for 30 min at 277 K and the expressed protein was initially purified by Ni^{2+} affinity chromatography. The protein was eluted from the nickel–nitrilotriacetic acid column (Qiagen) with a step-gradient method (100–500 m*M* imidazole in buffer *A* in 100 m*M* increments). The fractions containing BamC were pooled and concentrated to approximately 10 mg ml^{-1} using an Amicon ultracentrifugal filter device with a 10 kDa molecular-weight cutoff (Millipore) and then further purified by size-exclusion chromatography (Sephacryl S-100 HiPrep 26/60 column; flow rate of 0.5 ml min^{-1}) in buffer *A* on an ÄKTA Prime system (GE Healthcare).

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Crystal parameters	
Space group	*H*3
Unit-cell parameters (Å)	$a = b = 78.9$, $c = 52.9$
Data-collection statistics	
Wavelength (Å)	0.98058
Resolution (Å)	28.7–1.5 (1.6–1.5)
Total reflections	110632 (16027)
Unique reflections	19576 (2877)
R_{merge}†	0.075 (0.378)
Mean $I/\sigma(I)$	10.3 (3.2)
Completeness (%)	99.7 (100.0)
Multiplicity	5.7 (5.6)
Refinement statistics	
Protein chains in asymmetric unit	1
Residues	120
Water molecules	99
Total No. of atoms	1029
R_{cryst}‡/R_{free}§ (%)	16.1/18.3
Average *B* factor (all atoms) (Å^2)	28.0
R.m.s.d. on angles (°)	0.016
R.m.s.d. on bonds (Å)	1.60

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an individual reflection and $\langle I(hkl)\rangle$ is the mean intensity of that reflection. ‡ $R_{cryst} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively. § R_{free} is calculated using 5% of the reflections randomly excluded from refinement.

2.3. Limited proteolysis

BamC (200 µl at 1 mg ml^{-1}) was digested with chymotrypsin (1000:1 BamC:chymotrypsin ratio by mass) overnight at room temperature. Aliquots of 10 µl were taken at time points t = 0, 5, 10, 15, 30, 45, 60 and 120 min and overnight. The aliquots were mixed with SDS–PAGE loading dye, boiled for 3 min to stop the reaction and then run on SDS–PAGE for analysis.

2.4. Crystallization

Crystals were grown by the hanging-drop vapour-diffusion method. The crystallization drops were prepared by mixing 1 µl protein solution (30 mg ml^{-1}) in buffer *A* with 1 µl reservoir solution and the drops were equilibrated against 1 ml reservoir solution. The proteolytically resistant fragment of the BamC_SERM protein yielded crystals belonging to space group *H*3, with unit-cell parameters a = 78.9, b = 78.9, c = 52.9 Å. The crystals contained one molecule in the asymmetric unit with a Matthews coefficient of 2.4 Å^3 Da^{-1} (49.4% solvent). The optimal crystallization reservoir condition was 0.1 *M* NaCl, 0.1 *M* HEPES pH 6.5 and 1.6 *M* ammonium sulfate. Crystallization was performed at room temperature (∼295 K). The cryosolution consisted of 0.1 *M* NaCl, 0.1 *M* HEPES pH 6.5, 1.6 *M* ammonium sulfate and 30% glycerol. Crystals were washed in the cryosolution before being flash-cooled in liquid nitrogen.

2.5. Data collection

Diffraction data were collected on beamline 08ID-1 at the Canadian Macromolecular Crystallography Facility (CMCF) of the

Canadian Light Source (CLS) using a MAR Mosaic Rayonix MX300 CCD X-ray detector. The X-ray wavelength used was 0.98058 Å. The crystal-to-detector distance was 180 mm. A total of 180 images were collected with 0.35° oscillation and each image was exposed for 0.5 s. The diffraction data were processed with the programs *iMOSFLM* (Battye *et al.*, 2011), *POINTLESS* (Evans, 2006) and *SCALA* (Evans, 2006); see Table 1 for data-collection statistics.

2.6. Structure determination and refinement

Although the full-length BamC_SERM construct (Ser26–Lys344) was used for crystallization, the protein was cleaved to a smaller fragment in the crystallization drop during incubation. The determined unit-cell parameters and symmetry of the crystals are inconsistent with the full-length BamC_SERM construct fitting into the crystal lattice based on its molecular mass. Molecular-replacement trials with the N-terminal domain of BamC (PDB entry 2yh6; Albrecht & Zeth, 2011) as a search model failed. However, phases were obtained by molecular replacement using the program *Phaser* v.2.1 (McCoy *et al.*, 2007) when the recently reported C-terminal domain structure of *E. coli* BamC (PDB entry 2yh5; Albrecht & Zeth, 2011) was used as a search model. Clear electron density can be seen for residues Ala224–Lys343, which correspond to $BamC_C$. The structure was refined using restrained refinement in *REFMAC*5 (Murshudov *et al.*, 2011), and further manual adjustments to the atomic coordinates were performed with the program *Coot* (Emsley & Cowtan, 2004). The final model was obtained by running restrained refinement in *REFMAC*5 with TLS restraints obtained from the *TLSMD* server (Painter & Merritt, 2006). The refinement statistics are shown in Table 1. The atomic coordinates and structure factors have been deposited in the RCSB Protein Data Bank (PDB entry 3sns).

2.7. Structural analysis

Secondary-structural analysis was performed with the program *DSSP* (Kabsch & Sander, 1983). The program *Coot* (Emsley & Cowtan, 2004) was used to overlap coordinates for structural comparison. The stereochemistry of the structure was analyzed with the program *PROCHECK* (Laskowski *et al.*, 1993) and the *DALI* server (Holm *et al.*, 2008) was used to find proteins with similar protein folds. The surface electrostatics analysis was performed with the adaptive Boltzmann–Poisson solver plug-in using dielectric constants of 2 and 80 for the solute and solvent, respectively. The online programs *CASTp* (Liang *et al.*, 1998) and *PISA* (Krissinel & Henrick, 2007) were used to detect potential protein-binding pockets and measure solvent-exposed molecular surface and cavities and to identify protein–protein interaction interfaces, respectively.

3. Results

3.1. Purification, limited proteolysis and crystallization

Our limited proteolytic analysis of purified full-length $BamC_{26–344}$ revealed a proteolytically resistant fragment of approximately 25 kDa in size (Fig. 1). This cleavage pattern is consistent with secondary-structure predictions and recently reported NMR data, which suggest that the first 75 residues of the protein are unstructured (Warner *et al.*, 2011; Knowles *et al.*, 2009) while the rest of the protein folds into two separate but structurally related domains ($BamC_N$ and $BamC_C$; Fig. 1; Albrecht & Zeth, 2010, 2011). Interestingly, when full-length BamC was screened for crystallization the protein appeared to experience proteolytic cleavage in the crystallization drop. Crystals of $BamC_C$ (Ala224–Ser343) formed in ammonium sulfate at pH 6.5 in space group *H*3. These crystals diffracted to beyond 1.5 Å resolution.

3.2. Overall protein fold and molecular-surface properties

$BamC_C$ consists of a six-stranded twisted antiparallel β-sheet, three α-helices and one 3_{10}-helix (Figs. 2*a* and 2*b*). Two α-helices (α1 and α3) pack against one side of the β-sheet *via* hydrophobic interactions and a shorter helix α2 is positioned adjacent to the other two helices away from the β-sheet. A short 3_{10}-helix, η1, is part of a loop region between β-strands β3 and β4. The twisting of the β-sheet creates a concave surface on the solvent-accessible side of the sheet, pre-

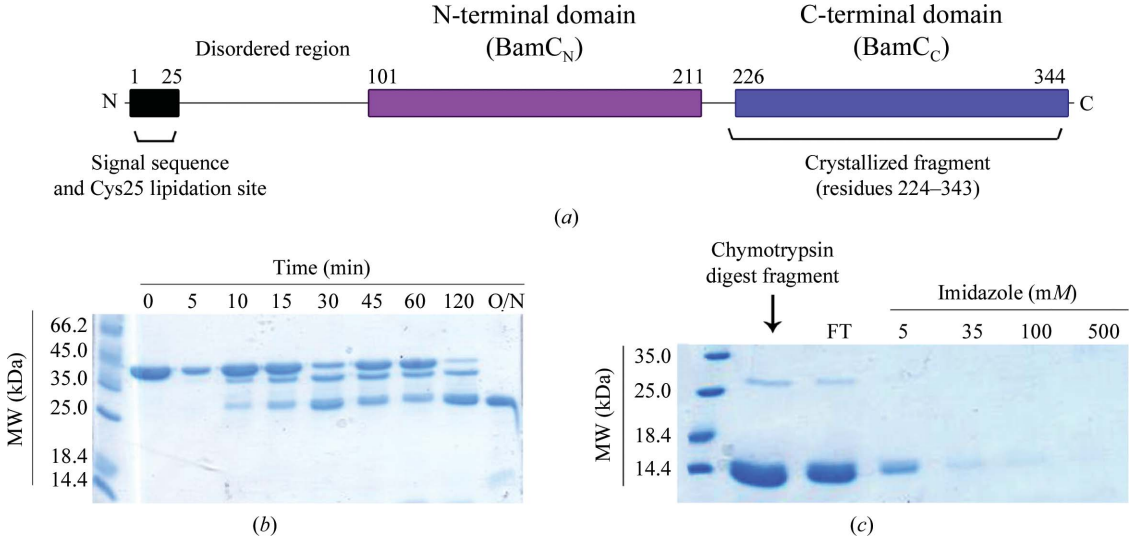


Figure 1
Structural regions of BamC as determined by limited proteolysis. (*a*) A schematic diagram of the BamC domains. (*b*) Following purification, the full-length BamC (35 kDa) was subjected to limited proteolysis with chymotrypsin. A 25 kDa fragment started to appear on the gel after 10 min of proteolytic reaction. After overnight incubation, it can be seen that all of the full-length BamC has been completely cleaved by the protease and that only the 25 kDa protease-resistant fragment remained. (*c*) The 25 kDa BamC fragment was subsequently run on an Ni^{2+}-affinity column to determine whether the N-terminal His_6 tag was lost owing to proteolysis. At this point, most of the 25 kDa species can be seen to have cleaved even further to a smaller ~12 kDa species. Both the remaining 25 and 12 kDa protein fragments of BamC did not bind to the column, suggesting that proteolysis had occurred at the N-terminus of the protein.

dominantly formed by $\beta2$, $\beta5$ and $\beta6$. This groove has a surface area of approximately 485 Å^2 and is lined with negatively charged residues. Analysis of the overall electrostatic properties of BamC$_C$ reveals a predominantly negatively charged surface (Fig. 2*c*) consistent with its theoretical isoelectric point of 4.7.

3.3. Conserved residues

Sequence comparisons of *E. coli* BamC with functional homologues indicates that although there are conserved residues throughout the protein, the majority of the conserved blocks of sequence reside in the unstructured N-terminal region (Ser26–Thr100) and within BamC$_C$ (Fig. 3*a*). When the conserved residues within BamC$_C$ are mapped onto our structure, all of the invariant residues are found to be solvent exposed; many of them (Arg245, Asp309, Asn312, Arg313 and Ser315) are located on $\beta5$ and $\beta6$ and in close proximity to each other such that they form a continuous surface (Fig. 3*b*). This conserved patch on the BamC$_C$ surface forms part of the negatively charged concave surface mentioned above.

3.4. Comparison of BamC$_C$ crystal (space groups *H*3 and *P*2$_1$) and solution structures

Our crystal structure of BamC$_C$ in space group *H*3 (pH 6.5, ammonium sulfate) superimposes with a previously determined BamC$_C$ crystal structure in space group *P*2$_1$ (pH 6.5, PEG 1000) with an overall main-chain r.m.s.d. of 0.7 Å (PDB entry 2yh5; Albrecht & Zeth, 2011) and with the lowest energy NMR solution structure of BamC$_C$ with an overall main-chain r.m.s.d. of 0.9 Å (PDB entry 2laf; Fig. 4*a*; Warner *et al.*, 2011). Although there is a slight variation in the positions of the $\alpha1$ and $\alpha2$ helices, the three structures show a very close resemblance overall. The structure of the β-sheet is especially well conserved between the compared coordinates (Figs. 4*b* and 4*c*). Comparing the *B*-factor distributions of the two crystal structures shows a well ordered β-sheet and flexible N- and C-termini in both structures (Fig. 4*d*). Unlike in the BamC$_C$ structure in space group *P*2$_1$, the structure in space group *H*3 suggests that $\alpha2$ and the residues immediately following $\alpha2$ are the regions of highest flexibility.

The crystal-packing protein–protein interaction surfaces observed in the *H*3 structure are significantly different from those seen in the

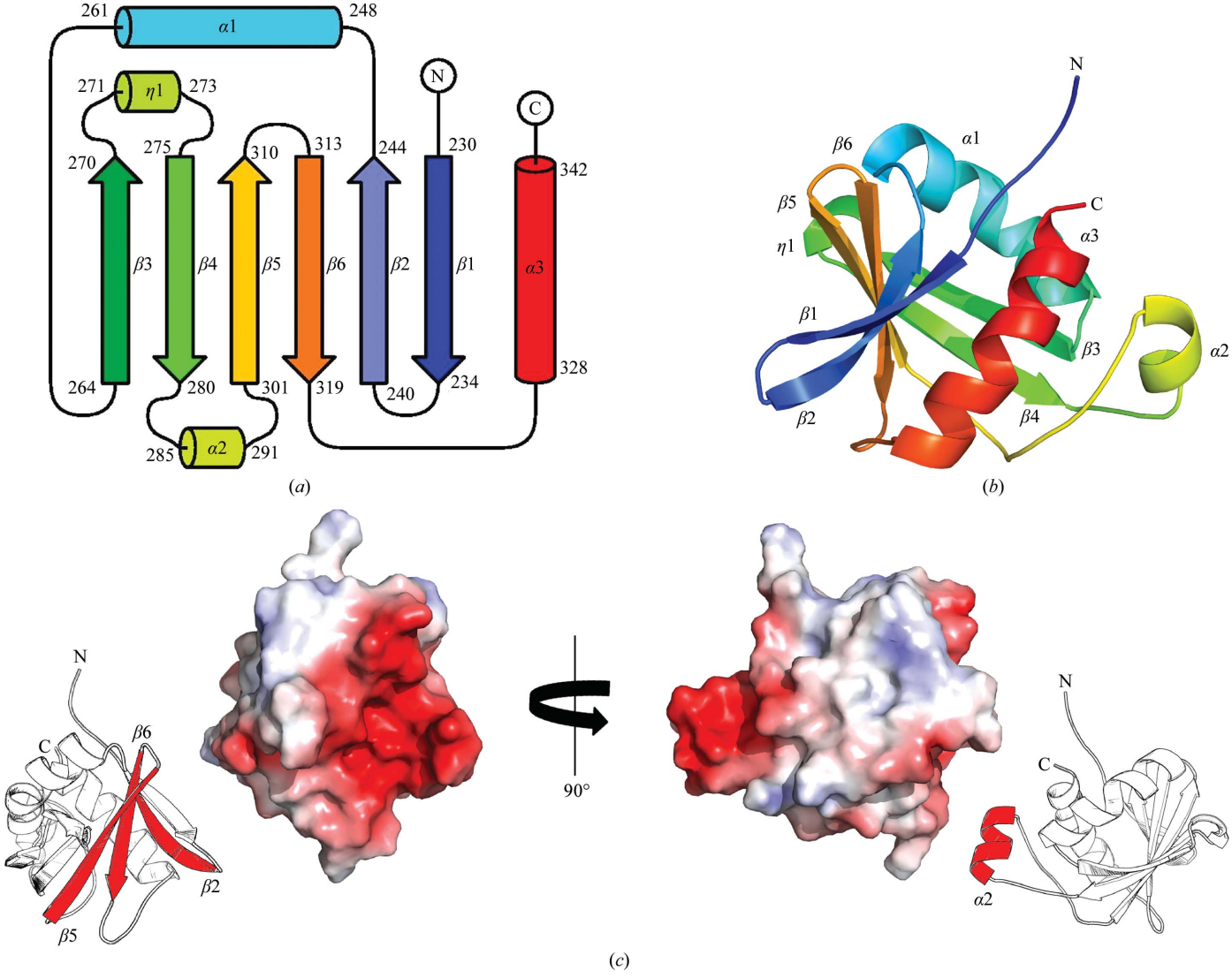


Figure 2
BamC$_C$ fold and surface features. (*a*) A topology diagram of BamC$_C$ is shown with strands represented as arrows, helices as cylinders and loops as lines. (*b*) A ribbon diagram of BamC$_C$; the colour changes progressively from the N-terminus (blue) to the C-terminus (red). (*c*) The electrostatic potential is mapped onto the surface of BamC$_C$. Red and white represent negative and hydrophobic potentials, respectively.

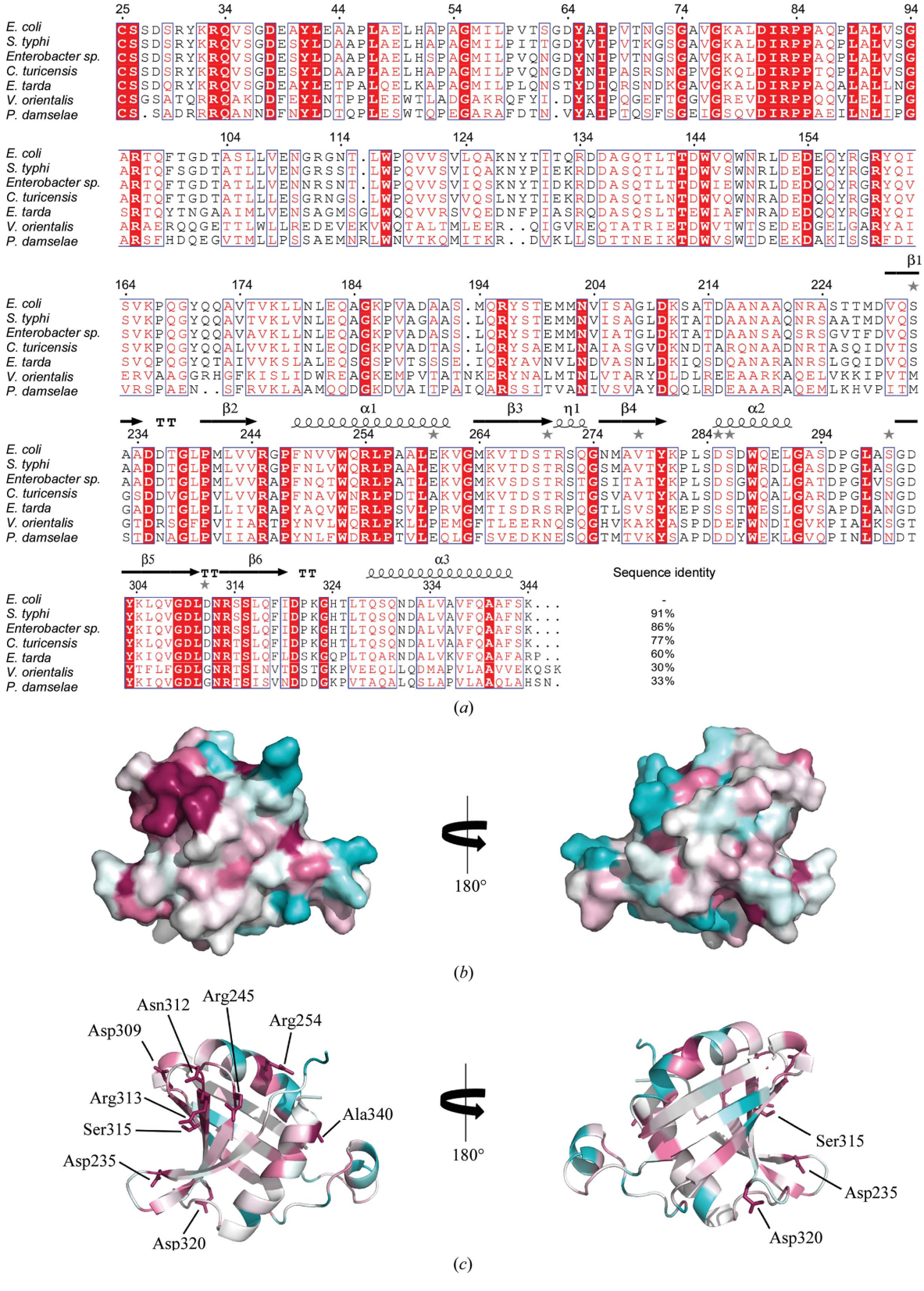


Figure 3
BamC sequence conservation. (*a*) A multiple sequence alignment of *E. coli* BamC with homologous proteins is shown. All sequences start at the cysteine residue just after the signal peptide. The secondary structure of *E. coli* $BamC_C$ is shown above the alignment. Absolutely conserved residues are shown in red boxes, similar residues in red text and stretches of residues that are similar across the group of sequences in blue boxes. The protein sequences were acquired from the Swiss-Prot database: *E. coli*, P0A903; *Salmonella typhi*, Q8Z4R9; *Enterobacter* sp., A4WD57; *Cronobacter turicensis*, C9XXL0; *Edwardsiella tarda*, E0T3S3; *Vibrio orientalis*, C9QLI3; *Photobacterium damselae*, D0YWF8. (*b*) A view of $BamC_C$ conservation mapped onto the *E. coli* $BamC_C$ surface generated using the above alignment. Absolutely conserved residues are shown in maroon and highly variable residues are shown in cyan. (*c*) A ribbon diagram showing conserved residues in stick representation.

$P2_1$ structure (Fig. 5). In the $P2_1$ structure the largest crystal contact has an interface area of 503 Å^2, while the largest interface in our $H3$ structure has an area of 875 Å^2. A detailed comparison of the protein–protein interaction surfaces (crystal contacts) observed in the $P2_1$ and $H3$ BamC$_C$ crystals is summarized in Table 2.

3.5. Potential protein–protein interaction site

Each BamC$_C$ molecule in the crystalline $H3$ lattice makes three major crystal contacts (Fig. 5*b*). The largest interface (875 Å^2) is formed by helix α1 lying within the negatively charged groove of the neighbouring molecule (predominantly involving β5 and β6 of the β-sheet) in the crystalline lattice (Figs. 6*a* and 6*b*). The interaction is mostly mediated *via* hydrogen bonds and electrostatic interactions (Table 2 and Fig. 6*a*, inset). The molecular surface involved in this interaction is the previously mentioned negatively charged cavity in which many of the conserved residues are located.

To examine how other proteins with BamC$_C$-like folds interact with their binding partners, we performed a search for structural homologues using the *DALI* server (Holm *et al.*, 2008) and identified several proteins that have similar protein topology and architecture to BamC$_C$. One such protein is the *Homo sapiens* AP2 complex

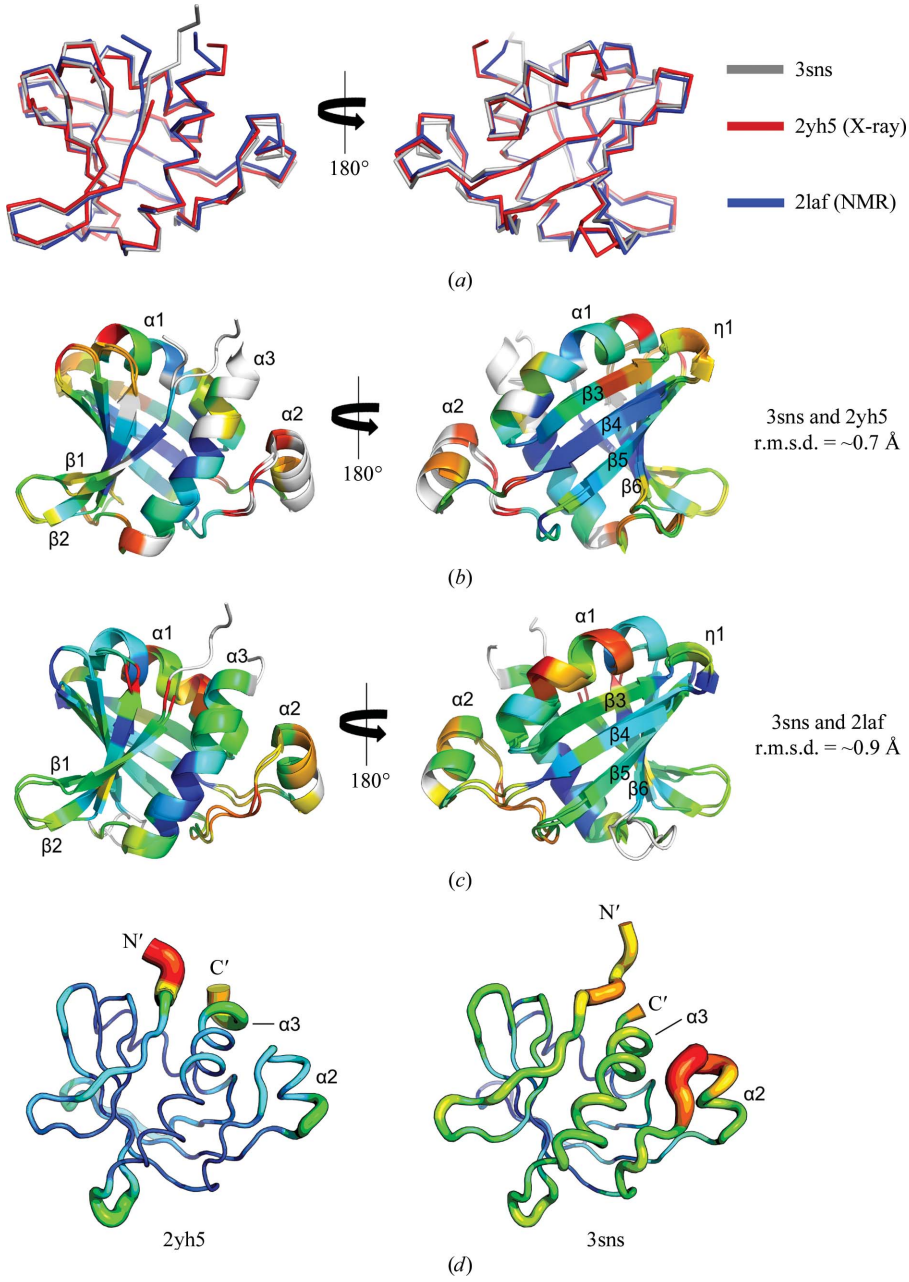


Figure 4
A comparison of BamC$_C$ crystal and solution structures. (*a*) The BamC$_C$ crystal structure (space group $H3$; PDB entry 3sns; grey) is shown superimposed with the crystal structure in space group $P2_1$ (PDB entry 2yh5; red) and the NMR solution structure (PDB entry 2laf; blue). (*b*) The structures in crystallographic space groups $H3$ (3sns) and $P2_1$ (2yh5) are aligned and coloured by r.m.s.d. Dark blue represents close alignment and orange/yellow/red represents increasing levels of deviation. Residues not used for alignment are coloured white. (*c*) The crystal structure in space group $H3$ (3sns) and the NMR solution structure (2laf) are aligned and coloured by r.m.s.d. as in (*b*). (*d*) 'Sausage' representation of the main-chain *B*-factor distribution of the BamC$_C$ crystal structure in space groups $H3$ (3sns) and and $P2_1$ (2yh5) is shown. Higher *B*-factor values are indicated by thicker lines and warmer colours.

β-appendage domain, which has previously been cocrystallized with a peptide that mimics a region of β-arrestin, the substrate (Schmid *et al.*, 2006; PDB entry 2iv8). A close examination of the cocrystal structure shows that the binding region of the substrate, which is a short α-helical peptide, binds to the concave surface formed by the twisted β-sheet of the AP2 complex β-appendage domain (Fig. 6*c*); this is reminiscent of the major crystal-packing interaction observed in the *H*3 $BamC_C$ crystal. When the structures of the AP2 complex β-appendage domain and $BamC_C$ are superimposed (r.m.s.d. of 3.0 Å), it can be seen that the binding pockets of the two proteins align closely (Fig. 6*d*).

Table 2
Crystal contacts.

Interfaces are numbered as in Fig. 5.

Space group	Interface	Interface area (Å^2)	Interfacing residues	Hydrogen bonds	Salt bridges
$P2_1$ (PDB entry 2yh5)	1	503	31	8	8
	2	278	21	5	0
	3	250	20	2	0
*H*3 (PDB entry 3sns)	4	875	53	12	4
	5	310	21	4	0
	6	115	12	2	2

4. Discussion

In *E. coli*, the BAM complex plays an essential role in OMP folding and assembly. The multi-component complex is made up of the integral membrane protein BamA and the four lipoproteins BamB, BamC, BamD and BamE (Hagan *et al.*, 2011; Anwari *et al.*, 2010). Mature BamC is tethered to the periplasmic face of the outer membrane *via* an *N*-acyl diacyl glyceryl group which is attached to the thiol group of Cys25. BamC contains three separate structural regions: an unstructured region (residues 26–100) and two copies of a globular domain ('helix-grip fold'; Iyer *et al.*, 2001) referred to as $BamC_N$ (residues 101–211) and $BamC_C$ (residues 226–344). The sequence identity of $BamC_C$ and $BamC_N$ is only 12%, yet they superimpose with an r.m.s.d. of 2.7 Å (Albrecht & Zeth, 2011). Here, we present the crystal structure of $BamC_C$ in space group *H*3 and compare it with a previously solved crystal structure in space group $P2_1$ (Albrecht & Zeth, 2011) as well as an NMR solution structure (Warner *et al.*, 2011).

Analysis of the conserved residues, molecular-surface properties, crystal-packing interfaces and comparisons with proteins with a similar fold all point to the same region of $BamC_C$ as a potential protein-binding site. This negatively charged groove formed by the solvent-exposed surface of the $BamC_C$ β-sheet has a shape that can

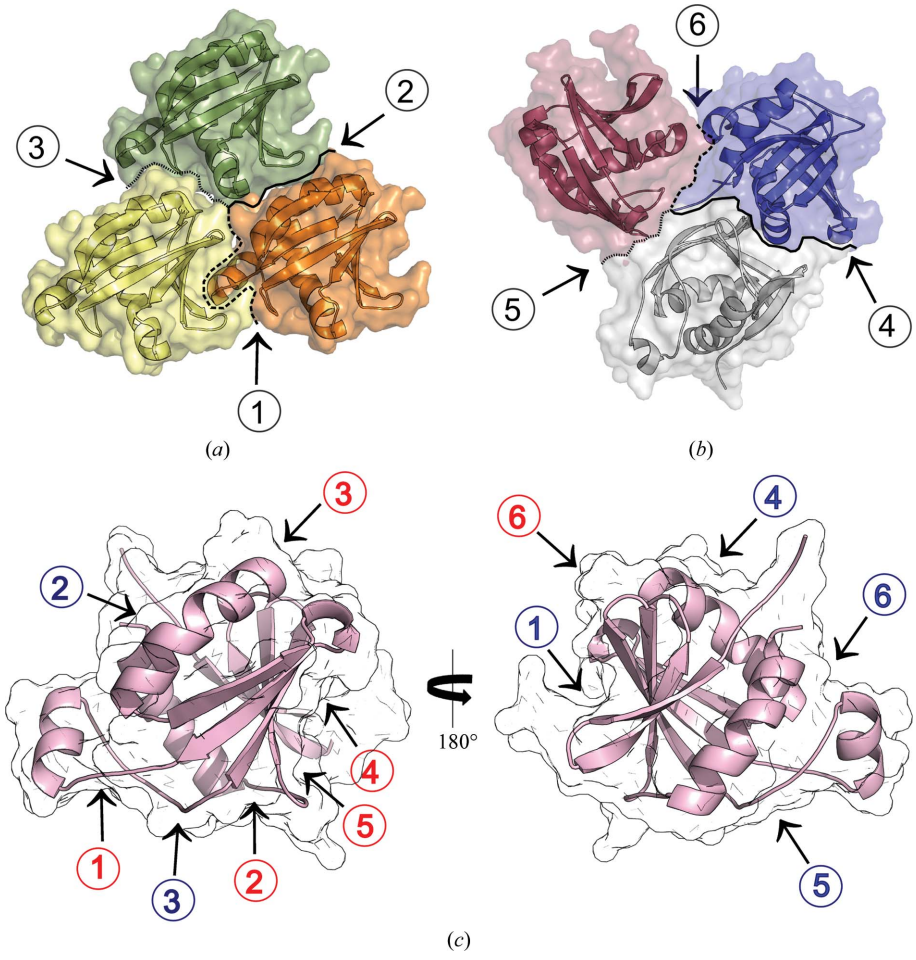


Figure 5
Crystal contacts. (*a*) The three largest crystal contacts observed in the $P2_1$ $BamC_C$ crystal structure (PDB entry 2yh5) are shown. The interfaces are numbered 1, 2 and 3 in order of decreasing interface area. (*b*) The three largest crystal contacts observed in the *H*3 $BamC_C$ crystal structure (PDB entry 3sns) are shown. The interfaces are numbered 4, 5 and 6 in order of decreasing interface area. (*c*) The $BamC_C$ structure is shown in two different views. Crystal contact regions observed in both the *H*3 and $P2_1$ crystal forms of $BamC_C$ are indicated with the same numbering as in (*a*) and (*b*). For each crystal contact, one interface surface is labelled in red and its partner interface surface is labelled in blue. Crystal contacts are summarized in Table 2.

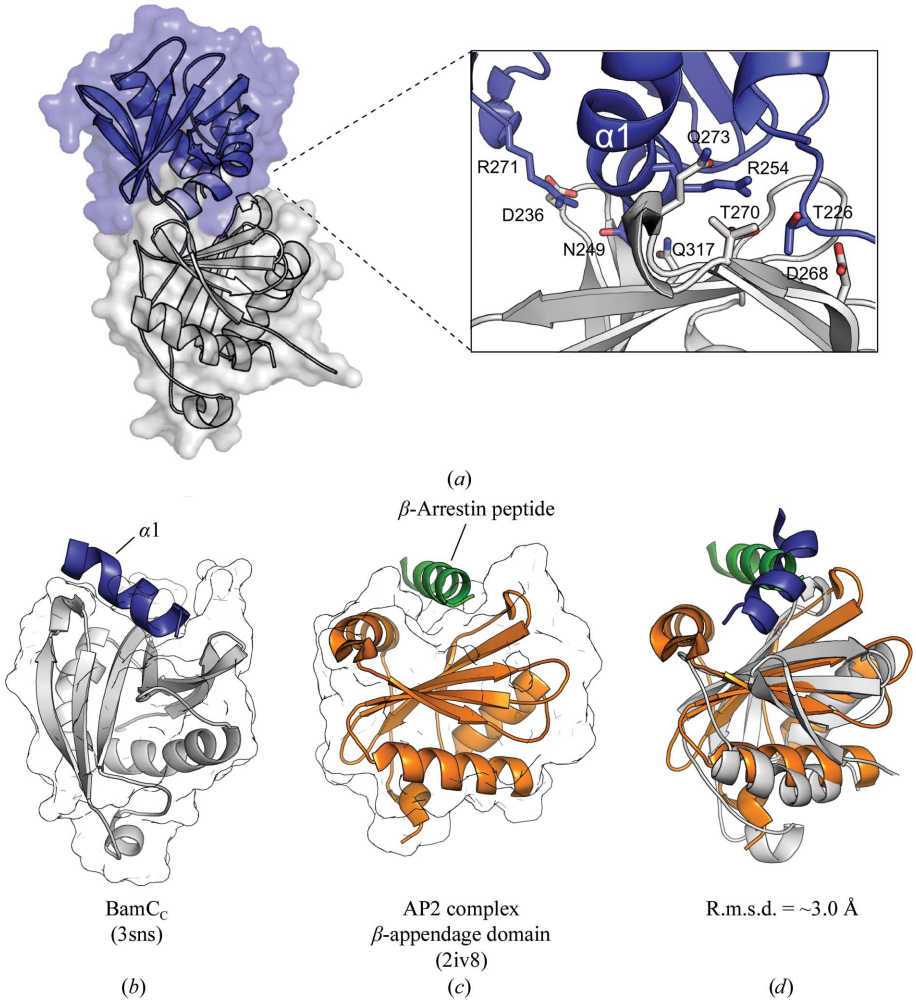


Figure 6
A potential protein-interaction surface on BamC$_C$. (*a*) The most extensive protein–protein interaction within the *H*3 BamC$_C$ crystalline lattice is shown. A close-up view of the residues at the interface is shown (inset). (*b*) A different view of the interface shows how BamC$_C$ (grey) interacts with α1 of a neighbouring BamC$_C$. Only the α1 helix of the interacting BamC$_C$ monomer (blue) is shown for clarity. (*c*) A structural homologue of BamC$_C$, *Homo sapiens* AP2 complex β-appendage domain (orange), is shown with bound substrate, a peptide that mimics a region of β-arrestin (PDB entry 2iv8; green). (*d*) The structure of the AP2 complex β-appendage domain (orange) bound to β-arrestin peptide (green) is superimposed on the structures of BamC$_C$ (grey) and α1 (blue) of a neighbouring BamC$_C$.

accommodate an α-helix. A similar-shaped cavity also exists in the N-terminal domain, although this region is not as well conserved.

If the conserved negatively charged cavity of BamC$_C$ mentioned above is actually involved in a biologically important protein–protein interaction, one candidate protein that could bind to this region of BamC$_C$ is BamD, which is mostly α-helical and has been proposed to interact with BamC *via* its C-terminal region (Malinverni *et al.*, 2006). The proposed binding surface of the BamD C-terminal region is lined with several positively charged residues which would complement the negatively charged BamC$_C$ binding cleft. However, the possibilities of other proteins such as BamE or even substrate proteins binding to this region of BamC$_C$ should not be ruled out.

There are still many aspects of the structure and function of BamC that need to be addressed. For example, what is the functional significance of having two domains with the same fold connected by a flexible linker? If each domain of BamC participates in protein binding, a flexible linker between them could allow the molecule to bind to its interaction partner by wrapping around it or to bind independently to two separate interaction partners. This may be how BamC interacts with BamD and possibly with substrates. Another unique structural feature of BamC is the unstructured N-terminal region, which is ~75 residues long. Interestingly, the N-terminus is the most conserved part of BamC, but there is currently a lack of data to help elucidate its role, if any, in β-barrel protein assembly. Continued progress in the structural and functional analysis of the BAM complex and its individual components will enhance our understanding of how β-barrel proteins assemble and insert into the outer membrane.

This work was supported in part by the Canadian Institute of Health Research (to MP), the National Science and Engineering Research Council of Canada (to MP), the Michael Smith Foundation for Health Research (to MP) and the Canadian Foundation of Innovation (to MP).

References

Albrecht, R. & Zeth, K. (2010). *Acta Cryst.* F**66**, 1586–1590.
Albrecht, R. & Zeth, K. (2011). *J. Biol. Chem.* **286**, 27792–27803.
Anwari, K., Poggio, S., Perry, A., Gatsos, X., Ramarathinam, S. H., Williamson, N. A., Noinaj, N., Buchanan, S., Gabriel, K., Purcell, A. W., Jacobs-Wagner, C. & Lithgow, T. (2010). *PLoS One*, **5**, e8619.

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Bos, M. P., Robert, V. & Tommassen, J. (2007). *Annu. Rev. Microbiol.* **61**, 191–214.
Charlson, E. S., Werner, J. N. & Misra, R. (2006). *J. Bacteriol.* **188**, 7186–7194.
Clantin, B., Delattre, A. S., Rucktooa, P., Saint, N., Méli, A. C., Locht, C., Jacob-Dubuisson, F. & Villeret, V. (2007). *Science*, **317**, 957–961.
Delcour, A. H. (2009). *Biochim. Biophys. Acta*, **1794**, 808–816.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Fitzpatrick, D. A. & McInerney, J. O. (2005). *J. Mol. Evol.* **60**, 268–273.
Gatsos, X., Perry, A. J., Anwari, K., Dolezal, P., Wolynec, P. P., Likić, V. A., Purcell, A. W., Buchanan, S. K. & Lithgow, T. (2008). *FEMS Microbiol. Rev.* **32**, 995–1009.
Gatzeva-Topalova, P. Z., Walton, T. A. & Sousa, M. C. (2008). *Structure*, **16**, 1873–1881.
Gatzeva-Topalova, P. Z., Warner, L. R., Pardi, A. & Sousa, M. C. (2010). *Structure*, **18**, 1492–1501.
Gentle, I. E., Burri, L. & Lithgow, T. (2005). *Mol. Microbiol.* **58**, 1216–1225.
Gentle, I., Gabriel, K., Beech, P., Waller, R. & Lithgow, T. (2004). *J. Cell Biol.* **164**, 19–24.
Hagan, C. L., Silhavy, T. J. & Kahne, D. (2011). *Annu. Rev. Biochem.* **80**, 189–210.
Heijne, G. von (2011). *Annu. Rev. Biochem.* **80**, 157–160.
Heuck, A., Schleiffer, A. & Clausen, T. (2011). *J. Mol. Biol.* **406**, 659–666.
Holm, L., Kääriäinen, S., Rosenström, P. & Schenkel, A. (2008). *Bioinformatics*, **24**, 2780–2781.
Iyer, L. M., Koonin, E. V. & Aravind, L. (2001). *Proteins*, **43**, 134–144.
Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.
Kim, K. H., Kang, H.-S., Okon, M., Escobar-Cabrera, E., McIntosh, L. P. & Paetzel, M. (2011). *Biochemistry*, **50**, 1081–1090.
Kim, S., Malinverni, J. C., Sliz, P., Silhavy, T. J., Harrison, S. C. & Kahne, D. (2007). *Science*, **317**, 961–964.
Kim, K. H. & Paetzel, M. (2011). *J. Mol. Biol.* **406**, 667–678.
Knowles, T. J. *et al.* (2011). *EMBO Rep.* **12**, 123–128.
Knowles, T. J., Jeeves, M., Bobat, S., Dancea, F., McClelland, D., Palmer, T., Overduin, M. & Henderson, I. R. (2008). *Mol. Microbiol.* **68**, 1216–1227.
Knowles, T. J., McClelland, D. M., Rajesh, S., Henderson, I. R. & Overduin, M. (2009). *Biomol. NMR Assign.* **3**, 203–206.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Liang, J., Edelsbrunner, H. & Woodward, C. (1998). *Protein Sci.* **7**, 1884–1897.
Malinverni, J. C., Werner, J., Kim, S., Sklar, J. G., Kahne, D., Misra, R. & Silhavy, T. J. (2006). *Mol. Microbiol.* **61**, 151–164.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Misra, R. (2007). *ACS Chem. Biol.* **2**, 649–651.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Noinaj, N., Fairman, J. W. & Buchanan, S. K. (2011). *J. Mol. Biol.* **407**, 248–260.
Painter, J. & Merritt, E. A. (2006). *J. Appl. Cryst.* **39**, 109–111.
Paschen, S. A., Neupert, W. & Rapaport, D. (2005). *Trends Biochem. Sci.* **30**, 575–582.
Sandoval, C. M., Baker, S. L., Jansen, K., Metzner, S. I. & Sousa, M. C. (2011). *J. Mol. Biol.* **409**, 348–357.
Schleiff, E. & Soll, J. (2005). *EMBO Rep.* **6**, 1023–1027.
Schmid, E. M., Ford, M. G., Burtey, A., Praefcke, G. J., Peak-Chew, S.-Y., Mills, I. G., Benmerah, A. & McMahon, H. T. (2006). *PLoS Biol.* **4**, e262.
Silhavy, T. J., Kahne, D. & Walker, S. (2010). *Cold Spring Harb. Perspect. Biol.* **2**, a000414.
Sklar, J. G., Wu, T., Gronenberg, L. S., Malinverni, J. C., Kahne, D. & Silhavy, T. J. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 6400–6405.
Sklar, J. G., Wu, T., Kahne, D. & Silhavy, T. J. (2007). *Genes Dev.* **21**, 2473–2484.
Vanini, M. M., Spisni, A., Sforça, M. L., Pertinhez, T. A. & Benedetti, C. E. (2008). *Proteins*, **71**, 2051–2064.
Volokhina, E. B., Beckers, F., Tommassen, J. & Bos, M. P. (2009). *J. Bacteriol.* **191**, 7074–7085.
Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M. & Tommassen, J. (2003). *Science*, **299**, 262–265.
Voulhoux, R. & Tommassen, J. (2004). *Res. Microbiol.* **155**, 129–135.
Warner, L. R., Varga, K., Lange, O. F., Baker, S. L., Baker, D., Sousa, M. C. & Pardi, A. (2011). *J. Mol. Biol.* **411**, 83–95.