Fall 2013

An Investigation of the Function of the Periplasmic Domain of BamA, an Essential Protein in Gram-Negative Bacteria

Danielle Andrea Castagneri
University of Colorado Boulder

An Investigation of the Function of the Periplasmic Domain of BamA, an Essential Protein in Gram-Negative Bacteria

Danielle Andrea Castagneri

Department of Chemistry and Biochemistry
University of Colorado at Boulder

Defense November 8th, 2013

Undergraduate Honors Thesis Committee

Primary Thesis Advisor:
Dr. Marcelo Sousa, Dept. of Chemistry and Biochemistry
Committee Members:
Dr. Joseph Falke, Dept. of Chemistry and Biochemistry
Dr. Michael Klymkowsky, Dept. of Molecular, Cellular, and Developmental Biology

Contents

Abstract

Gram-negative bacteria make up the majority of dangerous health-care associated infections (HAIs), and are becoming increasingly resistant to standard antibiotics. The Beta Barrel Assembly Machine (BAM) is a protein complex in Gram-negative bacteria composed of BamA, an essential member of the Omp85 family, and four associated lipoproteins B-E. The BAM is present in all Gram-negative bacteria and has a pivotal role in outer membrane protein (OMP) insertion, making it an attractive target for study and the development of new antibiotic agents. This investigation probes the importance of conserved residues in BamA, the central component of the BAM complex, and suggests the hypothesis that it may form higher order assemblies important for outer membrane biogenesis. The ConSurf bioinformatics server was used to identify highly evolutionarily conserved surface amino acids that could have putative functional importance. Tryptophan scanning and complementation assays were carried out to characterize potentially functionally important residues directed by this bioinformatics analysis. In addition, single cysteine residues were introduced in the periplasmic domain of BamA and used in crosslinking experiments in an attempt to trap and identify interacting partners. One tryptophan mutant, R36W, failed to complement, suggesting mechanistic importance. Strong crosslinking was observed in cysteine mutants T53C and R160C, supporting a model where BAM organizes into higher order assemblies to mediate folding and insertion of beta-barrel outer membrane proteins.

Acknowledgements

Firstly, thank you to Marcelo Sousa, without whose guidance and generosity this project would have been impossible. Dr. Sousa's instruction and expertise greatly facilitated my learning and was the most pivotal part of my undergraduate education within the Biochemistry department. Also many thanks to members of the Sousa Lab who generously provided protocols and technical tips, particularly Katarina Jansen, under whom I was "apprenticed" for a semester, and Cristina Sandoval, who helped troubleshoot my crosslinking protocol. Michelle Turco, Sandra Metzner, Myeongseon Lee, Alex Hopkins, and Arden Doerner also provided invaluable technical knowledge. I am grateful to have had the opportunity to spend the last year learning to be a biochemist in the Sousa Lab—it was the capstone of my education in the Biochemistry department. Also thank you to committee members Michael Klymkowsky and Joseph Falke, who generously offered their time on short notice.

Introduction

The Dangers of Antibiotic Resistance

Antibiotic resistance is a reality of modern medicine, and as such it constitutes a significant public health risk. Bacteria, mycobacteria, viruses, parasites, and fungi are treated with antimicrobials, but microbes themselves continually evolve new resistance mechanisms-- in many ways evolutionarily pressured by the very treatments that are designed to eradicate them [1]. This is a grave and cascading problem in today's healthcare system, in which a patient admitted to the hospital for one illness may die of secondary complications arising from the drug-resistant, pathogenic microorganisms that often inhabit health care settings. Nosocomial infections are consistently in the top ten leading causes of death in the United States; yearly, nearly 100,000 deaths are attributed to healthcare associated infections (HAIs): this is comparable with the number of accidental traumatic deaths, which includes automobile accident mortalities [3, 4]. Not only do HAIs cause many deaths, but these are deaths that may once have been preventable. With increasing antibiotic resistance, health care professionals may be forced to provide supportive care for a drug-resistant infection rather than treating the infection itself, and they must also be wary of spreading a particular strain of resistant infection to other vulnerable patients.

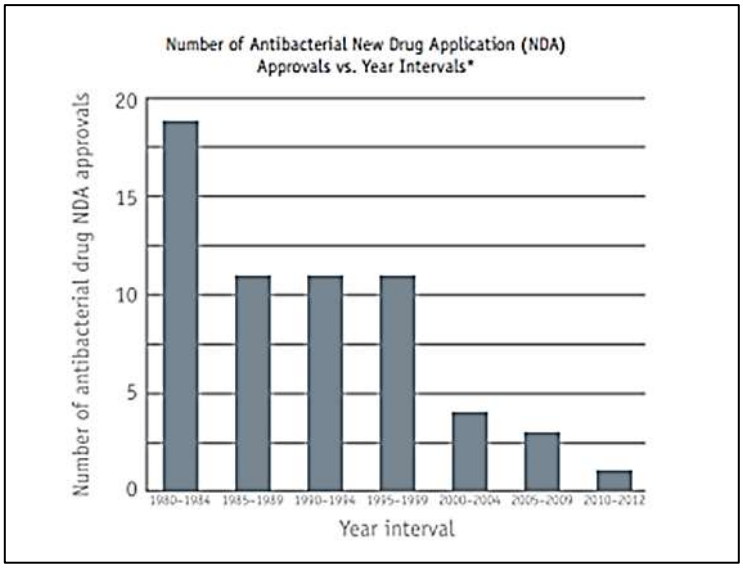


Figure 1: Number of antibacterial New Drug Application approvals per year interval. Despite the increasing demand for new antibiotics, NDAs for new antimicrobial agents have declined greatly since the early 1900's. Figure from [2].

"Antimicrobial" is a general term that encompasses the drug field of antibiotics, agents that kill or inhibit the replication of bacteria. The focus of this research is to study an

essential protein complex in Gram-negative bacteria, the Beta Barrel Assembly Machine (BAM), with the ultimate hope of informing the design of a novel class of antibiotic. Antibacterial resistance is a non-trivial problem, particularly in hospitals and retirement homes, and particularly because the discovery of novel classes of antibiotics has slowed dramatically [2]. Gram-negative bacteria, so named because the cells fail to take up the eponymous Gram dye, distinguish themselves from Gram-positive bacteria by being more typically (and intrinsically) prone to antibiotic resistance [5]. Gram-negative bacteria make up an estimated 70% of all ICU infections, making Gram-negative species a significant concern in health care settings [6]. In a Canadian survey, *E. coli* and *P. aeruginosa* (two of the most common Gram-negative HAIs) were responsible for more than 12% of all multidrug-resistant isolates, which require aggressive treatment with so-called "last-chance" antibiotics, an action that is not always successful [7]. In hospital settings, Gram-negative bacteria are one of the most dangerous elements of patient care: Gram-negative bacteria predominate in cases of ventilator-associated pneumonia and hospital-acquired urinary tract infections, and make up 30% of bloodstream infections acquired in health care settings [6, 8]. The scale of the problem of drug-resistant Gram-negative enterobacteria like *E. coli* is such that, for instance, the United States Center for Disease Control lists carbapenem-resistant *E. coli* as a threat to public health requiring "urgent and aggressive action" due to their resistance to nearly all antibiotics [2].

Currently, more than one hundred different resistance genes in bacterial plasmid DNA have been described. Due to selection pressure caused by the advent and institutionalization of antibiotics in the 1940s, there has been a significant present-day increase in the number of bacterial plasmids containing resistance genes versus the pre-antibiotics era [9, 10]. This has very real repercussions. Admittance for an inpatient hospital procedure for a range of diverse procedures has a typical length of stay of two to four days; ten years ago, the average hospital stay time was twice that [11]. Today, a long hospital stay presents increased opportunities for infection with drug-resistant microorganisms-- making it safer to send patients home sooner, where the likelihood of encountering drug-resistant bacteria is less probable. Any interaction with health-care

settings now brings a secondary threat of possibly antibiotic-resistant nosocomial infection.

Hospitals and healthcare establishments clearly perpetuate the problem of antibiotic resistance; previous and current medical practice espouses the over-prescription of antibiotics, thus providing evolutionary selection criteria for bacteria. A sobering example concerns the often unnecessary prescription of antibiotics for upper-respiratory infections, which are nearly always viral, and for which nearly half of patients in the United States are issued an antibiotic prescription [11]. Antibiotics are also used in animal husbandry, resulting in low antibiotic levels in food supplies, soil, and water runoff from feedlots and slaughterhouses. Tap water servicing residential areas downstream of pharmaceutical production has even demonstrated appreciable levels of antibiotics and other pharmaceutical agents [12].

The low levels of antibiotic drugs chronically present in a developed environment put forth an additional worry regarding antibiotic resistance; the trace amounts of antibiotics in drinking water and soil and the use of household antibiotics in the home, in daycares, and elsewhere exert a "non-lethal" pressure on bacteria, resulting in a potentially larger selection window for resistance than when lethal concentrations of antibiotic agent are used [13]. The implications if this finding are large: not only must the spurious use of antibiotics be curtailed (and protocols have been initiated to do so), but new antibiotics must be discovered in order to stay ahead of cascading resistance [14].

In light of the significant public health problem presented by Gram-negative infections, efforts are needed in the development of drugs that may inhibit resistance, as well as identification of targets for novel classes of antibiotics to treat resistant infections. This research aims to understand the protein BamA, a component of the Beta-Barrel Assembly Machine (BAM) complex, a protein complex that is integral in outer membrane protein biogenesis. Outer membrane protein biogenesis is an essential process in Gram-negative bacteria, and understanding this process may provide a new antibiotic drug target.

Gram-negative Bacteria Overview

Gram-negative bacteria include *Escherichia coli*, *Salmonella typhimurium*,, and other bacteria responsible for a variety of illnesses as diverse as pneumonia and hemorrhagic colitis. Gram-negative bacteria are unique due to their asymmetrical outer membranes. The inner membrane, composed of a typical phospholipid bilayer, is fluid and relatively permeable. The outer membrane, by contrast, is composed of lipopolysaccharides (LPS) in the outer leaflet, with the most common lipid attachment being Lipid A. The role of LPS in the outer membrane is to decrease fluidity in the outer membrane. Lipid A attachments can further form self-associations through electrostatic interactions of divalent cations, forming a charged permeability barrier against possibly hostile molecules [15]. Alterations to the polysaccharide attachment of the O antigen as well as Lipid A help some bacteria evade immune cells [15, 16]. The increase in the permeability threshold due to LPS in the outer membrane gives Gram-negative bacteria a degree of intrinsic antibiotic resistance. Between the two membranes is a peptidoglycan wall which, though it does present a target for the beta-lactam class of antibiotics, is extremely permeable to most molecules, allowing the passage of compounds as large as 30-57 kDa [17, 18].

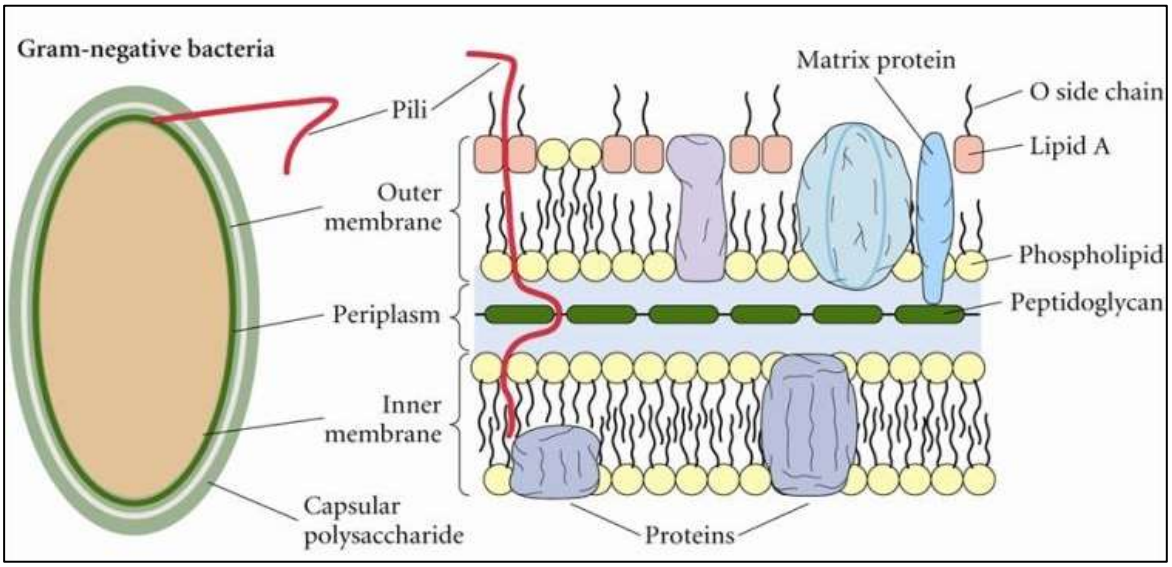


Figure 2: Gram-negative bacterial overview. The asymmetric outer membrane of gram negative bacteria decreases the organism's permeability to antibiotic agents. Figure adapted from [19].

The Beta-Barrel Assembly Machine (BAM)

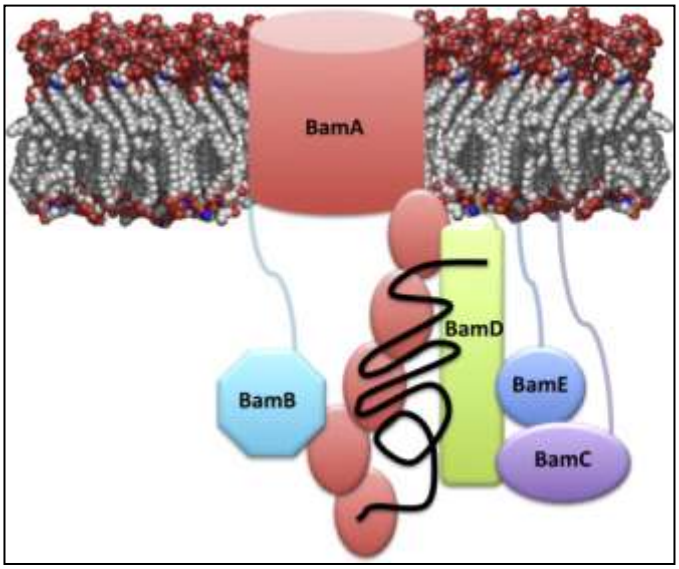


Figure 3: The Beta Barrel Assembly Machine. The BAM is composed of BamA and four associated lipoproteins, BamB, C, D, and E. Figure from [26].

Outer Membrane Proteins (OMPs) in gram negative bacteria are synthesized in the cytoplasm and cross the inner membrane via machinery like the SecYEG translocon [20, 21]. Nascent proteins move through the periplasmic space to be inserted into the outer membrane by a mechanism that is only beginning to emerge. This insertion process is essential to import of nutrients for the survival of Gram-negative bacteria, including the human pathogens *Escherichia* and *Neisseria* [22]. In addition, bacteria unable to carry out this mechanism efficiently have been shown to be more susceptible to antibiotics, a change in membrane permeability related to secondary effects on the machinery that introduces LPS to the outer leaflet of the outer membrane [23]. The multi-protein complex called the beta-barrel assembly machine (BAM) has been described as a crucial element of OMP insertion [24]. As shown schematically in Figure 3, the BAM is composed of one essential protein in the Omp85 family, BamA, as well as four lipoproteins, BamB, C, D, and E, of which only BamD is essential in *E. coli* [23]. Structures have been solved for BamA, B, C, D, and E [25-35]. BamA is composed of a beta barrel that traverses the outer membrane, as well as a periplasmic domain that likely behaves as an anchor for the associated lipoproteins [15, 30, 31]. Notably, BamA is found in all Gram-negative bacteria and is therefore an interesting target for study [34]. The periplasmic domain of BamA contains five polypeptide transport-associated (POTRA) domains. These water-soluble domains are thought be involved in protein-protein interactions, to initiate beta-strand formation in unfolded OMPs, and to have chaperone-like activity in OMP folding and insertion [32, 34, 35].

Studies of the POTRA domains have indicated that the POTRA "arm" has a crucial role in OMP insertion; deletions of POTRA subunits result in varied phenotypes depending on the specific POTRA domain being deleted. Only POTRAs 3-5 appear to be essential in *E. coli*. However, when only POTRAs 1 or 2 are eliminated there are significant decreases in the correct insertion of OmpA, a predominant outer membrane protein, reflecting decreased efficiency in OMP biogenesis. DegP, a protease for misfolded proteins in the periplasm, is upregulated, presumably to process the increase in poorly folded OMPs [34].

Hypothesis and Rationale: A Model for OMP Insertion

The current proposed mechanism for OMP insertion relies on the behavior of BamA and the POTRA domains as a scaffold. In this model, the actual transfer of nascent proteins from the cytoplasm occurs via protein chaperones Skp, DegP, and SurA as the protein emerges from the SecYEG translocon and other transporters in the inner membrane [36]. The SecYEG translocon is an ATP-dependent Type II secretion system transporter that brings unfolded OMPs via signal peptide recognition from the cytoplasm [37].

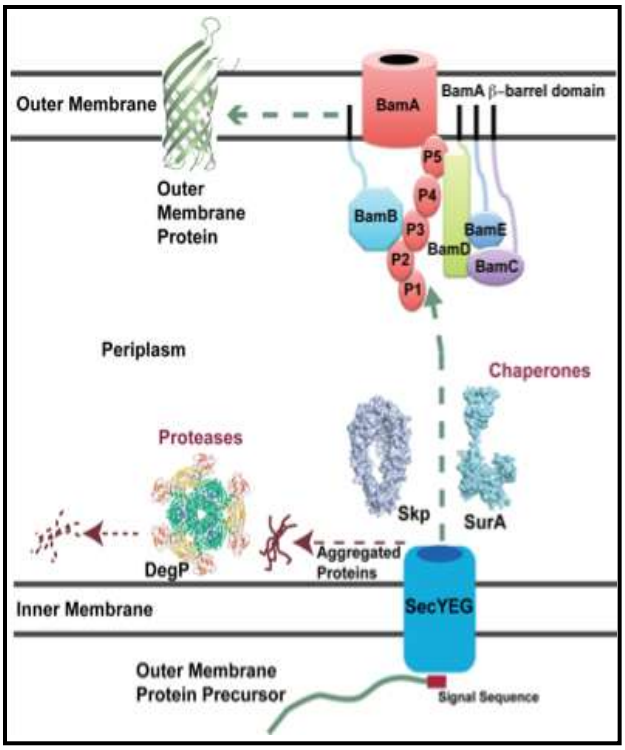


Figure 4: OMP Biogenesis. SecYEG translocates nascent OMPs from the cytoplasm to the periplasm. Chaperones SurA, Skp and DegP together with BamABCDE transport, fold and insert OMPs into the outer membrane. The mechanism of insertion is unknown. Figure from [38].

Figure 4 shows an unfolded protein as it is translocated through SecYEG, after which chaperones Skp and SurA bring the protein to the BAM complex, which folds and releases the protein into the outer membrane by an unknown mechanism. In this model, DegP degrades incorrectly-folded proteins that fail to be properly inserted. The POTRA domains (labeled P1-P5) do not

have a well-defined role within the current model for OMP insertion, although it has been shown that they can bind peptides and lipoproteins [35, 39]. POTRA deletion studies suggest that POTRAs 3-5 bind BamB and the BamCDE complex [34]. It has also been suggested that the POTRA domains, which have been shown to bind short peptides originating from OMPs, can initiate beta strand formation through beta-augmentation, thus aiding nascent OMP folding before insertion [35, 39- 41].

It has been suggested that SurA is the primary chaperone aiding BamA in the folding and transport of OMPs [42, 43]. It has been shown that SurA interacts with BamA *in vivo* [44]. This leaves DegP and Skp with a redundant role degrading misfolded proteins and possibly rescuing proteins lost by SurA. In SurA deletion mutants, DegP and Skp are both upregulated, presumably to compensate for the SurA deficiency [42]. However, double mutants depleting both Skp and SurA are not always lethal, notably in *Neisseria meningitidis*, implying the existence of a third redundant mechanism beyond the action of the Skp and SurA pathways [43].

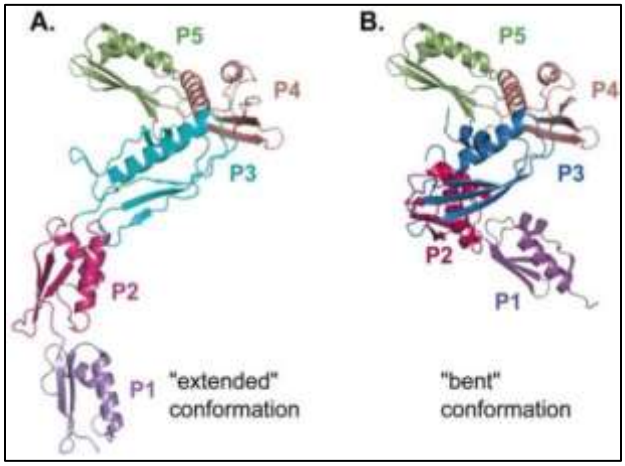


Figure 5: Bent and extended conformations of periplasmic POTRA domains of BamA. A "hinge" between P2 and P3 allows for a dynamic conformational change of the periplasmic domain of BamA. Figure from [34].

The Sousa lab has determined the structure of BamA's periplasmic POTRA arm, albeit in a truncated form lacking POTRA 5 due to difficulty purifying the full periplasmic domain [32, 35]. Later work defined the structure of the complete periplasmic domain containing all five POTRA domains through a combination of X-ray crystallography, Nuclear Magnetic Resonance (NMR) and Small Angle X-ray Scattering (SAXS) [25]. Previous work uncovered the finding that the POTRA arm of BamA adopts at least two conformations via a hinged region between POTRA 1 and POTRA 2; SAXS data has confirmed that the POTRA arm exists preferentially in both bent and extended conformations (see Figure 4) [35]. The

POTRA arm in its extended conformation is about 140 angstroms, a length that is consistent with the distance between the two membranes in the periplasmic space [35]. **Thus, it is physically possible for the POTRA arm to reach and interact with inner membrane components. This informs my hypothesis, that in addition to the model for OMP folding and insertion proposed previously, the POTRA domains of BamA interact physically with inner membrane components, allowing for the direct transfer of OMPs from the inner to the outer membrane.** The periplasmic arm has a number of peptide binding motifs, suggesting that the arm could function as a transperiplasmic folding envelope for nascent OMPs [35]. As mentioned, deficiencies of the chaperones Skp and SurA do not always result in a lethal phenotype, suggesting that a third compensating mechanism of OMP biogenesis exists [43]. A higher-order folding envelope formed between the two membranes by the periplasmic domain of BamA would be an elegant solution to a currently incomplete model for OMP biogenesis.

Specific Goal

To test my hypothesis, namely that BamA's POTRA domains interact with the inner membrane to form a trans-periplasmic protein folding envelope, I will mutate several surface amino acids via site-directed mutagenesis to search for residues involved in possible protein-protein interactions with inner membrane components. This will be achieved by

1) *identifying residues of functional interest based upon their degrees of phylogenetic conservation;*

2) *attempting to disrupt protein-protein interactions through mutations to tryptophan, a large, sterically hindering amino acid, and observing the effects via complementation assay; also*

3) *through the use of chemical crosslinkers, attempting to "trap" interactions between mutant versions of BamA containing the amino acid cysteine and inner membrane proteins.*

Results

Bioinformatics Analysis of BamA POTRA Domains

The ConSurf server uses phylogenetic relationships between homologs to estimate the rate of evolutionary change of amino acid residues and positions. The benefits of ConSurf over other algorithms are that ConSurf's calculations take into account amino acid chemical similarity as well as relationships between homologous proteins in different phylogenetic branches [45]. The server provides a variety of user-mandated options. In this case, the Sousa Lab crystal structure for BamA POTRAs 1-4 was uploaded to the ConSurf server via its Protein Database identification name 3EFC. ConSurf then extracted the amino acid sequence from this structure and searched for homologous sequences via CS-BLAST (Context-Specific Basic Alignment Search Tool).

CS-BLAST is differentiated from BLAST on the basis that its calculations take into account the environment of specific residues; for instance, a lysine exposed on the surface will be less likely to mutate than a lysine buried inside a protein, increasing entropic disorder. This difference enables CS-BLAST to find more homologs with fewer iterations and comparable error to its BLAST precursor [46]. ConSurf also utilizes Li and Godzik's CD-HIT program to cluster similar sequences, providing a threshold to discard highly similar homologs [47]. This increases accuracy of the entire estimate, as sequences that come from different organisms and phylogenetic branches offer more compelling data if residues are functionally conserved. CD-HIT constructs the Non-Redundant (NR) database, discarding duplicates of the seed sequence of the cluster, thus minimizing the possibility of false positives due to a particular sequence's overrepresentation in any pool of homologs. This is particularly important for an organism like *E. coli*, for which there are many genomic sequences available. The maximum number of homologs was set to 300, with identity of 95% (see Methods section).

Homologs were then organized into a Multiple Sequence Alignment (MSA) using default program MAFFT (Multiple Alignment Using Fast Fourier Transform) [48]. A phylogenetic tree was then built from this alignment using Rate4Site [49]. However, ConSurf improves this algorithm by computing conservation scores with regard to

position using Bayesian algorithms rather than Rate4Site's original Maximum Likelihood principle. It has been shown that using the Bayesian method reflects more accurate phylogenetic branch lengths and site-specific evolutionary rates in a wider range of conditions, making it an attractive method for this project, in which the phylogenetic branch lengths were unknown [50]. These calculations also generate confidence scores, which necessarily become smaller (implying more accuracy) when more varying homologous sequences are input. The JTT evolutionary model is used as a default by the ConSurf server [51].

Preliminary analysis of the POTRA domains of BamA suggests that portions of POTRAs 1 and 2 (those which hypothetically could interact with the inner membrane) offer fruitful residues to investigate. Figure 6 shows graded conservation scores of surface amino acids in the periplasmic portion of BamA; there is a high degree of conservation between POTRAs 1 and 2, suggesting some evolutionarily conserved function.

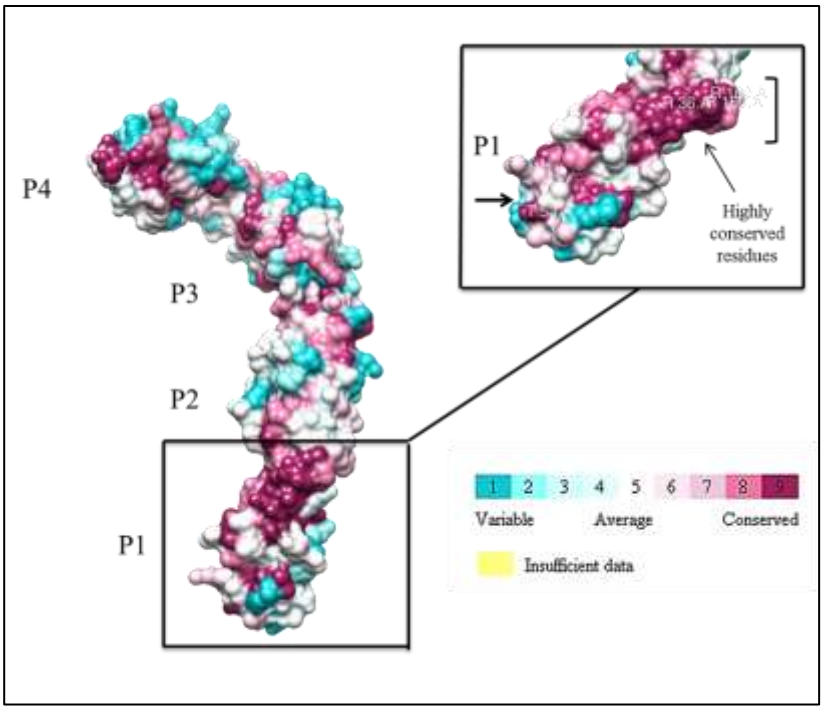


Figure 6: ConSurf analysis of POTRA domains 1-4 of BamA. ConSurf bioinformatics analysis shows a number of highly conserved amino acid residues between POTRAs 1 and 2, shown in dark magenta. Residues chosen to mutate are mapped as shown. Residues of interest are R36, R160, and R162 in the loop linking P1 and P2, as well as T53 at the C-terminal end of P1.

Residues to be modified by point mutation to tryptophan or cysteine were chosen primarily according to their conservation score given by the ConSurf bioinformatics tool. This was to direct experiments to optimize chances of success. Ideally, these single amino acid mutations would affect protein-protein or protein-phospholipid interactions

between the inner membrane and the POTRA domains, but would not affect the solubility, expression, or net folding of BamA.

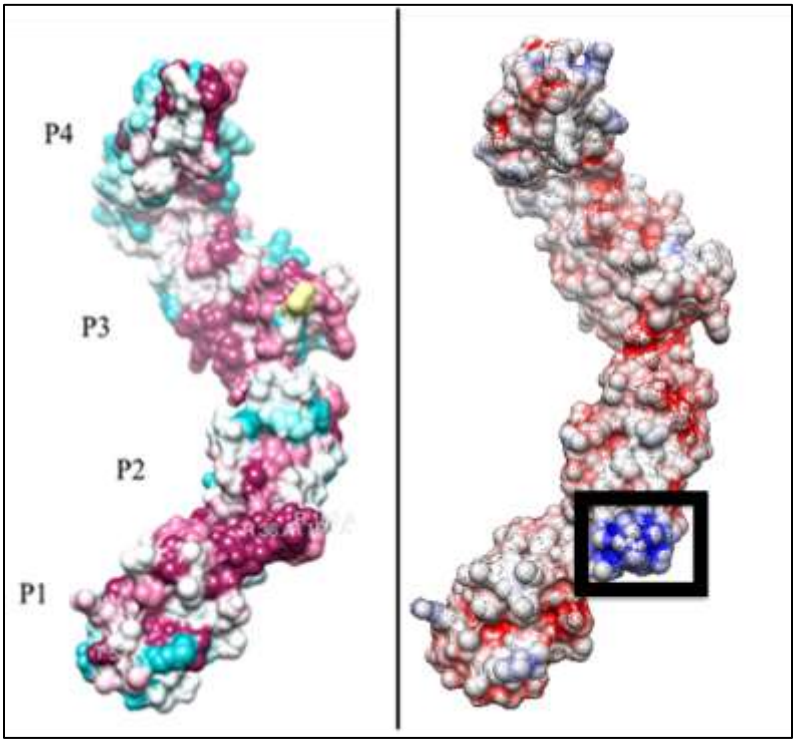


Figure 7: Illustration of positively charged area of POTRA1-2 link. Positive charges are shown in blue, with negative charges in red. The conserved area between POTRAs 1 and 2 corresponds with a unique area of positive charge that may mediate protein-protein interactions.

An interesting outcome of the ConSurf analysis is that the conserved patch between POTRAs 1 and 2 is largely composed of arginines, asparagines, and glutamines. This gives a positively charged cluster linking POTRAs 1 and 2, a pattern of similar adjacent residues that appears to be exaggerated only in this region (Figure 7).

The conserved residues observed in this region are capable of forming electrostatic interactions and hydrogen bonds, potentially aiding in protein-protein interactions or interactions with the phospholipid heads of the inner membrane. The ConSurf bioinformatics results therefore offered several targeted amino acid residues for study to attempt to discover functional importance.

As shown in Figures 6 and 7, three residues were chosen to mutate in the conserved area between POTRAs 1 and 2. Additionally, a residue was chosen at the end of POTRA1 based upon its intermediate conservation score (4) as well as its presumed proximity to elements of the inner membrane when the periplasmic domain of BamA is in its extended conformation. This identified, in total, four residues of interest: R36, R160, R162, and T53.

Cloning of BamA Mutants

Site-Directed Mutagenesis: The desired mutations chosen via conservation score or supposed proximity to the Sec machinery were introduced to an available pET41a(+) variant encoding full-length, hexahistidine-tagged BamA including an N-terminal signal sequence targeting the protein to the outer membrane. This plasmid was chosen due to its suitability for cloning and ease of purification, but it was not intended for protein expression owing to the insertion of the BamA gene backwards with respect to the T7 promoter.

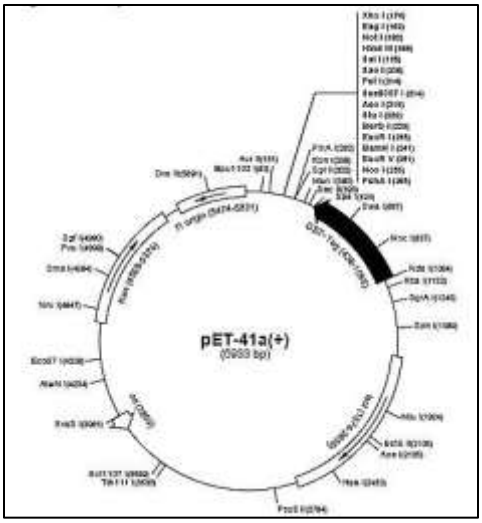


Figure 8: pET-41a(+) vector map. A parent vector consisting of the above pET-41a(+) expression system including the gene for BamA cloned into the multiple cloning site was used to create tryptophan and cysteine mutants for complementation and crosslinking assays. Figure from [53].

Tryptophan scanning complementation experiments required single mutations to tryptophan in the periplasmic domains of BamA. Three residues were chosen for mutation based upon their conservation score obtained from ConSurf: R160, R162, and R36, all of which are situated in the bridge between POTRAs 1 and 2. These residues were mutated to tryptophan via site-directed mutagenesis using the Stratagene Quickchange protocol (see Methods).

Tryptophan scanning was choses versus mutations to other amino acids (i.e. alanine or leucine) due to the residue's large size and intermediate hydrophobicity. The size of tryptophan may sterically hinder interactions between POTRAs 1 or 2 and other proteins. Its mildly hydrophobic character may also prevent interactions between the soluble POTRA domains and the soluble, periplasm-facing portions of the inner membrane proteins or phospholipids with which BamA is most likely able to interact. Alternatively, although tryptophan is more favorably buried in protein interiors, one tryptophan residue exposed to solvent is unlikely to have an aggregating destabilizing entropic effect.

Typtophan scanning has been used effectively to demonstrate structure and mechanism of various proteins [52].

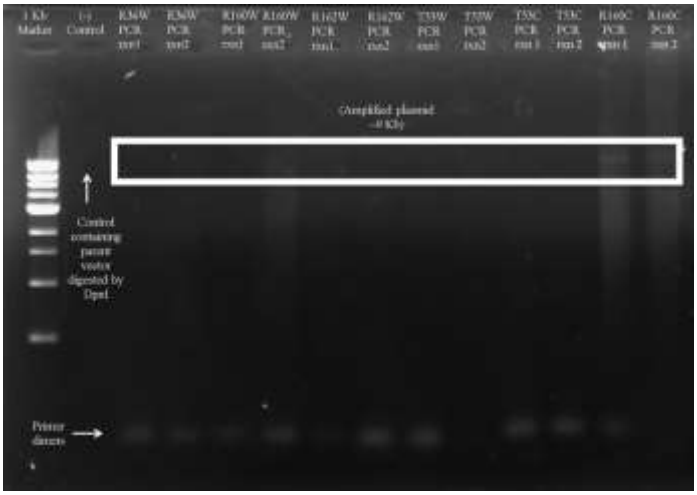


Figure 9: PCR amplification of pET plasmids encoding BamA mutants. 0 .8% agarose gel in 0.5% TAE containing EtBr loaded with 10 µL PCR reaction and run at 130V for 30 minutes. Plasmid amplification can be seen at ~9 Kb. The Stratagene protocol for site-directed mutagenesis was used to amplify DNA (see Methods), after which methylated parent DNA was digested 16 hours at 37 °C with DpnI.

To create the mutations necessary for complementation and crosslinking experiments, primers were designed to be ~30 nucleotides long with at least 15 nucleotides flanking the desired mutation, complementary to the coding and noncoding strands of the gene for BamA. Primers additionally had Guanine/ Cytosine content greater than 40%, ensuring melting temperatures within range for efficient annealing. Similar primers were designed to create cysteine mutants of conserved residues close to the terminal end of POTRA1, intended for cross-linking experiments with amine-to-sulfhydryl specific crosslinkers.

These plasmids were amplified via Polymerase Chain Reaction (PCR), digested with DpnI to eliminate the parent vector, and transformed into XL-10 gold *E. coli* cells, which were miniprepped to provide DNA for downstream applications. This resulted in the three tryptophan mutants, R36W, R160 W, and R162W, and two cysteine mutants, T53C and R160C. A fourth tryptophan mutant, T53W, was unable to be transformed successfully. Plasmids were sequenced to confirm mutation of interest as well as to screen for additional unintentional mutations arising from polymerase proofreading errors.

Transfer of BamA Tryptophan Mutants to a low-copy number plasmid:

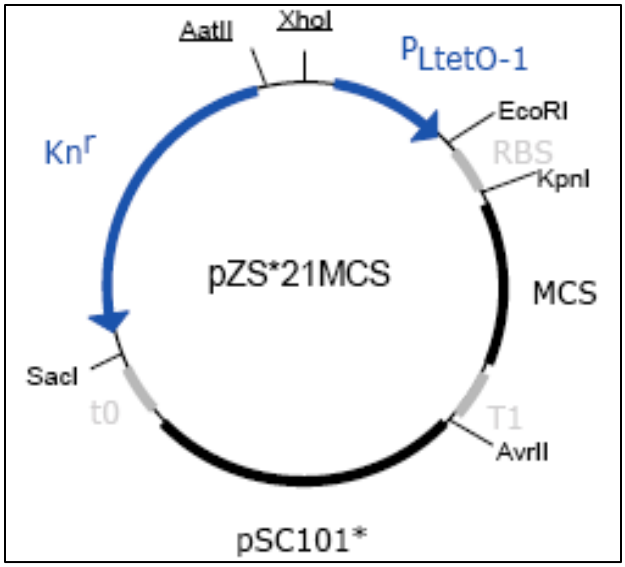


Figure 10: pZS21 plasmid vector map. The pZS21 vector encodes a kanamycin resistance cassette and is low-copy. Genes cloned into the multiple cloning site are transcribed under a tetracycline-inducible promoter. Figure from [54].

A complementation assay was used to test functionality of the tryptophan mutants described. A mutant strain of *E. coli* cells encoding the wild-type BamA gene under an arabinose-inducible promoter was used to this end. To complete this assay, it was necessary to move mutant genes to a low copy number plasmid encoding BamA mutants under a constitutive *E. coli* promoter, such that functional BamA expressed from the plasmid gene could "rescue" JCM166 cells grown in the absence of arabinose. The vector chosen for this purpose was a pZS21 plasmid (Figure 10).

Gene transfer was achieved through enzyme digest and ligation of BamA inserts into a pZS21 backbone. A pZS21 variant encoding GFP was used as the parent vector. This parent was digested with the enzymes EcoRI and XbaI (as were BamA inserts from site-directed mutagenesis), excised, purified, and ligated with BamA inserts with complementary sticky ends (Figure 11). Resulting ligations were verified by sequencing.

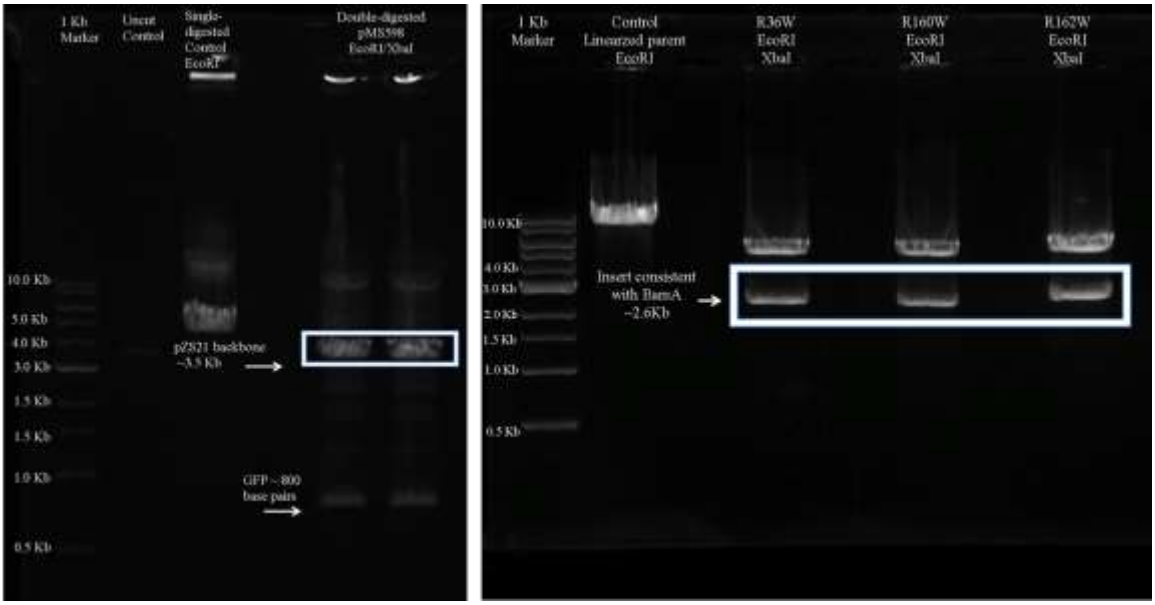


Figure 11: Enzyme digestion of pET mutants and pZS21 vector for ligation. 0.8% agarose gel in 0.5% TAE containing EtBr loaded with 10 µL digest reaction and run at 130V for 30 minutes. The sequenced BamA tryptophan mutant plasmids were digested overnight at 37°C with restriction enzymes EcoRI and XbaI, as was the parent pZS21 variant encoding GFP. The respective insert and backbone portions were excised and purified, resulting in a pZS21 vector and BamA inserts with complementary sticky overhangs.

Cysteine Mutants of BamA Intended for Crosslinking Experiments: Though the original mutations made in a pET41a cloning vector were optimal for cloning and sequencing purposes, they were unsuitable for crosslinking experiments, which necessarily require protein expression. Thus, it was necessary to design primers to amplify the BamA mutant gene while adding restriction digest sites to ligate the gene of interest into a different existing pET vector suitable for expression. Primers were designed to be complementary to the leading and lagging strands of the BamA gene until 5' and 3' restriction sites were added. In this case, primers included sites for NdeI (5' end of coding strand of BamA) and EcoRI (3'), plus an additional 6-nucleotide "tail" to facilitate enzyme recognition and digestion. Using these primers, the gene encoding BamA was amplified with the added restriction sites. This insert was digested, purified, and ligated into similarly doubly-digested pMS282, an available pET41a(+) variant encoding a kanamycin resistance cassette (see Figure 8) and BamC.

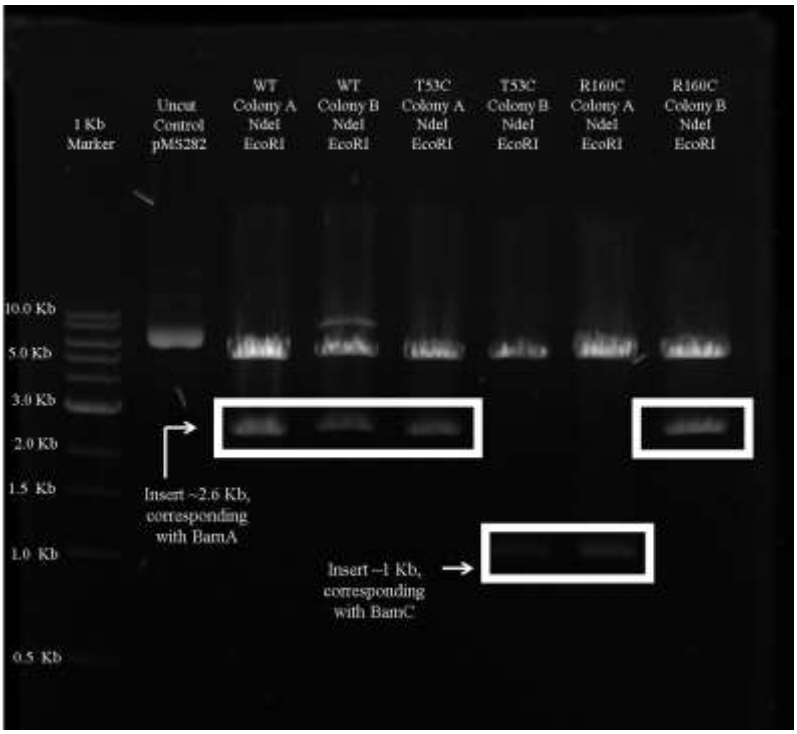


Figure 12: Colony screen by restriction digest of BamA cysteine mutant ligations into pET vector. 0.8% agarose gel in 0.5% TAE containing EtBr shows insert of interest (~3.6 Kb) is present in WT colonies A and B, T53C colony A, and R160C colony B. T53C colony B and R160C colony A returned the BamC insert present in the parent vector. Ligation colonies were miniprepped and digested with restriction enzymes NdeI and EcoRI for 1 hour at 37°C. Resulting digestion was loaded 10 µL per well and run at 135V for 30 minutes.

Transformants from this ligation were screened by enzyme digest; the parent vector produced an insert less than one kilobase in size, while BamA is about 2,600 base pairs long, allowing for unambiguous identification of the correct insert (Figure 12). Colonies containing the BamA insert were verified by sequencing. The resulting plasmids were two mutations of BamA, T53C and R160C, which were suitable for expression and therefore also for crosslinking experiments.

Complementation Assay

Single mutations to tryptophan were introduced in locations with high degrees of sequence conservation with the intent to target residues directly involved in the putative interaction between inner membrane elements and POTRAs 1 and 2 of BamA. Versions of BamA altered in this way are said to complement if they can "rescue" BamA-deficient *E. coli.* We acquired a useful strain of *E. coli*, JCM166, from the Silhavy Lab of Princeton University. The JCM166 strain encodes in its genome wild-type BamA under an arabinose promoter. In the presence of arabinose, JCM166 cells grow normally. In the absence of arabinose, JCM166 cells eventually die owing to their lack of functional BamA, which can be complemented by plasmid DNA encoding the functional protein.

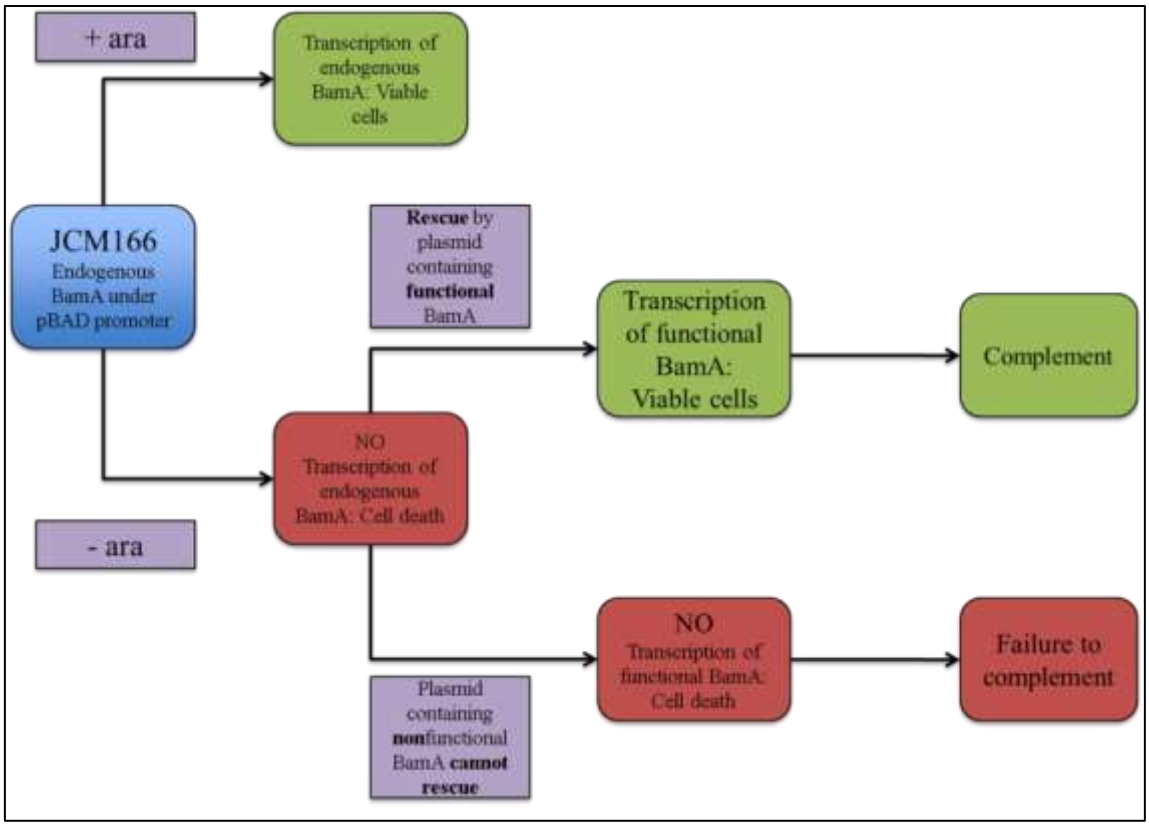


Figure 13: Complementation assay schematic. Complementation with JCM166 *E. coli* cells allows for the determination of residues of functional importance by screening for mutations that affect cell viability due to impaired activity.

Inability to fully complement would suggest that the plasmid-encoded protein is functionally compromised. This assay thus identifies mutants for further study. With this in mind the three tryptophan mutants R162W, R160W, and R36W were transformed into JCM166 cells and grown overnight in the presence of L-arabinose, after which growth media was switched to media containing the control sugar D-fucose and the cells were passaged three times, to deplete endogenous BamA. The effect of this change on optical density at 600 nm was monitored over a period of five hours. Figure 14 shows the results of this complementation assay.

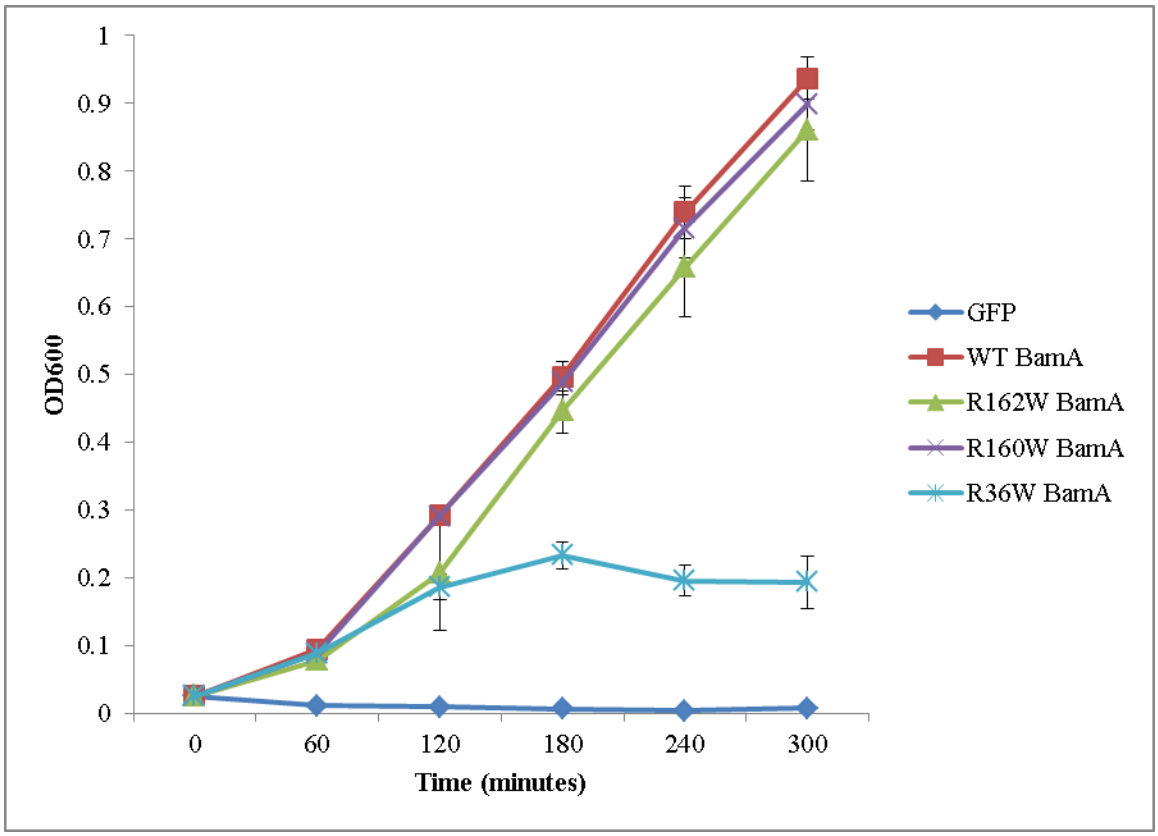


Figure 14: Complementation assay results. Optical density at 600nm was measured hourly after endogenous BamA was no longer expressed in complementation cells. After three passages into media containing the control sugar D-fucose, JCM166 *E. coli* cells containing plasmid encoding the control protein GFP failed to divide, indicating the depletion of endogenous BamA in all cells. Tryptophan mutants R160W and R162W appear to fully complement as their growth rate is indistinguishable within measurement error from the wild type positive control. However, tryptophan mutant R36W fails to complement unambiguously. Error bars represent standard deviation of triplicate trials.

Mutants R160W and R162W complement the endogenous BamA deficiency within measurement error. However, the mutant R36W fails to complement somewhat drastically. It demonstrates more growth than the negative control strain containing GFP, but it is demonstrably less viable than the wild-type strain ($P< 10^{-5}$).

To test if the failure of R36W to complement BamA depletion was due to impaired expression, an identical number of cells expressing WT or BamA mutants were analyzed by Western blot using a primary antibody against BamA. This blot returned bands for BamA in all samples using an anti-BamA probe (Figure 15). BamA is present in all strains, wild type and mutant, and in similar levels by inspection of the Western blot.

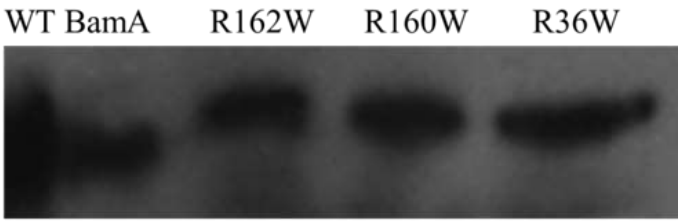


Figure 15: Western Blot for BamA. Samples were spun down and resuspended according to their OD_{600} such that they would have the same number of cells per load volume. These samples were run on a 10% SDS-PAGE gel at 100V for ~2 hours, transferred to a PVDF membrane, and probed with a primary antibody against BamA diluted 1:20,000. Blots were visualized via a secondary antibody conjugated to horseradish peroxidase, which reacts with a chemiluminescent substrate. BamA appears to be present in all samples.

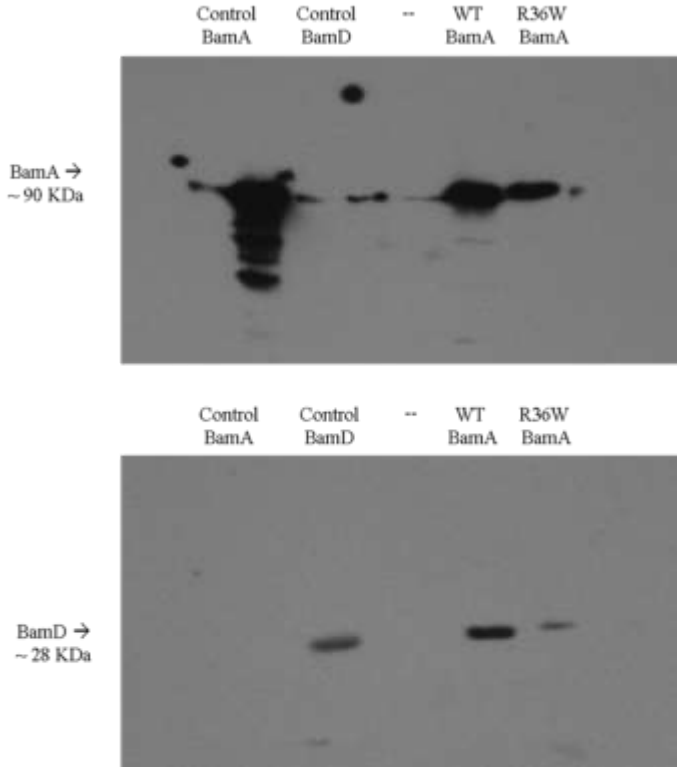


Figure 16: Western Blot of Ni-NTA pulldown showing the probable construction of the intact BAM complex. Cells were lysed and loaded on a Ni-NTA column and eluted with imidazole. Elution was used to run a Western blot as described. Primary antibodies against BamA and BamD were used at dilutions of 1:20,000 and 1:5,000, respectively. Control wells are purified aliquots of BamA and BamD.

Inconsistency in the running molecular weight of the wild-type control is thought to be an experimental artifact due to anomalous mobility in SDS-PAGE, as this result was not seen in subsequent Western blots.

In order to verify that BamA is both properly folded and properly localized in the outer membrane forming the Beta-barrel Assemble Machine, histidine-tagged plasmid-originating BamA was purified by Ni-NTA affinity column. This process purified BamA as well as whichever proteins it was associated with, i.e. components of the BAM complex. Western blots of antibodies against BamA and BamD, the two essential components of the BAM complex, showed that both BamA and BamD were present in the purified sample. This indicated that the construction of the BAM complex was proceeding normally, suggesting that the R36W mutant's failure to complement was unrelated to lack

of expression, misfolding, or mislocalization. A decrease in the mutant R36W's amount of both BamA and BamD in Figure 16 is attributed to protein loss in purification.

Cross-Linking Experiments

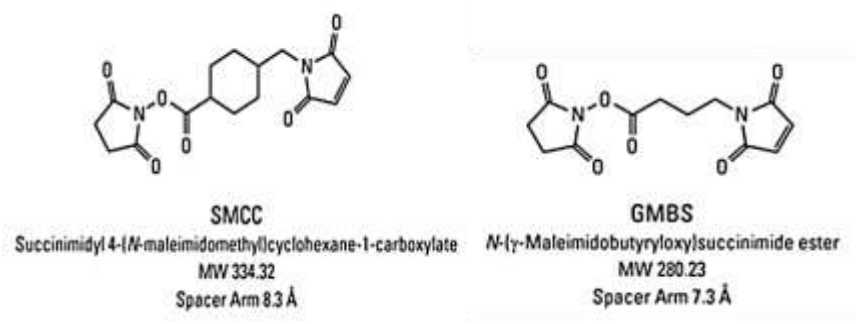


Figure 17: Membrane-permeable amino-to-sulfhydryl crosslinkers applied to *in-vivo* cysteine mutant crosslinking experiments. Crosslinkers were obtained from Thermo Scientific/Pierce Protein Biology Products. Images from [55, 56].

Plasmids encoding cysteine mutants of BamA were transformed into BL-21/DE3 *E. coli* cells and grown without IPTG induction to minimize transcriptional levels of BamA, thus making any crosslinking of cysteine mutants a larger ratio of the total and therefore easier to observe. Crosslinkers GMBS and SMCC (Figure 17) were chosen on the basis of membrane permeability, stability in reducing conditions, and the ability to react with sulfhydryls and amines. In this way, crosslinking can be largely directed to the specific location of the single cysteine residue in the cysteine-mutant BamA variants and whatever proteins they may be interacting with. Cells were diluted to the same optical density and processed for use in Western blots such that each sample had the same number of cells per load volume. These blots were probed with an antibody against BamA, as previously. Western blotting showed the presence (or absence) of high-molecular weight species detected by an anti-BamA antibody, indicating the presence or absence of crosslinking. Figure 18 shows this resulting Western blot, which was again carried out with whole cell lysates.

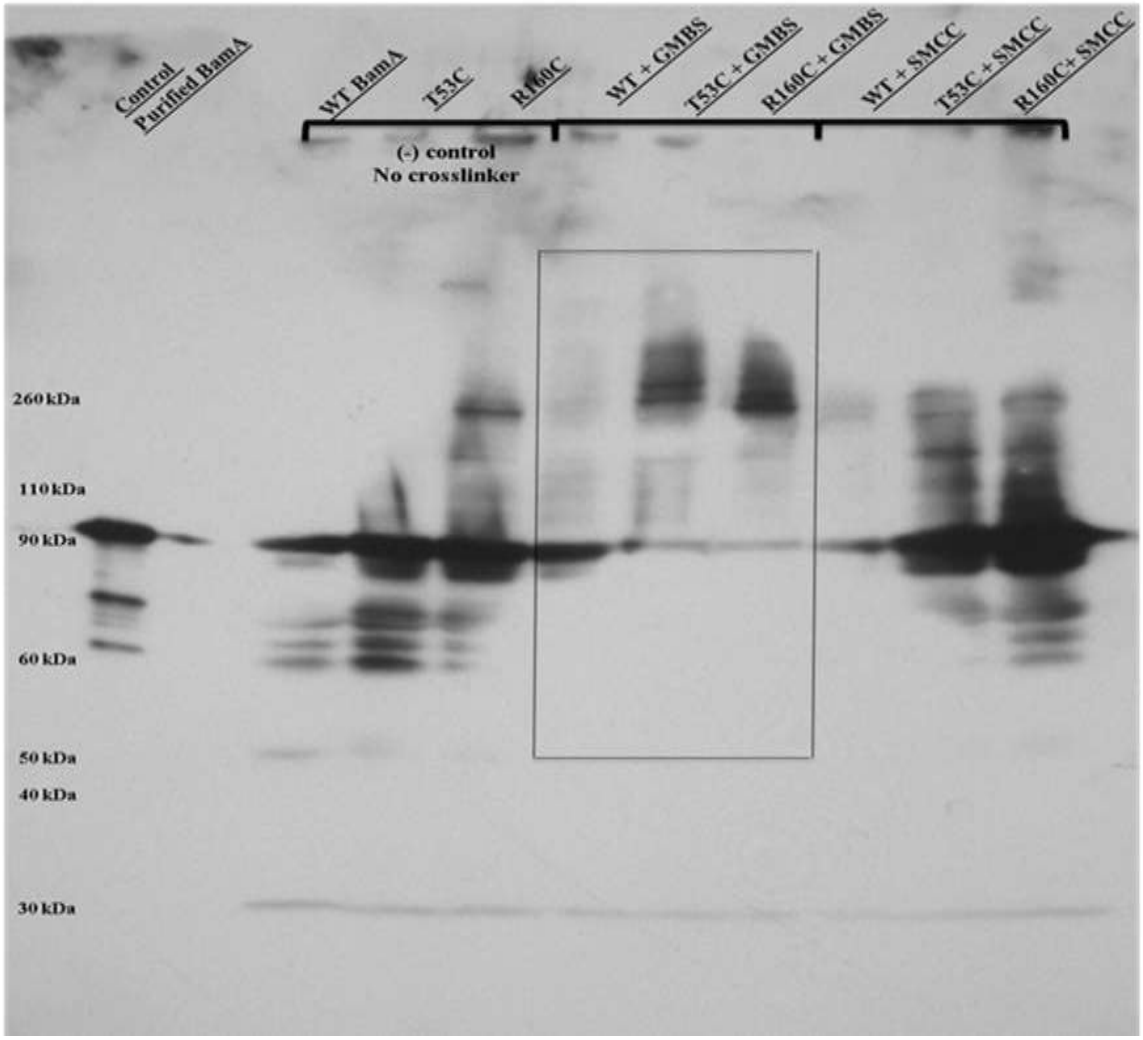


Figure 18: Western blot for BamA in crosslinked cells. Samples containing equal volumes with equal optical densities were run on a 10% SDS-PAGE gel at 100V for ~2 hours, transferred to a PVDF membrane, and probed with a primary antibody against BamA diluted 1:20,000. Blots were visualized via a secondary antibody conjugated to HRP reacting with a chemiluminescent substrate.

The Western blot for BamA in crosslinked cells appears to show crosslinked BamA at high molecular weight, > 200 kDa with the crosslinker GMBS (boxed in Figure 18). This band appears only in the cysteine mutants and not the wild type control; additionally, bands showing BamA at its expected molecular weight (~90 kDa) is much decreased in the cysteine mutants, supporting the notion that BamA is crosslinking with very high efficiency. There is similarly a noticeable difference between the wild type and mutant strains when subjected to the crosslinker SMCC, but high molecular weight bands are so faint that efficiency is likely low. There is a high molecular weight band in the R160C mutant of the negative control, itself composed of the two mutants and the wild type

strain without crosslinker. This may be the result of BamA dimers forming between cysteine mutants due to inadequate reduction by dithiothreitol (DTT) present in the loading buffer.

Discussion

This project was highly directed by the results of the bioinformatics analysis undertaken with ConSurf. Encouragingly, evolutionary conservation scores calculated by the ConSurf server illuminated an area of highly conserved surface residues that includes clustered residues capable of mediating protein-protein interactions. This provided the rationale for the project and aided in choosing specific amino acids for mutagenesis.

The fact that the conserved amino acids are localized to an area spanning POTRAs 1 and 2, combined with the fact that this particular area of conservation is so densely populated with amine- and amide- containing amino acids, is highly suggestive of some mechanistic function. The side groups of arginine, asparagine, and glutamine are capable of participating in electrostatic interactions and hydrogen bonding, in addition to other noncovalent interactions. It has been shown that evolutionarily conserved arginines are common in protein-protein interfaces [57]. Arginines characteristically can form many types of interactions that aid in protein-protein contacts; in addition to hydrogen bonding (five hydrogen bonds can be accepted or donated in the case of arginine, a large number of interactions), arginines can form salt bridges and exhibit pi electron delocalization that enables them to form cation-pi interactions, largely with aromatic residues [58]. Additionally, arginine's long hydrocarbon tail enables some hydrophobic interactions. The clustering of these residues does not point to a definitive explanation for their functional importance, but suggests that this area may serve as a "docking station" for nascent OMPs, chaperones, or the inner membrane proteins or phospholipids predicted by the transenvelope hypothesis.

Original site-directed mutagenesis in the pET vector suitable only for cloning was successful, despite the potential for mispriming or lack of annealing in primers that are by

necessity non-complementary to the site intended to be mutated. Sequencing with T7 promoter and terminator primers returned sequences that demonstrated the desired mutation and verified that no secondary, unintentional mutations were created via PCR cycling iterations.

The preliminary mutagenesis pET vector was chosen because it was suitable for cloning and was sufficiently high copy number to allow for high yields of DNA purification for downstream applications. However, this plasmid was not the original choice for mutagenesis. An effort was made to mutagenize a previously created pZS21 plasmid encoding full-length, wild-type BamA, which could have been used directly for both complementation and crosslinking experiments. Site-directed mutagenesis by PCR of this plasmid was not successful, likely due to difficulty amplifying the parent vector, a process that is necessary in the Quickchange mutagenesis protocol. Thus, it was necessary to create mutations in the pET cloning vector described and then move these genes into more suitable plasmid backbones. PCR amplification of the cysteine mutants and restriction digest and ligation of all inserts into both pET and pZS21 expression systems was successful, as shown by restriction digest check (in the case of the cysteine mutants in pET plasmids) and subsequently sequencing with appropriate primers (in the case of both the pET cysteine mutants and the pZS21 tryptophan mutants).

In the complementation assay, it was particularly important to ensure that any loss of function was due exclusively to impairment of an interaction in which the residue of interest participated, rather than lack of expression, solubility, or mislocalization. Recall that the use of tryptophan scanning to investigate functional importance of the residues identified for study was justified on the basis that mutations to a large, sterically hindering amino acid would effectively block interactions between the BamA POTRA domains and other proteins. In contrast to alanine or leucine scanning, single mutations to tryptophan are more informative and potentially catastrophic to function in interactions between proteins; however, such a dramatic change could potentially affect the proper folding and insertion of BamA.

Western blotting for BamA in all mutants showed appreciable levels of the protein, indicating that expression was unaffected (Figure 15). A limitation of this particular test is that the antibody used for the blot was non-specific for histidine-tagged BamA; in order to unequivocally show that JCM166 *E. coli* cells are rescued (or not) by plasmid-originating DNA, it would be necessary to blot using a more specific antibody, like an anti-histidine tag probe.

Western blots of Ni-NTA purified BamA effectively circumvented this limitation, as the purification process is also specific for histidine-tagged, plasmid-originating BamA, and they showed that BamA assembles with BamD in the BAM complex regardless of the R36W mutation. This shows that the R36W mutation does not cause aggregation or difficulties inserting into the outer membrane. A more complete test of the construction of the BAM complex would require blotting with antibodies against all five proteins; BamA and BamD were chosen in this case because they are the only essential components of the complex. Alternatively, purification of the outer membrane and subsequent Western blotting would robustly show that the R36W mutant is in fact inserted into the outer membrane and associates with BamB-E. However, the Western blot of Ni-NTA purified BamA showed that BamA was interacting with BamD, which would be highly improbable unless both proteins were properly folded and localized. This makes the finding that the R36W mutant fails to complement very significant, and implies that this particular residue is mechanistically important for the function of BamA.

This fact, that a mutation between POTRAs 1 and 2 impairs the function of the entire periplasmic domain, is an interesting finding because deletions of the entire POTRA1 and 2 domains have been shown to be non-lethal in *E. coli*, although the sigma^{E} SOS response is upregulated. [34]. Membrane permeability is altered, and these mutant phenotypes tend to display changes in outer membrane composition and an increase in the number of mis-folded OPMS, but POTRA1 and 2 deletions do not halt replication of *E. coli.* Thus, the fact that the R36W mutation in the bridge between POTRAs 1 and 2 fails to complement could lead to new understanding of the mechanism of OMP insertion. A possible explanation is that, while a truncated version of the periplasmic domain of BamA can be

compensated for by the protein chaperones or other mechanisms, introducing a tryptophan mutation in this location causes the POTRA arm to "stall," blocking its mechanism in a way that is catastrophic to OMP insertion.

The identification of a residue that appears critical for BamA function is not necessarily supportive of the hypothesis that the POTRA domains forms a trans-periplasmic folding envelope while interacting with inner membrane components. An alternative possibility is that R36 is involved in the conformational change of the POTRA arm (see Figure 5). It would be beneficial to repeat SAXS experiments comparing the wild-type and R36W mutant to see if conformations are altered or impaired, although the placement of R36 makes this possibility unlikely. Another possibility is that SurA, which has been shown to interact with BamA via crosslinking studies, is somehow blocked from docking by this tryptophan mutation [44]. These crosslinking studies could be repeated with the R36W mutant with an additional cysteine, to attempt to see if crosslinks can still be created. Again, this possibility is unlikely given that deletions of SurA do not give a lethal phenotype. The same is true for other BAM proteins (other than BamD, which can be ruled out by Figure 16) that may be blocked from scaffolding to BamA

Alternatively, R36W could be impeding the formation of the hypothetical transperiplasmic folding envelope. This in part supports why crosslinking experiments were successful. These crosslinking experiments could be repeated with R36W-cysteine mutants to see if the substitution of tryptophan at position 36 impairs the formation of such an envelope.

Crosslinking experiments with T53C and R160C also suggested areas for further experimentation. Figure 18 is highly suggestive of crosslinks forming between both cysteine mutants with the crosslinker GMBS, which is more flexible than the alternative SMCC. T53 is on the tip of POTRA1 and was chosen due to its hypothetical proximity to the inner membrane when the POTRA domains assume their extended conformation. T53 is also a residue that has a conservation score of 4, meaning it is somewhat conserved among homologs. By contrast, R160C is located in the highly conserved patch of

arginines between POTRAs 1 and 2, and it has a top conservation score of 9. Despite this difference, both mutants show high molecular weight bands, although T53 has a band slightly higher than R160C (Figure 18).

One possible explanation for this high molecular weight species is the formation of disulfide bonds between cysteine-mutant BamA, resulting in oligomers. Although the reducing agent DTT was present in the loading dye used to run the SDS-PAGE gel for Western blotting and samples were boiled prior to loading (an action necessary for the denaturation of membrane beta barrels), it is possible that disulfide bonds were not fully reduced. Alternatively, non-reducible crosslinks may be occurring between multiple BamA subunits. This would be an interesting outcome, as it has been suggested in *in vitro* studies that elements of the BAM complex oligomerize [59].

However, there are multiple possible explanations for the crosslinked species seen in Figure 18. It is known that POTRA 5 scaffolds the BamCDE subcomplex mediated by the essential lipoprotein BamD; POTRAs 3 and 4 also appear to have a role in scaffolding [34, 35]. The molecular weight marker identifies the presumed crosslinked bands to be more than 200 kDa in size. The complete BAM complex in a 1:1:1:1:1 ratio is not this large, and neither is any one component of the BAM complex linked with BamA a sufficient explanation for bands at molecular weights in excess of 200 kDa—BamA, B, C, D, and E have molecular weights of ~90, 40, 30, 25, and 12 kDa, respectively—and heat denaturation would break apart any non-crosslinked interactions between BAM complex components

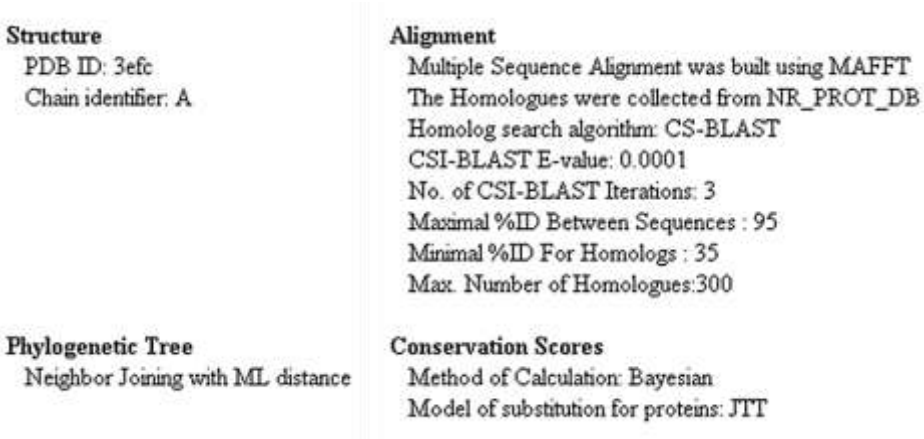
It is tempting to propose that the Sec translocon interacts directly with POTRA 1-2 of BamA. This would be the simplest form of a transenvelope assembly, as SecYEG translocates OMPs to the periplasm and has been suggested to partner with the BAM complex in two-partner secretion [21, 60]. However, The SecYEG monomer weighs 75 kDa. This, too, does not satisfactorily explain the high molecular weight bands seen in Figure 18. It is possible that the crosslinked species observed have anomalous mobility in the SDS-PAGE gel due to crosslinks causing uneven coating by SDS or other factors

affecting charge-to-mass ratio or steric barriers to migration through acrylamide pores. This makes it difficult to rule out the possibility of crosslinking with any known protein. Thus, crosslinked bands will be excised and digested by trypsin for analysis by mass spectrometry, enabling the comparison of the crosslinked entity to a library of possible proteins. In this way, the crosslinked species may be identified.

In conclusion, the experiments described here provide tantalizing evidence that the phylogenetically conserved patch of residues in POTRAs 1 and 2 that include R36 play a crucial role in BamA function. This is in stark contrast to the prevailing view in the field that POTRA1 and 2 are "disposable." Furthermore, the efficient crosslinking observed with single cysteine BamA mutants, particularly when using GMBS, suggests that BAM does indeed form higher order assemblies. Whether the assembly is with inner membrane components or due to oligomerization of BAM awaits the mass spectrometry identification of the crosslinked products.

Materials and Methods

ConSurf Bioinformatics analysis: The following parameters were run using the ConSurf Server:



Site-directed mutagenesis: Parent plasmid pMS629, a pET41a(+) variant encoding full-length BamA created by Troy Walton, was used in the Quickchange method for site-directed mutagenesis (Stratagene protocol). Synthetic mutagenic oligonucleotide primers were designed and ordered from Integrated DNA Technologies (IDT), Inc. (See Table 1).

Mutation	Primer sequence, 5'-3'
R36W	5'- CGCACCAACGGCGAC**cca**CTGAAGGCCTTCGAAATG -3' 5'- CATTTCGAAGGCCTTCAG**tgg**GTCGCCGTTGGTGCG-3'
R160W	5'-CACCAGTTTTAGGTCAAC**cca**GTTGCGCGGCAGCGGGGTC-3' 5'-GACCCCGCTGCCGCGCAAC**tgg**GTTGACCTAAAACTGGTG-3'
R162W	5'-CAGTTTTAGGTCAACACGGTT**cca**CGGCAGCGGGGTCACGACAGC-3' 5'- GCTGTCGTGACCCCGCTGCCG**tgg**AACCGTGTTGACCTAAAACTG -3'
T53C	5'- GATATCTTCATCATTAAC**aca**GTCGCCTGTGCGCACCGG-3' 5'- CCGGTGCGCACAGGCGAC**tgt**GTTAATGATGAAGATATC-3'
R160C	5'-CACCAGTTTTAGGTCAAC**aca**GTTGCGCGGCAGCGGGGTC-3' 5'-GACCCCGCTGCCGCGCAAC**tgt**GTTGACCTAAAACTGGTG-3'

Table 1: Primer sequences for site-directed mutagenesis of BamA

These oligonucleotide primers were used in the following 50 µL PCR reactions: 37 µL diH_2O + 5 µL 10x pfu Turbo Cx polymerase buffer (New England Biolabs, NEB) + 2.5 µL each upper and lower primers (5 µM) + 1.5 µL dNTP mix (10 mM, NEB) + 1 µL pMS629 parent (100 ng/µL) + 0.5 µL pfu Turbo Cx polymerase (NEB). Thermocycling parameters were as follows: 1) 95°C for 5 minutes 2) 95°C for 1 minute 3) 55°C for 1 minute 4) 68°C annealing temperature for 3 minutes; repeat 2-5 eighteen times 5) 55°C 1 minute, 6) 68°C 30 minutes, and 7) hold at 4°C until further experimentation. 1 µL per 50µL PCR reaction of restriction enzyme DpnI (NEB) was added to PCR products and incubated at 37°C overnight to digest methylated parent vector DNA, leaving mutant plasmid copies for use in further experiments. To verify appreciable amplification of mutant plasmid, 10µL of post-DpnI digested PCR product was run on 0.8% agarose gel in 0.5x tris-acetate-EDTA (TAE) buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA) at 135 V for 30 minutes and stained with ethidium bromide (EtBr) (an intercalator) and visualized at wavelength 302nm.

Remaining DpnI-digested PCR products were used to transform XL-10 gold *E. coli* competent cells (Agilent Technologies): 10µL of PCR product were introduced to 100µL aliquots of chemically competent XL-10 gold cells and placed on ice for 30 minutes. Cells were heat-shocked at 42°C for 45 seconds, placed on ice for 1 minute, and were suspended to a final volume of 1 mL in luria broth (Sigma-Aldrich) before incubation at 37°C, with agitation, for 45 minutes. Cell suspension was spun at 3000 rpm for 10 minutes,

after which supernatant was discarded, cells were resuspended in 200uL LB, and plated on agar plates containing 50 μg/ml kanamycin, a selecting antibiotic. Colonies resulting from this transformation were used to inoculate 5mL aliquots of LB containing 50 μg/mL kanamycin for use in Quiagen miniprep kits. Purified DNA from said miniprep kit (average concentration 100ng/μL) was sent for sequencing (ACGT, Inc.) to verify mutation and frozen at -20°C for downstream applications.

Digestion of tryptophan mutants for ligation into pZS21 vector: pMS598, a pZS21 variant encoding GFP and a kanamycin resistance cassette, was doubly digested with EcoRI and XbaI (NEB) for 12 hours at 37°C in the Cutsmart buffer system (NEB). Sample reaction: 7 μL pMS598 (100ng/μL) + 1 μL EcoRI + 1 μL XbaI + 1 μL Cutsmart buffer. Vector was gel-purified using Quiagen gel extraction kit. BamA pET mutants were digested and purified in the same fashion. Ligations were done in a 1:3 vector: insert ratio for 16 hours at 4°C. Sample reaction: 3 μL insert + 3 μLvector + 1 μL T4 Ligase buffer (NEB) + 1 μL T4 Ligase (NEB) + 2 μL diH_2O. Resulting colonies of XL-10 gold cells plated on Kanamycin-containing agar were then miniprepped as described and sequenced using Upper and Lower 566 BamA sequencing primers via ACGT, Inc.

Digestion of cysteine mutants for ligation into pET41a(+) vector for expression: pMS282, a pET plasmid encoding BamC and a kanamycin resistance cassette created by Jen Liddle was doubly digested as described, though using the restriction enzymes NdeI and EcoRI, which are compatible in the NEB 3 buffer system. PCR amplification primers were designed to amplify BamA cysteine mutants directly from pET cloning vector:

5'-TGC TTA **GAA TTC** TTA CCA GGT TTTACC GAT G-3'	Adds EcoRI site to 3' end of coding strand
5'-TAA GCA **CAT ATG** GCG ATG AAA AAG TTG CTC-3'	Adds NdeI site to 5' end of coding strand

Table 2: Primers intended to amplify BamA and add NdeI and EcoRI sites.

PCR program was run as described. Amplified inserts were digested with DpnI, again as described, and digested with NdeI and EcoRI after PCR purification (Quiagen kit) to remove DpnI. Ligations were done in as previously and transformed into XL-10 gold cells on kanamycin selective plates. Resulting colonies were miniprepped and screened via enzyme digest (as described) with NdeI and EcoRI; visualization via UV imager at 302nm of 0.8% agarose gel in 0.5% TAE containing EtBr and 10 μL each miniprepped plasmid showed which colonies contained insert of interest after running at 130V for 30 minutes. Plasmids with correct insert were sent to ACGT, Inc. for sequencing with T7 promoter and terminator primers.

Complementation assay: Tryptophan mutants mutations R162W, R160W, and R36W were transformed into JCM166 *E. coli* cells (Silhavy Group, Princeton University) according to the protocol outlined above, and streaked onto LB-agarose plates containing 50 μg/mL Kanamycin and 20% (w/v) agarose. Single colonies were selected from these plates and grown overnight at 37°C with agitation in 5 mL LB + 0.05% (w/v) L-arabinose. These cultures were diluted to OD_{600} = 0.025 in new, warmed media (5 mL) containing the control sugar D-fucose. Cultures were incubated at 37°C with aeration provided by a rotating shaker and allowed to grow to OD_{600} = 0.6, with OD_{600} measurements hourly. At OD_{600} = 0.6, trials were subcultured again to OD_{600} = 0.025, and monitored as described. This pattern was repeated until GFP control exhibited no growth and stable optical density, which in practice translated to three passages from overnight culture. Experiments were performed in triplicate.

Western blotting of complementation samples: Culture samples were taken following the culmination of the assay for use in western blotting: 100mL of each strain was grown and passaged as above. After five hours after the final passage, OD_{600} was measured and cells were diluted such that the final load for western blots would result in comparable levels of protein across all trials. Cells were resuspended in 200μL 1x SDS-PAGE loading dye containing 50mM dithiothreitol (DTT, Thermo), boiled 10 minutes, and spun at 14,000xg for 10 minutes before loading on 10% SDS-PAGE gel and run at 100 V for ~2 hours. Protein from gel was transferred using Pierce Fast Semi-Dry Transfer protocol on polyvinylidine fluoride (PVDF, Pierce) membranes with transfer taking place at 20V for 20 minutes. Membranes were blocked 16 hours at 4°C in 5% milk in TBS (10 mM TrisCl, pH 7.5, 150 mM NaCl), washed with TBS containing 0.5% v/v Tween-20 and 0.2% Triton X-100 (Sigma), and probed with UCO180 anti-BamA antibody diluted

1:20,000 in 5% milk in TBS for 1 hour at ambient temperature (antibody obtained from Petia Gatzeva-Topalova). TBS-Tween-Triton washes continued, and finally blots were exposed to Horseradish Peroxidase-conjugated AffiniPure Donkey anti-Rabbit IgG (Jackson Research) at dilution 1:10,000 in 7% milk in TBS. Immobilon Western Chemiluminescent HRP Substrate (Millipore) was used to visualize bands according to manual.

Crosslinking experiments: Cysteine mutants T53C and R160C were transformed into BL-21/DE3 cells following the same protocol described. Single colonies resulting from this transformation were grown overnight at 37°C with agitation in LB medium containing 50 μg/mL kanamycin. This culture was used to inoculate 25 mL LB containing kanamycin. Cells were grown to OD_{600} = 0.6 at 37°C with shaking and diluted to equal OD_{600} to minimize loading errors. Cells were then spun at 3,000 rpm, and washed twice with GMBS/SMCC crosslinking buffer (20 mM Sodium phosphate pH 7.5, 150 mM NaCl, 1 mM EDTA) and treated with crosslinkers GMBS or SMCC (Thermo) to a final concentration of 0.2 mM after resuspension in 5 mL crosslinking buffer. Cultures were incubated with crosslinker 30 minutes at room temperature with mild agitation. After incubation, crosslinking reaction was quenched with 20mM L-cysteine for 5 minutes shaking at room temperature. Cells were pelleted by centrifugation at 3,000 rpm and resuspended in 200 μL SDS-PAGE loading buffer containing 50mM DTT. Cells were vortexed, boiled 10 minutes at 95°C, cooled on ice, vortexed again, and spun 10 minutes at 14,000 rpm to achieve cell lysis. Samples were stored at -20°C until Western blot experiments were begun, according to the same protocol outlined above.

Ni-NTA Pulldown of BamA from Complementation Cells: Complementation cells were processed such that resuspension would result in the same number of cells/volume. Cells were resuspended in BugBuster lysis buffer (Millipore) and allowed to incubate for 1 hour at room temperature. The supernatant from this mixture when spun at 3000 rpm for 30 minutes was loaded on an equilibrated Ni-NTA (Quiagen) column at pH 7.5. Histidine-tagged BamA was eluted with 200mM imidazole after washes with low imidazole buffer. Elution was added to SDS-PAGE 2x loading buffer containing 50mM DTT and boiled for 10 minutes at 95°C before use in Western blot. Blotting for BamA took place as described. Blotting for BamD used a first-bleed antibody, UCO177, against BamD in dilutions of 1:5,000 in 5% milk for 1 hour at room temperature (antibody from Petia Gatzeva-Topalova). Secondary antibody and chemiluminescent reagent for visualization were added as described.

References

1. Freiden, T. *Antibiotic resistance and the threat to public health.* Centers for Disease Control and Prevention, US Department of Health and Human Services: Testimony to Committee on Energy and Commerce, 28 April 2010. Available from: http://www.cdc.gov/drugresistance/pdf/FriedenTestimony42810.pdf.
2. Centers for Disease Control and Prevention, US Department of Health and Human Services. *Antibiotic resistance threats in the United States, 2013.* Available from http://www.cdc.gov/drugresistance/threat-report-2013/pdf/ar-threats-2013-508.pdf.
3. Klevens RM, Edwards JR, Richards CL Jr, et al. *Estimating health care-associated infections and deaths in U.S. hospitals*. Public Health Report, 2002; 2007. (122):160-166.
4. Murphy S. et al. *Deaths: Final data for 2010.* National Vital Statistics Reports, 2013; 61(4):1-99.
5. L. Chopra, C. Schofield, M. Everett, et al. *Treatment of health-care - associated infections caused by gram-negative bacteria: A consensus statement.* Lancet infect. dis., 2008; (8):133-139.
6. Gaynes R, Edwards JR. *Overview of nosocomial infections caused by gram-negative bacilli.* Clin Infect Dis. 2005 (41):848-854.
7. Zhanel, G. et al. *Antimicrobial-resistant pathogens in intensive care units in Canada: Results of the CAN-ICU study, 2005-2006.* Antimicrobial Agents and Chemotherapy, 2008; 52(4):1430-37.
8. Hidron AI, Edwards JR, Patel J, et al. *NHSN annual update: antimicrobial-resistant pathogens associated with healthcare-associated infections: annual summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2006-2007*. Infect Control Hosp Epidemiol, 2008; (29):996-1011.
9. O'Brien T. *The Global Epidemic Nature of Antimicrobial Resistance and the Need to Monitor and Manage it Locally.* Clin Infect Dis, 1997. Supp1:S2-8.
10. Hughes V, and Datta N. *Conjugative plasmids in bacteria of the 'pre-antibiotic' era.* Nature, 1983; 302(5910):725-6.
11. Levy, S. *Antibiotic and antiseptic resistance: Impact on public health.* Pediatric Infectious Disease Journal, 2000; 19(10):S120-S122.
12. Larsson G and Fick J. *Transparency throughout the product chain: A way to reduce pollution from the manufacturing of pharmaceuticals?* Regulatory Toxicology and Pharmacology, 2009; 53(3):161-163.
13. Andersson D and Highes D. *Evolution of antibiotic resistance at non-lethal drug concentrations*, 2012; 15(3):162-72.
14. National Center for Emerging and Zoonotic Infectious Diseases/ Centers for Disease Control and Prevention. *National Antimicrobial Resistance Monitoring System (NARMS): Enteric Bacteria.* 2011.
15. Malinverni J and Silhavy T. *An ABC transport system that maintains lipid asymmetry in the gram-negative outer membrane*. Proc Natl Acad Sci USA, 2009; 106(19): 8009-14.
16. Bates J.M. et al. *Intestinal Alkaline Phosphatase Detoxifies Lipopolysaccharide and Prevents Inflammation in Response to the Gut Microbiota.* Cell Host and Microbe, 2007; **2** (6): 371–382.
17. Ehmann, David E. et al. *Avibactam is a covalent, reversible, non-β-lactam β-lactamase inhibitor*. PNAS, 2012; 109(29):11663-11668.
18. R. Scherrer, P. Gerhardt. *Molecular Sieving by the Bacillus megaterium cell wall and protoplast.* J. bacteriol., 1971; 107: 718-735.
19. *Pathogen Profile Dictionary: Gram Negative Bacteria*. J. Ugrad Biol S, 2010. Available from http://www.ppdictionary.com/gnbac.htm.

20. Gatsos X, Perry AJ, Anwari K, Dolezal P, Wolynec PP, Likic VA, et al. *Protein secretion and outer membrane assembly in Alphaproteobacteria.* FEMS Microbiol Rev. 2008;32(6):995-1009. PMCID: 2635482.
21. Pugsley, A. P. *The complete general secretory pathway in gram-negative bacteria.* Microbiol. Rev, 1993; 57:50–108.
22. Knowles T. et al. *Membrane protein architects: The role of the BAM complex in outer membrane protein assembly.* Nat Rev Microbiol, 2009; 7(3):206-14.
23. Vuong P, Bennion D, Mantei J, Frost D, Misra R. *Analysis of YfgL and YaeT interactions through bioinformatics, mutagenesis, and biochemistry.* J Bacteriol. 2008;190(5):1507-17. PMCID: 2258660.
24. Ruiz N, Kahne D, Silhavy TJ. *Advances in understanding bacterial outer-membrane biogenesis.* Nat Rev Microbiol, 2006; 4(1):57-66.
25. Noinaj N. et al. *Structural insight into the biogenesis of beta-barrel membrane proteins.* Nature, 2013; 501: 385-390.
26. Sandoval, C. M., Baker, S. L., Jansen, K., Metzner, S. I. & Sousa, M. C. *Crystal structure of BamD: an essential component of the β-barrel assembly machinery of gram-negative bacteria.* J. Mol. Biol, 2011; 409: 348–357.
27. Kim, K. H., Aulakh, S. & Paetzel, M. *Crystal structure of β-barrel assembly machinery BamCD protein complex.* J. Biol. Chem, 2011; 286: 39116–39121.
28. Knowles, T. J. *et al. Structure and function of BamE within the outer membrane and the beta-barrel assembly machine.* EMBO Rep, 2011; 12:123–128.
29. Kim K. and Paetzel M. *Crystal structure of Escherichia coli BamB, a lipoprotein component of the beta-barrel assembly machinery complex.* J Mol Biol, 2011; 406(5):667-78.
30. Hagan, C. L., Silhavy, T. J. & Kahne, D. *β-Barrel membrane protein assembly by the Bam complex.* Annu. Rev. Biochem, 2011; 80: 189–210.
31. Hagan, C. L., Kim, S. & Kahne, D. *Reconstitution of outer membrane protein assembly from purified components.* Science, 2010; 328: 890–892.
32. Gatzeva-Topalova PZ, Warner LR, Pardi A, Sousa MC. *Structure and flexibility of the complete periplasmic domain of BamA: the protein insertion machine of the outer membrane.* Structure, 2010;18(11):1492-501. PMCID: 2991101.
33. Knowles TJ, Jeeves M, Bobat S, Dancea F, McClelland D, Palmer T, et al. *Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes.* Mol Microbiol, 2008;68(5):1216-27.
34. Kim S, Malinverni JC, Sliz P, Silhavy TJ, Harrison SC, Kahne D. *Structure and function of an essential component of the outer membrane protein assembly machine.* Science. 2007;317(5840):961-4.
35. Gatzeva-Topalova PZ, Walton TA, Sousa MC. *Crystal structure of YaeT: conformational flexibility and substrate recognition.* Structure. 2008;16(12):1873-81.
36. Sklar JG, Wu T, Kahne D, Silhavy TJ. *Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in Escherichia coli.* Genes Dev. 2007;21(19):2473-84.
37. Johnson T. et al. *Type II Secretion: From structure to function.* FEMS Microbiology Letters, 2006; 255(2):175-186.
38. Sousa, M. (2012). *Bam grant proposal: Revision*. Informally published manuscript, Department of Biochemistry, University of Colorado at Boulder.
39. V. Robert, E.B. Volokhina, F. Senf, M.P. Bos, P. Van Gelder, J. Tommassen *Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif***.** PLoS Biol., 2006; 4: e377-87.
40. Heuck A., Schleiffer A., and Clausen T. *Augmenting beta-augmentation: Structural basis of how BamB binds BamA and may support folding of outer membrane proteins.* J Mol Biol, 2011; 406(5):659-66.

41. Knowles TJ, Jeeves M, Bobat S, Dancea F, McClelland D, Palmer T, et al. *Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes.* Mol Microbiol. 2008;68(5):1216-27.
42. Rizzitello A., Harper J., and Silhavy T. *Genetic evidence for parallel patheways of chaperone activity in the periplasm of E. coli.* J. Bacteriol, 2001; 183(23):6794-800.
43. Volokhina E. et al. *Role of the periplasmic chaperones Skp, SurA, and DegQ in outer membrane protein biogenesis in* Neisseria menigitidis. J. Bacteriol., 2011; 193(7):1812-1821.
44. Bennion D. et al. *Dissection of beta-barrel outer membrane protein assembly pathways through characterizing BamA POTRA1 mutants of E. coli.* Molec. Microbiol, 2010; 77(5):1152-71.
45. Glaser F. et al. ConSurf: *Identification of functional regions in proteins by surface-mapping of phylogenetic information.* Bioinformatics, 2003;19(1):163-164.
46. Angermüller, C., Biegert, A., and Söding, J. *Discriminative modelling of context-specific amino acid substitution probabilities.* Bioinformatics, 2012; 28 (24): 3240–7.
47. Li,W. and Godzik,A. *CD-Hit: A fast program for clustering and comparing large sets of protein or nucleotide sequence.* Bioinformatics, 2006; 22:1658-1659.
48. Katoh, K.; Misawa, K.; Kuma, K.; and Miyata, T. *MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform.* Nucleic Acids Research, 2006; 30 (14): 3059–66.
49. Pupko T., Mayrose I., Glaser F., and Ben-Tal N. *Rate4Site:An algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues.* Bioinformatics, 2002; 18 Suppl 1:S71-7.
50. Mayrose, I., Graur, D., Ben-Tal, N., and Pupko, T. *Comparison of site-specific rate-inference methods: Bayesian methods are superior.* Mol. Biol. Evol., 2004; 21(9):1781-1791.
51. Jones D., Taylor W., and Thornton J. *The rapid generation of mutation data matrices from protein sequences.* Computer Applications in the Biosciences, 1992; 8: 275-282.
52. De Feo C., Mootien S., and Unger V. *Tryptophan scanning analysis of the membrane domain of CTR-copper transporters.* J. Membr. Biol., 2010; 234(2):113-23.
53. Novagen. pET41a(+) vector map. 2010. Available from http://www.emdmillipore.com/docs/docs/PROT/TB239.pdf.
54. Expressys. *pZS Expression System: pZS21.* 2010. Available from http://www.expressys.com/available_vectors.html.
55. Thermo Scientific. *SMCC website entry*. Available from http://www.piercenet.com/product/smcc.
56. Thermo Scientific. *GMBS website entry*. Available from http://www.piercenet.com/product/gmbs.
57. Bogan, A. and Thorn, S. *Anatomy of hotspots in protein-protein interfaces*. J. Mol. Bio., 1998; 260: 1-9.
58. Pednekar D., et al. Electrostatics-defying interactions between arginine termini as a thermodynamic driving force in protein-protein interaction. Proteins, 2009; 74(1): 155-63.
59. Hagan C. and Kahne D. *The reconstituted Escherichia coli BAM complex catalyzes multiple rounds of beta barrel assembly.* Biochemistry, 2011; 50(35):7444-6.
60. Jacob-Dubuisson, F. et al. Two-partner secretion: As easy as it sounds? Research in Microbiology, 2013; 164(6): 583-595.