Structural Characterization of *Escherichia coli* BamE, a Lipoprotein Component of the β-Barrel Assembly Machinery Complex[†]

Kelly H. Kim,[‡] Hyun-Seo Kang,[§] Mark Okon,[§] Eric Escobar-Cabrera,[§] Lawrence P. McIntosh,*,[§] and Mark Paetzel*,[‡]

[‡]*Department of Molecular Biology and Biochemistry, Simon Fraser University, South Science Building, 8888 University Drive, Burnaby, British Columbia, Canada V5A 1S6, and* [§]*Department of Biochemistry and Molecular Biology, Department of Chemistry, and The Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3*

ABSTRACT: In *Escherichia coli*, the BAM complex catalyzes the essential process of assembling outer membrane proteins (OMPs). This complex consists of five proteins: one membrane-bound protein, BamA, and four lipoproteins, BamB, BamC, BamD, and BamE. Despite their importance in OMP biogenesis, there is currently a lack of functional and structural information on the BAM complex lipoproteins. BamE is the smallest but most conserved lipoprotein in the complex. The structural and dynamic properties of monomeric BamE (residues 21−133) were determined by NMR spectroscopy. The protein folds as two α-helices packed against a three-stranded antiparallel β-sheet. The N-terminal (Ser21−Thr39) and C-terminal (Pro108−Asn113) residues, as well as a β-hairpin loop (Val76−Gln89), are highly flexible on the subnanosecond time scale. BamE expressed and purified from *E. coli* also exists in a kinetically trapped dimeric state that has dramatically different NMR spectra, and hence structural features, relative to its monomeric form. The functional significance of the BamE dimer remains to be established. Structural comparison to proteins with a similar architecture suggests that BamE may play a role in mediating the association of the BAM complex or with the BAM complex substrates.

In Gram-negative bacteria, the outer membrane functions as a molecular sieve that regulates the traffic of solutes into and out of the cell. Residing in the outer membrane lipid bilayer are β-barrel proteins, collectively known as the *outer membrane proteins* (OMPs),[1] that fulfill a diverse range of biological functions such as nutrient uptake, antibiotic resistance, cell adhesion, and maintenance of membrane selective permeability (*1−3*). Proper folding and membrane insertion of these proteins are essential for maintaining the structural and functional integrity of the outer membrane and thus also for cell viability (*1*, *4*). In *Escherichia coli*, this process of OMP folding and membrane insertion is catalyzed by the multicomponent BAM (*b*eta-barrel *a*ssembly *m*achinery) complex, which was formerly known as the YaeT (Omp85) complex (Figure 1) (*4−10*).

The BAM complex in *E. coli* consists of the following five proteins with their commonly known names within parentheses: BamA (YaeT/Omp85), BamB (YfgL), BamC (NlpB), BamD (YfiO), and BamE (SmpA) (*7*, *10*). BamA is a β-barrel integral membrane protein (i.e., it is an OMP), and the remaining members of the complex are lipoproteins anchored to the periplasmic face of the outer membrane via a lipid moiety (*11*, *12*). Previous studies have shown that the loss of a gene encoding BamA or BamD completely halts OMP biogenesis and leads to cell death (*8*, *13−16*). In contrast, deletion of the genes coding for BamB, BamC, or BamE results in decreased efficiency of OMP folding and assembly (*7*, *17*, *18*). Although progress has been made (*12*), the detailed molecular mechanism of how these five different proteins work together as a BAM complex to assist in OMP folding and membrane insertion remains poorly understood.

In addition to ongoing genetics and biochemical analysis of the BAM complex, structural analyses of the individual members and the complex as a whole could provide important functional and mechanistic insights. Thus far, only the structure of the periplasmic domain of BamA (PDB: 3EFC, 3OG5) and the structure of BamB (PDB: 3P1L) from *E. coli* have been crystallographically determined (*19−22*). Currently, efforts are underway by various groups to determine the structures of the rest of the BAM complex components (*23−25*). The crystal structure of a BamA homologue in *Bordetella purtussis* (FhaC) is also available (*26*), as well as the NMR-derived structure of a BamE homologue in *Xanthomonas axonaopidis* pv. citri (OmlA) (*27*). The currently available crystal structures of BamA and its homologue revealed that this protein forms a β-barrel with a long periplasmic tail consisting of five *p*olypeptide *tr*ansport *a*ssociated (POTRA) motifs (*26*, *28*). The BAM complex is

*To whom correspondence should be addressed. M.P.: e-mail, mpaetzel@sfu.ca; phone, 778-782-4320; fax, 778-782-5583. L.P.M.: e-mail, mcintosh@chem.ubc.ca; phone, 604-822-3341; fax, 604-822-5227.

[1]Abbreviations: HSQC, heteronuclear single-quantum correlation; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; BAM, beta-barrel assembly machinery; SmpA, small protein A (BamE); OMP, outer membrane protein; BamE[21−113], the BamE construct used in this study. It represents the full-length mature BamE protein after cleavage of its first 20 residues (the signal sequence). The cysteine residue at position 20, which serves as the lipid attachment site, was removed to prevent potential intermolecular disulfide bond formation.
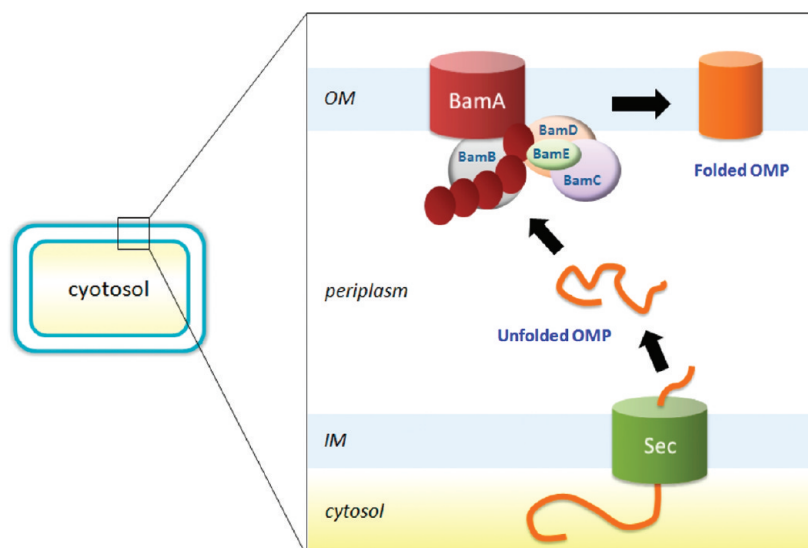
FIGURE 1: A schematic diagram of the OMP secretion and assembly pathway in *E. coli*. Following their synthesis in the cytosol, OMPs (orange) are translocated across the inner membrane in an unfolded state via the Sec translocation system (green). The OMPs are then released into the periplasm and subsequently escorted by chaperones to the BAM complex, which is a multicomponent protein complex consisting of BamA, BamB, BamC, BamD, and BamE in a yet undefined stoichiometry. By an unknown molecular mechanism, the BAM complex facilitates the assembly and insertion of the OMPs into the outer membrane lipid bilayer.

currently visualized as a large molecular machine in which BamA is the major structural and functional component, with BamB, BamC, BamD, and BamE serving as its accessory proteins to enhance its efficiency. The POTRA motifs of BamA are predicted to serve as docking sites for the lipoproteins BamB−E (*7, 21*). The POTRA motifs also seem to be important for initial substrate (i.e., unfolded OMPs) recognition and chaperone-like activity (*22, 29, 30*).

In comparison to BamA, the lipoprotein components of the BAM complex are much less well characterized. The gene encoding BamD (YfiO) is essential for viability of *E. coli* and is found ubiquitously in all Gram-negative bacteria. BamE (SmpA) is also conserved in all *Proteobacteria*. BamB (YfgL) and BamC (NlpB) are conserved in many Gram-negative bacteria, albeit to a lesser extent than BamD or BamE (*11*). At present, there is a lack of experimental evidence to make a prediction about the specific roles these lipoproteins play in the BAM complex, but it has been speculated that they could be involved in stabilization of the complex structure and/or in increasing the functional efficiency of BamA in OMP folding and membrane insertion (*7*). A recent study has also suggested that the homologous lipoproteins in *Caulobacter crescentus* may interact with Pal, a peptidoglycan binding lipoprotein postulated to anchor the BAM complex to the peptidoglycan layer of the cell wall (*31*).

To gain insights into the spatial organization of the BAM complex, we have initiated structural and biochemical studies on its constituent proteins. This paper specifically focuses on BamE, the smallest but most conserved lipoprotein subunit of the BAM complex. Previous studies have emphasized the importance of BamE in maintaining membrane integrity and normal levels of OMPs, as well as its role in bacterial stalk growth and stabilizing the BAM complex structure (*7, 31, 32*). Here, we present the structural and dynamic properties of *E. coli* BamE (SmpA) obtained by NMR spectroscopy.

EXPERIMENTAL PROCEDURES

Cloning. A 279 base pair DNA fragment, coding for residues 21−113 of *E. coli* BamE (SmpA), was amplified from *E. coli* K-12

genomic DNA using the forward primer 5-ATACATATGTC-CACTCTGGAG and the reverse primer 5-TATACTCGAGT-TATTAGTTACCACTC that contain the restriction sites *Nde*I and *Xho*I, respectively. The PCR product was ligated into vector pET28a (Novagen), and the resulting $His_{\times 6}$-$BamE^{21-113}$ construct encodes BamE (residues 21−113) with a cleavable N-terminal hexahistidine affinity tag. Subsequent DNA sequencing (Macrogen) confirmed that the BamE insert matched the sequence reported in the Swiss-Prot database (P0A937).

Protein Expression and Purification. The expression plasmid was transformed into *E. coli* BL21(λDE3). Uniformly ^{15}N-labeled $His_{\times 6}$-$BamE^{21-113}$ was expressed in M9 media supplemented with 1 g/L $^{15}NH_4Cl$ (Sigma-Aldrich). Uniformly $^{15}N/^{13}C$-labeled $His_{\times 6}$-$BamE^{21-113}$ was expressed in M9 media containing 3 g/L $[^{13}C_6]$glucose (Sigma-Aldrich) and 1 g/L $^{15}NH_4Cl$. For both isotopically labeled samples, cultures were grown at 37 °C to an OD_{600} of 0.6 and induced with 1 mM IPTG overnight at 25 °C. Cells were harvested by centrifugation and subsequently lysed using an Avestin Emulsiflex-3C cell homogenizer in buffer A (20 mM Tris-HCl, pH 8.0, 100 mM NaCl). The resulting lysate was clarified by centrifugation (45000*g*) for 30 min at 4 °C and loaded on a Ni^{2+} affinity chromatography column (Quiagen). The protein was eluted with a step gradient of 100−500 mM imidazole in buffer A in 100 mM increments. The fractions containing BamE were pooled, followed by incubation with thrombin (GE Healthcare) overnight for cleavage of the N-terminal hexahistidine tag. The digested protein sample was concentrated to approximately 10 mg/mL using an Amicon ultracentrifugal filter device (Millipore) with a 3 kDa MW cutoff and was then further purified by size-exclusion chromatography (Sephacryl S-100 HiPrep 26/60 column) on an AKTA Prime system (GE Healthcare). In this last step of protein purification, the buffer was also exchanged to 20 mM Na_2HPO_4/NaH_2PO_4, pH 6.8, and the monomeric and dimeric forms of BamE were resolved. The final protein (Ser21−Asn113 with four remnant N-terminal residues, Gly-Ser-His-Met, resulting from cloning and thrombin cleavage) is 96 residues in length and has a calculated molecular mass of 10562 Da and a calculated isoelectric point of 6.9. The purified
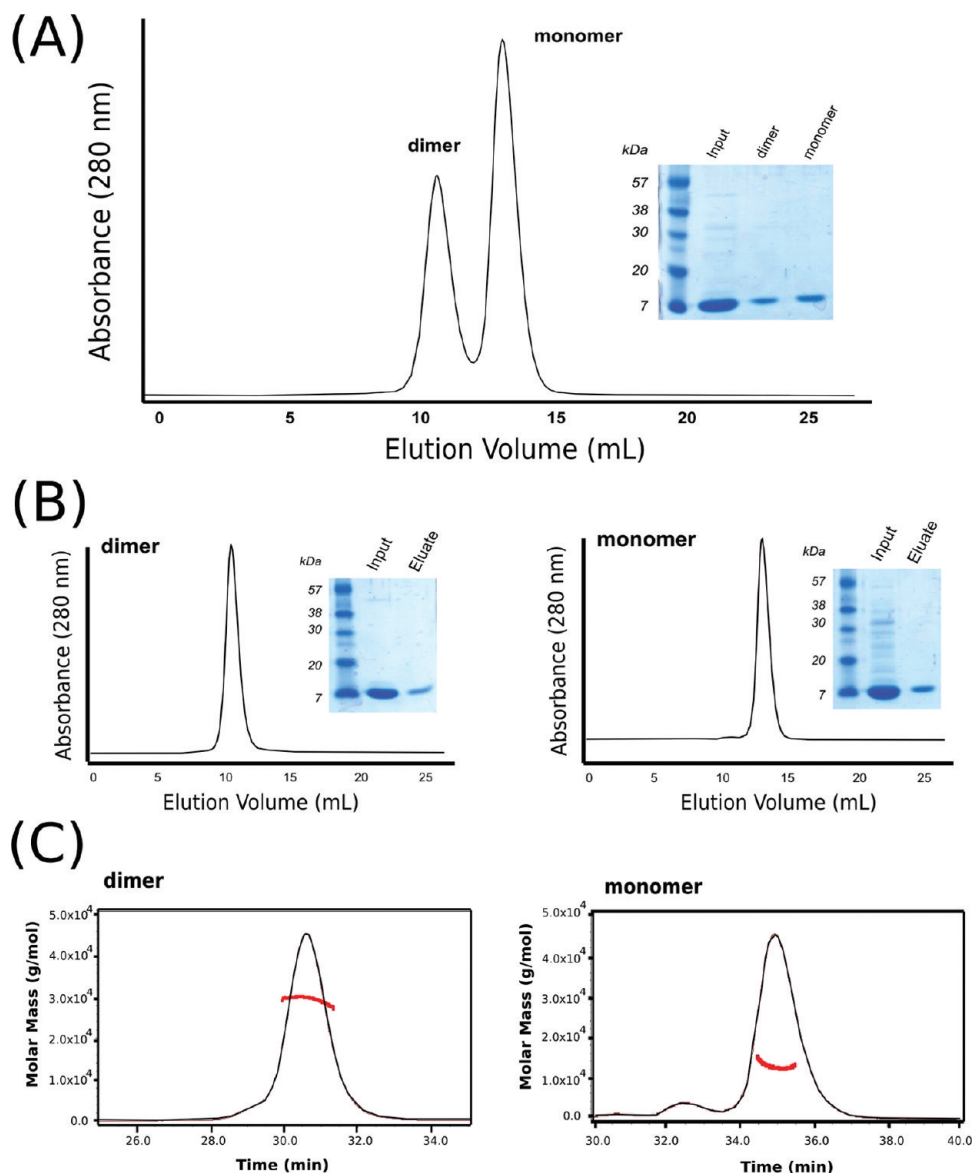
FIGURE 2: Isolated *E. coli* BamE$^{21-113}$ exists as a monomer and dimer. (A) Following nickel affinity chromatography, His$_{\times 6}$-BamE$^{21-113}$ was subjected to thrombin digest for tag removal and subsequently to size-exclusion chromatography using a calibrated Superdex 75 HR 10/30 column. Two major peaks were observed on the chromatogram, one eluting at an elution volume that is expected for an approximately 10−15 kDa species and the other for a 25−30 kDa species. However, fractions corresponding to each peak yielded a single band on SDS−PAGE with the apparent molecular mass expected for monomeric BamE$^{21-113}$ (~11 kDa). No other significant proteins of higher molecular weight were observed. Thus the expressed BamE$^{21-113}$ exists in both monomeric and dimeric states. (B) The dimer and the monomer fractions from (A) were collected separately, pooled, and subsequently subjected to a second size-exclusion chromatography run to determine whether there is a concentration-dependent monomer/dimer equilibrium. A single peak was observed in both chromatograms, demonstrating that the dimeric and the monomeric species do not interconvert under the conditions or time scale of this experiment (over the period of approximately 1 week). The SDS−PAGE of the input and the eluate samples are shown beside each chromatogram. (C) After purification by size-exclusion chromatography, the molecular masses of the monomeric and dimeric forms of BamE$^{21-113}$ were verified by multiangle dynamic light scattering analysis. The chromatogram from an in-line gel filtration column is shown in black and the calculated molecular mass in red. The calculated values were 12.4 ± 0.75 and 28.5 ± 3 kDa for the monomer and the dimer, respectively. Note also that the two peaks are monodisperse.

protein sample was stored in 4 °C until further use. The final protein concentrations of the samples used for NMR data acquisition were ~0.5 mM.

Analytical Size-Exclusion Chromatography. Apparent molecular mass of purified BamE$^{21-113}$ was determined by gel filtration chromatography using a calibrated Superdex 75 column (GE Healthcare). A sample of 200 μL of BamE$^{21-113}$ (5 mg/mL) was injected, resolved, and analyzed at a flow rate of 0.5 mL/min in buffer A. The oligomeric state of BamE$^{21-113}$ was also monitored using gel filtration chromatography under different conditions of sample pH values (CH$_3$COONa, pH 3.5, MES pH

6.5, Tris-HCl, pH 8.0, CAPS, pH 10) and salt concentrations (0 mM, 100 mM, 300 mM, 500 mM, and 1 M NaCl). An unlabeled protein sample, produced using *E. coli* grown in LB media, was used for the oligomeric state analysis.

Multiangle Light Scattering Analysis. The oligomeric state of purified BamE$^{21-113}$ was determined by gel filtration chromatography (Superdex 200 column; GE Healthcare) in-line with a multiangle light scattering system (Wyatt Technologies Inc.). A sample of 100 μL of purified BamE$^{21-113}$ (5 mg/mL) was injected and resolved at a flow rate of 0.5 mL/min in buffer A. Molecular masses of the monomeric and dimeric form of BamE
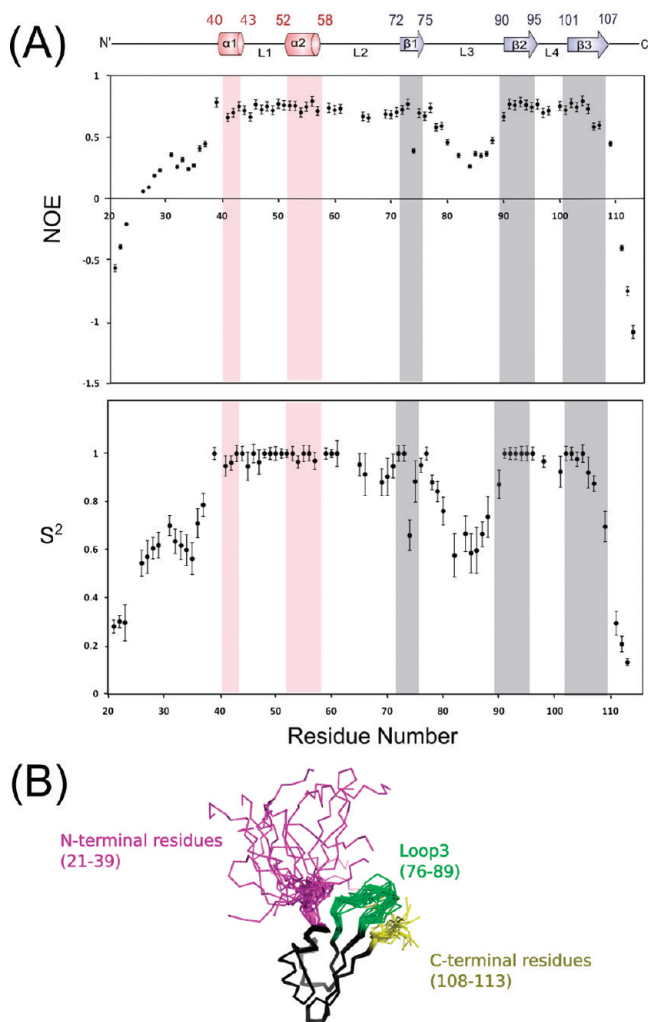
were determined by a multiangle light scattering (MALS) Dawn-EOS instrument with a 684 nm laser (Wyatt Technologies, Inc.) coupled to refractive index instrument (Optilab Rex; Wyatt Technologies, Inc.). The molar mass of the protein was calculated from the observed light scattering intensity and differential refractive index (*33*−*35*) using ASTRA v5.1 software (Wyatt Technologies, Inc.) based on Zimm fit method using a refractive index increment, $dn/dc = 0.185 \, \text{L g}^{-1}$.

NMR Data Acquisition. NMR spectra were recorded at 15 °C on Varian Unity 500 and 600 MHz NMR spectrometers. The low temperature was used to ensure stability of the protein sample during data collection. All samples consisted of ~0.5 mM protein in 20 mM Na_2HPO_4/NaH_2PO_4, pH 6.8, and ~10% D_2O. Spectra were processed using NMRPipe (*36*) and analyzed using Sparky (*37*). NMR chemical shifts were referenced directly or indirectly to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS).

Spectral Assignments and Structure Calculation. Using an extensive set of multidimensional NMR experiments, the backbone and side chain ^1H, ^{13}C, and ^{15}N chemical shifts of BamE were assigned by standard methods (*38*). These spectral assignments agree with those reported recently for a similar BamE construct (*25*). The BamE structure was then calculated and refined using ARIA 2.2 with CNS 1.2 (*39*). NOE-derived distance restraints were obtained from simultaneous regular and

FIGURE 3: NMR spectroscopy demonstrates that the monomeric and dimeric forms of ^{15}N-labeled BamE$^{21−113}$ have significantly different structures. (A) The ^{15}N-HSQC spectrum of BamE$^{21−113}$ monomer is shown with peaks assigned. The well-dispersed signals from $^1H-^{15}N$ groups confirm that the monomeric form of the protein is stably folded and a good candidate for further structural analysis. (B) The superimposed ^{15}N-HSQC spectra of the BamE$^{21−113}$ monomer (red) and dimer (gray) show very little peak overlap, indicating distinctly different backbone conformations. The samples used in panel B retained the His$_{\times 6}$ tag, accounting for the extra sharp peaks relative to panel A.

Table 1: NMR Restraints and Structural Statistics for BamE$^{21−113}$ Ensemble

summary of restraints		
NOEs		
intraresidue		726
sequential		316
medium range ($1 < |i − j| < 5$)		139
long range ($|i − j| \geq 5$)		299
total		1480
dihedral angles (ϕ, ψ, $\chi 1$)		55, 55, 0
hydrogen bonds		10×2
residues in allowed regions of Ramachandran plot, %[a]		98.4
mean energies, kcal mol^{-1}		
E_{vdw}		-247.2 ± 19.0
E_{bonds}		43.1 ± 2.6
E_{angles}		153.2 ± 9.8
E_{impr}		73.2 ± 8.5
E_{NOE}		181.9 ± 15.5
E_{cdih}		3.3 ± 1.0
rms deviation, Å		
	structured elements[b]	all[c]
backbone atoms	0.22	0.94
all heavy atoms	0.52	1.38

[a]Calculated with Procheck-NMR (*50*) and summed over most favored, allowed, and generously allowed regions. [b]Core structured region identified from Promotif (*46*), DSSP (*45*), SSP (*69*), and MOLMOL (*56*). α1, 40−43; α2, 52−58; β1, 72−75; β2, 90−95; β3, 101−107. [c]All of the atoms except the flexible N- and C-terminal regions, 40−107.

FIGURE 4: The NMR-derived structural ensemble of *E. coli* BamE$^{21−113}$. (A) A topology diagram with strands shown as arrows, helices as cylinders, and loops as lines. (B) A ribbon diagram of the lowest energy BamE$^{21−113}$ structure with least restraint violations. (C) An ensemble of 20 structures. Colors change progressively from the N-terminus (blue) to the C-terminus (red).

FIGURE 5: Backbone dynamics of *E. coli* BamE$^{21-113}$ from amide ^{15}N relaxation analysis. (A) Plots of heteronuclear NOE (upper panel) and fit isotropic model-free S^2 values (lower panel) versus sequence are shown. Smaller NOE and S^2 values, indicative of significant subnanosecond time scale backbone motions, are observed for both the N- and C-termini, as well as the loop L3. (B) These dynamic regions correspond to regions of the BamE$^{21-113}$ structural ensemble with the highest rms deviations.

constant time methyl 3D ^{13}C- and ^{15}N-NOESY-HSQC spectra, all with 100 ms mixing times (*40*). An initial set of NOE cross-peaks was assigned manually, and the remaining signals were assigned automatically by ARIA. Backbone dihedral angles were determined from ^{13}C$^\alpha$, ^{13}C$^\beta$, ^{13}C′, ^{1}H$^\alpha$, and ^{1}H^N chemical shifts using TALOS (*41*). A limited set of hydrogen bond distance restraints were included for selected amides located in β-strands, as determined via manual inspection of NOE patterns and chemical shift information. The chemical shifts and structural coordinates of the BamE$^{21-113}$ ensemble have been deposited in the BioMagResData Bank (accession number: 16926) and RCSB Protein Data Bank (accession number: 2kxx), respectively.

Backbone Amide Relaxation Measurements. Backbone amide relaxation data of ^{15}N-labeled BamE were acquired on a 500 MHz spectrometer at 28 °C (*42*). ^{15}N T_1 and T_2 lifetimes and heteronuclear ^{1}H−^{15}N NOE values were fit using Sparky (*37*) and analyzed according to the model-free formalism with TENSOR2 (*43*). The predicted global tumbling time was calculated using the program HYDRONMR (*44*).

Structural Analysis. The secondary structural analysis was performed with the program DSSP (*45*). Intramolecular interaction and fold analysis was performed with PROMOTIF 3.0 (*46*). The programs Coot (*47*) and PDBeFold (*48*) were used to overlap coordinates for structural comparison. Volume and surface area calculations were performed with UCSF CHIMERA (*49*). The stereochemistry of the structures was analyzed with the program PROCHECK (*50*). The DALI (*51*), CATH (*52*), and FAT-CAT (*53*) servers were used to find proteins with similar folds. The surface electrostatics analysis was performed with the adaptive Boltzmann−Poisson solver plug-in (*54*) within PyMol (*55*) using dielectric constants of 2 and 80 for solute and solvent, respectively.

Figure Preparation. Figures were prepared using PyMol (*55*) and MolMol (*56*). The alignment figure was prepared using the programs CLUSTALW (*57*) and ESPript (*58*), and the protein topology diagram was prepared using the program TopDraw (*59*).

RESULTS AND DISCUSSION

BamE Oligomerization State Analysis. We have produced a soluble construct of BamE that encompasses the entire wild-type sequence immediately following the cleavable N-terminal signal sequence and the conserved lipidation residue Cys20 (Ser21−Asn113). The purified BamE$^{21-113}$ was found to exist in both monomeric and dimeric states, as determined by analytical gel filtration chromatography (Figure 2A). The dimeric and monomeric fractions were separately collected, and each sample was run through the size-exclusion column again to test whether the two states exist in a concentration-dependent equilibrium. Our result reveals that both BamE$^{21-113}$ dimer and monomer remain in their original oligomeric states and do not interconvert under the conditions and time scale of these measurements (over the period of approximately 1 week) (Figure 2B). The homogeneity and molecular mass of each form of BamE$^{21-113}$ were confirmed by light scattering analysis (Figure 2C).

Additional analytical gel filtration chromatography was performed to determine whether dimer formation or dissociation is affected by conditions. Neither pH (3.5, 6.5, 8.0, and 10), salt concentration (0 mM, 100 mM, 300 mM, 500 mM, and 1 M NaCl), nor the presence of a detergent (0.01% *n*-dodecyl β-D-maltoside) induced dimerization of BamE$^{21-113}$ monomers or dissociation of the dimer (data not shown). Thus self-association is not due to simple electrostatic or hydrophobic interactions. Also, since the protein lacks cysteine residues, dimerization of BamE$^{21-113}$ cannot be due to disulfide bond formation.

To investigate further the self-association of BamE$^{21-113}$, we recorded the ^{15}N-HSQC spectra of the two purified forms of the ^{15}N-labeled protein (Figure 3). The spectrum of the monomer shows well-dispersed signals, indicative of a stable, folded protein. In contrast, the dimeric form yielded a spectrum with signals of significantly differing intensities, suggestive of both ordered and disordered regions or regions undergoing conformational exchange on a millisecond to microsecond time scale. More importantly, the spectra of the two forms of BamE$^{21-113}$ show remarkably little overlap, suggesting that the monomeric and dimeric forms have substantially different structures. Combined with the lack of observable interconversion, we thus hypothesize that BamE$^{21-113}$ can adopt a kinetically trapped intertwined or perhaps domain-swapped dimeric conformation (*60*, *61*). It is presently not clear which form exists within the BAM complex or
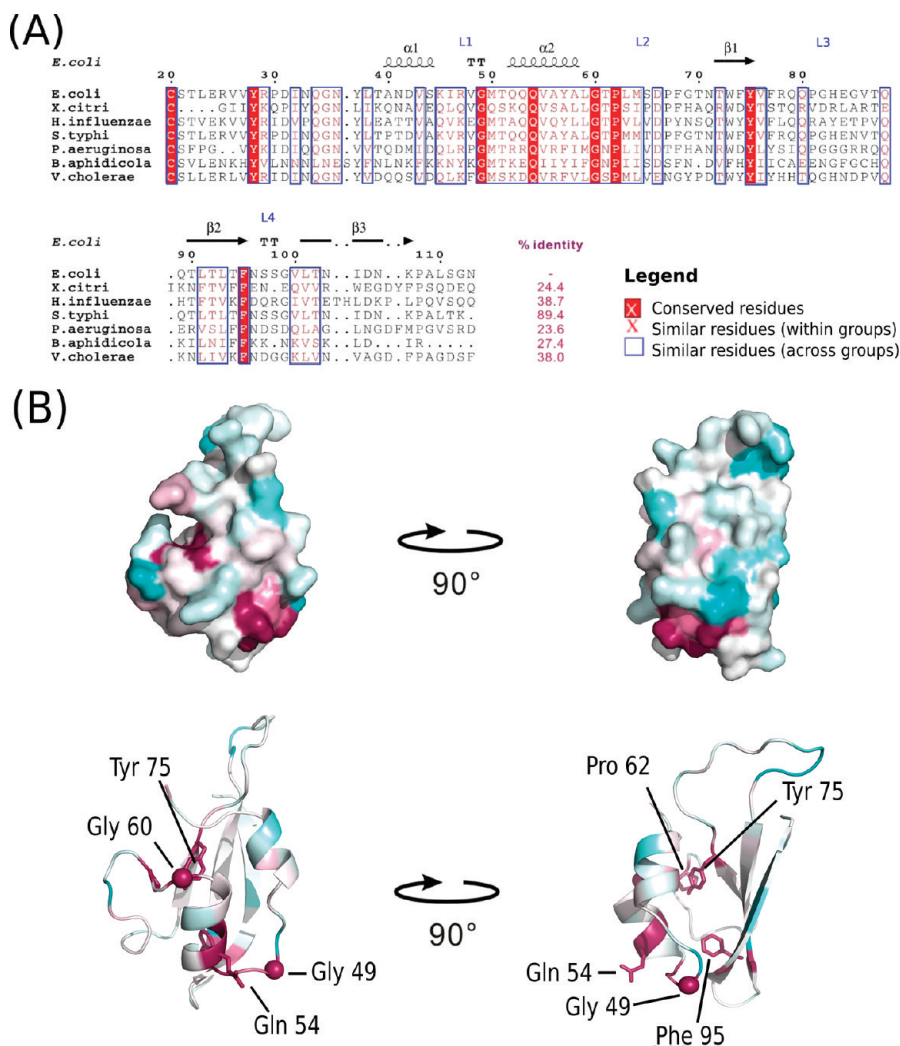
FIGURE 6: Conserved amino acids within BamE homologues in Gram-negative bacteria. (A) Sequence alignment starting from the invariant N-terminal cysteine residue (the preceding signal sequence is cleaved off in the mature protein). The NMR-derived secondary structure of *E. coli* BamE$^{21-113}$ as classified by DSSP (*45*) is shown above the alignment. Invariant residues are shown in red boxes, similar residues in red text, and stretches of amino acids that are similar across the group of sequences in blue boxes. The protein sequences were acquired from the Swiss-Prot database: *E. coli* (P0A937); *X. citri* (Q8PMB6); *H. influenza* (P44057); *S. typhi* (Q8XF17); *P. aeruginosa* (O68562); *B. aphidicola* (Q8K9V7); *V. cholerae* (P0C6Q9). (B) A view of BamE sequence conservation mapped onto the BamE$^{21-113}$ surface (top). Individual amino acid residues are colored according to the degree to which they are conserved; absolutely conserved residues are shown in maroon, while highly variable residues are shown in blue. In the ribbon diagram (bottom), the conserved residues are shown in stick representation.

if the BamE dimerization even holds a functional significance. Accordingly, all subsequent structural and dynamics analyses described in this study were carried out with the monomeric form of BamE$^{21-113}$.

NMR-Derived Structure of BamE$^{21-113}$. Using an extensive set of NOE-derived distance and chemical shift-derived dihedral angle restraints, we calculate the structural ensemble of monomeric BamE$^{21-113}$ with the program ARIA (Figure 4). Table 1 shows a summary of the NMR data and structural statistics. The root-mean-square (rms) deviations between the 20 lowest energy structures for the helical and strand regions of the protein were 0.22 Å (backbone atoms) and 0.52 Å (all heavy atoms).

BamE$^{21-113}$ has a well-structured core that is made up of two N-terminal antiparallel α-helices (α1, Ala40−Val43; α2, Gln52−Ala58) and a C-terminal twisted antiparallel β-sheet consisting of three β-strands (β1, Thr72−Tyr75; β2, Thr90−Phe95; β3, Leu101−Lys107) (Figure 4). Residues Pro67−Gly69 also form a helical-like turn. Collectively, these secondary structural elements yield a two-layer sandwich fold with α1 and α2 packing

against the β-sheet. Together, residues 40−107 have the approximate dimensions of 22 Å × 46 Å × 24 Å with a surface area of ~5000 Å^2 and volume of ~8200 Å^3. In contrast to the well-ordered core, the N- (residues 21−39) and C- (residues 108−113) terminal segments of BamE$^{21-113}$ and the 14 residue loop L3 (residues 76−89) joining β1 and β2 appear disordered with high rms deviations in the calculated ensemble due to a lack of structural restraints. This mobility was confirmed by ^{15}N-relaxation measurements, as discussed below.

Backbone Dynamics of BamE$^{21-113}$. In parallel with the structural analysis of BamE$^{21-113}$, we investigated the dynamic properties of the protein using ^{15}N T_1, T_2, and heteronuclear NOE relaxation measurements (Figure 5). Fitting the T_1 and T_2 data for the ordered main-chain amides (i.e., with heteronuclear NOE values > 0.65) by the model-free formalism yielded a correlation time of approximately 10 ns for the global isotropic tumbling of BamE$^{21-113}$. This is somewhat slower than predicted for the lowest energy NMR-derived structure of monomeric BamE$^{21-113}$ using the program HYDRONMR (8.6 ns), yet faster than expected for a globular 21 kDa dimer (~12 ns) (*66*). This
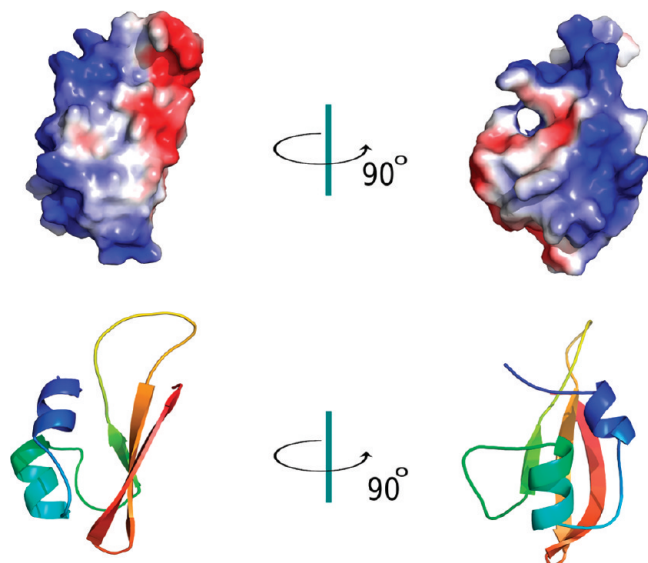
FIGURE 7: Electrostatic properties of the BamE[21−113] molecular surface. The electrostatic potential is mapped onto the solvent-accessible surface of BamE[21−113] (upper panel). The red, blue, and white represent negative, positive, and neutral potentials, respectively. The protein is also shown in ribbon diagram (lower panel) in the same orientation as the surface diagram, with the same coloring as in Figure 4.

difference may reflect weak self-association, although the ^{15}N-HSQC spectra of monomeric BamE[21−113] do not show any concentration-dependent changes upon diluting the protein from 0.5 to 0.15 mM (not shown). Alternatively, the disordered termini and large L3 loop may lead to an increased effective hydrodynamic size. This is consistent with the gel filtration studies in which the BamE[21−113] monomer was observed to elute from the column slightly earlier than expected, at a volume corresponding to a protein species of approximately 15 kDa in size rather than 11 kDa.

In addition to reflecting global motions, amide ^{15}N relaxation provides insights into the local backbone motions of a protein. The residue-specific [1]H−[15]N heternuclear NOE values and fit model-free order parameters S^2 of BamE[21−113] indicate that indeed both the N- and C-termini are highly flexible on the nanosecond to picosecond time scale (Figure 5). However, the N-terminal residues proceeding α1 may not be entirely unrestricted. Some local order is suggested by the NOE and S^2 values in this region that are intermediate between those of the more distal, highly flexible terminal residues and of those of the ordered helices and strands. The extended loop L3 is also conformationally flexible on this fast time scale, although its motions are dampened relative to those of the terminal regions.

Conserved Amino Acids and Molecular Surface Properties. Comparison of the sequence of *E. coli* BamE to those of its homologues from various Gram-negative bacterial species reveals a number of conserved amino acids (Figure 6A). The majority of the conserved residues in the core of BamE (Gly49, Gly60, Pro62, Tyr75, and Phe95) reside on the loops or turns (Figure 6B). Gly49 is located in L1 (between α1 and α2) where it participates in a type II β-turn, whereas Gly60 and Pro62 are found at turning points of L2 (between α2 and β1). Two conserved aromatic residues, Tyr75 and Phe95, are found as the last residues of β-strands β1 and β2, respectively. The side chains of both these residues point toward the interface between the helices and the β-sheet. Another conserved residue, Gln54, is

found in α2. When these conserved residues are mapped onto the surface view of the BamE[21−113] structure (Figure 6B), they are clustered in two separate patches. Analyzing the electrostatic properties of solvent-accessible molecular surface of BamE showed that the protein has positively charged residues clustered on the surface formed by the two N-terminal α-helices (Figure 7). On the other hand, the V-shaped surface formed by α1 and β3 is hydrophobic (Figure 7). Further experiments are needed to verify whether these regions of BamE are involved in interaction with other proteins (e.g., other components of the BAM complex or with substrate proteins) or if they are important mainly for the structural stability and folding of this protein.

Structural Comparison with OmlA and Other Homologues Provides Clues to the Function of BamE. The structure of BamE[21−113] closely resembles that of OmlA (ref 27; PDB: 2pxg), a BamE homologue found in *X. axonopodis* pv. citri (24.4% sequence identity). Both possess similar secondary structural elements and an overall tertiary fold, and the backbone atoms of the α-helices and the β-sheet can be superimposed with an rms deviation value of 2.66 Å (Figure 8D). Although quite similar in architecture, three notable differences were observed between the BamE[21−113] and OmlA structures. (1) Residues corresponding to α1 in BamE[21−113] are disordered in OmlA. (2) The angle between the α2 helix and the C-terminal β-sheet is more acute in the OmlA structure. (3) The flexible N- and C-termini of OmlA are significantly longer than in BamE[21−113].

A search for structural homologues using the DALI (62), CATH (52), and FATCAT servers (53) identified several additional proteins that have a significant degree of similarity in topology and architecture with BamE. Proteins with a BamE-like fold include *Streptomyces clavuligerus* BLIP (β-lactamase inhibitor protein) (refs 63 and 64; PDB: 2g2u) (Figure 8A,B), the dimerization domain of an *E. coli* disulfide bond isomerase known as DsbC (ref 65; PDB: 1eej) (Figure 8C), *Thermus thermophilus* TTHA1718, a putative heavy metal binding protein (ref 66; PDB: 2roe) (Figure 8E), and *Hirudo medicinalis* eglin C, an elastase (a serine protease) inhibitor (ref 67; PDB: 1cse) (Figure 8F).

Surprisingly, the search results from all three databases indicate that BamE shares more structural similarity with BLIP, a protein that inhibits a variety of class A β-lactamase enzymes, than with its homologue, OmlA. Structural comparison of BamE and BLIP suggests that BLIP has a tandem repeat of BamE-like folds, as each of the N- and the C-terminal domains of BLIP superimpose well onto the BamE structure with rms deviation values of 1.91 and 3.34 Å, respectively (Figure 8A,B). It is interesting that BLIP exists as a tandem repeat and that our gel filtration and light scattering data suggest that BamE can form stable dimer as well as monomer in solution. It is also interesting to note that the loop L3 of BamE, which was observed to be mobile from our NMR relaxation experiment, is found in a structurally equivalent position as the active site binding loop found in both domains of BLIP (64). L3 is also topologically equivalent to an active site binding loop in eglin C, a protein-based inhibitor of the serine protease elastase (67). Therefore, we postulate that L3 of BamE may serve a similar function as a protein binding motif. In BLIP, Asp49 found in the active site binding loop serves as a key residue involved in the interaction with the β-lactamase enzymes (64). Vanini et al. (27) observed that Asp62 of OmlA and the functionally important Asp49 of BLIP are found in a structurally equivalent position in both proteins. In our *E. coli* BamE structure, a glutamate (Glu84) residue is found at the equivalent position within L3.
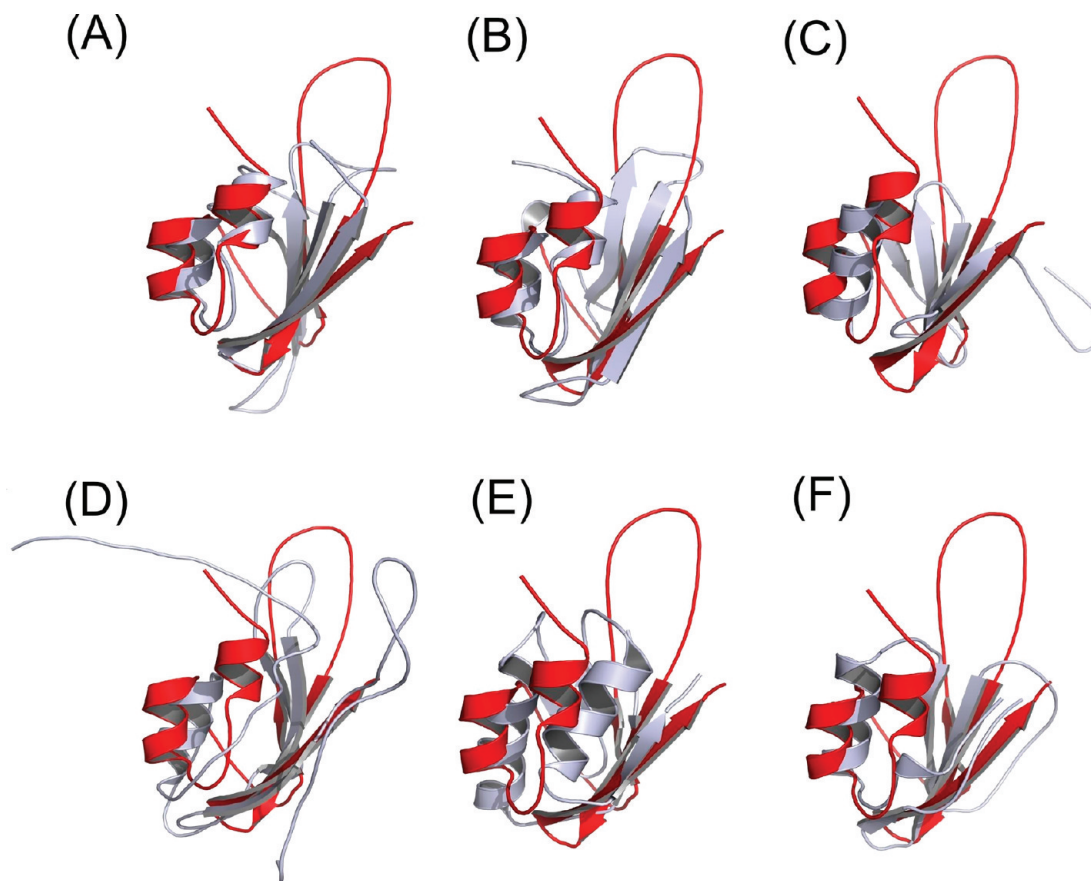
Figure	Superimposed Protein	Organism	PDB ID	RMSD (Å)	Seq ID (%)
(A)	BLIP (domain 2)	*S. clavuligerus*	2g2u	1.91	17.8
(B)	BLIP (domain 1)	*S. clavuligerus*	2g2u	3.34	19.8
(C)	DsbC (N-terminal domain)	*E. coli*	1eej	3.82	8.1
(D)	OmlA	*X. citri*	2pxg	2.66	24.4
(E)	TTHA1718	*T. thermophilus*	2roe	3.15	13.3
(F)	Eglin C	*H. medicinalis*	1cse	2.44	7.5

FIGURE 8: *E. coli* BamE$^{21-113}$ (red) is superimposed on the structures of proteins (white) with similar topology and architecture. The rms deviation values were calculated against the backbone atoms of the α-helices and β-sheets of the lowest energy BamE$^{21-113}$ structure.

Based on previous studies, it is now known that *E. coli* BamE participates in various protein−protein interactions with other members of the BAM complex, namely, BamA, BamC, and BamD (*7*, *11*, *12*). Sklar et al. (*7*) suggested that BamE plays an important role in the stabilization of the BAM complex structure, in particular strengthening the interaction between the C-terminal POTRA motif of BamA and BamD. Based on the overall structural similarity of BamE with BLIP and the observed structural flexibility of L3, it is possible that this loop is involved in the interaction of BamE with the other members of the BAM complex in a similar fashion BLIP interacts with β-lactamase enzymes. It is interesting also that there is an architectural similarity between BamE and the pro-segment or intramolecular chaperone within the serine protease subtilisin (*68*). This suggests a possibility that BamE may provide a chaperone function within the BAM complex. These hypotheses should be answered through continuing progress on the structural analysis of the β-barrel assembly machinery.

REFERENCES

1. Bos, M. P., Robert, V., and Tommassen, J. (2007) Biogenesis of the gram-negative bacterial outer membrane. *Annu. Rev. Microbiol. 61*, 191–214.
2. Costerton, J. W., Ingram, J. M., and Cheng, K. J. (1974) Structure and function of the cell envelope of gram-negative bacteria. *Bacteriol. Rev. 38*, 87–110.
3. Delcour, A. H. (2009) Outer membrane permeability and antibiotic resistance. *Biochim. Biophys. Acta 1794*, 808–816.
4. Gentle, I. E., Burri, L., and Lithgow, T. (2005) Molecular architecture and function of the Omp85 family of proteins. *Mol. Microbiol. 58*, 1216–1225.
5. Doerrler, W. T., and Raetz, C. R. (2005) Loss of outer membrane proteins without inhibition of lipid export in an *Escherichia coli* YaeT mutant. *J. Biol. Chem. 280*, 27679–27687.
6. Fitzpatrick, D. A., and McInerney, J. O. (2005) Evidence of positive Darwinian selection in Omp85, a highly conserved bacterial outer membrane protein essential for cell viability. *J. Mol. Evol. 60*, 268–273.

7. Sklar, J. G., Wu, T., Gronenberg, L. S., Malinverni, J. C., Kahne, D., and Silhavy, T. J. (2007) Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A. 104*, 6400–6405.

8. Voulhoux, R., and Tommassen, J. (2004) Omp85, an evolutionarily conserved bacterial protein involved in outer-membrane-protein assembly. *Res. Microbiol. 155*, 129–135.

9. Walther, D. M., Rapaport, D., and Tommassen, J. (2009) Biogenesis of beta-barrel membrane proteins in bacteria and eukaryotes: evolutionary conservation and divergence. *Cell. Mol. Life Sci. 66*, 2789–2804.

10. Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J., and Kahne, D. (2005) Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli. Cell 121*, 235–245.

11. Gatsos, X., Perry, A. J., Anwari, K., Dolezal, P., Wolynec, P. P., Likic, V. A., Purcell, A. W., Buchanan, S. K., and Lithgow, T. (2008) Protein secretion and outer membrane assembly in Alphaproteobacteria. *FEMS Microbiol. Rev. 32*, 995–1009.

12. Knowles, T. J., Scott-Tucker, A., Overduin, M., and Henderson, I. R. (2009) Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat. Rev. Microbiol. 7*, 206–214.

13. Bos, M. P., Robert, V., and Tommassen, J. (2007) Functioning of outer membrane protein assembly factor Omp85 requires a single POTRA domain. *EMBO Rep. 8*, 1149–1154.

14. Gentle, I., Gabriel, K., Beech, P., Waller, R., and Lithgow, T. (2004) The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. *J. Cell Biol. 164*, 19–24.

15. Malinverni, J. C., Werner, J., Kim, S., Sklar, J. G., Kahne, D., Misra, R., and Silhavy, T. J. (2006) YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli. Mol. Microbiol. 61*, 151–164.

16. Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M., and Tommassen, J. (2003) Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science 299*, 262–265.

17. Charlson, E. S., Werner, J. N., and Misra, R. (2006) Differential effects of yfgL mutation on *Escherichia coli* outer membrane proteins and lipopolysaccharide. *J. Bacteriol. 188*, 7186–7194.

18. Lewis, C., Skovierova, H., Rowley, G., Rezuchova, B., Homerova, D., Stevenson, A., Sherry, A., Kormanec, J., and Roberts, M. (2008) Small outer-membrane lipoprotein, SmpA, is regulated by sigmaE and has a role in cell envelope integrity and virulence of *Salmonella enterica* serovar Typhimurium. *Microbiology 154*, 979–988.

19. Gatzeva-Topalova, P. Z., Warner, L. R., Pardi, A., and Sousa, M. C. (2010) Structure and flexibility of the complete periplasmic domain of BamA: the protein insertion machine of the outer membrane. *Structure 18*, 1492–1501.

20. Gatzeva-Topalova, P. Z., Walton, T. A., and Sousa, M. C. (2008) Crystal structure of YaeT: conformational flexibility and substrate recognition. *Structure 16*, 1873–1881.

21. Kim, S., Malinverni, J. C., Sliz, P., Silhavy, T. J., Harrison, S. C., and Kahne, D. (2007) Structure and function of an essential component of the outer membrane protein assembly machine. *Science 317*, 961–964.

22. Knowles, T. J., Jeeves, M., Bobat, S., Dancea, F., McClelland, D., Palmer, T., Overduin, M., and Henderson, I. R. (2008) Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes. *Mol. Microbiol. 68*, 1216–1227.

23. Albrecht, R., and Zeth, K. (2010) Crystallization and preliminary X-ray data collection of the *Escherichia coli* lipoproteins BamC, BamD and BamE. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun. 66*, 1586–1590.

24. Knowles, T. J., McClelland, D. M., Rajesh, S., Henderson, I. R., and Overduin, M. (2009) Secondary structure and (1)H, (13)C and (15)N backbone resonance assignments of BamC, a component of the outer membrane protein assembly machinery in *Escherichia coli. Biomol. NMR Assign. 3*, 203–206.

25. Knowles, T. J., Sridhar, P., Rajesh, S., Manoli, E., Overduin, M., and Henderson, I. R. (2010) Secondary structure and ^{1}H, ^{13}C and ^{15}N resonance assignments of BamE, a component of the outer membrane protein assembly machinery in *Escherichia coli. Biomol. NMR Assign. 4*, 179–181.

26. Clantin, B., Delattre, A. S., Rucktooa, P., Saint, N., Meli, A. C., Locht, C., Jacob-Dubuisson, F., and Villeret, V. (2007) Structure of the membrane protein FhaC: a member of the Omp85-TpsB transporter superfamily. *Science 317*, 957–961.

27. Vanini, M. M., Spisni, A., Sforca, M. L., Pertinhez, T. A., and Benedetti, C. E. (2008) The solution structure of the outer membrane lipoprotein OmlA from *Xanthomonas axonopodis* pv. citri reveals a protein fold implicated in protein-protein interaction. *Proteins 71*, 2051–2064.

28. Misra, R. (2007) First glimpse of the crystal structure of YaeT's POTRA domains. *ACS Chem. Biol. 2*, 649–651.

29. Meli, A. C., Hodak, H., Clantin, B., Locht, C., Molle, G., Jacob-Dubuisson, F., and Saint, N. (2006) Channel properties of TpsB transporter FhaC point to two functional domains with a C-terminal protein-conducting pore. *J. Biol. Chem. 281*, 158–166.

30. Robert, V., Volokhina, E. B., Senf, F., Bos, M. P., Van Gelder, P., and Tommassen, J. (2006) Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol. 4*, e377.

31. 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., and Lithgow, T. (2010) A modular BAM complex in the outer membrane of the alpha-proteobacterium *Caulobacter crescentus. PLoS One 5*, e8619.

32. Ryan, K. R., Taylor, J. A., and Bowers, L. M. (2010) The BAM complex subunit BamE (SmpA) is required for membrane integrity, stalk growth and normal levels of outer membrane {beta}-barrel proteins in *Caulobacter crescentus. Microbiology 156*, 742–756.

33. Meyer, M., and Morgenstern, B. (2003) Characterization of gelatine and acid soluble collagen by size exclusion chromatography coupled with multi angle light scattering (SEC-MALS). *Biomacromolecules 4*, 1727–1732.

34. Wyatt, P. J. (1993) Light scattering and the absolute characterization of macromolecules. *Anal. Chim. Acta 272*, 1–40.

35. Zimm, B. H. (1948) The scattering of light and the radial distribution function of high polymer solutions. *J. Chem. Phys. 16*, 1093–1099.

36. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277–293.

37. Goddard, T. D., and Kneller, D. G. (2004). Sparky 3. (University of California, San Francisco).

38. Sattler, M., Schleucher, J., and Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. NMR Spectrosc. 34*, 93–158.

39. Habeck, M., Rieping, W., Linge, J. P., and Nilges, M. (2004) NOE assignment with ARIA 2.0: the nuts and bolts. *Methods Mol. Biol. 278*, 379–402.

40. Zwahlen, C., Gardner, K. H., Sarma, S. P., Horita, D. A., Byrd, R. A., and Kay, L. E. (1998) An NMR experiment for measuring methyl-methyl NOEs in C-13-labeled proteins with high resolution. *J. Am. Chem. Soc. 120*, 7617–7625.

41. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR 13*, 289–302.

42. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry 33*, 5984–6003.

43. Dosset, P., Hus, J. C., Blackledge, M., and Marion, D. (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR 16*, 23–28.

44. Garcia de la Torre, J., Huertas, M. L., and Carrasco, B. (2000) HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson. 147*, 138–146.

45. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers 22*, 2577–2637.

46. Hutchinson, E. G., and Thornton, J. M. (1996) PROMOTIF—a program to identify and analyze structural motifs in proteins. *Protein Sci. 5*, 212–220.

47. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 2126–2132.

48. Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr 60*, 2256–2268.

49. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem. 25*, 1605–1612.

50. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR 8*, 477–486.

51. Daley, D. O. (2008) The assembly of membrane proteins into complexes. *Curr. Opin. Struct. Biol. 18*, 420–424.

52. Orengo, C. A., Michie, A. D., Jones, S., Jones, D. T., Swindells, M. B., and Thornton, J. M. (1997) CATH—a hierarchic classification of protein domain structures. *Structure 5*, 1093–1108.

53. Ye, Y., and Godzik, A. (2003) Flexible structure alignment by chaining aligned fragment pairs allowing twists. *Bioinformatics 19* (Suppl. 2), ii246–ii255.

54. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A. 98*, 10037–10041.

55. DeLano, W. L. (2002) The PyMOL Molecular Graphics System.

56. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics 14* (51−55), 29–32.

57. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics 23*, 2947–2948.

58. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics 15*, 305–308.

59. Bond, C. S. (2003) TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics 19*, 311–312.

60. Bennett, M. J., Schlunegger, M. P., and Eisenberg, D. (1995) 3D domain swapping: a mechanism for oligomer assembly. *Protein Sci. 4*, 2455–2468.

61. Liu, Y., and Eisenberg, D. (2002) 3D domain swapping: as domains continue to swap. *Protein Sci. 11*, 1285–1299.

62. Holm, L., Kaariainen, S., Rosenstrom, P., and Schenkel, A. (2008) Searching protein structure databases with DaliLite v.3. *Bioinformatics 24*, 2780–2781.

63. Strynadka, N. C., Jensen, S. E., Alzari, P. M., and James, M. N. (1996) A potent new mode of beta-lactamase inhibition revealed by the 1.7 Å X-ray crystallographic structure of the TEM-1-BLIP complex. *Nat. Struct. Biol. 3*, 290–297.

64. Reynolds, K. A., Thomson, J. M., Corbett, K. D., Bethel, C. R., Berger, J. M., Kirsch, J. F., Bonomo, R. A., and Handel, T. M. (2006) Structural and computational characterization of the SHV-1 beta-lactamase-beta-lactamase inhibitor protein interface. *J. Biol. Chem. 281*, 26745–26753.

65. McCarthy, A. A., Haebel, P. W., Torronen, A., Rybin, V., Baker, E. N., and Metcalf, P. (2000) Crystal structure of the protein disulfide bond isomerase, DsbC, from *Escherichia coli*. *Nat. Struct. Biol. 7*, 196–199.

66. Sakakibara, D., Sasaki, A., Ikeya, T., Hamatsu, J., Hanashima, T., Mishima, M., Yoshimasu, M., Hayashi, N., Mikawa, T., Walchli, M., Smith, B. O., Shirakawa, M., Guntert, P., and Ito, Y. (2009) Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature 458*, 102–105.

67. Bode, W., Papamokos, E., and Musil, D. (1987) The high-resolution X-ray crystal structure of the complex formed between subtilisin Carlsberg and eglin c, an elastase inhibitor from the leech *Hirudo medicinalis*. Structural analysis, subtilisin structure and interface geometry. *Eur. J. Biochem. 166*, 673–692.

68. Chen, Y. J., and Inouye, M. (2008) The intramolecular chaperone-mediated protein folding. *Curr. Opin. Struct. Biol. 18*, 765–770.

69. Marsh, J. A., Singh, V. K., Jia, Z., and Forman-Kay, J. D. (2006) Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. *Protein Sci. 15*, 2795–2804.